Voice Interactive Indoor Navigation Application for the Visually Impaired on Android Phones with Real-Time Obstacle Detection using Augmented Reality with ARCore

Hayath T M[1], Udayshankar Ravikumar[2]
[1]*Assistant Professor, Department of CSE, Ballari Institute of Technology and Management, Ballari, Karnataka*
Email: hayath@bitm.edu.in
[2]*Chief Technology Officer, Urav Advanced Learning Systems Private Limited*
Email: udayshankar.cto@uralstech.in

Abstract: An Augmented Reality voice interactive assistive indoor navigation application for the visually impaired is described here using ARcore on Android phones. Human–machine interaction mechanism is voice activated with turn-by-turn continuous guidance. The app uses ARCore supported Android smartphones to acquire robust computer vision-based localization. Path creation is done using local anchors, which are then uploaded to the cloud. The ARCore Depth API is used for real-time obstacle detection and the app gives voice warning to the user. A prototype named 3rDi 4 All is implemented with the intended functionality and published on google PlayStore.

Keywords: augmented reality: AR; assistive technology; ARCore; depth sensor; indoor; voice; android; phones; navigation aids; visual impairment; SLAM; obstacle detection

I. INTRODUCTION

It is estimated that in 2020, about 596 million people had vision impairment worldwide. Of these, about 43 million were blind [1]. This number is projected to increase significantly with an aging world population and also due to the increase in the noncommunicable disease burden. People with vision impairment are facing difficulties with independent mobility, especially in unfamiliar environments, including obstacles and potential navigation paths, for the purpose of mobility. Since most of the directions and signages are based on visual cues, the visually challenged are prevented from effectively utilizing the existing navigation infrastructure. Indoor public spaces like bus and train terminals, schools and universities, airports and hospitals, shopping malls and public offices do not offer navigation aids for this specific condition. This severely restricts their mobility and prevents them from independently becoming a productive part of the workforce. The two aspects for proper mobility are path setting and obstacle avoidance. A particular static path between two points of interest can be made without obstacles. Since real world situations are dynamic, there is a chance that obstacles in the form of people or objects may be present in the static path. For mobility in the visually impaired it is necessary to "show" the path through cues other than sight and also warn the user of any obstacles on the path thus allowing the person to navigate without any trouble. All this has to be done in such a way that it is easily portable, comfortable to use and economically feasible.

Outdoor navigation is possible with the help of Global Positioning System (GPS) and such methods, but is more of a problem indoors where robust positioning systems may not be available. Some of the older methods in use for helping the visually challenged were white cane, human assistance and guide animals which were expensive or imperfect in giving unrestricted mobility. Indoor Positioning System (IPS) using available WiFi infrastructure is not a success due to the high variability of the Received Signal Strength Indicator (RSSI) [2]. Building walls and other indoor obstacles reduce the usefulness of this method. Another method of IPS used is the fingerprinting technique with training (offline) and positioning (online) phases. These use the supervised machine learning techniques to train a classifier, such as the k-Nearest Neighbours (kNN) [3] or Hierarchical Navigable Small World (HNSW) [4]. Though the accuracy has been improved, these methods still require a large training dataset for implementation. The recent advances in mobile applications on the ever-ubiquitous smartphones now allow the development of easily portable, innovative solutions to assist visually impaired people.

Fig.1 Graphical Abstract

A recent study [5] listed the various sensor based assistive devices for visually impaired with the advantages and drawbacks of each method. This study also classified the different types of mobility aids proposed by different authors. Systems form smart canes, haptics and laser range finder to using Kinect depth cameras, among many others, are mentioned in this study. Other methods use beacons or QR code to position the user within an environment and thus guide them in the navigation. The paper clearly mentions that none of the techniques were 100% efficient.

This paper describes a smartphone application that can be used through voice interaction to navigate a pre-set path through turn-by-turn instructions and also gives real time audible warning to the user of any obstacles. This application is implemented using ARCore Depth API on supported Android smartphones. ARCore is a software development kit (SDK) that has been developed by Google, which allows for Augmented Reality (AR) applications to be built. The depth sensor API is a feature of ARCore that allows for real-time depth mapping and 3D modelling of a user's environment. ARCore uses Simultaneous locating and mapping technology (SLAM), which allows for centimetre-level accuracy in indoor positioning. The prototype application, 3rDi 4 All, has been published in Google Playstore.

