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

D'Souza, Jason C.
University of San Agustin
Balabag, Pavia, Iloilo, Philippines
jasouzax@gmail.com
ORCID: 0009-0003-6062-7921

Wang, ChenLin
University of San Agustin
San Pedro, Molo, Iloilo, Philippines
cwang@usa.edu.ph

Pabito, Ethel Herna C.
University of San Agustin
Aurora Subd., Iloilo City, Philippines
ehpabito@gmail.com

Daywan, Vince Ginno B.
University of San Agustin
Sta. Justa, Tibiao, Antique, Philippines
vgbalitao-saan@usa.edu.ph

Guanzon, Glenda
University of San Agustin
Sta. Justa, Tibiao, Antique, Philippines
vgbalitao-saan@usa.edu.ph

***Abstract*—Our primary objective is aligned with Sustainable Development Goal (SDG) 10 on reducing inequalities, specifically to mitigate disparities faced by visually impaired individuals (VIP) by addressing a key challenge in sensory substitution devices (SSDs): sensory limitations. We hypothesize that enabling user-controlled selection of substituted sensory information via intuitive inputs like hand gestures could facilitate efficient subconscious pattern-building in the brain, thereby minimizing overload, fatigue, and integration issues.**

To test this, we developed a wearable prototype using a Raspberry Pi 5 equipped with a UPS, IMU sensor, camera, and earphones, designed as a head-mounted "hat" to distribute weight away from sensitive facial areas. The system generates depth maps using MiDaS AI for environmental perception, recognizes hand gestures via MediaPipe AI for user control, and employs a novel algorithm to produce targeted audio channels representing selective visual data by emulating human sensory processes in psychology, plus a primitive connection to a server through NodeJS for future modular features.

Validation involved empirical assessments of sensory limitations through metrics including sensory overload (e.g., interpretation accuracy), fatigue (e.g., sustainable usage duration), integration (e.g., reaction times and learning curves), and overall device performance. Extensive optimizations were required, such as model downscaling, information chunking into digestible segments, and consideration of alternatives like stereo vision to enhance real-time efficiency. Results indicate that users favored brief, intermittent audio outputs over prolonged sessions, demonstrated improved environmental perception, and experienced reduced sensory limitations. In conclusion, our approach partially resolves SSD challenges, validating the potential of controllable substitution, though further refinements and longitudinal studies are essential for broader applicability and impact on SDG 10 targets.

***Index Terms*—sensory substitution device, SDG 10, visual impairment, controllability, accessibility, assitive robotics, industrial automation**

I. Introduction

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 [1]. 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 [2]. This thesis focuses on SDG 10, which aims to reduce inequalities within and among countries based on income, sex, age, disability, sexual orientation, race, class, ethnicity, religion, and oppotunity [3]. More specifically our thesis focuses on visual impared people or VIPs under disability, VIPs does not pertain to total blindness and can be identified and categorized based on presenting visual acuity [2].

Navigation is one of the problems VIPs encounter; they often have difficulty crossing the road and tracking their location, as well as experiencing accidents, falls, and collisions [4]. In addition, VIPs experience social isolation and reduced access to information, limiting their educational and employment opportunities [5]. There are already existing solutions or aids for VIPs, for daily assistance, which include guide dogs, white canes, and electronic travel aids [4]. Additionally, one of the technologies VIPs depend on for navigation is assistive devices, which use sensory substitution to convert visual information into auditory or tactile cues to provide spatial awareness often incorporate IoT sensors to capture real-time environmental data, providing navigational assistance in complex environments [6].

Sensory substitution devices or SSDs are assistive wearable devices that can be eyewear, a vest, or a hat, allowing VIP users to perceive their surroundings by converting sensory information from one modality to another [7]. For our thesis we will focus on developing a SSD that is a hat. These SSDs are embedded with sensors that allow obstacle detection, navigation, and object recognition, enhancing VIPs' daily independence and reducing their reliance on traditional aids [8]. However, global adoption of these devices is limited due to cognitive overload performance, extensive training needs, ergonomic discomfort design, and the lower processing bandwidth of non-visual senses compared to vision [9]. This is what our thesis will focus on, to experiment to see if by implementing controllable selective substitution that we could overcome this sensory limitation.

