

GLYKE A Voice Command Assistive App for the Blind

Dennis C Tenerife* Glyzel Faith M Palcon, Keannumar S Aleta and Dennis J Lapong

*Department of Information Technology, College of Information Technology, Mindanao State University Buug, Buug, Zamboanga Sibugay, Philippines.

Received April 03, 2024; Revised April 08, 2024; Accepted April 11, 2024

ABSTRACT

Blindness is a condition where in a person cannot see a single thin-g even light itself and Visual Impairment is when a person's eyesight cannot see things as they were. Having this condition, limits them from doing certain activities specially when using smartphones. To help them, the developers made GLYKE a voice command application to supplement in some of their needs in information. This project's objective is to create an application where in the users will be able to know certain information with the use of their voice by adding a voice command function and text-to-speech for the application to be able to speak aloud the information needed. The features that this application have are being able to set alarms, asking about the current time, voice calculator, asking the temperature, asking the current battery percentage of the phone, using camera to classify object, and object detection to sense objects at close range.

Keywords: Voice command, Blind, Text-to-speech, Visually impaired

INTRODUCTION

Project Context

Blindness is an inability to see things which means a person cannot see anything even light and vision impairment is a condition in which a person's eyesight cannot see things as they used to be. People with this condition will have a difficult time coping with daily tasks or activities that demand vision since their eyes cannot see as clearly as they used to or their vision has entirely deteriorated [1]. Most individuals with this condition are left neglected since those who should be looking after them have tasks to do, leaving the affected person unattended.

Refractive error, cataract, age-related macular degeneration, glaucoma, diabetic retinopathy, corneal opacity, trachoma are the leading causes of visual impairment. Vision loss can affect a person's mobility and independence. It may also lead to injuries, falls, and even worsened status in mental health, social function, employment, cognition, and educational attainment.

There are over 284 million people who are visually impaired, with the number of 39 million blind people. It is possible to say that 60% of the world's population who are visually impaired can be cured and 20% could be prevented [2].

55 percent of people living with blindness were women, amounting to nearly 24 million women. 75 percent of the 19 million blind men were over the age of 50. This statistic

depicts the number of blind people worldwide in 2020, broken down by age group and gender [3].

Every blind individual face difficulty but through the easy availability of beneficial assistive devices, such as canes and electronic mobility aids, a blind individual's difficulties were lessened. When we look around, we could distinguish how visual the information is around us. Signs showing the correct route or a potential hazard are some examples of visual information that we see on a daily basis. The majority of this information is inaccessible to the visually impaired, limiting their freedom.

When it comes to carrying out a variety of tasks, technology proficiency is expected. Standard technologies are now more accessible to people with disabilities with the use of assistive technologies which are devices and software that enables people with disabilities to use technology.

"GLYKE" is a Voice Command-based and an assistive

Corresponding author: Dennis C Tenerife, Department of Information Technology, College of Information Technology, Mindanao State University Buug, Zamboanga Sibugay, Philippines, Tel: (062) 955-4952; E-mail: dennis.tenerife@msubuug.edu.ph

Citation: Tenerife DC, Palcon GFM, Aleta KS & Lapong DJ. (2024) GLYKE A Voice Command Assistive App for the Blind. Int J Clin Case Stud Rep, 6(2): 245-252.

Copyright: ©2024 Tenerife DC, Palcon GFM, Aleta KS & Lapong DJ. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

android application that is created for individuals who are visually impaired. It uses Voice Recognition to help them maneuver the application's features through the use of their voice.

Purpose and Description

The project is an application-based system which generally intends to help individuals who are visually impaired. It will likewise aid with blind individuals with the help of the features available in the app.

This application has various features namely: to set alarms, to ask the current time, to ask the current battery percentage, to calculate, to ask the temperature, to use camera to classify objects and object detection which will help the user.

Objective

The main goal of this project is to develop GLYKE App, a Voice Command Assistive App that will help visually impaired or blind individual to manage their time.

Scope and Limitations

GLYKE is a voice command app that helps blind or visually impaired individuals. The features of the app are: set alarms, asking about the current time, voice calculator, asking the temperature, asking the current battery percentage of the phone, using camera to classify object, and object detection to sense objects at close range.

Its limitations are the following:

- a. Cannot guarantee that the application will be free from bugs and errors.
- b. Cannot guarantee that the application is compatible with all devices.
- c. Cannot guarantee that the Voice Command will be able accurately to input and process the user's command.

Definition of Terms

The key terms used in this study have been defined for the sake of clarity.

Visual Impairment: When a person's eyesight cannot be corrected to a "normal" level, they have vision impairment.

