Auditory Spatial Awareness Device for the Visually Impaired

By
Arnav Das
Darius Haery
Noah Salk

Final Report for ECE 445, Senior Design, Spring 2020
TA: Dhruv Mathur

8 May, 2020
Project No. 25

Abstract

The product designed and developed for the second half of senior design was an assistive device for the blind that seeks to aid in the navigation of their environment. By taking inspiration from a flashlight and a white cane, the proposed product is a handheld device with an embedded ultrasonic sensor which measures the distance of the nearest object and transmits 3D audio feedback to the user via earphones. With 3D audio feedback the user will be able to pinpoint the spatial placement and relative distance of objects in their environment, allowing for general circumnavigation.

Table of Contents

1 Second Project Motivation

1.1 Updated Problem Statement

There are 36 million people who suffer from blindness around the world, and require mobility aids to help them effectively navigate their surroundings and avoid obstacles [1]. The most common tools used by victims of blindness include white canes and guide dogs [2]. However, both of these tools are problematic for the following reasons:

1) White canes are cumbersome to use in environments with many obstacles. The user may knock down precariously placed items indoors, accidentally hit pedestrians on a crowded sidewalk, etc. These canes only provide information about the environment within a tight radius.

2) Guide dogs require intensive training and a substantial amount of maintenance, so they may not be affordable to all victims of blindness.

1.2 Updated Solution

We aim to develop a cost effective electronic tool that can be intuitively used by blind people to navigate their surroundings. We propose a two-part device consisting of a headset and a handheld device. The handheld device will resemble the body of a flashlight and will use an ultrasonic sensor to detect objects up to 3 meters away from the user. A 3D audio signal (a short beep) will be transmitted to the user via the headset to indicate the presence of an obstacle, and will be varied based on two factors: the distance of the object from the sensor and the angle (horizontal and vertical) of the device relative to the user's head. The distance of an object from the sensor will determine the speed of the signal; shorter distances will result in more frequent beeps. The horizontal and vertical angles of the device will be used to determine a filter that will be applied to the audio signal, so that the source appears to match the physical position of the obstacle.

Since blind people rely heavily upon hearing to create a mental image of their environment, we hope that a 3D audio feedback will enable users to navigate their surroundings in a more natural manner [3]. Moreover, the usage of this device will be similar to that of a white cane so it will be relatively easy for users to learn how to use the device. Its operation is analogous to navigating a dark room with a flashlight.

In general there are more products utilizing advanced technologies to aid the hearing impaired then there are for the visually impaired. Blind people have a harder time navigating their surroundings than the deaf and there are very few products that offer a solution beyond walking sticks and guide dogs. The proposed solution is inspired by the flashlight and the ability it provides for the user to navigate a dark room by only illuminating a section of the room at a time. The images of the previously illuminated sections are burned into the memory of the user. For the blind, their whole world is one dark room, and we believe that the concept of a flashlight can be ported over to the auditory domain.

1.3 Updated High-level Requirements

- The handheld component of the device should be lightweight and portable.
- The sensor should provide general object awareness information within 10 ft of the user.
- The source of the audio feedback should appear to align with the physical location of the object.

