



Voice Assistant Android Application

Prathmesh Godse
Student

Information Technology
Pillai College of Engineering
Panvel, Maharashtra

Ronak Kothari
Student

Information Technology
Pillai College of Engineering
Panvel, Maharashtra

Mahendra Sutar
Student

Information Technology
Pillai College of Engineering
Panvel, Maharashtra

Gayatri Hegde
Faculty

Information Technology
Pillai College of Engineering
Panvel, Maharashtra

Abstract : Smartphones are a vital part of life these days due to the connectivity they provide. One of the things that make smartphones so important in our daily lives is their efficiency. The speed with which you can do tasks on a smartphone is almost unmatched. The major drawbacks of smartphones were their touch operations which create difficulties in doing daily mobile related tasks for visually impaired people. The Voice Assistant Android Application interacts with visually impaired people using voice command, and assists as well as provides the support required to complete their task easily. Speech recognition API provided by google, allows users to verbally command the application to carry out ordinary tasks such as searching on the web, making a call, adding contact, reading and composing messages, reading and composing email, Optical Character Recognition(OCR), adding Notes, looking for Weather, searching Nearby Places, live location information, health tracker, setting an alarm, etc.

Keywords : visually impaired, Google Speech API, Android, Optical character recognition.

I.INTRODUCTION

A. Android Studio : Android studio is an IDE(Integrated Development Environment) for Android Application development and developed by Google. The Programming languages supported are JAVA and Kotlin. Android studio helps to build and manage android apps efficiently and fast. Android studio has auto generated code editor support which means that developers need not code from scratch. One can run an android app from android studio by connecting android phone via USB or can use Android emulator to run and debug the android app.

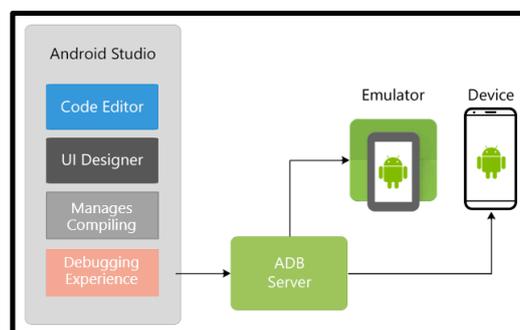


Fig. 1.Android Studio Architecture

The main objective of the application is to use voice commands to access necessary and some basic functionalities like calling, messaging, creating notes, etc.

II. LITERATURE REVIEW

In paper [1] the major purpose of this system was to provide a system so that the blind and physically Disabled population can easily control many functions of a smart phone via voice. But this application is only designed for calling. In paper [2], Mobile Digital Assistants are mobile applications that act as software agents/ middle men to allow easier interaction between users and their mobile devices. These applications can perform tasks, or services, for an individual based on user input, location awareness, and the ability to access information from a wide variety of online sources such as news, stock market, traffic updates, user schedules and retail prices among others. Use of databases to store data, as more and more data comes, it increases the memory space in mobile which can decrease the performance of mobile is the downside of this technology. In paper [3], they have presented the system designs and use cases of JustSpeak, a universal voice control assistant on the Android operating system. The contributions of JustSpeak are twofold. First, it is the first voice control application which synthesizes the commands set on screen context and provides enhancements to all applications running on a mobile system. Secondly, it allows one block command to run multiple commands which enables more seamless interaction experience. Unfortunately, application developers often assign lower priority to accessibility than to other features and then simply forget providing alternative text or content descriptions to visual interface elements such as images or icons. In paper [4], the application is a chat-bot and responds to the user queries through voice, so that the user can listen to the search results. The application mainly consists of five features. The Visiting card feature lets the user scan a visiting card and provides an option to the user to save the contact. Using Google Cloud Vision API (Optical Character Recognition), this application scans the visiting card to extract the contact details. Basic and important features like Calling and messaging are not included. And other features are CryptoCurrency Tracker, Twitter Tweets Reader, Image Reader and Finding information from Wikipedia are also included. But all these features are useful when there is touch operation, but visually impaired peoples can not access these features.

