使用 McCall 模型評估視障者基於 ROS 的個人輔助設備

[*1] 施金波, [1]Welsey Daniel C. Advincula, [1]Leonheart Van M. Costillas, [1,2]Karmelo Antonio Lazaro Carranza

[1] 南臺科技大學電機工程系、[2] 菲律賓德拉薩爾大學製造工程與管理系

*aaronsee@stust.edu.tw

摘要

個人輔助設備對於視障人士和盲人（VIB）在提供安全機動性和包容性方面都非常重要。許多研究和商用輔助設備都是針對特定任務或針對特定環境開發的。但是，在發展中國家，這些特定於任務和價格昂貴的設備是無法實現的。因此，該研究旨在為基於機器人操作系統（ROS）的模塊化體系結構上的視障人士和盲人（VIB）平台開發個人輔助設備（PAD）。這將使本地開發人員能夠快速開發可定制以供本地社區使用的輔助設備。該研究將使用 McCall 模型評估輔助設備中軟件的質量。隨後，使用 ROS 框架開發了 PAD，它可以執行對象檢測，讀取助手和點雲數據收集。然後，為 VIB 用戶開發了一個手機應用程式，VIB 的應用程序將用作 PAD 的主要界面，並進一步提供其他基於移動應用程序的系統功能，以增強用戶的安全性和移動性。該研究證明了 ROS 平台在開發量身定制的輔助設備中的潛力。

關鍵詞：視障人士和盲人、個人輔助設備、機器人操作系統、手機應用程式

Using the McCall Model to Evaluate ROS-based Personal Assistive Devices for the Visually Impaired

[*1]Aaron Raymond See, [1]Welsey Daniel C. Advincula, [1]Leonheart Van M. Costillas, [1,2]Karmelo Antonio Lazaro Carranza

[1]Department of Electrical Engineering, Southern Taiwan University of Science and Technology, Tainan, Taiwan

[2]Manufacturing Engineering and Management Department, De La Salle University, Manila, Philippines

Abstract

Personal assistive devices are important for the visually impaired and blind people (VIBs) in providing both safe mobility and inclusivity. Many research and commercially available assistive devices were task-specific or developed for a certain environment. However, in developing countries, these task-specific and highly priced devices are unattainable. Therefore, the research aims to develop a personal assistive device (PAD) for the visually impaired and blind (VIB) platform on a robot operating system (ROS) based modular architecture. This will allow local developers to quickly develop assistive devices that can be customized for use in local communities. The research will use the McCall model to evaluate the software quality use in assistive devices. Subsequently, a PAD was developed using the ROS framework that can perform object detection, reading assistant and point cloud data gathering. Subsequently, a mobile application was developed for the VIB user. The application for the VIB would be used as the main interface for the PAD and further provide additional mobile app-based system features to enhance the safety and mobility of the user. The research demonstrated the potential of the ROS platform in developing tailor-fit assistive devices.

Keywords: Visually Impaired and Blind (VIB), Personal Assistive Device (PAD), Robot Operating System (ROS), Mobile Application

Received: Dec. 25, 2020; first revised: Jan. 29, 2021, accepted: Feb. 2021.

Corresponding author: A.R. See, Department of Electrical Engineering, Southern Taiwan University of Science and Technology, Tainan 710031, Taiwan.

I. Introduction

International organizations, including the World Health Organization (WHO), have highlighted the growing number of visually disabled and blind people (VIBs). The cause of the increase can be attributed to a range from accidents, by birth or the trending ageing population in both developing and developed countries. It is projected that by 2050, around 550 million people would have mild to severe visual impairment in relation to the current population of 200 million people this year with this disability [1]. This alarming scenario is especially true in countries which do not have the facilities to cater to the disabled. Cities like Berlin [2], Milan [2], Denver [2], and Tokyo [3] are among the few exceptions that have the tools necessary to be considered accessible to people with disabilities. In those cities, there are facilities such as wheelchair ramps, VIB tactile pave ways, chirping sounds for an indication of crossing the streets, and handrails which are extremely useful for the VIBs. However, with the case of most other cities, it is not present. Hence, researchers and the industry are creating assistive technologies to mitigate the VIBs' struggles. There are various solutions currently attempting to tackle this problem, from traditional white canes used by the VIB as an extended arm [4-6], to wearable devices [7] and implants [8] that utilizes many different technologies to help the VIBs in terms of orientation and mobility, and also mobile applications to enhance their context awareness of the environment and objects around them. These technologies are helpful to the VIBs in a broad sense where the industry and the researchers would cater to most of the population. However, the VIBs start to struggle when they encounter difficulties in adapting to the new system. It is common for users to abandon assistive devices especially if it does not improve their quality of life to a significant level [8–9]. Even in developed devices with the target market, an individual may not have the same preference or understanding as the next person, hence it would be challenging for them to adapt on the same use case making them reluctant to use new technology [10–11].

Assistive technology has been gaining relevance in maintaining and even improving the quality of life of the VIBs, for most of them it is already impossible to attain independence without any form of assistive device. It has long been defined that this field of technology is the prime example of user-centric technology [12–13] because the way the user interacts with it greatly affects its effectiveness and usability. Usability of assistive devices can directly be affected by the change in functioning that is achieved through it. One factor is the quality of functioning which refers to the efficiency of achieving a certain task. Another factor is the quantity of functioning which refers to the number of tasks influenced by additional functions. However, most assistive devices are task-oriented and are sold separately and for a VIB individual there would be a need to purchase multiple assistive devices just for reading, safe mobility, education, and recreation [14–15]. Other than the high cost of having to buy multiple devices, there is also the burden of carrying multiple devices for the daily living related functions needed for an independent individual. Novel assistive devices have the hurdle in being adopted by the target users on a large scale due to insufficient functionalities. Furthermore, different cities and countries have a unique culture, landscape and needs. Products currently available in the market may be useful for those living in Europe but may not work well for those living in Asia [16]. Thus, there is a need for a customizable and scalable system to allow local developers and even academic institutions to assist in the development of these assistive devices. However, different developers may be accustomed to various programming languages and frameworks thus, in recent years, the Robot Operating System (ROS) has gained popularity because it serves as a middleware providing a modular approach in developing solutions for both the industry [17] and academia [18]. Therefore, in this article, the use of ROS in developing personal assistive devices (PAD) for the VIBs will be presented as a platform for developers in developing countries to make projects for the VIBs in their community and improve their quality of life through safe mobility and inclusivity.

