

Article

A Smartphone-Based Mobility Assistant Using Depth Imaging for Visually Impaired and Blind

Aaron Raymond See , Bien Grenier Sasing and Welsey Daniel Advincula 

Department of Electrical Engineering, Southern Taiwan University of Science and Technology,
Tainan 71005, Taiwan; ma82b209@stust.edu.tw (B.G.S.); ma82b208@stust.edu.tw (W.D.A.)

* Correspondence: aaronsee@stust.edu.tw

Abstract: The means of assisting visually impaired and blind (VIB) people when travelling usually relies on other people. Assistive devices have been developed to assist in blind navigation, but many technologies require users to purchase more devices and they lack flexibility, thus making it inconvenient for VIB users. In this research, we made use of a mobile phone with a depth camera function for obstacle avoidance and object recognition. It includes a mobile application that is controlled using simple voice and gesture controls to assist in navigation. The proposed system gathers depth values from 23 coordinate points that are analyzed to determine whether an obstacle is present in the head area, torso area, or ground area, or is a full body obstacle. In order to provide a reliable warning system, the research detects outdoor objects within a distance of 1.6 m. Subsequently, the object detection function includes a unique interactable feature that enables interaction with the user and the device in finding indoor objects by providing an audio and vibration feedback, and users were able to locate their desired objects more than 80% of the time. In conclusion, a flexible and portable system was developed using a depth camera-enabled mobile phone for use in obstacle detection without the need to purchase additional hardware devices.

Keywords: visually impaired and blind (VIB); mobile application; depth imaging analysis; obstacle detection; object detection



Citation: 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>

Academic Editor: Liang-Bi Chen

Received: 7 February 2022

Accepted: 3 March 2022

Published: 9 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The number of visually impaired and blind (VIB) people has been increasing at an alarming rate with the World Health Organization (WHO) highlighting this concern. For the next three decades it is projected that individuals with moderate to severe visual impairment will rise to more than 550 million people up from approximately 200 million individuals in 2020 [1,2]. There are numerous causes for the increase in visual impairment and blindness, which can be attributed from different sources or instances such as genetics, accidents, diseases, or the trending aging population from both developing and developed countries. The level of severity for visual impairment and blindness varies among genders, and an individual's and their country's economic income contributes to how they respond to such a crisis, whether it be through access to the latest research in vision correction or medication that could prevent or reduce the worst case scenarios of this disability beforehand [3,4].

Visually impaired and blind people are met with difficulties in their daily lives as they lack the ability to perceive visual information. This limits their capability to process their surroundings and interact with society, hindering their day-to-day activities, which decreases their quality of life (QoL) [5]; thereby, VIB people are in most need of assistance for their mobility constraints, which requires the adaptation of assistive devices and accessible infrastructure [6]. The evolution of assistive technologies for VIB people has gradually increased over the years, and electronic travel aids for the blind have been developed by researchers globally in a response to address their complications. However, using the

current assistive technologies comes with challenges, as the flexibility and the portability aspect remains as an issue due to hardware and usability limitations [7]. Traditional assistive devices such as eyeglasses, tactile symbols, magnifiers, walking canes, etc., are used by the visually impaired and blind as it allows them to get by with basic daily tasks. However, these devices can be further expanded by exploring and applying smart sensing technologies [8].

In this paper, the researchers developed an application that caters to both cognitive and spatial awareness using an embedded 3D depth camera on a smartphone, which analyses depth images to process and calculate distances for obstacle and object detection. An inclusive user interface is developed for VIB people to navigate with ease by using both gesture and a voice command feature.

2. Related Studies

The inconvenience of the current assistive devices for VIB people has always been a great interest of development for various researchers around the world, but some of these technologies are not portable, too technical, or impractical to use. Advancements in technology allows for the improvement of these limitations, and existing studies on different problem aspects in terms of spatial awareness, cognitive awareness, and inclusive user interface design are discussed in this section.

2.1. Spatial Awareness

Assistive devices are continuously evolving and expanding with the rapid development of technologies and the extensive research to improve the QoL of VIB people. Earlier researches on assistive technologies focused mainly on mobility and orientation, cognitive and context awareness, obstacle detection, etc. Navigation is a huge concern for the VIB as safety concerns arise with the lack of visual information. Obstacles, terrain, and overall environmental inconsistencies have to be taken into account when travelling from one destination to another [9]. Sighted people perceive spatial awareness through the means of direction, acquiring one's specific location and relative position [10]. However, VIB people lack the capability to do so, hence, assistive navigating technologies for spatial awareness have been developed. The following studies were conducted to alleviate these challenges.

A previous study by See et al. developed a personal wearable assistive device with a modular architecture that was based on a robot operating system that detects and warns about obstacles using the Intel RealSense camera for outdoor navigating scenarios [11]. A similar study by Fernandes et al. developed a multi-module navigational assistant for blind people that generates landmarks through the implementation of various points of interest, which adjusts based on the environment. Orientation and location are provided to the user by means of audio feedback and vibration actuators, allowing a non-intrusive and reliable navigation. However, the research is still in progress and actual testing is yet to be delivered [12]. Another approach utilized an eBox 2300 TM equipped with a USB camera, ultrasonic sensor, and a headphone for detecting obstacles of up to 300 cm through the implementation of ultrasound-based distance measurement. It also includes an additional human presence algorithm that is programmed to detect face, skin, and cloth within 120 cm [13]. The developed module has proven to have promising result as it detects obstacles with 95.45% accuracy, but it is mounted on a helmet as a sensor unit while carrying the eBox 2300 TM weighing at about 500 g, which compromises the portability and flexibility aspect. Research by Li et al. developed an application called ISANA on the Google Tango tablet by taking advantage of the embedded RGB-D camera that provides depth information, enabling navigation and a novel way-point path finding. It also has a smart cane that handles interaction from the tablet to the user, which outputs vibration feedback and tactile input whenever an obstacle is detected [14]. The developed application is a robust system that also implements semantic map construction and a multi-modal user interface; however, the support for the Google Tango tablet has already shut down back in 2018 because of a newer augmented reality system [15].

