Boston University
Electrical & Computer Engineering
EC463 Capstone Senior Design Project

Problem Definition and Requirements Review

Visually Impaired AI Wearable

Submitted to

Michael Hirsch
mhirsch@bu.edu

By

Team #32
Mimir

Team Members

Louis Jimenez-Hernandez louisjh@bu.edu
Heather Li hli9753@bu.edu
Dylan Ramdhan dylram01@bu.edu
Houjie Xiong xhj@bu.ed

Submitted: October 18th, 2024

Client Sign-Off _____

Visually Impaired AI Wearable

Table of Contents

Project Summary

Our wearable device aims to assist the visually impaired with real-time guidance regarding their surroundings. Utilizing a camera, speaker, and artificial intelligence, the device will be able to recognize people, obstacles, and the environment around and dictate the information back to the user. We will also incorporate a mobile application that will add additional capabilities such as image analysis or text dictation. Through a focus on comfort and functionality, we aim to empower visually impaired individuals to be able to navigate the world more easily utilizing an intuitive and unobtrusive design.

1. Need for this Project

For visually impaired individuals, navigating daily life without external assistance can present numerous challenges, such as avoiding obstacles, identifying key environmental features, or recognizing individuals. While there exists various products that aim to resolve this issue, many are quite cumbersome to utilize, limited in functionality, or expensive. Some assistive devices are quite bulky and uncomfortable creating a poor user experience. Other devices are fairly limited in their ability to detect and provide real-time feedback. In addition, many devices are also limited to more controlled environments, thus lacking adaptability to the dynamic environments that individuals face daily. While more advanced devices with complex

capabilities do exist, they are largely locked behind large price tags thus creating a significant barrier to entry for many visually impaired individuals.

Thus, there exists the demand for an intuitive, unobtrusive, cheaper assistive device that has advanced capabilities. To achieve a more intuitive and unobtrusive design, our device focuses on a compact, simple and tactile design delivering a more user friendly experience. To that end, we aim to create a sleek and lightweight form allowing individuals to incorporate the device into their daily life without hassle. In addition, we will utilize verbal feedback allowing for more real time and detailed information being provided to the user. The use of machine learning algorithms will provide our device with the ability to better process the information the device's camera takes in. To assist that capability, we will utilize cloud servers to give our device greater processing abilities, thus allowing more advanced functionality without compromising our design. We will incorporate a mobile app providing further capabilities such as analyzing images from the phone's camera. In conjunction, the application will allow users to have text be dictated further assisting them in everyday tasks. Overall, our design will provide a practical and accessible design enabling easier and more confidence navigation for visually impaired individuals.

2. Problem Statement and Deliverables

2.1 Problem Statement

Visually impaired individuals can struggle with daily navigation without the use of assistive technology. Many available options are currently lacking either in functionality, user experience or are too expensive to be accessible. From researching other alternative products on the market, a lot of them excel in one of these areas at the expense of others. Our proposal aims to solve this issue through a streamlined, ergonomic design prioritizing comfort while maintaining advanced capabilities.

Our design incorporates a slim form factor involving buttons with a tactile design on them creating an easy to operate device. To help maintain the compact nature of this device, we chose to have our device be battery powered and offloaded computational processing onto a cloud server thus allowing us to bolster functionality despite our lightweight design. Through the combination of the speaker and camera, our device is able to handle visual input and provide vocal feedback. Through the use of machine learning algorithms, our device will be able to handle facial recognition while maintaining a small database of individuals for the user. Additionally, the AI will be able to process the environment allowing the user to detect obstacles in their path or receive insight into key features of their surroundings allowing them smoother navigation.

Further pushing the capabilities of our product, we want to incorporate a mobile application that will assist users in everyday tasks. A key feature of the app will be text-to-speech, allowing users to point their phone camera at text and have it read aloud. We also

aim to integrate analysis of images on the user's phone so they can gain insight on photos taken. To provide a better user experience, we want to allow voice commands for easier navigation of the application.

2.2 Deliverables:

Device that implements:

1. Speaker for Audible Feedback
2. Camera for Image Capture
3. Battery for Power Supply
4. Tactile buttons
5. Comfortable to hold and wear

