ECE 445 - SENIOR DESIGN LABORATORY
PROJECT PROPOSAL
02.13.2025

Audio Augmented Reality Glasses
AARG

Team #14
Nikita Vasilyev (nvasi2)
Sunny Chen (sunnyc3)
Evan Chong (eschong2)

TA: Aishee Mondal

Table of Contents

Section 1: Introduction

Section 2: Design

Section 3: Ethics

Section 4: References

1 Introduction

1.1 Problem

Have you ever seen a plant in nature or an animal in the wild that piqued your interest, but you didn't have an efficient way of researching what it was? Repeatedly searching online to identify the subject can be a lengthy and tedious task, and this is the problem we seek to address. Our solution is meant to enlighten our user of unknown plants, animals, or objects in any setting they are observing. Think of it as an auditory tour guide!

1.2 Solution

Our project idea stems from the surge of AR prototype glasses being introduced over the past year. We are planning to create our own glasses but in contrast to those on the market, ours will focus on the audio experience of the user. These glasses will have the explicit capability of capturing images of objects and relaying this information to an application that will process these images in the backend. The application will then send an explanation of the object back to an audio device on the glasses (either a speaker or bone-conducting device). The glasses will essentially work as a digital tour guide, with the explanation of the object being auditory rather than visual. The use case we have decided to tackle is a botanical tour guide, but the purpose is to create a platform that other applications can utilize for their objectives.