II. Related Studies

The assistive device industry has been a field of interest of multiple researches in the world. Other researchers have developed devices that can solve problems of people with disabilities and this includes VIBs. There are existing devices that tackle their mobility and orientation, cognitive awareness, and context awareness problems. These existing studies are discussed in Sections 1 to 3. A proposed approach for development of assistive devices would be using the ROS framework to enable easy integration of different features to a singular system. This approach is mainly used in the fields of self-driving cars and robot navigational systems to enable scalability and upgradability. The existing research using ROS framework is discussed in Section 4.

1. Mobility and Orientation

In a study by Elmannai, et al., they tackled the different problems of mobility (such as navigation, safety, awareness, etc.) of the VIBs. The researchers grouped the devices into electronic travel aids (ETAs), electronic orientation aids (EOAs), and position locator devices (PLDs). The ETAs are for determining information about the environment using sensors to determine whether there are obstacles or obstructions in front of the user and provide for essential direction instructions to the VIBs. EOAs are for providing directions and navigation towards unknown places to the VIBs. They must determine the best path for the VIB to take in accordance with the current location of the user and provide in depth instructions to guide the user within the environment. In terms of functionality according to the VIBs, the devices must provide with several features such as clear information or indicators within a short timeframe (seconds), consistent performance within different daytimes (night or day), capable in either indoor or outdoor scenarios, detects obstacles within 5 meters, and detects static or dynamic objects. It is shown in their evaluation that there is no single system that provided satisfactory results in 100% of the criteria set by the VIBs in terms of most needed features. Not only did all of them have each own special feature, but none of them supported all the features deemed necessary. One of the main problems seen by the Elmannai, et al. was that researchers of assistive devices were trying to add a novel feature to their device before nailing in the fundamental features of their device [19]. In a study by Mocanu, et al., they utilized ultrasonic sensors and computer vision to perform a more efficient obstacle detection by gathering information about the environment, interpret the information, and inform the users about the obstacles within the environment. This implementation uses an auditory feedback interface for the user while being hands-free and ears-free due to the device being a belt attachment and the audio interface to be using bone conducting earphones. The device shown that the obstacles missed by each of the sensing platform is accommodated by the other. The main requirement the VIBs recommended to the researchers was a device should be hands-free and ears-free to be considered useful in terms of mobility [20]. Other researches also focused on the development of ETAs [21–22], EOAs [23–25], and PLDs [26–28].

2. Cognitive Awareness

In a study by Yang et al., VIBs have a significant problem in trying to understand, identify, and locate objects in the space around them. One main problem within this is about recognizing the banknotes they use for payment in shops. In a currency, there are multiple bill and coin denominations that can prove confusing to the VIBs, especially the bills wherein they are mostly the same sizes. Usually, VIBs either fold bills in a different way to designate its value or they try to keep it in different pockets. Yang, et al. developed a computer vision-based application that helps VIBs to recognize U.S. banknotes on an Android platform. The application is composed of a camera module, a banknote recognition algorithm, a database, a text-to-speech engine, and a user interface. The banknote recognition algorithm used the Speeded-Up Robust Features (SURF) algorithm and interest point detection, which are stable, reliable, and accurate algorithms even in different lighting conditions, angles, and environments. Their testing indicated the robustness of the system in different rotation angles, backgrounds, distances, and illumination except notably in the long distances of about 12 inches from the camera [29]. In a paper by Rahman, et al., they developed an Android-based banknote verification system. This paper focuses on checking the legitimacy of the banknotes used where it is difficult to discern the difference between the genuine and counterfeit banknotes. This is a viable application that can be used by VIBs in their everyday lives. They utilized support vector machine on the statistical and textural roughness of the banknotes to determine the security features of the genuine banknotes [30]. Dunai, et al. [31] used a similar method to the research of Yang et al. [29] where they used the SURF method and interest point detection using the Fast-Hessian matrix which are matched to the information within the database [31].

Elgendy, et al. provided an overview of the different mobile assistive devices that can help VIBs read information in on shopping tags. The implementations discussed in this paper is divided into tag-based, computer vision-based, and hybrid systems. Tag-based systems do not read information on the products itself but use the database in which the information is stored by using the RFID or NFC tags embedded on the barcode labels. Some of these devices provides an auditory interface to provide verbal directions. Computer vision-based systems have different implementations as well where they could be quick response-based (QR-based) or non QR-based. QR-based systems rely on a database where the information is stored and can be tagged onto each product. Non QR-based shopping assistants use image processing and neural networks to determine what is written on the product labels similar to a reading assistant and optical character recognition (OCR) systems. Hybrid systems use computer vision techniques with tag systems to improve the accuracy of the results in terms of information and location of the user [32]. A reading assistant for the blind was developed by Sabab, et al. where they convert the visual information within the document into voice feedback for the VIBs to read electronic documents or e-books. It was implemented through a touch display where one would touch a line for the assistant to read and then convert into an auditory format [33].

3. Context Awareness

Zhong, et al. proposed a pedestrian signal system where the signal lights account for the information about the VIBs' position in determining the timing, acts as pushbutton using the touch screen interface to let the system know that they would like to cross the street, and informs the VIBs of the status through auditory cues using text-to-speech technology [34]. Another implementation of providing context awareness to the VIB is through the object detection where the user is told about the objects in the field of view of the camera using computer vision, image processing, and neural network interfaces. Tapu, et al. introduced the DEEP-SEE framework that detects, tracks, and recognizes objects in real-time. The framework is integrated into an assistive device to improve the context awareness of the VIBs where the device would provide acoustic feedback in terms of urgency of the obstacle based on its perceived distance from the camera. They limited the device to reduce the frequency rate as not to overwhelm the users with information. They propose to integrate additional functionalities to their system to help the VIBs in their everyday lives [35]. Lillo, et al. used image processing to provide useful information about a local environment from the Google Street View. Image segmentation and object recognition were the tools used to achieve navigation in new environments to enhance the knowledge about the static scene [36]. Hu, et al. introduced different auditory sonification methods to indicate different information about the scene. Using different scene detection algorithms such as super-pixel down sample, obstacle detection, and traversable area detection, the algorithm outputs different information, depth images, obstacle distribution, and pathway directions to indicate audio feedback to image, obstacle, and path sonification methods, respectively [37]. These researches provide information on VIB travel.