Mature Fine-Tuned Edge AI Model

1. Facial recognition and small database of saved individuals
2. Obstacle detection
3. Environmental detection

Mobile Application

1. Analysis of camera images
2. Text-To-Speech
3. Voice command navigation

3. Visualization

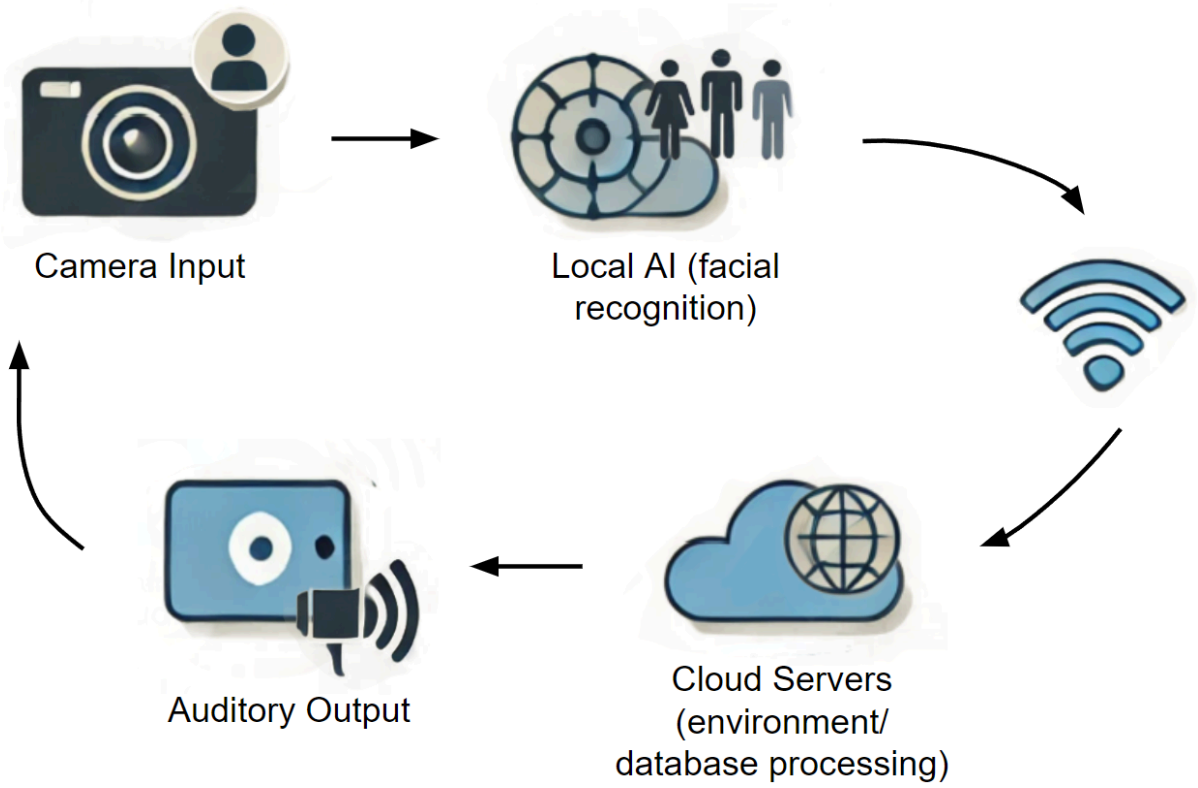

Figure 1. This flowchart illustrates the operational stages of our device that processes the user's surroundings and provides auditory feedback for navigation and awareness. This system consists of five key components: the camera that captures visual data (objects, obstacles, people, and environmental features), local processing on the device using AI algorithms, transmission over bluetooth for datasets that require in-depth analysis or to offload more complex computational tasks, data processing on cloud servers, and generation of the appropriate auditory feedback.

