

BME 400 Fall 2022 iPhone Virtual Reality Training Model Complete Notebook

PDF Version generated by

ALEXANDER VAZQUEZ

on

Oct 12, 2022 @03:08 PM CDT

Table of Contents

Project Information	2
Team contact Information	2
Project description	3
Previous Lab Archives	4
2022/09/15 - BME 300 FA21	4
2022/09/15 - BME 301 SP22	5
Team activities	6
Meeting Notes	6
2022/09/16 - Meeting Minutes	6
2022/09/09 - Meeting Minutes	8
2022/09/30 - Meeting Minutes	9
2022/10/05 - Meeting Minutes	10
Materials and Expenses	11
4/28/2022 Materials Spreadsheet	11
2022/09/23 - PDS	12
2022/09/30 - Design Matrix	17
Sam	18
Design Progress	18
2022/09/23 Mathematical Basis of Image Transposition	18
2022/09/29 The Single Frame View Design	19
2022/09/30 The Combined Frame View Preliminary Design	20
2022/10/10 Reserach on Anaglyph Filters	22
Haochen Wang	23
Design	23
2022/09/30 - Design ideas to combine hardware and software design	23
2022/10/05 - Python script	26
Research	28
2022/09/21 - Python image processing blog	28
2022/09/23-3D photography	31
2022/09/29-3D photography using context-aware layered depth inpainting	33
Alex Vazquez	35
Design Ideas	35
2022/10/02- Potential Ideas	35
Research	36
2022/09/12- Background Project Research	36
2022/09/24-iPhone Design Research	38
2014/11/03-Template	39

**Team contact Information**

HAOCHEN WANG - Oct 12, 2022, 12:26 PM CDT

Last Name	First Name	Role	E-mail	Phone	Office Room/Building
Block	Walter	Advisor	wfblock@wisc.edu		
Shaffrey	Ellen	Client	shaffrey@wisc.edu		WIMR
Samuel	Poore	Client	poore@surgery.wisc.edu		WIMR
Neuman	Samuel	Leader & BWIG	ssneuman@wisc.edu		
Vazquez	Alexander	Communicator & BPAG	arvazquez@wisc.edu		
Wang	Haochen	BSAC & BWIG	haochen.wang@wisc.edu		
Albiter	Frida	BWIG	albiter@wisc.edu		



Project description

HAOCHEN WANG - Oct 12, 2022, 12:28 PM CDT

Course Number: BME 400

Project Name: iPhone Virtual Reality Training Model for Microsurgical Practice

Short Name: iPhone VR

Project description/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.

About the client:

Dr. Ellen Shaffrey and Dr. Samuel Poore are plastic surgeons at the Wisconsin Institute for Research and are both very passionate about the project of designing an alternative for microsurgical practice. They have worked with multiple teams to design this project and are currently supporting the fourth semester of the project.



2022/09/15 - BME 300 FA21

HAOCHEN WANG - Sep 15, 2022, 12:31 PM CDT

BME Design-Fall 2021 - Nicholas Jacobson Complete Notebook

PDF Version generated by
HAOCHEN WANG
on
Sep 15, 2022 @ 11:58 AM CDT

Table of Contents

Project Information	2
Team/Project Information	2
Project Description	3
Team/Students	4
Client Meetings	4
2021/09/17 - First Client Meeting	4
2021/09/22 - Second Client Meeting (Lab Tour)	7
2021/10/01 - First Client Meeting	9
2021/10/08 - Second Client Meeting	10
2021/10/22 - 3rd Client Meeting	11
2021/11/12 - 4th Client Meeting	12
Advisor Meetings	13
2021/09/18 - Advisor Meeting	13
2021/09/17 - Second Advisor Meeting	14
2021/10/01 - Third Advisor Meeting	16
2021/10/28 - Fourth Advisor Meeting	16
2021/11/01 - 5th Advisor Meeting	17
2021/11/16 - Show and Tell and Advisor meeting	18
2021/11/22 - Sixth Advisor Meeting	19
Design Process	20
Materials and Expenses	21
Materials and Expenses	21
Fabricator	22
2021/11/18 - Testing on most feasible parameters	22
2021/10/28 - Fabricator Printer	26
Fabricator's Results	30
Protocols	30
2021/10/04 - Measurement on the belly	30
2021/10/04 - Modifications on the calculator and formula	32
Experimentation	33
2021/10/08 - Measurement on the belly	34
Project Files	37
Model/Requirements	38
Research Notes	38
Biology and Physiology	38
2021/10/01 - Microscopy in Stubby Population	38
2021/10/13 - Relationship between Risk for Injury - Research Techniques	39
Comparing Designs	40
2021/10/29 - Hand Motion Analysis Tool	40
2021/10/28 - Depth Perception with Single Camera	41
Microscopy Training	42
2021/10/14 - Microscopy Training Recommendations	42
Design Object	43
2021/10/21 - Our Vision Header Design	43
Tracing Documentation	44
Testing	45

[Download](#)

