Innovative Assistive Technology: Speech-Based Obstacle Detection for the Blind

Sangita Roy, Moupali Roy, Arpita Santra

ECE Department, Narula Institute of Technology, Kolkata, India

Abstract The development and deployment of a speech-based blind obstacle detection system aim to enhance the mobility and independence of individuals with visual impairments by providing real-time obstacle detection and audio feedback. The system integrates advanced sensor technologies such as ultrasonic and infrared sensors to detect objects, determine their location and distance, and provide speech cues. A microcontroller processes the sensor data and translates it into verbal alerts using a text-to-speech module, ensuring clear and concise communication. Designed for portability and ease of use, the system is lightweight, wearable, and features an efficient power management mechanism. It adapts to diverse environmental conditions, maintaining accuracy and reliability. By combining advanced hardware with intelligent software, this solution offers a cost-effective, accessible alternative to traditional mobility aids, marking progress in assistive technology to enhance the quality of life and independence of users.

Introduction The increasing prevalence of visual impairment underscores the need for innovative assistive technologies to improve mobility and independence. Traditional aids such as canes and guide dogs, while useful, have limitations in obstacle detection, adaptability, and accessibility. A speech-based blind obstacle detection system leverages advancements in sensor technology, real-time processing, and audio communication to overcome these challenges. By providing real-time vocal feedback, the system empowers visually impaired individuals to navigate safely. Utilizing ultrasonic or infrared sensors, a microcontroller for data processing, and a text-to-speech module for audio feedback, the system ensures precise and timely information delivery. Its compact and ergonomic design enhances portability and usability, while its adaptability to different environmental conditions makes it more effective than traditional mobility aids.

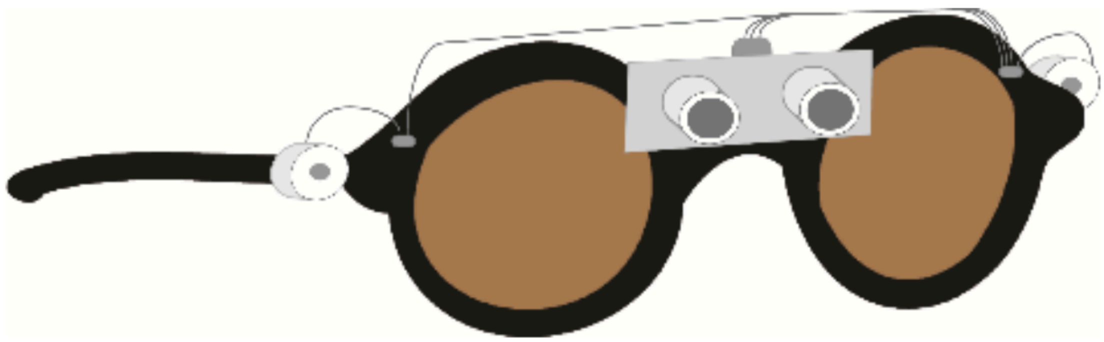

Figure 1. Ultrasonic smart glasses [10]