Blind: Blind are people who are unable to see.

Assistive Devices: Assistive devices are a tool that helps a person with a disability to do a certain task.

Assistance: Assistance is the act of assisting; helping; aiding; supporting. Likewise giving assistance to blind and visually impaired individuals.

Proficiency: Proficiency is the fact of having the skill and experience for doing something. Having the skills and experience in accessing technology.

Voice recognition: Voice recognition is a software program or hardware device with the ability to decode the human voice, it is commonly used to operate devices, perform commands and write without having to use a keyboard, mouse, or press any button.

Individuals: Individuals is a people or things, especially when compared to the group or set in which they belong such as blind and visually impaired individuals.

Feature: Refers to the typical quality or an important part of the Voice Command Based Assistive App for the Blind.

REVIEW OF RELATED LITERATURE AND SYSTEMS

This chapter presents the Review of Related Literature of the application. This chapter discusses the Technical Background, Related studies, and Related Systems.

Technical Background

This research was based on multiple sources to supplement the knowledge and the needs of this study. With the progress of technology, numerous technical elements that may be utilized to aid blind and visually impaired individuals in a variety of scenarios have been developed. Development of capabilities that turn visible text into read-aloud speech, voice-based navigation, and many more. The following is a synopsis of the said feature.

Related Studies

Text-to-Speech

This function is a type of voice synthesis that translates any visible digital text into spoken speech output. Which is audible via the audio output. It is accessible via cellphones, desktops, and tablets. Which are already pre-programmed in every personal electronic device [4].

This was created to help those with visual impairments by using computer-generated speech to interpret the screen for the user to listen to. The way Text-to-Speech assists persons who are unable to read, particularly blind people, is a significant function [5].

Voice Recognition/Voice Command

Voice recognition/command refers to a device's capacity to receive and comprehend spoken instructions, as well as interact and respond to human orders. It is also a feature that allows the hands-free navigation of a device. Mostly used controls are Swiping left or right, scrolling up or down, turning up the volume, muting sound, opening apps, playing music, making emergency calls, taking screenshots, and Typing texts. This feature uses natural language processing and speech synthesis in order to help users [6].

Image Recognition

The ability of software to recognize the writing, objects, places, and actions in images and people is referred to as

image recognition. To recognize images, computers can employ machine vision technology in conjunction with a camera and artificial intelligence software. Many machine-based visual tasks, such as tagging image content with meta-tags, performing image content search, and directing autonomous robots, self-driving automobiles, and accident-avoidance systems, rely on image recognition [7].

Alarm Clock in Smartphones

This is a feature that allows a smartphone to function as an alarm clock, which helps in waking up people or as a reminder for a scheduled activity but with greater flexibility. Currently, all mobile phones with varied feature sets have this functionality. Most gadgets, for example, allow you to set an unlimited number of alarms that will repeat daily or weekly [8].

Calculators in Smartphones

Calculators help users to easily perform calculations of numbers with the use of their phones. Calculators have fundamental operations (addition, subtraction, multiplication, and division), while others include more complex features like square root or trigonometric functions [9].

Related System

The study was also based on:

Speech Services

In this paper developers of Google uses Google Cloud Text-to-Speech which is powered by WaveNet, a software that was developed by Google's UK-based AI Company DeepMind, which was acquired by Google in 2014 with various AI features.

Google created it for its Android operating system. It enables programs to read out (say) the text on the screen in a variety of languages. Text-to-Speech may be utilized by apps like Google Play Books to read books aloud, Google Translate to read aloud translations that provide helpful insight into word pronunciation, Google Talkback, and other spoken feedback accessibility-based apps, and third-party apps. Each language's speech data must be installed by the user [10].

This paper described the Text-to-Speech architecture that blind people utilize to simply and effectively access numerous apps. We can use this to build and implement Text-to-Speech interfaces like TalkBack.

Google Assistant

In this paper, the developers of Google Assistant Service use a low-level API that allows you to alter the audio bytes of an Assistant request and response directly. For all systems that support gRPC, bindings for this API may be built-in languages such as Node.js, Go, C++, and Java. () [10].

Google produced a mobile and smart home app that is primarily available on mobile and smart home devices. Unlike the company's previous virtual assistant, Google Now, the Google Assistant can hold two-way conversations.

This paper provides a prototype of a voice application integrated with voice recognition. We may create designs with the aid of this and utilize voice user interfaces such as voice commands.

