



Android Based Speech Recognition

Sonali Thite¹, Archana Gore², Sagar Yelmar³, Yogesh Lonkar⁴

Dept. of Computer-Science, Karmayogi Engineering College, Shelve Pandharpur¹.

Dept. of Computer -Science, Karmayogi Engineering College, Shelve Pandharpur²

Dept. of Computer -Science, Karmayogi Engineering College, Shelve Pandharpur³

Asst. Prof. Dept. of Computer -Science, Karmayogi Engineering College, Shelve Pandharpur⁴

ABSTRACT—

Every day a Smartphone user may look for a new application dedicated for his need. Android makes it easier for consumers to get and use new content and applications on their Smart phones. This paper presents an extremely on-demand, fast and user friendly Android Application voice recognition. Voice recognition is an alternative to typing on a keyboard. Put simply, you talk to the mobile and your words appear on the screen. For the past several decades, designers have processed speech for a wide variety of applications ranging from mobile communications to automatic reading machines. However, with modern processes, algorithms, and methods we can process speech signals easily and recognize the text. The software has been developed to provide a fast method of writing on mobile and can help people with a variety of disabilities. It is useful for people with physical disabilities who often find typing difficult, painful or impossible.

Voice-recognition software can also help those with spelling difficulties, including users with dyslexia, because recognized words are almost always correctly spelled. This paper is about to develop an on-line speech-to-text engine. The recognized text can be stored in a file. We are developing this on android platform using eclipse workbench. Our speech-to-text system directly acquires and converts speech to text. A speech-to-text system can also improve system accessibility by providing data entry options for blind, deaf, or physically handicapped users. Voice SMS is an application developed in this work that allows a user to record and convert spoken messages into SMS text message. User can send messages to the entered phone number. Speech recognition is done via the Internet.

Keywords— Android, DVM, Speech Recognition, Short Message Service (SMS), Neutral Network.

1. INTRODUCTION

People love their mobile phones because they can actually stay in touch wherever they are. That means not just for talking, but e-mailing, texting, and so on. For constructing with the growth of mobile phone technology. Android is the most popular cellular OS in the world nowadays. According to the recent movement Android covers nearly 50% of the whole smart phone market. Android is approved not only for its OS's great functionality but also act's are amazing. Now a day's there is a Smartphone user may look for a new application dedicated to the user's need. Android makes it simpler for consumers to get and use new



content and applications on their mobile. This application presents a really on-demand, fast and user friendly Android Application voice recognition, i.e. by using this application It will achieve more work with less time and do many at that time. For that One of the best application is the speech to text. In this application just talk in front of mobile and the speech can be converted into text right away. This type of app just frees our hands from getting busy typing SMS taking notes etc. It is been reportedly found that we often text message or do other work with our phone while driving.

By doing this sometimes concentration may be loss on the road and accident happens. So in this context this type of application will help us a lot in doing our misc works with phone even calling someone without losing our concentration on road. This Paper analyzes voice recognition technologies and their markets. It recognizes the fact that software and hardware technologies act in tandem, building the momentum needed for its success. An overview of the voice recognition industry precedes that review the main voice recognition categories, discuss top supplier market. This paper has the speech recognizing and speech synthesizing capabilities though it is not a complete replacement of what we call a NOTEPAD but still a good text editor to be used through voice. This software also can open windows based software or services such as Notepad, Ms-paint and more are new technologies, and the unique challenges faced by each category in the future.

Developing speech-enabled applications for mobile devices typically requires comprehensive Analysis of several speech processing engines and different architectural approaches before the desired application can be built. In addition, framework and tool support is often insufficient, which makes developing applications difficult: speech recognition systems? Usually provide a native low-level API or support Microsoft SAPI or Java SAPI (yes [Mice] and [Sun98b]). Developing a speech-based application for mobile devices requires work upfront, since mobile devices and speech recognition systems vary dramatically in their capabilities. While mobile devices can concisely be classified by their processing power, memory, operating system and wireless network speed it is a bit trickier for speech recognition engines.

