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

COLLEGE OF TECHNOLOGY

Virtual World Navigation Hat – Development of Visual Sensory Substitution Device through Point Cloud Projection and IoT services for the Visually Impaired

An Undergraduate Design Project Presented to Faculty of the
Computer Engineering Department College of Technology
University of San Agustin

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

Engr. Glenda S. Guanzon

September 2025

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Background of the Study

Visual impairment affects a significant portion of the global population, with varying degrees of severity that impact daily life, mobility, and social integration. According to the World Health Organization (WHO), approximately 2.2 billion people worldwide experience near or distance vision impairment, of which at least 1 billion cases could have been prevented or remain unaddressed (W.H.O., 2023). Among these, 43.3 million individuals are fully blind (visual acuity worse than 3/60 in the better eye), while 295 million suffer from moderate to severe visual impairment (visual acuity between 3/60 and 6/18) (Lijuan Que, 2025). Common types include refractive errors such as 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 (Yon, 2025). In the Philippines, an estimated 2.17 million people live with visual impairment, including approximately 592,000 who are fully blind and over 2.1 million with low vision or moderate impairments (Cubillan, 2025). Refractive errors and cataracts are predominant, with blindness rates at about 0.89% and moderate to severe visual impairment (MSVI) at 4.71% (Norton, 2024). In Iloilo, part of Western Visayas, regional surveys indicate a blindness prevalence of around 2.6-3.0% in adults aged 50 and older, primarily due to avoidable causes like cataracts, aligning with national trends but with limited city-specific data (Cristina Eusebio, 2007).

Visual sensory substitution devices (SSDs) convert visual information into alternative sensory modalities, such as auditory or haptic feedback, to assist the visually impaired in perceiving their environment (Otilia Zvorișteanu, 2021). 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. 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 (Hoffmann, Spagnol, Kristjánsson, & Unnthorsson, 2018). Our study attempts to address these through three major solutions: creating virtual worlds for tracking physical and digital environments, emulating human visual processing techniques to reduce overload, and implementing a modular IoT-connected system for customization. The final objective is to advance SSDs and test their potential for societal-wide implementation, overcoming barriers to real-world use.

Rationale

Existing sensory substitution devices, particularly visual SSDs, face significant knowledge gaps and issues that limit their practical application. Key challenges include limited generalizability beyond controlled lab settings, where devices perform well but fail in dynamic real-world environments due to variability in lighting, noise, or user movement (Farina, 2013). User comfort, ergonomics, and aesthetics are also problematic, as many SSDs are bulky, interfere with mobility, or lack appeal, leading to low adoption (Paul Bach-y-Rita, 2003) (Jamie Ward, 2010). Cognitive overload arises from overwhelming the user's finite attentional capacity with constant sensory input, potentially reducing situational awareness ((Macpherson, 2018). Training demands are high, requiring weeks or months for proficiency, which deters users (Maidenbaum, 2015). Finally, sensory processing constraints mean auditory or haptic channels cannot fully match vision's detail and speed (Paul Bach-y-Rita, 2003).

Solving these issues would greatly improve the lives of visually impaired individuals, especially the fully blind, by enabling seamless navigation in physical worlds (e.g., avoiding obstacles, detecting changes) and digital worlds (e.g., accessing social media, maps). In the Philippines, where over 2 million face visual impairment and infrastructure support is inconsistent, this could foster independence and societal integration (Jr., 2024).

Our solutions differentiate from existing research: First, virtual worlds track physical (via point clouds) and digital (via IoT APIs) environments, allowing holistic navigation unlike device-specific approaches (e.g., AuralVision's depth-to-sound conversion) (AuralVision, n.d.). Second, we emulate human visual techniques—logarithmic scaling for change detection, focus prioritization (proportional to depth, inversely to peripheral changes), and equal-loudness contour adjustments—to minimize overload, contrasting with non-adaptive systems like Eyesynth's NIIRA (Eyesynth NIRRA, 2021). Third, modularity via a Raspberry Pi OS with apps for gestures, voice commands, and IoT integration customizes for impairments, extending beyond fixed-function devices like the Unfolding Space Glove (Jakob Kilian, 2022). Using a hat form factor temporarily addresses bulkiness until miniaturization

Objectives of the Study

General Objectives

To build a visual sensory substitution device through stereo photography into point cloud projection and into a spectrogram representation and finally audio.