2.2. Cognitive Awareness

Visually impaired and blind people struggle to partake in education due to the lack of visual information, although specialized educational materials do exist for learning. Tactile materials such as braille books, audio books, screen readers, refreshable displays, and many more are used to teach in blind schools [16]. However, perceiving and acquiring common daily objects on their own is a limitation that makes them dependent on other people. Independence for the VIB person is a valuable aspect as it reduces social stigma such as overly helpful individuals and other misconceptions that society assumes [17]. Cognitive awareness allows one to be aware of the surrounding environment and enables interaction with objects by utilizing different senses and reasoning [18]. This is an asset that VIB people struggle to achieve on their own; however, with an adaptive assistive technology that focuses on cognitive solution, independence can be achieved.

Joshi et al. developed an efficient multi-object detection method trained on a deep learning model with a custom dataset on a YOLO v3 framework installed on a DSP processor with a camera and distance sensor that captures images from different angles and lighting conditions and achieves a real-time detection accuracy of 99.69% [19]. The developed system is a robust assistive detection device that provides broad capability with its integration of artificial intelligence, although system maintenance seems to be difficult as each object has to be manually established into the system. Research by Rahman et al. developed an automated object recognition through internet of things-enabled devices such as a Pi camera, GPS module, Raspberry Pi, accelerometer, and more. Objects are detected through the installed laser sensors with a single-shot-detector model, which are defined by different directions such as front, left, right, and ground [20]. The developed system has an accuracy of 99.31% in detecting objects and 98.43% in recognizing its type, although it is currently limited to five types of objects and its size and weight can be further reduced in future work.

2.3. Inclusive User Interface Design

The demand for the integration of inclusive usability on technologies has escalated as these gadgets have become a necessary day-to-day means of communications, labor, and entertainment for society. In addition, according to the WHO, the number of people with disabilities is expected to increase over the years due in part to an aging population and increase in chronic health conditions [21]. A means of adaptability can bridge the complications for disabled people in using existing technology through the use of accessibility features. The accessibility features have been around for a long time and are not necessarily meant only for disabled people. Accessibility features are a set of options that can be activated for the reason of convenience, preference, or a means of easier navigation for those who cannot use the smartphone in their default setup [22]. In an article by Kriti, she stated that an accessible design is to capitalize on ease of use for all levels of ability for a specific goal of inclusivity for all kinds of users [23]. Different forms of accessibility such as touch, visual, hearing, and speaking are available regardless of whether it is for a permanent, temporary, or situational solution. Examples of accessibility in touch can be the fingerprint scanner, assistive touch, or even the capability to adjust button sizes. When it comes to visuals, increasing contrast, reducing motion, adjusting a larger text, using a magnifier, inverting colors, turning on subtitles and captions are accessible options, and even a guided access is available for personal use. For hearing and speaking, accessibilities such as VoiceOver or Text-To-Speech are popular examples [24]. These features are mostly available for smartphones and other similar devices.

An inclusive user interface can be constructed through designing a unified interface that accommodates the VIB person's special needs, requiring extensive research, especially on the design of the user experience, to provide the best user experience and usability. Existing assistive mobile applications for the VIB include a study by Nayak et al. that provides a solution for difficulty in appointing schedules, writing emails, and SMS reading on a smartphone completely based on voice commands [25]. However, the usability of the

application is limited with it being only a full voice command-based navigation, creating a possibility for complications with pronunciation and audibility of the said commands. Hence, the system development needs rigorous planning that aligns the user's needs and the technology [26].

3. Materials and Methods

The objective of this research is to develop a portable and flexible navigational solution that caters to both cognitive and spatial awareness of VIB people, thus, a smartphone with a depth camera is used. The smartphone chosen for the research is the Samsung Galaxy A80 (Samsung Corporation in Seoul, South Korea) running on Android v11 Operating System with a Qualcomm Snapdragon 730 G processor, 8 GB RAM, 6.7" infinity display, 48 MP + 8 MP rotating rear camera that supports 3D depth estimation with an embedded Time-of-Flight (ToF) sensor. The 3D depth camera has a pixel size of 10 μm with a field of view at 72° capable of calculating distances of up to 10 m away with a constant accuracy of <1%. The developed user interface for the mobile application is a combination of different accessibility features of a smartphone that helps users with vision impairment through the use of different gestures and voice commands. The main workflow of the system is shown in Figure 1 where the process starts off with specific voice commands to open the features. The grey area shows the 2 main components, which are the object detection and the obstacle detection features with an output of vibration and audio feedback.

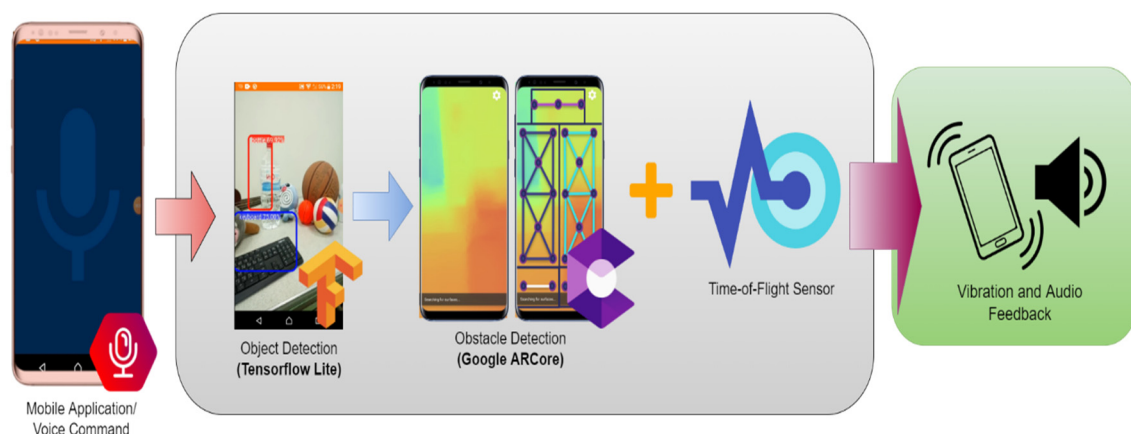


Figure 1. The proposed system workflow mobile application using voice command (**left**), mobile application interface and features (**middle**), and feedback to users (**right**).