1.4 Updated Visual Aid

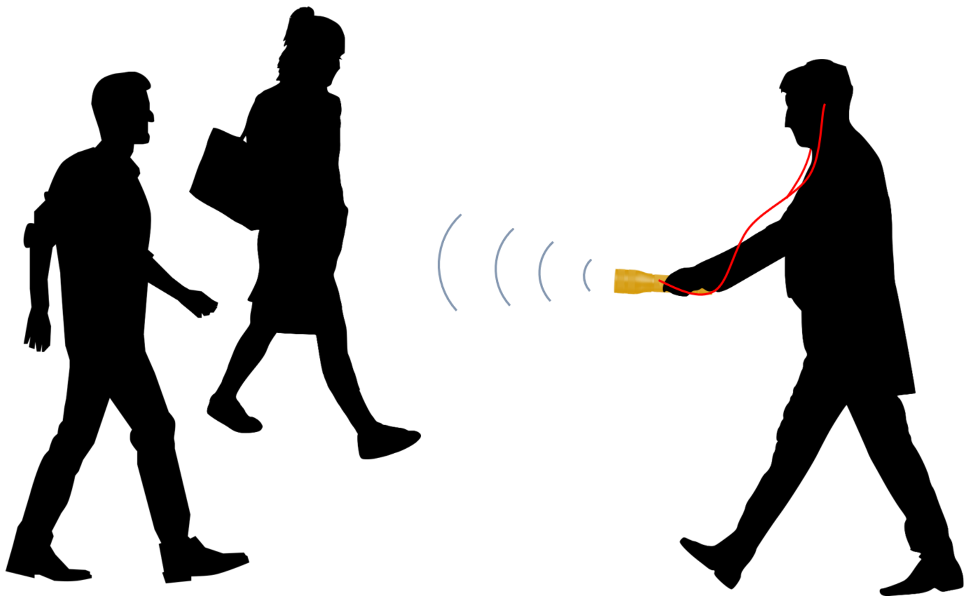


Figure . Typical use case visualization

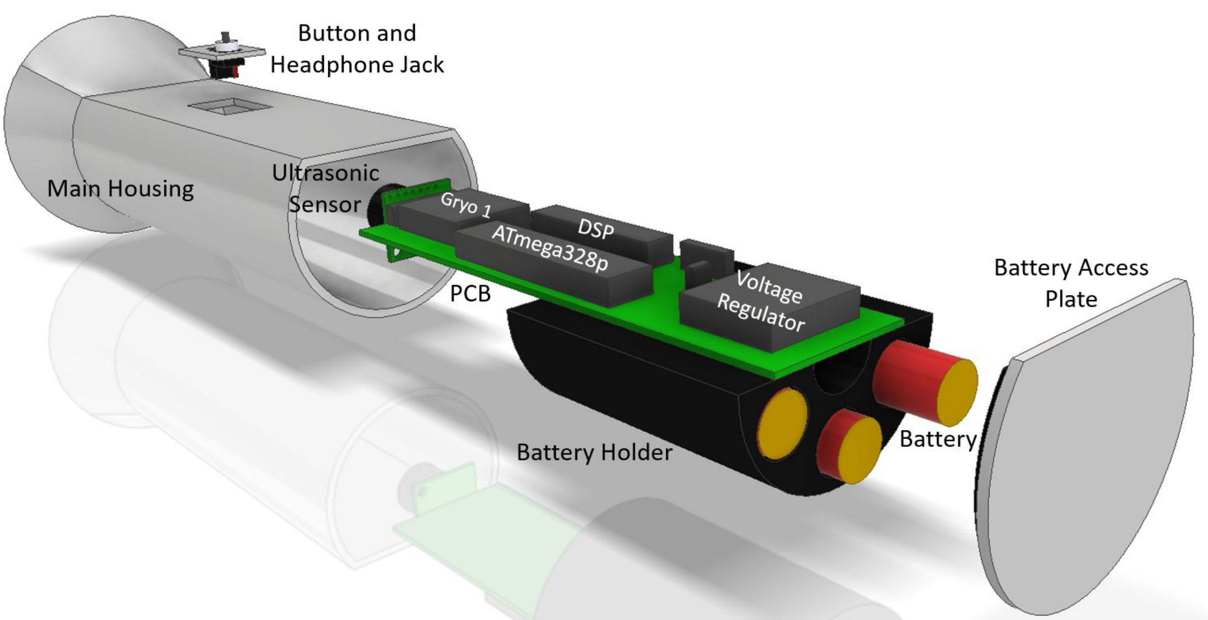
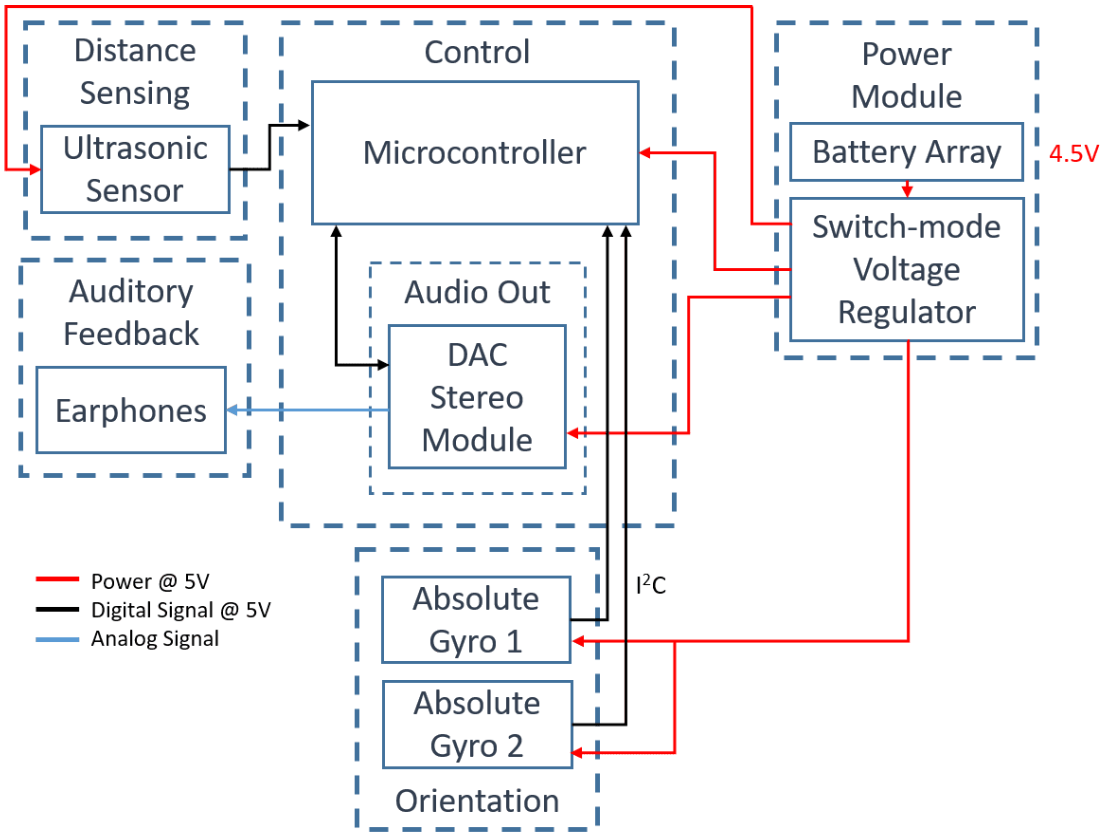