iPhone-VR_BME-300_Lab-Archive.pdf (88.6 MB) Lab Archive from BME 300 FA21



2022/09/15 - BME 301 SP22

HAOCHEN WANG - Sep 15, 2022, 12:32 PM CDT

BME Design-Spring 2022 - HENRY PLAMONDON Complete Notebook

PDF Version generated by
HAOCHEN WANG
on
Sep 15, 2022 @ 12:32 PM CDT

Table of Contents

Project Information	2
Team/Project Information	2
Project Description	3
Team/Specs/Notes	4
Client Meetings	4
2022/09/04 - First Client Meeting	4
2022/09/04 - Second Client Meeting	5
2022/09/05	6
2022/09/15 - Fourth Client Meeting	7
2022/09/29 - 8th Client Meeting	8
Advisor Meetings	9
2022/09/04 - First Advisor Meeting	9
2022/09/11 - Second Advisor Meeting	10
2022/09/18 - Third Advisor Meeting	11
2022/09/04 - Fourth Advisor Meeting	12
2022/09/29 - 8th Advisor Meeting	13
Design Process	14
2022/09/01 Drawing Material for Housing	14
2022/09/29 Brass and Teflon	15
Materials and Dimensions	16
478.0202 Material Specifications	16
Fabrication	17
Fabrication Results	18
Protocols	19
2022/09/08-Protocol for testing air depth perception	19
Experiment	20
2022/09/08 Test with Prototype 1.2	20
2022/09/08 Testing with prototype 1 to determine the angle of interest	22
2022/09/08 Test with charts	26
Project Files	26
2022/09/11 Product Design Specifications	26
Henry Plamondon	26
Research Notes	28
Biology and Psychology	28
Computing Diagrams	28
Optical Illusions	28
2022/09/01 Optics/Illusions	28
Design Notes	27
Timing Documentation	28
Hardware Notes	28
Research Notes	28
Biology and Psychology	28
Computing Diagrams	40
Materials	41
2022/09/01 - Research on Gasp Pro Brain	41
2022/09/01 - Research on Touch 1588	40

[Download](#)

iPhone-VR_BME-301_Lab-Archive.pdf (22.8 MB) Lab Archive from BME 301 SP22



2022/09/16 - Meeting Minutes

HAOCHEN WANG - Oct 12, 2022, 10:37 AM CDT

Title: Meeting notes

Date: 09/16/22

Content by: Haochen Wang

Present: Sam, Alex, Haochen, Prof. Block

Goals: To go through progresses for the semester

Content:

Advisor meeting

1. Presentation: design choice to make AR
2. Decide within few weeks
3. Describe options that you have tried - talk to MS HoloLens 2
4. **Order the 3D goggles and talk about this in the client meeting**
 - a. Test out the goggles
5. Consider real-time imaging processing in python
 - a. openCV
 - b. See sample code below

6. Test out with students on campus
 - a. Possibly from another country
7. Test the learning time compared to the exoscopes
8. The end point is a paper about the effectiveness of the system and reduce training complexity
 - a. The paper will be due next semester
9. phone -> laptop communication -> MP4 realtime code (python) -> image stream clip to left and right with filter -> add to output (timing)
 - a. Latency (time from input to output)
 - i. Come up with a requirement(30 fps, 0.1 s delay max)
 - b. Reliability (how often the laptop will fall behind during tasks)
 - i. Catch up with the real time processing if there is a time delay
 - ii. Priority in the task queue
10. Find real time processing software
11. Filtering is parallelizable
 - a. Analysis on the computational power (e.g: GPU)
12. **Learning image processing**
 - a. From python
 - b. Add filters to the images
13. Recommended restaurant: Stadium Takeout, Kosharie's

Client meeting

1. Goal is to use iPhone as a microscope with depth perception

- a. Used to use VR set, but there was too much lag time
2. Last year worked on making an attachment on the smartphone to split the view
 - a. Refine the mirror angle to achieve better results
3. Develop a larger prototype
4. Distance of the cameras determine the depth perception
 - a. Further distance, better depth perception
 - b. 63 mm apart
 - c. C920 cameras with raspberry pi**
 - d. Oculus web server for computational work
5. New monitor, not google cardboard.
6. Goals: have images stitched, complete astamimoses (?), ultimate goal is for med students/residents to do procedures.



2022/09/09 - Meeting Minutes

HAOCHEN WANG - Oct 12, 2022, 10:36 AM CDT

Title: Meeting notes

Date: 09/09/22

Content by: Haochen Wang

Present: Sam, Haochen, Prof. Block

Goals: To setup expectations for this semester

Content:

Progress report and what to report to the advisor

- Team activities

- What I did what I figured out/work out

- Team leaders break up the work/significance/gaps

Story to tell for the project

- Train sb with the design versus the microscope, what was the impact

- Fill the gap in commercial field (pinpoint)



2022/09/30 - Meeting Minutes

HAOCHEN WANG - Oct 12, 2022, 10:40 AM CDT

Title: Meeting notes

Date: 09/30/22

Content by: Haochen Wang

Present: Sam, Alex, Haochen, Prof. Block

Goals: To go through progresses for the semester

Content:

Advisor:

1. Augmented reality instead of VR (new specification)
2. Explain the two designs (merge/anaglyph filter/enhance depth perception with the python script)
 - a. We don't need high quality image because we don't operate on patients
 - b. Acquisition 2 pts, combine the channel, separating the channel
 - c. Constraints: just need to watch the screen during movie, but we need to augment
 - d. Time capacity when viewing the image to combine
 - i. Guy in Italy miss one eye
3. How 3D work
4. Brain adding two images to create depth perception
5. Keep the channels separate
6. Latency and frame rate (define numerically) go in to the design matrix
7. What's the display like, what are the requirements
8. More layer, better depth distinction
 - a. How RGB-D works

Client:

1. Consider the focus
2. Client used raspberry pi
3. Try magnification and check the processing time when zooming in
4. The time delay is still significant
 - a. Reduce the number of connections to further reduce the latency (ie.: do not connect to an extra monitor for display)
5. The goal is to have the deliverable ready in Feb 2023 before the clients go to South Africa
 - a. One unit is enough
6. Availability
 - a. Lab is available after 5 pm



2022/10/05 - Meeting Minutes

HAOCHEN WANG - Oct 12, 2022, 10:49 AM CDT

Title: Meeting notes

Date: 10/05/22

Content by: Haochen Wang

Present: Sam, Alex, Haochen, Prof. Block

Goals: To have a lab tour and have hand-on experience with the microscope

Content:

Client meeting:

1. Had a tour in the client's lab
2. Had experience viewing the microsurgery station from microscopes, to set up expectations on the intended image quality and magnification from the design



4/28/2022 Materials Spreadsheet

EMMA KUPITZ - May 03, 2022, 10:20 PM CDT



[Download](#)

Materials_Spreadsheet_-_Sheet1.pdf (55.4 kB)

**Title: Product Design Specification (PDS)****Date:** 09/23/2022**Content:**

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].



