ISSN XXXX XXXX © 2019 IJESC

Research Article Volume 9 Issue No.3

Voice Assistant for Visually Impaired

Gayathri .S[1], Porkodi Venkatesh[2], Pushpapriya Premkumar[3]
Assistant Professor[1], BE Student[2, 3]
Department of Computer Science and Engineering
Jeppiaar SRR Engineering College, Chennai, Tamil Nadu, India

Abstract:
In this modern society, visually impaired persons need helpful tools for operating digital devices. Our idea mainly focused on designing and implementing an assistive system for visually impaired persons to access the Android smart phones easily. The existing system is talk back which leads to ambiguity and it also produces unwanted voice. The proposed system is used to help the visually impaired to have access to the most important features of the phone. The aim is to design a low-cost and high-performance assistive device for daily activities of visually impaired persons.

Keywords: Voice Assistant, Text to speech, Mobile application, TTS Engine, Voice recognition.

I. INTRODUCTION

Assistive technology for visually impaired persons has been studied for many years [5]. In this information society, visually impaired persons need assistive tools to help them for operating digital devices so that they could get and apply digital information while learning, living and working. According to the report of Ministry of Interior, Taiwan, 80% of visually impaired persons are not originally blind and most of them have the ability of hearing. Therefore it arises that visually impaired persons operate digital devices by the help of voice guidance. A screen reader supports voice features to help users operating computers and dramatically reduces the difficulty of operating computers for visually impaired persons. For example, Jaws and NVDA are both famous screen readers on MS-Windows platforms. The smart phone is one of the commonly used digital devices and the smart phone with Android system is inexpensive and so popular and in the market. Nearly 50% of visually impaired persons are unemployed. This paper proposes and implements an integrated app Voice Assistant on the Android platform for visually impaired persons.

II. PROPOSED SYSTEM

Voice assistant is an application, installed in the mobile phones used to help the visually impaired to have access to the most important features of the phone. There is one time login to identify the particular user. The user can initiate this application by pressing the volume button by two times. The user is provided with a speaker he/she can speak out their needed features. It includes sending messages, accessing notes, making phone calls, and knowing the battery level just by their voice as their input. This system helps the user to read the contents of the message along with the sender and the date and time. This system tells the battery level in the phone helps the user to read the numbers in the call logs and also supports the voice dialler feature. This system helps the user to read the notes in the phone.

III. SYSTEM IMPLEMENTATION

More number of applications is available for blind people. Apart from the proposed system, visually challenged person can send / receive textual messages, can make a voice call to the receiver, and they can make notes on mobile application using voice input and also they can check battery level using voice input. They can initiate this application by pressing the Volume button by one, two, three, four times. And also received messages will be notified using voice. Initially the voice is given as an input that is recognized using the speech synthesis technology. The first step is user can speak out the needed facilities such as messages, calls, battery status or notes. Then in the selected facility the user has to select the viewport area. Once the view port is selected, the contents in the selected viewport is queued in the TTS queue and waiting for the process. TTS engine recognize the content in the TTS queue and then it converts the text content in to voice.

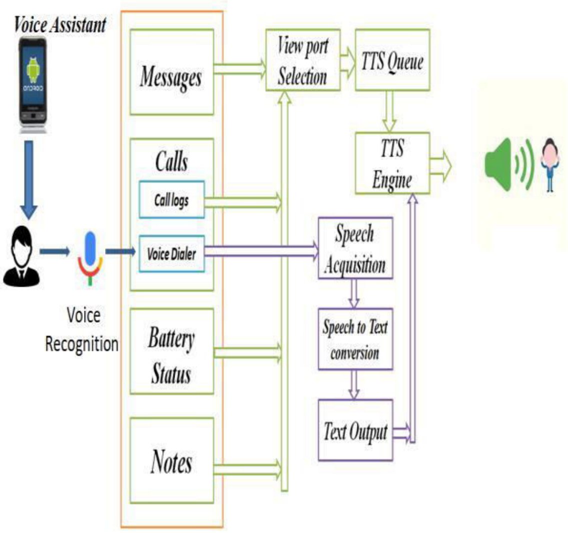

Figure .1. System Model

IV. VOICE RECOGNITION

Implementation of voice application in mobile phones has opened the marketplace for disabled people. Similar type of technology was apple's voice assistant siri. The siri helps users to make contacts, send email and plenty of additional practicality. Just like siri, Google has speech recognition (SR) API. The foremost distinction between each of them is that siri will perceive multiple phrases and words, Google voice majorly centered for specific phrases and keywords. Speech