A. *Problem Identification*

This thesis addresses a global problem aligned with the United Nations Sustainable Development Goal (SDG) 10, which is to reduce inequalities within and among countries based on income, sex, age, disability, sexual orientation, race, class, ethnicity, religion, and oppotunity [3]. Specifically, our target problem is to make it possible for visually impaired persons (VIPs) to have approximately the same level of opportunities as the general population [10] by minimizing the major limitations currently faced by sensory substitution devices (SSDs), thereby contributing to the achievement of several SDG 10 targets.

SDG Target 10.1 aims to achieve and sustain income growth for the bottom 40% of the population at a rate higher than the national average [11]. Improved sensory substitution devices (SSDs) can directly support this by expanding job opportunities for visually impaired individuals, thereby reducing income inequality [12]. Target 10.2 seeks to empower and promote the social, economic, and political inclusion of all irrespective of disability [11]; effective SSDs contribute by enabling visually impaired people to perform tasks and interact beyond their current sensory limitations [13]. Target 10.3 calls for ensuring equal opportunity and reducing inequalities of outcome through the elimination of discriminatory practices [11]; better SSDs help by providing practical capabilities that allow visually impaired individuals to access opportunities currently out of reach [14]. Target 10.4 encourages the adoption of fiscal, wage, and social protection policies to achieve greater equality [11]; when SSDs improve employability and workforce participation, they strengthen both the evidence and societal demand for such inclusive policies [15]. The remaining targets (10.5–10.b), which address global financial regulation, official development assistance, and migration policies, fall outside the direct scope of a technology-focused SSD intervention.

The inequalities the VIPs faced we specifically focused on are two major but often overlapping categories: navigational and social [16]. Navigational inequalities which are barriers in physical and digital infrastructure not designed for accessibility such as building that lack tactile guides or website and software without audio alterantives [17]. Social inequalitities which are limitations in communication, societal status, or workforce participations, for instance only about 44% of visually impaired individuals are employed compaired to 79% of those without disabilities, often due to employer attitudes and lack of accommodations [18]. These two categories of inequalities could be minimized if the VIPs partially have the ability they lack, this could be achieved through Sensory Substitution Devices [19].

B. *Research Objectives and Contributions*

The primary objective of this research is to minimize inequalities for visually impaired individuals by developing and testing substitution technologies, specifically sensory substitution devices (SSDs) that convert visual information into accessible formats like audio, thereby enabling more independent navigation and social interaction [4]. This approach prioritizes substitution to address core sensory gaps holistically, while acknowledging potential hybrids with assistive elements, and aligns with SDG 10 by promoting economic inclusion through expanded job opportunities (Target 10.1), empowerment via regained access and participation (Target 10.2), and equal opportunities by reducing discriminatory barriers (Target 10.3) [3].

The key novelty in solving the sensory limitation issues that SSDs face is through incorporating controllability for selective perception [20]. This idea is inspired by psychological principles where the human sense filters relevant information to avoid overload [21]. To accomplish this our device should allow the sensory translation to be controlled by the user like through hand gestures, allowing realtime customization of translation and addressing the Sensory limitation issue of SSDs like overload, fatigue, or integration [9]. By demonstrating that this is possible, it could pave the way for scalable applications to other sensory disabilities, enhancing the fields of assistive technology with user-centered, efficient substitution framework [22].

II. Related Work

There are existing solutions in providing VIPs a way to navigate the world, but from our research there are two major overlapping categories of these solutions which are Assistive and Substitution technologies [4]. Assistive technologies which primarily extend existing capabilities such as canes, guide dogs, or braille displays [26]. Substitution technology which primarily represent missing senses through alternatives such as screen readers and sensory substitution devices that convert visual data to audio feedback [20]. Since our thesis aims to minimize inequalities of the VIPs by providing them a way to navigate the physical and digital world so that they could have approximately the same opportunities as the general population, we focused on Substitution technology, more specifically visual substitution devices [27].

