ECE 445 - Senior Design Lab

Project Proposal

Smart Assistive Glasses for the Blind

TA Sanjana Pingali

February 7th, 2024

Team Members

Ahmed Nahas (anahas2)

Siraj Khogeer (khogeer2)

Abdul Maaieh (amaaieh2)

Table of Contents

I. Introduction

A. Problem

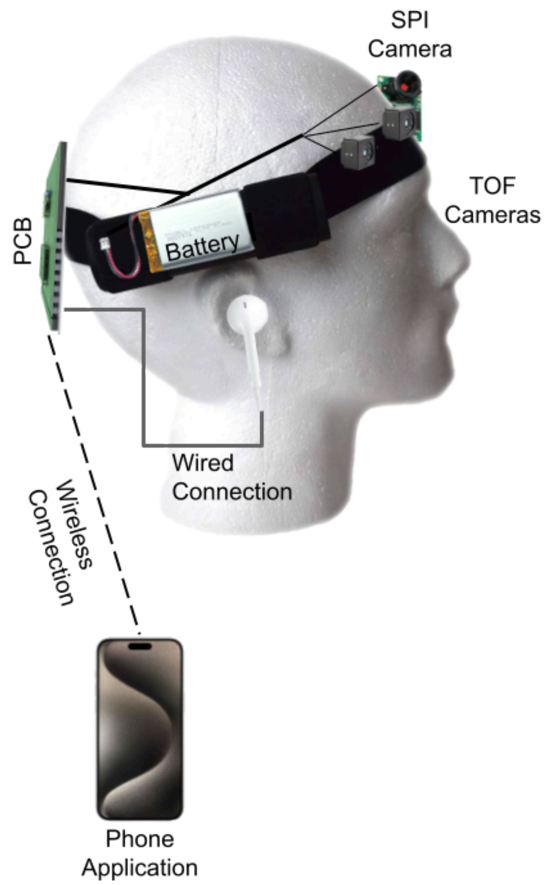
The underlying motive behind this project is the heart-wrenching fact that, with all the developments in science and technology, the visually impaired have been left with nothing but a simple white cane; a stick among today's scientific novelties. While the current solution, the cane, may help blind people walk, it is an outdated solution that only tells them there is an object if they touch it with the stick. This causes the blind trouble and sometimes it could put them in harm's way if they don't place the cane in the right place. This is a major problem that blind people deal with on a daily basis. On top of the fact that they must place the cane in the correct spot, they also have no way of knowing what it is in front of them. No way of having what's in front of them described without the help of another person. These lower the quality of life of blind people causing them more trouble throughout their life.

B. Solution

Our overarching goal is to create a wearable assistive device for the visually impaired by giving them an alternative way of "seeing" through sound. The idea revolves around glasses/headset that allow the user to walk independently by detecting obstacles and notifying the user, creating a sense of vision through spatial awareness. The general idea is to map the user's surroundings through depth maps and a normal camera, then map both to audio that allows the user to perceive their surroundings.

We'll use two low-power I2C ToF imagers to build a depth map of the user's surroundings, as well as an SPI camera for ML features such as object recognition. These cameras/imagers will be connected to our ESP32-S3 WROOM, which downsampled some of the input and offloads them to our phone app/webpage for heavier processing and ML algorithms.

C. Visual Aid



General concept visualization (may be built around headset or sports glasses). [1] [2] [3] [4] [5] [6] [7]