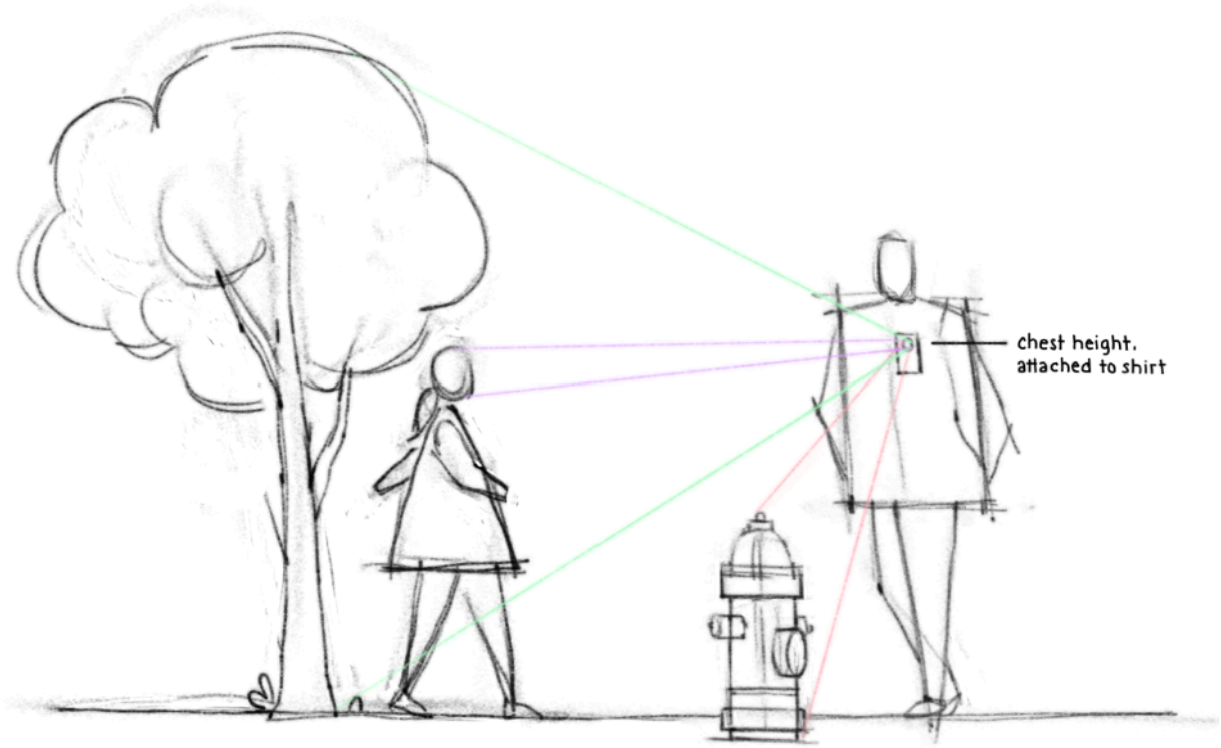

Figure 2. This diagram illustrates the practical use of our wearable device, emphasizing the device's portability and ease of use in everyday scenarios.

4. Competing Technologies

Various other products exist on the market which also aim to be assistive technology for visually impaired individuals during their daily navigation. Some of the biggest competitors are the Ray-ban Meta Smart Glasses and the Microsoft Seeing AI application.

4.1 Ray-Ban Meta Smart Glasses

Ray-Ban has developed smart glasses for visually impaired individuals. The glasses have a lightweight design meant to be reminiscent of their traditional Ray-Ban glasses look. They incorporate cameras into the front of the glasses as well as integration with the Meta-View companion app. The glasses have a built-in microphone which facilitates voice commands to allow for a smooth user experience. While this device is quite capable, the cost of the product being $300 poses a barrier to entry for many users. In addition, there are data and privacy concerns due to their connection to Metal which has faced privacy issues in the past.

4.2 Microsoft Seeing AI

Microsoft Seeing AI is an application published by Microsoft that seeks to assist the visually impaired by utilizing verbal feedback to describe the world around them. The

application has the ability to dictate out short text or documents that the camera captures. There is also facial recognition with the ability to recognize individuals and store that within the application. The application has environmental detection being able to provide details about the environment as well as information on color or lighting giving the user a better understanding of their surroundings. The application does have limitations such as being limited to IOS preventing Android users from accessing it. Due to the processing power required to do AI recognition, the application can be quite draining on the phone's battery.

5. Engineering Requirements

5.1 AI Facial Recognition Model

1. It should recognize people with at least 80% accuracy and within 5 seconds.
2. Constraints: Performance degradation is to be expected on untrained datasets.

5.2 Bluetooth Connection

1. The system should maintain a stable and reliable connection with the mobile device during operation.
2. Constraints: Bluetooth connectivity could be degraded if two devices' distance exceed 15 meters; the application cannot locally run the AI Environment-to-Context model, relying on server-side processing for analysis.

5.3 Battery

1. The battery should provide sufficient power to operate the device for a minimum of 3 hours under typical usage conditions.
2. Constraints: 3 hours is still considered a short duration compared with other portable devices.

5.4 Phone Application

1. While being able to receive images from devices' cameras, upload images to a server for further processing, and retrieve the results from the server, it should also be user-friendly and easily understood by our target user.
2. Constraints: Navigating the application can be tedious for the users depending on the layout and complexity.

5.5 Speaker

1. It should clearly vocalize the text received from the phone in both outdoor and indoor scenarios.
2. Constraints: It could be difficult to hear in noisy environments, possible solutions can include noise cancellation with signal processing.

Appendix A References

STORE, RAY-BAN. “Check out the Ray-Ban | Meta Wayfarer at Ray-Ban.com.” *Ray-Ban.com*, 2018, www.ray-ban.com/usa/electronics/RW4006ray-ban%20%7C%20meta%20wayfarer-black/8056597769440. Accessed 18 Oct. 2024.

Microsoft. “Seeing AI.” *Microsoft Garage*, www.microsoft.com/en-us/garage/wall-of-fame/seeing-ai/.