Siri

Siri is the brand name for the intelligent voice assistant included in practically every Apple product. It also encompasses all machine learning and on-device intelligence technologies used for smart suggestions. Apple's smart assistant is effortlessly triggered across all Apple devices with a tap, push, or wake command. You can make instructions or questions from practically any Apple device using the universal wake phrase "Hey Siri." When speaking, the spoken commands are processed locally to determine whether they can be performed on the device or not [9,10].

DESIGN AND METHODOLOGY

This chapter discusses the design and methodology of the application. It contains Requirements Analysis and Requirements Documentation.

Requirements Analysis

To fully utilize and optimize the application's main functions, a process of researching and analyzing the application's process was required. To access the app's features, voice recognition and text-to-speech are used. **Figure 1** displays the study's conceptual framework.

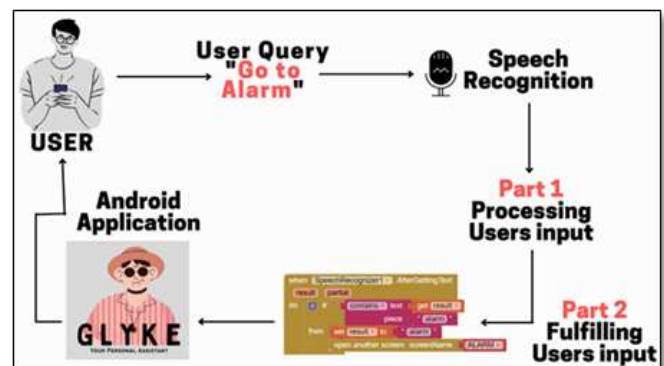


Figure 1. The study's conceptual framework.

Requirements Documentation

To successfully develop the application, the proponents follow a step-by-step process that includes the Planning Stage, Quality Assurance Stage, User Documentation, and Release of Voice Command Based Assistive App for the Blind, GLYKE - Your Personal Assistant, as shown in **Figure 2** Each phase is discussed below.

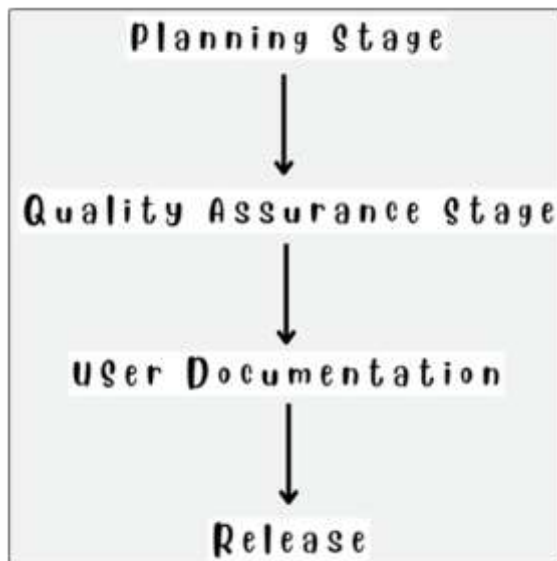


Figure 2. Methodology Flow.

Planning Stage

The developers are starting to plan how the application should look and function. Coding standards, design patterns, user flows, and mental maps must all be established.

Quality Assurance Stage

To create a successful application, developers create testing standards, tasks, test flows, and test types. Defining the testing subject. Testing results must be documented.

User Documentation

The developers create application usage and installation instructions such as system documentation, end-user guides, and installation guides.

Release

If the application has already passed all the stages the app is now then released, for all the blind and visually impaired individuals to access and use.

Application Criteria

The respondents of this research are people who are blind, visually impaired individuals, and even people who are not blind or visually impaired (Table 1).

Table 1. System Evaluation Criteria.

Criteria	Description
Presentation of Contents	The minimum content requirements of the user's expectation to be displayed in the system.
User Friendliness/ User Interaction System Functionality	This assesses whether the users find the system easy to use with respect to their experience in the site. The system should be easy to learn and navigate that allows users to achieve the goals and objectives.
System Functionality Look and Feel	This assesses the functionality of the system whether the links are functional and response to a request of the users.
Look and Feel	This assesses the appearance and responsiveness of the system in terms of the movement and speed of navigation.
Overall Performance	This assesses the system performance such as the reliability, usability and efficiency of the system to perform the task during user interactions.

DEVELOPMENT AND TESTING

This chapter presents the design and development used in the project. It includes the Description of Prototype, Testing and Development.

Description of Prototype

The project is designed specifically for the blind or visually impaired, and it includes buttons as features. The right side of the screen is completely covered with a button, and the left side is similarly covered. When clicked, the right-side button

accepts the user's voice command, while the left side button returns the user to the main screen. The following are the screens associated with each application feature.