The object detection feature utilizes TensorFlow Lite framework with a custom trained COCO SSD MobileNet v2 model and an additional unique interactable feature for VIB people to use. The obstacle detection takes advantage of the generated depth map on ARCore by Google that is overlaid with different coordinates on multiple directions to locate nearby obstacles. The output for both features will be a vibration and audio feedback. The application requires Android version 7 as the minimum SDK version. Figure 2 shows the proposed wearable system. The smartphone is inserted on a portable, foldable fabric sling pouch that is hung around the neck of the user.



Figure 2. Proposed wearable mobility assistive system by using Android smartphone (right) and sling pouch (left).

3.1. Obstacle Detection Integration

The obstacle detection module is achieved by 3D depth map generation of the ARCore Depth Lab API by Google, which is a software library that includes various depth-based UI/UX paradigms. It includes rendering of occlusion, shadows, collisions, physics, anchors, and other necessary augmented reality assets, which allow 3D objects to interact dynamically on real-life environment. Depth Lab consists of 39 geometry-aware AR features that range from entertainment purposes to educational purposes. We utilized its tracking functions and the generated depth information for analysis. Depth map generation only requires an RGB camera that is able to calculate distance of each pixel on the screen; however, ToF sensor can improve the performance. It starts by acquiring the raw depth map of the original image. The raw depth map is an unprocessed depth rendering that focuses on the accurate estimate of most pixels rather than processing speed, which results in aliased edges and obvious black spots. To prevent the black spots, a full depth map is generated, which understates some accuracy and calculates the depth estimation per pixel in order to accelerate the processing speed that covers the whole image. Bilinear filter is applied to smooth out the edges, but aliased parts are still visible. Hence, a depth-guided fast approximate anti-aliasing is proposed by the researchers to further clean up the aliased edges [27]. The depth map generation is visualized in Figure 3.

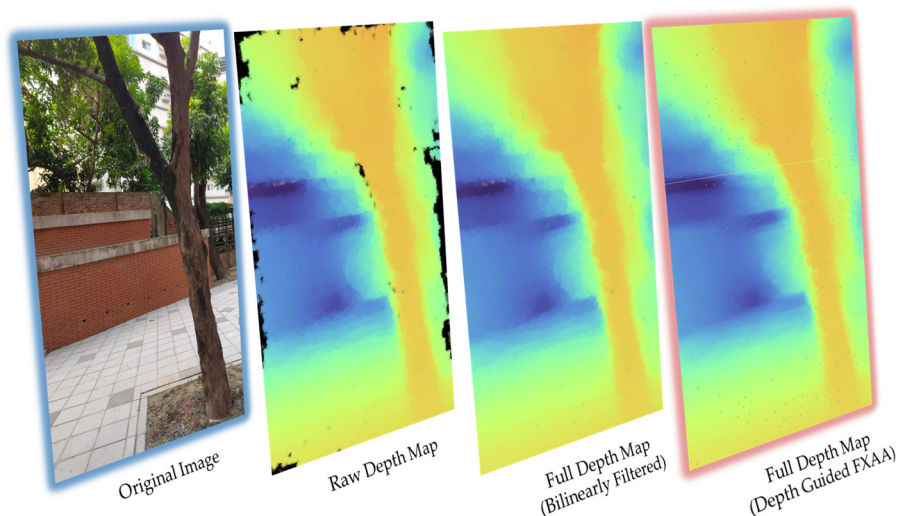


Figure 3. Processing visualization of full depth map generation starting from the original image to raw depth map to full depth map (bilinearly filtered), and, lastly, full depth map (depth-guided FXAA).

To extract distance from a depth image, a function by the ARCore is utilized which is the *getMillimetersDepth* function. Variables are required such as the initiated depth image and the desired x and y coordinates. Its concept states that each pixel contains distance value in millimeters parallel to the camera plane to which the aspect ratio of the depth image depends on the device [28]. In the case of the phone model used, an estimate of 260×120 resolution was used. Once distances are acquired, interconnected coordinates are overlaid on different locations as shown on the right picture in Figure 4 to cover most of the desired areas. A threshold of 1.6 m is set for the distance to reduce false detections and provide the optimal range for the users.

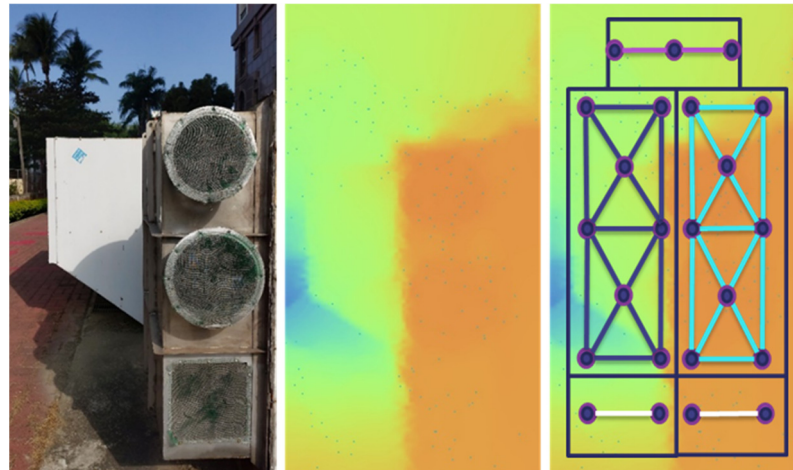


Figure 4. Original image (left), generated depth map at 1.6 m (middle), depth map with overlaid point coordinates on different locations (right).

