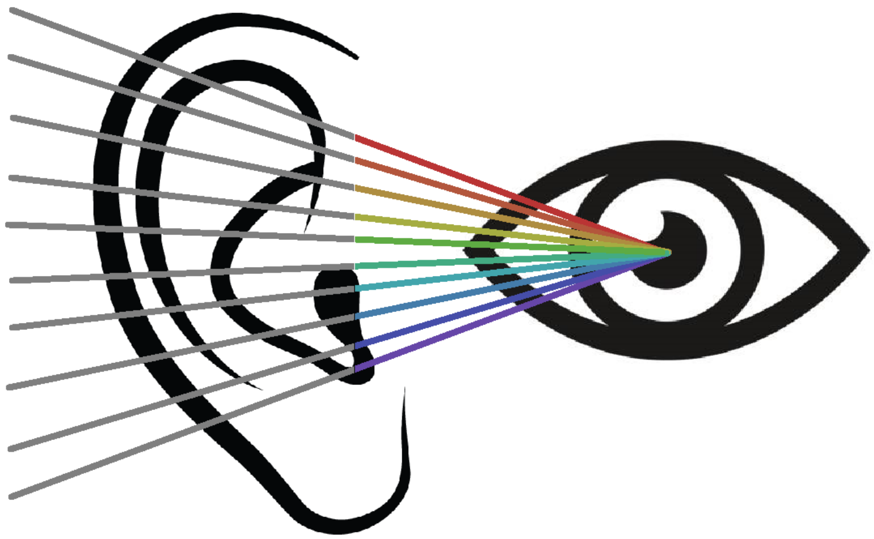
Initial Project Document V2

Chromesthesia Simulation Device



Department of Electrical Engineering and Computer Science

University of Central Florida

Dr. Lei Wei and Samuel Richie

Group 27

Angel Garcia	angelgarcia@knights.ucf.edu	CE
Brooke Roeder	brookeroeder@knights.ucf.edu	EE
Nicholas Alban	nickalban@knights.ucf.edu	EE
Wesley Ellery	ellerywesley@knights.ucf.edu	CE

Date: 10/10/2021

1. Project Description

This project aims to simulate/recreate Chromesthesia - an audio-visual variant of the phenomenon called Synesthesia. Synesthesia is when the stimulation of one sensory or cognitive pathway leads to involuntary experiences in a second sensory or cognitive pathway. Chromesthesia is the involuntary association of sounds to visual effects such as shapes and colors.

The aim of this project is to make a device that will take audio and visual inputs and produce a visual output that simulates/recreates the visual experiences that a person with Chromesthesia would experience. The current plan is to combine audio input devices (i.e. microphones) and an optical capture device (i.e. a camera) on a single platform that can be oriented and pointed in any direction. This platform will receive audio and visual inputs from the direction it is pointing, and then perform a computational process to recreate the same imagery from the visual input edited with visual effects influenced by the audio input. The visual output will be displayed to a screen for the user to view. Ideally, the platform containing the audio and visual input devices, the PCB/computational hardware, and the display will all be integrated together into one device that the user can carry and thus, aim the microphones and camera in any direction and receive a corresponding visual output, as if looking around with their own eyes and experiencing Chromesthesia.

There are two designs in mind - a worn design, and a handheld design. The two designs are fundamentally the same, but the worn design will be in the form of a headset the user wears and aims by turning their head, and the handheld design would be wimed with the user's hand(s) - similar to a camcorder.

If the device is to be wearable, it would need to be lightweight. It should be comfortable to wear and easy to use - ideally, the user would only have to look around and aim with their head to operate the device. The device could be portable, however, power demands have not yet been determined, so power supply might be too great to use a battery. Because creating a visual display that is wearable is a large task, we reserve the possibility of making the device not wearable, and instead having the output display and input platform separate. In this case, the user would aim the platform with their hands and see the correlating output displayed on a stationary screen. At the moment, this is only a possibility if creating a wearable device becomes too complicated, expensive, or time-consuming.

Since this project will have a complicated processing algorithm, it might require something with more processing power than average microcontrollers unless we want to spend a lot of time optimizing it for the memory and clock cycles. Instead, we are looking towards utilizing "Single board computers" (i.e. SoC or "System on Chips" such as a Raspberry Pi since they are more powerful and have the capabilities to process video easily.

This project is an exciting feat of engineering to pursue; Virtual/Augmented reality, computer vision, image processing, and electronic entertainment are all new fields of engineering. Exploring these fields would not only be an exciting challenge but also provide useful skills and experience for industry. In addition to the technical intrigue, another motivation is that only a fraction of the population experiences some sort of synesthesia. This project would enable people, who would otherwise be unable, to experience Chromesthesia, and educate them about the phenomenon.

2. Project Goals and Objectives

This section outlines the main goals and objectives that our team wishes to accomplish in our device. These will be the main ideas and concepts guiding our design choices as the project develops and nears completion. These goals and objectives will be constantly referenced and used to determine important design choices such as what component to buy and what functions should be prioritized.

2.1. Accuracy - In order to simulate Chromesthesia, our device needs to be accurate in determining audio frequencies and be able to determine the exact location of the audio source. Any inaccuracies will cause incorrect visuals and disrupt the recreation of the phenomena. Processing of the audio inputs will be especially important.

2.2. Speed - the device needs to produce an image in real time. Ideally, the time between the user aiming the device and the device creating an output image will be as short as possible to simulate “looking” with their own eyes. This means the processes need to be fast and efficient. This will be heavily dependent on the software optimization and hardware design and quality.

2.3. Software - The software must be able to communicate with the microprocessor and be able to perform complicated algorithms including noise processing, 3D sound localization, and image processing. These algorithms must be as quick and efficient as possible to determine what color correlates to the audio frequency and how it should be displayed and produce a near real-time picture.

2.4. Hardware - The hardware will consist of a microprocessor, multiple microphones, and a camera able to display images of quality 720p or higher. These will all be the key components to the Chromesthesia device. The hardware should be chosen carefully to optimally perform the calculations needed to process audio and visual inputs so the device can operate at the desired speed.