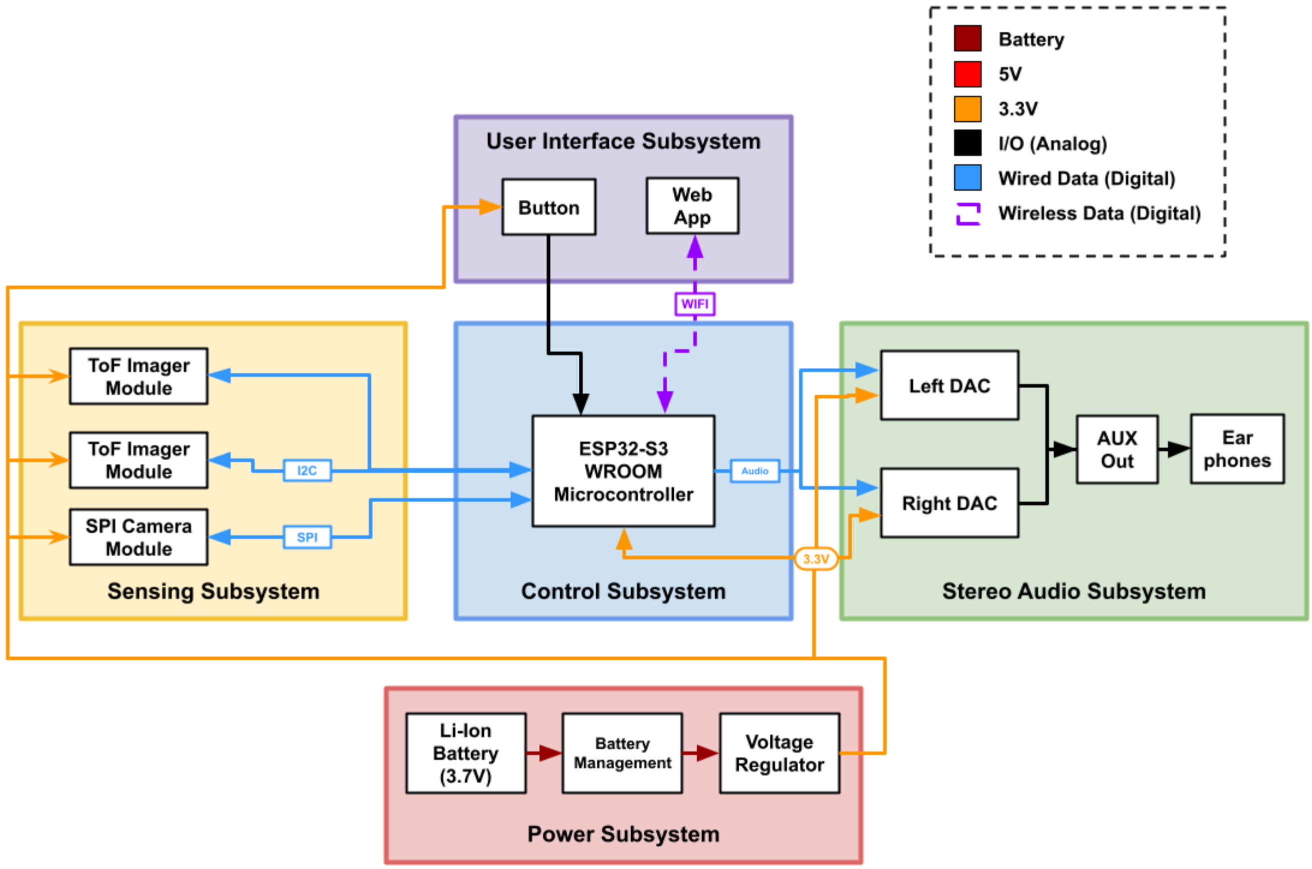
D. High Level Requirements

➔ The device will allow the user to differentiate between an obstacle that is 1 meter away vs an obstacle that is 3 meters away [Spatial Awareness].

- ➔ The stereo audio output will allow the user to successfully differentiate between obstacles upto 40° to the right and upto 40° to the left side of the user [Spatial Awareness].
- ➔ The device will correctly identify an object upto 1 meter in front of the user and communicate that to the user once prompted [Object Recognition].

II. Design

A. Block Diagram



B. Subsystem Overview and Requirements

Subsystem 1: Control

Overview:

The control subsystem consists of an ESP32-S3 WROOM that acts as the main hub for the other subsystems to connect to. The sensor subsystem will have two connections from it to the ESP32. First, the Tof Sensors will connect to it using the I2C bus to receive a 2x8x8 depth map array. Secondly, the camera will be connected via SPI, and when commanded by the ESP32, it will take a 2MP JPEG image and transmit it to the ESP32. The ESP32 will then use the data from these sensors, and send it via WIFI to the APP/web server. After processing it, the APP will send 8-bit unsigned stereo audio back to the ESP32 which it will then send to the stereo audio subsystem through 16 GPIO pins. The ESP32 is also in charge of programming the sensors if needed, and changing the control settings as needed. This will be done using the same SPI and I2C busses.