recognition adds another dimension to the classic keyboard input which is very easy for the user to access i.e. manipulation of text is much easier than the classic methodology. This application uses the Google API that uses the Hidden Markov Models (HMM) methodology to send sms. In this sms application the user must speak out the numbers to contact any person so that sms will be sent or messages will be received from the receiver. This application also ensures that user will solely input numeric character for contact data, i.e. the safety validation for range is finished. SR can hear input and convert numeric to text and can be displayed on contact data to verify. If any user try and insert the other character into the data a slip would be displayed e.g. if user speaks his name for contact, it'll be displayed as invalid contact. The message box will settle for any character. To use the speech recognition user must be loud and clear in order that command is correctly handled by the system. The system shown here can use SR with Google server that uses HMM methodology. The temporary description of how the speech is recognized is as follows. First the speech is inputted, sound may be unsteady set of signals that is measured and recorded. These signals depend on speaker and also on his/her voice quality and phrase of the language. The computer file is split into words and phrases, i.e. command is split into many components. Lastly in the process section system understands command and executes it. Speech Recognition stands majorly on 5 pillars that contain square measure, feature extraction; acoustic models information that is constructed and supported the coaching information, dictionary, language model and also the speech recognition algorithmic rule. The computer file i.e. voice is 1st regenerated to digital signal and square measure sampled on time and amplitude axis. This digitalized signal is then processed. For process the signal is split into tiny intervals, which depend on the algorithmic rule used. The generalized timestamp is twenty ms. this division relies on the options of knowledge as those options square measure is compared with information part. Information part contains data of feature of the word found and accordingly the command is made. The basic part may be a sound for continuous speech or word for isolated words recognition. The lexicon or information is employed to attach the frequency model i.e. the vocals with actual vocabulary word. The signal particularly speech has its constraints as same speech ought to match that means of matter language brain created. The HMM model uses words for modeling. States sequence isn't a command that's SR system typically assumes that the signal is realization of message that is encoded as a sequence of symbols. Here symbols square measures the words sampled. To result the reverse operation of recognizing the underlying image sequence given in a spoken auditory communication, the continual speech is 1st regenerated to a sequence of equally spaced distinct parameter vectors.
Vectors of speech characteristics consist principally of MFCC (Mel Frequency Cepstral Coefficients), standardized by the Eu Telecommunication Standards Institute for speech recognition. The MFC may be simply created. The analysis is performed based on sample i.e. divided information is converted to variable information measure, triangular filters square measure, placed in conjunction with the Mel frequency scale and energies square measure is calculated by spectrum.

V. MODULES

The System module is categorized into four sub-modules namely,

Module 1: Mobile client registration and start up process.
Module 2: Voice recognition for selecting messages, calls, notes and battery status.
Module 3: Auto fetching of numbers in voice dialer.
Module 4: View port selection and producing voice output.

1. Mobile client registration and start up process Mobile client is an android application which is created and installed in the user's android mobile phone. The application home page consists of the user registration process. Next process is to create the user login page by dragging button and text field class in the android tools panel. Once the application is designed and then it is implemented using code. Once the creation of full mobile application is done, it will generate as android platform kit (APK) file. This APK file will be installed in the user's mobile phone as an application. The user can initiate this application by pressing the Volume button by two times.
2.Voice recognition for selecting messages, calls notes and battery status. The user selects messages, calls, notes or battery status by voice commands. Google API voice will recognize commands given by user. And it will play voice output to the blind people for understanding. In the message block, there are two categories such as sent items and inbox messages. In the calls block, the system consists of voice dialer and call log details. When user selects the battery status block, the system speaks out the battery level in the user's mobile phone. The user notes stored in the mobile phone read by the system when they get into notes block**.**
3.Auto fetching of numbers in voice dialer The user speaks out the number or tells the name of the person in the contact list to make a call. This system will create the main requirement of every speech to text conversion system in a database which will compare speech with frequencies. User gives the voice input to fetch the number and that is converted to text format which is displayed in dial pad.
4. View port selection and producing voice output The view port selection method is applied for messages, call log and notes. If the user touches any notes in the list then the system reads out the content in it. Similarly, this is applicable for inbox, sent items and call logs. This system produces the voice output which improves the performance.

VI. TEXT TO SPEECH

This system additionally helps the user to scan the contents of the message, call logs, battery standing and notes. Automaton permits you exchange your text into voice. Not solely you'll be able to convert it however it additionally permits you to talk text in style of completely different languages.

