University of San Agustin
General Luna St., 5000 Iloilo City, Philippines
www.usa.edu.ph

COLLEGE OF TECHNOLOGY

World Navigation Hat - Development of a Wearable Navigation Aid using AIoT for the Visually Impaired

An Undergraduate Design Project Presented to Faulty of the Computer Engineering Department of College of Technology University of San Agustine

In Partial Fulfillment of the Requirement of the Course

CPE 413 - CpE Practice and Design I

By:

D'Souza, Jason C.

Pabito, Ethel Herna C.

Wang, ChenLin

Daywan, Vince Ginno B.

Submited to:

Engr. Glenda S. Quanzon

November 10, 2025

Table of Contents

List of Figures

List of Tables

Introduction

Background of the study

Visual impairment is a global health problem that significantly affects individual's daily lives by impeding independent navigation, social interaction, and overall quality of life (Theodorou et al., 2023). There is an estimation of 2.2 billion people globally who are identified as visually impaired and this number could still increase to 2.5 billion by 2050 as stated by the World Health Organization (World Health Organization, 2023). In the Philippines, 2.17 million Filipinos are identified as visually impaired as quantized by reports from the Philippines Eye Research Institute (PERI) and Department of Health (DOH) (Shinagawa Lasik & Aesthetics, 2025). Visual impairment does not pertain to total blindness. According to World Health Organization, 2023, visual impairment can be identified and categorized based on the presenting visual acuity

(1) Mild Vision Impairment, visual acuity is better than 6/18

(2) Moderate Vision Impairment, visual acuity is worse than 6/18 but better than 6/60

(3) Severe Vision impairment, visual acuity is worse than 6/60 but better than 3/60

(4) Blindness, visual acuity is worse than 3/60

  (4.a) Blindness with light perception, individuals can only perceive light

  (4.b) Total Blindness, individuals who have no light perception

Other common types include astigmatism, near-sightedness (myopia) and far-sightedness (hyperopia), which account for a large share of impairments, alongside conditions like cataracts, glaucoma, and age-related macular degeneration (AMD), with AMD alone affecting 8.06 million people globally in 2021 . Visually impaired individuals face safety cite concerns and challenges such as accidents, falls, and collisions, as well as difficulties with Visually impaired individuals face safety concerns and challenges such as accidents, falls, and collisions, as well as difficulties with navigation, including road crossings and destination location (Gao et al., 2025; Ikram et al., 2024; Muhsin et al., 2023). Beyond these practical difficulties, social isolation and reduced access to information further compound the challenges, often limiting educational and employment opportunities (Arvind, 2023). Current existing solutions or aid for visually impaired individuals are guide dogs, white canes, and electronic travel aids (Muhsin et al., 2023).

Sensory substitution devices (SSDs) allow users to perceive their environment by converting sensory information from one modality to another, particularly aiding those with visual impairments (Mishra et al., 2025). These devices help with obstacle detection, navigation, and object recognition, enhancing independence for users relying on traditional aids (Tokmurziyev et al., 2025). However, global adoption is limited by issues such as cognitive overload, extensive training needs, ergonomic discomfort, and the lower processing bandwidth of non-visual senses compared to vision (Hou et al., 2025).

