Boston University
Electrical & Computer Engineering
EC464 Senior Design Project

User's Manual

Mimir

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.edu

Submitted by
04/18/2025

Mimir
User’s Manual

Table of Contents

Executive Summary

Our wearable device aims to assist the visually impaired with real-time guidance regarding their surroundings. Utilizing a camera, speaker, microphone and artificial intelligence, the device will be able to recognize people, obstacles, and the environment and dictate the information back to the user. We will also incorporate cloud computing to assist in our computational power without sacrificing form factor. 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.

Written by Louis

1. Introduction

Cooking is an essential and enjoyable activity, but for the visually impaired, it poses a myriad of challenges. Tasks such as identifying ingredients, reading nutrition labels, taking measurements, operating touchscreen appliances, or following recipes all often have visual cues making them inaccessible to individuals with limited vision. As technology has gotten more modernized, many kitchens are relying more and more on complex and touchscreen heavy technology which makes them less visually impaired friendly. As it stands, the lack of accessibility makes it difficult for the visually impaired to be independent and safe as they cook without the use of accessibility technology.

Generally, navigating daily life and performing daily tasks without external assistance can present numerous challenges for the visually impaired, such as navigation, identifying key environmental features, or the aforementioned cooking. While there exists various products that aim to resolve these issues, 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 or have very limited use cases causing them to not be adaptable to daily life. While more advanced devices with complex capabilities do exist, they are largely locked behind large price tags and are more focused on navigation. Thus, there exists the demand for an intuitive, unobtrusive, assistive device that has empowers the visually impaired as they focus on daily tasks, particularly cooking.

To address this issue, we are proposing Mimir, an AI bodycam which serves to assist the visually impaired in the kitchen. With Mimir, our goal is to provide real-time guidance and feedback for various cooking tasks giving users more independence and confidence as they cook. Our device will feature a compact, simple and tactile design delivering an intuitive and unobtrusive user experience.

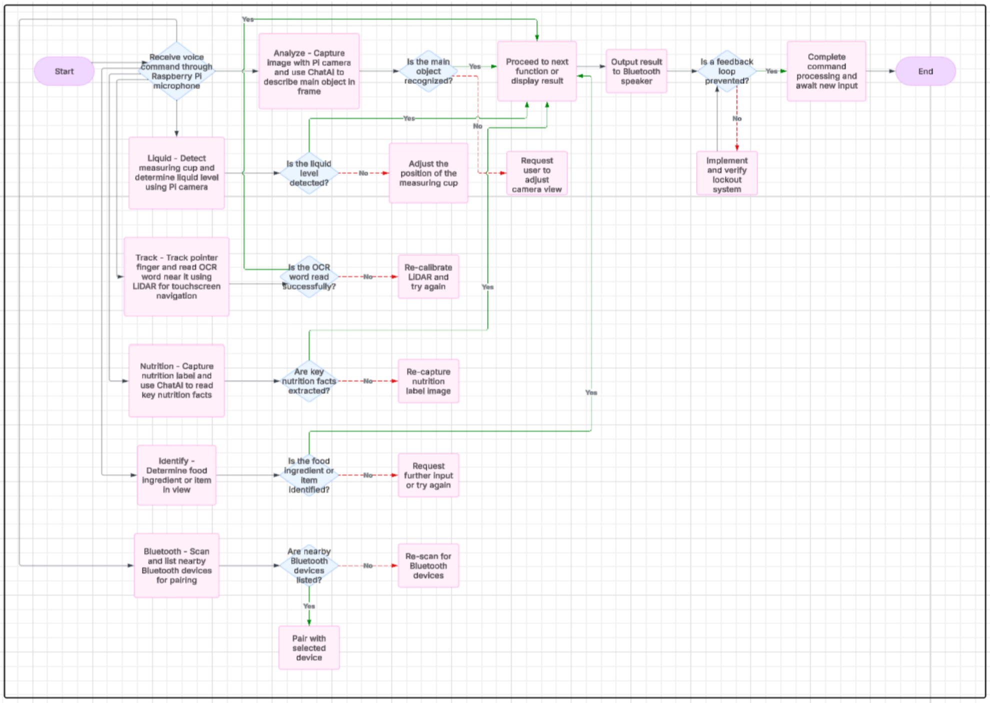
To that end, we aim to create a sleek and lightweight form allowing individuals to incorporate the device into their daily life without hassle. In addition, we will utilize verbal feedback allowing for real time and detailed information being provided to the user. With the use of computer vision and machine learning algorithms, we will provide our device with the ability to better process the information the device’s camera takes in. By leveraging these technologies, Mimir will be able to identify text or ingredients, as well as assist with measurements such as depth, level or distance as the user cooks.

To assist these capabilities, we utilize cloud servers. Through the incorporation of cloud servers, we can provide our device greater processing abilities, thus allowing more advanced functionality without compromising our small form factor. Overall, our design will provide a practical and accessible design enabling easier and more confident navigation for visually impaired individuals.

Written by Louis

2. System Overview and Installation

2.1 Overview Block Diagram



2.2 User Interface

The user interface for the Visually Impaired AI Wearable is designed to be fully accessible, intuitive, and reliable for individuals with little to no vision. It combines hardware buttons with voice-controlled interaction and audio feedback, removing the need for visual menus, screens, or manual software navigation.

Upon powering on, the device automatically boots into an idle listening mode. A dedicated physical button must be pressed to activate the microphone for listening. This intentional design prevents accidental triggering of commands from ambient noise or the device’s own audio responses. Once the microphone is active, the system listens for specific voice commands such as “analyze,” “track,” “liquid,” “identify,” “nutrition,” “bluetooth,” and “echo.” Each of these triggers a unique, locked-out function, ensuring that only one feature is active at a time.

After a command is received and confirmed, the system disables the microphone to avoid misrecognition during audio feedback. For example, during the “analyze” function, the Raspberry Pi camera captures an image of the user’s surroundings and uses an AI model to describe the main object, which is then read aloud through a Bluetooth-connected speaker. Similarly, in “track” mode, the LiDAR camera follows the user’s fingertip on a touchscreen and reads the word closest to it using OCR. In “liquid” mode, the system identifies measuring cup markings and gives an estimate of the liquid level using standard kitchen units.

To test or confirm that the interface is responsive, the “echo” command will cause the system to respond with a simple “echo” playback—useful for debugging, demonstration, or setup checks. Because the interface is centered around clear physical and auditory interactions, users are not required to navigate screens or visual elements. All responses are delivered via text-to-speech, and all inputs are either spoken or initiated via physical buttons.

2.3 Physical Description

Mimir’s modular hardware design is centered around a Raspberry Pi 5, which acts as the main processing unit and is enclosed within a lightweight, 3D-printed case for protection and structure.

Mounted directly onto the Raspberry Pi’s 40-pin GPIO header is a battery management board, providing power from a rechargeable battery pack. This stacked configuration reduces cable clutter and allows the system to operate independently of an external power source or monitor. The battery pack is integrated into the enclosure, creating a self-contained unit that powers all onboard components.

Attached to the bottom of the enclosure is a LiDAR camera, positioned to face forward and downward. This placement allows the camera to effectively track hand and finger movement, especially useful for the touchscreen navigation functionality. A Raspberry Pi Camera is mounted on the front of the device, giving it a clear forward-facing view to capture objects, ingredients, or environments for analysis and interpretation by AI models. This is particularly important for functions such as “analyze,” “identify,” and “nutrition.”

The entire unit is designed to be worn at the collar using a sturdy rear-mounted attachment clip. This clip allows the device to be fastened to a shirt, jacket, or lanyard so that the camera and sensors remain aligned with the user’s line of sight and hand gestures. The positioning is intentional to ensure optimal audio pickup, camera framing, and LiDAR sensing.

Two tactile push-buttons are located on accessible parts of the enclosure one on the top of the device and one on the left when worn.These buttons are placed for quick access and have an audible and tactile response for users with limited vision.

2.4 Installation, Setup, and Support

The Visually Impaired AI Wearable is designed to run as a fully self-contained system, requiring no monitor, keyboard, or manual intervention after initial deployment. All software and hardware dependencies are pre-installed and configured, allowing the device to launch automatically on startup and function independently in real-time.

The system is built on a Raspberry Pi 5 with a Raspberry Pi Camera and an OAK-D Lite LiDAR camera. These peripherals are physically mounted to the board and enclosed in a custom 3D-printed shell, which also secures a mini USB microphone for user interaction. Power is supplied via a battery pack connected to the GPIO header, allowing for portable and continuous use without reliance on external power. Once assembled, the user simply powers on the Raspberry Pi using a toggle or button. Upon startup, the operating system boots into a headless mode and automatically launches the main Python script via a configured system service.

This main script initializes all critical components—including voice activation, camera input, LiDAR positioning, and Bluetooth audio output—without requiring user input. The software stack includes pre-installed libraries such as depthai, vosk, openai, pyttsx3, opencv-python, and numpy. For offline use or improved privacy, the system supports running a local Ollama container hosting the LLaMA model, which is also launched automatically on boot if configured.

Voice interaction serves as the primary user interface. The microphone continuously listens for specific trigger words like "analyze," "liquid," or "track." When a valid command is detected, the system enters a lockout mode to execute the desired function in isolation—whether that be identifying ingredients, reading touchscreen labels, or analyzing a liquid's volume—then returns to its idle listening state once complete. To prevent accidental re-activation during audio feedback, speech output is automatically muted during command execution and re-enabled once the function concludes.

Users do not need to connect a monitor, install software, or input commands to operate the wearable. If a reset is required or issues arise, the device can be safely restarted using the built-in power button. In case of Bluetooth or audio output problems, users can re-pair the speaker by holding the pairing button during boot, which re-triggers the pairing mode.

Written by Heather

3. Operation of the Project

3.1 Operating Modes 1: Normal Operation

In its normal operating mode, the Visually Impaired AI Wearable functions as a fully autonomous, voice-activated assistive device. Upon power-up, the system runs headlessly—meaning it begins operation automatically without the need for a monitor, keyboard, or mouse. The device remains in an idle standby state until the user presses the primary button, which activates the microphone. Once activated, the system listens for specific voice commands such as "track," "analyze," "liquid," "nutrition," "identify," "bluetooth," or "echo."

Each command initiates a distinct function:

- Track: enables the system to detect the user's fingertip using LiDAR and reads aloud the nearest recognized word via OCR
- Analyze: captures an image with the Pi Camera and uses the OpenAI API to describe the most prominent object in the scene
- Liquid: identifies a measuring cup and estimates the liquid level in cups using visual features and text recognition
- Nutrition: captures a nutrition label and extracts important dietary facts using AI
- Identify: names or categorizes the ingredient or object in view
- Bluetooth: scans for nearby devices for audio pairing
- Echo: simply responds with "echo" to confirm that the voice loop is functional

Each command locks out other functions until it finishes execution. The secondary button can be used to interrupt the current routine or re-enable the listening mode if needed. Once the task completes, the system returns to idle and awaits the next button press to begin the cycle again. All AI processing is handled onboard or via minimal calls to external APIs, ensuring a responsive and low-latency user experience.

3.2 Operating Modes 2: Abnormal Operations

In the event of an abnormal operating condition—such as loss of Bluetooth audio output or incomplete API response—the system is designed to safely revert to its idle state. Errors such as camera disconnection or microphone failure are handled via exception detection in the Python script, and the user is notified via audio cues. The system will not attempt to execute a function if the necessary hardware is not detected or initialized properly.

Another potential abnormal condition includes accidental re-triggering of a command during audio playback. To prevent recursive input or misinterpretation of output as a command, the system disables the microphone while speaking. If commands are not recognized due to unclear input or ambient noise, the system will time out and await another button press to reactivate listening, rather than entering an undefined state.

3.3 Safety Issues

While the wearable is a low-voltage, battery-powered device, several safety considerations have been addressed. The use of a UPS battery pack with safe shutdown support ensures that the Raspberry Pi is not abruptly powered off, reducing the risk of SD card corruption or component damage. Additionally, the GPIO stacking method used to connect the power supply has been tested for secure mechanical and electrical fit to avoid shorts or instability.

From a user safety perspective, the device is lightweight and clips onto clothing without obstructing movement or requiring contact with skin. The camera and LiDAR are passively cooled. Because the system avoids excessive wiring or heat generation, it is safe for extended use in indoor environments like kitchens. All audio outputs are played at safe volume levels, and no physical interaction with the internal electronics is required during operation.

Written by Houjie

4. Technical Background

The Visually Impaired AI Wearable operates as a compact, embedded AI system designed to run multiple perception tasks autonomously on a low-power edge device. The central processor is a Raspberry Pi 5, chosen for its quad-core CPU and expanded RAM, which enables on-device computation of both computer vision and natural language tasks without relying on external servers.

From an embedded systems perspective, the device is optimized for startup autonomy, booting directly into the voice-activated service pipeline without requiring user interaction with a monitor or input peripherals. Hardware modules—including a LiDAR camera, Pi camera, USB microphone, Bluetooth speaker, and battery-backed UPS board—are integrated directly into the Raspberry Pi’s GPIO and USB interfaces, forming a cohesive wearable system that runs headlessly.

The software pipeline combines traditional computer vision with AI-based language processing. Visual data captured from the Pi camera is processed using OpenCV for image preparation and framing. Depending on the active voice command, the image is either analyzed using Tesseract OCR for reading text near a pointed finger (enabled by the LiDAR’s depth mapping), or sent through a language generation model via the OpenAI API to identify and describe objects in context—especially useful for grocery items or nutrition labels.

Audio input is processed locally using Vosk, a lightweight offline speech recognition model ideal for embedded applications. Once a keyword is detected, the device enters a locked-in function state, executing a predefined routine (e.g., tracking, reading, identifying) to ensure processing resources are focused and user feedback is consistent. After completion, the system reverts to an idle state, ready for the next button-initiated voice trigger.

To deliver responses, the system uses pyttsx3, a text-to-speech engine that operates entirely offline, converting AI-generated or OCR-based outputs into spoken feedback without latency or connectivity concerns. This approach ensures that all core functionality—including AI interpretation, speech recognition, and visual feedback—operates locally on the embedded system, allowing the wearable to function reliably in kitchens or environments where internet access may be intermittent or unavailable.

Written by Dylan

5. Relevant Engineering Standards

Mimir is designed in accordance with the IEEE 360-2022 standard, which provides a comprehensive framework for wearable consumer electronic devices. This standard defines terminology, categorization, and technical requirements for wearables, including aspects such as security, suitability for wear, and functional testing methods. The device’s architecture and components are selected and assembled from original equipment manufacturers (OEMs) whose parts comply with relevant international and industry standards.

All wall-socket power supplies used with Mimir are certified and approved to meet regional electrical safety standards. The power system is designed to optimize battery life and ensure safe operation, incorporating over-current, over-voltage, under-voltage, and over-temperature protection features. Battery selection and integration strictly follow manufacturer specifications to prevent hazards and ensure longevity.

Regarding communication protocols, Mimir employs standard wireless communication protocols, including IEEE 802.11 (Wi-Fi) and IEEE 802.15 (Bluetooth), ensuring interoperability with smartphones, cloud servers, and other smart devices.

Software libraries used in Mimir are sourced from reputable, approved repositories and comply with open-source or commercial licensing requirements. Embedded software is documented according to established guidelines to facilitate maintenance, reuse, and consistent interface design. All software and firmware updates are thoroughly tested to ensure compatibility and security.

The device’s tactile and auditory feedback mechanisms are designed to be intuitive and accessible, following user-centered design principles and standards for assistive technology. The design process incorporates feedback from visually impaired users to ensure that the device meets real-world needs and usability expectations.

Subsystems and assemblies are constructed from standardized, approved parts to ensure seamless integration and serviceability. Components are selected for compatibility and compliance with relevant standards, minimizing the risk of failure and maximizing product lifespan.

All aspects of the device, including hardware schematics, PCB layouts, bill of materials, and software, are thoroughly documented. Documentation is maintained in a structured and accessible format to support future updates, regulatory compliance, and user support.

All of our team members were encouraged to stay informed about relevant standards and best practices throughout the product development. By rigorously adhering to these engineering standards, Mimir aims to deliver a safe, reliable, and user-friendly assistive device that empowers visually impaired individuals with greater independence and confidence in the kitchen and beyond.

Written by Heather

6. Cost Breakdown

Project Costs for Production of Beta Version			
	Quantity	Description	Unit Cost
1	1	Raspberry Pi 4	$62.99
2	1	Arducam High Definition Raspberry Pi Camera	$55.99
3	1	Oak-D Lite Fixed Focus LiDAR Camera	$149.00
4	1	SunFounder USB-C Microphone	$7.99
5	1	DEVMO Mount Push Buttons	$10.99
6	1	Geekworm Power Management Board	$43.00
7	1	CPZZ 18650 Rechargeable Battery	$9.59
		Total Cost	$339.55

7. Appendices

7.1 Appendices A – Specifications

Requirement	Value, Range, Tolerance, Units
Dimension	10x10x5cm
Weight	185g

7.2 Appendices B – Team Information

Louis Jimenez-Hernandez (louisjh@bu.edu)

Heather Li (hli9753@bu.edu)

Dylan Ramdhan (dylram01@bu.edu)

Houjie Xiong (xhj@bu.edu)