2.5. Convenience - Ideally, we aim for the final device to be lightweight and wearable. Our vision for the project is for it to be a pair of goggles or any other similar item that you can wear and one would be able to see through them and experience Chromesthesia. Alternatively, the device will be handheld and aimed with the user's hands like a camcorder. Both designs demand ease of use with minimal inputs from the user (i.e. the user only needs to aim the device with their head or hands) and a light weight (ideally under 5 lbs).

2.6. Experience - The device should adequately represent the phenomenon of chromesthesia, and the final visual output should have a semblance of the experience, as if "through the eyes" of someone with chromesthesia. Chromesthesia varies between people, and it has nuances in how it is experienced. We hope to do some research and get a better definition of how exactly we will represent chromesthesia visually on the display. Ultimately, we hope to recreate the phenomenon in a way that has the perfect balance between accurate and fun. People should be excited to use the device and want to experiment aiming it at different sound sources with different pitches.

3. Features

Our device will have three defined features. This section will go into detail what each feature is and why it is important to implement.

3.1. Basic features

Basic features define the minimum functionality of the device. For this project, there are two basic features; one is the ability to process audio inputs with a microprocessor and perform a basic algorithm that can display colors correlating to the frequency of the soundwaves from the audio input. The second basic feature is the ability to capture visual images in real time, parallel with detecting and processing the audio inputs.

PURE SONOCHROMATIC SCALE		
(invisible)	Ultraviolet	Over 717.591 Hz
	Violet	607.542 Hz
	Blue	573.891 Hz
	Cyan	551.154 Hz
	Green	478.394 Hz
	Yellow	462.023 Hz
	Orange	440.195 Hz
	Red	363.797 Hz
(invisible)	Infrared	Below 363.797 Hz

SONOCHROMATIC MUSIC SCALE (basic 12/360)		
	Rose	**E**
	Magenta	**D#**
	Violet	**D**
	Blue	**C#**
	Azure	**C**
	Cyan	**B**
	Spring	**A#**
	Green	**A**
	Chartreuse	**G#**
	Yellow	**G**
	Orange	**F#**
	Red	**F**

Figure 1: Example of correlation between wavelength of sound and color

Figure 2: Example of captured image from camera

3.2. Extended features

Extended features build upon the basic features and add some sophistication to the device's abilities. The next logical step from the aforementioned basic features would be to create some sort of neural network or other algorithm to better process the audio inputs, allowing the device to perform 3D sound localization and combat noise. Additionally, a visual output should be produced, which combines the raw visual input from the camera with the color determined by the algorithm that uses the frequency to determine the color. Figure 2 is a theoretical example of what the output would be, if the theoretical audio input's frequency was calculated to correspond to the color blue.

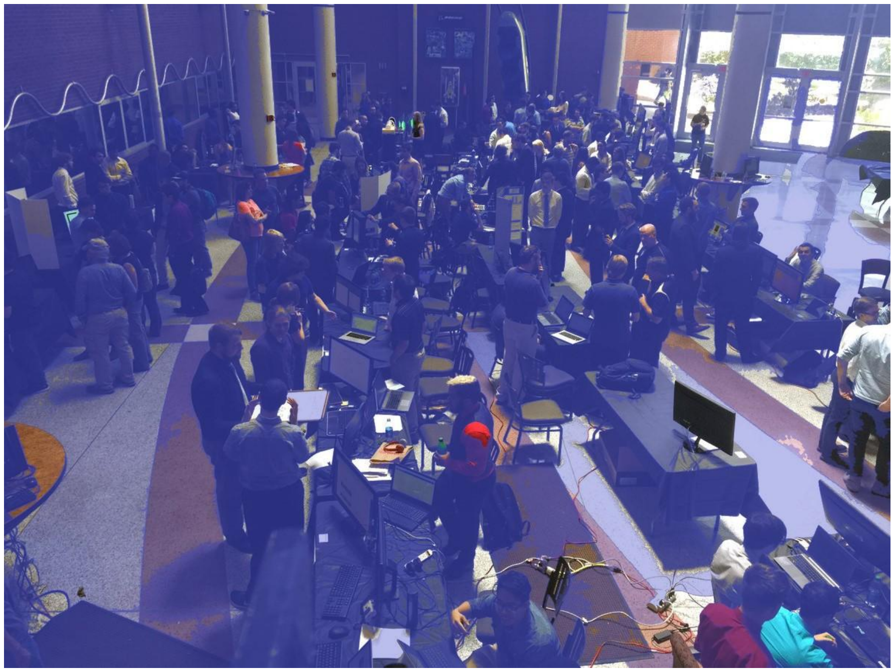


Figure 3: Example of outputting color of detected sound on same output image as visual input

3.3. Advanced features

Advanced features go above and beyond what is the fundamental functionality of the device, further sophisticating existing features, or adding additional advanced features. Here, we have considered using 3D sound localization to determine the spatial location of the sound sources in the acquired image, and overlaying the corresponding frequency color to that location. Figure 4 demonstrates what this might look like. This might not be a realistic example, but it demonstrates how different sounds will be coming from different locations and the frequency of those different sound sources would display a different color.

Figure 4: Example of outputting color of detected sound on same output image dependent on location in image

4. Designs

There are currently two designs being considered. Both designs are functionally the same; the device is aimed to collect audio and visual inputs, the inputs are processed, and then a visual output is generated and displayed in real time to allow the user to experience chromesthesia. Both designs will satisfy the project goals and objectives, and will have the same basic, extended, and advanced features. The main difference between the two designs is how the user interacts with the device.

4.1. Design A

4.1.1. Design description:

A custom platform will be made to hold the microprocessor, microphones, and camera. The platform will be shaped for handheld use so the user can use their hands to aim it. Whichever direction the user aims the device, the microphones and camera will obtain inputs dependent on the direction of aiming. The microphones will be positioned on the platform in a way that 3D sound localization can be performed. Once It processes the audio, the audio and visual inputs will be sent to the microprocessor, which will determine what color will correlate to the frequency of the audio input, and generate an augmented visual output of the captured scene. The display screen may or may not be fixed to the device. This will be dependent on the weight and cost of the screen itself. A smaller screen will be more expensive but potentially enable our device to fit onto one handheld platform. A larger, more traditionally sized screen will be less expensive, but heavier. A heavy screen will most likely need to be placed on a table. In the case of a stationary screen, the device will still function as intended, however, it will diminish the experience of “looking” around with the device.