Specific Objectives

- The device is able to use the point cloud to accurately construct a virtual environment which could effectively convey necessary information to the user by emulating human-like virtual shortcuts to minimize sensory overload while mapping the area to keep track of items and destinations.
- The device allows easy navigation in a modular operating system such that even the blind could navigate through non-visual actions like hand gestures or voice commands, and manage other applications.

Significance of the Study

Visually impaired individuals, especially the blind, face profound challenges in navigating physical and digital worlds. In the physical realm, they encounter obstacles like uneven terrain, traffic, or indoor hazards, often relying on limited aids that provide incomplete feedback. Digitally, barriers include inaccessible apps, websites, and devices, hindering communication and information access. In the Philippines, with 2.17 million affected and inconsistent infrastructure (e.g., limited braille signage in public spaces), social isolation and dependency are exacerbated (Cubillan, 2025).

Existing solutions include mobility aids like white canes, guide dogs, and electronic travel aids (e.g., ultrasonic canes), and digital tools such as screen readers (e.g., JAWS), voice assistants (e.g., Siri), and WCAG-compliant apps. However, gaps persist: Public implementation is uneven, with many environments lacking blind support; our device translates visual information (e.g., signs, objects) to audio, reducing dependency on external accommodations. Cost is a barrier, as high-end aids exceed PHP 50,000; our Raspberry Pi-based hat aims to be cheaper than smartphones (under PHP 10,000), serving as a "blind person's phone." Inconsistent WCAG/ADA implementations limit access; our modular apps use automation/APIs (e.g., for Messenger, Google Maps) to interact with non-compliant sources, enhancing independence.

Conceptual Framework

The conceptual framework outlines the device's input-process-output model, incorporating variables aligned with human sensory processing. Inputs include stereo images from two OV5640 USB cameras (5MP, 160° FOV, 10cm apart, 30fps), yielding depth (disparity map) and color data via stereo computations. User orientation from the IMU (accelerometer + gyroscope) corrects for head movements. The point cloud output is scaled and mapped to a virtual digital environment using error correction from prior frames or IMU data.

Necessary information is processed: Voice data (e.g., commands, messages) is directly output to speakers; depth information becomes a spectrogram (frequency for Y-axis, binaural audio for X-axis, volume for Z-axis/depth). The focus point is centered in the user's forward view, with focus radius inversely proportional to focus change speed (biased on X-axis for horizontal priority; shape as a tall oval to emphasize vertical threats like steps or overhead objects). Variables emulate human theories: Logarithmic scaling filters constant movements over time; focus scoping prioritizes central depth over peripherals, reducing data volume. Outputs are audio/haptic cues via earphones, enabling navigation in physical (obstacle avoidance) and digital (app interactions) worlds.

Figure 1: IPO Model of the Controlled System

Theoretical Framework

This research is grounded in theories of human cognitive senses, such as logarithmic scaling (Weber-Fechner law), where sensory perception is relative to stimulus change, allowing automatic filtering of static elements to prevent overload (Fechner, 1960). Focus scoping draws from attentional theories, prioritizing foveal (central) over peripheral vision to manage information (Orienting of Attention, 1980). We theorize that emulating these processes in SSDs can resolve sensory overload and fatigue by delivering intuitive, context-aware feedback, unlike raw data in traditional devices.

Additionally, modularity is based on user-centered design theory, positing that customizable systems better accommodate diverse impairments and preferences (Norman, 2013). By integrating IoT for generalized situations (e.g., API-driven apps), the framework hypothesizes improved adoption, solving lab-to-real-world gaps through adaptive, brain-like processing.

Figure 2: IOT System Architecture of the whole system

Figure 3: Basic Architecture of and Embedded System (whole System)

Figure 4: Flowchart of the AI in the System

Figure 5: Process flow of the whole System

Scope and Delimitations

Our device should work indoors and outdoors to avoid inconvenience, but to prevent scope creep, we focus primarily on sensory substitution via virtual worlds for visual-to-audio translation, ensuring interpretability. Secondarily, we emphasize modularity for user-specific processes and tracking physical/digital environments. Delimitations include potential power and weight issues as a wearable prototype, excluding non-essential features like advanced AI beyond basic recognition.