Requirements:

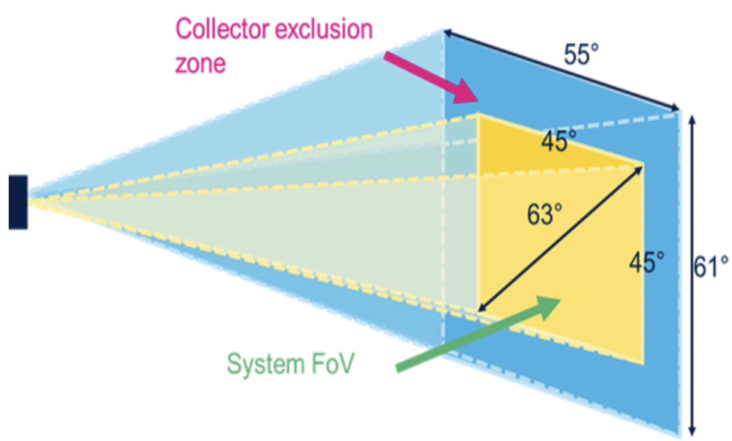
- MCU successfully receives data from ToF sensor and sends it to the APP within 1s.
- MCU sends a capture signal to the camera when a button is pressed on the application. Receives signal from APP, Captures 2MP JPEG image through SPI bus and sends it to APP through WIFI, all within 2 seconds
- MCU can successfully receive 8-bit audio from APP and sends it to the audio subsystem through its GPIO pins.

Subsystem 2: Sensing Subsystem

Overview:

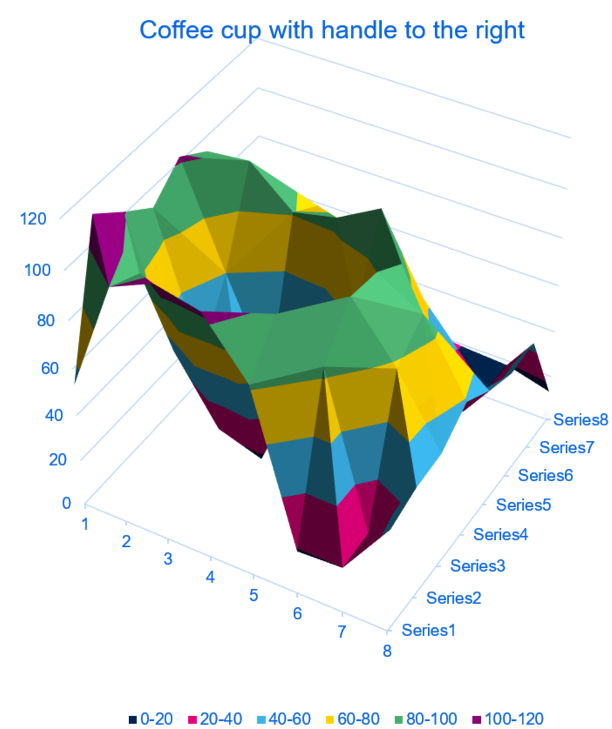
The sensor subsystem is separated into two sections; the time-of-flight (Tof) sensors, and the SPI camera.

The Tof sensors are used to capture a depth map in front of the user. Two Tof sensors will be used to provide a wide field of view (FOV). Each sensor has a 8x8 array of pixels, and an IR transmitter which maps the distance of objects in front of it. This data allows a person to be able to tell the distance to obstacles around them, helping them navigate without sight. The depth map data is then sent to the ESP32 using I2C to be transformed into spatial audio. The Tof sensors operate at 15 Hz when at 8x8, and draw around 100mA. The data format of the sensors can differ, but minimally, it's a 8x8 array of unsigned int, with some header information.



Tof Sensor FoV [15]

As for the SPI Camera, it will allow us to capture a colored image of the user's surroundings. A captured image will allow us to implement egocentric computer vision, processed on the app (more details highlighted in the user interface subsystem). This is exciting as having such an input will allow for other ML features/integrations that can be scaled drastically beyond this course. The camera will typically be idle, until a capture signal is sent by the ESP32, triggering it into capturing an image. Camera commands and image data are sent using the SPI bus. The camera has multiple formats, but in our case a 1920x1080 JPEG image will be sent to the ESP32



Example of Tof Sensor Visualization [22]

Requirements:

- The Tof sensor successfully maps the surroundings into a 8x8 array and sends it to the ESP32.
- SPI camera captures a 1920x1080 image when a button is pressed, and sends the data to ESP32 within 2s

Subsystem 3: Stereo Audio Subsystem

Overview:

The stereo audio subsystem will take in two channels of 8-bit audio from the ESP32 and convert them to an analog signal. This signal will be stepped down to audio safe levels and isolated using op-amps to be connected to an AUX. The user will then use the AUX port to connect earphones for onboard stereo audio. The signals will be a total of 16 bits from the GPIO pins of the ESP32. These represent an 8-bit unsigned number and are at a standard sample rate of 44.1 Khz.