III. PROPOSED ARCHITECTURE

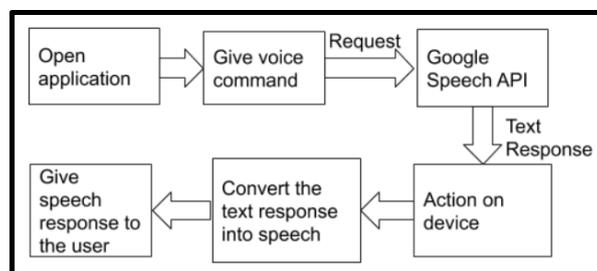


Fig. 2. Proposed Architecture

This application works on users voice commands. User gives the voice command to the application. This command is then sent to google speech API where the data is processed and converted into the text. Then this text response comes back to the application. If the text response matches any module in the application, then the application responds to the text accordingly and converts the result received after performing the action into speech which is delivered to the user through the speaker or headphones of the device. And if the text response does not match any module then the application does nothing, that is no speech output is delivered.

IV. IMPLEMENTATION DETAILS

Voice Assistant Android Application consists of multiple features which are based on voice commands. Features include: Phone Module, Message Module, Web Search Module, Help Module, Live Location Module, Mail Module, OCR Module, Date and Time Module and Live Weather Module. The Output of these modules are provided in two forms: Text and Speech. Phone Module consists of the following sub modules : call history, contact details, search contact number and call on contact/number. These above commands do all the basic functionality of the Call and Contact feature of Smartphone. Message Module helps to send and read messages from the user's smartphone using voice commands such as read message and send message. Web Search Module consists of the following sub modules : wikipedia search module and google search module. Users can use either of these modules to fetch relevant information from the Web and also output in two forms: Text and speech. Help Module helps users to navigate within the application. This module is embedded within other modules such as phone module and message module which helps the user to know the voice commands required within that module. Live Location Module provides the user current location information. This module is implemented using Location manager and geocoder which are provided by android studio. Mail Module helps the user to compose or read mail within the application. Mail Module consists of following submodules such as compose mail and read mail. This module is implemented using JAVA Mail API. Live Weather Module gives the weather information of the user's current location. It also provides the weather information of different locations within the application. Date and Time Module helps the user to know the current time and date. Also User can add notes and set an alarm using voice commands. Using Nearby Place feature, User can know the location of nearby places such as hospital, cafe, stores, bank, atm, restaurant. Using the Health Tracer feature, users get to know the Number of steps covered and calories burnt. This Application also includes a news section where users can find the latest top headlines of the day. OCR Module helps User by speaking out the text from text formatted documents using the camera source.

4.1. Algorithm

Voice Assistant Android Application uses Google Speech Recognition API for controlling voice inputs and outputs and Optical Character Recognition Algorithm for live reading of texts. Google Speech Recognition API uses two algorithm: Speech-to-text Recognition Algorithm and Text-to-Speech Recognition Algorithm:

4.1.1 Speech-to-Text Recognition : A Speech-to-text recognition API takes voice commands as inputs. Then this speech input is segmented into smaller chunks of word which is known as segmentation. Then the token of segmented words is passed through a Acoustical Model which checks the audio signals of basic speech units taking into account speech variability of the speaker. Then it is passed through a pre-trained lexical model which checks the semantics of the given input.

4.1.2 Text-to-Speech Recognition : Text-to-speech algorithm provides people to implement human sounding speech audio. One can use this audio in various applications that require speech output. Text-to-speech converts the text on the screen into speech which is then delivered to the user using their device speakers, Bluetooth, headset etc. Text-to-speech allows the user to convert any string, paragraph, document, etc. into human sounding speech output.

4.1.3 Optical Character Recognition(OCR) : Optical character Recognition allows us to extract data from different types of documents such as scanned documents, PDF files or images captured by a camera and then it converts this data into machine encoded or machine readable text.

It works on 3 steps :

1. Image Pre-Processing in OCR: In Image Pre-Processing the image is processed to remove unwanted distortions and to enhance the unique characteristics of the image which help in recognizing the character..
2. Character Recognition in OCR: Character Recognition extracts the words from images or videos received after Preprocessing is done.
3. Post-Processing in OCR: Post-processing is an error correction technique which ensures that the detected character is correct and accurate. The accuracy of the character that is to be returned can be increased by training the model with a large data set of distinct characters.

4.2. Use Case Diagram

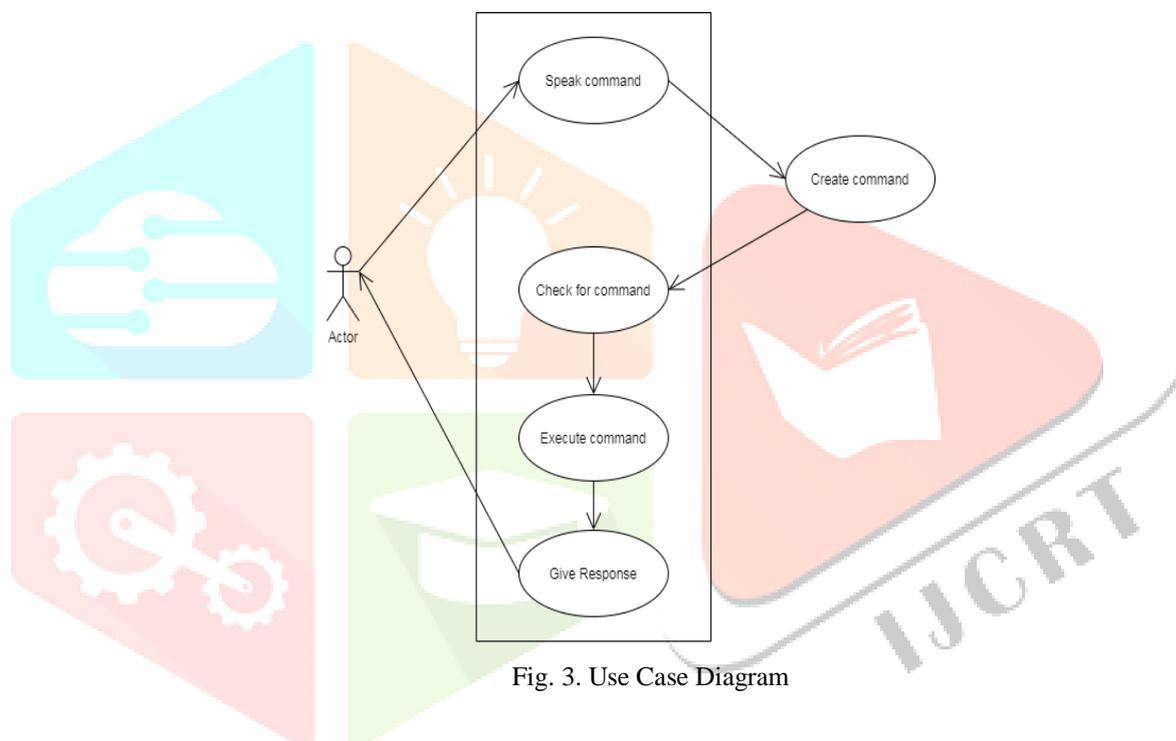


Fig. 3. Use Case Diagram

1. *User (Actor)* : There will be one user per application, who will manage the application through his voice.
2. *peak command* : User can activate the application manually or by his voice commands, then the user speaks a specific command for which he wants an answer from the application for e.g : one can command the application for weather Information.
3. *Create command* : The voice command is then converted to text using speech-to-text API.
4. *Check for command* : In this section the application will check for the module that is based on the text command received from speech-to-text API.
5. *Execute command* : The command received is executed in the application.
6. *Give response* : As soon as the data is fetched the text data is converted into speech and it is given as response to the user.

V. ANDROID APPLICATION

Following are the commands/Modules in the application:

- ‘Call Logs or call history’ — List out the call logs/ history
- ‘Contact Details’ — Displays all contact details
- ‘Search contact contact name’ — shows the contact detail of particular person
- ‘Call contact name’ — Call on specified name by user
- ‘Send message’ — get input from user via voice and send message
- ‘Read message’ — Read all the messages from user device
- ‘Wikipedia’ — User provide input via voice and application will get the required information from wikipedia and displays it to the user
- ‘compose mail’ — opens the mail module where user can able to send mail to others

- 'Read mail' — Read mails from inbox
- 'Location' — Application provide users current location
- 'Weather' — initially provides the user's current location and when providing the name of another city it will display the weather information of that location.
- 'add contact' — user can create new contact
- 'scanner' — Open OCR module which help to detect text from documents, provided with speech output
- 'health or health track or pedometer' — tracks number of steps user walk based on number of steps application will calculate how much calories are burned and distance travelled
- 'notes' — allow user to create new notes and will read already create notes after providing read command
- 'alarm' — user can set alarm, application will guide user to set alarm
- 'places' — this command will open Near by places module UI and will display nearby places according to user input
- 'news' — will display the top news headlines
- Miscellaneous- 'Battery status', 'wifi on/off'

Here are few test results that we achieved :

- 1) **OCR module** : In this module we open the camera using the 'Scanner' command. Then the user has to say 'capture', which will enable the use of Optical Character Recognition(OCR) algorithm to detect text from documents, and then with 'Read' command text is delivered to the user by speech output as well as displayed on the device.



Fig. 4. Output of OCR module

- 2) **NearBy Places Module** : We can access this module by giving 'Places' command in the application. In this module we can get the name and address of nearby places by speaking the command related to or which signifies the same meaning as the location. For example, if a user gives 'hospital' command then names and addresses of all the nearby hospital with respect to user location will be given as output.



Fig. 5. Output of Places Module

- 3) **Alarm module** : In this module, a user can set an alarm using the User interface (by adjusting the time over UI clock) or just by saying time for setting an alarm such as "6.45 pm".

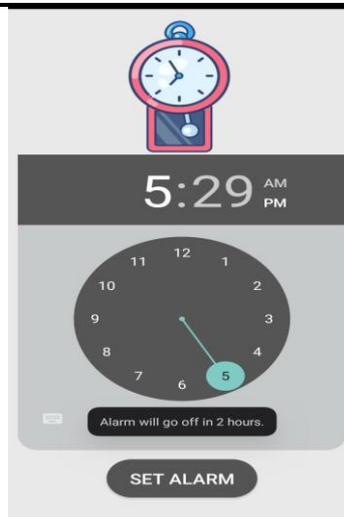


Fig. 6. Output of Alarm module

- 4) **Weather Module** : In this module, the user can get weather information of its current location by giving command as “weather” and the user can also get weather information of different cities of the world.