In the next section, we present our related work. The design philosophy including the software and hardware used is explained later. Finally, we discuss the limitations and conclusions.

II. RELATED WORK

Assistive technologies for the visually impaired have been tried out many times in the past with the most basic ones being white canes, guide dogs and a person helping the blind. A 2019 literature review [6] of existing electronic mobility aids for persons who are visually impaired, identified and classified 146 products, systems, and devices. The main thrust of the assistive technologies so far has been in the field of navigation and real-time obstacle detection. For obstacle detection, two major approaches have been sensor-based obstacle detection, the simplest being ultrasonic sensors on white cane helping to warn about obstacles and computer vision-based obstacle detection. Data is collected from a single of multiple sensors mounted on a stand-alone device or making use of the sensor of a multipurpose device. This data is then compared with a threshold value for obstacle detection. In computer-based vision for obstacle avoidance, machine learning or artificial intelligence are used. One of the papers of note for sensor based indoor obstacle detection [7] was one by Huy-Hieu Pham et al., which described the use of Microsoft Kinect Sensor for obstacle detection. Mueen et al. [8] described an IoT device for obstacle avoidance using fog computing.

Augmented Reality on the Microsoft HoloLens was used for spatial mapping in [9]. This was done using the audible 3D mesh of the surroundings generated by the device, whereby a soundscape was implemented letting the users hear the names of the objects around them. ARCore based assistive navigation system was described by Zhang et al. [10] which was implemented on a smartphone, and using affordable haptic accessories, it helped the visually impaired people to navigate indoor without using GPS and wireless beacons. ARCore is used to acquire computer vision-based localization, adaptive artificial potential field (AAPF) is used for path planning with audio and haptic interface for human-machine interaction. While in [11], only Depth Imaging API in a smartphone with ARCore was used for navigation. Many commercial applications have been attempted using the varied technologies available. Of these, some of the smartphone applications are ARIANNA [12], which uses AR to visualise a set of paths with markers and WAYMAP [13], which uses Trace, the indoor location IP of Cambridge Consultants.

In this paper, we have developed a smartphone application that caters to both obstacle avoidance and path planning using ARCore supported devices. Individual paths are made of cloud anchors and the path setting in the indoor environment can be done by the facility owners. Varelas et al. [14] in a 2020 paper showed the accuracy of an AR IPS using anchors. Metadata attached to the cloud anchors helps to describe the nature of the environment of the anchor. It uses the camera on a smartphone, which analyses depth images to process and calculate distances for obstacle and warns the user of the obstacle if it is within a preset distance. A voice interactive user interface is developed for the visually impaired people for turb-by-turn navigation using acoustic cues as per the ITU-T F.921 [15] standards for digital navigation for vision-impaired persons. The application will run on ARCore supported devices and as of November 2022, about 86% of devices supports Depth API [16].

III. DESIGN

The Block Diagram of proposed system is depicted bellow:

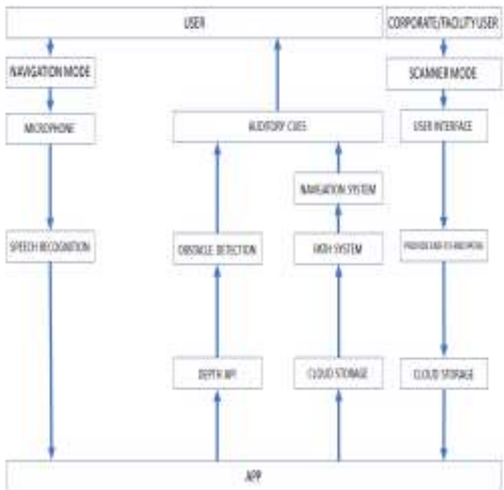

Fig.2 Block Diagram

Software Used

Unity Engine is used as the framework of the app. Unity's ARFoundation helps to wrap all functionalities of the ARCore SDK and ARKit SDK under one unified workflow. So essentially, the app uses the ARCore SDK on Android. Cloud Storage for Firebase is also used to add cloud storage functionality for the app. ARCore is Google's platform for Augmented Reality (AR) experiences for both Android and iOS devices. The ARCore SDK helps in tracking user position, tracked plane detection and raycasting (used for placing virtual objects in the real world). It also helps in camera image access and multithreaded rendering for AR. The depth detection is provided by the ARCore Depth API. Local anchors are used for tracking position and orientation of virtual objects in the real world. Cloud anchors (provided by the ARCore Extensions for ARFoundation) is used for uploading local anchors to Google Cloud to be accessed by any user with a unique ID.