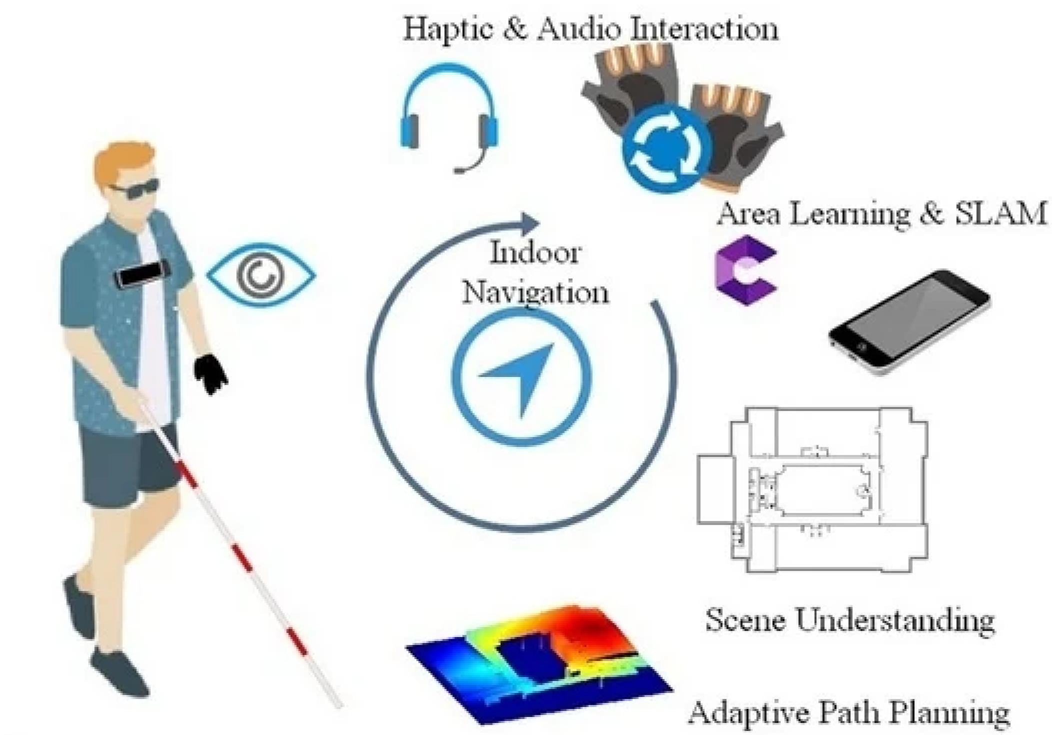

Figure 2. The components of the proposed ANSVIP system [11].

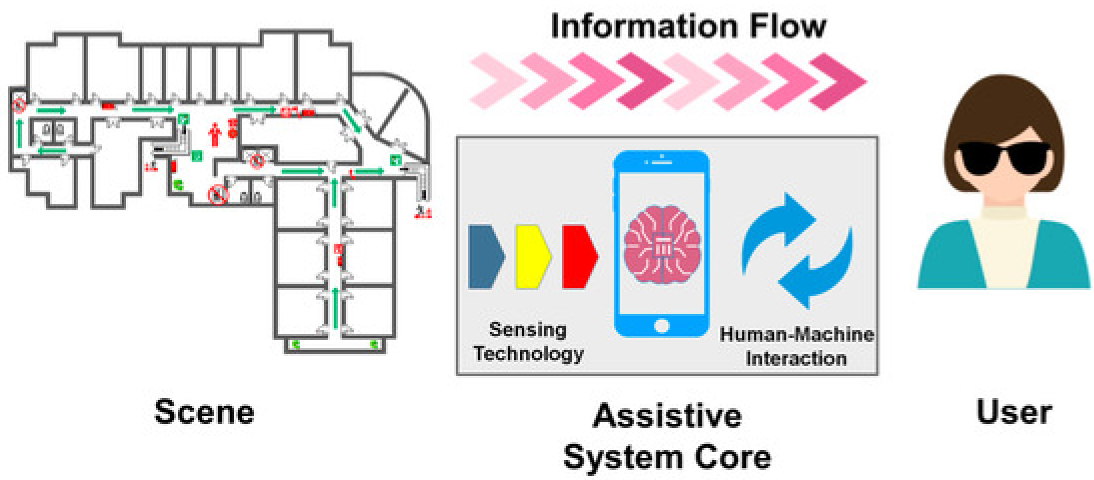

Figure 3. The information flow in the assistive system: The assistive system core aims to understand the world and translate the essential understanding to the user [11].

1. System Architecture The proposed system consists of four key components:

1.1 Sensors: Sensors play a crucial role in detecting and identifying obstacles. Ultrasonic sensors measure distances by emitting high-frequency sound waves and analyzing reflected signals. Infrared sensors provide close-range detection, enhancing portability. The sensor array is strategically positioned to detect obstacles at different heights and angles, ensuring comprehensive coverage.

- **Ultrasonic Sensors:** Detect obstacle distances.
- **LiDAR or Cameras:** Capture environmental data for object classification.
- **Infrared Sensors:** Identify close-range obstacles and uneven surfaces.

1.2 Processing Unit: A microcontroller or microprocessor processes sensor data, determines obstacle location and distance, and converts the information into real-time auditory feedback.

- **Object Detection & Classification:** Machine learning models classify obstacles.
- **Direction Estimation:** Algorithms identify obstacle position.
- **Distance Calculation:** Real-time computation of obstacle proximity.

1.3 Speech Synthesis: Processed data is converted into clear audio feedback using text-to-speech technology. The system generates real-time speech cues, ensuring prompt and precise navigation guidance.

Text-to-Speech (TTS): Converts processed data into spoken instructions, e.g., "Obstacle ahead, 3 meters away."

1.4 User Feedback: User feedback ensures system usability and effectiveness. Customizable speech settings improve clarity, while iterative testing refines response accuracy.