4. ROS Framework

The ROS framework is a common environment in the field of robotics where they aim to develop scalable applications. This enables different programs running simultaneously even using different programming languages while being easy to integrate into a single architecture because it uses a message-passing type of communication. A node publishes a language-neutral variable onto a topic for any other node within the architecture to use. This is the message-passing type of communication that allows ROS to be a multi-language architecture [38–39]. Gutierrez et al. developed a control system for autonomous road vehicles that is scalable and robust using the ROS framework. Some modules relevant to their study were available however this did not fully satisfy their requirements so further upgrades were needed. The researchers utilized existing packages of the "Navigation" stack as the input topics to their controller node. They were able to implement a package that can be easily integrated into different navigation ROS architectures such as robot mobility applications [40] Wilson et al. developed a floor surface profiling robot using low-cost sensors and the ROS framework. ROS was used as the framework for internal communication, computation, and control of the robot and since there are existing solutions to common problems within the community, it is easy to start and integrate without programming from scratch. They compared a 2D laser scanner and RGB-D to provide an accurate surface profiler for a better navigation application [41].

Although the assistive technology market is a mature field wherein there is an abundance of devices that can cater to a specific functionality requirement of the VIBs, there are still aspects in which they lack usability and adaptability. This is where the researchers would like to propose a system wherein, modularity in mind, can customize the functionality of a single device to add different hardware and software using the ROS framework without needing to start from a new device.

III. Materials and Methods

1. Designing a Modular Architecture for Personal Assistive Devices

It is evident in the related research that there is no existing device that can cater to all the different main problems of the VIBs. Most devices are single use case where they have to carry all those devices to be able to have safe navigation, cognitive awareness (e.g., checking bills using to pay or getting change), and context awareness (e.g., checking what objects are in their surroundings). The researchers want to propose a system wherein the VIBs can choose which functionalities are needed and add only these, but they are not limited in the future where an upgrade in terms of functionality is possible without needing an entirely new device. The researchers will be using the ROS framework to develop this system where in modules (or functionalities) can be plugged in or removed without affecting the whole system. This provides a system based on modularity where it would be easy to develop different functionalities without starting from scratch. The proposed system architecture is shown in figure 1. The ROS framework is composed of different nodes and topics. Nodes are modules used to compute for a certain value or gather information from a sensor and publishes the output values (e.g. RGB image, object position, etc.) into a topic in which other nodes can also use the values stored within each topic for other processes. The system developed has user defined features that were taken from the VIBs where mostly they needed safe mobility [19–20], object identification [35], and reading assistants [29–33]. Meanwhile, the other nodes developed were intended for backend processes of data gathering, computing, and communications (Bluetooth serial, sensor launcher). For further evaluation of the software, the researchers will follow the McCall software quality model. Although the ROS environment is already implemented with regards to this, it is worth reevaluating it in the perspective of developing an assistive technology. This model attempts to bridge the gap between developers and users through product revision, product operations, and product transitions [42].

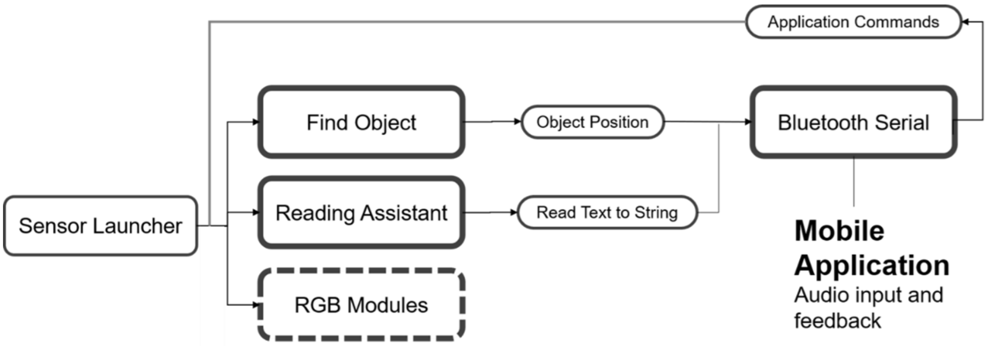

Fig.1 Proposed ROS node-based system architecture for modular development of personal assistive devices and seamless communication between embedded system and mobile applications.

2. System Configuration

The main board powering the sensors and containing the software developed was a Raspberry Pi 4 (RPi4). This provided the researchers with enough processing power to run the ROS architecture developed to control the different nodes. The RPi4 served as the hub for the instructions and data gathering, to be sent to a separate device for the computations needed within each node, in our case a desktop computer (running with an Intel i7-8700, 8GB of RAM, and a GeForce GT1030) which is always connected online and stationed in our laboratory. Data gathering for the purpose of the current modules require specific information such as RGB and depth images. The researchers utilized the Intel RealSense D435i as it provided all the necessary information (RGB sensors, depth sensors, inertial measurement unit, etc.) while being compact and portable. Figure 2 shows the prototype of the proposed system.

Fig. 2 Prototype of the Proposed System.

The main requirements of the researchers for the mobile application development platform would be its capability to maximize the hardware APIs of the mobile phone. It would utilize features such as voice commands, Text-to-Speech (TTS) and Speech-to-Text (STT) features, navigation, geolocation (GPS), and communication to the personal assistive device for the audio interface of the reading assistant and object detection feature of the whole system. The researchers favored Android (Google) over the iOS (Apple) platform because Android being an Open System platform allows developers to gain access to more features making it highly flexible and capable to higher degree of customization unlike the iOS platform, thus meets the main requirements mentioned above. Android Studio was used as the Integrated Development Environment (IDE) for it being Google's official Android IDE and known for its robustness and active advancement compared to other IDEs [43].