Title: Design Matrix for Methods to Merge Hardware and Software Design

Date: 09/30/22

Content:

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	4/5	20	5/5	20
Cost	15	3/5	9	2/5	6
Compatibility	10	4/5	8	4/5	8
Frame Rate	5	4/5	4	2/5	2
Durability	5	4/5	4	5/5	5
Total	100	26/35	76	28/35	86

1. Effectiveness(Time Lag): this is given the highest weight, because the primary goal for this semester is to reduce the time delay in streaming the video to the user. Since both design would use a monitor with wired connection to the smartphone, the limiting factor for reducing the time delay will be the running time complexity. Since the Combine Frame View design requires an extra step for combining the images, it has a lower score than Single Frame View design.

2. Sensitivity (Depth Perception): This is given the second highest weigh due to the importance of depth perseption. This gives the trainee the requisite knowledge of the location of their hands/tools during the mock procedure. Any inconsistency or misinput given to the user could cause issues in the training sequence that could translate to an actual procedure. When combining the individual views into one frame, Combined Frame View design may result in interference in the RGB signal of the object, which can then undermine the depth perception. Because the microsurgery practice is performed at a millimeter level of precision, we deduce that the interference can significantly impact the depth perception of the design. Therefore, Combined Frame View receives a lower score for the criteria.

3. Ease of Use: Ease of use is determined by the time to implement our deisgn to the microsurgery practice prosedure. Since the Single Frame View design requies adjustment on the program to match the frame rate of the monitor for display, it is given a lower score for the criterion.

4. Cost: cost includes the cost for mimium rerquired processing power of the smartphone for video capture, the laptop for processing the video, and possibly the cost for the display monitor that is compatible with the design. Because the team expect relative high proessing speed for video processing, neither of the design receives a perfect score. However, because the single frame design requies at least 2x fps than the Combined Frame View design to reach equivalent viewing results,

5. Compatibility: This criteria gauges the designs ability to be used with current training equipment and any other hardware components involved

6. Frame Rate: This is a quantitative measurement of the Frames per second that is output to the monitor. A higher number indicates a faster* processing of the image and results in a better user experience.

7. Durability: Refers to the ability of the design to withstand any wear/tear of use during normal operation.



Title: Mathematical basis of image transposition for manipulating images from 2 input sources

Date: 23 September 2022

Content by: Sam

Present: n/a

Goals: Make a forward mathematical proof of the transformations of the images that would occur in the software.

Content:

Mathematical Basis of Image Transposition for Surgical Practice Device

Samuel Neuman

Presented 23 September 2022

This report is intended to be the beginning of a discussion on the manipulations that would need to be made to two images taken from different angles and focused on a single object to combine them into one single image. This is neither a definitive discussion nor the final decision of the team's design direction.

Proof. We can represent two matrices that can be seen as modifications of the identity matrix. The number of each matrix is equal to the number of pixels in each of the two images. The first matrix gives the x and y coordinates of each pixel in an image:

$$\begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix} \quad (1)$$

The second matrix is a linear transformation that is meant to be applied to the position matrix:

$$\begin{bmatrix} \cos(\theta) & 0 \\ 0 & 1 \end{bmatrix} \quad (2)$$

Where theta is the angle of deflection of the second set of mirrors of the angle that the imaginary focal plane of the image makes with the plane orthogonal to the direction of the object being imaged.

Notice that when we multiply these two matrices, certain characteristics apply.

$$\begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix} \times \begin{bmatrix} \cos(\theta) & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} x\cos(\theta) & 0 \\ 0 & y \end{bmatrix} \quad (3)$$

Where

$$\det \begin{bmatrix} x\cos(\theta) & 0 \\ 0 & y \end{bmatrix} = xy \cos \theta = Area \quad (4)$$

Taking the determinate of matrix result of Eq.3, we see that this is always nonzero with reasonable parameters. This means that there is no nullity in any of the rank or column, meaning that there is no loss of dimension of resolution in the compression. \square

Functionally, our computations will be limited by two factors. The first is the computing power of the machine doing these calculations; a mild compression of the pixel resolution may be needed to be able to produce truly real-time images that have no frustrating lag-time between reality and virtual reality. The second is a possible reduction in the computed frame rate. It is known that human eyes only perceive 16 frames/second; thus, if the user's cell phone camera captures images at a higher frame rate, it may be beneficial to have the program occasionally skip frames to save on computing power.

Conclusions/action items:

In this entry, I propose the mathematical basis for manipulations that could be made to two images to combine them into one image. This derivation relies on the fact that the device used to take the pictures will involve a single camera and an apparatus of mirrors to split the image to a distance that is roughly the distance between human pupils. When implementing this math into image manipulations, it will also require matrices that carry the hexadecimal values of the color of each pixel. More, more manipulations will be required to average the euclidian distance (?) of the RGB profile of each pixel during compression, and the overlapping regions of the pixels in the "center" of the two images will likely need to be cropped and removed from the image entirely.



Title: Overview of the single frame view design

Date: 29 September 2022 (edited 11 October 2022)

Content by: Sam

Present: Sam

Goals: Outline the design concept of the single frame view design.

Content:

The Single Frame View design presents 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. Hypothetically, the user's device may need to operate at a lower fps (I believe the human eye is only capable of resolving 16 fps - need to check this). 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 features.

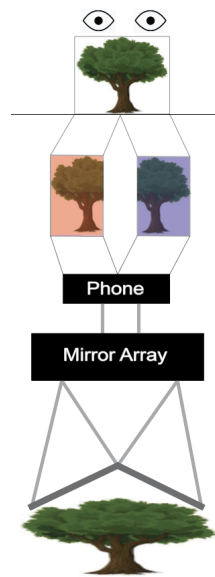


Figure 1. The schematic of the single frame view. First, the camera acquires the images of the two inclined planes of the object. The phone first applies the linear transformation to flatten the image and crops out the overlapping regions of the two halves of the image. Finally, the phone applies an 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.