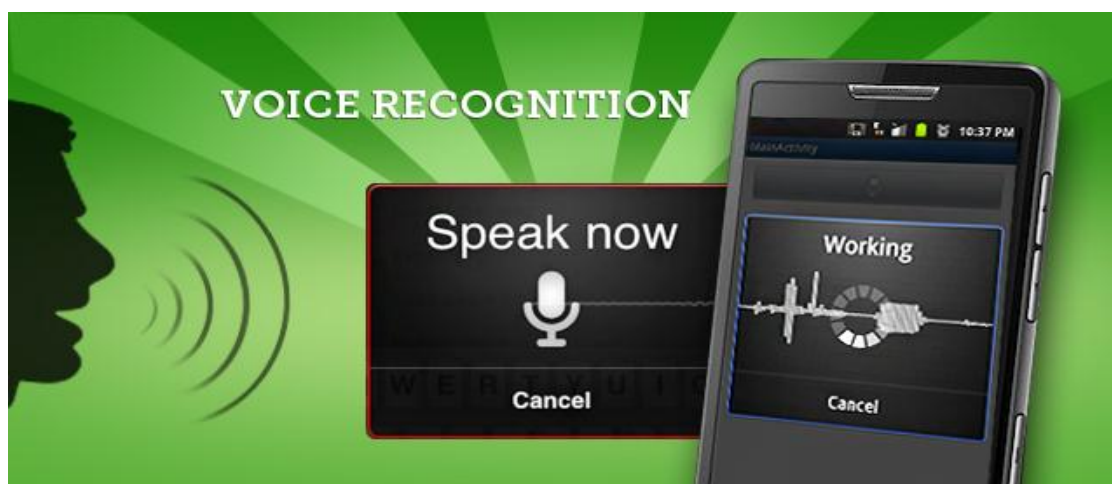


Figure 1 Conversion for speech to text



This paper presents a comprehensive approach that comprises a profound classification of speech recognition systems for mobile applications and a framework for mobile and distributed speech recognition. In this paper we are developing such application which is wholly useful for user such as give input as speeches word and converting it into text. As well as written Spelled word is converted into speech. This is very useful for the user for making many work at a time and your time will be save with more output. As well as it is very helpful for the blind person. And also helpful for people who having difficulties in expressing the ideas in speech.

2. ANDROID

Operating system has developed a lot in last 15 years. Starting from black and white phones to recent android mobile or mini computers, mobile OS has come far away. One of the most widely used mobile OS these days is ANDROID. Android is a software environment for mobile devices that includes an operating system, middleware and key applications [1]. In 2005 Google took over the good company Android Inc., and two years later, in collaboration with the group the Open Handset Alliance, presented Android operating system (OS). Main features of Android operating system are:

- Enables free download of development environment for new application development.
- Free use to all user and adaptation of operating system to manufacturers of android mobile devices.
- Equality of basic core applications and additional new applications in access to resources.
- Optimized use of memory and automatic control of applications which are being executed.
- Quick and easily development of applications using development tools in android and rich database of software libraries.
- High quality of audiovisual content, it is possible to use vector graphics, and most audio and video formats

Dalvik Virtual Machine (DVM) is the important layer of the android. Forms the main part of the executive the any application in system environment of android. Virtual machine is used to start the core libraries written in the Java programming language. Unlike Java's virtual machine, which is based on the stack, DVM bases on registry structure and it is intended for android mobile devices.

3. SPEECH RECOGNITION

Speech recognition is a process of converting speech to text as well as text to speech. It takes word as the input and converts it into text by using android SDK. Also if the written text is converted into speech. Accuracy of speech recognition systems differ in vocabulary size and



confusability, speaker dependence vs. independence, modality of speech (isolated, read or spontaneous speech), task and language constraints [2]. It is beneficial for the blind people such as they can only speak in front of mobile and it converted into text and send to their friend or any important message of any company by using android SDK.

3.1 Types of speech recognition

Speech recognition systems can be separated in several different classes by describing what types of utterances they have the ability to recognize the text. These classes are based on the fact that one of the difficulties of ASR is the ability to finding when a speaker starts and finishes an utterance. Most packages can fit into more than one class, depending on which mode they're using.

3.1.1 Isolated Words

Isolated word recognizers usually require each utterance to have quiet (lack of an audio signal) on BOTH sides of the sample window. It doesn't mean that it accepts single words, but does require a single utterance at a time. Often, these systems have "Listen/Not-Listen" states, where they require the speaker to wait between utterances (usually doing processing during the pauses). Isolated Utterance might be a better name for this class.

3.1.2 Connected Words

Connect word systems (or more correctly 'connected utterances') are similar to Isolated words, but allow separate utterances to be 'run-together' with a minimal pause between them.

3.1.3 Continuous Speech

Continuous recognition is the next step. Recognizers with continuous speech capabilities are some of the most difficult to create because they must utilize special methods to determine utterance boundaries. Continuous speech recognizers allow users to speak almost naturally, while the computer determines the content. Basically, it's computer dictation.