4.1.2. Pros and Cons of Design A:

Pros of Design A	*Cons of Design A*
Platform design will be more appropriate for our skill levels	Less immersive
Simpler platform design	Potentially heavier screen
Easier to pass between users	
Easier to troubleshoot real-time	

Table 1: Pros and Cons Table of Design A

4.2. Design B

4.2.1. Design description:

Design B is our aim, and it will be pursued if everything is on time as expected. Everything from Design A will be translated to the context of a pair of VR goggles or some sort of wearable vision device. Ideally, the device will use some third party wearable screen display and instead of the microphone/camera/PCB platform being designed to be handheld, it will be attached to the goggles. This way, wherever the user physically aims their head, the device will collect audio and visual inputs from that direction. This is the ideal design for the device because it will maximize the experience by making users feel like they are looking and experiencing chromesthesia with their own eyes and senses.

4.2.2. Pros and Cons of Design B:

Pros of Design B	*Cons of Design B*
Much more immersive	Need to interface with third party VR headset
No need to create ergonomic platform unlike Design A	More expensive (VR headset)
	Potential issues connecting platform to headset
	Concerns with sanitization when switching between users

Table 2: Pros and cons Table of Design B

5. Constraints & Requirements

Table 3 contains all the constraints and requirements that need to be met and considered during the development of the project. These requirements all take into consideration the details discussed in the aforementioned designs A and B, and the goals and objectives of the device.

5.1. Wearable devices should be as lightweight as possible. Industry VR headsets range from 0.84 lbs to 2.5 lbs. Because these are mass-produced and professionally designed, we can arbitrarily set a limit to double their standards: 5 lbs. This will be our aim in weight limitations +/- 1 lbs

5.2. 4 I2S microphones

5.3. 1 camera

5.4. Camera resolution = 720p/1080p

5.5. Humans can hear 20 Hz - 20 kHz, microphones will need to be able to detect in a similar range.

5.6. Noise will be a big factor for the microphones, especially since presentation day will take place in a loud atrium with hundreds of people talking. Audio processing will arguably be the most important factor in the signal processing ability of our device.

5.7. Microphones come in a variety of “cones”. Since we plan on having an aimable device, we will probably need to use microphones with a narrower cone. Ideally, the microphones will have a field of view identical to the camera.

5.8. 1 display screen of identical resolution to the camera (720p/1080p)

  5.8.1. Will either be a small display if wearable or a full sized computer screen if stationary.

5.9. 3D printed platform to hold the microphones and camera

5.10. 1 PCB to do some type of functionality

5.11. Utilize a microprocessor that can do the necessary calculations for our algorithm - must be able to do audio processing and potentially image processing.

  5.11.1. One viable option is the Raspberry Pi 4. There are three different versions of RAM: 2GB, 4GB, and 8GB LPDDR4. It features a quad-core 64-bit ARM Cortex-A72 CPU and is a high-performance general-purpose computer. The processor speed is 1.5Ghz. Other key features are: two micro-HDMI ports that are capable of supporting resolutions up to 4K, two USB3.0 and two USB2.0 ports.

5.12. The total cost of the project is currently unknown due to not knowing exactly what the requirements are and what materials we will be needing. The potential for the device to go over our budget is definitely possible depending on the microprocessor we will be using, but there is a high chance that we will be able to stay on budget.