Related Studies

Several studies inform our work:

- **AuralVision** (AuralVision, n.d.): Converts depth via ToF lasers to stereo sound; provides depth encoding insights but lacks modularity.
- **Eyesynth's NIIRA** (Eyesynth NIRRA, 2021): Uses RealSense D415 for spatial audio; highlights semantic integration, aiding our point cloud approach.
- **Unfolding Space Glove** (Jakob Kilian, 2022): Haptic feedback for hand-object distance; informs ergonomics but is hand-specific.
- **Virtual Whiskers** (Junchi Feng, 2024): AI-driven vibration for paths; supports obstacle density mapping in our virtual world.
- **Depth Sonification** (Louis Commère, 2023): LiDAR-to-repetition rate; guides our spectrogram design for depth.

- **SoundSight App** (Giles Hamilton-Fletcher, 2021): Mobile color/depth to audio; inspires multi-feature translation.
- **Smart Hat** (Memoona Sami, 2024): Object detection to voice; aids IoT integration.
- **Clearway Companion** (Akshay Vijay Panchal, 2024): Haptic/audio with motion; informs user feedback.
- **Multi-faceted SSD** (Ligao Ruan, 2025): Audio-based curb detection; enhances real-time alerts in our system.
- **Multi-path Navigation** (Zaidao Han, 2025): Low-vision mobility; supports virtual pathfinding.
- **BrainPort Vision Pro** (Yiming Zhao, 2025): Functional evaluation; guides our testing protocols.
- **Smart Vision Glasses** (Devi Udayakumar, 2025): AI-powered glasses; informs gesture/voice interfaces.

These provide foundations for depth processing, modularity, and user testing, filling gaps in our emulation and IoT focus.

Review Literature

Reviewed studies emphasize processes, outcomes, and materials relevant to ours. AuralVision uses ToF lasers for depth-to-stereo sound via bone conduction, achieving lab-based navigation but limited by sensor range (AuralVision, n.d.). Eyesynth's NIIRA integrates RealSense cameras for semantic audio, outperforming in object recognition but causing overload without human emulation (Eyesynth NIRRA, 2021). Unfolding Space Glove employs haptic motors for distance feedback, effective for grasping but non-wearable for head-based navigation (Jakob Kilian, 2022). Virtual Whiskers uses cameras/AI for vibration belts, reducing collisions by 30% in tests but lacking digital integration (Junchi Feng, 2024). Depth Sonification recommends repetition rates for depth, improving intuitiveness over raw tones (Louis Commère, 2023). SoundSight translates LiDAR/thermal to timbre/volume, enabling color/depth perception on mobiles but app-limited (Giles Hamilton-Fletcher, 2021). Smart Hat detects objects via camera for voice output, cost-effective but non-modular (Memoona Sami, 2024). Clearway uses PIR sensors for haptic alerts, user-friendly but sensor-constrained (Akshay Vijay Panchal, 2024). Newer studies like multi-faceted SSDs show 85% accuracy in curb detection, aligning with our audio focus (Ligao Ruan, 2025). Our work builds on these by adding brain emulation and IoT for superior overload reduction and versatility

METHODS AND MATERIALS

Methods

Research Design/Methods

The main process translates visual/environmental data to audio/haptic via stereo projection to point clouds, spectrograms, and binaural output. 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.

Materials

Materials and Processes

Inputs Visual via cameras, audio/commands via microphone. Processes Convert to 3D environment/point clouds, emulate human shortcuts (logarithmic changes, focus prioritization), apply equal-loudness contours, Fourier transform to audio. Outputs Selected audio slices. Data flow is from Stereo images → Depth map (Q matrix) → Point cloud (X,Y,Z) → Scaled virtual map → Prioritized info → Spectrogram/audio.