It is our estimate that the anaglyph filter will be a large drawback for the client. The implementation of 3D glasses will make some colors more difficult to see, like reds, blues, and greens. This would be less than useful in all contexts.

Conclusions/action items:

This entry aims to document the single frame view design of our preliminary designs. It involves taking a picture of the object through the mirror array before applying the linear transformation outlined in the previous entry, cropping the images, applying an anaglyph filter of red and blue to the left and right images respectively, and then stitching them to create the anaglyph shadow in the combined images. This single image is then presented to the user at optimally 60 fps. We anticipate that the requirements of 3D glasses for this design will be unattractive due to the loss of color resolution by the user as well as obstruction of peripheral vision by the glasses frames.



2022/09/30 The Combined Frame View Preliminary Design

Sam Neuman - Oct 11, 2022, 11:53 PM CDT

Title: The Combined Frame View Preliminary Design

Date: 30 September 2022 (edited 11 October 2022)

Content by: Sam

Present: Sam

Goals: Give an overview of the combined frame view preliminary design prior to the preliminary deliverables

Content:

The combined frame, like the single frame design, involves the simultaneous capture of both hemispheres through the mirror array. Note that the images taken in both hemispheres will have some degree of overlap (this is what helps them become 3D). 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 images 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.

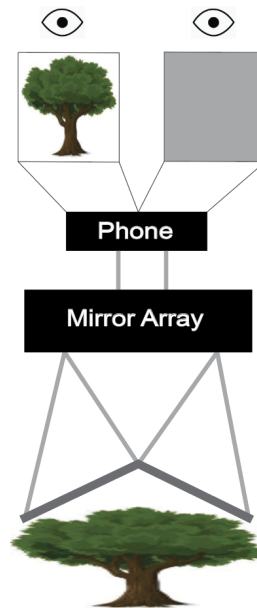


Figure 1. The combined view preliminary design. Both hemispheres of the images will be acquired simultaneously and cropped by the software to contain the same landmarks and thus be the same shape and encompass the same features. The phone screen will display only one of the left or right images to both 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.

This design is based off of the 3D movies used in the early 2000s in Europe. People reported feeling nauseated by watching them. We are apprehensive towards this design for this reason - we believe that the alternative viewpoints, with a large angle theta between the image planes and the reference plane, will require the brain to amalgamate images with large disparities between them. Probably nauseating to be honest.

Conclusions/action items:

The combined frame view is a 'more extreme' version of the single frame view with the benefit of not having the same computational demands as the single frame view design. The combined frame view design, like the single frame view design, acquires a single image of both hemispheres through the mirror array. The phone does not apply the linear transformation, only cropping the images such that both images include the same

landmarks and are the same size as each other. The screen will then rapidly alternate between the left and right hemisphere images, creating the illusion of retinal disparity, allowing the user to subconsciously perceive three dimensions in the images. We expect that this design will be nauseating and unappealing to the clients for this reason as well as unrealistic to operate at 120 fps based on the hardware available to us.



2022/10/10 Reserach on Anaglyph Filters

Sam Neuman - Oct 12, 2022, 12:07 AM CDT

Title: Reserach on Anaglyph Filters

Date: 10 October 2022

Content by: Sam

Present: Sam

Goals: Document my notes while reading a paper about creating and implementing anaglyph filters to 2D images

Content:

- Anaglyph filters are the red/blue filters used at movie theaters to create the perception of 3 dimensions. They require the user to wear red/blue glasses for it to work.

-This is based off of certain Gestalt principles (which are very abstract) and the idea that the brain takes the information perceived by the two eyes at the same time and tries to combine them into a meaningful image.

-The shadows produced by the filter have pixels whose RGB values are the sum of the original image's pixels RGB values and the RGB value of the desired color for the filter (i. e. blue).

-Thus, for us to use this, pixel/voxel math would be needed. We're not sure how difficult this would be for the computer to do in real-time.

-One way to do this, actually, is tensor algebra in which a tensor contains information about the pixel location, RGB values, and transformed shadow RGB values. This is a very bad way to do this but easy to code.

- Often 2D images don't contain any information about the 3D information. Actually this is entirely rare. Some AI can use computer vision principles to take a bad 'guess' as to the relative distances of objects in the image.

-Based on this lack of information alone, it is unlikely we would be able to use an anaglyph filter meaningfully.

-AI would be possible long-term, but not for our project. This paper gives a good way of how to do it. We also would need server-quality computer power for a few weeks and hours of surgery footage to make this work.

[1] E. Dubois, "A projection method to generate anaglyph stereo images," *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.01CH37221)*, 2001, pp. 1661-1664 vol.3, doi: 10.1109/ICASSP.2001.941256.

Conclusions/action items:

This paper gave a nice review of how anaglyph filters work and a mathematical background of how to make our own. The way to do this (most commonly used by professionals) is to create a tensor that contains all the information needed for the transformation, including the pixel/voxel original position, RGB values, and color change RGB values. What we didn't realize is that the anaglyph filter has variations in shadow size and strength based on the relative position of the image in the close/far direction (owing to the fixed distance of human pupils at around 63 mm). This could be discerned by some sort of AI framework, but it isn't feasible for our project. The anaglyph filter is really cool but simply not realistic for our project this semester.