1.3 Visual Aid

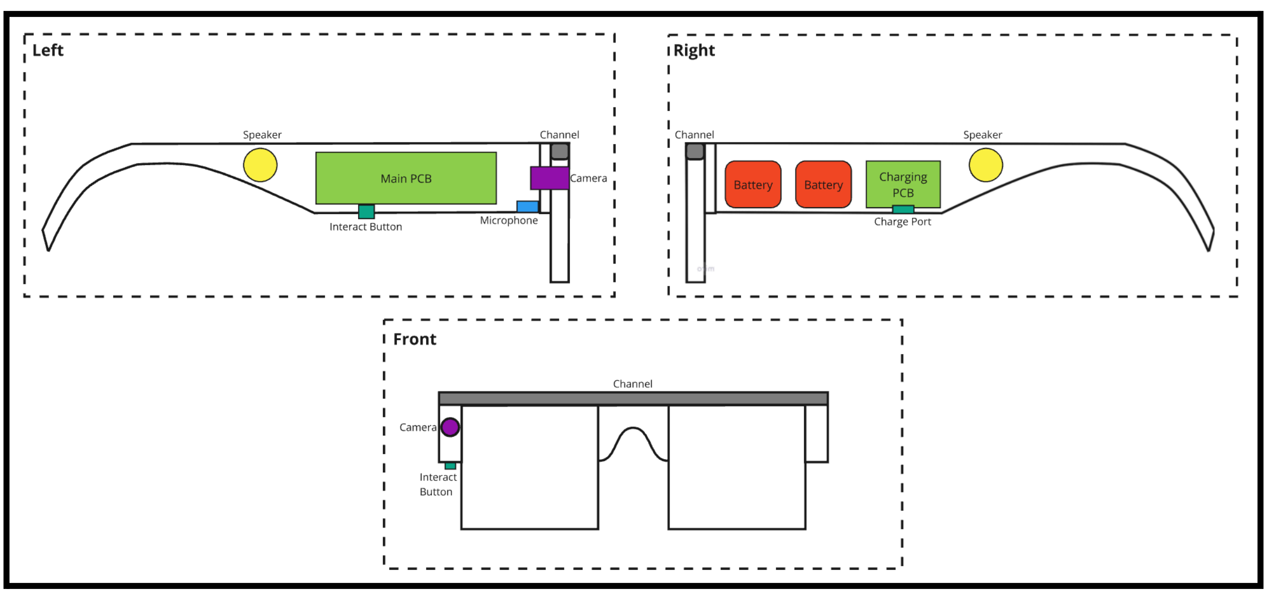

Figure 1: Physical Layout

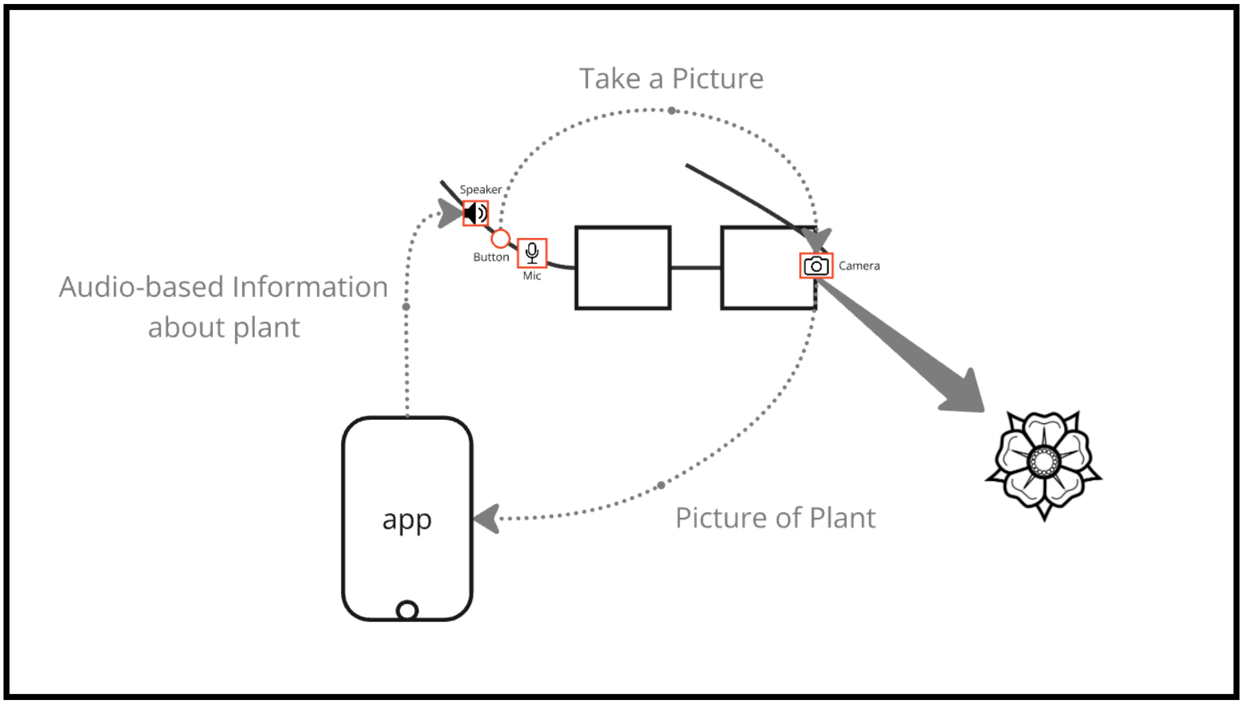

Figure 2: Visual Depiction of Example User Flow

1.4 High-Level Requirements

- **Successful User Flow**
    - The user should be able to look at a plant, press the interact button, and then wait for the system to return the audio of the plant description. Essentially, all components are powered, working, and interfacing correctly with one another.
- **Accuracy**
    - The final prototype should be able to correctly identify plants 75% of the time. This will be based on our application subsystem, which will use a model to determine the plant's classification. We will test the device on 12 plants and expect to get 9 out of the 12 correct. In the case of poor camera framing, we will allow for each plant to be captured up to 3 times.
- **Strong Bluetooth Connection**
    - There should be an uninterrupted Bluetooth connection between the glasses and the mobile device within a 15-foot range. This will be tested by measuring the distance away from the connected device when tests are run.