Hardware Material Specification

The main controller is the Raspberry Pi 5; it is powered by a PiSugar 2 which is a UPS Power supply. It inputs image data from two OV5640 USB Camera of 5MP and 160 degrees of FOV, audio data from microphone in the earphones, orientation data from IMU, and internet access and GPS data from the SIM808 module. It then outputs everything to the speakers in the earphones with binaural audio, better to be replaced with bone-conducting features.

Software Specifications

To format the SD card, a bash script is executed through Ubuntu a Debian-based linux operating system to install Raspberry PI into the SD card, the main program for the device uses C++ has the main program running OpenCV to process the images and virtual worlds. In the server the device is connected to, it runs DenoJS and Puppeteer to keep track of IoT data and run automation scripts connecting to other services like Meta, or Ollama of AI

Library and Board Managers

REFERENCES

(n.d.). Retrieved from AuralVision: https://auralvision.org/

(1980). In M. I. Posner, *Orienting of Attention* (pp. 32(1), 3-25). Quarterly Journal of Experimental Psychology.

Akshay Vijay Panchal, M. C. (2024, April 27). *Clearway Companion - an Ai Powered Aid for Visually Impaired*. Retrieved from SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4809356

Cristina Eusebio, H. K. (2007, December). *Rapid assessment of avoidable blindness in Negros Island and Antique District, Philippines*. Retrieved from ResearchGate: https://www.researchgate.net/publication/6268823_Rapid_assessment_of_avoidable_blindness_in_Negros_Island_and_Antique_District_Philippines

Cruz, B. D. (2023, March 13). *Population of visually impaired in Iloilo City*. Retrieved from Freedom of Information Philippines: https://www.foi.gov.ph/requests/population-of-visually-impaired-in-iloilo-city/

Cubillan. (2025, May 27). *Perfect Vision and Eye Health in the Philippines*. Retrieved from Shinagawa Lasik & Aesthetics: https://shinagawa.ph/perfect-vision-and-eye-health-in-the-philippines/

Devi Udayakumar, S. G. (2025, May 30). *Artificial intelligence-powered smart vision glasses for the visually impaired*. Retrieved from National Center for Biotechnology Information: https://pmc.ncbi.nlm.nih.gov/articles/PMC12178407/

Eyesynth NIRRA. (2021). EE Journal. Retrieved from RealSense: https://realsenseai.com/case-studies/powering-sonic-vision-for-the-blind/

Farina, M. (2013, April 4). Neither touch nor vision: sensory substitution as artificial synaesthesia? In M. Farina, *Biology & Philosophy* (pp. 32(4), 639-655). Pringer Nature. Retrieved from PhilArchive: https://philarchive.org/rec/FARNTN

Fechner, G. T. (1960). *Elemente der Psychophysik.* Breitkopf und Härtel.

Fleck, A. (2024, January 4). *Vision Loss Predicted to Surge 55% by 2050*. Retrieved from Statista: https://www.statista.com/chart/31502/expected-number-of-people-with-vision-loss-globally/

Giles Hamilton-Fletcher, J. A. (2021, July 2). *SoundSight: a mobile sensory substitution device that sonifies colour, distance, and temperature*. Retrieved from Home | SpringerLink: https://link.springer.com/article/10.1007/s12193-021-00376-w

Hoffmann, R. M., Spagnol, S. P., Kristjánsson, Á. P., & Unnthorsson, R. P. (2018, September). *Evaluation of an Audio-haptic Sensory Substitution Device for Enhancing Spatial Awareness for the Visually Impaired*. Retrieved from Optometry and Vision Science:

https://journals.lww.com/optvissci/fulltext/2018/09000/evaluation_of_an_audio_haptic_sensory_substitution.10.aspx

Jakob Kilian, A. N. (2022, Febuary 26). *The Unfolding Space Glove: A Wearable Spatio-Visual to Haptic Sensory Substitution Device for Blind People*. Retrieved from PubMed: https://pubmed.ncbi.nlm.nih.gov/35271009/

Jamie Ward, P. M. (2010). Visual experiences in the blind induced by an auditory sensory substitution device. In P. M. Jamie Ward, *Consciousness and Cognition* (p. March). ELSEVIER.