The point coordinates work together in detecting an obstacle's distance. The number of point coordinates for each direction are listed in Table 1 and a single point in any direction is required to activate the warning. Furthermore, a combination of point coordinates activates warnings for full right, full left, full ground, and full body obstructions. The output for detecting obstacles is audio feedback that is based on the direction of the obstacle.

Table 1. List of regions where obstacles are detected, the number of point coordinates per direction and audio warning feedbacks.

Region Where Obstacles Are Detected	No. of Point Coordinates	Audio Warning Feedback
Head level	3	"Head Level"
Left torso	8	"Left Torso"
Right torso	8	"Right Torso"
Left ground	2	"Left Ground"
Right ground	2	"Right Ground"
Full right obstruction	(Right Torso + Right Ground) 10	"Full Right"
Full left obstruction	(Left Torso + Left Ground) 10	"Full Left"
Full ground obstruction	(Right Ground + Left Ground) 4	"Full Ground"
Full body obstruction	(All Directions) 23	"Full Body"

As previously stated, generating the depth map only requires a single moving RGB camera, although performance improvement can be achieved by initializing a dedicated depth sensing technology such as ToF sensor, which instantly provides depth map without calibrating the camera motion. ToF works by calculating the distance travel time of the light source emitted onto a certain object then bounces back to the camera as shown in Figure 5, formulating a mathematical equation as $\text{distance} = \frac{\text{Speed of Light} \times \text{Time}}{2}$ [29].

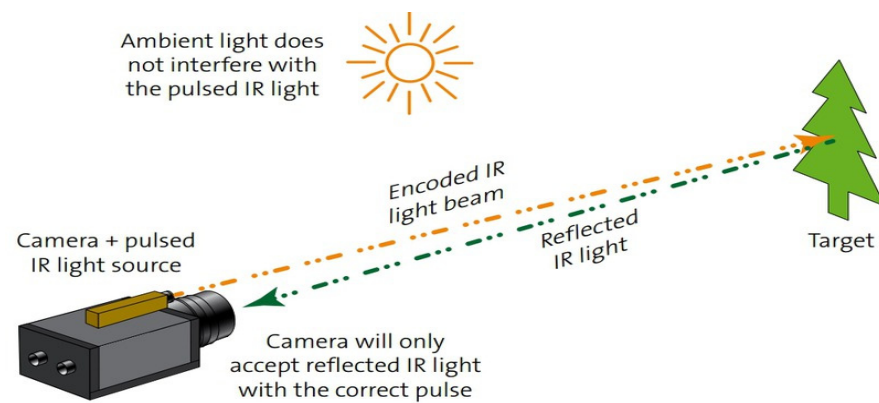


Figure 5. Time-of-Flight concept. Measuring travel time distance from IR light to the target and back to the camera [29].

3.2. Object Detection Integration

The object detection module utilizes the TensorFlow Lite framework, which is an on-device inference for different kinds of machine learning models typically applied on IoT devices. For the proposed system, a custom COCO SSD MobileNet v2 model is trained specifically to provide relevance for a VIB person's daily struggle of finding their desired objects, enabling over 90 different classes of objects such as walking cane, handbag, umbrella, tie, eyeglasses, hat, etc. It also includes street assets such as signage, traffic lights, bench, pedestrian lane, and many more for future outdoor usage implementation. Figure 6 shows the user interface for the voice command on the left image and the object detection result with confidence values on the right image.



Figure 6. Object detection module with the voice command interface (left) and sample object detection result (right).

A unique interactable feature is developed onto the object detection module. Once activated, it launches the camera, a text-to-speech prompt, and a tactile vibration function whenever the user's gesture is within the boundaries of the detected object. This feature will give VIB people the ability to select their desired objects and know its location that provides tactile feedback. Figure 7 shows the concept of the interactable feature where the edge of the bounding box is compared with x-axis and y-axis coordinates when a gesture is detected. In Table 2, a complete list of voice commands and touch gestures with corresponding functions and feedbacks are shown.