6. Block Diagrams

6.1. Block Diagram of Whole Project

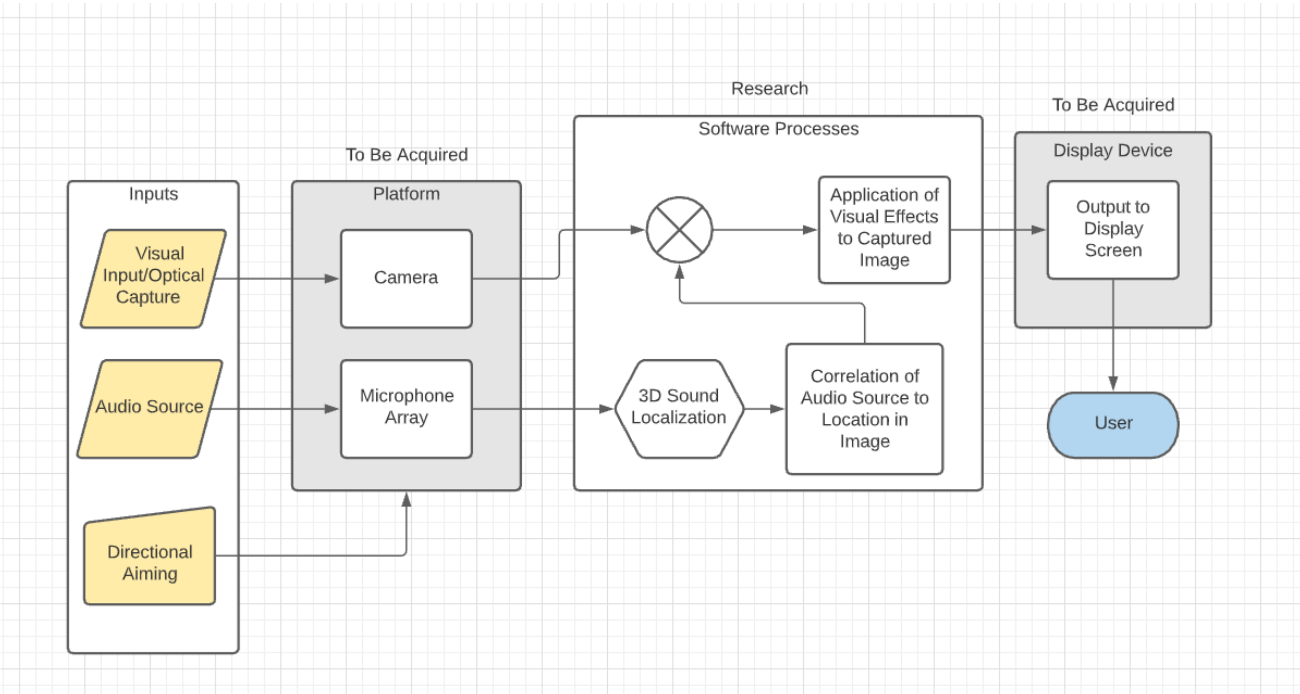


Figure 5: Block Diagram Overview

6.2. Block Diagram of Software

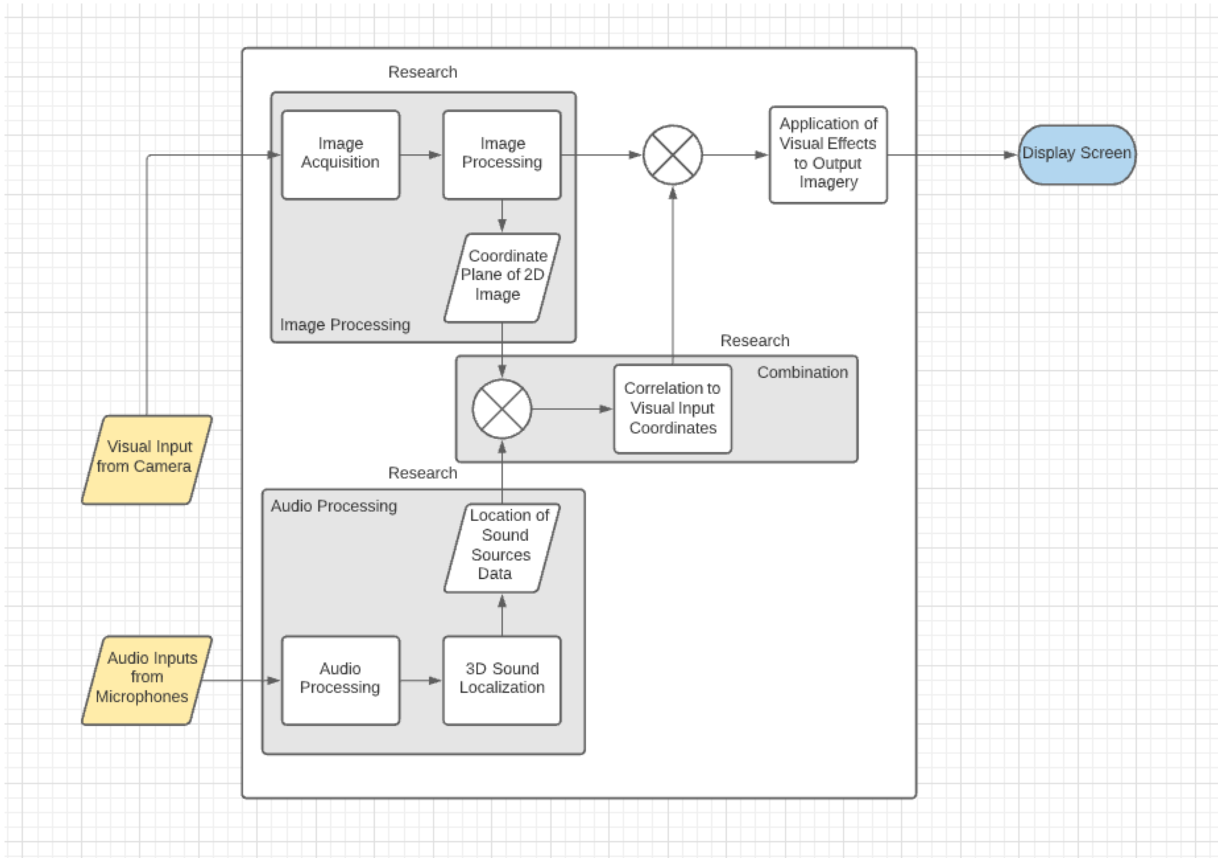
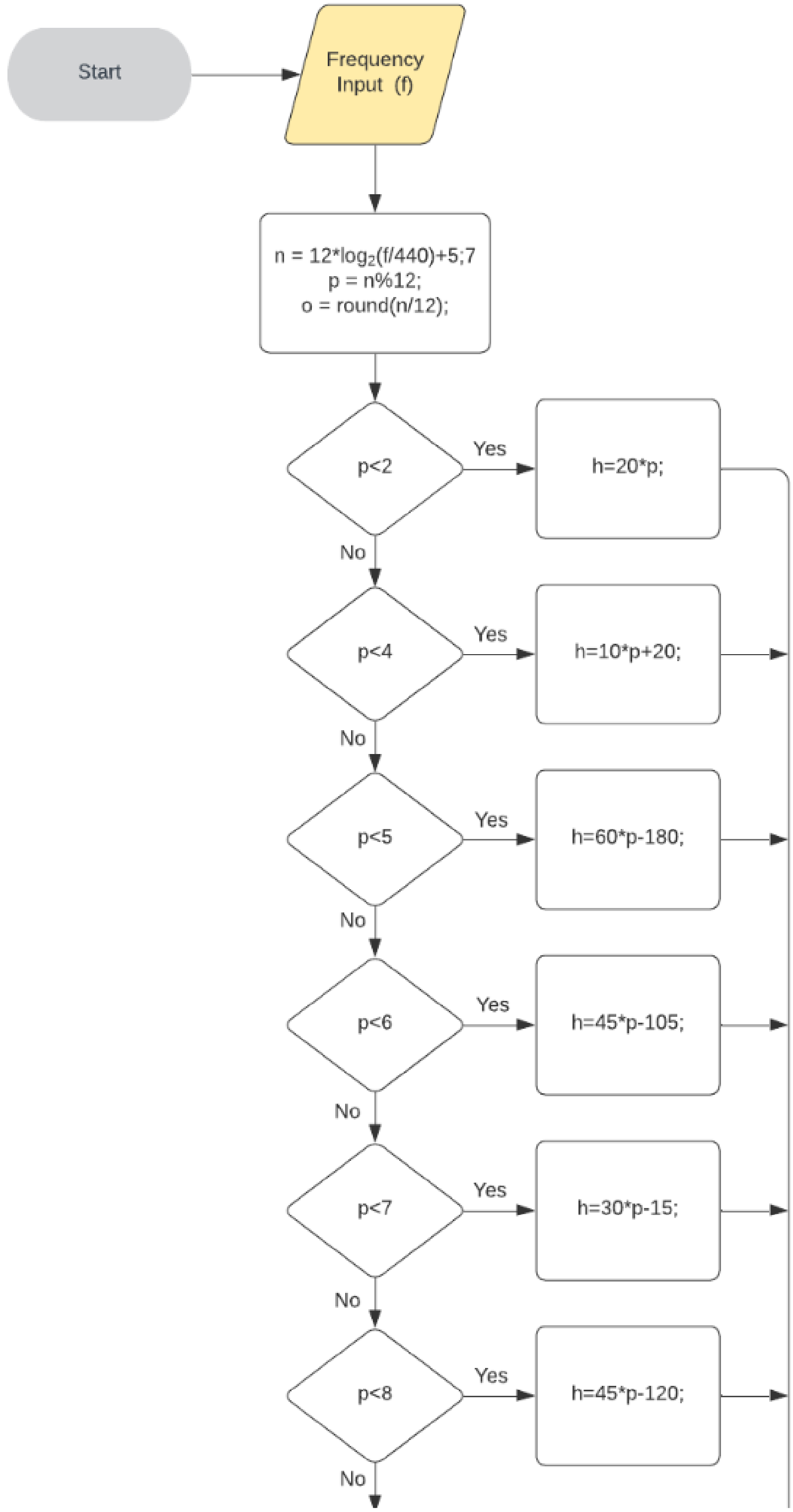


