

***SPEECH RECOGNITION SYSTEM IN ANDROID
(Voice App)***

Ms. Perachi muthu
Reg No: 950716621320
Department of Computer Applications
Francis Xavier Engineering College
sperachimuthu@gmail.com

Mr. A. Arul Amalraj
Assistant Professor
Department of Computer Applications
Francis Xavier Engineering College
arulamalraj@gmail.com

Abstract-Speech recognition technology is one from the fast growing engineering technologies. It has a number of application in different areas and provides potential benefits. Nearly 20% people of the world are suffering from various disabilities; many of them are blind or unable to use their hands effectively. The speech recognition systems in those particular cases provide a significant help to them, so that they can share information with people by operating computer through voice input. This paper is discussed about speech recognition system. This system is mainly used for blind people and handicap peoples because this project is develop in android application and speech recognition system is executed in desktop system also. This system input is human natural speech. Speech is the natural and most important form of communication for human begins. In this system recognition is the kind of technology its using computer to transfer the voice signal to an associated text or sound are used identification and understand for user. This system fully depends on the frequency of speech. Speech recognition is involves in many fields of physiology, psychology, linguistic, computer science and signal processing. This system goal is naturally communication between man machine and in future is related to persons body language.

Keyword: android , HMM, Speech recognition

I. INTRODUCTION

Mobile phones are part of our everyday life. Smartphones offer all availability and many methods to interact with their phones. In android

new version have actions and iphones are applications that enable control of a mobile phone using voice, such as calling businesses and contacts, sending texts and email ,listening to music ,browsing the web and completing common taks Mobile phones have become an inevitable part of our daily lives. It is difficult to think of a day without having our mobile phone by our side. The evolution of mobile phones is witnessed by all of us, touch screens being the latest amongst all of them. While touch screens were once rare, touch screen-based interfaces are now present across a wide range of everyday technologies, including mobile devices, personal computers etc. As touch screens have become mainstream, it is crucial that touch screenbased interfaces be usable by people with all abilities, including blind and visually impaired people. Until recently, most touch screens provided few or no accessibility features, leaving them largely unusable by blind people. Unlike fully capacitated or sighted people, blind people cannot read messages displayed on the smart phones or cannot use the basic functions like calling, messaging etc. of the phone. Hence, Interaction via mobile devices is a challenge for blind users, who often encounter sever accessibility and usability problems. The main issues are due to the lack of hardware keys, making it difficult to quickly reach an area or activate functions, and to interact via touch screen. A touch screen has no specific reference points detectable by feel ,so a blind user cannot easily understand exactly where he is positioned on the interface nor readily find a specific item/function. In this study we, therefore, investigate enriching the user interfaces of touch

screen mobile devices to facilitate blind users' orientation called "NEW VISION", an Android Smart Phone application for the visually challenged. Through this application Calling, Messaging, Retrieving Position, Battery Level and Time is made possible without the need to see the phone. Accessible touch screens still present challenges to both users and designers. Users must be able to learn new touch screen applications quickly and effectively, while designers must be able to implement accessible touch screen interaction techniques for a diverse range of devices and applications. Because most user interface designers are sighted, they may have a limited understanding of how blind people experience technology. We therefore argue that accessible touch screen interfaces can be improved substantially if designers can better understand how blind people actually use touch screens..There are some people who are physically challenged and needs special measures and devices to overcome their disability and merge themselves in the society. Although technology is increasing its potential level, therefore accessibility issues are almost resolved for a routine user; however smart phones are still inaccessible to visually impaired people. Visual impairment, also known as vision impairment or vision loss, is a decreased ability to see to a degree that causes problems not fixable by usual means, such as glasses. Some also include those who have a decreased ability to see because they do not have access to glasses or contact lenses. Of the 37 million people across the globe who are blind, over 15 million are from India. One of the most common problems that many blind and visually impaired people experience is their day-to-day challenge in coping with their impairment. Equipment such as Braille, reading glasses, or a walking stick are just some of the few things that help visually impaired people get along with their lives. With the advancement of technology, a common Android smart phone equipped with specific applications can aid visually impaired and blind people in functioning. Presently, the most enhanced choice of smart phones among visually

impaired users is either an iPhone with a very worthy tool called Voice Over or an inexpensive selection of one of the Android-based smart phones. Using a modern smart phone has an advantage that they provide a wide range of services such as digital camera, speech recognizer, etc. We have delivered an Android based smart phone as a system for image processing and object recognition module which work on images captured by a visually impaired user using a built-in camera. Android 8.0 Features and APIs Android 8.0 (API level 26) introduces a variety of new features and capabilities for users and developers. This document highlights what's new for developers. Make sure to also check out Android 8.0 Behavior Changes to learn about areas where platform changes may affect your apps