2 Design

2.1 Block Diagram

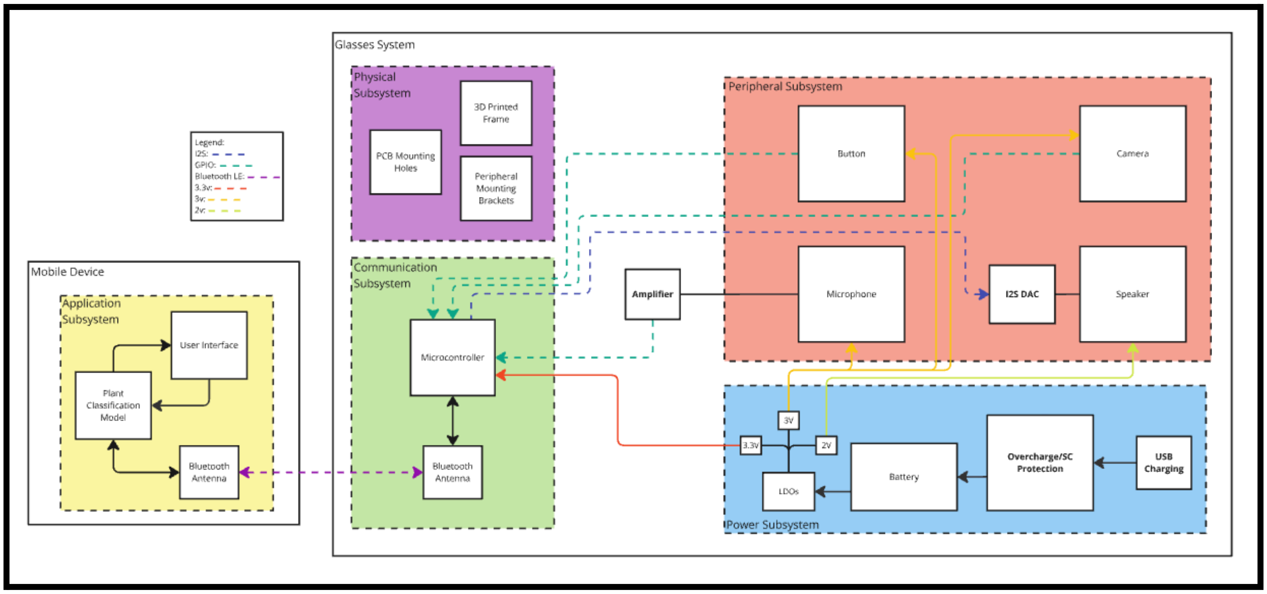

Figure 3: Subsystem Block Diagram

2.2 Power Subsystem

Functionality:

The power subsystem consists of a self-charging battery circuit, and a generic power circuit that will power our peripheral subsystem and communication subsystem. The self-charging circuit will comprise of a rechargeable battery (ASR00011) and a CC/CV linear charger for the battery (TPB4056B2X-ES1R), which is an integrated solution IC that prevents overcharge, over-discharge, and short-circuiting when charging the battery. This IC also supports CC/CV charge from a USB port. LEDs will also be integrated into this part of the circuit to indicate whether or not the battery is being charged. For powering our peripherals and microcontroller, we need to know the operating voltage levels of all our components. The following list is comprised of the operating voltage levels of several parts in our peripheral, communication, and power subsystems:

- Battery (ASR00011): 3V to 4.2V, average rated at 3.7V
- Microcontroller (ESP32-S3-WROOM-1-N16R8): 3V to 3.6V
- Speaker (ADS01008MR-LW100-R): 2V (max)
- Microphone (CMA-4544PF-W): 3V to 10V
- Camera (CAMERA-OV7670): 1.7V to 3V
- Button (B3U-1100P): 3V to 12V

- Audio Stereo DAC (PCM5100APW): 3.3V

Since all of our components that need to be powered are below our battery's output power, we do not need a boost converter to increase output voltage. However, we do need Linear Voltage Regulators (LDOs) to decrease our battery's voltage to our peripherals' operating voltages. The XC6206P302MR CMOS step-down voltage regulator can reduce the voltage going into our four peripherals and microcontroller to a desired amount between 1.5V to 5V.

Contributions:

The power subsystem will contribute mainly to our first high-level requirement: successful user flow. In order for the user to be able to interact with the button, for the camera to take a picture, and for the speaker to output volume, a consistent power source must be able to function and interact with our peripheral subsystem. Additionally, the power subsystem will also contribute to our third high-level requirement: strong bluetooth connection. The microcontroller in our communication subsystem must have an adequate power supply in order to ensure that a Bluetooth connection between the application and the microcontroller is feasible.