Requirements:

- Converts 8-bit audio from ESP32 to analog signal at AUX out.
- Plays audio from ESP32 through AUX port using commercial earphones at normal and safe levels.
- Stereo audio is clear and spatial distinction can be made from left vs right originating audio

Subsystem 4: User Interface and ML Subsystem

The User Interface and ML subsystem will consist of a web app, as well as a button. Since we aim to implement one ML feature as a baseline for this project (one of: scene description or object recognition). This will only be given as feedback to the user once prompted by a button on the PCB: when the user clicks the button (along with the corresponding circuit) on the glasses/headset, they will hear a description of their surroundings. Therefore, we don't need real time object recognition, as opposed to a higher frame rate for the depth maps which do need lower latency.

The web app will carry all the heavy processing for (a) the spatial awareness algorithm as well as (b) the object recognition or scene description algorithms. We plan to use python to develop both (a) and (b), and use React Native to build a user-friendly interface (while this might not be used by the user, it will be helpful with testing). For (a), we will be building over research papers that effectively translate 2D scenes to "audioscapes", which allow the users to 'see' through sound, which is no simple feat, and no effective solution has yet been established. Our algorithm will build over these past approaches and introduce 3D depth maps to the scope, allowing for an exciting avenue to explore and innovate, which we do expect to be quite challenging and complex. As for (b), the Software Design section explains our approach.

Requirements:

- Successfully connects to the Control Subsystem through WIFI and receives data packs at up to 2 seconds of latency as an upper limit for higher resolution images

- Sends back audio data successfully upon processing to the Control Subsystem to be outputted by the Stereo Audio Subsystem with 50ms of latency as an upper limit

- Pressing the button will successfully prompt the ML algorithm, the object upto 1 meter away is successfully identified, and communicated back to the user

Subsystem 5: Power Subsystem

Overview:

This sub-system will supply power to all other subsystems except the app. The power subsystem will contain a rechargeable lithium ion battery pack that will be mounted alongside the rest of the circuits. The battery is regulated to the working voltage of 3.3V and a battery management system will be included to ensure short circuit and undervoltage protection. Battery will be encased to allow for easy removal and charging. Two linear regulators (LDO) will be used separately. One for the MCU and Stereo Audio subsystems, and one for the Sensor subsystems. This is to reduce heat dissipation, and allow for modular placement of sensors.

Requirements:

- Regulate 4.2-3V Li-Ion battery to 3.3V +/- 5% at 1A max while ensuring under-voltage protection
- Ensure that short circuit of battery is stopped to ensure safety of user
- Ensure that reverse polarity connection is stopped to ensure safety of circuit

C. Tolerance Analysis

One aspect of our design that might pose a risk to the project is the power draw of the circuits. There are many power hungry components which may be too much for the regulator to handle. All components work using 3.3V from the regulator. The MCU draws under 400mA, Tof sensors under 200mA, and Camera under 150mA. The audio subsystem draws a negligible amount of current. A liberal estimation for the max current draw is 800mA, which can be handled using a single high current regulator, but we chose to use two due to temperature concerns. The formula below can be used to estimate the temperature change of the linear regulator. [10]

$$T_j = i_{out}(v_{in} - v_{out})(\Theta_{jc} + \Theta_{ca}) + T_a$$

Parameter	**Value**
V_{IN} (V)	3.7 V
V_{IN}max (V)	4.2 V
V_{OUT} (V)	3.3 V
T_j max (C)	125C
Θja = Θjc + Θca	62 C/W
Current Draw (mA)	800mA

Calculation:

1 LDO $Typical\ T_j = 0.8 * (3.7 - 3.3) * 62 + 25C = 45\ C$

2 LDO $Typical\ T_j = 0.4 * (3.7 - 3.3) * 62 + 25C = 35C$

1 LDO $Max\ T_j = 0.8 * (4.2 - 3.3) * 62 + 25C = 70C$

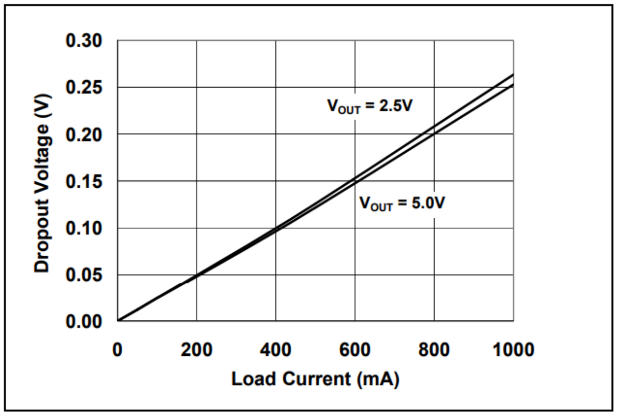
2 LDO $Max\ T_j = 0.4 * (4.2 - 3.3) * 62 + 25C = 48C$