Figure . Physical design and layout of the hand-held component

1.5 Updated Block Diagram



2 Second Project Implementation

2.1 Implementation Details and Analysis

2.1.1 Power Module

2.1.1.1 Battery Array

For accessibility and simplicity reasons, commercially available battery cells will be used in an array as the main energy source to power all devices. The battery module will be housed in the hand-held device. Along with a current rating (not to be exceeded), each battery has a nominal voltage that is assumed constant on a first-order design but actually declines over time as the cell is drained; this is one of the factors that explains the need for a voltage regulation circuit. By stacking consecutive battery cells in series, the overall array output voltage will be the individual cell voltage multiplied by the number of series cell connections. Adding parallel branches will proportionally increase the array's rated current output.

To properly size the battery array, it is necessary to have voltage rating and current consumption information for each component in the device. That information was extracted from selected component datasheets (more information on component selection can be found in later sections) and listed in the following table.

Component	Maximum Current Rating (mA)	Voltage Ratings (V)
Gyroscopes x2	7.8 (total)	2.375 – 3.46
Earbuds	10	0.447 – 1.736
ATmega328p Microprocessor	10	2.7 – 5.5
Ultrasonic Sensor	2	2.8 – 5.5

Table I. Component voltage and current ratings

The battery array was designed to approach the microprocessor and ultrasonic sensor's nominal voltage of 5 V. The most common commercial battery cells AA and AAA both supply 1.5 V. Three of these cells stacked in series gives a 4.5 V output voltage. The remaining 0.5 V is supplied by the voltage regulator (more information in next section). The gyroscopes will be powered by the ATmega328p's onboard 3.3 V supply and the earbuds will be driven with a stereo DAC module that will handle voltage modulation. Maximum power draw is overestimated using the simple equation of the sum of the products of maximum voltage and maximum current draw rating for each component. This computation yields 111 mW of maximum power draw which equates to roughly 25 mA from a 4.5 V stacked voltage array. This current draw is well within the rating of a single AAA battery cell. Using three AAA batteries stacked in series yields a worst case scenario of 40 operational hours (as analyzed from the data shown in Fig. 3) of use before the pack voltage drops too low to be properly regulated.

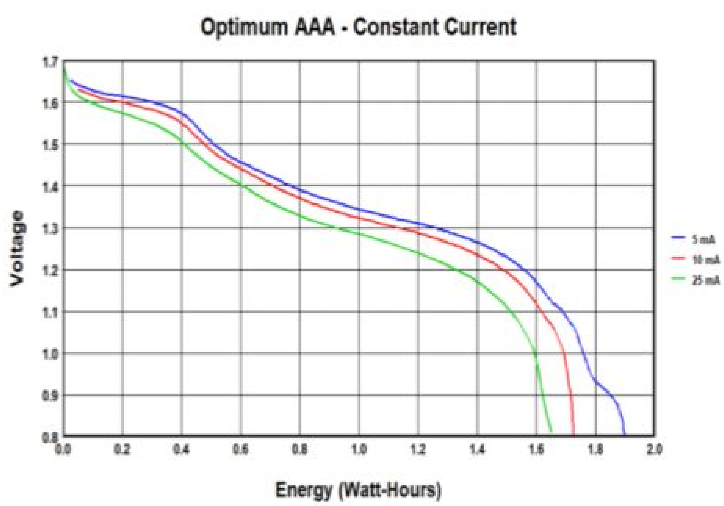


Fig. 3. AAA performance data [4]

2.1.1.2 Voltage Regulation

As mentioned in the previous section, voltage regulation is required to maintain constant supply voltage for the components in the device. Not only is the battery array's output voltage not at the desired 5 V, but it also declines as the batteries are drained. For low power applications, such as this, there are two approaches typically used to regulate voltage: linear voltage regulation and non-linear/switch-mode DC-DC voltage regulation. The former has the advantages of simplicity and cost, but lacks in efficiency and also requires a stacked input voltage to be at least 2x that of the desired output voltage; it can only convert down voltage using a minority carrier transistor (thus lack in efficiency). The latter option is more complex and costly, but is also more efficient and can convert down or up voltage (buck-boost).

An integrated circuit (IC) switch-mode buck-boost voltage regulator was chosen for this application. The Linear Technology's LTC3204-5 has a fixed 5 V output and can handle input voltages of 2.7 - 5.25 V and is overrated for an output current of 150 mA. A circuit schematic for the typical application of the voltage converter was taken from the chip's datasheet and is shown in Figure . The chip will automatically shut off in under voltage situations which provides safety and battery health.