II. LITERATURE SURVEY

Yogita H. ghadage **et al**[1] proposed a multilingual language for speech to text conversion. Multilingual language means use of more than one languages ,either by an individual speaker by a community of speakers .multilingual language means one common language is used by all type of community. English is the multilingual language so SR is maximum developed by English language.english language is easy to convey all type of people like rural communities or illiterate people. This paper is used MFCC,SVM, and Minimum distance classifier for developing the speech recognition. **Nehashatma et.al** [2] proposed the realtime speech recognition . this paper is used kalman filter technique it means linear quadratic estimation and kalman filter is remove the unwanted sound from background or otherwise.kalman filter is ideal for system which is continuously changing. This real time SR is used two techniques LPC and MFCC.LPC-linear predicted coefficient ,MFCC-Melfrequency cepstrum coefficient This speech recognition used bidirectional for communication text to speech and speech to text.[2]

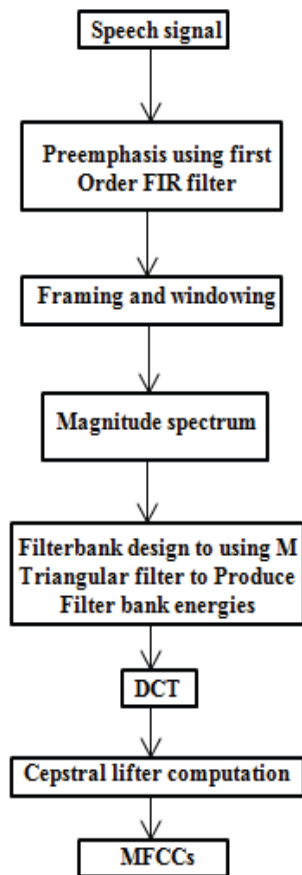


Figure 1: Architecture Diagram

Sheriff yacoub et al[3] Proposed on emotion of human being. The emotions are input of the recognition. Emotions like sadness, happy, angry, boredom, cold anger. This paper used IVR technologies. IVR-interactive voice response. This SR also used two types of database training and testing database.

This is used wizard of oz uses a program that interact with the actor and drives him into a specific emotion situation and then records his responses[3].

B. Raghavendhar et.al proposed online speech to text engine. Speech recognition is developed on the android platform using eclipse workbench. In this speech recognition for application voice sms is done on google server, using the HMM algorithm. Speech recognition system can be divided into several blocks: features extraction, acoustic models database which

are build based on the training data, dictionary language mode and SR algorithm[4].

III. THEORY

Android is developed by open handset alliance, led by google and other company. It is a software environment for mobile devices that includes an operating system, middleware and key application. It is used on smart phones like google own google nexus programs for android also called apps. android program is extension of .apk. Android offers unified approach to application development for mobile devices which means devices which means developers need only develop for android and their applications should be able to run on different devices powered by android.

Main features of android OS:

- Enables free download of development environment for application development.
- Free use and adaption of operating system to manufactures of mobile services.
- Beautiful UI-android os basic screen provides a beautiful and intuitive user interface.
- Connectivity: GSM/EDGE, IDEN, CDMA, U-MIS, bluetooth, wifi, NFC, wimax.
- Storage: SQLite, a lightweight relational DB is used for data storage purpose.

the fifth section gives the future enhancement.

A. ANDROID STUDIO

Android Studio is the official integrated development environment (IDE) for the Android platform. It was announced on May 16, 2013 at the Google I/O conference. Android Studio was in early access preview stage starting from version 0.1 in May 2013, then entered beta stage starting from version 0.8 which was released in June 2014. The first stable build was released in December 2014, starting from version 1.0.

Speech recognition techniques:

Software: SR systems have mostly been designed for windows operating system. SR programmes are available for Mac OS X and windows 7. A SR used acoustic signals for identify a sequence of words uttered by a speaker. Voice

application include a software microphones through USB.