- **Audio Output:** Delivered through headphones or bone-conduction devices to preserve environmental awareness.

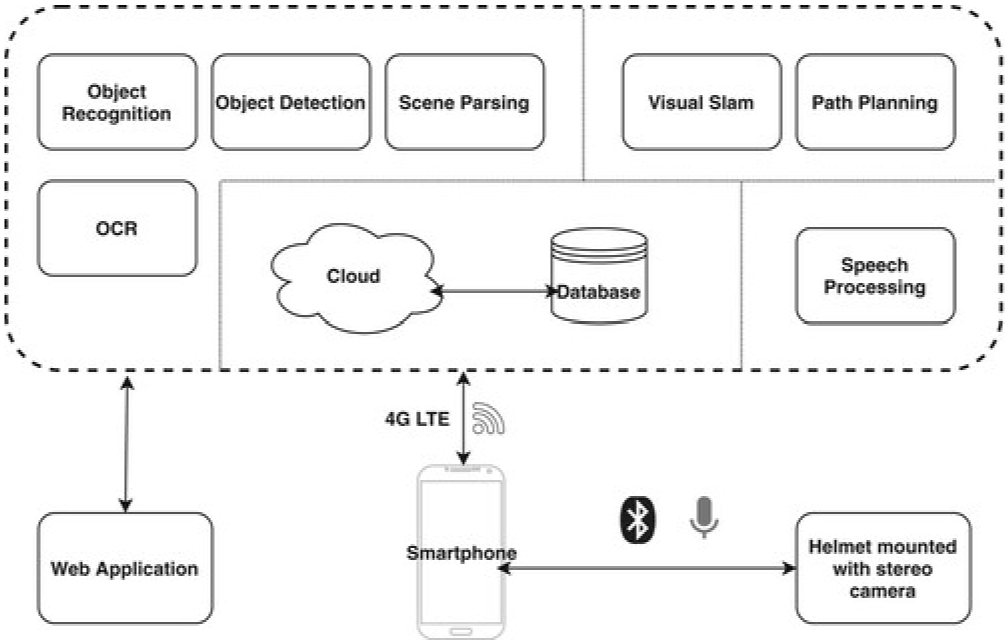

Figure 4: Cloud and vision-based navigation system [12]

2. System Flowchart and Block Diagram 2.1 Flowchart:

1. **Start**: Power on and initialize components.
2. **Sensor Data Collection**: Gather input from ultrasonic, LiDAR, and infrared sensors.
3. **Data Processing**: Detect, classify, and determine obstacle distance and position.
4. **Obstacle Detection**: Identify obstacles and their characteristics. If none, continue monitoring.
5. **Decision-Making**: Classify objects and estimate proximity.
6. **Speech Feedback**: Generate verbal cues based on object type and location.
7. **User Alert**: Provide directional audio feedback.
8. **Repeat**: Continuously process new data.

2.2 Block Diagram:

- **Input Sensors:** Ultrasonic Sensors, LiDAR, Infrared Sensors.
- **Processing Unit:** Object Detection, Direction Estimation, Distance Measurement.
- **Speech Generation:** Text-to-Speech Module.
- **Output Feedback:** Audio signals via headphones or bone-conduction devices.

3. Methodology

3.1 Data Collection: Real-time environmental data is gathered using sensor modules such as ultrasonic and LiDAR sensors. The data is **processed** to remove noise and extract relevant obstacle features.

- **Continuous Sensor Monitoring:** Captures obstacle dimensions, positions, and movement.

3.2 Data Processing: Raw sensor data is processed into meaningful feedback. Noise is filtered out, and machine learning **models** improve object classification accuracy.

- **Object Classification:** Neural networks (e.g., YOLO) identify obstacles.
- **Direction Detection:** Algorithms determine obstacle angle relative to the user.
- **Proximity Assessment:** Ultrasonic sensors measure object distance.

3.3 Speech Feedback Generation: Processed obstacle data is transformed into real-time speech output. Customizable settings allow for user-specific adjustments in speech clarity and volume.

- **Structured Voice Alerts:** "Person approaching on the right."
- **TTS Engine:** Converts data into real-time speech cues.

4. Implementation

4.1 Hardware Components:

- **Sensors:** Ultrasonic and LiDAR for precise detection.
- **Microcontroller:** Raspberry Pi or Arduino for data processing.
- **Audio Output:** Headphones or speakers.