The modular architecture using the ROS framework will run on a Raspberry Pi 4 embedded system. The researchers made use of Ubiquity Robot Image that runs on Ubuntu 16.04 based image that already has ROS pre-installed. Most of the nodes were developed using Python language and made use of both ROS and Python 2.7 dependencies. Two object identification nodes would be developed for the system but due to the hardware limitations of a Raspberry Pi 4, only the node using TensorFlow Lite will be tested in it. However, the other node using Darknet will be implemented in a personal computer equipped with a CUDA enabled NVIDIA GeForce 1030 graphics card that will communicate wirelessly with the master node hosted in the Raspberry Pi 4. Subsequently, the system made use of the Intel RealSense D435i camera (Intel Corporation, California, U.S.A.) for its visual inputs. This camera enables the RGB stream used in the different modules of the device.

Following the concept of ROS nodes that can function independently or collaborate with other nodes, the system was able to implement the concept of "modules" and each module would correspond to a task-oriented function that would help the VIB user accomplish the predetermined task. In addition to these nodes that are feature specific for the end-user, there are back-end nodes that support them. In this setup, the researchers made use of the realsense2_camera node to launch the RealSense D435i camera and publish its streams to topics. Another back-end node used was the Bluetooth node, which sets up the Bluetooth server connection with the Mobile Application. This node is responsible for listening to task-oriented nodes or module commencement and sends the necessary code output back to the mobile application. The code output corresponds to audio feedback played using the mobile application. Thus, the focus of this research is to develop the ROS node-based modules and perform usability tests. The subsections are the identified task-oriented ROS node-based modules necessary for independent travels, basic orientation, mobility tasks, and reading.

(1) Object Identification Node

The researchers implemented an object identification node to enable context awareness on the part of the VIB. This would provide them context about the status of the surroundings and whether an object is present at that given instance of time. The project made use of the TensorFlow Lite neural network to be able to detect objects within the frame using the machine learning model on the graphical processing unit (GPU) of the remote device. TensorFlow Lite is a set of tools to help developers run TensorFlow models on mobile, embedded, and IoT devices. It enables on-device machine learning inference with low latency and a small binary size, in this research a pre-trained model (COCO SSD MobileNet v1) was used. This model can detect over 75 different everyday objects such as person, bicycle, car, motorcycle, airplane, bus, train, truck, tennis racket, bottle, wine glass, cup, fork, knife, spoon, bowl, banana, mouse, remote, keyboard, cell phone, microwave, etc. [44].

In this module, the node assists the user to identify and find objects in their surroundings. To use the node, the user must first issue a command which is sent through a voice input through the mobile application. Then, it triggers the RGB-D camera to start and initializes the image processing and filtering algorithms to be used by the TensorFlow Lite model in object identification. Lastly, the model searches and identifies the object if it is present in the frame or not. When the object is existing within the frame, the node then publishes necessary strings for the Bluetooth node to use in sending the audio cues to the audio output which is located within the mobile phone. These strings are specific audio feedbacks as shown in Table 1. The goal of the "find object" modules would be to center the object that is being located to be able to guide the user and locate it in front of the user. By providing a specific "object" stated through the mobile application, the object detection modules were able to single out the object within the environment and guide the user whether it is on the right or left of the frame. Furthermore, once the object is centered, an alert would be sent to provide notification that the object is directly in front of the user. The TensorFlow Lite has about 2 to 5 seconds frame latency when trying to display the stream, but while the stream is disabled it performs at a latency of less than a second. Figure 3 presents the object identification tests.

Table 1 Find Object Module Audio Feedback

Command	Audio Feedback
Find “Object”	The object is not in frame, keep moving
	Move left
	Move right
	Grab object in front

(a) (b) (c)

Fig. 3 The object detection was tested by a user with the PAD as shown in (a) and the image captured by the camera uses the object detection node with the Darknet Node and TensorFlow Lite to detect the objects within the image frame as shown in (b) and (c) respectively.

(2) Reading Assistant Node

The researchers developed the reading assistant node to help the VIBs to have cognitive awareness of the things around them especially labels, currency, and signs. This would give them the valuable information for their everyday tasks like shopping, paying, travelling, etc. The researchers used the LSTM-based Tesseract engine to enable text localization, text detection, and optical character recognition from the color image frame [45–46]. The node locates the text within the image frame and its bounding box is formed using the text detection algorithm and cropped to remove unnecessary information from the image frame. Finally, the optical character recognition algorithm determines the text within the image. The system takes advantage of this algorithm for the reading assistant function to help the VIB in understanding visual textual information in the environment. The reading assistant function outputs an audio feedback for the VIB to hear the word displayed within the frame.

(3) Mobile Application Development

The researchers developed a mobile application for the VIB and their PAD. The mobile application will provide extended features to the system and will also serve as an interface for the PAD. The mobile APP will communicate to the PAD via Bluetooth. The VIBs application was developed through the native Android SDK and coded in Java. For easy adaptation of the VIB user, a simplified workflow is implemented as exhibited in figure 4. The researchers made use of the Android hardware APIs and capabilities of the Android phone Samsung Galaxy A70 SM-A705MN/DS (Samsung, Seoul, South Korea) in performing tasks such as voice commands, text-to-speech features, maps, and geolocation.