Jr., M. J. (2024, April 8). *Country report from the Philippines to the mid-term regional general assembly, Thailand, 27-29 November 2023*. Retrieved from World Blind Union - Asia Pacific: https://wbuap.org/archives/1587

Junchi Feng, G. H.-F.-R. (2024, August 26). *Haptics-based, higher-order Sensory Substitution designed for Object Negotiation in Blindness and Low Vision: Virtual Whiskers*. Retrieved from arXiv: https://arxiv.org/abs/2408.14550

Ligao Ruan, G. H.-F.-R. (2025, Febuary 1). *Multi-faceted sensory substitution using wearable technology for curb alerting: a pilot investigation with persons with blindness and low vision*. Retrieved from Taylor & Francis Online: https://www.tandfonline.com/doi/full/10.1080/17483107.2025.2463541

Lijuan Que, Z. Q. (2025, April 29). *An analysis of the global, regional, and national burden of blindness and vision loss between 1990 and 2021: the findings of the Global Burden of Disease Study 2021*. Retrieved from Frontiers | Publisher of peer-reviewed articles in open access journals: https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2025.1560449/full

Louis Commère, J. R. (2023, April 11). *Evaluation of short range depth sonifications for visual-to-auditory sensory substitution*. Retrieved from arXiv: https://arxiv.org/abs/2304.05462

Macpherson, F. (2018). Sensory substitution: A new perceptual experience? In F. Macpherson, *Sensory Substitution and Augmentation* (pp. 162-174). Oxford University Press.

Maidenbaum, S. (2015, October 26). *Sensory Substitution Training for Users Who Are Blind with Dynamic Stimuli, Games and Virtual Environments*. Retrieved from ACM Digital Library: https://dl.acm.org/doi/10.1145/2700648.2811324

Memoona Sami, D. A. (2024, December). *Efficient Object Detection and Voice-Assisted Navigation for the Visually Impaired: The Smart Hat Approach*. Retrieved from Sukkur IBA Journals: https://journal.iba-suk.edu.pk:8089/SIBAJournals/index.php/sjcms/article/view/1535/469

Norman, D. (2013). *The Design of Everyday Things.* New York: Basic Books.

Norton, H. (2024, September 12). *The most and least visually impaired countries worldwide*. Retrieved from Compare the market: https://www.comparethemarket.com.au/health-insurance/features/visually-impaired-countries-2024/

Otilia Zvorișteanu, S. C.-G. (2021, July 6). *Sensory Substitution for the Visually Impaired: A Study on the Usability of the Sound of Vision System in Outdoor Environments*. Retrieved from MDPI - Publisher of Open Access Journals: https://www.mdpi.com/2079-9292/10/14/1619

Paul Bach-y-Rita, S. W. (2003). Sensory substitution and the human–machine interface. In S. W. Paul Bach-y-Rita, *Trends in Cognitive Sciences* (pp. 7(12), 541-546). Cell Press.

Shachar Maidenbaum, S. H.-R.-T. (2014, December 1). *The "EyeCane", a new electronic travel aid for the blind: Technology, behavior & swift learning*. Retrieved from Sage Journals: https://journals.sagepub.com/doi/abs/10.3233/RNN-130351

W.H.O. (2023, August 10). *Blindness and vision impairment*. Retrieved from World Health Organization: https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment

Yiming Zhao, S. Y. (2025, April 29). *Evaluation of the Efficacy of an Assistive Device for Blind People: A Prospective, Non-Randomized, Single Arm, and Open Label Clinical Trial*. Retrieved from Taylor & Francis Online: https://www.tandfonline.com/doi/full/10.1080/02713683.2025.2495212

Yon, D. K. (2025, June 25). *Global burden of vision impairment due to age-related macular degeneration, 1990–2021, with forecasts to 2050: a systematic analysis for the Global Burden of Disease Study 2021*. Retrieved from PubMed Central: https://pmc.ncbi.nlm.nih.gov/articles/PMC12208786/

Zaidao Han, S. L. (2025, August 27). *Multi-path sensory substitution device navigates the blind and visually impaired individuals*. Retrieved from ScienceDirect: https://www.sciencedirect.com/science/article/pii/S0141938225002379