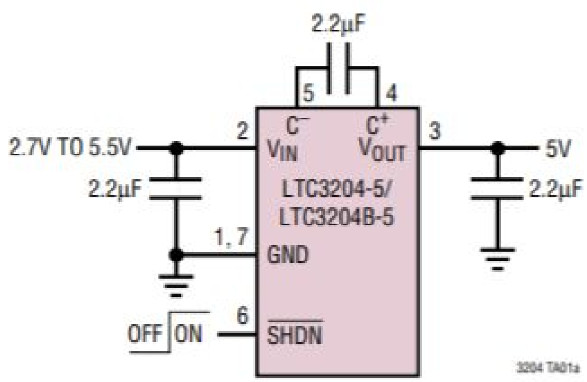


Fig. 4. Typical application of the LTC3204-5 [5]

2.1.2 Distance Sensing

2.1.2.1 Ultrasonic Sensor

With the high-level requirement of general object sensing within 10 feet of the user, an ultrasonic sensor falls perfectly in this category. Compared to infrared and light sensors (outside of LIDAR), ultrasonic sensors have the longest range. They are slow (at roughly 20 Hz) but sufficiently fast enough for this application, and have low power consumption. Ultrasonic sensors transmit sonar waves and receive these back at the point of transmission, measuring time from transmit to receive. The acoustic media means that the reading is slow compared to EM and light based methods and also that the signal experiences dispersion at a significant angle. This angle affects the reading range and fidelity in aurally depicting the environment, however is advantageous in that the user cares about general objects in their path and not fine details that are impossible to discern from the hearing sense.

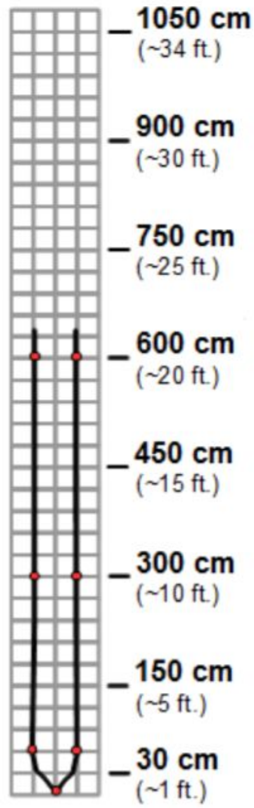


Fig. 5. LV-MaxSonar-EZ MB1030 distance reading capabilities on 30 cm grid [6]

A ultrasonic sensor that meets the required distance reading capabilities while maintaining a compact form factor was chosen (LV-MaxSonar-EZ MB1030). Its reading capabilities can be seen in Figure which was extracted from the components datasheet. After dispersion, the beam width is around 60 cm and can read up to 20 feet. Closer objects (within 2 feet) can be discerned with more detail.

2.1.3 Feedback and Orientation

2.1.3.1 Motion Tracking

We will be using two MPU-9250's, one to determine the orientation of the headset and the other the orientation of the handset. These orientations will be fed to the signal processing unit for the 3D audio generation. The MPU-9250 has a gyroscope, accelerometer, and a magnetometer. This device can connect to the Arduino microcontroller using an I2C connection, relying on a Master-Slave model. There are four pins, VDD, GND, SCL (Serial Clock), and SDA (Serial Data). Since we are relying on two chips that both need to send data to the microcontroller, one MPU will connect to the other MPU by I2C, which will then connect to the microcontroller by I2C. The MPU-9250 allows this configuration, forwarding the signals to and from the microcontroller to the other chip. It determines this by the slave address the microcontroller specifies, which will differ by 1 bit depending on whether it wants to communicate with sensor 1 or 2. In Figure 6, the optional sensor is where the second MPU-9250 would be connected and the system processor is the Arduino microcontroller.

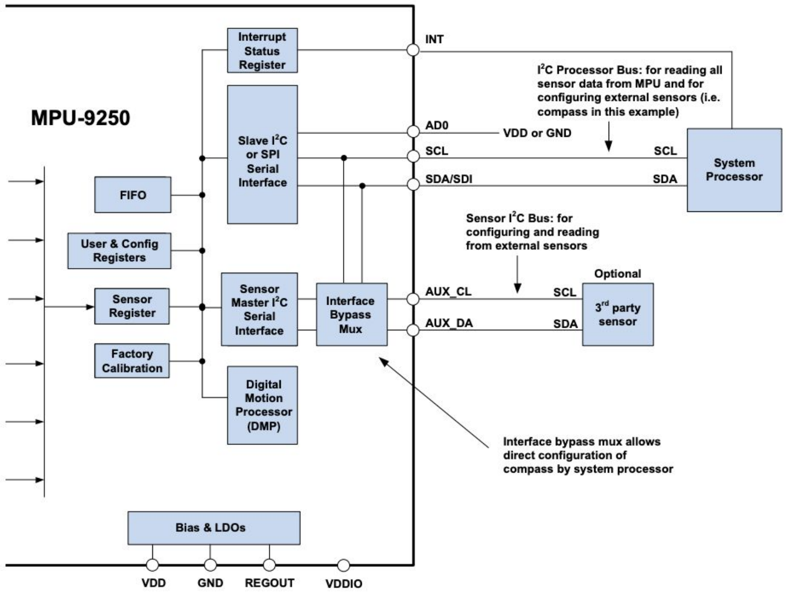


Fig. 6. MPU-9250 [7]

The MPU-9250 has a 3-axis gyroscope. This gyroscope has registers which store the X, Y, and Z values of axis angles. As the documentation states, these values are 16-bit in size and can be sampled at any chosen value in the range of 4 to 8,000 Hz. We will set the sampling rate to be around that of the ultrasonic sensor. The maximum detectable speed of the sensor is programmable as well. The options are +/- 250, 500, 1000, and 2000 degrees per second. We will be going with the range of +/- 250 because it is within the expectations of a human's range of motion and choosing a lower number will maximize the fidelity of the sensor. The magnetometer has a 3-axis silicon monolithic Hall-effect sensor that has a measurement range of +/- 4800 µT. Figure 7 below shows the physical representation of these values [8].

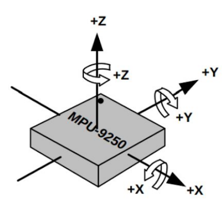


Figure 4. Orientation of Axes of Sensitivity and Polarity of Rotation for Accelerometer and Gyroscope

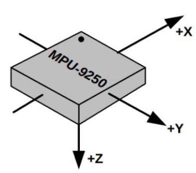


Figure 5. Orientation of Axes of Sensitivity for Compass

Fig. 7. Gyroscope and Compass Axis Depiction [7]

2.1.3.2 Earphones

Commercially available, cheap stereo in-ear earbuds are sufficient for this application. These earphones have a 3.7 mm headphone jack that will be used to transfer audio signals to the earphone drivers. To access the four ports on a 3.7 mm interface, a breakout board supplied by Sparkfun (BOB-1157) has a female headphone jack input and provides pin access to each ring on the headphone jack. A breakdown in [9] of Apple's wired earpods brings to light the electrical properties of a typical low end earbud. A driver impedance of roughly 100 Ω at audible frequencies and an average supply voltage of 1 V_{rms} results in current consumption of 10 mA.

2.1.4 Control, Software, and Auditory Feedback

2.1.4.1 Microcontroller

The microcontroller will be responsible for computing the distance of an obstacle from the ultrasonic sensor readings, finding the angle differential between the device and the user's head from the gyroscope sensors, and retrieving the appropriate 3D audio signal to play through stereo headphones.

The ATmega328P was determined to be the processor of choice for the control unit due to its ease of use, prevalence, and the designers predisposition to the Arduino platform which uses this chip. Moreover, the ATmega328P has 14 I/O pins which is more than enough to receive inputs from our gyroscope and ultrasonic sensors and send an audio output through a DAC to our headphones. The ATmega328P, like most microcontrollers, has very limited SRAM space (only 2 kilobytes). However, it is compatible with SD Card reader modules which will enable the microcontroller to easily access stored audio files during runtime [10].

2.1.4.2 3D Audio Generation

The 3D audio feedback will create an effect that causes the user to perceive the source of the signal to be aligned with the position of the obstacle. A 3D audio signal can be generated by convolving an input sound with a pair of head-related impulse responses (HRIR), a filter that characterizes how a sound wave reaches each ear from a point source [11].

We plan to first obtain a collection of filters consisting of several HRIR pairs corresponding to a large array of different positions (various pairs of horizontal and vertical angles) around the user. The vertical angles will range from 0 to 180 degrees while the horizontal angles wil range from 0 to 360 degrees. Due to limited storage, we will sample both of the angles in 10 degree increments resulting in a total of 648 possible positions and filter pairs.

We will then generate a sine wave with a frequency of 880 Hz for about 10000 samples (corresponding to about .25 seconds at the standard sampling rate of 44.1kHz). We will apply each of the filter pairs on the signal, resulting in a total of 1296 different audio signals. Assuming we use 8-bit audio, the collection of audio signals will take about 12.9 megabytes (calculation shown below).

$$1\,bytes\,*\,10,000\,samples\,*\,1,296\,filters = 12,960,000\,bytes$$

At runtime, the microcontroller will compute the horizontal and vertical angles relative to the user's head and the closest matching 3D audio signal will be retrieved from the SD card to be played through the headphones. The audio signals will be generated offline, using a high level programming language such as Python.

2.1.4.3 Micro SD Card Reader

Due to the high volume of audio files we plan to store and the limited SRAM storage on most microcontrollers, a 2GB SD card is necessary. To enable our microcontroller to access the SD card, we will use a Memory Card Shield Module with a SPI Reader Micro SD Memory Card (shown below). This module is compact (4.1x 2.4cm) and is compatible with the Atmega328p microcontroller [12].

Fig. 8. SPI Card Reader Shield

3 Second Project Conclusion

3.1 Implementation Summary

3.1.1 3D Audio

We completed the first step of the offline 3D audio generation by collecting a large set of HRIR pairs and convolving them with sinusoids. In Figure 9, we plot a HRIR pair where the source is 45° to the left of the user's head, and on the same elevation (0°).

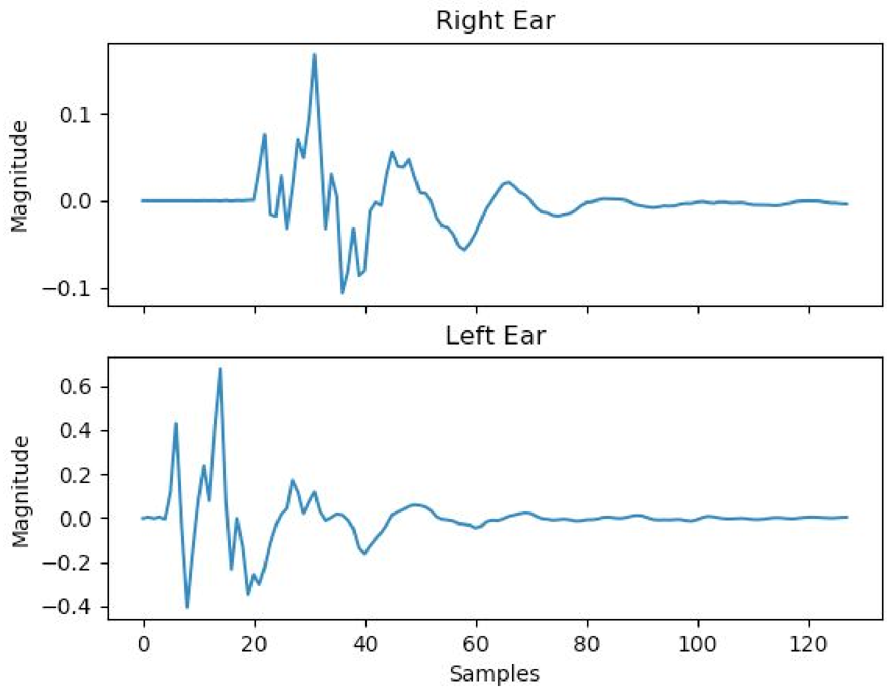


Fig. 9. An example of a HRIR pair

It is interesting to note that, though much of the filter is uninterpretable, the HRIR pair incorporates a time delay between the two ears. More specifically, the filter corresponding to the right ear peaks a few samples to the right of the left ear filter. This is due to the fact that the virtual source is to the *left* of the user, and a sound wave originating from the specified source would reach the left ear first.

Since the time delay is one of the predominant effects of the HRIR's, we found that applying these filters on a pure sinusoid did not create the desired 3D audio effect. Due to the periodic nature of pure tones, a time delay effect imparted by the HRIR pair was practically imperceptible when tested. Therefore, we decided to utilize an exponentially decaying sinusoid as our base signal. When we tested the HRIR pair on the new base signal, the effect improved substantially. We show part of the base signal and the 3D audio signal in the figures below (truncated for visual clarity).

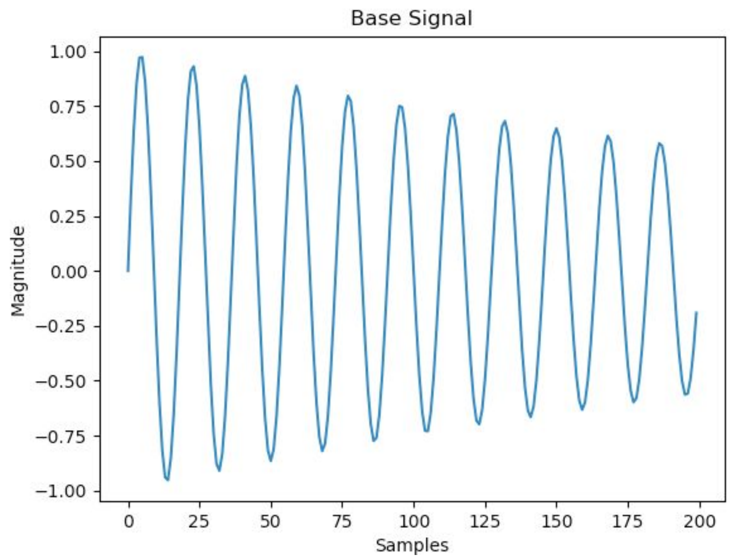


Fig. 10. Exponentially decaying tone

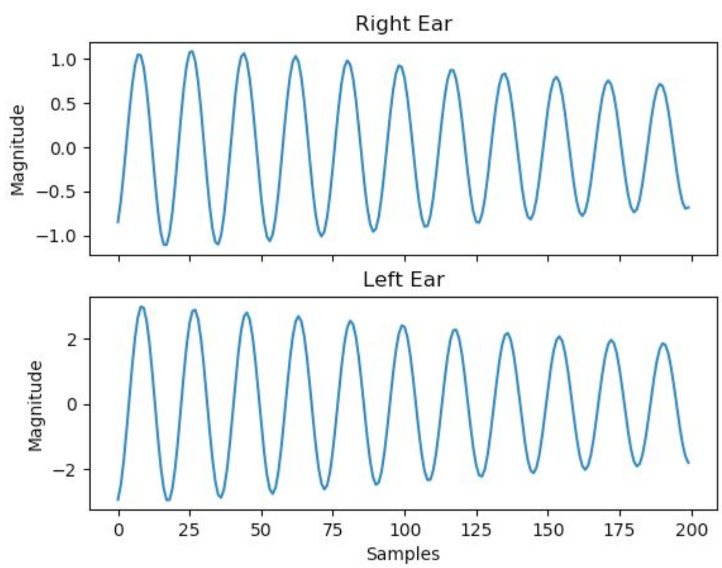


Fig. 11. Base signal after convolved with HRIR pair

We suppose that the 3D audio effect could be further improved through generating a more natural sounding base signal. In the future, such a signal could be generated using techniques from electronic music synthesis. However, the 3D audio signals we generate from the exponentially decaying tone are sufficient for providing general spatial awareness to the user.

3.2 Unknowns, Uncertainties, Testing Needed

The completion of this project rests on hardware development, testing, and validation. All in house 3D audio can be developed however the HRIR filters may need adjustments that can only be done after human trial. The PCB completion will inherently require validation and debugging, using standard lab equipment such as multimeters, function generators, power sources, and potentially even trace cutting equipment. The actual PCB assembly will use delivered components from DigiKey and requires soldering equipment to mount the components to the PCB. All components in the product rely on the voltage regulator to provide steady 5V DC power and therefore it is imperative that we validate the output and operation of the chosen IC before inserting it into the product.

The middle ear has an impulse response that may vary from person-to-person. This introduces the greatest amount of uncertainty in our project as to the usability of the product. It is possible that the 3D audio HRIR filters will need to be adjusted case-to-case. As we cannot analytically estimate each user's middle ear acoustic characteristics, this requires trial and experimental measurements.

Finally, the handheld component rests on a 3D printed housing that is only possible to print in the ECEB labs or at another facility with 3D printing capabilities. Outside of general testing, the rest of the work can be and was done remotely.

3.3 Ethics and Safety

Since the device we propose is intended to be used as a primary source of navigation, malfunctions in the device could seriously harm the user. As such, thorough testing of the object sensors and the audio feedback system is necessary in order to ensure that the device never fails to detect the presence of an object and provides accurate feedback to the user. It is paramount for us to be forthcoming about any potential bugs or limitations in the device to make sure that our product will not result in any injuries, in accordance with Rules #3 and #9 in the IEEE Code of Ethics [15].

Safety risks involved in the designed product can be incredibly consequential. If the device doesn’t correctly identify an object in close proximity to the user, that object has serious potential to harm the user. One method of risk mitigation is through the use of an ultrasonic sensor as the main sensor. This type of sensor has a large beam angle compared to other types of proximity sensors (infrared, LIDAR, etc.) and therefore gives a buffer for the user to identify all objects incident upon that beam. The tighter the beam angle, the more the user has to sweep the device to get a reading and consequently the more likely they are to miss a fine detail that may pose danger to the user.

3.4 Project Improvements

If we were given an entire year to develop this product, the following improvements would be made to produce a better result:

1. Fully explore different types of distance sensors (Lidar and various ultrasonic sensors) with human trials to determine whether a smaller or larger beam angle is more comprehensive for the user to navigate their environment. A comparison can be made between the different sensors analytically, but the user's opinion on which type is easier to use is a more psychological and subjective variable.
2. Delve into the physical design to incorporate ergonomics and user feedback for button placement, headphone jack placement, gyroscope orientation, weight distribution in the handheld portion etc. The potential for earphones to inhibit the blind's sense of hearing is real and would be explored to make sure we can reduce that effect. Audio pass through or non-sealing earbuds could be a solution for this. In addition, bluetooth truly wireless earphones may be preferred by the customer and this should be looked into as an option.
3. Create an ASIC chip with an integrated controller, power converter, and storage unit for filter data instead of using commercial ICs whose resources are not completely utilized. This requires a final result and would only be done after determining previous prototypes have been debugged fully and the next stage is commercialization. The main reason for this improvement is to cut down on mass production costs. The ATmega328p is inexpensive however the resources of the chip aren't fully utilized in our product and so a more streamline ASIC will provide just the amount of computation we need and no more leading to a reduction in costs.

4 Progress Made on First Project

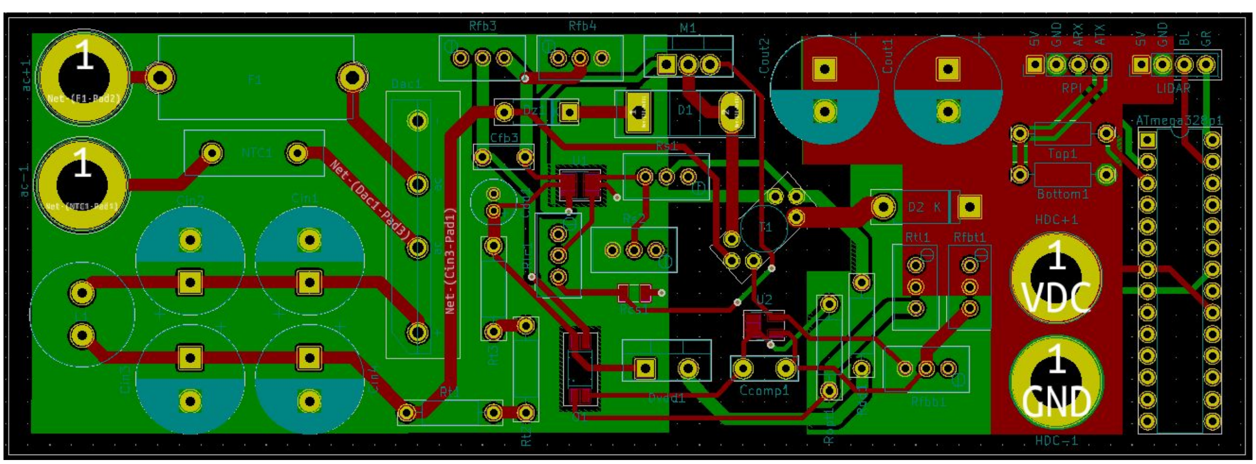


Figure . PCB layout of the first project including power converter stage and microcontroller

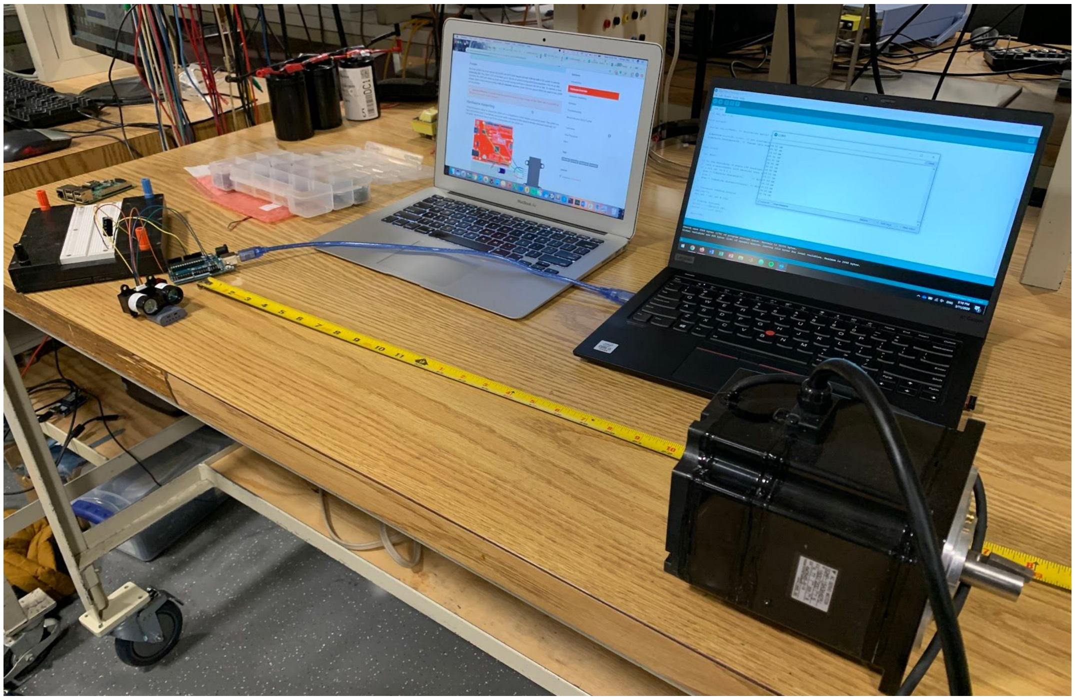


Figure . Lidar circuit setup and distance component validation

5 References

[1] “Global Vision Impairment Facts.” [Online]. Available: https://www.iapb.org/vision-2020/who-facts/

[2] [Online]. Available: http://www.rehab.research.va.gov/jour/74/11/2/73.pdf

[3] “Sight vs. Sound: Can Blind People Hear Better?” [Online]. Available: https://www.signia-hearing.com/blog/do-blind-people-have-a-better-hearing/

[4] [Online]. Available: https://d2ei442zrkqy2u.cloudfront.net/wp-content/uploads/2016/03/01160557/OP2400_US_OP1.pdf

[5] "LTC3204 Datasheet and Product Info | Analog Devices", *Analog.com*, 2020. [Online]. Available: https://www.analog.com/en/products/ltc3204.html.

[6] "LV-MaxSonar® EZ™ Series Datasheet - MaxBotix Inc. | DigiKey", *Digikey.com*, 2020. [Online]. Available: https://www.digikey.com/en/datasheets/maxbotix-inc-lv-maxsonar-ez_datasheet.

[7] [Online]. Available: https://learn.sparkfun.com/tutorials/gyroscope/all

[8] “MPU-9250 Product Specification.” [Online]. Available: https://invensense.tdk.com/wp-content/uploads/2015/02/PS-MPU-9250A-01-v1.1.pdf

[9] "Earphones | CDW", *Cdw.com*, 2020. [Online]. Available: https://www.cdw.com/search/Electronics/Headphones-Earphones-Accessories/Earphones/?w=EE1.

[10] [Online]. Available: https://www.microchip.com/wwwproducts/en/ATmega328P

[11] “Head-Related Impulse Responses Repository.” [Online]. Available: https://research.steinhardt.nyu.edu/marl/research/head_related_impulse_responses_repository

[12] [Online]. Available: https://www.elliottelectronicsupply.com/micro-sd-storage-board-mini-tf-card-memory-reader-shield-module-spi-for-arduino.html

[13] "Solid angle", *En.wikipedia.org*, 2020. [Online]. Available: https://en.wikipedia.org/wiki/Solid_angle.

[14] [Online]. Available:
https://randomnerdtutorials.com/complete-guide-for-ultrasonic-sensor-hc-sr04/

[15] “IEEE Code of Ethics.” [Online]. Available:
https://www.ieee.org/about/corporate/governance/p7-8.html