A. *Assistive and Substitution Technologies*

Reference [20] designs their SSD for blindness using touch and audio, similarily following design princicples like task-focused information conveyance to avoid redundancies and match the sensory bandwidths, highlighting the issues faced by SSDs like sensory overload from exessive data so there must be a selective process of key environmental cues to mitigate it. The bandwidth mismatches does lead to persistent fatigue and high training requirements delay usability and integration. Reference [23] introduces a multi-device combination SSD, lowers overload by distributiong cognitive load, enchances integration through perceptual reweighting, and improving long-term navigation accuracy. The initial dual-use however, increases fatigue without extended training, highlighting integration challenges for less experienced users. Reference [24] reviews vision substitution via hearing and touch, implements preprocessing filters to reduce overload by focusing on essential data, plasiticity aids long-term integration, and enables better nativation equity. The bandswith limits still causes persistent information loss and fautigue, often struggling with controllability and requires user adaption. Reference [25] device similarly is also a wearable portable smart hat, is low-cost, minimizes navigational inequalitiess, and provides real-time feedback to improve integration and learning. It lacks selective processing and can lead to sensory overloading in cluttered environments and can cause fautigue over prolonged use.

B. Gaps in Addressing Sensory Limitations

These SSDs highlights and face a common issue which we can generally call as Sensory Limitation, these include sensory overload where the translated information is just too much for the user to actually process either through auditory bandwidth limit or interferance, sensory fautigue where hearing audio repetively can cause discomfort and less sensory perception, and finally sensory integration where the translated audio representation cant be subconsciously integrated by the user's perception of their environment making learning harder and daily use tiring [20]. Our thesis aims to minimize sensory limitation through high controllable selective translation where specific environment cues can be selectively filtered by the user directly and for any period of time, minimizing overload, fautigue, and allowing more controllable integration like allowing modular learning when using it for the first time [23].

III. Proposed Approach / Methodology

Our approach is to generate a depth map from a camera using MiDAS, the hand guesture is also captured from the same camera image using MediaPipe. The hand gesture specifically the back of the right hand determines what is to be exectuted by the command, for this instance it has a few pre-installed gestures:

- **Open hand gesture** which translates the depth map into audio by cycling between right and left side where each slice of the image is translated into audio with the vertical position representing frequency and depth/proximiting representing volume represented by the equation

$$A_c(T,s) = \sum_{n=0}^{B-1} \frac{D_{t(T)}\left[B-1-n\right]}{255 \times 10} \cdot G_c(t(T)) \cdot \sin\left(2\pi f_n \frac{s}{R}\right) \tag{1}$$

  Where:
  - c - Channel number, 0 for right and 1 for left
  - A_c - Audio output for channel c
  - T - Time in seconds
  - C - Cycle duration in seconds *(Configurable by "-cd=", default* 4.0*)*
  - $t(T)$ - Normalizes time T to be between 0 and 1 occilating triangularly
  - $D_t[n]$ - Depth map vertical slice at horizonal position t, frequency index n *(*0 *to* 255*)*
  - B - Number of frequency bins *(Configurable by "-nb=", default* 64*)*
  - n - Frequncy bin index *(*0 *to* $B-1$*)*
  - $G_c(t)$ - Channel gain factors $t - 2ct + c$
  - s - Sample index within the current audio block *(*0 *to* $S-1$*)*
  - S - Block size in samples *(Configurable by "-bs=", default* 512*)*
  - R - Sample rate in Hz *(Configurable by "-sr=", default* 44100*)*
  - f_n - Frequency in Hz for bin n, logarithmically spaced $f_n = 100 \cdot \left(\frac{8000}{100}\right)^{\frac{n}{B-1}}$
- **Closed hand gesture** which is similar to open hand gesture but only translate the middle section of the depth map allowing the user to control the scanning at their own control, represented by the forumula

$$K_c(T,s) = A_{\frac{1}{2}}(T,c) \tag{2}$$

  Where:
  - K_c - Audio output for channel c
  - A_c - Original audio output *(open hand gesture)* for channel c
- **Call me hand gesture** calls a specific user in facebook, just to test the server API and to ensure that future continuations would allow for modular or additional applications.