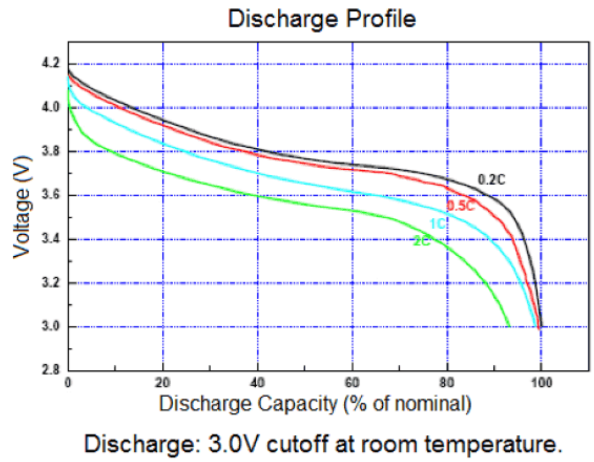
As seen above, while using 1 LDO is typically fine since Tj < 125C, at worst case the temperature can rise above 60C, which is the maximum recommended temperature for the Li-Ion battery. This is why using two LDOs would be better to ensure better heat management.

The second area of concern is the dropout voltage of the regulators. Linear regulators need to have an input voltage that is higher than the output by at least the dropout voltage to work properly. This is the main reason we chose the MCP1826S since at 400mA each it has a Vdrop = 0.1V. Furthermore, the fuse and P-MOSFET used in the battery management system will have a small voltage drop that we need to take into account. This drop is around 0.1V at max current draw, this value was approximated using the on resistance of the MOSFET and the fuse.

$Vcutoff = Vout + Vfuse + Vmosfet + Vdropout, Vcutoff = 3.5V$

This means that the regulators will work consistently as long as the battery voltage doesn't go below 3.5V. As seen in the discharge diagram below, the Li-Ion battery would have been discharged by 80-90%, which is sufficient depending on the battery size. If the drop-out voltage was higher, then the cut-off voltage for the under-voltage detector would need to increase, cutting down the battery capacity.





Dropout Voltage vs. Load Current[20]

Discharge Profile[21]

The third area of concern is the amount of data generated by the sensors, and if the ESP32 can store and transmit the data. The vl53l5cx packet size is 3.36KB, at a bandwidth of 50.909KB using the 8x8 15 hz mode [15]. Two of these sensors mean that the packet size is 6.72KB with a 101.82KB/s bandwidth. The SPI camera has a bandwidth of 4-8MB/s with a compressed 2MP image being 150KB [16]. This gives us a packet size of 157 KB and max bandwidth of 252 KB/s. And given that the ESP32-S3 has a 512KB SRAM, and 16MB PSRAM data should be limiting.

III. Safety and Ethics

A. Safety

1. This device is created as a way to increase the safety, awareness, and lifestyle of the blind. This device will be relied on by people who need it every day. With that in mind, it must be created and adapted to be the most accurate it can possibly be. Inaccuracies can potentially cause harm to those who are relied on. The IEEE Code of Ethics states “To uphold the highest standards of integrity, responsible behavior, and ethical conduct in professional activities. To hold paramount the safety, health, and welfare of the public.” This holds us to guarantee the safety of those using and relying on our device for guidance within their daily lives.

2. The ACM Code of ethics also states “An essential aim of computing professionals is to minimize negative consequences of computing, including threats to health, safety, personal security.” This includes the head mounting of batteries and other electronic components. We will ensure that all components are up to the ACM Code with safety measures, such as insulating all batteries and other high temperature components, as well as ensure thorough safety testing, thermal and otherwise, once the project is complete. We will fully inspect any technical components that have the potential to cause harm thoroughly in order to guarantee the safety of anyone using our device. Our components will also be enclosed in order to protect the user as an additional security measure.