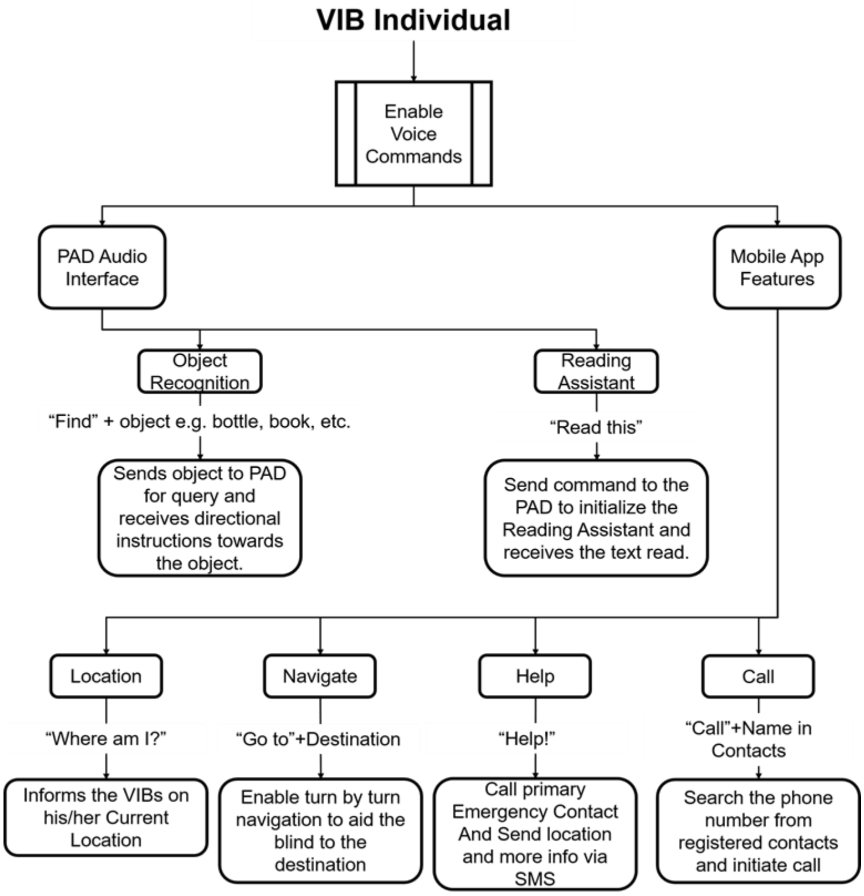

Fig. 4 VIB mobile application workflow starting from enabling voice commands for the personal assistive device and to the use of various functions such as location, navigation, help, and call.

a. Voice Recognition

It is difficult for machines to understand the natural speech patterns of human beings and without proper structuring, the meaning behind the words may be misinterpreted by machines. Natural language processing (NLP) deals with the challenges presented by non-uniform structures of speech by using statistical models or machine learning models in analyzing and predicting heard speech and responding to them [47]. The algorithm uses the utterances heard to be able to compare this to existing syllables, letters, or words in terms of the highest probability which uses a vast amount of data used to be able to be accurate. Using a NLP algorithm such as found in the Google API, it provides the VIBs a user-centric interface. This is used to enable the command functionality on what the device should do and the instructions relaying capabilities on how the user must react based on what the data sees [48].

SpeechRecognizer API, which is used in the VIB mobile application to trigger specific commands and perform operations based on the detected voice command. The SpeechRecognizer API supports multiple languages; however, the development and implementation of the mobile application are mainly focused on the English language. Thus, the voice recognition component was set to recognize the English language inputs. Operations such as PAD interface, location query, turn by turn navigation assistance, call initiation, and emergency assistance are the main functions of the VIB APP and each of these features are activated through a specific voice command.

b. Text-to-Speech

The VIB mobile application utilizes text-to-speech in any cases where it is needed to provide audio feedback to the VIB person [49]. There are two fundamental components for a typical text-to-speech model, text analysis and speech synthesis. On the text analysis, the raw text containing symbols like numbers and abbreviations are converted into an equivalent written out words. This is often coined as text normalization or tokenization. Each word or token is assigned with a phonetic transcription, divides, and marks the text into prosodic units we normally know as phrases, clauses and sentences. Phonetic transcription is a representation of a sound in speech. This process is known as text-to-phoneme conversion. Then on the speech synthesis, the synthesizer converts the symbolic linguistic representation into sound and processes it to be intelligible to make the output sound easily understood and natural like that of human speech. Subsequently, with the use of Android development TextToSpeech API, the data strings sent via Bluetooth by the object detection and reading assistant node of the PAD to the mobile application are converted to sensible audio and played to the mobile phone's speaker or to a headphone where the phone is connected.

c. Mobile Application Design and Operation

The application was designed for ease of use thus it has only one page to serve as an audio interface and intermediates for the mobile phone and the PAD as shown in figure 5. With just a single tap anywhere on the user interface of the application, the user can then speak a specific voice command to initiate a certain application feature. After each voice command input, an audio feedback is initiated so that the VIB user will know the mobile application's action based on the voice command; whether the command is being confirmed or it was not recognized. When the command is unrecognized it will ask the user to input the voice command again. The VIB application was able to successfully perform each of the task-specific features initiated by its corresponding voice command. For the location query, the application was able to inform the user with his/her current location address through audio feedback. The turn-by-turn navigation was successfully called every time the user initiated the navigation feature of the application. The application also was able to initiate calls and send SMS messages each time the "call" and "help" features are triggered. For the PAD related feature like the reading assistant and object recognition; the APP interacts with the PAD for the necessary information on object detection. Codes are sent from the APP to the PAD, and vice versa which is then decoded into a specific audio feedback. In the initial testing, the voice recognition component of the application performs as expected with low recognition error rate even at noisy environments since Google's voice recognition API is already optimized to detect speech regardless of any background ambient noise. In addition, positioning the microphone closer to the user would further minimize error. Moreover, the performance of the Text-to-Speech component is desirable and can be easily understood. However, to effectively perform in places with ambient noises, adjusting the phone's speakers, or using headphones should be considered. The usability of the mobile application as interface to the PAD is further evaluated in the next section.

Fig. 5 The VIB mode of the APP is reflected as a single GUI for ease of use

IV. Results and Discussions

1. ROS-Based Modular Architecture Evaluation

The focus of research is to develop a setup to allow modular architectures for personal assistive devices using the ROS platform. The researchers made use of the McCall quality model to evaluate the said implementation. It is a guide for the developers to have a perspective of the users' as well as their own to better improve the quality of the system. Table 3 displays the itemized evaluation of product revision and product operation based on the McCall Model [42].