Google's Cloud Storage for Firebase SDK for Unity is used to upload metadata about all paths in a building to the cloud. These metadata files contain information about the paths, such as the name, cloud anchor IDs and time of creation. These metadata files are then downloaded on start-up of the app.

We also use GPS and Google Maps Platform's Geocoding API to access the user location and convert the output coordinates into an address. The address would then be hashed and used as the name of the metadata file uploaded to Firebase Cloud Storage. As the hashing process is not reversible, the users' location would remain private. When searching for the metadata file, the app would simply compare the hash of the users' current location with the name of the metadata file.

For the production version of the application, we use the Firebase Authentication SDK to help authenticate the corporate / facility user. This is essential as not everyone should have access to create or override paths.

Jimmy To's Speech and Text asset for Unity is used for Android and iOS Speech-To-Text (STT) and Text-To-Speech (TTS) integration. The asset allows Unity to use the native STT and TTS services available in Android (android.speech for STT and android.speech.tts for TTS) and iOS (Speech Framework for STT and Speech Synthesis API for TTS).

Hardware Used

The application can be used by any Android smartphone which supports the ARCore Depth API. The Depth API algorithm is based on the processing power of the devise and hence needs mid to high end smartphones. It provides a 3D view of the world by creating depth images using a depth-from-motion algorithm. This means that each pixel in a depth image is associated with a measurement of the distance of the scene from the camera. Multiple images from different angles are taken by the algorithm and comparisons made to estimate the distance to every pixel as the phone is moved by the user. Even with minimal motion from a user, machine learning is used selectively to increase depth processing. Additional hardware available on the device, such as a ToF dedicated depth camera, is also used to increase the accuracy by automatically merging the date from all available sources. This enhances the existing depth image, thus enabling depth even when the camera is not moving. Additional hardware will also provide accurate depth estimation, especially on plain surfaces such as white walls or in dynamic situations with moving people or objects.

For the application development we used a Google Pixel 3a. This is an ARCore Depth API supported device.

User Interface (UI)

The application is used in two modes. The scanner and path creation mode are used to scan and create path by the user of the indoor facility. The navigation mode is the one used by the visually impaired person. Both functionalities are incorporated in the same application thus promoting a minimalist design philosophy. The application starts on the Navigation mode on start up. On double tapping the screen, voice input is necessary to go to the Scanner Mode or in the case of a blind user the destination in input as a voice command (Figure 3).

Fig.3 Scanner Mode

The Scanner Mode UI consists of a debug console, a button to switch to and from the Path Creation and Path Editing sub-modes and a button to switch to Navigation Mode. The Path Creation sub-mode UI, which is shown as default, contains a list of all created paths, a text input area for typing the name of a new path and two buttons for hosting all paths to Firebase Cloud Storage and creating a new path. Each element in the list of paths consists of two buttons to select the path and delete the path and text to show if the path is selected or not. The Path Editing sub-mode UI contains buttons to create, delete, host (to Google Cloud) and resolve anchors and a reticle to show where an anchor would be created (Figure 4).

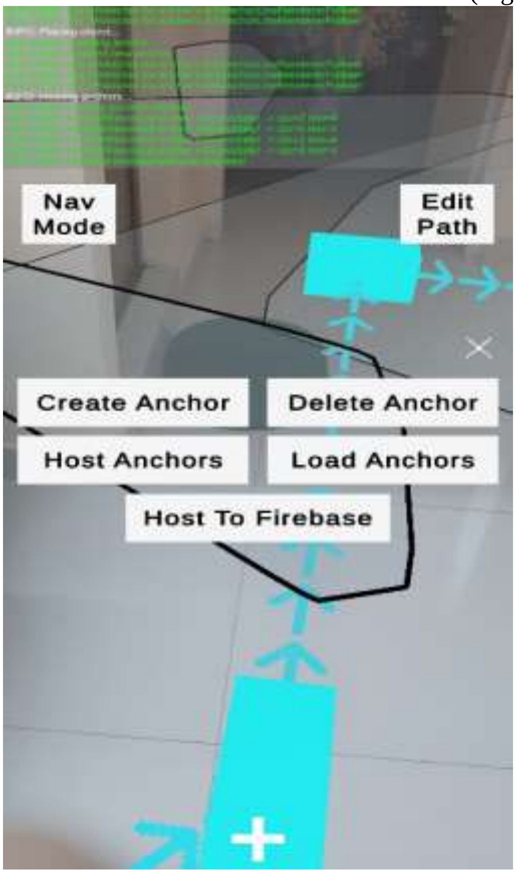

Fig.4 Path Editing sub-mode

We allow the corporate / facility user to add metadata to each created path. This can enable the app to inform the user about special properties of the current path being navigated. For example, if a path ends at an elevator, the corporate / facility user can add metadata to the path to warn the user about the elevator at the end.

The application, when in Navigation Mode, has a very simple UI. It consists of a debug console, text showing if an obstacle has been detected or not, text showing the distance to the nearest object in Centimetres and an image showing the depth map. For STT, a pop-up, used by Androids' STT service takes in the users' reply. The UI also has text which shows the users' reply. All other text in the Navigation Mode UI is for developer debugging.

Human Machine Interaction

In Scanner Mode, the user interacts with the application using the UI. The application interacts with the user using logs in the UI's debug console. In Navigation Mode, the user mainly interacts with the app through tapping the screen. When the user taps the screen twice, the application asks them which path to take. The application communicates with the user exclusively through TTS audio giving them turn by turn navigation. If there is an obstacle in front of the user, it warns about the obstacle and takes them around it to arrive at their destination (Figure 5).

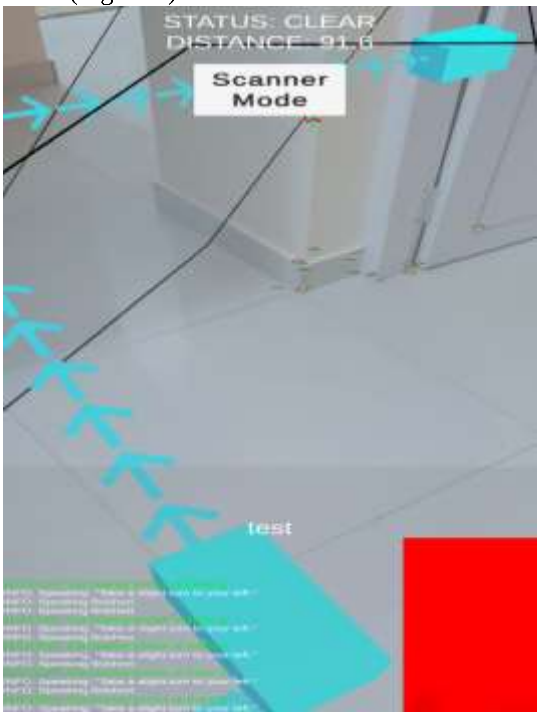

Fig.5 Scanner Mode

IV. LIMITATIONS

An implementation of both obstacle detection integrated with indoor navigation is attempted in this paper through the use of Augmented Reality. A prototype of the application has been developed that can integrate new features as per user feedback. As of now, the application is limited to smartphones which supports Google ARCore Depth API. Another limitation is that the sensor effectiveness can vary depending on performance accuracy, scene, and light conditions. With more robust computer vision with dedicated depth sensing technology such as ToF sensor that can instantly provide depth map without the need to calibrate camera motion, we hope to increase the sensitivity of object detection and also enable robust navigation in indoor spaces.

V. CONCLUSIONS

A smartphone application for indoor mobility assistance of the visually impaired has been successfully developed that provides users with an easily portable, multipurpose navigation device with obstacle detection as part of a single package. It can be controlled through a voice command-enabled user interface. The obstacle detection uses ARCore Depth API for real-time obstruction detection and warns users. The navigation paths can be implemented with metadata about the nature of surroundings. The application can be further improved with the use of different technologies like ToF sensors or LiDAR which will lead to better accuracy. This solution has been evaluated using blindfolder volunteers. The results indicate that the application is suitable for turn-by-turn navigation of visually challenged people through indoor spaces and also gives accurate real-time obstacle detection and warning. A prototype named 3rDi 4 All is implemented with the intended functionality and published on google PlayStore.

VI. REFERENCES

[1] Burton MJ, Ramke J, Marques AP, Bourne RR, Congdon N, Jones I, et al. The Lancet Global Health commission on Global Eye Health: vision beyond 2020. Lancet Glob Health. 2021; 9(4):e489–e551.

[2] Li, G.; Geng, E.; Ye, Z.; Xu, Y.; Lin, J.; Pang, Y. Indoor Positioning Algorithm Based on the Improved RSSI Distance Model. Sensors 2018, 18, 2820

[3] Lin, Tsung-Nan, and Po-Chiang Lin. "Performance comparison of indoor positioning techniques based on location fingerprinting in wireless networks." 2005 international conference on wireless networks, communications and mobile computing. Vol. 2. IEEE, 2005.

[4] Lima, M.; Guimarães, L.; Santos, E.; Moura, E.; Costa, R.; Levorato, M.; Oliveira, H. A Small World Graph Approach for an Efficient Indoor Positioning System. *Sensors* 2021, *21*, 5013. https://doi.org/10.3390/s21155013

[5] Elmannai W, Elleithy K. Sensor-Based Assistive Devices for Visually-Impaired People: Current Status, Challenges, and Future Directions. Sensors (Basel). 2017 Mar 10;17(3):565. doi: 10.3390/s17030565. PMID: 28287451; PMCID: PMC5375851.

[6] Roentgen, U. R., Gelderblom, G. J., Soede, M., & de Witte, L. P. (2008). Inventory of Electronic Mobility Aids for Persons with Visual Impairments: A Literature Review. Journal of Visual Impairment & Blindness, 102(11), 702–724. https://doi.org/10.1177/0145482X0810201105

[7] Huy-Hieu Pham, Thi-Lan Le, Nicolas Vuillerme, "Real-Time Obstacle Detection System in Indoor Environment for the Visually Impaired Using Microsoft Kinect Sensor", Journal of Sensors, vol. 2016, Article ID 3754918, 13 pages, 2016. https://doi.org/10.1155/2016/3754918

[8] Mueen A, Awedh M, Zafar B. Multi-obstacle aware smart navigation system for visually impaired people in fog connected IoT-cloud environment. Health Informatics Journal. 2022;28(3). doi:10.1177/14604582221112609

[9] Augmented Reality for Visually Impaired People, MIMS Final Project 2019 by UC Berkeley School of Information, https://www.ischool.berkeley.edu/projects/2019/augmented-reality-visually-impaired-people

[10] Zhang, X.; Yao, X.; Zhu, Y.; Hu, F. An ARCore Based User Centric Assistive Navigation System for Visually Impaired People. *Appl. Sci.* 2019, *9*, 989. https://doi.org/10.3390/app9050989

[11] See, A.R.; Sasing, B.G.; Advincula, W.D. A Smartphone-Based Mobility Assistant Using Depth Imaging for Visually Impaired and Blind. *Appl. Sci.* 2022, *12*, 2802. https://doi.org/10.3390/app12062802

[12] D. Croce, L. Giarré, F. G. La Rosa, E. Montana and I. Tinnirello, "Enhancing tracking performance in a smartphone-based navigation system for visually impaired people," 2016 24th Mediterranean Conference on Control and Automation (MED), 2016, pp. 1355-1360, doi: 10.1109/MED.2016.7535871.

[13] https://www.cambridgeconsultants.com/press-releases/helping-visually-impaired-people-navigate-indoor-spaces

[14] H. Tm, U. K, M. Shafiulla and Dadapeer, "An Overview of SQL Optimization Techniques for Enhanced Query Performance," 2023 International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE), Ballar, India, 2023, pp. 1-5, doi: 10.1109/ICDCECE57866.2023.10151265.

[15] Varelas, Thomas & Pentefountas, Alexandros & Georgiadis, Charalabos & Kehagias, Dionysios. (2020). AN AR INDOOR POSITIONING SYSTEM BASED ON ANCHORS. MATTER: International Journal of Science and Technology. 6. 10.20319/mijst.2020.63.4357.