1ZJ-03

AiDAR for VI: A LiDAR and AI Based Smartphone Application for Explaining Surroundings to the Visually Impaired

Gurung Sahas† Jee-Eun Kim† Masahiro Bessho†

Department of Information Networking for Innovation and Design, Toyo University†

1. Introduction

Visually impaired individuals face significant challenges when navigating in an indoor environment that they are not used to as mobility, safety, and overall well-being are compromised. Recent advancements in modern smartphone such as the integration of cameras and LiDAR technology have opened new solutions to this critical issue. The main objective of this research is to use the LiDAR scanner and camera of a smartphone and generative AI to create an application that helps visually impaired individuals to perceive the indoor environments.

Blind navigation system that uses Bluetooth Low Energy (BLE) beacons [1] fails when exposed to new environment that does not use BLE beacons. Real-time Obstacle Detection systems [2] can successfully detect obstacles and notify the user but still lack in classifying the obstacles and explaining it to the user. Projects like Be My Eyes [3] can interpret images with text but lacks the capability to detect the distance of the obstacle. Our system provides solution to above mentioned problems by being able to use the application in any situation if connected to the internet and explaining the surrounding in human language while alerting the distance of the obstacle to the user, that have not been solved by existing systems.

2. AiDAR for VI

2.1. Approach

AiDAR for VI is a smartphone application that helps visually impaired person to understand the surroundings in real-time by explaining the surrounding and alerting the user of immediate obstacle by measuring the distance of objects from the device. The application has two main functions:

- Using GPT-4 Vision to explain the details of surrounding to the user
- Using LiDAR Scanner of Smartphone to measure distance and alert the user of immediate obstacle

The application can be controlled by simple 4 gestures, single tap, double tap, swipe-up and swipe-down, and are very easy to learn.

2.2. Explaining Surroundings

GPT-4 with Vision's API was used to add this function to the application. When the user swipes up or down, a snapshot is taken by the camera and is send to GPT-4 and depending on whether the user swiped up or down, an appropriate prompt sent to GPT-4 for analysis. If swiped up, "Describe the picture to a visually impaired person in 20 words." (in Japanese) is sent to GPT-4 and if swiped down, " Describe the picture to a visually impaired person in details." (in Japanese) is sent. Then the application waits for the response from GPT-4 and after the response is received from GPT-4,

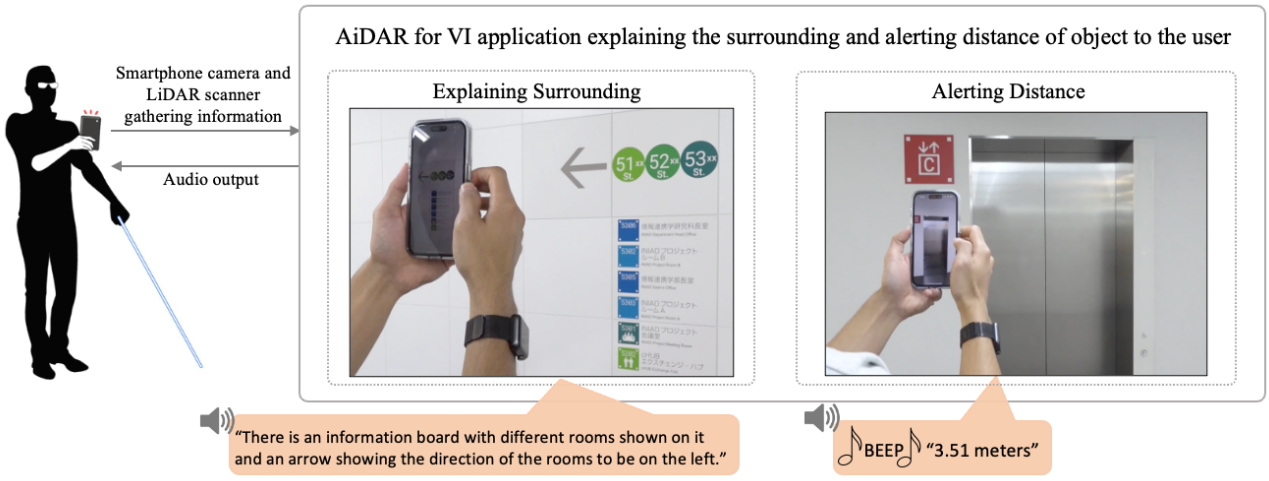

Figure 1 AiDAR for VI is an application that helps visually impaired person to understand the surroundings in real-time.