Table 3 Evaluating the ROS PAD Contents using the McCall Model [42]

Major Perspective	Quality Factors	Quality Criteria	ROS-based Setup
Product Revision	Maintainability	Simplicity	Yes
		Conciseness	Yes
		Self- Descriptiveness	Yes
		Modularity	Yes
	Flexibility	Self-Descriptiveness	Yes
		Expandability	Yes
		Generability	Yes
	Testability	Simplicity	Yes
		Instrumentation	Yes
		Self-Descriptiveness	Yes
		Modularity	Yes
Product Operation	Correctness	Traceability	Yes
		Completeness	No
		Consistency	Yes
	Efficiency	Execution efficiency	Yes
		Storage efficiency	Yes
	Reliability	Consistency	No
		Accuracy	No
		Error tolerance	Yes
	Integrity	Access control	Yes
		Access audit	Yes
	Usability	Operability	Yes
		Training	Yes
		Communicativeness	No

(1) Product Revision

The ROS setup on the modules were implemented independently from each other and it allows for parallel development and troubleshooting. Whenever there are problems encountered in a specific node, troubleshooting is done on an isolated environment without affecting the other working nodes. With this, it would also be very easy to maintain the whole system through isolating a specific module that is encountering issues. The modular setup also allows easy improvement and change of specific nodes for revision. In fact, it is the design of architecture to allow more modules or remove modules from the setup. The different programs are already organized and separated from each other through packages. These factors in the setup proves that it is a maintainable, flexible, and testable setup [42, 50].

(2) Product Operation

Ease of use was considered through making an audio interface and throughout the design of the mobile application workflow. The PAD is relatively easy to use through saying the correct command to initiate operations. However, some of the nodes do not require any initiation as it is understood that it should always be operating. Overall, the researchers can say that the product is not self-descriptive. This is due to the non-visual interface which cannot instantly explain the features to the VIB. The personal assistive device was designed to have two specific feedbacks as mentioned, vibro-tactile and audio feedback. Hence, there is a need for training to intuitively understand the vibration configuration and what means, this was initially tested through a tutorial which explains the meaning of vibration with audio feedback. Subsequently, to also address the interpretation for vibration, the research also developed a mode where vibrations will be accompanied with audio feedback that is intended during the early stages of using it.

The multi process computing was handled well in the ROS setup and the storage and program execution can be considered efficient as all the nodes were just handled within the 1GB RAM. Although a lot could still be improved in trying to optimize the RAM usage of the setup. Subsequently, the accuracy of the object identification node can still be further improved. There were two main issues in the two methods of object identification. TensorFlow Lite can be handled by the Raspberry Pi 4 however it was not as accurate as using the DarkNet. Nonetheless, DarkNet was very accurate but needs a CUDA enabled GPU to have usable frames per second.

(3) Product Transition

The setup only makes use of catkin_make which is a tool to build and install a workspace or a ROS package containing the Python node developed. The package was tested in a different ROS environment in a computer and can be considered stable. In the case of change in the operating system, ROS was primarily tested on Ubuntu and Mac OS X systems and the ROS community has been contributing support for Fedora, Gentoo, Arch Linux and other Linux platforms. However, it is very important to consider hardware and its implications on the intended features when choosing a microprocessor. The ease of installation of the setup is an attribute particularly analyzed in implementing ROS as a middleware and in this setup also proves to be relatively easy.

The ROS framework is already a flexible architecture through the concept of nodes and topics. The setup in this research allows the end-user to adjust, mix and match different modules according to the setting, infrastructure in their place, their skills, age, etc. As mentioned in the previous section, this ROS environment has already a wide compatibility with a variety of operating systems and therefore can be considered reusable in different platforms. Lastly, the setup designed did not face any problem in communicating with other systems. This factor refers to the system's interoperability. Specifically, in this setup the assistive device is required to communicate with a mobile phone or a desktop computer. During the development stage, all visualization for the node processes was done through the desktop computer. On the operation perspective, the assistive device needs to communicate with the Android mobile application installed in a smartphone. This was easily achieved through developing a Bluetooth node which is implemented through Python. The ROS setup developed in this research allows for easy interface with other systems through an SSH or local network and a Bluetooth protocol connection.

VI. Conclusions

The lack of assistive devices that have flexible user cases is one of the barriers for usability and adoption. This could be attributed to highly specialized design of assistive devices which are very task specific. These assistive devices who have fixed designs also face this barrier in being implemented abroad because of the totally different context and settings in different places. Thus, there is a need for a customizable and scalable system to allow an extent of flexibility in the design that could potentially tailor-fit to a wider end user. A modular architecture for a personal assistive device would also allow local developers and even academic institutions to assist in the development of these assistive devices.

Aiming to fill this gap, the goal of this research is to be able to provide a personal assistive device which is able to cater to more than one task-specific function needed by the VIB. The researchers were able to achieve this through implementing a ROS-based modular architecture for the personal assistive device. Along with this, the research was able to achieve the following: 1. ROS node-based modules which are task-oriented for object identification, reading assistant, and photograph capture. 2. Mobile application for the VIB user which will serve as an audio interface with the personal assistive device and provides additional voice command-activated features to enhance the VIB's mobility and safety.

The researchers evaluated the implementation of the ROS-based system through the McCall Software Quality Model. In the product revision perspective, the setup was able to perform well because of its already modular nature; the different ROS node-based modules are easily manageable for troubleshooting, improvement, or maintenance. On the product operations perspective, which is the second perspective of the model, the setup also performed well even if the setup was not considered self-descriptive. Nonetheless, it was easy to use because it had a straightforward audio interface. The personal assistive device task-specific modules are already reliable even in outdoors or indoors situations any time of the day. Finally, evaluating the setup through the last perspective of the model, which is product transition, the ROS environment proves its benefits because of relatively easy installation in a completely different computer and it can be easily implemented in a variety of operating system and hardware specifications. Through implementing a Bluetooth node in the setup, it can easily communicate with a mobile application. In the future, wearable devices and assistive devices can be implemented in a cross-platform environment to advance the developments in developing countries.