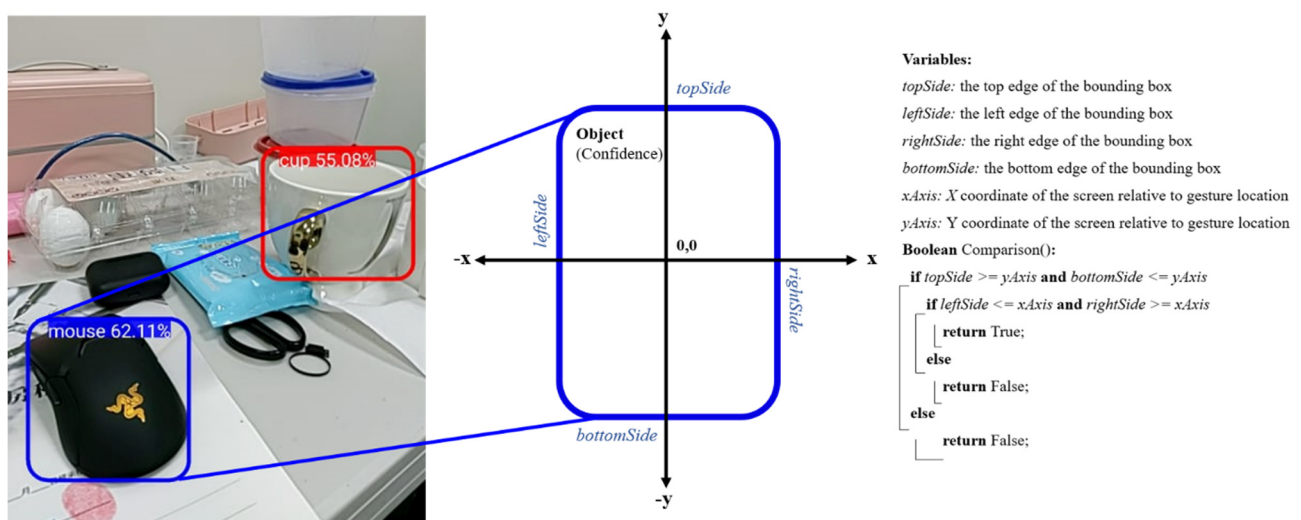


Figure 7. Object detection unique interactable feature that activates whenever gesture input is detected within the bounding box.

Table 2. List of different touch gestures and voice commands with its function and feedback.

Touch Gestures/Voice Commands	Function	Audio Warning Feedback
Single tap on main UI	Activate voice commands	<ul style="list-style-type: none"> • Beep sound effect after tapping • If command is wrong; speaks out “cannot recognize voice input. Please try again”
Voice command “open camera”	Opens the Camera	Speaks out “Opening Camera” then changes the Surface View of the screen to Camera1.
Voice command “object detection”	<ul style="list-style-type: none"> • Activates the object detection node on the view • Deactivates the obstacle detection node 	Speaks out “activating object detection”
Voice command “obstacle detection”	<ul style="list-style-type: none"> • Activates the obstacle detection node on the view • Deactivates the object detection node 	Speaks out “Activating Obstacle Detection”
Voice command “close camera”	Stops any camera activity	<ul style="list-style-type: none"> • Speaks out “closing camera” • Returns to the main UI

4. Results and Discussions

The concept of the study revolves around the integration of the obstacle and object detection in one application to provide a portable and flexible mobility assistive solution for VIB people in the future. In order to test the system functions, we tested the modules with different settings. A total of five blindfolded individuals aged 21 to 26 years old participated in the experiments. Due to the limitations during the pandemic, only limited participants were able to perform the experiments. In the future, further tests can be implemented.

4.1. Evaluation of the Hardware Device

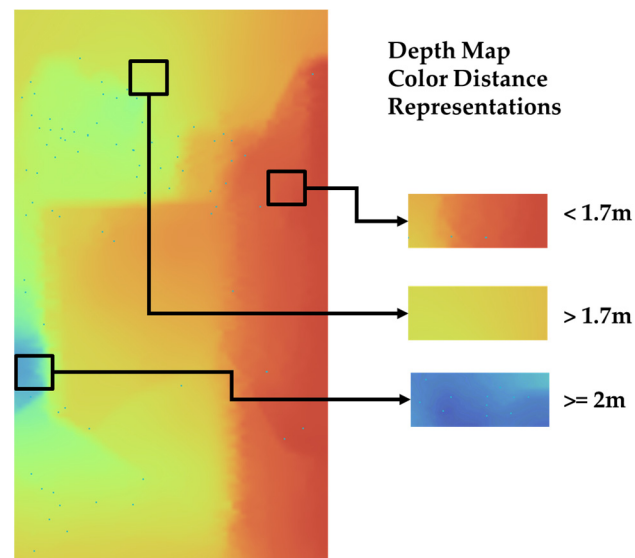
The Samsung Galaxy A80 is rated as a mid-end device that carries a 3700 mAh battery capacity, lasting for long hours when running the application, although depending on the activated connections the battery life duration might be reduced as seen in Table 3. The smartphone is equipped with a supercharging 25 W charger and cable that is convenient to use as it only takes few hours to fully charge the phone. Most of these connections such as Bluetooth, internet, and location services are only optional if other features within the phone are not used. The only requirement to run the application is permission to open the camera.

Table 3. Battery life duration evaluation based on activated functions while running the mobile application.

Activated Functions	Battery Life Duration
Camera Only	8 h 43 min
Camera + Bluetooth	7 h 17 min
Camera + Bluetooth + Internet	4 h 11 min
Camera + Bluetooth + Internet + Location Services	3 h 26 min

4.2. Evaluation of Obstacle Detection Module

The obstacle detection module calculates distances per pixel from 0 m to 8 m away via depth estimation and are represented in different colors as shown in Figure 8. Red and orange areas mean closer to the camera with a range of less than 1.7 m, yellow and green areas are greater than 1.7 m, and, lastly, blue areas are greater or equal to 2 m.

**Figure 8.** Depth map color distance representations with red distance as <1.7 m, yellow or green > 1.7 m, and blue ≥ 2 m.

While the application is capable of detecting distances up to 8 m, a distance of less than 2 m is considered and tested as it has been observed that a distance of greater than 2 m is not practical for detection and evasion for VIB people and creates complications in navigating. In Figure 9, depth image distances of <1.7 m, >1.7 m, and ≥ 2 m are displayed and evaluated. It has been observed that at a distance of 1.6 m, the detection capability is strong and was confirmed as an optimal range for the VIB to detect the obstacle as the walking cane was just centimeters away from the object, allowing confirmation of the detected obstacle at a safe range.