3.1.4 Spontaneous Speech

There appears to be a variety of definitions for what spontaneous speech actually is. At a basic level, it can be thought of as speech that is natural sounding and not recited. An ASR system with spontaneous speech ability should be able to handle a variety of natural speech features such as words being run together, "ums" and "ahs", and even slight stutters. This function performs the searching and sorting of the similar data items.



3.2 Speech to text conversion system

Elements of Speech to Text 1) Speech Acquisition 2) Speech to text conversion 3) SMS transmission. This system allows us to give our voice as input and produce text as an output. Systems use "training" where an individual speaker reads sections of text. These systems analyze the person's specific voice and use it to fine tune the recognition of that person speech, resulting in more accurate reproduction. The speech to text conversion system is the ability of Smartphone to identify the words which is in the spoken language and translation of it into the readable form that is in the form of text. Many speech recognition software available in the market which is available for limited vocabulary of words or if the words spoken very clearly.

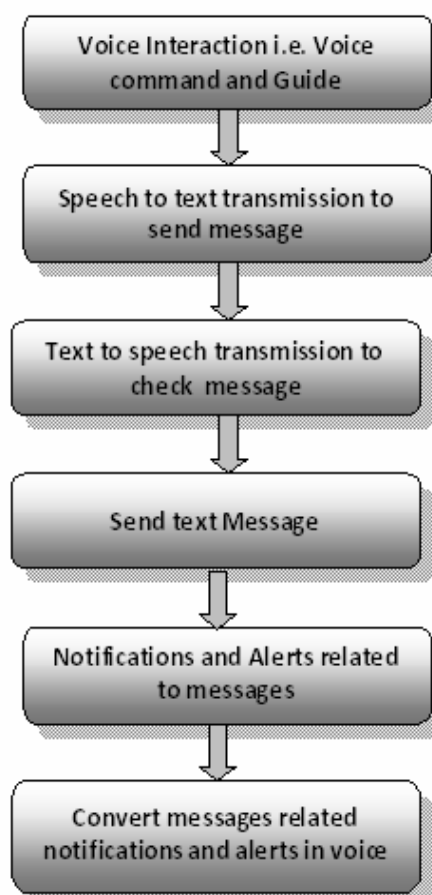


Figure 2 Application for sender side.



4. SYSTEM ANALYSIS

4.1 Existing System

In the past for managing the large amount data or typing the text is very time consuming. And it is not hand able to all people. Some people having the grammatical problem so, if they send message to other people, they cannot understand what they actually say. In market there are many new applications or services but for blind there is no any application or service. For that it is very useful for the blind people such as they only speck in front of their mobile and automatically text will written. Every day a Smartphone user may look for a new application dedicated for his need. Android makes it easier for consumers to get and use new content and applications on their Smart phones. And using this application we can do many work at a time such as if you are busy in work that time any important message or mail of company is coming that time we simply speck in front of your mobile and text will automatically written and Text is converted into voice.

4.2 Proposed System

In proposed system a speech to text and text to speech is implemented by using the eclipse and android SDK .By using that we can achieve more work at minimum time. Every day a Smartphone user may look for a new application dedicated for his need. Android makes it simple for consumers to get and use new content and applications on their Smart phones. It contain the three parts such as:

4.2.1 Android

Android is a software stack and mobile operating system that includes the operating system for portable devices, middleware, user interface, and a standard application , multimedia message service (MMS). Android developers were able to write applications in the Java ,it is a runtime library that can run the compiled byte code. In addition, it provides the required application through the Android Software Development Kit (SDK) to develop a variety of tools and APIs. Android also works on the Linux kernel and the Android system uses C as well as C++ libraries for developing application, etc. are included. Android, unlike existing Java virtual machines, uses an Java application made of Dalvik Virtual machine the important layer of android that runs on a separate process.

4.2.2 GPS(Global Positioning System)

GPS is a radio navigation system using satellites and it is developed by USA Department of Defense for military use navigation but it can be used by citizens with a limited range. It predicts radio coverage from satellites to a receiver, and then it shows the correct 3D location, speed and time.



4.2.3 LBS(Location-Based System)

LBS service indicates a wireless contents service that provides certain information based on the location change of the user. Designers of mobile handset have voluntarily tried to install LBS within their devices on mobile. However, LBS were beginning developed by telecommunication companies and mobile contents providers. Main benefit of the system is the fact that the users don't have to directly insert location as they move. GPS positioning technology is one of important technologies that allow simpler excess of wireless networking Service.

5. SPEECH RECOGNITION SOFTWARE

5.1 XVoice

XVoice is a dictation/continuous speech recognizer that can be used with a variety of XWindow applications for recognizing the text or dictating the text. This software is primarily for users.