References

[1] T. Mazumdar. (2017, Aug. 3). *Global blindness set to 'triple by 2050'.* BBC News. https://www.bbc.com/news/health-40806253

[2] D. Ruhm. (2016, Aug. 24). *6 of the world's most disability-friendly travel destinations*. Goodnet. https://www.goodnet.org/articles/6-most-disabilityfriendly-travel-destinations

[3] J. Grisdale. (2016, Apr. 26). *Accessible Japan: Can you visit Japan with a disability?* https://www.insidejapantours.com/blog/2016/04/26/accessible-japan/

[4] R. Bourne, S. Resnikoff, and P. Ackland. (n.d.). *GBVI - changes in numbers & prevalence.* IAPB Vision Atlas. http://atlas.iapb.org/global-burden-vision-impairment/gbvi-changes-numbers-prevalence/

[5] I. Attia and D. Asamoah. (2020). The white cane. Its effectiveness, challenges and suggestions for effective use: The case of Akropong School for the Blind. *J. educ. soc. behav. sci.*, 47–55. http://dx.doi.org/10.9734/JESBS/2020/v33i330211

[6] A.D.P. dos Santos, F.O. Medola, M.J. Cinelli, A.R.G. Ramirez, and F.E. Sandnes. (2020). Are electronic white canes better than traditional canes? A comparative study with blind and blindfolded participants. *Univers. Access Inf. Soc.,20*, 93–103.

[7] J.A. Hossack, J.S. Ha, and T.S. Sumanaweera. (2001, May 30). Quantitative free-hand 3D ultrasound imaging based on a modified 1D transducer array. In *Medical Imaging 2001: Ultrasonic Imaging and Signal Processing, 2001, vol. 4325: International Society for Optics and Photonics* (pp. 102–112). San Diego, CA, United States. https://doi.org/10.1117/12.428230

[8] A. Kusnyerik, M. Resch, H.J. Kiss, and J. Nemeth. (2018). Vision restoration with implants. In E. Pissaloux and R. Velazquez (Eds.), *Mobility of Visually Impaired People: Fundamentals and ICT Assistive Technologies* (pp. 617–630). Springer International Publishing.

[9] H.Y. Day, J. Jutai, W. Woolrich, and G. Strong. (2001). The stability of impact of assistive devices. *Disabil. Rehabil., 23*(9), 400–404.

[10] A.G. Money, J. Barnett, J. Kuljis, M.P. Craven, J.L. Martin, and T. Young. (2011). The role of the user within the medical device design and development process: medical device manufacturers' perspectives. *BMC Med. Inform. Decis. Mak., 11*(1), 15.

[11] P.J. Ogrodnik. (2019). *Medical device design: Innovation from concept to market*. Academic Press.

[12] N. Sachdeva and R. Suomi. (2013). Assistive technology for totally blind− barriers to adoption. *SOURCE IRIS: Selected Papers of the Information Systems Research Semina, 47*. https://www.mn.uio.no/ifi/english/research/news-and-events/events/conferences-and-seminars/iris2013/groups/iris36_submission_16.pdf

[13] A.T. Sugawara, V.D. Ramos, F.M. Alfieri, and L.R. Battistella. (2018). Abandonment of assistive products: assessing abandonment levels and factors that impact on it. *Disabil. Rehabil.: Assist. Technol., 13*(7), 716–723.

[14] S. Arthanat, S.M. Bauer, J.A. Lenker, S.M. Nochajski, and Y.W.B. Wu. (2007). Conceptualization and measurement of assistive technology usability. *Disabil. Rehabil.: Assist. Technol., 2*(4), 235–248.

[15] I. Asghar, S. Cang, and H. Yu. (2018). Usability evaluation of assistive technologies through qualitative research focusing on people with mild dementia. *Comput. Hum. Behav., 79*, 192–201.

[16] C. Lamour, C. De La Robertie, N. Towers, and H. Kotzab. (2016). Prescribed consumption and consumers' decision-making styles: A cross-cultural comparison between Europe and Asia. *International Journal of Retail & Distribution Management, 44*(3), 266–283.

[17] O. Stasse, T. Flayols, R. Budhiraja, K. Giraud-Esclasse, J Carpentier, J. Mirabel, A. Del Prete, P. Soueres, N. Mansard, F. Lamiraus, J.-P. Laumond, L. Marchionni, H. Tome and F. Ferro. (2017. Nov. 15–17). TALOS: A new humanoid research platform targeted for industrial applications. In *2017 IEEE-RAS 17th International Conference on Humanoid Robotics (Humanoids)*(pp. 689–695). Birmingham, UK.

[18] H.M. Do, C.J. Mouser, and W. Sheng. (2012, Feb. 17). Building a telepresence robot based on an open-source robot operating system and android. In *Third Conference on Theoretical and Applied Computer Science (TACS 2012)*. Stillwater, OK, United States.

[19] W. Elmannai and K. Elleithy. (2017). Sensor-based assistive devices for visually-impaired people: Current status, challenges, and future directions. *Sensors, 17*(3), 565.

[20] B. Mocanu, R. Tapu, and T. Zaharia. (2016). When ultrasonic sensors and computer vision join forces for efficient obstacle detection and recognition. *Sensors, 16*(11), 1807.

[21] S. Jin, M.U. Ahmed, J.W. Kim, Y.H. Kim, and P.K. Rhee. (2020). Combining obstacle avoidance and visual simultaneous localization and mapping for indoor navigation. *Symmetry, 12*(1), 119.

[22] M. Elgendy, T. Guzsvinecz, and C. Sik-Lanyi. (2019). Identification of markers in challenging conditions for people with visual impairment using convolutional neural network. *Appl. Sci.,* 9(23), 5110.

[23] D. Plikynas, A. Žvironas, A. Budrionis, and M. Gudauskis. (2020). Indoor navigation systems for visually impaired persons: Mapping the features of existing technologies to user needs. *Sensors, 20*(3), 636.