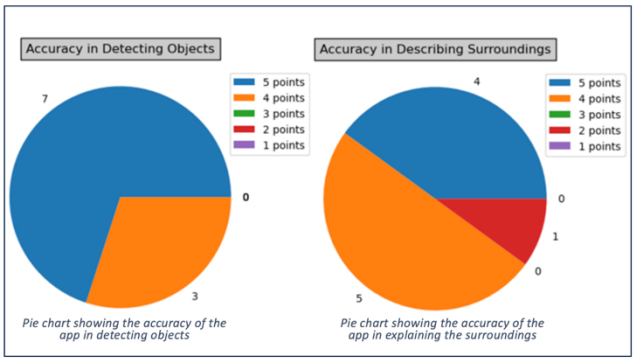

Figure 2 Result of Experiment 1

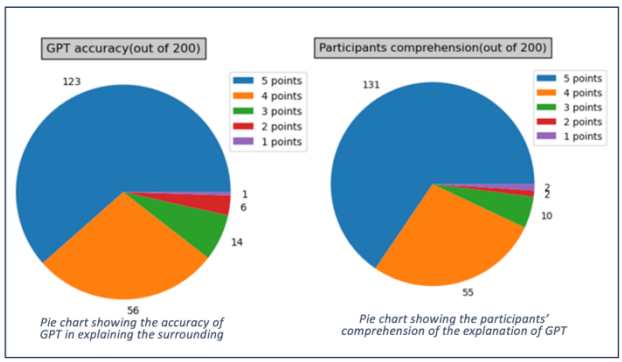

Figure 3 Result of Experiment 2

it is outputted through the speaker.

2.3. Alerting Distance

LiDAR scanner is used to determine the distance of the object from the device for accurate measuring. The distance of the object from the device is measured in real-time and when the distance is obtained, it is conveyed to the users through the audio output. The measured distance is also used to determine which beep alarm sound to play from the three different Beep alarm sounds, each having different frequency and pitch. As the measured distance decreases, the frequency and pitch of the alarm increases.

3. Evaluation

3.1. Procedure

We conducted two experiments that aims to evaluate the application's performance in assisting users with real-time navigation, alerting distance, and surrounding understanding. 10 participants were equipped with a blindfold to simulate a visually impaired scenario. In Experiment 1, participants navigated through a predefined indoor space using the application's alerting distance and explaining the surrounding function. In Experiment 2, 20 images of different indoor scenes were displayed on a projector screen. Each participant was tasked with commanding the app to explain 10 images briefly and 10 images in detail. Points from 1 to 5 were assigned based on the accuracy of the application's explanations and the comprehension of the participants, 1 being the lowest and 5 being the highest.

3.2. Result

From the data, we can see that most of the participants has given 4 or 5 points for the fields of the app experience. The participants were able to be aware of the obstacles most of the time while using this app even when they were blindfolded, and they were able to visualize the surroundings with the explanation of the application.

From the data, we can see that the GPT accuracy and the participants comprehension has received 4 or 5 points for most of the questions. The app was able to correctly explain the surroundings in most of the cases. The participants were also able to understand the explanation of the application. Lower points were received in few cases as GPT sometime gave wrong explanation especially when the images were blurry or hard to read. Some participants were also struggling with length and speed of the detailed explanation.

4. Conclusion

We proposed 'AiDAR for VI', a LiDAR and AI based smartphone application for explaining surroundings to the visually impaired. This application has the function of taking a photograph of the surroundings and explaining the details in the photograph using GPT-4 with Vision and measuring the distance of objects that are in front of the user and alerting the user if the object is too close to prevent collision with the object. Then, a real-world user study was performed to check the potential of the app. Results showed that AiDAR for VI was useful for the participants to understand the surroundings and avoid obstacles.

References

1. D. Ametovicand et al., "NavCog: A Navigational Cognitive Assistant for the Blind," Association for Computer Machinery, Sept. 6, 2016.
2. S. Kayukawaand et al., "BBeep: A Sonic Collision Avoidance System for Blind Travellers and Nearby Pedestrians," Association for Computer Machinery, Paper No.: 52, May 2, 2019.
3. Open Ai, "GPT-4V(ision) System Card," Open AI, Sept. 2023. [Online]. Available: https://cdn.openai.com/papers/GPTV_System_Card.pdf.