2022/09/30 - Design ideas to combine hardware and software design

HAOCHEN WANG - Oct 12, 2022, 10:31 AM CDT

Title: Design ideas to combine hardware design and software design

Date: 09/30/22

Content by: Haochen Wang

Present: N/A

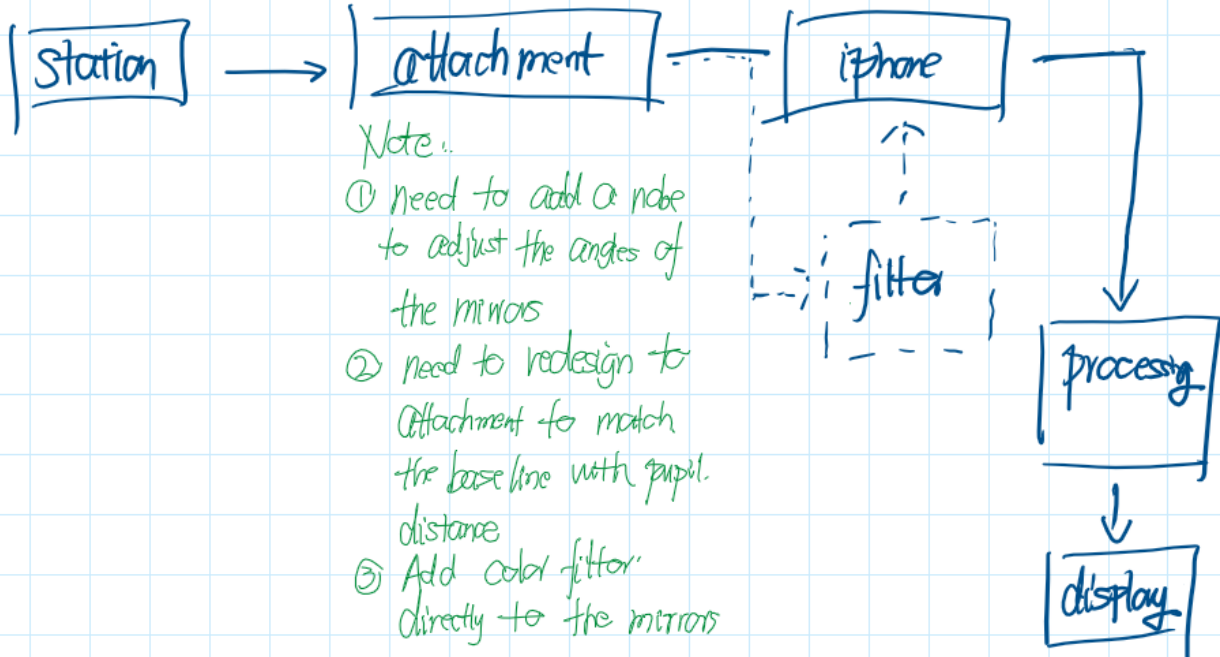
Goals: To provide possible solutions for combining designs from previous years

Content:

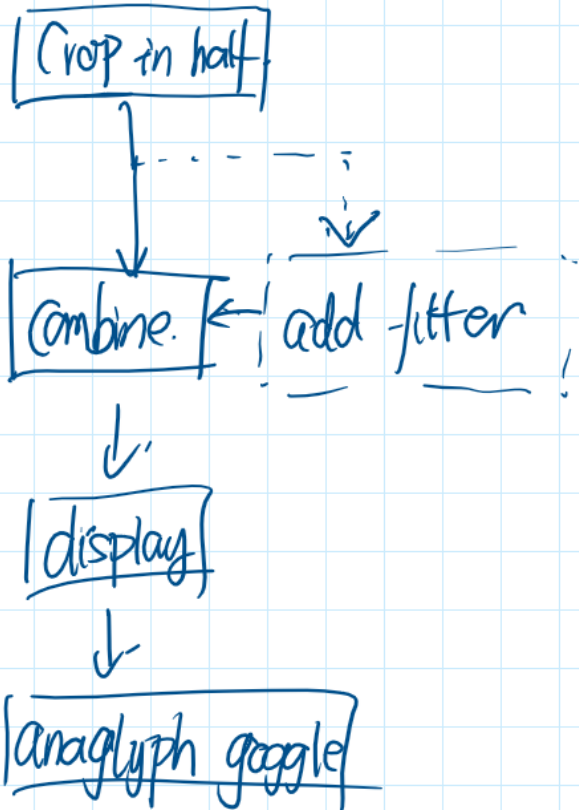
Design matrix

Friday, September 30, 2022

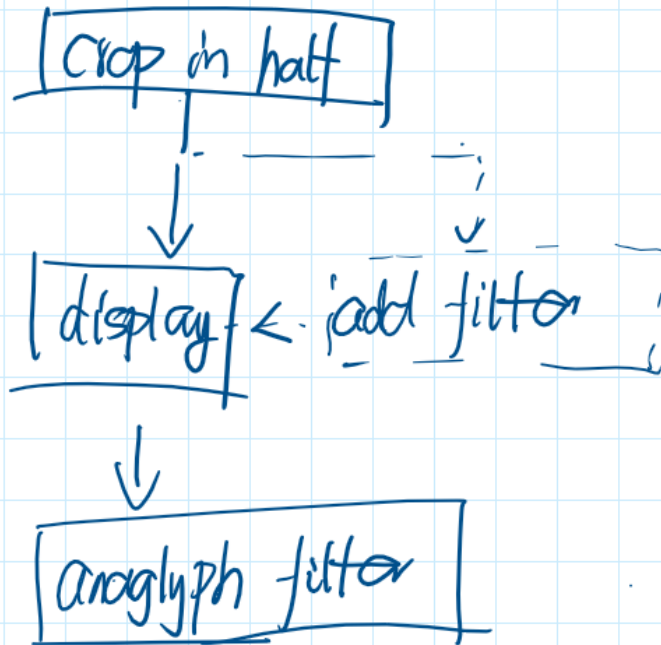
1:58 PM



① combined view



② shutter:



Note: the display here is an alternative between left and right eye view with $\text{fps} \geq 30$

Note: dotted line for adding filter can be either implemented as a physical filter attached to mirrors in the smartphone attachment, or it can be set as one of the video processing speed in python script

Conclusions/action items:

Discuss these ideas during the advisor meeting and team mates



2022/10/05 - Python script

HAOCHEN WANG - Oct 12, 2022, 12:13 AM CDT

Title: Python Script 1.0

Date: 10/05/22

Content by: Haochen Wang

Present: N/A

Goals: To develop the first python script for video processing

Content:

```
import cv2
import sys
import numpy as np
```

```
cap = cv2.VideoCapture(r'C:\Users\Wang-\Desktop\BME400\test.mp4')
```

```
while(cap.isOpened()):
    ret, frame = cap.read()
```

```
// add processing here
```

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

```
cap.release()
cv2.destroyAllWindows()
```

Conclusions/action items:

Tried loading the microsurgery practice video from last semester. There seemed to be high consumption of cpu/graphic card power in my laptop (used for 3 years). Need to have a closer look into the method calls to reduce computation requirement, which is to leave space for actual video processing.