A. System Architecture

B. Hypothesis and Design Principles

C. Implementation Details

IV. Experimental Evaluation

A. Experimental Setup

B. Evaluation Metrics

V. Results

A. Quantitative Results

B. Qualitative Insights

VI. Discussion

A. Implications and Impact

B. Limitation and Future Directions

VII. Conclusion

Acknowledgment

References

[1] P. Theodorou, K. Tsiligkos, and A. Meliones, "Multi-sensor data fusion solutions for blind and visually impaired: Research and commercial

navigation applications for indoor and outdoor spaces" *Sensors*, vol. 23, no. 12, p. 5411, Jun. 2023, doi: 10.3390/s23125411. Available: https://doi.org/10.3390/s23125411

[2] World Health Organization: WHO, "Blindness and visual impairment" Aug. 10, 2023. Available: https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment

[3] Martin "Goal 10: Reduce inequality within and among countries" *United Nations Sustainable Development*, Nov. 25, 2025. Available: https://www.un.org/sustainabledevelopment/inequality

[4] Z. J. Muhsin, R. Qahwaji, F. Ghanchi, and M. Al-Taee, "Review of substitutive assistive tools and technologies for people with visual impairments: recent advancements and prospects" *Journal on Multimodal User Interfaces*, vol. 18, no. 1, pp. 135–156, Dec. 2023, doi: 10.1007/s12193-023-00427-4. Available: https://doi.org/10.1007/s12193-023-00427-4

[5] A. Arvind, "A deep neural architecture search Net-Based wearable object classification system for the visually impaired" in *Communications in computer and information science*, 2023, pp. 198–213. doi: 10.1007/978-3-031-46338-9_15. Available: https://doi.org/10.1007/978-3-031-46338-9_15

[6] S. Maidenbaum, S. Abboud, and A. Amedi, ""Sensory substitution: Closing the gap between basic research and widespread practical visual rehabilitation" *Neuroscience & Biobehavioral Reviews*, vol. 41, pp. 3–15, Nov. 2013, doi: 10.1016/j.neubiorev.2013.11.007. Available: https://doi.org/10.1016/j.neubiorev.2013.11.007

[7] A. Mishra, Y. Bai, P. Narayanasamy, N. Garg, and N. Roy, "Spatial Audio Processing with Large Language Model on Wearable Devices" *arXiv.org*, Apr. 11, 2025. Available: https://arxiv.org/abs/2504.08907

[8] I. Tokmurziyev, M. A. Cabrera, M. H. Khan, Y. Mahmoud, L. Moreno, and D. Tsetserukou, "LLM-Glasses: GenAI-driven Glasses with Haptic Feedback for Navigation of Visually Impaired People" *arXiv.org*, Mar. 04, 2025. Available: https://arxiv.org/abs/2503.16475

[9] Y. Hou, Q. Xie, N. Zhang, and J. Lv, "Cognitive load classification of mixed reality human computer interaction tasks based on multimodal sensor signals" *Scientific Reports*, vol. 15, no. 1, p. 13732, Apr. 2025, doi: 10.1038/s41598-025-98891-3. Available: https://doi.org/10.1038/s41598-025-98891-3

[10] J. D. Steinmetz et al., "Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: the Right to Sight: an analysis for the Global Burden of Disease Study" *The Lancet Global Health*, vol. 9, no. 2, pp. e144–e160, Dec. 2020, doi: 10.1016/s2214-109x(20)30489-7. Available: https://pubmed.ncbi.nlm.nih.gov/33275949/

[11] United Nations, "The Sustainable Development Goals Report 2025". New York, NY, USA: United Nations, 2025. Available: https://unstats.un.org/sdgs/report/2025/The-Sustainable-Development-Goals-Report-2025.pdf

[12] Eurostat, "Employment gaps for women & people with disabilities" *Eurostat*, May 27, 2025. Available: https://ec.europa.eu/eurostat/en/web/products-eurostat-news/w/ddn-20250527-1

[13] V. Alcaraz-Rodríguez, D. Medina-Rebollo, A. Muñoz-Llerena, and J. Fernández-Gavira, "Influence of Physical Activity and Sport on the Inclusion of People with Visual Impairment: A Systematic Review" *International Journal of Environmental Research and Public Health*, vol. 19, no. 1, p. 443, Dec. 2021, doi: 10.3390/ijerph19010443. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC8744778/

[14] J. Shaw, M. Wickenden, S. Thompson, and P. Mader, "Achieving disability inclusive employment – Are the current approaches deep enough?" *Journal of International Development*, vol. 34, no. 5, pp. 942–963, Jul. 2022, doi: 10.1002/jid.3692. Available: https://doi.org/10.1002/jid.3692