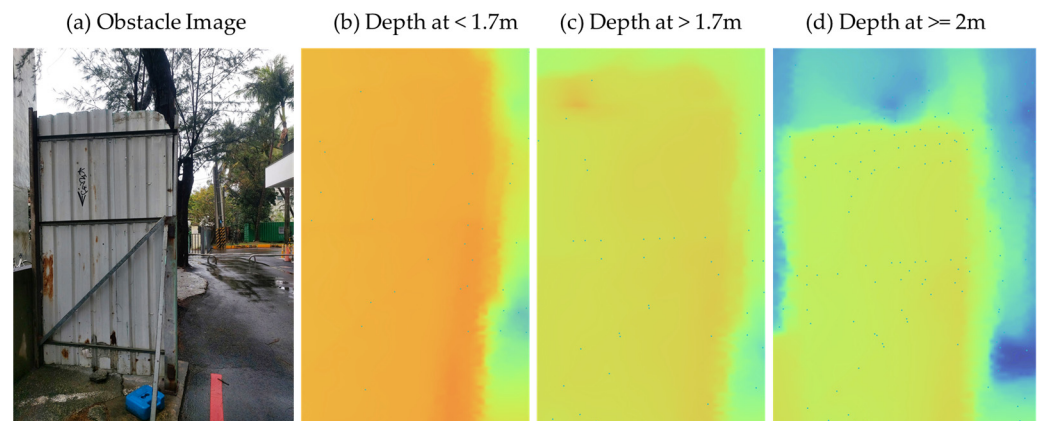


Figure 9. Depth map comparison of fan obstacle at different distances.

The smartphone is placed with the backside facing outside and placed vertically inside a sling pouch. During the testing, the participants were blindfolded and given a walking cane as it is recommended to use this together with the application to determine their direction and the terrain of the environment safely as seen in Figure 10. The obstacle course had some natural obstacles such as light posts, elevated areas, trees, and others. We also included randomly placed obstacles along the way to further test the system.



Figure 10. Blindfolded participant (left) with walking cane walking through the obstacle course (right) with randomly placed obstacles along the way.

The obstacle detection test results are listed in Table 4. This includes the finish time, the number of obstacles successfully detected and evaded, and the number of obstacles that were not detected by the application. The five participants were asked to navigate at their own pace. It was observed that when the person was walking too fast, the process time of the application could not keep up with the user's quick movement as some parts of the area take time to focus and update the distance calculation.

Table 4. Obstacle detection evaluation with finish score, successfully detected and evaded, and detection fails of the application.

Obstacle Detection Test Results	Finish Time	No. of Obstacles Successfully Detected and Evaded	No. of Obstacles the APP Failed to Detect
Participant 1	4:57 min	8	2
Participant 2	6:53 min	7	4
Participant 3	4:08 min	3	6
Participant 4	5:58 min	6	3
Participant 5	5:15 min	7	3

4.3. Evaluation of Object Detection Module

The object detection experiment took place indoors where selected common objects were placed randomly on the table for the blindfolded participants to find by using the unique interactable feature. Different angles, lighting, shapes, and other variables of an object were observed during the testing, which affected the detection capability, as seen in Table 5 where the average confidence score and the successful grab counts for each object are listed. The confidence score and the grab count for the bottle and cup scored the highest as its position and its overall visibility was consistent throughout the testing. It was followed by the mouse and AC remote, which scored lower as it is quite a flat object, resulting in a misinterpreted object. The scissors were observed to score the lowest with an average confidence score of 65.23%, and it was observed that aside from being flat on the table the metal part also reflected light resulting in a complication during detection. This limitation can be fixed by training a more robust object detection model that can consider the different variables of an image such as angle and lighting.

Table 5. Object detection evaluation with average confidence score and successful grab counts for each object.

Objects	Average Confidence Scores	Successful Grab Counts for Each Participant (Max of 5)
AC Remote	72.31%	4
Bottle	96.59%	5
Cup	92.01%	5
Scissors	65.23%	2
Mouse	86.72%	4

4.4. Usability Experience of the Mobile Application

After the experiment, the participants were asked to evaluate the experience in using the assistive application, and the survey results can be seen in Table 6. Feedback for the obstacle detection experiment include that the device was very lightweight and comfortable to use. The audio warning was helpful but some users found the looping audio warning annoying to listen to at times. The overall assistive capability of the obstacle detection was relatively safe, but caution and using the walking cane is always recommended as blind spots still exist with its detection range limitation. Feedback on the object detection states that it is a reliable feature, but it takes a few practices and learning to use effectively.

Table 6. Overall usability experience of the mobile application.

Survey Questions	Average Rating
<i>Obstacle Detection</i>	
1. I found the system comfortable to carry	4.8
2. I found the audio legible enough in a crowded place	4.2
3. I found the audio warnings helpful	4.8
4. I thought the loop for audio warnings were annoying	3
5. I thought the overall application is safe enough for blind navigation	4
<i>Object Detection</i>	
1. It was easy to navigate the app	4.2
2. I thought the voice command was useful	5
3. I found the gestures were responsive	4.4
4. I thought looking for objects with this app was easy	3
5. I found the overall application needs a high learning curve to use	3.6

Note: Each item is rated as 1 = Strongly Disagree, 5 = Strongly agree.

During the development of the system, a spectrum of vision loss was extensively researched where it was found out that there is a difference in perceiving and adapting information depending on the severity of blindness [30]. For example, individuals who are born blind have no perception of colors and basic concept of shapes at first. Thus, they need more supervision as compared to those whose vision was gradually lost or worsened over time as they already have the basic concepts in mind. Numerous researches have explored different approaches of technologies and methods to achieve an inclusive navigational aid, such as the previously mentioned related works. Most of the mentioned studies require all sorts of sensors, wiring, or external mechanisms in order to function well to its purpose. While it does not mean a drawback on the system, it is just inconvenient to setup at times as it requires expertise or comprehension of the developed system, whereas utilizing a depth sensing application with an inclusive user interface on a smartphone only is a more accessible option. Smartphone-based navigational aids have already existed for a long time, with examples such as an obstacle detection and classification system using points of interest [31], obstacle collision detection using emitted acoustic signal beep from a microphone on a smartphone [32], obstacle avoidance using smartphone camera and emitted laser light triangulation [33], and many more. These studies provide promising results as a navigational aid to VIB people; however, these implementations have inconsistencies due to hardware and software limitations. The current improvement means that using a mobile phone with a 3D depth camera provides portability and does not require additional costs to purchase expensive hardware and burden the user with extra devices.