Interfaces:

- Peripheral Subsystem – The XC6206P302MR voltage regulator will convert our battery's voltage output to each of the four peripheral's operating voltages as described above.
- Communication Subsystem – The XC6206P302MR voltage regulator will convert our battery's voltage output to each of the four peripheral's operating voltages as described above. In the case that the ESP32 is drawing significant current and the LDO is wasting too much energy as heat, we might revert to a buck converter instead during prototyping (i.e: LM1117-3.3)

Requirements:

- The USB Charging port should be able to charge the ASR00011 battery through the TPB4056B2X-ES1R integrated chip. An LED will be integrated into this part of the subsystem to indicate whether or not the charging circuit is working as intended.
- The ASR00011 battery should be charged to 4.2V.
- The XC6206P302MR voltage regulators should be able to reduce the node voltage to 3.3V, 3V, and 2V for different peripherals and the microcontroller.
    - In the event that the 3.3V output is inefficient for a microcontroller, a buck converter will be used instead (LM1117-3.3).
- Functioning peripherals and microcontrollers will be a final check as to whether or not the power subsystem is working sufficiently.

2.3 Peripheral Subsystem

Functionality:

The peripheral subsystem works as a way of gathering information to be processed by the application system and receiving information that will be passed back to the user. The subsystem will consist of microelectronic devices that the user will interact with such as a button, speaker, microphone, and camera. The information from the camera and microphone will be passed into the communication subsystem corresponding to a push of the button. Once the application system is done processing the information and has data to send back, the communication subsystem will pass back the appropriate audio to the peripheral system's speakers.

Contributions:

The peripheral subsystem helps to satisfy our first high-level requirement: user flow criteria. The push button, speaker, and microphone all interact directly with the user, and the success of this high-level requirement hinges on the functionality of the individual peripherals.

Interfaces:

- Communication Subsystem – The microelectronic components will communicate using wired connections on the PCB to the GPIO pins of the microcontroller
- Power Subsystem – The power subsystem will supply adequate voltages to the microelectronic components in the peripheral system

Requirements:

- The button and microphone must instigate the start of the glasses system: no other process should begin the system
- The speakers must output audible sound. The sound level must be measured as at least 50 decibels
- The microphone must only begin the glasses system with certain prompts: background noise should not be interpreted as a prompt to begin the system

2.4 Application Subsystem

Functionality:

The application subsystem will interface with the communication subsystem to receive and send data for app functionality. The application will be on an iOS device and developed using Xcode tools. The application will utilize the mobile device's onboard Bluetooth to send and receive data. The application itself will receive the image from the communication subsystem, process the image using a plant classification algorithm, generate an audio stream, and then relay this audio stream to be played by the speaker on the glasses.

Contributions:

This subsystem contributes to the user flow and accuracy high-level requirements. The application does much of the computing regarding the user experience. This is because it processes the desired image and sends the audio stream to the speaker on the glasses. In terms of accuracy, the application subsystem is largely responsible for this requirement. It is the functionality of the application to send accurate information.

Interfaces:

- Microcontroller – using Bluetooth LE (Low Energy) to send image and audio data

Requirements:

- Maintain a user interface to relay requested data and establish a connection to glasses.
- Be able to correctly identify 75% of presented plant images.

2.5 Communication Subsystem

Functionality:

The communication subsystem is intended to act as an interface between the peripherals of the glasses and the mobile application. This subsystem will consist of the ESP32-S3-WROOM-1-N16R8 and its onboard PCB Bluetooth LE (Low Energy) antenna. This microcontroller will then communicate with the microphone, speaker DAC, interact button, camera, mobile application, and charging circuit. This communication will occur across a variety of protocols.

Contributions:

This subsystem will contribute to user flow and Bluetooth connection high-level requirements. The communication and interfacing performed by this subsystem will facilitate the experience and direction of the user flow. This is done by being prompted by the interact button, retrieving information from the camera sensor, relaying data to the application, and then sending the I2S stream to the speaker DAC.