- Software is android studio3.0.1.
- Android platform.
- Android 5.1lollipop

B. HARDWARE REQUIREMENTS

COMPONENT	SPECIFICATION
CPU	Intel Dual 1 Core
RAM	2GB DDR2
Hard Disk	160 GB
Display	Wide VGA (Video Graphics Array)
Input	Keyboard and Mouse
Optional Devices	Android Phone (Jelly Bean 4.2 or higher)

Table.1

C. SOFTWARE REQUIREMENTS

COMPONENT	SPECIFICATION
Front End	PHP, Android
Back End	MySQL
IDE	Dreamweaver, Android Studio
Language	Xml,java
Platform	Windows 7 or later

Table 2

IV. EXPERIMENTAL SETUP

A. SUPPORT VECTOR MACHINES

SVM is one of the powerful state-of-the-art classifiers for pattern recognition which uses a discriminative approach. Optimised margin, between the samples and the classifier border, helps to generalise unseen patterns. SVMs use linear and nonlinear separating hyper-planes for data classification. However, since SVMs can only classify fixed length data vectors, this method cannot be readily applied to task involving variable length data classification. The

variable length data has to be transformed to fixed length vectors before SVMs can be used. It is a generalized linear classifier with maximum-margin fitting functions. This fitting function provides regularization which helps the classifier generalized better. SVM controls the model complexity by controlling the VC dimensions of its model rather than controlling model complexity by using a small number of features. This method is independent of dimensionality and can utilize spaces of very large dimensions spaces, which permits a construction of very large number of non-linear features and then performing adaptive feature selection during training. By shifting all non-linearity to the features, SVM can use linear model for which VC dimensions is known. Sendra et al have worked on a pure SVM-based continuous speech recogniser by applying SVM for making decisions at frame level and a Token Passing algorithm to obtain the chain of recognized words. The Token Passing Model is an extension of the Viterbi algorithm meant for continuous speech recognition so as to manage the uncertainty about the number of words in a sentence. The results achieved from the experiments have concluded that with a small database, recognition accuracy improves with SVMs but with the large database, same result is obtained at the expense of huge computational effort. Automatic speech recognition system involves two phases: Training phase and recognition phase. A rigorous training procedure is followed to map the basic speech unit such as phone, syllable to the acoustic observation. In training phase, known speech is recorded, pre-processed and then enters the first stage i.e. Feature extraction. The next three stages are HMM creation, HMM training and HMM storage. The recognition phase starts with the acoustic analysis of unknown speech signal. The signal captured is converted to a series of acoustic feature vectors. Using suitable algorithm, the input observations are processed. The speech is compared against the HMM's networks and the word which is pronounced is displayed. An ASR system can only recognize what it has learned

during the training process. But, the system is able to recognize even those words, which are not present in the training corpus and for which sub-word units of the new word are known to the system and the new word exists in the system dictionary.

B. ACOUSTIC MODEL

Acoustic model is the main component for an ASR and it accounts for most of the computational load and performance of the system. It is used to link the observed features of the speech signals with the expected phonetics of the hypothesis sentence. The Acoustic model is developed for detecting the spoken phoneme. Its creation involves the use of audio recordings of speech and their text scripts and then compiling them into a statistical representation of sounds which make up words. At present Gaussian mixture models are the dominant technique for modelling the emission distribution of hidden Markov models for automatic speech recognition. HMM suffers two major drawbacks. Strong Independency assumption in HMM states that frames are independent, given a state. As a result, it lacks an ability to deal with a feature which straddles over several frames. Features such as delta coefficient, segmental statistics and modulation spectrum have been developed which can deal with phenomena of straddling. Secondly, HMM is a generative model and fails to discriminate sequences. This weakness has aroused due the maximisation of maximum likelihoods (MLI) instead of maximum..

C. SPEECH OUTPUT

Activities and services can use instances of TextToSpeech to dictate and pronounce content. As of Android 8.0 (API level 26), your app can obtain more precise timing information about when a text-to-speech engine begins speaking individual synthesized words, as long as the engine provides this information. You can use this functionality to call attention to specific words as the text-to-speech engine speaks them. To use these text-to-speech engine improvements in your app, register an instance of UtteranceProgressListener. As part of the

registration process, include a handler for the `onRangeStart()` method.

The text-to-speech engine calls `rangeStart()` to record the point in time at which it expects audio playback of a specific range of text to start. When the audio for that text range starts playback, your app's `onRangeStart()` method executes. Your app can then respond to this callback, such as by highlighting the text range that's associated with the utterance.

Speech recognition activity is startup activity defined as launcher in `AndroidManifest.xml` file. `REQUEST_CODE` is static integer variable, declared on the beginning of activity and used to confirm response when engine for speech recognition is started. `REQUEST_CODE` has positive value. Results of recognition are saved in variable declared as `ListView` type.[8]

D. WI-FI AWARE

Android 8.0 (API level 26) adds support for Wi-Fi Aware, which is based on the Neighbor Awareness Networking (NAN) specification. On devices with the appropriate Wi-Fi Aware hardware, apps and nearby devices can discover and communicate over Wi-Fi without an Internet access point. We're working with our hardware partners to bring Wi-Fi Aware technology to devices as soon as possible. For information on how to integrate Wi-Fi Aware into your app, see [Wi-Fi Aware](#).