Main Screen

Figure 3 shows the GLYKE Main Screen. The application prototype opens with a welcoming phrase introducing itself and informing the user of the necessary actions in the main menu screen which is shown above.



Figure 3. GLYKE Main Screen.

Features Screen

Figure 4 shows the GLYKE Features Screen. The screen will display the features of the application while also using Text-to-Speech to read aloud the said features together with the instructions.

Alarm

Figure 5 shows the GLYKE Alarm Screen. The alarm screen where in users will be able to set an alarm mainly using their voice to command while the system does the process of setting the alarm.



Figure 4. GLYKE Features Screen.



Figure 5. GLYKE Alarm Screen.

Time

Figure 6 shows the GLYKE Time Screen where the application speaks the current time for the user to hear, the screen does not display the time, it only speaks it aloud.



Figure 6. GLYKE Time Screen.

Calculator

Figure 7 shows the GLYKE Calculator Screen. In this screen users will be able to do calculations with voice command.

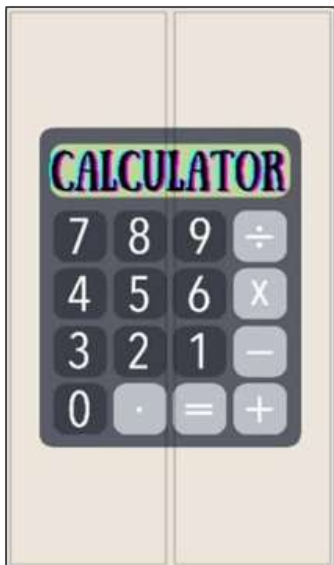


Figure 7. GLYKE Calculator Screen.

Object Classification

Figure 8 The above figure shows the screen displaying image where the camera is positioned. When clicking the right side of the screen, the camera will classify the object in front of it and speaks aloud the result of classification process.

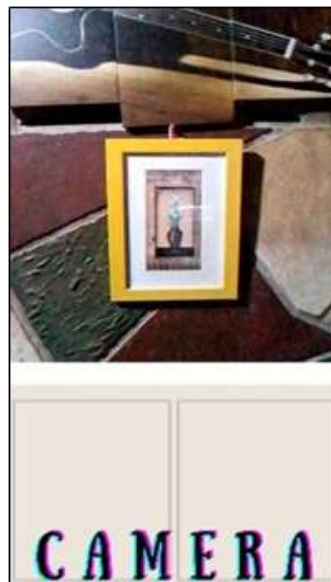


Figure 8. GLYKE Camera Screen.

Temperature

Figure 9 shows the GLYKE Temperature Screen. In this screen, when the user asks about the current temperature, the application will process the command and then it will speak the temperature. In this feature the user can also ask about the sunrise time.

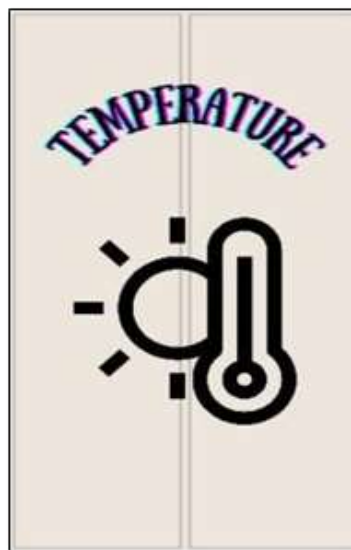


Figure 9. GLYKE Temperature Screen.

Battery

Figure 10 shows the GLYKE Battery Screen. The figure shown is the screen for asking the current battery percentage, in which the application speaks aloud the information.



Figure 10. GLYKE Battery Screen.

Object Sensor

Figure 11 shows the GLYKE Object Sensor screen. The figure shown below is a feature where the application can detect whether there is an object in front of the user.

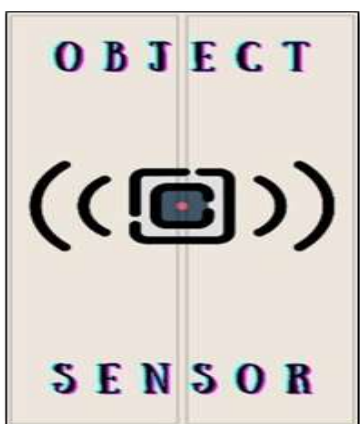


Figure 11. GLYKE Object Sensor Screen.

Testing

The developers tested the system to identify various errors in order to see which part of the application needs improvements. The developers acted as blind and used the application without using their eyes.