Figure 6: Block Diagram Software Overview

6.3. Block Diagram of Frequency-to-Color Algorithm



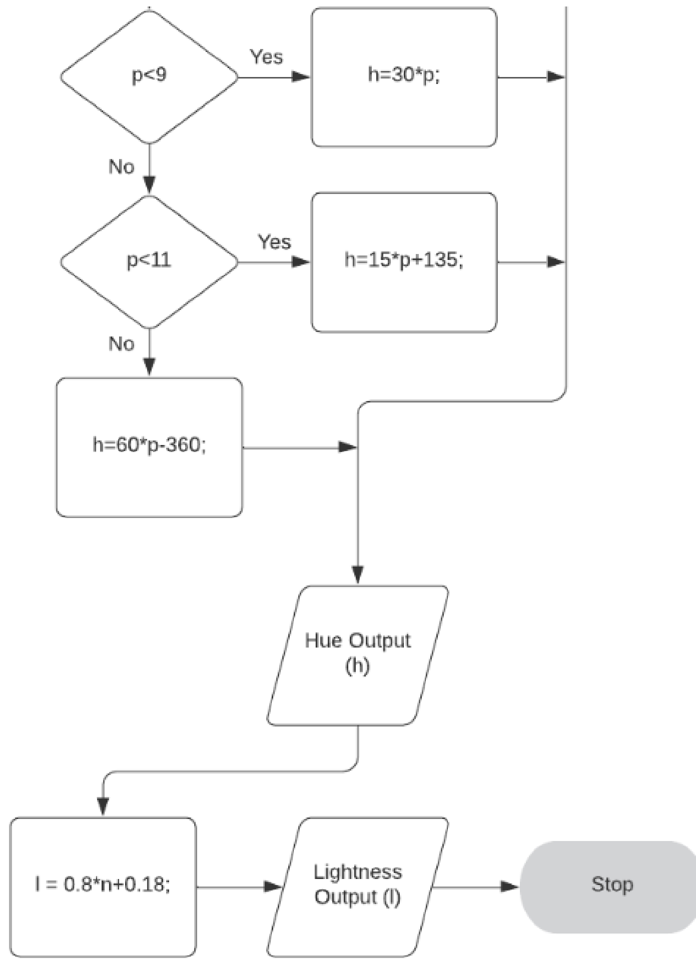


Figure 7: Block Diagram of frequency to color algorithm

6.4. Key for Block Diagram

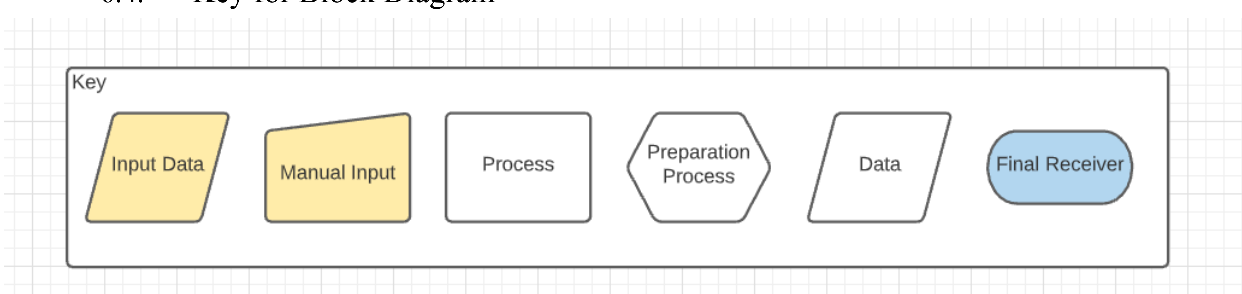


Figure 8: Block Diagram Key

6.5. List of roles categorized by group member

Member	Leading	Co-Leading/Assisting
Angel Garcia	• PCB • Microcontroller	• Product Hunting and Identification
Brooke Roeder	• Pitch-color algorithm method • Image Processing	• Testing and review processes
Nicholas Alban	• Audio Processing software • Constructing platform	• PCB
Wesley Ellery	• 3D sound localization	• Testing and review processes

Table 3: List of roles by group member

7. Research

7.1. Pitch to Color Algorithm

A 2017 study published in nature.com titled "Musical pitch classes have rainbow hues in pitch class-color synesthesia" was the basis for my research and algorithm development. The study was done on 15 subjects who possessed "pitch class-color synesthesia" and had moderate to high levels of absolute pitch.