Rationale

Visual Impairment encompasses a spectrum of conditions, from moderate vision loss to complete blindness, which significantly impacts individual's ability to perceive their environment and perform daily activities (Kumar et al., 2025). This condition also affects the individual's social interaction and overall quality of life (Que et al., 2025). More than 2.17 million people in the Philippines live with visual impairment, which restricts their perception, mobility, navigation, and accessibility to information—depriving them of work opportunities and independence (Chavarria et al., 2025). To mitigate the challenges encountered by the VIP, assistive technologies and rehabilitation strategies have been developed (Skulimowski, 2025). Traditional assistive aids like white canes offer limited feedback, initiating the development of electronic travel aids that deliver richer environmental information and improve autonomous navigation (Chandra et al., 2025) (Kim, 2024). Another solution is a wearable SSD, an assistive technology that an individual can wear on their body to help users navigate both their physical surroundings and digital applications, and emerging as a potential avenue to restore or gain a sense of spatial perception or awareness and navigational ability (Skulimowski, 2025). Additionally, SSDs promotes social interaction that aligns with one of the United Nations' Sustainable Development Goals for health and equality (Xue et al., 2025). In previous studies, SSDs have been limited in their applicability, as existing devices often perform poorly in real-world situations, encountering cognitive overload problems that overwhelm users with excessive information, making navigation difficult (Casanova et al., 2025). Another identified concern is that the SSD design is uncomfortable, as it is bulky and requires extensive training, which limits its practicality (Olaosun et al., 2024). In addition, many of these devices lack digital integration as capabilities, focusing mainly on obstacle detection and not connecting well with digital tools (Makati et al., 2024). These deficiencies underscore the pressing need for higher SSDs that can outperform and enhance existing devices. Therefore, a need for development of optimized smart navigation device improves the functional efficiency and performance of SSDs without overlooking the previous concerns like overloading data processing. In this study the researchers aim to address these matters through the development of a wearable navigation hat integrated with AIoT, which introduces a smart navigation hat that uses 3D point cloud data to give audio feedback and connects to IoT services for navigation in both physical and digital environments. Furthermore, it enhances sensory processing, focusing on reducing cognitive overload in the system and making navigation simpler for VIP users. Additionally, the hat features customizable options, including a modular operating system that allows users to choose between gesture or voice commands.

General Objective

To develop an AIoT visual sensory substitution hat device that converts real-time environment information through emulating human sensory processes into audio cues for the visually impaired people (VIP)

Specific Objectives

(1) To develop the main pipeline to represent environmental information to audio cues

- (1.a) Working photogrammetry (2D to 3D) system from the cameras
- (1.b) Uses help from the IMU to create a virtual environment
- (1.c) Toggable modes to prioritize what information is cued into audio
- (1.d) Allows insertion anywhere in the pipeline for more custom preferences

(2) To develop the device to be portable and comfortable

- (2.a) Uses battery instead of needing direct connection to power
- (2.b) Can be charged through UPS instead of disposable batteries
- (2.c) All packed into a hat to avoid putting strain on sensitive areas like eyes, nose, and ears

(3) To develop an interface that is not visually dependent nor over stimulating

(3.a) Command and instructions are derived from hand guestures through hand recognition

(3.b) The commands and instructions are easy to understand and navigate

(3.c) The audio output is controlled minimizing unessary information

(4) To test if users develop unsubconscious visual senses of the environment

(4.a) They can navigate indoor through rooms/doors/furnitures

(4.b) They can detect movement and even catch trown objects

Significance of the study

The proposed AIoT navigation hat has the potential to escalate the performance of existing assistive devices by:

(1) Enhancing independence and mobility

(2) Improving cognitive comfort and ease of use

(3) Bridging physical and digital accessibility participation

(4) Promoting Social Inclusion and Employment Opportunities for the VIP

In addition, this system innovates the research for Assistive Technology in the Philippines.

(1) VIP can navigate safely and confidently with less assistance

(2) The system simplifies understanding spatial information, reducing mental strain (cognitive overload) and allowing quick interpretation without long training

(3) With AIoT integration, users can access navigation apps and online services hands-free using voice or gesture commands

(4) As VIP gain independence, they can engage more in education, work, and community activities, boosting their self-esteem and reducing reliance on others

The project's success may inspire further studies on brain-inspired computing and wearable assistive devices, paving the way for more innovative tools. This project device will specifically benefit to the following:

(1) **Visually Impaired and Blind Individuals** - They are the primary beneficiaries, gaining increased mobility, safety, and access to digital and social spaces through sensory substitution and IoT assistance.

(2) **Families and Caregivers** - With the device promoting user independence, caregivers will experience reduced physical and emotional strain while maintaining peace of mind about the user's safety.

(3) **Educators and Accessibility Advocates** - Teachers and advocates working with visually impaired students can integrate the technology into learning environments to enhance inclusivity and accessibility. Such as allowing them to access resources and infrastructure that did not take the visually impaired into account (Sites, Media, etc.)

(4) **Medical and Rehabilitation Specialists** - Eye health professionals and rehabilitation centers can use the system as a tool for sensory training and mobility rehabilitation programs, especially in fill in aspects the patent lacks or have trouble with.

(5) **Researchers and Engineers in Assistive Technology** - The project offers a novel framework that combines virtual world modeling, IoT, and sensory emulation, providing a foundation for future innovations in human-computer interaction and embedded systems.

(6) **Government and NGOs for Disability Support** - Organizations involved in disability welfare can adopt or fund similar low-cost solutions to support national accessibility programs and meet inclusive development goals.

Conceptual Framework

Figure 1: IPO Model of the Conceptual Frame of the Study

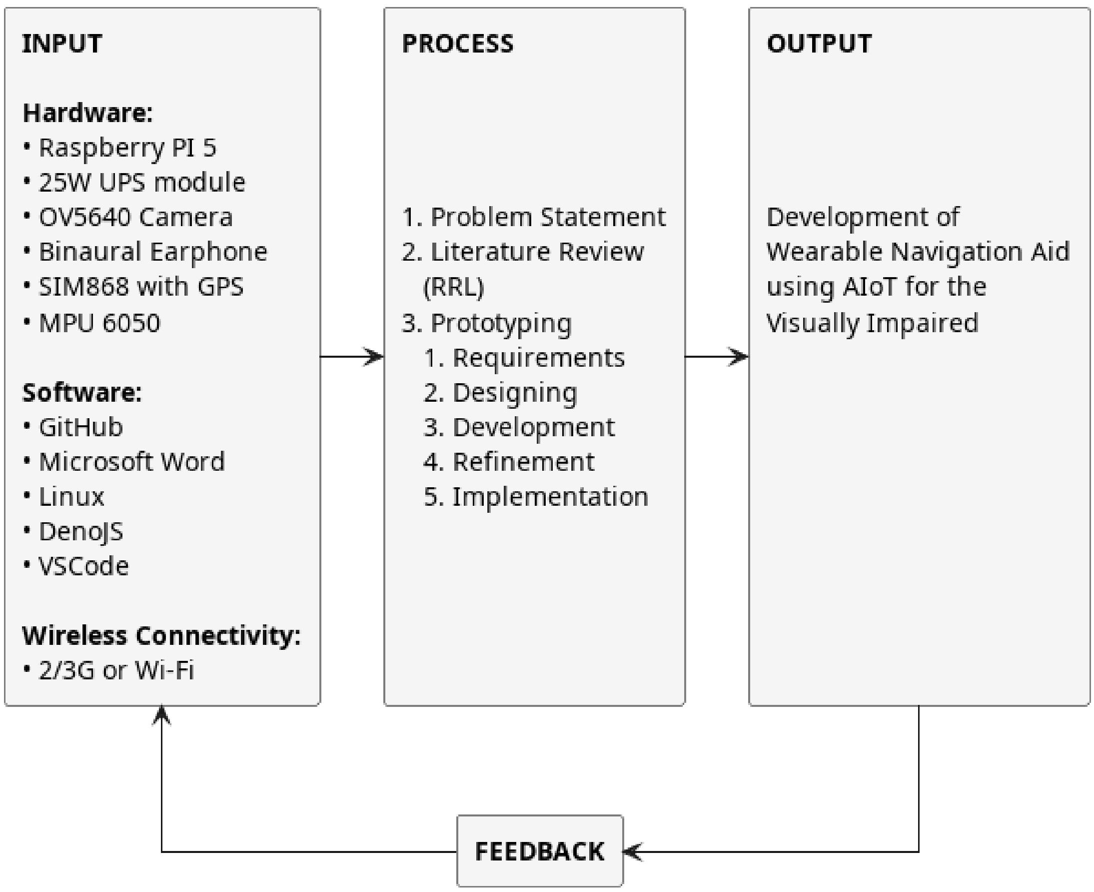

Figure states the conceptual framework of our research in a form of an Input-Process-Output (IPO) model. We start with three kinds of required input such has Hardware for our Microcontroller, Sensors, Actuators, and SIM Module, Software for our Project management, Client-side device, and Server-side device, and the Wireless Connectivity used in our device which can be either 2/3G which is used on deployment doors, or Wi-Fi which is used deployment indoors or development. Our method of research mimics typical development cycle with addition to literature review and research, such has literature review, prototyping, designing, etc. By the end of each iterative development we are to expect to be closer to the end goal, with a feedback from either advisor or mentor to start the cycle again until our end goal of an wearable navigation aid is developed.

Theoretical Framework

This research attempts to solve an issue facing most Visual Sensory Substitution Devices which is primarily about the issue on translating a high bandwidth visual information to low bandwidth audio information without causing overstimulation, our novel solution is the emulate the human sensory mechanism to carry the processing load from the user and ensure that substituted information is minimal and nessary to the user. The theoretical principles includes:

(1) **Focus and Spatial Resolution (Foveation)** - The human vision has a tiny high-resolution fovea covering around 1 to 2 degrees of the visual field but accounts for around half of visual cortex (Krantz, 2012), outside this region is a much coarser peripheral vision where acuity drops rapidly (Iwasaki & Inomata, 1986).

Our device attempts to mimic this by using a focus point system where the user can contain the focus radius indicanting that area of the environment the user wants to be translated into audio and at what detail.

(2) **Selective Attention and Daliency** - Has the visual system cannot process all details at once, it selectively attends to salient or task-relavant features, this is done by implementing a bottom-up saliency and top-down goals to filter out redundant/irrelevant information (Kristjánsson et al., 2016). This is to mean that people focus on key objects ignoring uniform backgrounds as the nervous system "tunes out" repeated stimuli and amplifies novel/focused ones (Gershman, 2024).

Has the device translates environmental information, any stagnating/constant information must be tuned out over time leaving behind changes indicating motion.

(3) **Scene Gist and Gestalt Organization** - Human vision rapidly extract the gist (background) and figures (foreground) from a scene within the first fixation (36ms) with around 80% accuracy (Loschky, 2025). This process is done through Gestalt principles that groups elements to simplify complex images such as by similarity, proximity, common region, continuity, etc. (UserTesting, 2024)

This indicates that the device should be able to generally/primitively seperate the environmental data into background and foreground categories where foreground can then be seperated into groups, this is information to indicate the translated audio.

(4) **Parallel Motion vs. Detail Pathways** - Human visual systems process motion and detials in two seperate channels, the **magnocellular pathway** for fast motion and size but in less detail, and the **parvocellular pathway** for static object in more detail (Zeki, 2015).

This suggest that when our device switches to motion detection it should prioritize the speed, size, and direction of that motion.

(5) **Temporal Dynamics and Scanning** - Vision is not a singlular static snapshot but rather continuous sample of the world via eye movements has humans typically shift gaze 3-4 times a second (Kristjánsson et al., 2016).

Thus our device should rapidly provide updating audio to allow temporal integration rather than one large static soundscape all at once, this also solves the issue of cognitive overload and sensory fautigue.

(6) **Multisensory Integration** - The brain integrates additional senses like auditory, vestibular, and proprioceptive senses to form a coherent representation of the environment (Kristjánsson et al., 2016).

Thus our device should align with those senses to prevent conflicting senses that could often cause nausea. This could be in a form of aligning the virtual environment with IMU to align with vestibular senses, or lower the output audio when the user is focusing on something else like talking to others.

Scope and Delimitation

Our research focuses on Sensory Substitution Devices, more specifically on Vision Substitution. Our device aims to fill in what different variety of visually impaired individuals lacks, like if fully blind our device should be able to act as an artificial eye providing information the person wants, or if the person is near sighted it could read text from afar from the user and warn of objects rapidly approaching the user, and so on. It does this by having a modular system running the main pipeline, the main pipeline inputs images from the stereo camera, creates a point cloud, maps it to a virtual physical map, follows an algorithm to emulate human vision processes, and finally generate the necessary audio cues for the individual. This pipeline allows applications to take into account more varied needs such has hand gesture recognition to control the parameters of the main pipeline so the person can navigate the device without seeing anything, image recognition to read text near or far, face recognition to recognize friends and family, and even more personalized interest like keeping track of expenses.

This research is conducted in the University of San Agustin and thus the prototype is initially to be tested by the researchers who only half of them are little bit near sighted, our testing sample is to soon include members of Iloilo blind associations such as the Association of Disabled Persons - Iloilo Inc. The main pipeline is prioritized over other features has it is the central component and partially answers the research question of solving existing sensory substitution issues. Has a result our prototype doesn't take into much consideration the much varied range of visual imparities, and since our study prioritize the main function over the physical design, the prototype would be heavier than what it could be if optimized for weight. However, our main objective is to build the main pipeline and the capability of support applications anywhere in the pipeline (modular operating system) showing it is possible that some of the issues facing most sensory substitution devices could be solved.

Related Review Literature

FILL THIS IN

Methodology

System Architecture

Figure 2: System Architecture

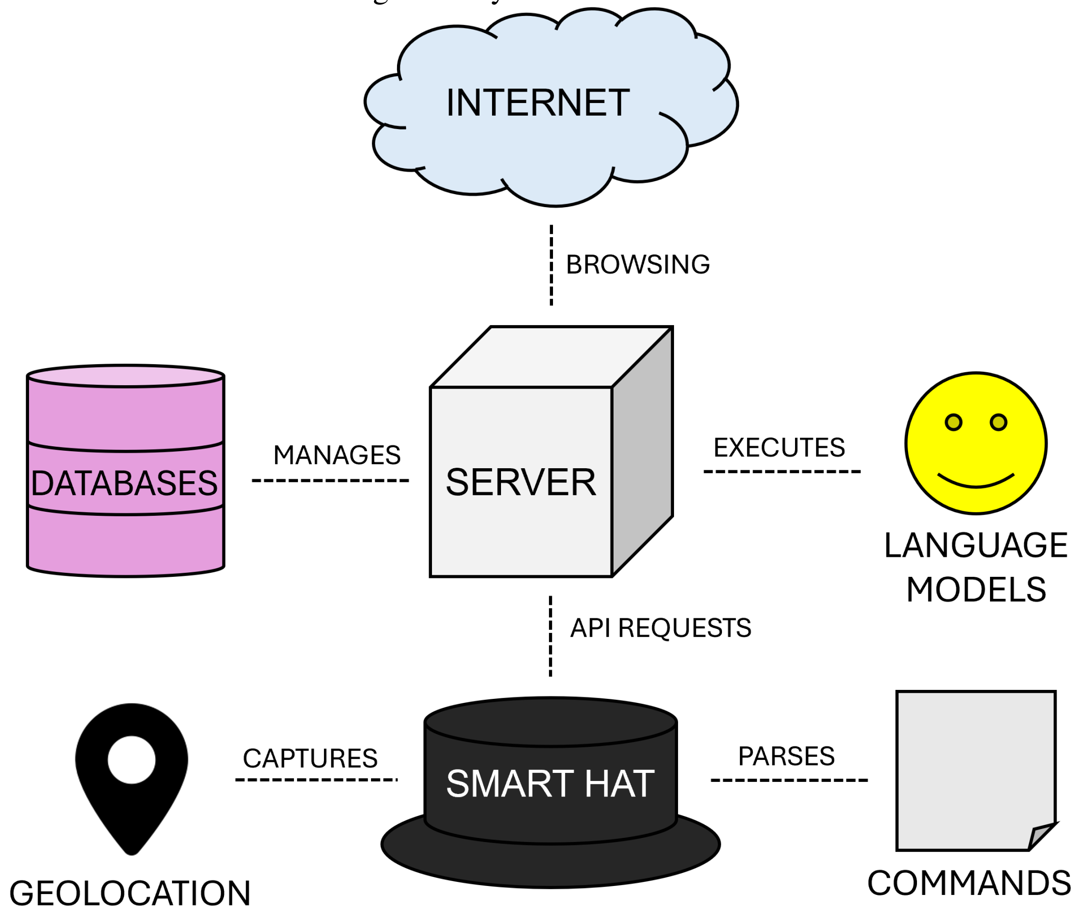

Figure 2 shows the architecture of the system, it first starts at the Smart Hat which is responsible for capturing direct sensor data like Camera, GPS, IMU, etc. and parsing/matching it to commands. Some of these commands require internet access or heavy computation which is where it makes an API request to the Server which is responsible for managing user's information like current location, credentials, etc. and also executing heavy computation like run AI/Language models. If the API request requires internet access like to find a path in google maps, or message/call people through platforms, etc.