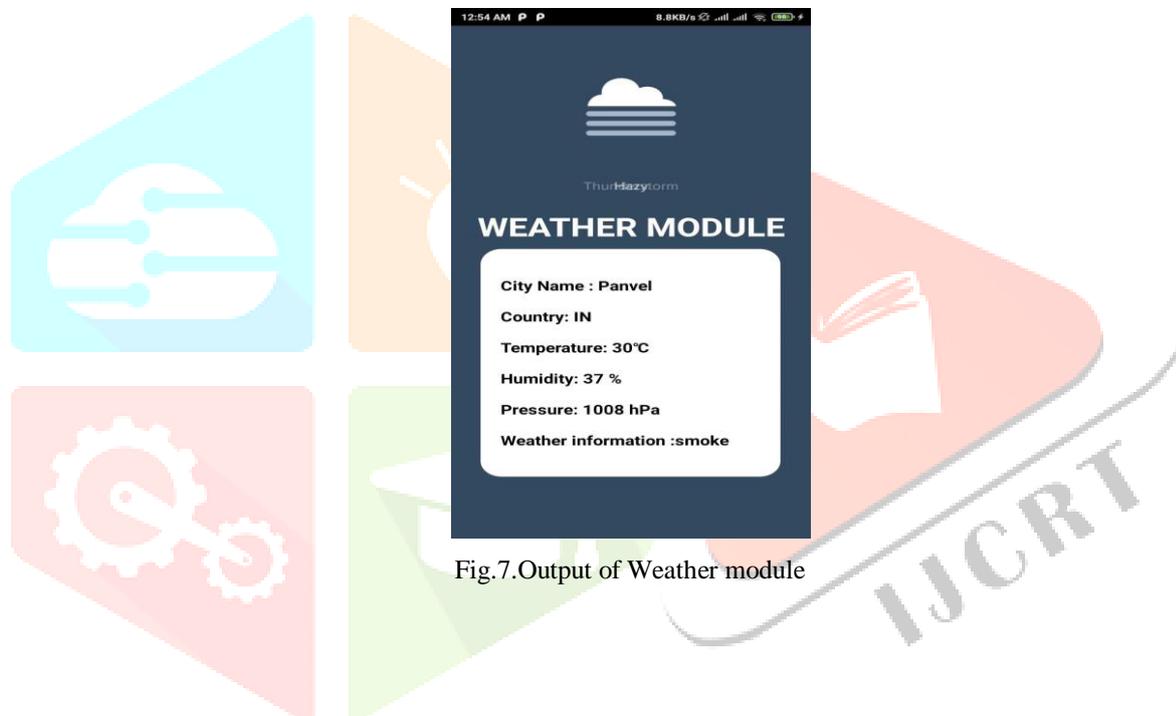


Fig.7.Output of Weather module

VI. FUTURE SCOPE

The application performs most of the important functions, but there are different areas which could be improved as there is a lot of room for expansion. There are many features which could be included in the project. Features such as :

- Product identification using the device camera.
- Real time object detection using the camera of the device which can help the user in walking in scenarios where he forgot his walking stick or stick gets stolen, etc.
- Music player.
- Directions from current location to a specific nearby location using the nearby places module.
- Can extract detailed information on specific news using the news module

VII. CONCLUSION

Voice Assistant Android Application is an android application developed for visually impaired people to assist them in their daily life mobile activities. The application is developed with multiple features which can help to ease the daily task for visually impaired people using voice commands. The main functionalities of the application is to give output response through speech. Main features include calling, messaging, web searching, reading out the text in the images, providing live location, etc . All these features can be accessed using voice commands without the need of touching the device screen.

VIII. REFERENCES

- [1] Renu Tarneja, Huma Khan, Prof.R.A.Agrawal, Prof. Dinesh. D. Patil, "VOICE COMMANDS CONTROL RECOGNITION ANDROID APPS", International Journal of Engineering Research and General Science Volume 3, Issue 2, March-April, 2015.
- [2] Bernard Shibwabo, Kevin Omyonga, "The Application of Real-Time Voice Recognition to Control Critical Mobile Device Operations", International Journal of Research Studies in Science, Engineering and Technology, Volume 2, Issue 7, · July 2015.
- [3] Yu Zhong , T.V. Raman, Casey Burkhardt, Fadi Biadisy and Jeffrey P. Bigham, "Enabling Universal Voice Control on Android", National Science Foundation, April 9, 2014.
- [4] Ravi Teja Neela, "Smart Assistant for VIP: Android Application", Department of Computing Sciences Texas A&M University-Corpus Christi Corpus Christi, Texas, June 2018.
- [5] M. A. Khan Shishir, S. Rashid Fahim, F. M. Habib and T. Farah, "Eye Assistant : Using mobile application to help the visually impaired," 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), Dhaka, Bangladesh, 2019.
- [6] Belinda, M.J. & Nandhagopal, Rupavathy & Mahalakshmi, N.R.. "A voice based text mail system for the visually impaired", International Journal of Engineering and Technology(UAE),2018 .