Interfaces:

- Interact Button – using GPIO (General Purpose In/Out) pins available on the ESP32 to poll when the button is pressed
- Camera – using a series of GPIO (General Purpose In/Out) pins to establish communication with the camera module according to its datasheet

- Microphone – receives a signal from the microphone after being processed by a supplementary circuit
- Speaker – using a IS2 (Inter-IC Sound) protocol DAC (Digital to Analog Converter) to send the desired audio stream
- Mobile Device – using Bluetooth LE (Low Energy) to interface with the native iOS application

Requirements:

- Have the ability to communicate with all peripheral components:
  - Detect when the button has been pressed
  - Have the ability to receive images from the camera
  - Receive an audio signal from the microphone
  - Send an audio signal to the speaker
- Establish Bluetooth communication with the application subsystem.

2.6 Physical Subsystem

Functionality:

The purpose of the physical subsystem is to create an enclosure for all electronic components such that they can fit into a pair of glasses. This will be done by designing and 3D printing a set of glasses with a hollowed-out frame and custom dimensions to accommodate for PCB and peripheral clearances. The mounting of the components will be facilitated by adding preplanned screw holes for each part.

Contributions:

This subsystem mainly contributes to the user flow high-level requirement. Proper execution of the physical subsystems allowed for the device to be worn properly and the user to interact with peripherals without issue.

Interfaces:

- PCB – will secure the PCB using pre-planned screw holes
- Peripherals – using cutouts in the frame and mounting brackets

Requirements:

- Components secured firmly in place: no internal movement when shaken at a rate of ½ meter per second
- No abrasive textures/sharp edges on the product. This will be measured by rubbing a paper towel against the edges and examining if the paper towel has any tears.
- Physical subsystem must weigh no more than 500 grams. This will be measured by weighing the physical subsystem before attaching any other components

2.7 Tolerance Analysis

2.7.1: Application Subsystem: Modelling Ambiguity

We have identified the subsystem with the greatest difficulty and risk of completion to be the application subsystem. There are several avenues from which we can approach the problem of plant classification. The two directions that stand out are running a model to classify the plant locally on the mobile device or outsourcing the classification to a GPT or similar model in the cloud.

The first approach would be sending all the information to a GPT that is focused on plant identification. This would be a relatively simple backend implementation since all of the computing will be done in the cloud. However, the main drawback is that processing images in a large model such as GPT can be lengthy and therefore delay the response to the user. Additionally, the accuracy of GPT in this context is not well tested and could be a problem for the accuracy high-level goal.

The alternative would be creating or finding a model to run locally on the mobile device. This can be done by sourcing a dataset online and utilizing CoreML. Then, more information about the plant could be generated using GPT or pulled from a pre-compiled database. This approach has the upside of being faster since there is less or no information that needs to be processed in the cloud by GPT.

To accommodate this potential shortcoming, we might expect a delay of 5-10 seconds between procuring the plant image and the audio dialogue being played to the user.

2.7.2: Power Subsystem: LDO vs. Buck Converter

As aforementioned, our battery will output a voltage that needs to be reduced to lower voltages for our peripherals' operating voltages. The microcontroller will interface in a similar manner, as the ESP32 has an operating voltage of 3.3V.

However, if the ESP32 is drawing significant current (300mA+ in WiFi mode), an LDO reducing voltage from a high 4.2V to 3.3V will waste energy as heat, reducing efficiency. This is different from our peripherals, as they will most likely draw low currents, in which case the LDO will be efficient. If in the future we deem that this power loss is too great, we may change from a generic LDO to a buck converter like the AMS1117-3.0 in order to send a lower voltage to the microcontroller.