[15] U.S. Bureau of Labor Statistics, "Labor force participation rate 24.2 percent for people with a disability in 2023" *The Economics Daily*, Oct. 2024. Available: https://www.bls.gov/opub/ted/2024/labor-force-participation-rate-24-2-percent-for-people-with-a-disability-in-2023.htm

[16] "Disability Employment Research: Key takeaways" *The American Foundation for the Blind*, 2021. Available: https://afb.org/research-and-initiatives/employment/reviewing-disability-employment-research-people-blind-visually

[17] A. T. Parker, M. Swobodzinski, J. D. Wright, K. Hansen, B. Morton, and E. Schaller, "Wayfinding tools for people with visual Impairments in Real-World Settings: A literature review of recent studies" *Frontiers in Education*, vol. 6, Oct. 2021, doi: 10.3389/feduc.2021.723816. Available: https://doi.org/10.3389/feduc.2021.723816

[18] M. C. McDonnall and Z. Sui, "Employment and Unemployment Rates of People Who Are Blind or Visually Impaired: Estimates from Multiple Sources" *Journal of Visual Impairment & Blindness*, vol. 113, no. 6, pp. 481–492, Nov. 2019, doi: 10.1177/0145482x19887620. Available: https://doi.org/10.1177/0145482x19887620

[19] M. M. Billah, Z. M. Yusof, K. Kadir, and A. M. M. Ali, "Sensory substitution for Visual impairments: A Technological review" *IntechOpen eBooks*, Dec. 2019, doi: 10.5772/intechopen.89147. Available: https://www.intechopen.com/chapters/69033

[20] Á. Kristjánsson et al., "Designing sensory-substitution devices: Principles, pitfalls and potential1" *Restorative Neurology and Neuroscience*, vol. 34, no. 5, pp. 769–787, Aug. 2016, doi: 10.3233/rnn-160647. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC5044782/

[21] K. C. MSEd, "How we use selective attention to filter information and focus" *Verywell Mind*, Dec. 18, 2023. Available: https://www.verywellmind.com/what-is-selective-attention-2795022

[22] T. Lloyd-Esenkaya, V. Lloyd-Esenkaya, E. O'Neill, and M. J. Proulx, "Multisensory inclusive design with sensory substitution" *Cognitive Research Principles and Implications*, vol. 5, no. 1, p. 37, Aug. 2020, doi: 10.1186/s41235-020-00240-7. Available: https://doi.org/10.1186/s41235-020-00240-7

[23] C. Jicol et al., "Efficiency of sensory substitution devices alone and in combination with Self-Motion for spatial navigation in sighted and visually impaired" *Frontiers in Psychology*, vol. 11, p. 1443, Jul. 2020, doi: 10.3389/fpsyg.2020.01443. Available: https://doi.org/10.3389/fpsyg.2020.01443

[24] J. M. Loomis, R. L. Klatzky, and N. A. Giudice, "Sensory substitution of vision: Importance of perceptual and cognitive processing" in *Assistive Technology for Blindness and Low Vision*, R. Manduchi and S. Kurniawan, Eds. Boca Raton, FL, USA: CRC Press, 2013, pp. 162-191. Available: https://umaine.edu/vemi/wp-content/uploads/sites/220/2016/06/Loomis-Klatzky-Giudice-in-press-sensory-substitution-chapter.pdf

[25] M. Sami, N. D. Agha, J. Baloch, A. Dewani, and K. D. Maheshwari, "fficient object detection and Voice-Assisted navigation for the visually impaired: the Smart Hat approach" *Sukkur IBA Journal of Computing and Mathematical Sciences*, vol. 8, no. 2, pp. 1–12, Feb. 2025, doi: 10.30537/sjcms.v8i2.1535. Available: https://journal.iba-suk.edu.pk:8089/SIBAJournals/index.php/sjcms/article/view/1535

[26] Moth, "List of the best assistive devices for the blind" *InviOcean*, Oct. 03, 2025. Available: https://inviocean.com/learn/best-assistive-technologies-for-visually-impared/

[27] E. Casanova, D. Guffanti, and L. Hidalgo, "Technological Advancements in Human Navigation for the visually Impaired: A Systematic review" *Sensors*, vol. 25, no. 7, p. 2213, Apr. 2025, doi: 10.3390/s25072213. Available: https://doi.org/10.3390/s25072213