Case number	Age	Sex	Music Training (ages)	AP Score	Colored hearing	Trigger	Acquired
1	22	f	5-18	97	yes	syllable, pitch	Painted colors on notes
2	22	f	2-22	100	yes	syllable, pitch	no idea
3	22	m	7-22	100	yes	pitch	no idea
4	21	f	5-21	86	yes	syllable	no idea
5	18	f	5-15	100	yes	syllable, pitch	no idea

6	21	f	6-20	100	yes	syllable	no idea
7	19	f	4-19	86	yes	syllable, pitch (inconsistent)	color stickers on keyboard
8	21	f	5-18	100	yes	syllable	color stickers on keyboard
9	20	f	4-20	100	yes	syllable	color stickers on keyboard
10	19	f	3-18	86	yes	pitch	no idea
11	19	f	11-19	100	yes	syllable	color stickers on keyboard
12	20	f	3-20	91	yes	syllable, pitch keyboard key	color stickers on keyboard
13	20	m	4-17	91	no	syllable	no idea
14	18	f	4-18	94	no	syllable	no idea
15	20	f	6-14	80	no	syllable	no idea

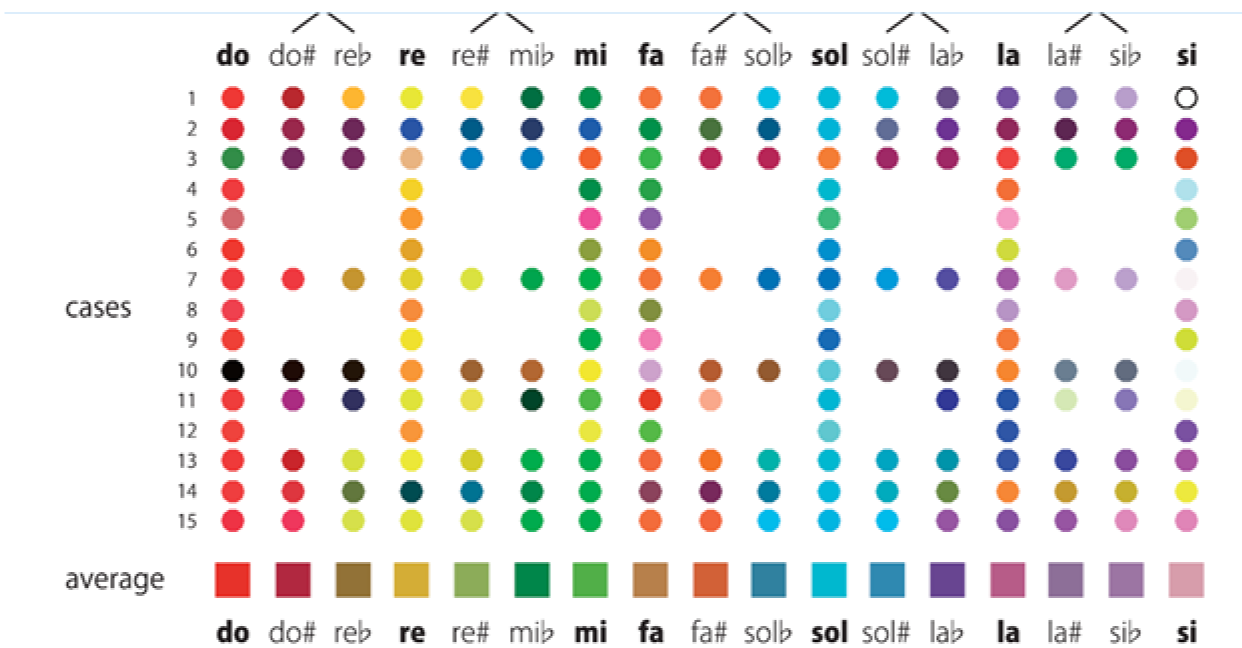


Figure 9: correlation of color to pitch chart

Observations: There are many similarities between subjects, especially those who had colored keys on their keyboard when learning music. Overall, the different pitches and color associations produce a rainbow.

Case 12:

The subject has an AP score of 91, is triggered by both the syllable and pitch, and acquired synesthesia from colored keys. This case seemed to best follow an increasing hue while the other pitch-triggered subjects who acquired synesthesia in a similar way have a sudden change in hue around “fa”.

Pitch #	Note	Color	R	G	B	H
1	C		239	67	61	2
3	D		248	150	58	29
5	E		233	231	65	59
6	F		85	185	71	113
8	G		96	199	207	184
10	A		47	86	166	220
12	B		122	81	161	271

Figure 10: Hue Graph

The case 12 pitch to color correlation is similar to D. D. Jameson's color scale, published in his book "Colour Music".

Colours and Semi-Colours, in their order.	
Red.	
	Red-orange.
Orange.	
	Orange-yellow.
Yellow.	
Green.	
	Green-blue.
Blue.	
	Blue-purple.
Purple.	
	Purple-violet.
Violet.	

Replacing Violet with Magenta creates the following color scheme:

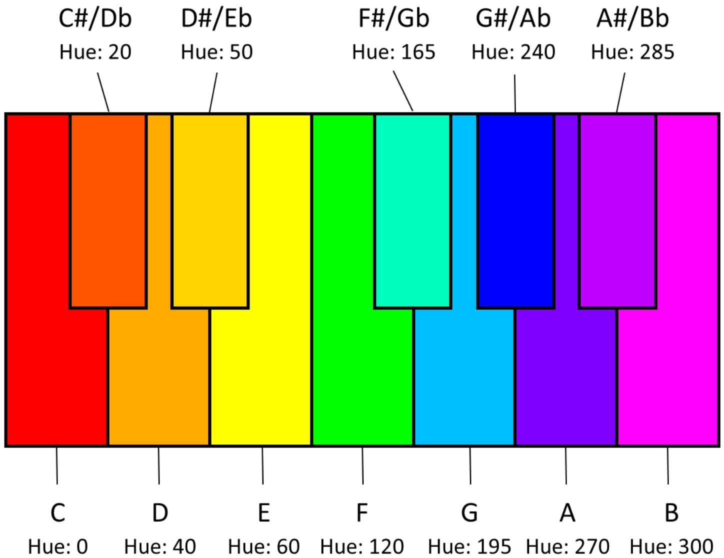


Figure 11: Hues correlating to piano keys

7.2. Components

Here is some research and identification of exact components and other important information we have acquired that will influence and be applied towards the development of the project.