**Title: Image and Video Processing in Python****Date:** 09/21/22**Content by:** Haochen Wang**Present:** N/A**Goals:** To learn how to do image processing with python**Content:**Website: <https://new.pythonforengineers.com/blog/python-for-scientists-and-engineers/>

1. Display an image
2. Blur and grayscale
 1. used gaussian blur
3. Edge detection
4. Count objects
 1. This seems to be the algorithm used by ImageJ
 2. Blue -> edge detection -> count
5. Face detection
 1. with cascades, from rough sections of blocks to more detailed check
 2. involves machine learning
6. *** use webcam
 1. Code:

```
1. import cv2
import sys

cascPath = "haarcascade_frontalface_default.xml"
face_cascade = cv2.CascadeClassifier(cascPath)
```

```
2. if len(sys.argv) < 2:
    video_capture = cv2.VideoCapture(0)
else:
    video_capture = cv2.VideoCapture(sys.argv[1])
```

```
3. while True:
    # Capture frame-by-frame
    ret, image = video_capture.read()
```

```
4. ret, image = video_capture.read()

    if not ret:
        break

    gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)

    faces = face_cascade.detectMultiScale(
        gray,
        scaleFactor = 1.2,
        minNeighbors = 5,
        minSize = (30,30)

    )

    #print("The number of faces found = ", len(faces))

    for (x,y,w,h) in faces:
        cv2.rectangle(image, (x,y), (x+h, y+h), (0, 255, 0), 2)

    cv2.imshow("Faces found", image)
```

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

```
6. # When everything is done, release the capture  
video_capture.release()  
cv2.destroyAllWindows()
```

```
7. python webcam_face_detect.py
```

```
8. python webcam_face_detect.py webcam.mp4
```

2. 3 => read returns the actual video frame read from each loop and a return code (which tells if we are running out of frames)

3. 5=> press key q to quit

4. 6=>> clean up

5. 7=> use webcam

6. 8=> view the video

7. Motion detection

1. Code

```

1.  #!/usr/bin/python

import cv2
import sys
import numpy as np

if len(sys.argv) < 2:
    video_capture = cv2.VideoCapture(0)
else:
    video_capture = cv2.VideoCapture(sys.argv[1])

2. # Read two frames, last and current, and convert current to gray.
ret, last_frame = video_capture.read()
ret, current_frame = video_capture.read()
gray = cv2.cvtColor(current_frame, cv2.COLOR_BGR2GRAY)

3. i = 0
while(True):
    # We want two frames- last and current, so that we can calculate the
    different between them.
    # Store the current frame as last_frame, and then read a new one
    last_frame = current_frame

4. ret, current_frame = video_capture.read()
gray = cv2.cvtColor(current_frame, cv2.COLOR_BGR2GRAY)

5. # Find the absolute difference between frames
diff = cv2.absdiff(last_frame, current_frame)

6. # If difference is greater than a threshold, that means motion
detected.
if np.mean(diff) > 10:
    print("Achtung! Motion detected.")

7. # Display the resulting frame
cv2.imshow('Video',diff)
if cv2.waitKey(1) & 0xFF == ord('q'):
    break
# When everything done, release the capture
video_capture.release()
cv2.destroyAllWindows()

```

2. 5=> note the use of cv2.absdiff to compare differences

Conclusions/action items:

1. check import numpy and a good IDE (like ancorda)
- 2.



Title: Photo inpainting python tutorial

Date: 09/23/22

Content by: Haochen Wang

Present: N/A

Goals: To understand the python algorithm for converting photos into 3D-like photos

Content:

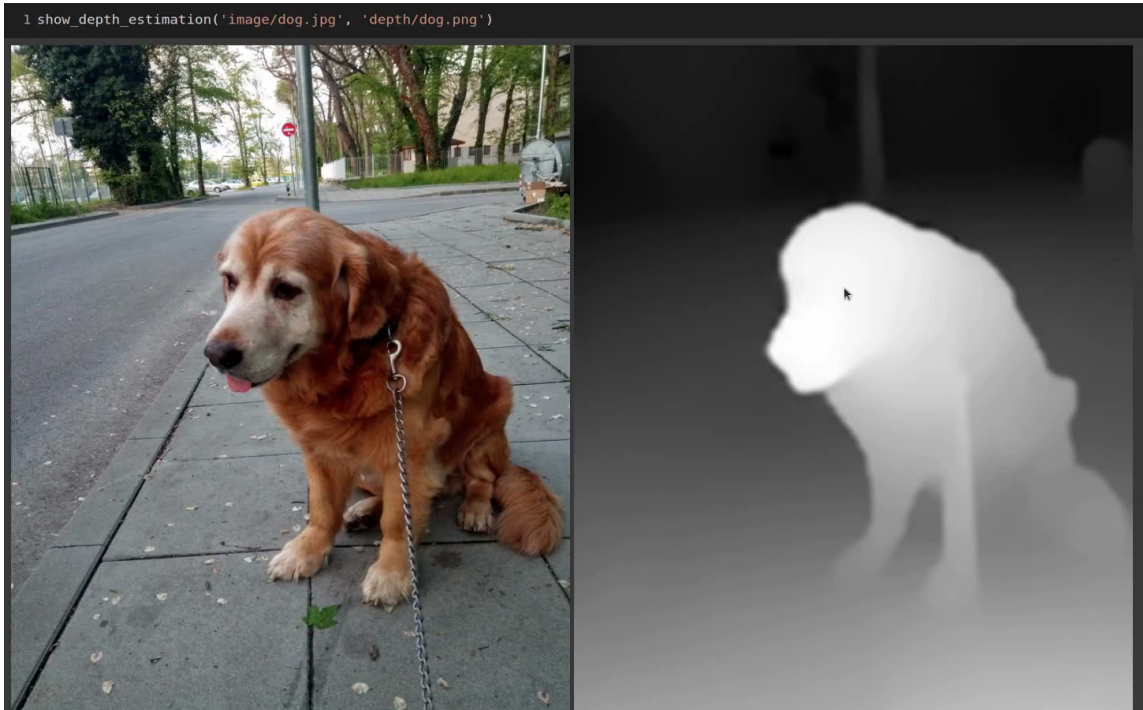
Video link: <https://youtu.be/x54GdkpW9JY>