4.2 Software Components:

- **Programming Language:** Python for sensor integration and speech synthesis.
- **Libraries:** TensorFlow/PyTorch for object detection, pyttsx3 for TTS.
- **Operating System:** Raspbian (for Raspberry Pi implementation).

5. Results and Discussion The system was evaluated in both controlled and real-world settings, demonstrating:

- **High Accuracy:** Reliable obstacle detection up to 5 meters.

- **Real-Time Feedback:** Minimal delay in generating speech alerts.
- **Positive User Experience:** Enhanced confidence in navigation among visually impaired participants.

Challenges include sensor range limitations and environmental interference. Future improvements will focus on optimizing detection precision and user experience.

6. Conclusion This paper presents a speech-based obstacle detection system integrating sensor technology, AI, and speech processing to enhance navigation for visually impaired individuals. Future enhancements will improve object classification and integrate GPS for expanded navigation capabilities.

Acknowledgment The research and development of this system were made possible through iterative testing and user feedback. Future developments will focus on refining obstacle detection and enhancing user adaptability for greater real-world effectiveness.

[1]. A. G. Nicholas, and G. E. Legge, "Blind navigation and the role of technology," in The Engineering Handbook of Smart Technology for Aging, Disability, and Independence, 2008, pp. 479–500.

[2]. A. Riazi, F. Riazi, R. Yoosfi, and F. Bahmeei, "Outdoor difficulties experienced by a group of visually impaired iranian people," *J. Curr. Ophthalmol.*, Vol. 28, no. 2, pp. 85–90, 2016. doi: https://doi.org/10.1016/j.joco.2016.04.002

[3]. R. Manduchi, S. Kurniawan, and H. Bagherinia, "Blind guidance using mobile computer vision: a usability study," in ASSETS, 2010, pp. 241–242.

[4]. B.-S. Lin, C.-C. Lee, and P.-Y. Chiang, "Simple smartphone-based guiding system for visually impaired people," *Sensors*, Vol. 17, no. 6, pp. 1371, 2017. doi: https://doi.org/10.3390/s17061371

[5]. Y. Zhao, E. Kupferstein, D. Tal, and S. Azenkot, “It looks beautiful but scary: How low vision people navigate stairs and other surface level changes,” in Proceedings of the 20th International ACM SIGAC-CESS Conference on Computers and Accessibility, ACM, 2018, pp. 307–320.

[6]. H. Devlin. Echolocation could help blind people learn to navigate like bats, Feb 2018.

[7]. Acoustical Society of America. Exploring the potential of human echolocation, Jun 2017.

[8]. L. Thaler, and M. A. Goodale, “Echolocation in humans: an overview,” *Wiley* Interdisciplinary *Reviews: Cognitive Science*, Vol. 7, no. 6, pp. 382–393, 2016.

[9].M. Srikulwong, and E. O’Neill, “Tactile representation of landmark types for pedestrian navigation: user survey and experimental evaluation,” in Workshop on using audio and Haptics for delivering spatial information via mobile devices at MobileHCI 2010, 2010, pp. 18–21.

[10]. M. C. Holbrook, and A. Koenig. History and theory of teaching children and youths with visual impairments. Foundations of Education. Volume I:, 2000.

[10]. Agarwal, R.; Ladha, N.; Agarwal, M.; Majee, K.K.; Das, A.; Kumar, S.; Rai, S.K.; Singh, A.K.; Nayak, S.; Dey, S.; et al. Low cost ultrasonic smart glasses for blind. In Proceedings of the 2017 8th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, Canada, 3–5 October 2017; pp. 210–213.

[11]. Zhang, X., Yao, X., Zhu, Y., & Hu, F. (2019). An ARCore Based User Centric Assistive Navigation System for Visually Impaired People. Applied Sciences, 9(5), 989. https://doi.org/10.3390/app9050989.

[12]. J. Bai, D. Liu, G. Su, and Z. Fu, "A cloud and vision-based navigation system used for blind people," in Proceedings of the 2017 International Conference on Artificial Intelligence, Automation and Control Technologies, ACM, 2017, pp. 22.

[13. Redmon, J., et al. (2016). YOLO: Real-Time Object Detection. [Paper Reference]

[14]. Goodfellow, I., et al. (2016). Deep Learning. MIT Press.

[15]. Google TTS Documentation: https://cloud.google.com/text-to-speech

[16]. TensorFlow Documentation: https://www.tensorflow.org

[17]. https://www.mdpi.com/

[18].https://www.researchgate.net/figure/Flowchart-for-moving-obstacle-detection_fig1_2959
19039