Since the device is going to have complicated processing algorithms, it might require something with more processing power than average microcontrollers. Otherwise, we would need to spend a lot of time optimizing for the memory/clock cycles. After researching microprocessors, a few promising ones were found. Some options include a Raspberry Pi 4, a STM32F3 series MCU, and a BPI-M2 Zero board. These are much better at processing video which is something the device will definitely do. For this reason, a Raspberry Pi 4 will be used as a placeholder for price range.

8. Estimated Budget

Below in Table 4 is an itemized list of identified and anticipated equipment we will need to acquire for this project, along with their anticipated prices. Some items have been determined and have exact prices listed, while other items do not. This means we have yet to determine an exact model, but we know we will need this component and based on our evaluation of the market, we can provide a price range.

Item	Price
Adafruit I2S MEMS Microphones	$6.95
Arducam 5MP Camera for Raspberry Pi	$10
PCB	$10-$50
Raspberry Pi 4	$100
Display	$30

Table 4: Itemized list of materials

9. Initial Milestones

9.1. Need to figure out how exactly to represent the synesthesia - will the screen change color? Are we going to have images/shapes permeate from the source of the sounds on screen? We need to figure out how closely we want to simulate synesthesia, and choose a way that is complex enough to be technically challenging, but also achievable within our constraints.

9.2. Determine a frequency to color algorithm

9.3. Determine the exact hardware we will need for the device

9.4. Create 2D audio localization as precursor for 3D sound localization

9.5. Initial breadboard design

9.6. Determine the requirements for the microcontroller - how advanced is the audio and image processing, how demanding will the graphical display be, etc.

10. House of Quality

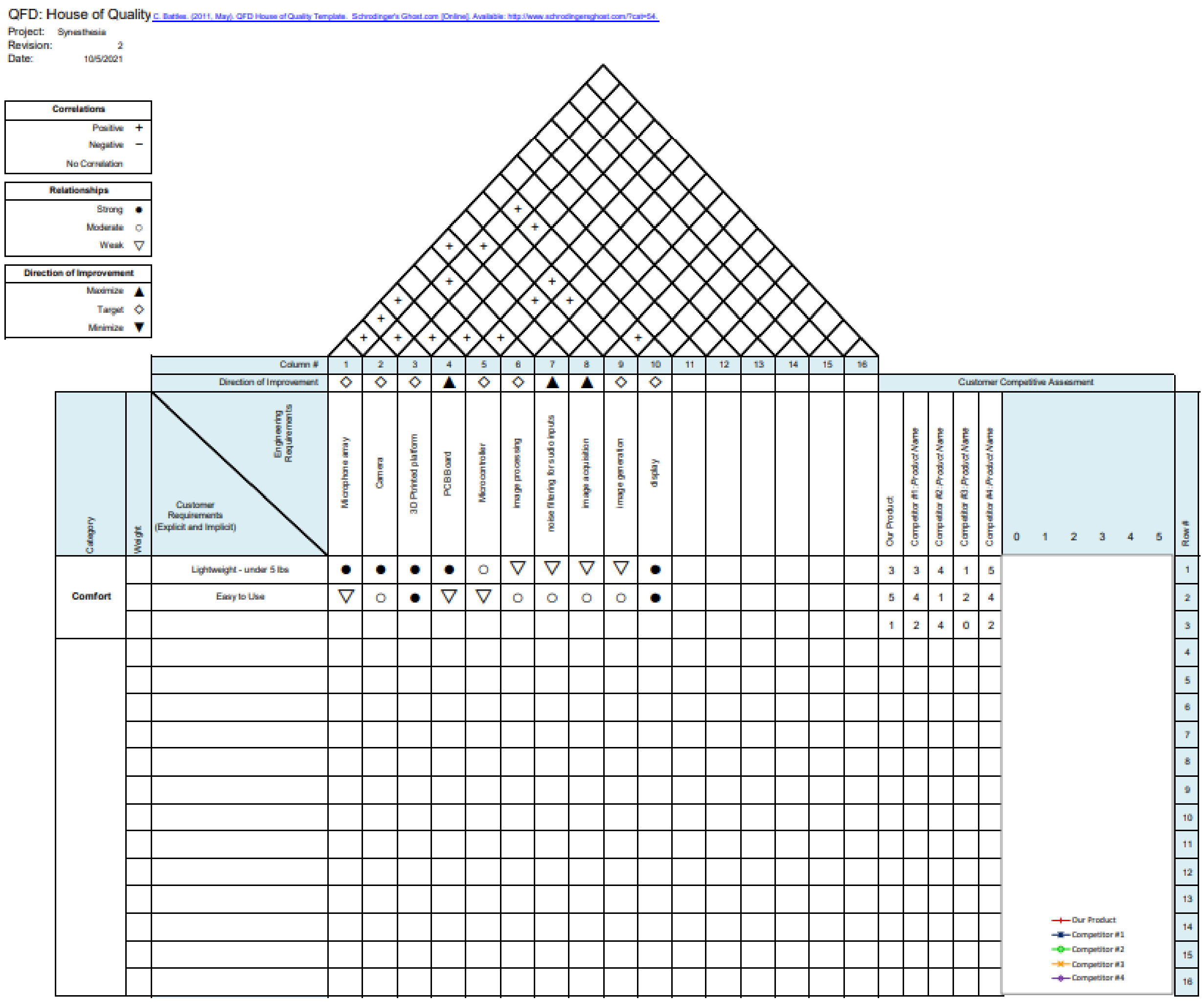


Figure 12: HOQ Diagram

11. Related Standards and Design Constraints

Standards are rules, guidelines, and specifications that are used to implement a product's design process. These standards give manufacturers and developers guidelines and information on how to implement and produce the project. This section will discuss the standards that will be applied to our Chromesthesia device.

11.1. Related Standards

The standards that will be used for our Chromesthesia device will be acquired from the Institute of Electrical and Electronics Engineers Standards Association, or IEEE-SA. This institute is a well-known functioning arm of IEEE that creates and distributes global standards for a wide range of global technology businesses. Some of these businesses include: the National Electrical Safety Code, power and energy, etc. Our Chromesthesia device is a consumer technology so the standards that will be used will be ones that refer to those.