5. Limitations

The current navigational capability of the mobile application is limited to locating and calculating any upcoming solid material as a general obstacle only, it is currently incapable of identifying the name and the type of the upcoming object. Thus, a synchronous implementation of both object detection and obstacle detection can be integrated for future works. The object detection is currently limited to the number of trained objects and a more robust deep learning model can be trained to include necessary objects that VIB people should recognize. The application is developed as an open system that can integrate new features quite easily as the developer wishes, this allows for easier access for future fixes and improvements.

6. Conclusions

A mobile application for mobility assistance was successfully developed that provides users with a portable and flexible navigational assistive device featuring obstacle detection and object detection in a single application. It can be controlled via gestures and a voice command-enabled user interface. The obstacle detection uses depth value from multiple

coordinates to generate an obstruction map and warns users of imminent danger. It was also found that the detection distance chosen was sufficient to warn the users and the object detection is capable of detecting relevant objects to assist VIB people in their daily lives. Furthermore, expansion of the platform is possible because of the versatile implementation of the Android system, especially within the software and hardware components. As smartphones continue to evolve overtime, the platform's ability to assist people with disabilities will be improved further, even possibly leading to a new generation of assistive devices. The platform can also be further improved with the use of different technologies such as stereo imaging [34], radar [35], and LiDAR [36], but some of these are expensive or will lead to a more complicated implementation.

Author Contributions: Conceptualization, A.R.S.; methodology, A.R.S. and B.G.S.; software, B.G.S.; validation, W.D.A., B.G.S. and A.R.S.; formal analysis, A.R.S. and B.G.S.; investigation, B.G.S.; data curation, B.G.S. and W.D.A.; writing—original draft preparation, A.R.S., B.G.S. and W.D.A.; writing—review and editing, A.R.S.; supervision, A.R.S.; project administration, A.R.S.; funding acquisition, A.R.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the Ministry of Science and Technology (MOST), R.O.C. Taiwan, under Grant MOST 109-2222-E-218-001-MY2, and the Ministry of Education, R.O.C. Taiwan, under grant MOE 1300-108P097.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We would like to thank the National Taichung Special Education School for the Visually Impaired, Resources for the Blind in the Philippines and Tainan City You Ming Visually Impaired Association that gave us insights and inspiration for the project.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. World Health Organization. Blindness and Vision Impairment. Available online: <https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment> (accessed on 6 January 2022).
2. Steinmetz, J.D.; Bourne, R.R.; Briant, P.S.; Flaxman, S.R.; Taylor, H.R.; Jonas, J.B.; Abdoli, A.A.; Abrha, W.A.; Abualhasan, A.; Abu-Gharbieh, E.G. 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. *Lancet Glob. Health* **2021**, *9*, e144–e160. [CrossRef]
3. United Nations Conference on Trade and Development. *The Impact of Rapid Technological Change on Sustainable Development*; United Nations: New York, NY, USA, 2020.
4. Wang, W.; Yan, W.; Müller, A.; Keel, S.; He, M. Association of Socioeconomics with Prevalence of Visual Impairment and Blindness. *JAMA Ophthalmol.* **2017**, *135*, 1295–1302. [CrossRef] [PubMed]
5. Khorrami-Nejad, M.; Sarabandi, A.; Akbari, M.-R.; Askarizadeh, F. The impact of visual impairment on quality of life. *Med. Hypothesis Discov. Innov. Ophthalmol.* **2016**, *5*, 96. [PubMed]
6. Litman, T. *Evaluating Accessibility for Transportation Planning*; Victoria Transport Policy Institute: Victoria, CA, USA, 2007.
7. Hoenig, H.; Donald, H.; Taylor, J.; Sloan, F.A. Does Assistive Technology Substitute for Personal Assistance Among the Disabled Elderly? *Am. J. Public Health* **2003**, *93*, 330–337. [CrossRef]
8. Senjam, S.S. Assistive Technology for People with Visual Loss. *Off. Sci. J. Delhi Ophthalmol. Soc.* **2020**, *30*, 7–12. [CrossRef]
9. Pfeiffer, K. Stress experienced while travelling without sight. *Percept. Mot. Ski.* **1995**, *81*, 411–417. [CrossRef]
10. Cognifit. What Is Spatial Perception? Cognitive Ability. Available online: <https://www.cognifit.com/science/cognitive-skills/spatial-perception> (accessed on 8 January 2022).
11. See, A.R.; Advincula, W.; Costillas, L.V.; Carranza, K.A. Using the McCall Model to Evaluate ROS-based Personal Assistive Devices for the Visually Impaired. *J. South. Taiwan Univ. Sci. Technol.* **2016**, *6*, 10–25.
12. Fernandes, H.; Costa, P.; Filipe, V.; Hadjileontiadis, L.; Barroso, J. Stereo vision in blind navigation assistance. In Proceedings of the 2010 World Automation Congress, Kobe, Japan, 12–23 September 2010; pp. 1–6.
13. Kumar, A.; Patra, R.; Manjunatha, M.; Mukhopadhyay, J.; Majumdar, A.K. An electronic travel aid for navigation of visually impaired persons. In Proceedings of the 2011 Third International Conference on Communication Systems and Networks (COMSNETS 2011), Bangalore, India, 4–8 January 2011; pp. 1–5.
14. Li, B.; Munoz, J.P.; Rong, X.; Chen, Q.; Xiao, J.; Tian, Y.; Ardit, A.; Yousuf, M. Vision-based mobile indoor assistive navigation aid for blind people. *IEEE Trans. Mob. Comput.* **2018**, *18*, 702–714. [CrossRef] [PubMed]

15. Express, T.I. Google Kills Project Tango AR Platform, as Focus Shifts to ARCore. Available online: <https://indianexpress.com/article/technology/mobile-tabs/google-kills-project-tango-ar-project-as-focus-shifts-to-arcore/> (accessed on 8 January 2022).
16. Rony, M.R. Information Communication Technology to Support and Include Blind Students in a School for all An Interview Study of Teachers and Students' Experiences with Inclusion and ICT Support to Blind Students. Master's Thesis, University of Oslo, Oslo, Norway, 2017.
17. Brady, E.; Morris, M.R.; Zhong, Y.; White, S.; Bigham, J.P. Visual challenges in the everyday lives of blind people. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris, France, 27 April–3 March 2013; pp. 2117–2126.
18. lumencandela. What Is Cognition? Introduction to Psychology. Available online: <https://courses.lumenlearning.com/wmopen-psychology/chapter/what-is-cognition/> (accessed on 12 January 2022).
19. Joshi, R.C.; Yadav, S.; Dutta, M.K.; Travieso-Gonzalez, C.M. Efficient Multi-Object Detection and Smart Navigation Using Artificial Intelligence for Visually Impaired People. *Entropy* **2020**, *22*, 941. [CrossRef] [PubMed]
20. Rahman, M.A.; Sadi, M.S. IoT Enabled Automated Object Recognition for the Visually Impaired. *Comput. Methods Prog. Biomed. Update* **2021**, *1*, 100015. [CrossRef]
21. World Health Organization. Disability and Health. Available online: <https://www.who.int/news-room/fact-sheets/detail/disability-and-health> (accessed on 12 January 2022).
22. Newell, A.F.; Gregor, P. Design for older and disabled people—where do we go from here? *Univers. Access Inf. Soc.* **2002**, *2*, 3–7. [CrossRef]
23. Krishan, K. Accessibility in UX: The Case for Radical Empathy. Available online: <https://uxmag.com/articles/accessibility-in-ux-the-case-for-radical-empathy> (accessed on 12 January 2022).
24. Nelson, B. 25 Smartphone Accessibility Settings You Need to Know about. Available online: <https://www.rd.com/article/accessibility-settings/> (accessed on 17 January 2022).
25. Nayak, S.; Chandrakala, C. Assistive mobile application for visually impaired people. *Int. J. Interact. Mob. Technol.* **2020**, *14*, 52–69. [CrossRef]
26. Patnayakuni, R.; Rai, A.; Tiwana, A. Systems development process improvement: A knowledge integration perspective. *IEEE Trans. Eng. Manag.* **2007**, *54*, 286–300. [CrossRef]
27. Du, R.; Turner, E.; Dzitsiuk, M.; Prasso, L.; Duarte, I.; Dourgarian, J.; Afonso, J.; Pascoal, J.; Gladstone, J.; Cruces, N. DepthLab: Real-Time 3D Interaction with Depth Maps for Mobile Augmented Reality. In Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology, Virtual, 20–23 October 2020; pp. 829–843.
28. ARCore. Use Depth in Your Android App: Extract Distance from a Depth Image. Available online: https://developers.google.com/ar/develop/java/depth/developer-guide#extract_distance_from_a_depth_image (accessed on 17 January 2022).
29. Imaging, S. 3D Time of Flight Cameras. Available online: <https://www.stemmer-imaging.com/en-ie/knowledge-base/cameras-3d-time-of-flight-cameras/> (accessed on 17 January 2022).
30. Perkins School for the Blind. Four Prevalent, Different Types of Blindness. Available online: <https://www.perkins.org/four-prevalent-different-types-of-blindness/> (accessed on 17 January 2022).
31. Tapu, R.; Mocanu, B.; Bursuc, A.; Zaharia, T. A smartphone-based obstacle detection and classification system for assisting visually impaired people. In Proceedings of the IEEE International Conference on Computer Vision Workshops, Sydney, Australia, 2–8 December 2013; pp. 444–451.
32. Wang, Z.; Tan, S.; Zhang, L.; Yang, J. ObstacleWatch: Acoustic-based obstacle collision detection for pedestrian using smartphone. *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.* **2018**, *2*, 1–22. [CrossRef]
33. Saffoury, R.; Blank, P.; Sessner, J.; Groh, B.H.; Martindale, C.F.; Dorschky, E.; Franke, J.; Eskofier, B.M. Blind path obstacle detector using smartphone camera and line laser emitter. In Proceedings of the 2016 1st International Conference on Technology and Innovation in Sports, Health and Wellbeing (TISHW), Vila Real, Portugal, 1–3 December 2016; pp. 1–7.
34. Kazmi, W.; Foix, S.; Alenyà, G.; Andersen, H.J. Indoor and outdoor depth imaging of leaves with time-of-flight and stereo vision sensors: Analysis and comparison. *ISPRS J. Photogramm. Remote Sens.* **2014**, *88*, 128–146. [CrossRef]
35. Nobis, F.; Geisslinger, M.; Weber, M.; Betz, J.; Lienkamp, M. A deep learning-based radar and camera sensor fusion architecture for object detection. In Proceedings of the 2019 Sensor Data Fusion: Trends, Solutions, Applications (SDF), Bonn, Germany, 15–17 October 2019; pp. 1–7.
36. Ton, C.; Omar, A.; Szedenko, V.; Tran, V.H.; Aftab, A.; Perla, F.; Bernstein, M.J.; Yang, Y. LIDAR assist spatial sensing for the visually impaired and performance analysis. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2018**, *26*, 1727–1734. [CrossRef] [PubMed]