Text-to-Speech (TTS), additionally called speech synthesis, in automaton is a straightforward nonetheless powerful feature you'll be able to use to supplement your apps in terms of benefiting your users in a very thoughtful method. Automaton TTS API offers multiple-language support, management of voice characteristics and options, file output, and so on. With simply a little range of lines of code, you'll be able to build your apps to reach a wider audience. Automaton provides Text To Speech category for this purpose. In order to use this category, you wish to instantiate the associated object of this category and additionally specify the init Listener. During this beholder, you need to specify the properties for TextToSpeech object, like its language, pitch etc. Language are often set by set Language () methodology. The Text-to-Speech API is ideal for any application that plays audio of human speech to users. It permits you to convert arbitrary strings, words, and sentences into the sound of someone speaking identical things With the Text-to-Speech API; you'll be able to convert that response string to actual human speech to replay to the user The process of translating text input into audio information is named synthesis and also the output of synthesis is named artificial speech. The Text-to-Speech API takes 2 varieties of input: raw text or SSML-formatted information. We make replacement or audio fill, to "synthesize "endpoint of the API is used. The speech synthesis method generates raw audio information as a base64-encoded string. Decipher the base64-encoded string into associate audio file before associate application will play it. Most platforms and operational systems have tools for decoding the base64 text into playable media file.

VII. WORKFLOW

The system workflow is depicted as follows:

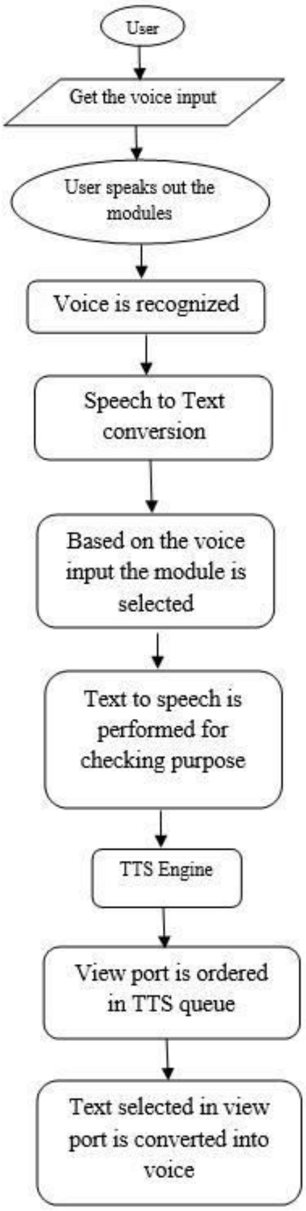

Figure.2. Work Flow

VIII. CONCLUSIONS

The proposed system is used to help the visually impaired to have access to the most important features of the phone ,enhancing the quality of the system making use of the different custom layouts and using text to speech. There are many similar applications in the market having a huge amount of success. This can be later improved by making the system to work, or understand the er's voice even in more disturbing environment. This can be also published, or released in Google play store.

IX. REFERENCES

[1] Fabio Bif Goularte, Siliva Modesto Nassar, Renato Fileto, Horacio Saggion, "A text summarization method based on fuzzy rules and applicable to automated assessment", Expert System with Applications in Elsevier Ltd.,27 July 2018.

[2] Ms. Sneha K. Upadhyay, Mr.Vijay N. Chavda, "Intelligent system based on speech recognition with capability of self-learning", International Journal for Technological Research in Engineering ISSN (Online): 2347 - 4718 Volume 1, Issue 9, May-2014.

[3] Aggarwal,L.,Gaur,V.& Verma, P., All-in-One Companion for Visually Impaired", International Journal of Computer Applications, vol. 79, no. 14, pp. 37-40, 2013.

[4] Al-Khalifa, H. S., "Utilizing QR code and mobile phones for blinds and visually impaired people", In Proceedings of the 11th international conference on Computers Helping People with Special Needs, pp. 10651069, Springer Berlin Heidelberg, 2008.

[5] Hersh, M., and Johnson, M. A. eds., "Assistive technology for visually impaired and blind people", Springer Science & Business Media, 2010.

[6] .B.Raghavendhar Reedy, E.Mahender, "Speech to Text Conversion using Android Platform", International Journal of Engineering Research and application, vol.3, Issue1, Januray-February 2013.

[7].Kuei-Chun Liu, Ching-Hung Wu, Shan-Yin Tseng and Yin-Te Tsai Voice helper:mobile Assistive System for visually Impaired persons", International Journal of Engineering Research and application, July 2018.

[8]. Prachi khilari, Prof.Bhope V.P, "Implementation of speech to Text Conversion, International Journal of Innovative Research in Science, Engineering and technology, vol.4, Issue 7, July 2015.

[9]. SonaliThite, Archana Gore, Sagar Yelmar, Yogesh Lonkar, "Android Based Speech Recognition", International Journal of Advanced Research in computer networking, wireless and mobile communication, vol.2, Issue 1, 09 Apirl 2015.

[10] .Mark Gales, Steve Young, "The Application of Hidden Markov Models in Speech Recognition", Foundation and trends in signal Processing, vol. 1, July 2015.