11.2. IEEE and Relevant Standards

11.2.1. Customer/user satisfaction: The user will be able to experience a very rare phenomena that only a small percentage of humans in the world can naturally experience. The target audience is anyone that wants to see the world how synesthetes experience it. This product could be used as an educational device to inform and educate people of the phenomenon, or purely for entertainment reasons.

11.2.2. Risk management: This is a process that is ongoing throughout the development of a device, system or service. It can be used to assess risks associated with system development, maintenance, and operation. Risk analysis must be addressed within the team to identify any potential risks and hazards, in order to establish the risks and contingencies that could possibly hurt the users and/or developers. For example, the Chromesthesia device could break if the user mishandles it and could potentially short circuit, or cause some other kind of injury. If mishandled, the device could also produce delayed outputs which would change the way the user would experience chromesthesia.

11.2.3. Durability: the device needs to be durable; it should be able to be used dozens of times, be handled by different users, and maintain the same performance throughout all simulations. By choosing a reliable power source, along with having a reliable build, while withstanding potential exterior forces, being durable is obtainable. Since this is an electronic device, the possibility of dust accumulating inside of the device is possible if not being used over a period of time, usually three months. With this in mind, any user needs to be mindful of dust buildup which could possibly cause malfunctions. Furthermore, the device will be susceptible to other hazards for electronics such as moisture and temperature.

11.2.4. Programming languages: This project will require fast and advanced algorithms to work. Therefore, a fast language such as C or C++ will be strongly considered.

11.3. Design Impact of Relevant Standards

This section will be discussing the effects on using these standards for the development of the chromesthesia device. This device's purpose is to generate interest into the world of synesthetes and see how they live their lives with this special phenomena that happens to them.

11.3.1. Design impact of User satisfaction

The main purpose of the chromesthesia device is to enable users to experience the phenomenon of Chromesthesia. The experience should have educational and entertainment benefits for the user

11.3.2. Design impact of Risk Management

Through proper risk management, The project would be handled in a dynamic manner and we would be able to

Every step of the design process we should be asking questions and referencing our goals and objectives of this project. The

11.3.3. Design impact of sustainability

The device should last about an hour on one charge. It should be sturdy enough to withstand a fall from the height of a person's head. Ideally, this device should be capable of being used more than 50 times between 50 different users. The microphones should have minimal interference, however due to the circumstances anticipated in the engineering atrium, the microphone array needs a very robust noise processing algorithm to handle the environment it will be used in.

11.3.4. Design impact of Programming languages

The microcontroller must be compatible with C languages.

11.4. Realistic Design Constraints

11.4.1. Economic and Time Constraints

Some hurdles can be found based on what is required for the project and the deadline (April 2022), allowing our group a total of 8 months to complete it. It is up to the group to overcome them and present a proper design and demo.

Economic Constraints

Some of the economic constraints the group has taken notice of are problems like not knowing the exact price of all the parts and pieces we will need for the project. There is no sponsor for this group, so finding funds is more difficult. There was no proper instruction in the beginning of how our budget for the device would be spent, although after conversing with group members it was determined that everything would be split equally. The members of this group want to demonstrate and showcase an amazing project but we have found out some pieces that would be the best parts that could possibly be purchased come out to be expensive and out of our budget. This results in the team members needing to take a step back and re-evaluate the situation and consider alternatives, which can cause a big delay in production. With all of this being said, all potential problems and solutions, such as manufacturing of materials, software, equipment costs, etc, are being reviewed constantly.

Time Constraints

Some of the time constraints the group has taken notice of are potential issues such as being a team project if one person falls behind, everyone falls behind. This turns into a matter of checks and balances and making sure everyone stays on track or needing to pick up the slack for other group members due to external circumstances. This project is of a large magnitude and needs to have certain time periods for project milestones to be completed. With our combined group's experience being low, there are many uncertainties in regards to knowing when certain things need to be completed on specific dates. Another constraint for our group is the availability of the individual group members, which makes having meetings more difficult than average. Other constraints that have been discussed were late delivery times, incorrect parts being delivered, or parts being the wrong size, not exactly what is needed, or possibly just not being what the project will need. Time management is key. If there is no sense of good time management throughout the team, this will inevitably cause a downward spiral in team morale and more importantly, project production.

11.4.2. Social, Environmental, and Political Constraints

Social Constraints

These consist of factors that can come up as a result of growing interest in a project or potential opposition. If a project uses public money, public concern and potentially pressure from media outlets can create tight constraints on a project which can result in changes to the product's original development roadmap.

Environmental Constraints

Any circumstance concerning the wellbeing of an individual person or the environment are real potential problems that need to be considered when determining constraints. Since this project is not The development of this project mainly concerns:

- Use of hazardous materials.
- Energy consumption
- Proper disposal and waste management
- Safety and wellbeing of users

Political Constraints

The policial constraint factors for our project primarily have to do with our use of microphones. While the design of this project will not utilize an internet connection, some questions might be brought forward about our algorithm, and the possibility of voice being recorded and the data being stored locally.

11.4.3. Ethical, Health and Safety Constraints

Ethical

Some ethical limitations that need to be considered relate to our handling of audio and visual recording that are required for our project. By utilizing a camera, we will have a visual archive of any individual who may be present during a demonstration. It is expected that footage not useful to a demonstration will not be saved, and promptly removed.

Health

As with any visual based process, consideration needs to be made for the health of a user. The design of this project incorporates visual based imagery, so it will be important to notify a potential user that imagery on the screen can cause issues for conditions such as epilepsy.

Safety

Safety considerations are important in both design, and demonstration of our device. While the focus of this device will mostly be software related, we need to consider the fact that faulty software can lead to unintended consequences, especially with a design based around heavy computations. One such issue can be the overheating of our Microprocessor. Depending on the severity, this can cause a fire hazard or injury if touched. Another safety consideration is making sure we keep to IEEE standards, as we will be using a Microprocessor and various other electronics with this device. Failure to follow these standards can result in dangerous exposure to electricity.