E. LOCATION AND MAPS

With the Google Maps Android API, you can add maps to your app that are based on Google Maps data. The API automatically handles access to Google Maps servers, data downloading, map display, and touch gestures on the map. You can also use API calls to add markers, polygons and overlays, and to change the user's view of a particular map area[8].

F. ANDROID SMS SERVICE

There'll be times when you'll want to allow the users to send SMS messages directly from your app to other numbers (destination). The Android SDK does support to capability of sending SMS/MMS messages in two ways (from your app):

- Invoke an SMS client app via Implicit Intents.

- Send the SMS directly yourself by using the SmsManager telephony class API.

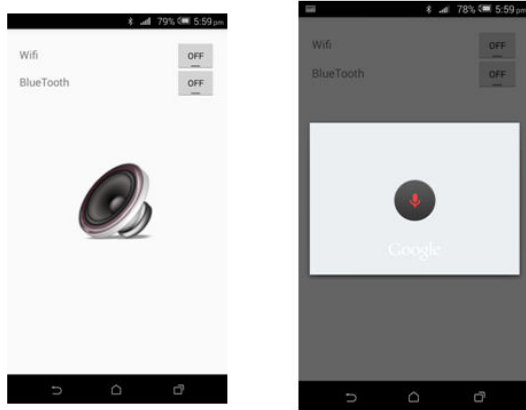


Figure 2: Speech recognition

Usages of speech recognition system:

- In this SR system is mainly used for blind peoples and handicap peoples. The blind people just speak in front of the mobile or computer it automatically converted into text or respective operation is performed. In android mobiles have special apps for speech recognition system.
- SR is commonly used to operate a device, perform commands, or write without having to use a keyboard, mouse, or press any buttons.
- Simplifies the workflow and notes available immediately, less pointing, clicking, and scrolling.
- Increase the productivity, can help with menial computer tasks, such as browsing and scrolling.

V. CONCLUSION

From the problems faced by the speech recognition, speech recognition systems in order to be widely used still have a lot of areas for improvement. However, it is foreseeable in the near future that, with the voice recognition technology continues to progress, the speech recognition system will be more in-depth, the

application of speech recognition systems will be more extensive [8]. A variety of speech recognition systems will appear in the market, people will adjust their speech patterns to adapt to a variety of

recognition system. Human beings in the short term is also impossible to create a people comparable to the speech recognition system, to build such a system is still a big challenge facing humanity, we can only forward step by step direction to improve the speech recognition system.

VI. FUTURE SCOPE

In future speech recognition is used on all fields. The SR inputs are taken by our body movement and SR is implemented in the classroom and hospital over the year. Main SR is comes in mobile games, games are controlled by our voice instruction. A speech app is implemented in wheelchairs. Like this many future hopes are available in speech recognition.

VII. REFERENCES

- [1] Yogita H. Ghadage, Sushama D. Shelke - Speech to Text Conversion for Multilingual Languages, International Conference on Communication and Signal Processing, April 6-8, 2016, India
- [2] Neha Sharma, Shipra Sardana-developed Real time speech to text conversion system using bidirectional kalman filter matlab 2016 Intl. Conference on Advances in Computing, Communications and Informatics (ICACCI), Sept. 21-24, 2016, Jaipur, India
- [3] Sherif Yacoub, Steve Simske, Xiaofan Lin, John Burns- Recognition of Emotions in Interactive Voice Response Systems, presented at Eurospeech 2003, 8th European Conference on Speech Communication and Technology, September 2003, Geneva, Switzerland © Copyright Hewlett-Packard Company 2003
- [4] Jianliang Meng, Junwei Zhang, Haoquan Zhao- Overview of the Speech Recognition Technology, 2012 Fourth

- International Conference on
Computational and Information Sciences
- [5] Douglas O'Shaughnessy- Automatic speech recognition
- [6] Gulbakshee Dharmale , Dr. Vilas Thakare, Dr. Dipti D. Patil- *Intelligent Hands Free Speech based SMS System on Android*, International Conference on Advances in Human machine Interaction (HMI - 2016)
- [7] March 03-05, 2016, R. L. Jalappa Institute of Technology, Doddaballapur, Bangalore, India
- [8] B. Raghavendhar Reddy , E. Mahender- Speech to Text Conversion using Android Platform, International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com Vol. 3, Issue 1, January - February 2013, pp.253-258
- [9] Lavin JalanI, Rahul MasraniII, Roshan JadhavIII Tejaswini PalavIV- Speech recognition based learning system , *International Journal of Engineering Trends and Technology- Volume4Issue2-2013*