5.2 ISIP

The Institute for Signal and Information Processing at Mississippi State University has made software for speech recognition engine is available. The toolkit combines a front-end, a decoder, and a training module. It is a utility toolkit. This software is primarily for developers.

5.3 Ears

Although Ears isn't fully developed, it is a good beginning point for programmers wishing to start in ASR. This software is mainly for developers.

6. FEATURE OF SPEECH RECOGNITION

Application Voice SMS accommodate direct voice word as input enabling user to record spoken information as text message, and send it as SMS /message.

- Able to write the text through both keypad and voice input.
- Voice recognition is for the different notepad commands such as open save and clear, exit.
- Open different windows software's, based on voice input.
- Requires less amount of consumption of time in writing text and sending.
- Provide important help for the people with disabilities.
- Lower operational costs.

7. COMPARISION

Last few year ago, there was mobile which having the keypad and by using that keypad we were send the SMS. If people are busy in the work suddenly SMS came then for replying it took too much time



to avoid this for that to create such application. For blind people there was no any service to write the message or see the keypad .As well as some person having the grammatical problem such as writing text.

Now a day there is a system such as speech recognition by using that we can speak only in front of your mobile and speak and the automatically text will written .As well as the text will automatically converted into speech. This application is also useful for the blind people for sending message. Who having grammatical problem for typing and also handicapped problem it will reduces by using this application.

This application was already developed for the desktop but it having some disadvantages such as mobility which is overcome by using this application.

8. CONCLUSION

In recent years, Text to Speech for disability and handicapped communication aids has become widely developed in Mass Transit. Text to Speech is also developing new applications outside the disability market in the world. This Paper work of speech recognition started with a briefly introduction the technology and its applications in different areas. At the later stage discussed different tools for bringing that idea into practical work. After the development of the software finally it was tested and results were discussed, few deficiencies factors were brought in front. After the testing work, advantages of the software were described and suggestions for further enhancement and improvement were discussed.

9. REFERENCE

- 1] M.J.F Gales, F. Diehl, C.K. Raut, M. Tomalin, P.C. Woodland, and K. Yu, "Development of a phonetic system for large vocabulary arabic speech recognition," in Proc. of ASRU, 2007.
- 2] Santos, J. ,Ciudad Universitaria, Madrid, Spain , Nimble , J. –"Text-to-speech conversion in Spanish a complete rule-based synthesis system"||Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '82.
- 3]"Android developers", <http://developer.android.com>
- 4] J.Tebelskis.Speech Recognition using Neural Networks, Pittsburgh: School of Computer Science, Camegie Mellon University, 1995.
- 5] "Speech recognition- The next revolution" 5th edition.
- 6]Ksenia Shalnova, "Automatic Speech Recognition"07 DEC2007Source:http://www.cs.bris.ac.uk/Teaching/Resources/COMS12303/lectures/Ksenia_Shalnova-Speech_Recognition.pdf



- 7] Brandon Ballinger, Cyril Allauzen, Alexander Gruenstein, Johan Schalkwyk, —On Demand Language Model Interpolation for Mobile Speech Input_, INTERSPEECH 2010, 26-30 September 2010, Makuhari, Chiba, Japan, pp 1812-1815
- 8] Ryuichi Nisimura, Jumpei Miyake, Hideki Kawahara and Toshio Irino, —Speech-To-Text Input Method For Web System Using JavaScript_, IEEE SLT 2008 pp 209-212
- 9] M. Tomalin, F. Diehl, M.J.F. Gales, J. Park & P.C. Woodland , —Recent Improvements to the Cambridge Arabic Speech-To-Text Systems_, ICASSP 2010 pp 4382- 4385
- 10] Santos, J. ,Ciudad Universitaria, Madrid, Spain ,Nombela, J. —“Text-to-speech conversion in Spanish a complete rule-based synthesis system”||Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '82.
- 11] Panikos Heracleous, Hiroshi Ishiguro and Norihiro Hagita, —Visual-speech to text conversion applicable to telephone communication for deaf individuals_ 18th International Conference on Telecommunication 2011. Pp 130-133
- 12]HomePage:<http://www.compapp.dcu.ie/~tdoris/Xvoice>/<http://www.zachary.com/creemer/xvoice.html>Project: <http://xvoice.sourceforge.net>
- 13] <http://www.isip.msstate.edu/project/speech>.
- 14] FTP site: <ftp://svr-ftp.eng.cam.ac.uk/comp.speech/recognition/>.