If the battery is at 4.2V, an LDO drops 0.9V of heat for the operating voltage of the ESP32 at 3.3V. Using Ohm's Law and the Power Dissipation Formula and assuming WiFi mode will draw a current of 300mA:

$$P_{lost} = VI$$
$$P_{lost} = 0.9V * 300mA$$
$$P_{lost} = 270mW$$

A buck converter steps down voltage efficiently by switching at high frequencies and utilizing an inductor to store and release energy. At an efficiency of 90% for the AMS1117-3.0 and 300mA drawn by the ESP32, the power input would be:

$$P_{output} = VI$$
$$P_{output} = 3.3V * 300mA$$
$$P_{output} = 990mW$$

$$P_{input} = P_{output} / Efficiency$$
$$P_{input} = 990mW / 0.9 = 1100mW$$

$$P_{lost} = P_{input} - P_{output}$$
$$P_{lost} = 1100mW - 990mW = 110mW$$

As we can see, at a high enough current, a buck converter saves much more power than a linear voltage regulator. Due to the limited size of our form factor, it is important to prioritize power consumption and physical space. However, LDOs are much cheaper and take up less space than a buck converter like the AMS1117-3.0, so the trade-off is only worth it if the microcontroller draws in significant current. We will test both of these options to see whether or not an LDO for the microcontroller or a buck converter would be better, leveraging the aforementioned tradeoffs.

3 Ethics and Safety

3.1 Ethics

To address ethical concerns in our project we consulted the IEEE Code of Ethics. In the IEEE Code of Ethics Section 7.8 subsection I.1 we are compelled to withhold the privacy of others when creating and testing our project [1]. To ensure this criterion is met, our developed application will not store any images taken to protect privacy.

According to subsection I.5, we are required to offer and accept honest criticism of our technical work [1]. This is an important factor since there could be numerous cases where we fail to see a potential safety or ethical concern within our project, and external opinions could prove helpful in alerting us to these oversights.

Subsection III.10 also compels us to ensure the IEEE Code of Ethics is upheld throughout the entire process of creating our project, and as a group, we will hold each other accountable for every action that could be in violation of the ethics code.

3.2 Safety

- Battery Temperature – During design runs of our product we will closely monitor and measure the temperature of our battery and charging circuit to verify that it never exceeds the appropriate range specified on the datasheet (60 degrees celsius)
- Water Resistance – As our project is designed to be worn by a user in various conditions, the created physical subsystem will contain a water-resistant layer (via a liquid O ring) to protect the electrical components of our design.

4 References

4.1 Datasheets

ASR00011 - Rechargeable Battery
Datasheet:
https://www.mouser.com/datasheet/2/855/ASR00011_70mAh-3078654.pdf

TPB4056B2X-ES1R - Li-Ion Battery Linear Charger
Datasheet:
https://static.3peak.com/res/doc/ds/Datasheet_TPB4056B.pdf

ESP32-S3-WROOM-1-N16R8 - Microcontroller
Datasheet:
https://www.espressif.com/sites/default/files/documentation/esp32-s3-wroom-1_wroom-1u_datasheet_en.pdf

ADS01008MR-LW100-R - Speaker
Datasheet:
https://www.mouser.com/datasheet/2/334/ADS01008MR_LW100_R-1916757.pdf

CMA-4544PF-W - Microphone
Datasheet:
https://www.sameskydevices.com/product/resource/cma-4544pf-w.pdf

CAMERA-OV7670 - Camera, CMOS Image Sensor

Datasheet:
https://www.olimex.com/Products/Components/Camera/CAMERA-OV7670/resources/OV7670.pdf

B3U-1100P - Button, Ultra-small Tactile Switch (SMT)
Datasheet:
https://omronfs.omron.com/en_US/ecb/products/pdf/en-b3u.pdf

PCM5100APW - Audio Stereo DAC
Datasheet:
https://www.ti.com/lit/ds/symlink/pcm5100a.pdf?ts=1739433553738&ref_url=https%253A%252F%252Fwww.ti.com%252Fproduct%252FPCM5100A

XC6206P302MR - CMOS Linear Voltage Regulator
Datasheet:
https://mm.digikey.com/Volume0/opasdata/d220001/medias/docus/5014/XC6206P362PR.pdf

LM1117-3.3 - Buck Converter
Datasheet:
https://www.ti.com/lit/ds/symlink/lm1117.pdf

4.2 Safety/Ethics Documentation

[1] IEEE. “”IEEE Code of Ethics”.” (2020), [Online]. Available: https://www.ieee.org/about/corporate/governance/p7-8.html (visited on 02/12/2025).