B. Ethics

When it comes to ethics, our project is fairly straight forward. Protect our users with our full ability. Their data must be protected. Our vision is to have no data transmitted out of the system onto a network. This will help protect their data and also can help prevent any malicious attacks. The ACM code of ethics states that honesty and trustworthiness are critical for development of a product and we hold that highly. While our ML algorithms will use visual data to identify the user’s surroundings, none of that data will be stored beyond real-time processing. We will be completely open and straightforward with our users in any circumstances that might require the proper information. We also want them to know what is happening with their data and that it is secure.

IV. References

[1] "Pov headband holder Elastic Head Mount for TACTACAM 4K solo 5.0 6.0 wide WIFI," Stuntcams, https://stuntcams.com/tactacam-4k-5-0-wide-wifi-pov-headband-holder-elastic-head-mount.html.

[2] Phunk, "BLM PCB side view 2," MIDIbox Forum, http://midibox.org/forums/gallery/image/350-blm-pcb-side-view-2/.

[3] "Verizon Apple iphone 15 pro max 512GB natural titanium," Walmart.com, https://www.walmart.com/ip/Verizon-Apple-iPhone-15-Pro-Max-512GB-Natural-Titanium/5055309496.

[4] "EarPods (lightning connector)," Apple, https://www.apple.com/shop/product/MMTN2AM/A/earpods-lightning-connector.

[5] "PKCELL flat 3.7V 2000mah rechargeable lithium polymer 803860 battery W," Rokland, https://store.rokland.com/products/pkcell-flat-3-7v-2000mah-rechargeable-lithium-polymer-803860-battery-with-jst-type-ph-2-0-plug?currency=USD&utm_source=google&utm_medium=cpc&utm_campaign=google%2Bshopping&gclid=CjwKCAiA8YyuBhBSEiwA5R3-Ey4whPgbrOMG6_z4ZGZ5wZIhXDMhlLOUQZO6LEmSkiu2qe-QqpAAVxoC6fcQAvD_BwE.

[6] "3D time of Flight Camera - DepthVista," e, https://www.e-consystems.com/3d-depth-cameras/tof-time-of-flight-camera.asp.

[7] "ArduCam-mini OV5642 5mpx 2592x1944px 120fps SPI -," BOTLAND, https://botland.store/arduino-rpi-cameras/6558-arducam-mini-ov5642-5mpx-2592x1944px-120fps-spi-camera-module-for-arduino-uno-mega2560-raspberry-pi-pico-5904422375133.html.

[8] "IEEE code of Ethics," IEEE, https://www.ieee.org/about/corporate/governance/p7-8.html.

[9] Code of ethics, https://www.acm.org/code-of-ethics.

[10] E. I. S. Services, "Propeller-less multi-rotor," :: ECE 445 - Senior Design Laboratory, https://courses.engr.illinois.edu/ece445/wiki/#/regulators/index.

[11] "ESP32-S3-WROOM-1-N8: Digi-Key Electronics," Digi, https://www.digikey.com/en/products/detail/espressif-systems/ESP32-S3-WROOM-1-N8/15200089.

[12] "TLC27L1ACP last Time Buy," TLC27L1A | Buy TI Parts | TI.com, https://www.ti.com/product/TLC27L1A/part-details/TLC27L1ACP.

[13] "AZ1117CD-3.3TRG1: Digi-key electronics," Digi, https://www.digikey.com/en/products/detail/diodes-incorporated/AZ1117CD-3-3TRG1/4470979.

[14] A. Industries, "Lithium ion polymer battery - 3.7V 2500mah," adafruit industries blog RSS, https://www.adafruit.com/product/328#technical-details (accessed Feb. 8, 2024).

[15] UM2887, https://www.st.com/resource/en/user_manual/um2887-a-software-integration-guide-to-implement-the-ultra-light-driver-of-the-vl53l5cx-timeofflight-8-x-8-multizone-ranging-sensor-with-wide-field-of-view-stmicroelectronics.pdf.

[16] "Arduino Camera: SPI camera module from Arducam," Arducam, https://www.arducam.com/spi-arduino-camera/#:~:text=The%20Arducam%20camera's%20SPI%20speed,recommended%20speed%20is%204Mbps~8Mbps.

[17] Pui Audio, https://www.digikey.com/Site/Global/Layouts/DownloadPdf.ashx?pdfUrl=36123F849FAA4FD9939A80185959CD4D.