[24] X. Zhang, H. Zhang, L. Zhang, Y. Zhu, and F. Hu. (2019). Double-diamond model-based orientation guidance in wearable human–machine navigation systems for blind and visually impaired people. *Sensors, 19*(21), 4670.

[25] F.R. Amorim and F.L.d.P. Santil. (2018). Prototype of mobile device to contribute to urban mobility of visually impaired people. *Big Data Cogn. Comput., 2*(4), 38.

[26] R. Velazquez, E. Pissaloux, P. Rodrigo, M. Carrasco, N. I. Giannoccaro, and A. Lay-Ekuakille. (2018). An outdoor navigation system for blind pedestrians using GPS and tactile-foot feedback. *Appl. Sci., 8*(4), 578.

[27] W.C. Simões, G.S. Machado, A. Sales, M.M. de Lucena, N. Jazdi, and V.F. de Lucena. (2020). A review of technologies and techniques for indoor navigation systems for the visually impaired. *Sensors, 20*(14), 3935.

[28] X. Liu, S. Zhang, J. Zeng, and F. Fan. (2019). Analysis and optimization strategy of travel system for urban visually impaired people. *Sustainability, 11*(6), 1735.

[29] Z. Yang, J.M. Schafer, and A. Ganz. (2014). Reliable and user friendly US banknote recognition application for visually impaired users on android smartphones. *Int. J. E-Health Med. Commun. (IJEHMC), 5*(3), 1–16.

[30] U.U. Rahman, A.B. Sargano, and U.I. Bajwa. (2017). Android-based verification system for banknotes. *J. Imaging 3*(4), 54.

[31] L. Dunai Dunai, M. Chillarón Pérez, G. Peris-Fajarnés, and I. Lengua Lengua. (2017). Euro banknote recognition system for blind people. *Sensors, 17*(1), 184.

[32] M. Elgendy, C. Sik-Lanyi, and A. Kelemen. (2019). Making shopping easy for people with visual impairment using mobile assistive technologies. *Appl. Sci. 9*(6), 1061.

[33] S.A. Sabab and M.H. Ashmafee. (2016, Dec. 18–20). Blind reader: An intelligent assistant for blind. In *2016 19th International Conference on Computer and Information Technology (ICCIT)*(pp. 229–234). Dhaka, Bangladesh

[34] Z. Zhong and J. Lee. (2020). Virtual guide dog: Next-generation pedestrian signal for the visually impaired. *Adv. Mech. Eng., 12*(3), 1–9.

[35] R. Tapu, B. Mocanu, and T. Zaharia. (2017). DEEP-SEE: Joint object detection, tracking and recognition with application to visually impaired navigational assistance. *Sensors, 17*(11), 2473.

[36] A.D. Lillo, A. Daptardar, K. Thomas, J.A. Storer, and G. Motta. (2012). Compression-based tools for navigation with an image database. *Algorithms, 5*(1), 1–17.

[37] W. Hu, K. Wang, K. Yang, R. Cheng, Y. Ye, L. Sun and Z. Xu. (2020). A comparative study in real-time scene sonification for visually impaired people. *Sensors, 20*(11), 3222.

[38] M. Quigley, B. Gerkey, K. Conley, J. Faust, T. Foote, J. Leibs, E. Berger, R. Wheeler, and A. Ng. (2009). ROS: an open-source robot operating system. In *ICRA workshop on open source software, 3*(3.2), 5. Kobe, Japan.

[39] V.A. Hax, N.L. Duarte Filho, S.S. da Costa Botelho, and O.M. Mendizabal. (2013). ROS as a middleware to internet of things. *J. Appl. Comput. Res., 2*(2), 91–97.

[40] R. Gutiérrez, E. López -Guillen, L.M. Bergasa, R. Barea, O. Perez, C. Gomez-Huelamo, F. Arango, J. del Egido and J. López -Fernandez. (2020). A waypoint tracking controller for autonomous road vehicles using ROS framework. *Sensors, 20*(14), 4062.

[41] S. Wilson, J. Potgieter, and K.M. Arif. (2019). Robot-assisted floor surface profiling using low-cost sensors. *Remote Sens.*, *11*(22), 2626.

[42] R.E. Al-Qutaish. (2010). Quality models in software engineering literature: An analytical and comparative study. *Am. J. Sci.*, *6*(3), 166–175.

[43] @TheASOTool (2018). Mobile App Development: Differences Between iOS & Android.

[44] TensorFlow Lite guide. (n.d.). https://www.tensorflow.org/lite/guide

[45] R. Smith. (2007, Sep. 23–26). An overview of the Tesseract OCR engine. In *Ninth international conference on document analysis and recognition (ICDAR 2007)*, 2, 629–633. Curitiba, Brazil

[46] S. Thakare, A. Kamble, V. Thengne, and U. Kamble. (2018, Dec. 1–2). Document segmentation and language translation using tesseract-OCR. In *2018 IEEE 13th International Conference on Industrial and Information Systems (ICIIS)*(pp. 148–151). Rupnagar, India

[47] P. Kłosowski. (2018, Sep. 19–21). Deep learning for natural language processing and language modelling. In *2018 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*(pp. 223–228). Poznan, Poland

[48] G. López, L. Quesada, and L. A. Guerrero. (2018). Alexa vs. Siri vs. Cortana vs. Google Assistant: A comparison of speech-based natural user interfaces. In *Advances in Human Factors and Systems Interaction*, (pp. 241–250). Springer International Publishing,

[49] S. Kaur and K.S. Dhindsa. (2019). Design and development of android based mobile application for specially abled people. *Wirel. Pers. Commun.*,*111*, 2353-2367.

[50] P. Berander, L.-O. Damm, J. Eriksson, T. Gorschek, K. Henningsson, P. Jonsson, S. Kagstrom, D. Milicic, F. Martensson, K. Ronkko, P. Tomaszewski, and L. Lundberg. (2005). Software quality attributes and trade-offs. *Blekinge Institute of Technology, 97*(98), 19.