Related article: 3D photography using context-aware layered depth inpainting, by Meng-Li Shih et al.

1. Install dependencies

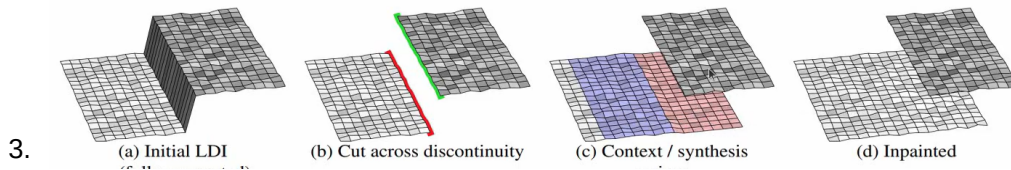
1. !pip install -q visp==0.6.4
2. !pip install -q moviepy==1.0.2
3. !pip install -q transforms3d==0.3.1
4. !pip install -q network==2.3
5. !pip install -q -U watermark
- 6.
7. !sudo apt install sed

2. the Midas package on gitHub includes functions for estimating depths (from the Intel-isl group)





- 2.
- 3. whiter pixel in the depth detection means that the depth is lowered (closer to the camera)
- 4. Can use MiDaS to obtained RGB-D images
 - 1. Could the mirrored view be used as the D-detection?
 - 2. The image quality from the mirror is really not as good, so probably can just serve as D-factor adding onto the direct view, which has better focusing and resolution



Conclusions/action items:

Consider how D-factor should be obtained and incorporated to generate the 3D effect

The processing time from the video demo, for one frame only, takes about 10 seconds. So a possible solution is to generate the depth map at the beginning of the microsurgery practice, with background, station and the skin/blood vessel only. Then keep using this depth map as the depth for all following frames.

***Check out the original paper from the authors of the code

2022/09/29-3D photography using context-aware layered depth inpainting

HAOCHEN WANG - Sep 30, 2022, 10:25 AM CDT

Title: 3D photography using context-aware layered depth inpainting

Date: 09/29/22

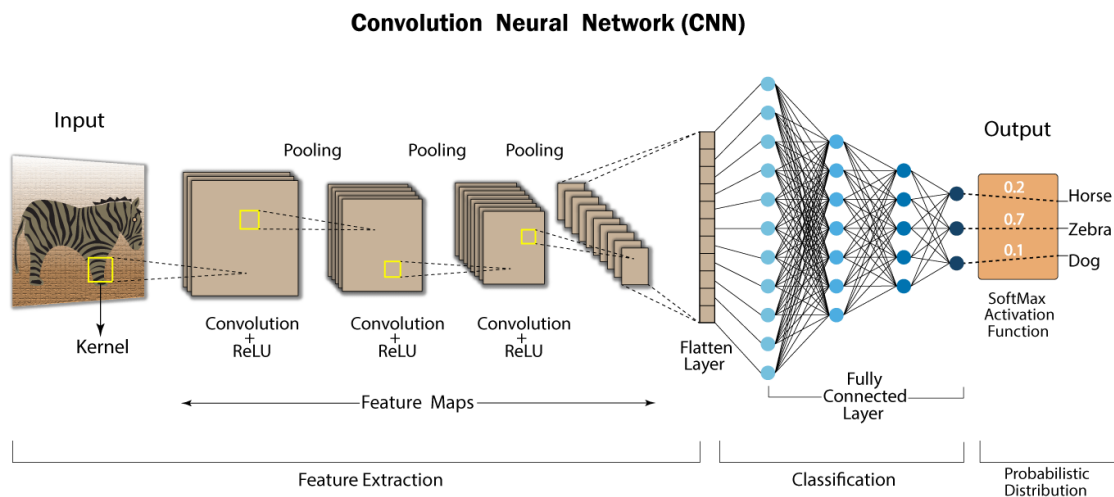
Content by: Haochen Wang

Present: N/A

Goals: To understand the mechanism of inpainting

Content:

1. Facebook 3D photos require capturing a single snapshot with a dual lens camera phone, which essentially provides an RGB-D input image
 1. uses layered depth image (LDI), converted into a light-weight mesh representation for rendering
 2. isotropic diffusion algorithm for inpainting colors
 3. but result in overly smooth results
2. Multi-plane representation (MPI), synthesized from the small-baseline dual camera stereo input, but produces artifacts on sloped surfaces
 1. also memory and storage inefficient
 2. costly
3. The paper uses algorithm based on spatially adaptive context regions, designing to the local depth-complexity of input and generates a varying number of layers across the image
4. View synthesis: light field, multi-plane image representation (fixed depth discretization, so not working for sloped surfaces), layered depth images [[Jonathan Shade, Steven Gortler, Li-wei He, and Richard Szeliski. Layered depth images. In Proceedings of the 25th annual conference on Computer graphics and interactive techniques, pages 231–242. ACM, 1998.](#)] can locally adapt to any depth complexity
5. Image based rendering: need sufficiently large baselines --> learning-based blending, soft 3D reconstruction, handling reflection, relighting, and reconstructing mirror and glass surfaces
6. Learning based view synthesis: most current work require pre-trained network
7. Image inpainting: CNN-based architectures (convolutional neural network)
 1. Found this image on google to explain that it is



2.