Development

The application is developed through MIT2 App Inventor using a block-based programming language. The application's design and features are developed with the help of ai2 app inventors' user interface feature, layout and media feature, Speech Recognition extension, and Text-To-Speech extension.

Below are the applications features development:

Main Screen: The main screen is developed by adding a text to speech feature in which it will read the programmed text to the user, a designated button to maneuver the application, and a speech recognizer to input the user's command. The speech recognizer is then programmed to do what the user commands.

Features: The feature is developed by adding a text to speech feature in which it will read the programmed text to the user, a designated button to maneuver the application. It then calls the text to speech feature in which it will read the programmed text to the user.

Alarm Screen: After the user commands the application to go to the alarm feature the application then calls the speech recognizer to do what the user commands. The Alarm feature is developed by adding the TaifunAlarm feature which will enable the application to set an alarm programmatically.

Time Screen: After the user commands the application to go to the Time feature the application then calls the speech recognizer to do what the user commands. The Time feature is developed by adding a clock sensor that provides instantly time using the internal clock of the user's phone.

Calculator: After the user commands the application to go to the Calculator feature the application then calls the speech recognizer to do what the user commands. The Calculator feature is developed by programming the calculator feature that helps the user to calculate numbers.

Camera: After the user commands the application to go to the Camera feature the application then calls a speech recognizer to do what the user commands. The Camera feature is developed by adding a WebViewer interface that can be used to properly communicate between your app and JavaScript code running on the WebViewer page. The Looktest extension is also added to classify video frames from the device camera.

Temperature: After the user commands the application to go to the Temperature feature the application then calls the speech recognizer to do what the user commands. The Temperature feature is developed by programming the feature to know the temperature and sunrise time in a specific location the user asks.

Battery: After the user commands the application to go to the Battery feature the application then calls the speech recognizer to do what the user commands. The Temperature

feature is developed by adding the TaifunBattery extension that reads the Capacity, Health, Status, Temperature, and Technology of the user's phone battery.

Object Sensor: After the user commands the application to go to the Detect feature the application then calls the speech recognizer to do what the user commands. The Detect feature is developed by adding a proximity sensor that can detect whether an object is in front of the user. Accelerometer Sensor is also added to turn off the object sensor.

CONCLUSION AND RECOMMENDATION

This chapter gives the developers' conclusions and recommendations based on their observations of the project.

Conclusions

With additional testing, the application performed as expected. The buttons that were added are performing well in terms of process execution and providing the expected output, which benefits the user in certain situations. It was also discovered that the application's Speech Recognizer does not always correctly input the user's command because loud noises can interfere with word processing.

Recommendations

The developers recommend that the study be improved by including more available features that benefit the Blind and Visually Impaired. The developers believe that as the prototype evolves, it will become better. Furthermore, developers recommend improving Speech Recognition by making it faster in processing commands and improving response time to better assist users.

References

1. Koyuncular B (2021) The Population of Blind People in the World! Available online at: <https://www.blindlook.com/blog/detail/the-population-of-blind-people-in-the-world>
2. Elflein (2021) Number of blind people worldwide in 2020 by age and gender (in millions). Retrieved from: <https://www.statista.com/statistics/1237876/number-blindness-by-age-gender/>
3. GSMarena (2022) AlarmClock definition. Available online at: <https://www.gsmarena.com/glossary.php3?term=alarm-clock#:~:text=This%20is%20a%20feature%20allowin,g,a%20daily%20or%20weekly%20basis>
4. GSMarena (2022) Calculator definition. Available online at: <https://www.gsmarena.com/glossary.php3?term=calculator>, <https://appleinsider.com/inside/siri>
5. Dakic M (2022) What Is Voice Recognition Technology and Its Benefits. Available online at:

<https://zesium.com/what-is-voice-recognition-technology-and-its-benefits/>

6. Quiller Media Inc (2022) Siri Sam Joehl Meaghan Roper and Jaclyn Petrow Understanding Assistive Technology How does a Blind Person use the Internet. Available online at: <https://www.levelaccess.com/understanding-assistive-technology-how-does-a-blind-person-use-the-internet/>
7. TechTarget Contributor (2021) Image Recognition. Available online at: <https://www.techtarget.com/searchenterpriseai/definition/image-recognition>
8. TechTarget Contributor (2021) Voice Control (Voice Assistance). Available online at: <https://www.techtarget.com/whatis/definition/voice-assistant>
9. The Understood Team (2022) What is text-to-speech technology (TTS) and-how-it-works.
10. World Health Organization (2021) Blindness and vision impairment. Retrieved from: <https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment>