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:

Jason C. D'Souza

Ethel Herna Pabito

ChenLin Wang

Vince Ginno Daywan

Submited to:

Engr. Glenda S. Quanzon

November 7, 2025

Table of Contents

List of Figures

List of Tables

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).

According to Mishra et al. (2025), sensory substitution devices (SSDs) are assistive technology that converts information typically perceived through one sensory modality into another, enabling individuals with sensory impairments to access environmental cues they lack. In the context of visual impairment, such devices translate visual data, such as depth or object presence, into haptic or auditory signals to assist the visually impaired in perceiving their environment (Jiayu & Liu, 2025). This intermodal conversion allows users to develop a novel form of perception, effectively substituting the impaired sense with an intact one, enhancing the understanding of the surroundings (Jiang et al., 2025). These devices could significantly enhance independence by enabling obstacle detection, navigation, and object recognition, particularly for the fully blind who rely on traditional aids like white canes or guide dogs (Tokmurziyev et al., 2025). However, SSDs are not widely adopted globally due to challenges such as cognitive overload from processing substituted sensory data, extensive training requirements for intuitiveness, ergonomic discomfort in wearables, and processing constraints of non-visual senses like hearing or touch, which have lower bandwidth than vision (Hou et al., 2025)

Rationale

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. Wearable SSD is an assistive technology that an individual can wear on their body to supplement a missing sense, helping users navigate both their physical surroundings and digital applications. Additionally, it promotes social interaction that aligns with one of the United Nations' Sustainable Development Goals for health and equality. Supporting local innovation by developing a low-cost, locally made device increases access to technology, thereby enhancing the technology sector.

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. Another identified concern is that the device design is uncomfortable, as it is bulky and requires extensive training, which limits its practicality. In addition, many devices lack digital integration as capabilities, focusing mainly on obstacle detection and not connecting well with digital tools. These deficiencies underscore the pressing need for higher SSDs that can outperform and enhance existing devices. Therefore, there is a need for a smart navigation hat that improves the functional efficiency and performance of SSDs. This navigation hat integrates AIoT, introducing 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.

Objectives of the study

Objectives of our study

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 a working photogrammetry (2D to 3D) system that:

- (1.a) Creates a depth map based on the stereo camera
- (1.b) Transforms it into a point cloud projection
- (1.c) Matches the point cloud project to existing virtual map with help from IMU
- (1.d) Can include additional information such as color and movement

(2) To develop a system of representing environmental information though audio cues that:

- (2.a) Includes different modes of representing audio based on human sensory processes
- (2.b) Each mode prioritizes a specific aspect of the environment compromising other aspects
- (2.c) The modes can transition from one mode to another

(3) To develop the modular operating system such that:

(3.a) Features can be added as applications even if the prototype only includes the core features

(3.b) Integrate these application with the central pipeline from visual to audio information

(3.c) Add one application that recognizes hand guestures and interpret it has instructions to navigate the system

(4) To develop an AIoT server that:

(4.a) Keep tracks of users' information such has location, credentials, etc.

(4.b) Uses that information to process API requests for the users

(4.c) API methods can include accessing other APIs like Google maps, or towards AI like Language Models

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 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 de-

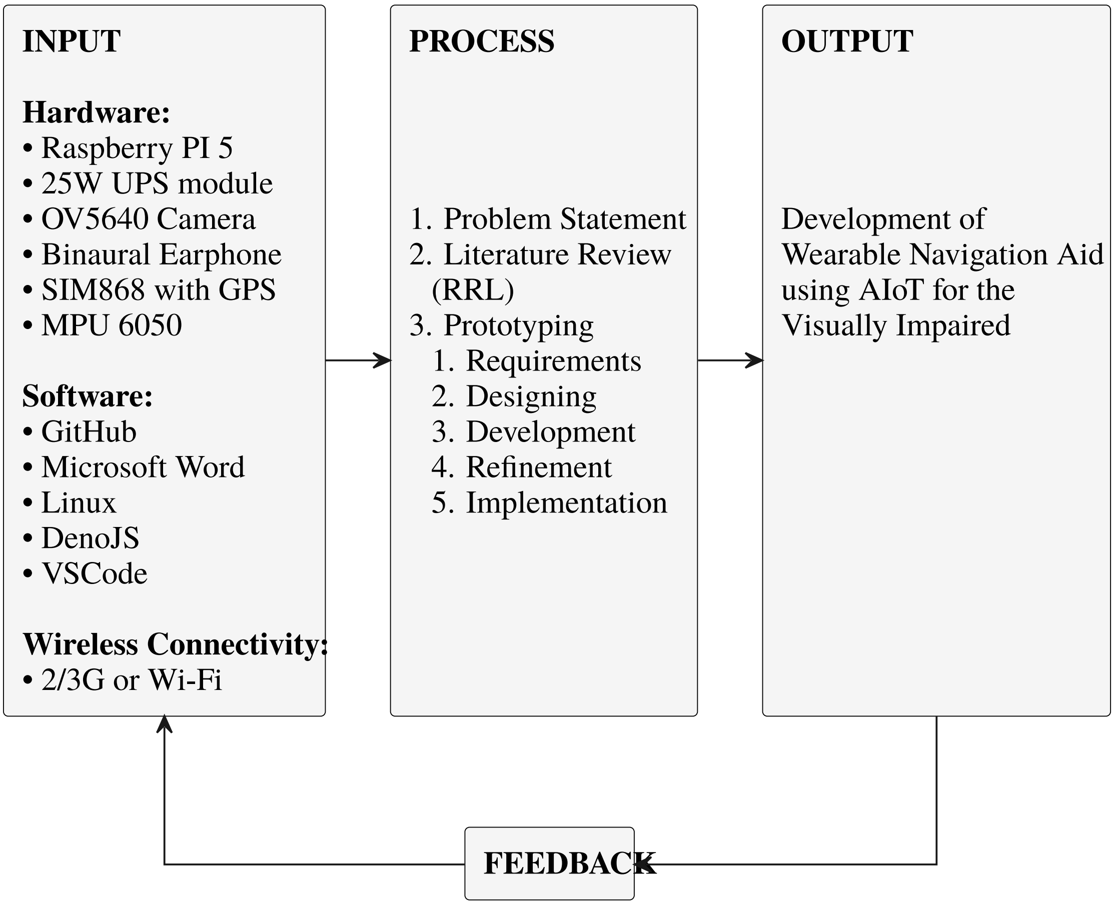

Figure 1: IPO Model of the Conceptual Frame of the Study

velopment 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**

(5) **Temporal Dynamics and Scanning**

(6) **Multisensory Integration**

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.

Methods and Materials

Research Design/Methods

The main process translates visual/environmental data to audio/haptic via stereo projection to point clouds, spectrograms, and binaural output through emulating human sensory processes. Secondary processes network users via IoT for sensor/actuator/API interactions (e.g., maps, weather). Qualitative research tests comfort, overload, and modularity through user experiments: Physical proficiency (obstacle navigation, movement detection, ball catching); Digital proficiency (communication, browsing). Timeline: Concept (Aug 2025), Prototype (Sep 2025), Testing (Oct 2025), Paper/Defense (Nov 2025). Ethical considerations: Informed consent, privacy via API security, inclusivity in design. For the main processes that emulates human sensory processes, these processes are grouped into modes that the user can gradually switch to through hand gestures but if there is no hand detected then there is no audio generated. Modes includes:

(1) **Nothing** - There is no hand so don't generate any audio

(2) **Complete** - The whole front view of the user is translated into audio cues, the depth is mapped to volume, the vertical placement is mapped to frequency, and the horizontal placement is mapped to Binaural channels. This audio map is however smoothened out through blurring to give a general and summarized environment (Triggered if a back whole hand is detected)

(3) **Focused** - The visual close to the center of the user is translated to audio cues and gradually fades off has it goes away from the origin point, here color is included has the audio texture (Triggered if a back closed hand is detected)

(4) **Movement** - Any movement across the user is transformed into audio cues (Triggered if a back closed hand with thumbs up is detected)

(5) **Text** - Reads the text infront of the user (Triggered if back closed hand with index finger up)

(6) **Face** - Gives a unique audio cue representing the user's face structure (Triggered if back closed hand with thumbs up and index finger is up)

Hardware Material Specification

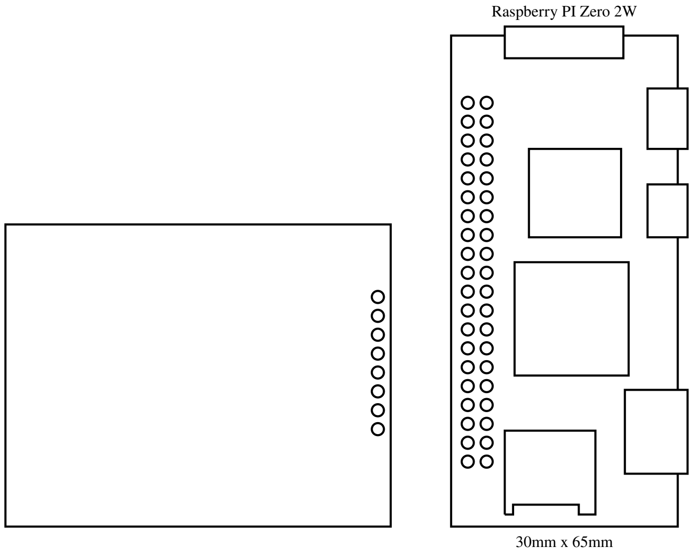

Figure 2: Skematic Diagram of Device

Quantity	Material	Description
1	Raspberry PI Zero 2W	Main microcontroller
1	USB-C 5A/3A UPS Module	Power supply
2	OV5640 USB Camera 5MP and 160°FOV	Image input
1	Wired binaural earphone w/ microphone	Audio input/output
1	SIM868 development board	Phone features (Internet, GPS, etc.)
1	MPU 6050	Orientation detection
1	Female audio jack module	Connects earphone with Raspberry pi
1	Micro SIM card	Connects to provider for internet
1	Micro SD card	Stores operating system and data
2	18650 Li-ion Rechargeable Battery	Power storage

Table 1: Hardware Specification

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.

Jiang, P., Rossiter, J., & Kent, C. (2025). Auditory and tactile frequency mapping for visual distance perception: A step forward in sensory substitution and augmentation (C. Papadelis, Ed.). *PLOS ONE*, *20*(3), e0318354.

Jiayu & Liu. (2025). Developing an ai-guided assistant device for the deaf and hearing impaired.

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.