8.

Conclusions/action items:

check these two articles (lowkey 3D capture)

<https://dl.acm.org/doi/10.1145/3197517.3201384>

<http://visual.cs.ucl.ac.uk/pubs/casual3d/>



2022/10/02- Potential Ideas

ALEXANDER VAZQUEZ - Oct 12, 2022, 2:13 PM CDT

Title: Flow of Data

Date: 10/2/22

Content by: Alex

Present: Alex

Goals: Come up with some potential flow ideas

Content:

Previous Team Idea:

- Image split by single pair of mirrors
- Image captured by iPhone camera
- Video feed sent over zoom call to user

Alternative Design 1:

- Image split by single pair of mirrors spaced rough human pupillary distance
- Image captured by phone camera
- Feed sent wirelessly to compute device where post processing occurs
- Device directly connected to display, method of switching between monitor display or google cardboard display

Alternative Design 2:

- Image split by single pair of mirrors spaced rough human pupillary distance
- Image captured by iPhone
- Code executed to add anaglyph filter over captured image
- Send to a display via a cord to reduce latency
- User wears anaglyph filter glasses to perceive depth

Alternative Design 3:

- Video capture done with iPhone Lidar sensor
- Use raw 3D image and stream/hardware to display directly
 - Alternatively we could use the result to create a depth mask give depth to the video feed
- Add any additional computer or other step to facilitate quick image processing

Conclusions/action items:

There are various ways that the clients conditions can be met. More research into the topics can give us insight into which is the most effective/realistic solution.



2022/09/12- Background Project Research

ALEXANDER VAZQUEZ - Oct 12, 2022, 8:59 AM CDT

Title: Project Catch Up

Date: 9/12/22

Content by: Alex

Present: Alex

Goals: Go through past material to get a better understanding of the project scope and what needs to be done.

Content:

Document 1: Final Report

Abstract/Introduction

- **Problem:** Medical students travel to larger hospitals to train with micro surgical equipment. This standard for training medical students is inconvenient and has become more difficult due to travel restrictions.

Background

- Surgical Microscopes:
 - Offer high resolution and depth perception
 - High magnification
 - Expensive, \$200k-\$1m
 - Sources Viewed: [1][2][3]

Preliminary Designs

- Grey pro resin: High accuracy, rigid material, inexpensive (Winner)
- Tough 1500 Resin: High accuracy, weaker material, not ideal for holding mirrors.
- Acrylic (Laser Cut): Highest accuracy, strongest material, expensive, difficult to assemble

Final Design

- Grey Pro Resin
- 1 mirror pair for image capture
- Image capture projected a outside display, or second smartphone in google cardboard

Results

- The final prototype completed several goals of the project; depth perception, high resolution/quality image, visible sutures
- Image not seamless, high delay in relaying picture
- Inaccurate pupil distance, adjustments could improve functionality of stereoscopic view

Thoughts:

- Main improvements that we should make are to the design and orientation of the mirrors/ housing
 - Simulate human pupillary distance
 - Ensure design is sturdy to prevent shifting
- Improve system of merging images to create a more seamless picture

- Find a streaming method that offers lower latency
- **New ideas:**
 - Different Methods of displaying picture (Multiple output options for flexibility)
 -

Document 2: PDS

- **Budget:** 500\$ not including; iPhone, stand, and attachment for camera
- Other criteria similar to what was in the report

Conclusions/action items:

I think I have a better idea of the goals of this project and where the team got last year. I think we can make meaningful improvements to the design this semester and provide a viable design for a training setting. Next, we should meet as a group to discuss how w want to proceed this semester and what design ideas we have.



2022/09/24-iPhone Design Research

ALEXANDER VAZQUEZ - Oct 12, 2022, 10:06 AM CDT

Title: iPhone Design Understanding

Date: 9/25/22

Content by: Alex

Present: Alex

Goals: Gain a better understanding of the design of the iPhone camera system as well as how these features can be leveraged for our design needs.

Content:

iPhone 13 Pro Camera:

- Offers 12MP camera with three different camera types. Telephoto, Wide and Ultra Wide
- Portrait mode that offers Depth Control
- Can record up to 4k 60 fps
- Offers a very powerful photography device in a small package

Biggest benefit from the iPhone is that it is a small but very powerful package. Its connectability makes it easier to move the video feed to different sources. The powerful camera and the Lidar sensor could prove to be useful.

Source:

“iPhone 13 Pro - Technical Specifications.” https://support.apple.com/kb/SP852?locale=en_US (accessed Oct. 24, 2022).

- iPhone Lidar detectors can be used to develop a 3D maps of objects or rooms:
 - Could be used to create a depth mask over the image.
- Laser scans millions of points and return the information to the sensor
- *Look into Lidar sensor applications: “Photogrammetric” and “Point Cloud”
- May be limited to pictures depending on process time

Source:

I. Aslan and N. Polat, “Availability of Iphone 13 Pro Laser Data in 3D Modeling,” *Advanced LiDAR*, vol. 2, no. 1, Art. no. 1, Mar. 2022.

Conclusions/action items:

The iPhone camera is indeed powerful enough to capture the quality image that our client needs. The main concern is the sense of depth. A split view can do this but the Lidar sensor on the iPhone could also prove to be a beneficial tool.



Title:

Date:

Content by:

Present:

Goals:

Content:

Conclusions/action items: