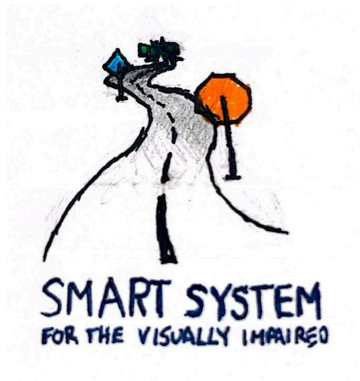


Boston University
Electrical & Computer Engineering
EC463 Senior Design Project

First Semester Report

Smart System for Visually Impaired



Submitted to

Michael Hirsch

mhirsch@bu.edu

8 Saint Mary's Street, Boston, MA 02215

by

Team 29

Smart Bears

Team Members

Lukas Chin

lchin10@bu.edu

Jake T. Lee

jaketlee@bu.edu

Shamir L. Legaspi

slegaspi@bu.edu

David Li

dav@bu.edu

Jason Li

jli3469@bu.edu

Submitted: 12/09/2024

Table of Contents

Executive Summary	2
1.0 Introduction	3
2.0 Concept Development	5
3.0 System Description	8
4.0 First Semester Progress	10
5.0 Technical Plan	13
6.0 Budget Estimate	17
7.0 Attachments	18
7.1 Appendix 1 – Engineering Requirements	18
7.2 Appendix 2 – Gantt Chart	19
7.3 Appendix 3 – Technical References	20
7.4 Appendix 4 – Drawing and Schematics	21

(right click on Table of Contents to update fields and page numbers automatically)

Executive Summary

Smart System for Visually Impaired
Team 29 - Smart Bears

In the United States, there are over 20 million people who are visually impaired, with about 1 million of those with blindness. Every day these visually impaired individuals have trouble navigating their surroundings with many needing to use white canes or guide dogs for assistance. While white canes and guide dogs provide great aid for everyday navigation, our Smart System for Visually Impaired offers even greater enhanced mobility through wearable technology that combines obstacle detection, real-time object recognition, live haptic and auditory feedback, and speech control. Smart System for Visually Impaired uses cutting edge computer vision techniques to detect everyday hazards, and it listens to the user for quick and automatic speech input. Our Smart System for Visually Impaired combines a Smart Cap and an Intelligent Wrist-Wearable wirelessly to provide the user with exceptional environmental awareness and navigation confidence.

1.0 Introduction

The visually impaired face significant challenges in navigating both indoor and outdoor environments, particularly in unfamiliar settings. These mobility challenges primarily involve avoiding obstacles, detecting ground level changes, and navigating street crossings. Despite having traditional aids like white canes and guide dogs, many visually impaired individuals still lack adequate accessibility to easily navigate independently. A guide dog comes with significant ongoing expenses, including specialized food that can cost up to \$150 monthly, regular veterinary care, grooming services, and unexpected medical expenses. Guide dogs also cannot serve indefinitely and will eventually need to retire. They may also face unforeseen health issues that could end their service prematurely. These limitations make it difficult to emotionally and physically rely on guide dogs for independent transportation and navigation. White canes on the other hand, while proven to be extremely effective in assisting visually impaired people in navigating their surroundings, also create unsafe situations as they can get stuck in sidewalk cracks and crevices, potentially harming the user. Additionally, they can not detect vehicles or obstacles that are further than the cane's reach.

We introduce, “Smart System for Visually Impaired” which aims to provide another solution for visually impaired individuals to navigate their surroundings better, and while the product is not designed to fully substitute current solutions, it strives to be a supplement to give the user even greater control and confidence when commuting every day. Our approach in creating “Smart System for Visually Impaired” is to create a portable, smart, and easy to use device that combines numerous methods of user feedback and environment detection to allow the visually impaired user to better navigate their

surroundings. We introduce a Smart Cap device and a Wrist-Wearable device that act together as a smart system pair. The Smart Cap combines computer vision techniques with real-time object recognition to provide comprehensive environmental awareness through relaying that information to the Wrist-Wearable device wirelessly. Both the Smart Cap and Wrist-Wearable feature different ways to alert the user ranging from haptic feedback vibration and speaker audio.

Smart System for Visually Impaired will address navigational challenges of visually impaired individuals through this comprehensive multi-modal approach. The system enhances environmental awareness through the Smart Cap object detection capabilities as well as the Wrist-Wearable haptic feedback mechanisms. Unlike traditional aids, our pair of smart devices offers dynamic, real-time environmental detection while maintaining hands-free operation.

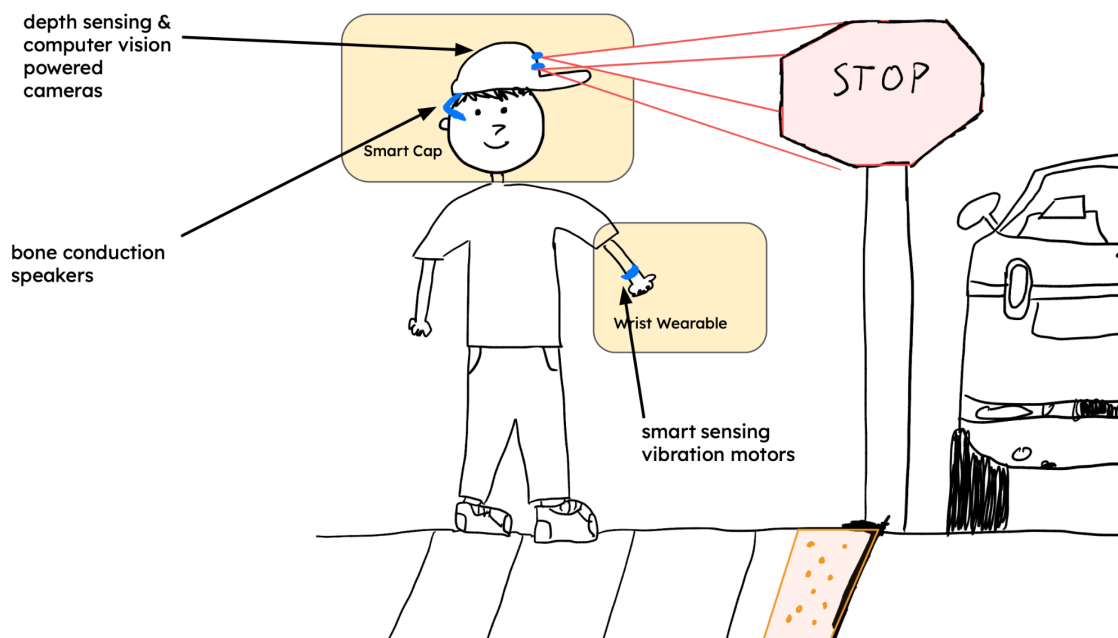


Fig 1. Mock-up of Smart System for Visually Impaired, showing Smart Cap and Wrist-Wearable use case.

2.0 Concept Development

Based on our customer's needs and requirements as concisely identified in Appendix 1, we've carefully created an initial proposed concept to not only tackle problems visually impaired individuals face but also tailor our design to fit within the constraints and preferences of our customer. Our customer requested a solution that would be hands-free, utilizes relevant object detection techniques, and is light and portable to use. Given these constraints, we honed in on the hands-free aspect of the design and as a group we ideated on creating a system that would replace a conventional white cane. Our wrist-wearable device instead will act like a white cane, utilizing numerous lidar and distance sensors to detect the ground below the visually impaired user. The measurements from the sensors will then be read by the microcontroller on the device and translated into haptic feedback for the user—essentially simulating a white cane but without the need of a white cane. With this approach, we believe we could revolutionize the way a visually impaired user interacts with their surroundings as now they can benefit from the feedback of a white cane while being hands-free, allowing them to gain greater independence on top of the improved confidence in navigation. This system would also involve the smart cap's multi-camera computer vision algorithm to provide even more information for the user to navigate every day.

However, after much deliberation on the possibility of this smart system combined with extensive feedback from Shark Tank, we had to go back to the drawing board, as replicating white cane feedback via lidar/distance sensors proved to be a more difficult task than first anticipated. The first realization is the polling rates of these sensors would have to be tremendously high in order to pick on instantaneous movement and changes to the ground below the user. The second realization is that the sensors

would need a completely separate support sensor to aid in rotational movement of the user's wrist because with every wrist movement, the sensors would have to recalibrate to adjust. Lastly, the lidar/distance sensors would have to be extremely accurate as well as portable and energy efficient enough to be worn as a wrist-wearable device. All of these circumstances led us to instead pivot from replacing the white cane to supplementing it. In doing this, our ideas of haptic feedback through environmental awareness is still maintained with the possibility of hands-free operation, but emphasis is placed on the benefits of conjunctional use with a white cane.

Our newly iterated solution combines two key components: a smart cap and a wrist wearable device. This dual-device approach addresses the problem through a visual processing system and feedback mechanics.

In the smart cap, we implemented the YOLOv3 object detection model, a dual Raspberry Pi camera setup, Whisper.cpp, and a feedback mechanism using bone conduction speakers. The YOLOv3 model is used through OpenCV for real-time object recognition, in which the user will be alerted when it sees signs, close objects, etc. We decided on adding two cameras in our PiCam setup for more accurate depth mapping and distance calculation. Whisper.cpp has speech-to-text capability, which allows for user queries through voice commands. Finally, the bone conduction speakers will communicate with the three visual processing systems and inform the user of the recognized objects in front of him/her, as well as any output from their spoken user query.

We plan to implement two additional feedback mechanics in our wrist wearable device: vibration motors and a high-visibility LED system. The vibration motors will be an additional alert from the OpenCV data; when objects are too close to the user, the

vibration motors will create a greater vibration. The LED system was added to our wrist wearable to allow for the user's safety.

3.0 System Description

We will provide a system with a smart cap and smart wrist wearable to increase the navigational abilities and safety of visually impaired users. At a high level, the product will consume visual information and transform it into audio and haptic signals that can be interpreted by visually impaired users.

The smart cap will capture images of the area in front of users, identifying navigational objects and obstacles with an 80% detection rate and a mapping to their relative distances. Users will be notified of relevant detected objects through audio signals transmitted through bone conduction speakers at 60-70 decibels, a range that is audible without causing ear damage. Users will also be able to query the system through speech to tailor what the system considers to be the most relevant objects. The responsiveness of the system will vary based on whether the user makes a query or not. Without a query, the system will have a response time such that a picture can be taken, processed, and converted to audio output at least once per second. When responding to a query, the response time will allow for pictures and audio to be captured, processed, and converted to output at least once every four seconds (plus the duration of the query made) to maintain conversational pace. The goal of providing users with descriptions of objects in front of them is to allow them to make navigation decisions. Because of this, relevant objects are defined as objects that help users make these decisions (stop signal at crosswalk -> stop, go signal at crosswalk -> go, dead end -> turn). Relevant objects will be detected at a rate of 65%.

The smart wrist wearable will augment the output of the smart cap, outputting signals that suggest a navigational action based on the relevant object identified by the cap. Given a relevant object and suggested action as input from the smart cap, the wrist

wearable will be able to output distinct signals that reflect the suggested action within 100 milliseconds. The wrist wearable will also house an intelligent visibility system. It will have an LED array capable of producing 50 lumens when a threshold is reached, detected by the output of a photoresistor.

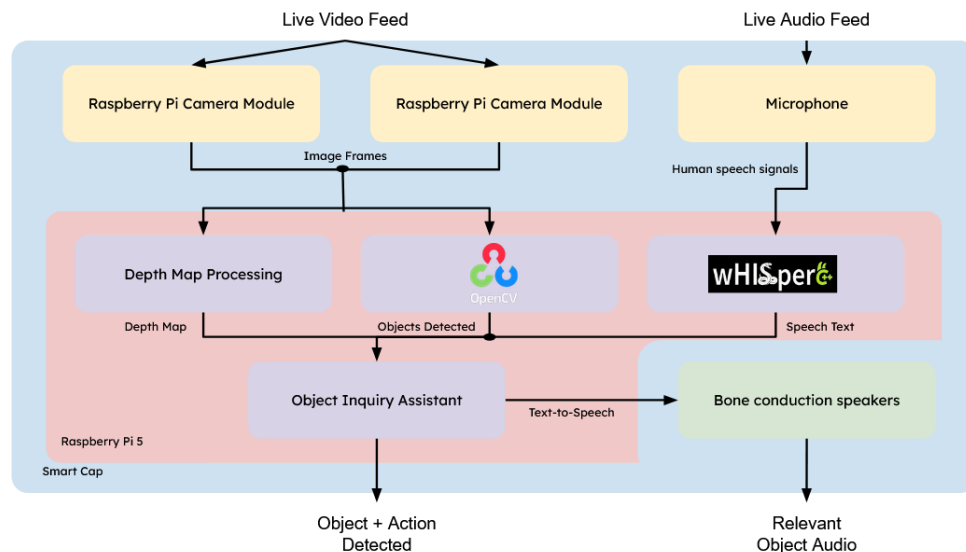


Fig 2. Block diagram of smart cap showcasing how video and audio are transformed into audio to be consumed by the user and signals to be used by the smart cap.

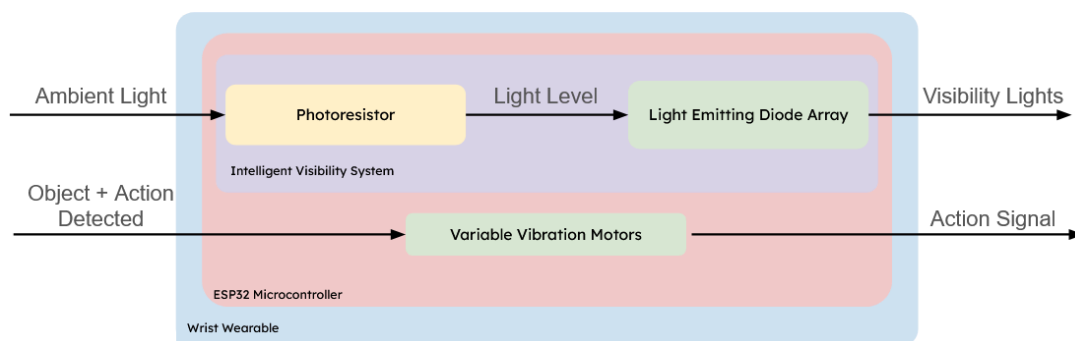


Fig 3. Block diagram of smart wrist wearable showcasing how light and detected objects are transformed into outputs for the user.

4.0 First Semester Progress

This semester, our team focused on establishing a proof of concept for the “Smart System for Visually Impaired.” We worked on integrating hardware components, improving communication across devices, and refining our computer vision capabilities. One key limitation we encountered was the lack of computing power in the Raspberry Pi Zero W, which made it challenging to handle the computational demands of real-time computer vision processing. To address this in the next phase, we plan to upgrade to a Raspberry Pi 5 and add an additional camera module to enable stereo vision and depth map calculations, enhancing the system's ability to detect obstacles and provide more accurate feedback.

Despite the limitations of the Raspberry Pi Zero W, we successfully established robust connections and communication between all our microcontrollers and auxiliary devices. This achievement was crucial for ensuring smooth operation and feedback from both the Smart Cap and the Wrist-Wearable device.

For the computer vision aspect, we found success with the YOLOv3 real-time object detection model, training it on the COCO dataset to detect common objects. While the model has shown promising results, we are continuing to fine-tune it to detect more specific objects, such as road signage and text on the street. Although we have not fully achieved this capability yet, our extensive research and testing have brought us closer to the goal.

A significant challenge we faced was minimizing the latency in streaming video from the Raspberry Pi Camera Module to our processing source. Initially, there was a 4-second delay in video streaming, which we aimed to reduce. In the prototype, we opted

to keep the Raspberry Pi lightweight by using the Raspberry Pi Zero W solely for video capture and streaming. The video feed was then sent to an external processing source, such as our laptop, for object detection. While the latency was an issue, this setup allowed us to evaluate our approach and consider improvements for future prototypes.

To ensure the accessibility of our product, we incorporated a text-to-speech library that announces the objects detected by the OpenCV model, making the system more intuitive for visually impaired users. This feature will be crucial as we continue to refine the user experience in future iterations.

Below are the key results from our prototype testing:

- **Hardware Communication and Integration:**

- All microcontrollers (Raspberry Pi Zero W, TinyPico) successfully established stable connections with the router and the server device.
- The Raspberry Pi Camera Module streamed video to the local computer with minimal interruption.

- **Computer Vision:**

- The YOLOv3 model successfully detected common objects from the COCO dataset with an accuracy of 80%.
- Preliminary research into fine-tuning the model for detecting road signage and street text is ongoing.

- **Latency and Streaming:**

- Initially, there was a 4-second delay in video streaming from the Raspberry Pi Camera Module to the laptop. Plans are in place to improve this latency in the next prototype.

- **Haptic Feedback:**

- The system successfully activated the two distinct vibration motors (on/off and varying intensity) based on proximity and obstacle detection.
- The vibration feedback patterns (triple pulse, gradual intensity increase, and constant vibration) were tested, with users successfully differentiating between them 90% of the time.

- **Text-to-Speech Accessibility:**

- Integrated a text-to-speech library to announce detected objects, improving accessibility for visually impaired users.

Overall, the first semester was marked by significant progress in hardware integration, communication, and computer vision research. Our next steps will focus on enhancing computing capabilities, improving object detection, and reducing latency to ensure a more responsive and efficient system for users.

5.0 Technical Plan

Task 1: Integration of Dual Pi Camera Modules

- **Description:** This task involves configuring and integrating two Raspberry Pi Camera Module v2 devices with the Raspberry Pi 5 to enable stereo vision. The cameras, each capable of 8MP resolution and 1080p video at 30 fps, will be connected to the Raspberry Pi 5 using CSI ports. The Raspberry Pi 5, powered by a Broadcom BCM2712 quad-core processor with 4GB RAM, will run Raspbian OS and utilize Python with OpenCV for camera interfacing. The process includes connecting the cameras, installing OpenCV, and configuring scripts for synchronized image capture. Testing will verify the synchronization, resolution, and frame rate performance, ensuring the cameras provide synchronized 1080p image streams.
- **Deliverable:** Both cameras connected and functioning, providing synchronized image capture.
- **Lead:** Jason
- **Assisting:** Jake

Task 2: Development of Stereo Vision Algorithm

- **Description:** This task focuses on implementing a stereo vision algorithm to compute depth maps from the dual camera setup. Using a Semi-Global Matching (SGM) or Block Matching algorithm implemented in Python with OpenCV, the system will produce depth maps with a resolution of at least 480×320 pixels and a processing time under 1 second per frame. The cameras will first undergo stereo calibration using OpenCV's calibration functions, followed by implementing and testing the algorithm on stereo image datasets, such as Middlebury or KITTI. The depth map accuracy will be optimized to support precise navigation for visually impaired users.
- **Deliverable:** A functional depth mapping system capable of determining distances between detected objects and the user.

- **Lead:** Shamir
- **Assisting:** David

Task 3: Smart Cap Housing Design

- **Description:** This task involves designing and fabricating a durable and lightweight housing unit for the smart cap, which will securely hold the Raspberry Pi 5 and two Pi Camera Modules. The housing will be modeled using CAD software like SolidWorks, with specific provisions for securing the hardware and ensuring heat dissipation. The design will be fabricated using a 3D printer with PLA material and tested for durability, fit, and weather resistance. After installing the hardware in the housing, adjustments will be made based on real-world testing feedback to ensure ergonomic use.
- **Deliverable:** A durable, lightweight housing unit that meets ergonomic and safety standards.
- **Lead:** Jake
- **Assisting:** Lukas

Task 4: Battery Housing Design

- **Description:** This task involves designing a rechargeable battery housing unit to power the Raspberry Pi 5, meeting specifications for weight, battery life, and heat dissipation. The power source will be a 3.7V, 5000mAh lithium-ion battery with a boost converter to provide a stable 5V output. CAD software will be used to design the housing, which will be fabricated using durable, lightweight materials like ABS plastic. Testing will include charging cycles, heat dissipation under load, and endurance tests to confirm 6+ hours of runtime for typical use.
- **Deliverable:** A tested battery housing capable of supporting the system's power needs under real-world conditions.
- **Lead:** Jake
- **Assisting:** Lukas

Task 5: Wrist-Wearable Design

- **Description:** This task includes designing and building an ergonomic wrist-wearable device to house the ESP32 microcontroller, haptic motors, LED array, and photoresistor sensor. The ESP32 will handle motor control and feedback, while the LED array will increase the user's visibility to vehicles and pedestrians. The wearable will be designed in SolidWorks, fabricated with a flexible TPU material for comfort, and tested for fit, durability, and functionality. Feedback patterns from the motors will be integrated to convey directional cues based on object detection data.
- **Deliverable:** A fully assembled, wearable device tested for comfort, durability, and functionality.
- **Lead:** Lukas
- **Assisting:** Jake

Task 6: Training a New YOLOv3 Model

- **Description:** This task involves fine-tuning and training the YOLOv3 object detection model to recognize road signage and parse street text. Using a dataset collected from public sources, including the COCO dataset and custom images, the model will be trained on an NVIDIA GPU using BU's Shared Computing Cluster. The goal is to achieve a detection accuracy above 90% for specific objects and text relevant to navigation. After training, the model will be optimized for deployment on the Raspberry Pi 5 using TensorFlow Lite to ensure efficient inference.
- **Deliverable:** A trained YOLOv3 model with improved recognition accuracy for signage and text.
- **Lead:** Jason
- **Assisting:** Shamir and Lukas

Task 7: Central Processing Unit Development

- **Description:** This task requires building a central processing unit (CPU) to integrate inputs from object detection, depth mapping, and speech processing. The

CPU, implemented on the Raspberry Pi 5, will run Python scripts to merge object detection and depth data and process user queries via a speech-to-text interface powered by Whisper.cpp. Outputs will be delivered through a text-to-speech system using the pyttsx3 library. Testing will ensure real-time processing (<1 second response time) and accurate feedback for user queries like “What is to my right?”

- **Deliverable:** A CPU capable of integrating inputs and providing accurate audio feedback.
- **Lead:** Lukas
- **Assisting:** Shamir and David

Task 8: Haptic Motor Tuning for Directional Feedback

- **Description:** This task involves calibrating the haptic motors to provide intuitive directional feedback based on object detection data. Using an ESP32 microcontroller, motor feedback will be mapped to object positions, ensuring motors on the user’s left or right activate based on the detected location of obstacles. Feedback intensity and duration will be adjusted using PWM signals for clarity. Testing will involve user trials to confirm the feedback is intuitive and effective for navigation.
- **Deliverable:** A calibrated system providing intuitive directional feedback to the user.
- **Lead:** David
- **Assisting:** Jake and Jason

Milestone	Completion Date	Associated Tasks
Dual Cameras Integrated	01/31/2025	Task 1.4
Housing Units Fabricated	02/14/2025	Tasks 1.7
Stereo Vision Algorithm Ready	02/24/2025	Task 2.8
YOLOv3 Model Trained	03/03/2025	Task 2.9
CPU Development Complete	04/07/2025	Task 3.5
Haptic Feedback Calibrated	04/13/2025	Task 4.3

6.0 Budget Estimate

Item	Description	Cost
1	Teensy40 ESP32 Microcontroller	\$23.80
2	TinyPICO ESP32 Microcontroller	\$130
3	Bone Conduction Headphones with Microphone	\$17.88
4	Mini Vibration Motors DC 3V (x30)	\$12.99
5	Vibration Motors Coreless DC (x6)	\$8.99
6	ATtiny85 ESP32 Microcontroller	\$1.66
7	ESP32 S2 Mini Microcontroller	\$8.99
8	Seeed Studio Microcontroller	\$9.90
9	Raspberry Pi Zero W + Raspberry Pi Camera Module v2	\$44.95
10	Raspberry Pi 5	\$88.00
11	Raspberry Pi Camera Module v2	\$17.91
12	Cap + Housing material	\$75.00
13	Wristband fabric material	\$30.00
14	LED rings (x2)	\$32.00
	Total Cost	\$394.02

Our vendors currently include Adafruit, Amazon, and PJRC Store, but may expand in the future as we continue to iterate and upgrade key components of our Smart System. The many microcontrollers purchased were used for extensive prototyping to examine which would best fit our use case (size, power, connectivity, ease of use), and given their relatively low price range, we opted to try a variety of microcontrollers to assess the best decision for our final product. Various electronic components, wires, soldering irons and solder, and cables were previously owned/borrowed from the ECE lab.

7.0 Attachments

7.1 Appendix 1 – Engineering Requirements

Team # 29 Team Name: Smart Bears

Project Name: Smart System for Visually Impaired

Requirement	Specification	Value, range
System Performance	Response Time	< 1 second
	Object Detection Accuracy	> 90%
	Environmental Adaptation	< 10% accuracy variation between environments
Physical Properties	Total Weight	< 1 pound
	Wearable Dimensions	Cap: Standard hat sizer (21-24 inches)
	Operating Temperature	-10°C to 40°C
Power Requirements	Battery Life	≥ 12 hours per charge
	Power Consumption	< 5W continuous operation
Signal Processing	Video Frame Rate	≥ 30 fps
	Audio Latency	< 100 ms
	Haptic Signal Resolution	≥ 8 distinct patterns
Reliability	Mean Time Between Failures	> 1000 hours
	Water Resistance	IP54 rating
	Drop Resistance	1.5 m onto concrete
User Interface	Training Time	< 2 hours for basic operation
	Audio Output	40-85 dB adjustable
	Haptic Feedback Strength	0.5-2.0 G adjustable

7.2 Appendix 2 – Gantt Chart

Smart Bears Gantt Chart

PROJECT TITLE Smart System for Visually Impaired

TEAM NAME Smart Bears Team 29

WBS NUMBER	TASK TITLE	TASK OWNER	START DATE	DUE DATE	DURATION	PCT OF TASK COMPLETE	01/20-01/26							01/27-02/02							02/03-02/09							02/10-02/16							02/17-02/23																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
							M	T	W	R	F	S	U	M	T	W	R	F	S	U	M	T	W	R	F	S	U	M	T	W	R	F	S	U	M	T	W	R	F	S																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																									
1	Hardware Refinement																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																

7.3 Appendix 3 – Technical References

- Ask Samie. (n.d.). Ara - Mobility Aid for the blind and visually impaired.
<https://www.asksamie.com/products/ara?variant=43683470442704¤cy=USD>
- Harrison, C., Horstman, J., Hsieh, G., & Hudson, S. (2012). Unlocking the expressivity of point lights. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1573-1582). Association for Computing Machinery.
- Jiang, F., Yang, X., & Feng, L. (2021). A Wearable Navigation Device for Visually Impaired People Based on the Real-Time Semantic Visual SLAM System. Sensors, 21(4), 1536.
<https://doi.org/10.3390/s21041536>
- Justia Patents. (2022). Wearable Device Enablement for Visually Impaired User.
<https://patents.justia.com/patent/20220084433>
- National Federation of the Blind. (n.d.). Blindness Statistics.
<https://nfb.org/resources/blindness-statistics>
- National University of Singapore. (2023). AiSee - Mobility Aid for the blind and visually impaired. Futurity. <https://www.futurity.org/aisee-visual-impairment-ai-device-3175252/>
- OrCam Technologies. (n.d.). OrCam MyEye 3 Pro.
<https://www.orcam.com/en-us/orcam-myeye-3-pro>
- Shen, J., Chen, Y., & Sawada, H. (2022). A Wearable Assistive Device for Blind Pedestrians Using Real-Time Object Detection and Tactile Presentation. Sensors, 22(12), 4537.
<https://doi.org/10.3390/s22124537>
- Silva, B., Costelha, H., Bento, L. C., Barata, M., & Assuncao, P. (2020). User-Experience with Haptic Feedback Technologies and Text Input in Interactive Multimedia Devices. Sensors, 20(18), 5316.
- Srivastava, P., Tiwari, A., & Yadav, R. (2021). Design and Development of a Wearable Assistive Device Integrating a Fuzzy Decision Support System for Blind and Visually Impaired People. PMC, 8466919. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8466919/>
- Wearable Technologies. (2018). These New Generation of Wearables are Empowering Blind and the Visually Impaired.
<https://wt-obk.wearable-technologies.com/2018/12/these-new-generation-of-wearables-are-empowering-blind-and-the-visually-impaired/>

7.4 Appendix 4 – Drawings and Schematics

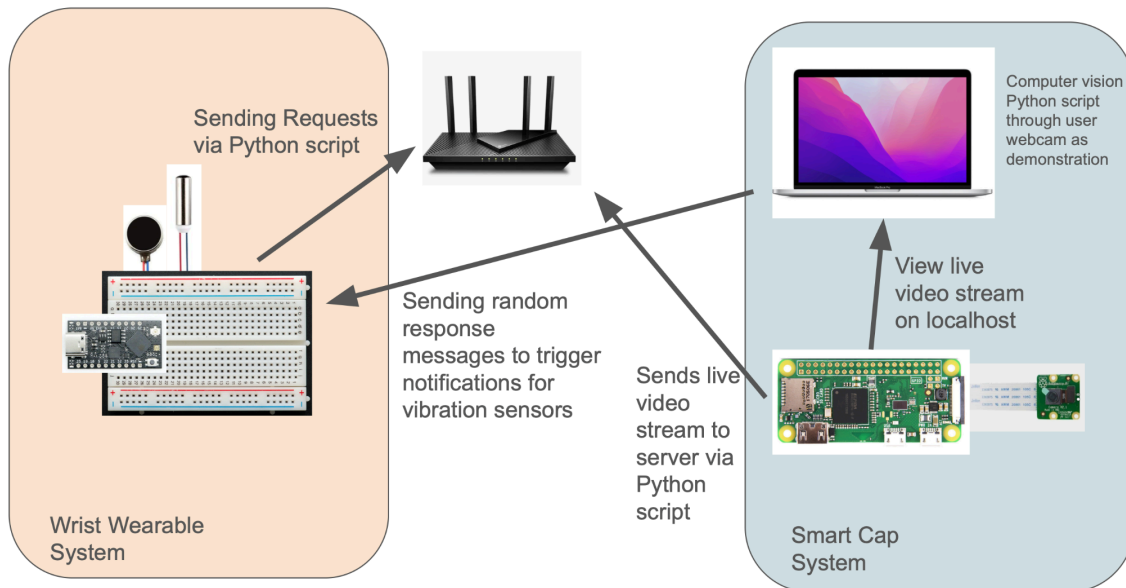


Fig 4. Initial block diagram mockup used for prototype testing.

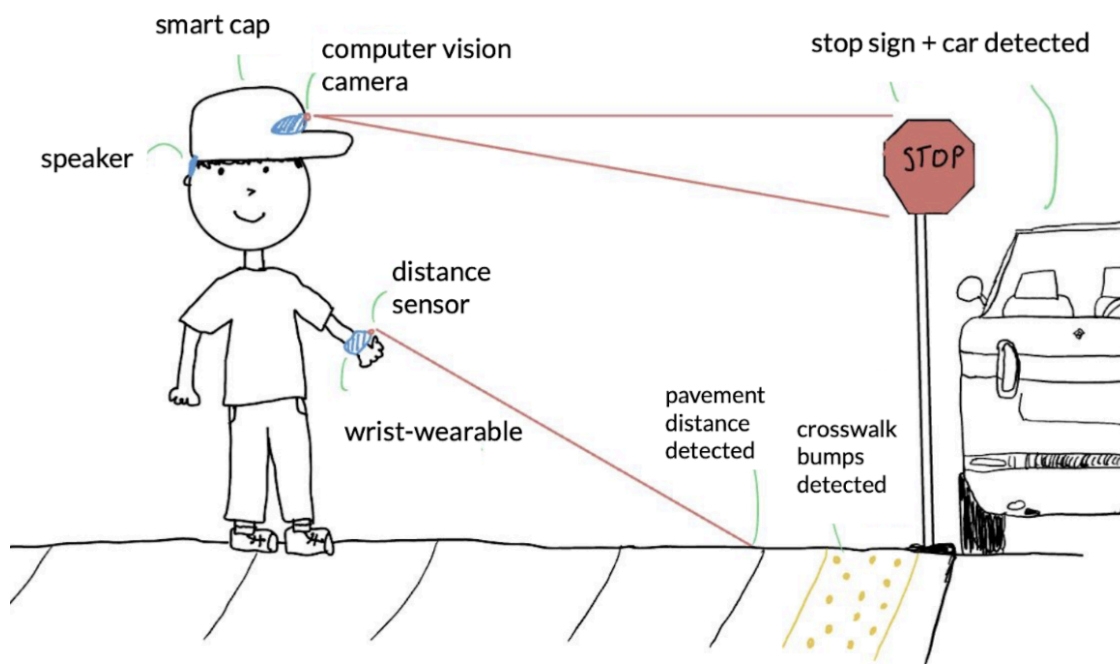


Fig 5. Initial design of Smart System for Visually Impaired showing wrist-wearable intent to replace a traditional white cane.

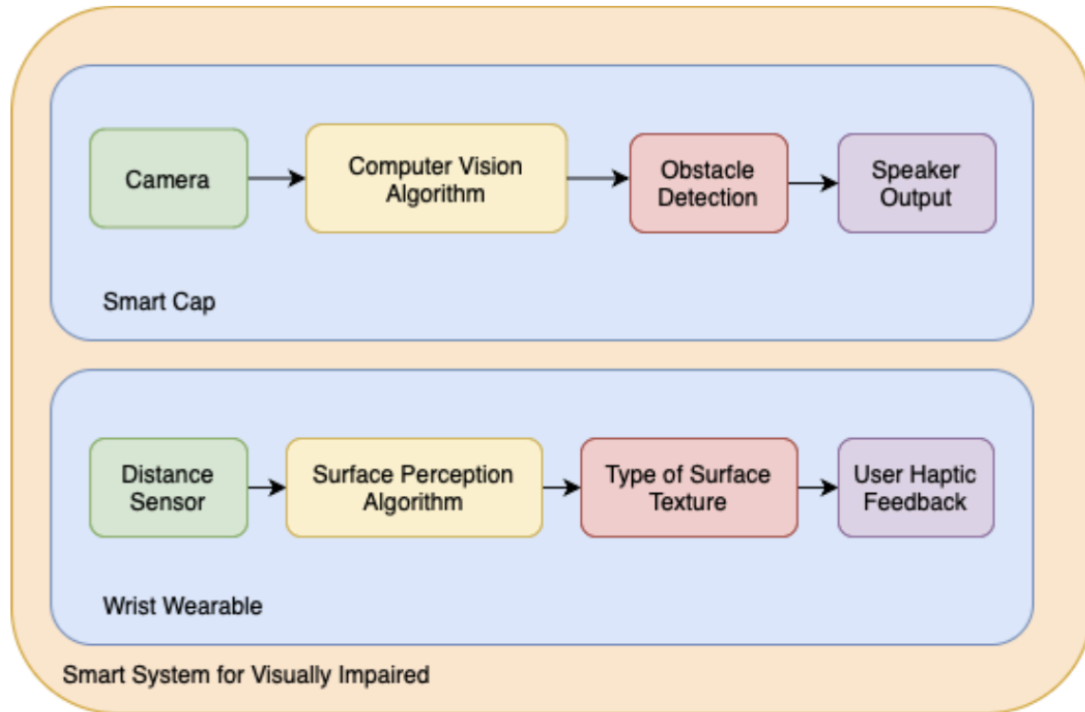


Fig 6. Initial simplified block diagram describing functionality of Smart Cap and Wrist-Wearable. Showcased at Shark Tank presentations.