then the server could make another API request to these providers. After the server is done computing what the user needs the user recieves the data and applies it to its virtual map or audio output.

Block diagram of the System

Figure 3: Block diagram of the System

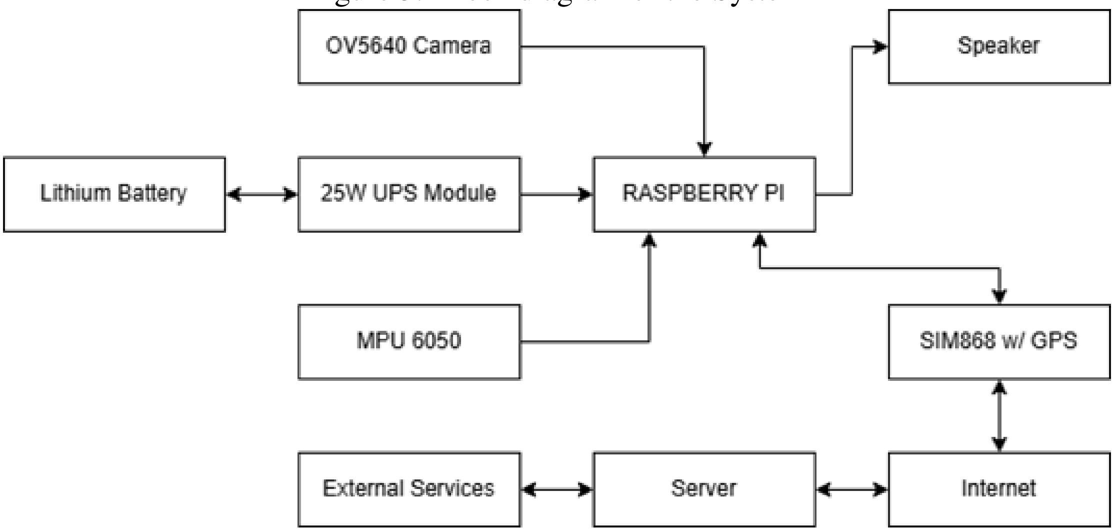

Figure explains how to build the setup by connections, it first starts with the Raspberry PI which is powered by a 25W (5V at 5A at high performance) UPS module that is connected to a lithium battery, it reads from two primary sensors the Camera for visual information, and the MPU 6050 to position the generated point cloud. After all is processed the raspberry pi and it additional applications can send the API request to the server through the SIM868, where the server can process the request through external services like AI models or external APIs like google maps, the response is then retrieved by the raspberry pi. After all is done the final generated audio cue is sent to the speakers.

Flowchart of device

Schematics of Device

Testing

References

Arvind, A. (2023, November). A deep neural architecture search net-based wearable object classification system for the visually impaired. In *Intelligent systems and pattern recognition* (pp. 198–213). Springer Nature Switzerland.

Gao, Y., Wu, D., Song, J., Zhang, X., Hou, B., Liu, H., Liao, J., & Zhou, L. (2025). A wearable obstacle avoidance device for visually impaired individuals with cross-modal learning. *Nature Communications*, *16*(1).

Gershman, S. J. (2024). Habituation as optimal filtering. *iScience*, *27*(8), 110523.

Hou, Y., Xie, Q., Zhang, N., & Lv, J. (2025). Cognitive load classification of mixed reality human computer interaction tasks based on multimodal sensor signals. *Scientific Reports*, *15*(1).

Ikram, S., Bajwa, P. I., Gyawali, S., Ikram, A., & Alsubaie, N. (2024). A iot-enabled obstacle detection and recognition technique for blind persons.

Iwasaki, M., & Inomata, H. (1986). Relation between superficial capillaries and foveal structures in the human retina. *Investigative Ophthalmology & Visual Science*, *27*(12), 1698–1705.

Krantz, J. H. (2012). Chapter 3: The stimulus and anatomy of the visual system. In *Experiencing sensation and perception*. Pearson Education.

Kristjánsson, Á., Moldoveanu, A., Jóhannesson, Ó. I., Balan, O., Spagnol, S., Valgeirsdóttir, V. V., & Unnthorsson, R. (2016). Designing sensory-substitution devices:

Principles, pitfalls and potential1. *Restorative Neurology and Neuroscience*, *34*(5), 769–787.

Loschky, L. C. (2025, October). Recognizing the gist of a scene [Accessed: 2025-11-06].

Mishra, A., Bai, Y., Narayanasamy, P., Garg, N., & Roy, N. (2025). Spatial audio processing with large language model on wearable devices.

Muhsin, Z. J., Qahwaji, R., Ghanchi, F., & Al-Taee, M. (2023). Review of substitutive assistive tools and technologies for people with visual impairments: Recent advancements and prospects. *Journal on Multimodal User Interfaces*, *18*(1), 135–156.

Shinagawa Lasik & Aesthetics. (2025, May). Perfect vision and eye health in the philippines.

Theodorou, P., Tsiligkos, K., & Meliones, A. (2023). Multi-sensor data fusion solutions for blind and visually impaired: Research and commercial navigation applications for indoor and outdoor spaces. *Sensors*, *23*(12), 5411.

Tokmurziyev, I., Cabrera, M. A., Khan, M. H., Mahmoud, Y., Moreno, L., & Tsetserukou, D. (2025). Llm-glasses: Genai-driven glasses with haptic feedback for navigation of visually impaired people.

UserTesting. (2024, April). 7 gestalt principles of visual perception: Cognitive psychology for ux.

World Health Organization. (2023, August). Blindness and visual impairment.

Zeki, S. (2015). Area v5â€”a microcosm of the visual brain. *Frontiers in Integrative Neuroscience*, *9*.

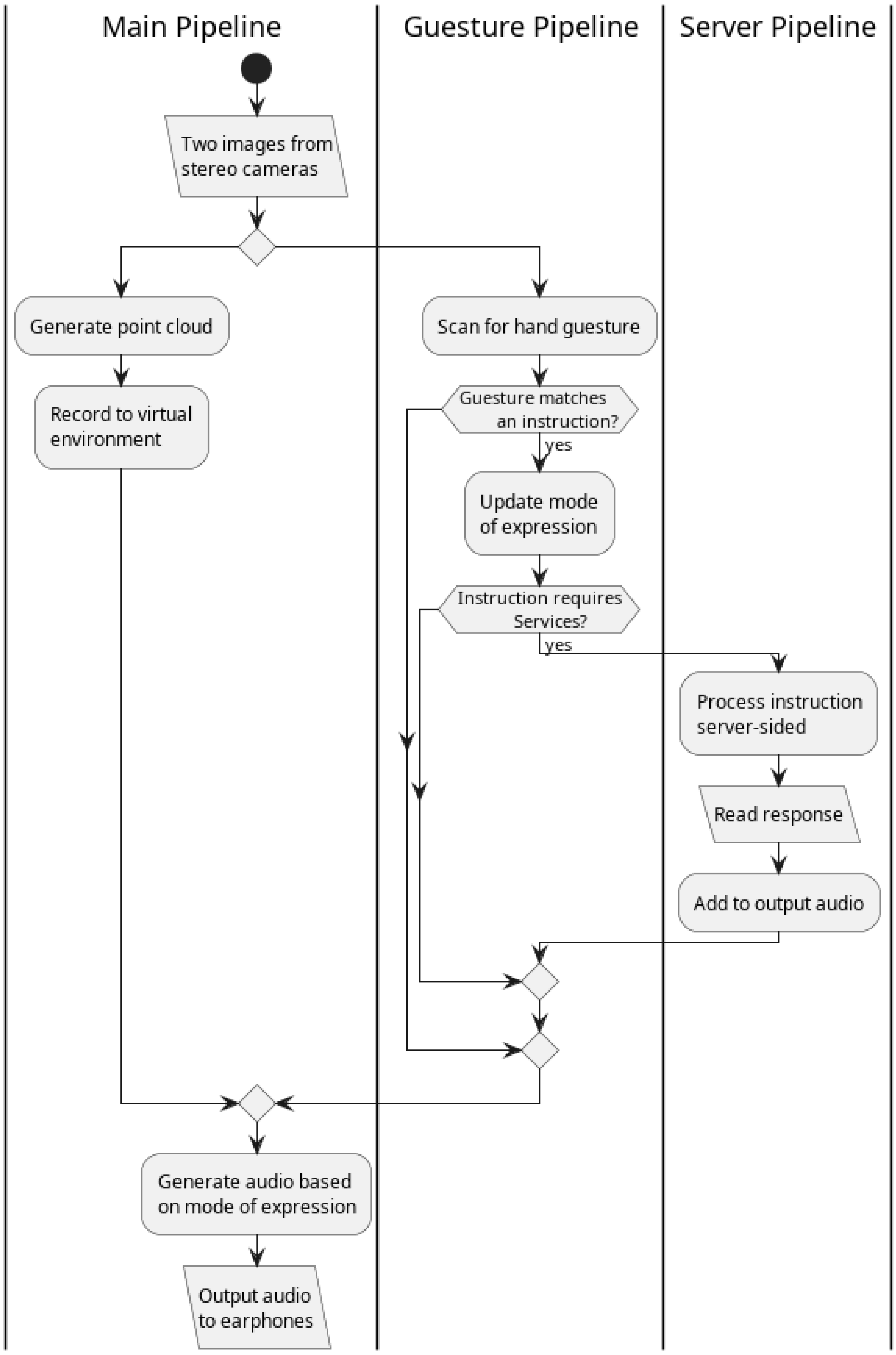

Figure 4: Flowchart of device's main pipeline

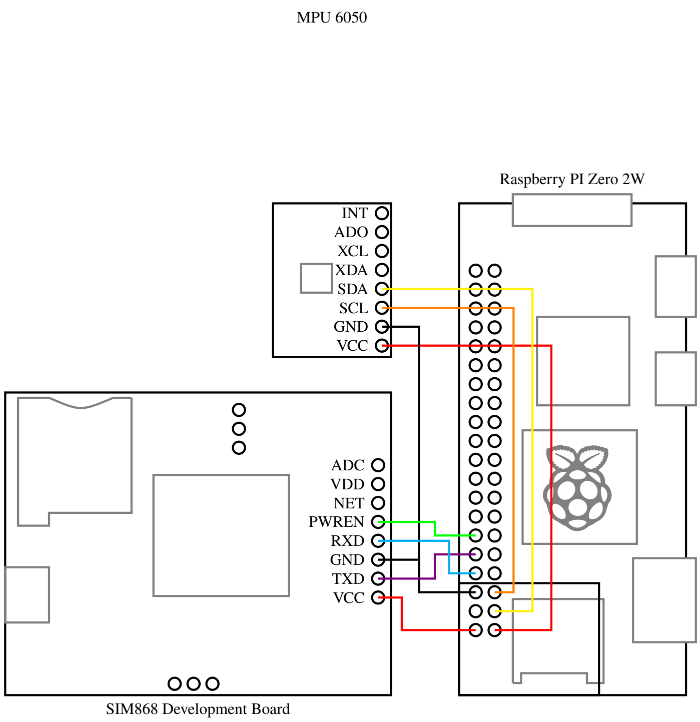

Figure 5: Skematic Diagram of Device