Intelligence Hands-Free Speech Based System on Android

Shaikh Shaheda Faiyaz
Department of Computer Engineering
Anjuman-I-Islam's Kalsekar Technical Campus
New Panvel, India
shaheedashaikh23@gmail.com

Khan Arshiya Rakeeb
Department of Computer Engineering
Anjuman-I-Islam's Kalsekar Technical Campus
New Panvel, India
arshiyak307 @gmail.com

Inamdar Mohsin Harun
Department of Computer Engineering
Anjuman-I-Islam's Kalsekar Technical Campus
New Panvel, India
inamdarm87@gmail.com

Prof. Ansari Mukhtar Amir
Department of Computer Engineering
Anjuman-I-Islam's Kalsekar Technical Campus
New Panvel, India
mukhtar.amir44@gmail.com

Abstract—SMS and Texting is an important feature of using Mobile phone and we also know that the mobile phone usage is spreading over the World rapidly and has gone through the number of features due to new techniques and Developers. This paper is based on creating an application that works on Google libraries and API's for conversion of Text-To-Speech and Speech-To-Text converter. It also works for Searching Contact with the Alphabets and Numeric read. Mainly the goal of the project is, it is for those who not be in the position of using mobile phones for texting ,surfing on web and dialing calls such kind of the communicating features ,so we called it as an Application that is useful for society .In other words,messaging can be completely based on speech recognition. The Application converts your text into the speech, speech into text, search a contacts manually from contact list or can be selected by taking name of a person which is voice based. We can select a multiple contacts for sending a message to multiple people at a time. Previous speech recognition system was difficult to use and it was having a lots of drawbacks, with leads in new technologies and techniques it is possible to generate a desire speech recognition system. This comes with lots of features by using an algorithm i.e Hidden Markov Model(HMM),which makes it possible to get a desire output. Another technologies are Android System, SR (speech recognition) libraries i.e. speech API's which is used in this paper.

Keywords—Speech-To-Text, Text-To-Speech Converter (Both Side), Contacts Selection with Numeric and Alphabets

I. INTRODUCTION

Now a day's Android try to make an applications more attractive for each and every categories of people. It try to cover every people in our society. Likewise, android try to improve a speech recognition system for comfort of people who are physically disabled, people who are having a less knowledge about language and to prevent a people from accidents. In this application user is able to access the services of smart phone with their SR (speech recognition) Command. This application is also developed for making a conversation in a very short period of a time. Speech can be processed faster than a text. Sender can send a message from their contact list as well as from a speaker, which automatically select a contact from a user list. This application contains a different services and functionality: speech to text, text to speech, and making a selection of contacts by using numbers, manually and by using a name of a recipient. Speech is a natural way of communication; conversations which are voice based are very clear and understanding. In message system their may be a misunderstanding between people just because accents are hidden from receivers. Sender should speak in a clear manner so that it can be understandable by a system. System uses different HMM models for every word of sentence. There are lots of HMM models are using for making a conversation possible. HMM models keep every state of words different from each other, so that it can make a correct sentence.

The Application which are going to make will use SR with Google server which uses HMM model. A detail description about working of this system is as follows: Initially speech taken as input and recorded by mice .When user speaks sounds will be fluctuating in a form signals, fluctuations of signal depends on users quality of voice. Input speech will divided into different set of words, which are in different set

of frames. First input is inputted, sound can be fluctuating set of signals which are recorded. Then these words process by system to execute it and convert it in desire text. Speech will recognize by using different methods such as feature extraction, acoustic model, dictionary, speech recognition algorithm and language model. Input speech first convert into digital signals, then it divided into small intervals. These digital signals then process by using an algorithm.

II. LITERATURE SURVEY

	Paper 1	Paper 2	Paper 3	Paper 4	Paper 5	Paper 6	Paper 7	Paper 8	OUR PROJJECT
Name of paper	Android Speech to Text Converter for SMS Application	Android text messaging application for visually impaired people	Intelligent Hands Free Speech based SMS System on Android	A REVIEW ON SPEECH TO TEXT CONVERSION METHODS	Free Model Speech Recognition System Using MFCC Model	Intelligent Hands Free Speech based SMS System on Android	Intelligent Hands Free Speech based SMS System on Android	Smart Voice Assistant: a universal voice control solution for non-visual access to the Android operating system	Intelligent Hands Free Speech based SMS System on Android.
Features									
1.speech to text	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2.Text to speech conversion.	No	Yes	No	No	No	Yes	Yes	Yes	Yes
3.contact selection with voice command	No	Yes	No	No	No	No	Yes	Yes	Yes
4.multiple recipient.	No	No	No	No	No	No	Yes	No	Yes
5.notification	No	No	No	No	No	No	Yes	No	Yes
Reminders	No	No	No	No	No	No	No	Yes	No
Drawbakcs.	Can't perform text perform.	Can't perform convertors at both the side.	Can't perform text conversion.	Can't perform Text to speech.	Can't support multiple language.	Cant't perform text Convertor.	Can't uses HMM model.	Cant't support multiple langauge.	Can't support multiple language.
Future scope	Text-speech -convertor.	Sender-side speech-to-text convertor.	Text-Recognsition.	Convertor to speech to text.	Can support multiple language.	Can perform text conversion.	Support HMM model.	Can use Multiple language.	Can use mulitple languagae for conversion.

III. OVERVIEW OF SYSTEM

For the visually impaired peoples it is not easy to handle that particular speech icon so here it is a problem to implement such kind of the application is accessible for impaired peoples. The technologies and algorithm use for this application are such as: HMM (Hidden Markov model), MFCC (Mel Frequency Cepstrum coefficient), android, forward algorithm, SMS manager class, java script, N-gram Database and Artificial Intelligence. Speech and text to be understood by the system is now popularly called as Speech recognition (SR).Different types of speech are as follows: Connected Words: Separate utterances together with minimum pause are input requirement of this system. Continuous Speech: A dictation by computer to the speaker, it is the most difficult recognizers to create. Spontaneous speech: Speaker's natural speech acts as the input for the system. It needs careful speaker, otherwise it generates excessive error.

A. Existing System

Speech recognition adds tremendous changes Into the classic keyboard input which leads to the manipulation of text is easier the the classic method. This application uses the Google API which uses the hidden Markov models (HMM) method. HMM use to send message to receivers in this application the speech is recorded and user selects the contacts from their list of contacts and then send a SMS to specific person. First Software was developed 1994 was dictation software, which is based on discrete speech. Discrete speech works slowly and not a natural means of communication, after every word spoke, it needs a pause.2nd speech based software developed by IBM which is based on continuous speech. Continuous speech based system was very flexible and a natural conversation, but it was too expensive and needs a costly PC's.

B. Proposed System

The application here will use the SR with Google server which uses HMM method. The description of how the speech recognized are as follows. Initially a speech inputted and sound fluctuates which can be represented by set of signals. Signals which are generates application is depends on quality of sound .If sound quality is high then signal level increases at a high level. Speech recorded a recorder .After a recording a done, speech divided into set of frames or words and every words and phrases works as independently Additional sounds comes with speech is filtered by a MFCC model, so that it can be easily understood by a system .Background voice and low quality voice all should be filter to convert it into desired text. Then algorithm is used for making a conversion from speech to text at sender site. These converted texts send to receivers.

IV. SYSTEM ARCHITECTURE

A. Working of the System

When application initiates, it ask contact number from user. User can provide contact number from two way; either it can be manual or voice based. Below contact filed their be a message field which we are going to sent a receiver. Sender speaks a message that would be converted at a sender site into text.

Message input is converted into speech to text and vice versa by using message converter. HMM is most successful and most flexible approach to speech recognition. It is used to send SMS,

HMM is state independent i.e current state, past state and future state are calculated independently. HMM method is basically used for recognition of speech. It converts a speech into text. This method is more flexible and efficient method if we used it properly. It keeps different states of HMM independent from each other to make a proper or desire pattern. MFCC is used for extracting a feature. This method is used to filter a speech, so that it must be understandable by a system.
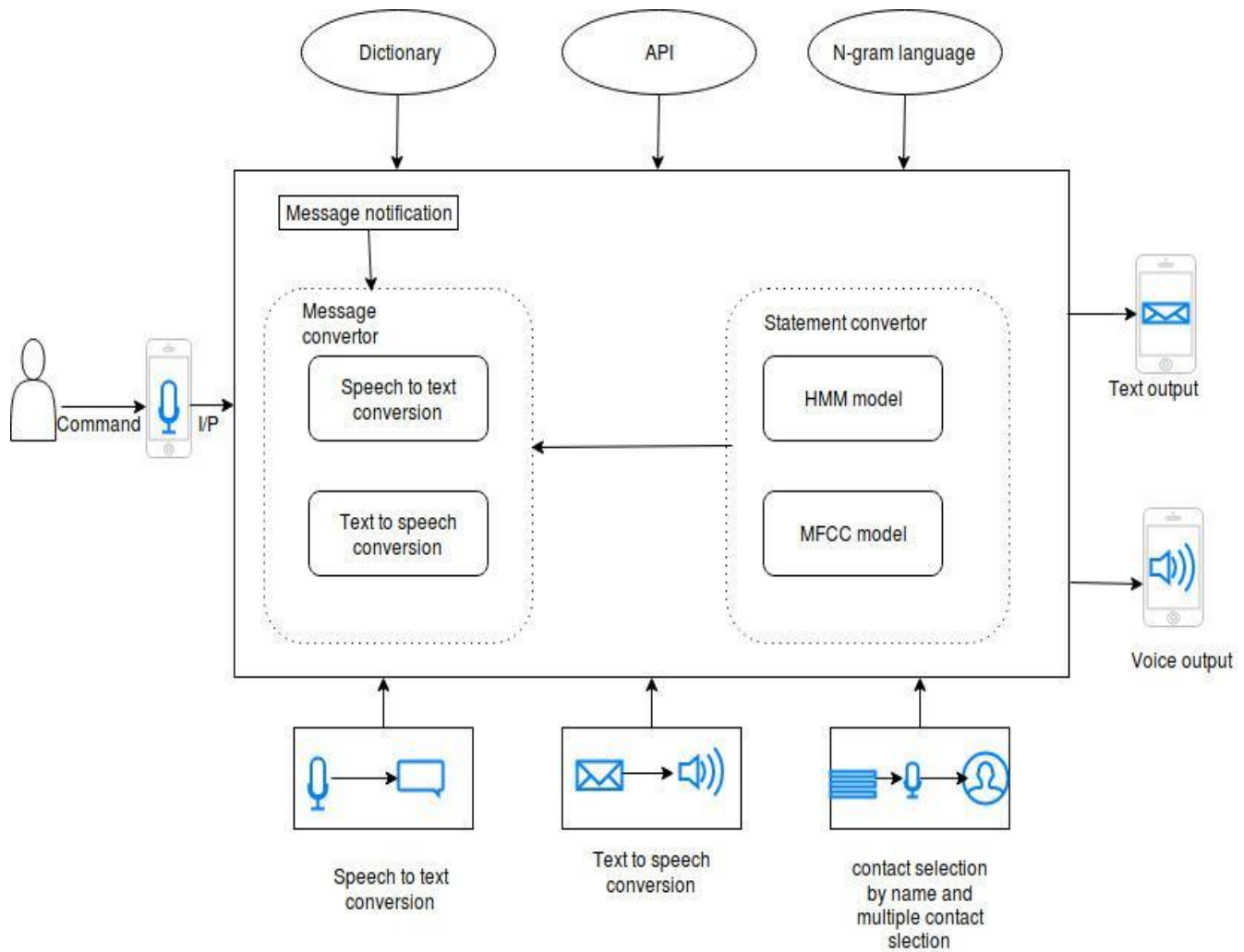
Fig. 1 System Architecture

Dictionary database contain all the words related to particular language. When user speaks, it checks words from dictionary and extract the work from database. N-gram Data-Base is performs error correction mechanism and data can be collected from world web pages and Internet document

B. Working of Speech-To-Text Recognition

First the speech is taken as the input, now it analyzed by the speech analysis with the help of speech dictionary or speech to text conversion database and then it further checking by the vocabulary database database by the selection of words ,phrases according to the sound and ascent of the user then it finally converts all the speech into the text and can send this speech by the text message.
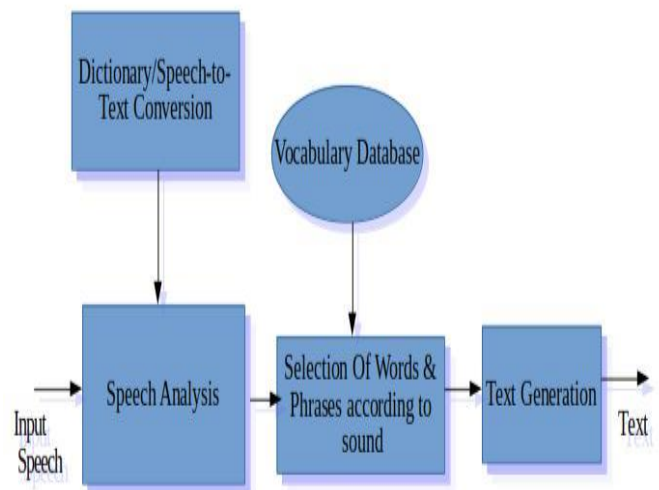

Fig 2: Speech-To-Text Recognition.

C . *Text-To-Speech Recognition*

First the input is taken as the text it analyzed by text analysis with the help of text dictionary or text to Speech converter and it sends to the speech database which selects the units of words spoken on the mike now it further sends speech generation module and on the basis of this process text is converted into the text.
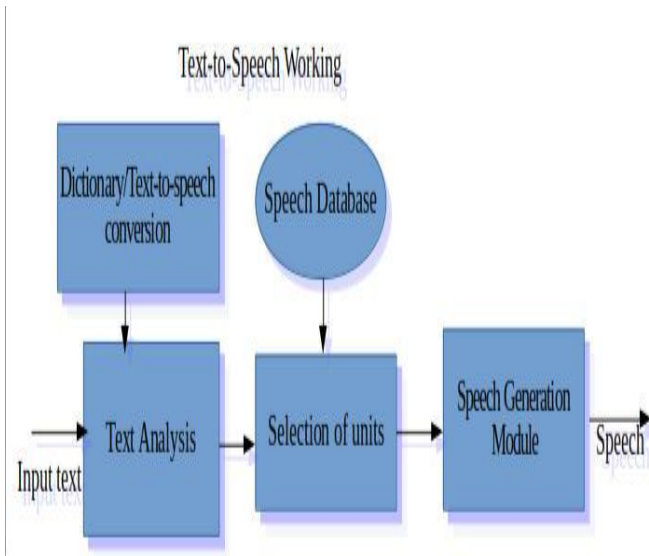

Fig 3: Text-To-Speech Recognition

D. Hidden Markov Model

1) Why Using HMM For Speech Conversion?

HMM are having very good calculations structures. This model are using in a wide range of applications, If we use this model properly it works in an efficient manner. It can perform a complex calculations are very rich in mathematical structure and hence can form the theoretical basis for use in a wide range of application.

2) Hidden Markov Model elements are as follows

HMM elements can be categorized as   follows:
1. Number of state N
2. Number of distinct observation symbol per state M, V
$= V_1, V_2, \cdots, V_M$

3. State transition probability, $a_i = P[q_{t+1} = S_i | q_t = S_j]$, $1 \leq i, j \leq N$
4. Observation symbol probability distribution in each state j, $B_j(K) = P[V_k$ at $t | q_t = S_j]$

5. The initially state distributions are: $\pi = \pi_i$ where $\pi_i = P[q_1 = S_i]$ $1 \leq i \leq N$ [2]
For appropriate value of N,M,A,B and π, HMM Model work as a generator. This generator will give an observation sequence.
$O = O_1 O_2 O_3 \cdots O_T$

3) Types of HMM:

1. Context-Independent Phoneme HMM:

* Number Of State : d-state HMM for each phoneme (d is normally equal to 3)

* Accuracy : not accurate in continuous speech recognition

* Compact : d-state HMM lead to less parameter to be calculated
* General : Yes, we can build HMM for new word using existing phoneme HMM

2. Context-Dependent Trip hone HMM:

* Number Of State : d-state HMM for each phoneme
* Accuracy : Accurate, as it has left-right phoneme relation.
* Compact : Each phoneme has immediate left-right relation, more parameter needs to be calculate
* General : Yes

3.Whole-Word HMM

* Number Of state : No phoneme generation

* assign number of state to model a word as whole

* Accurate: It is accurate and having large no of training data. It works for small vocabulary.

* Compact: It is not compact; it requires many states as vocabulary increases.

General: this HMM can't make new. Words HMM are having very good calculations structures. This model are using in a wide range of applications, If we use this model properly it works in an efficient manner. It can perform a complex calculations are very rich in mathematical structure and hence can form the theoretical basis for use in a wide range of application.
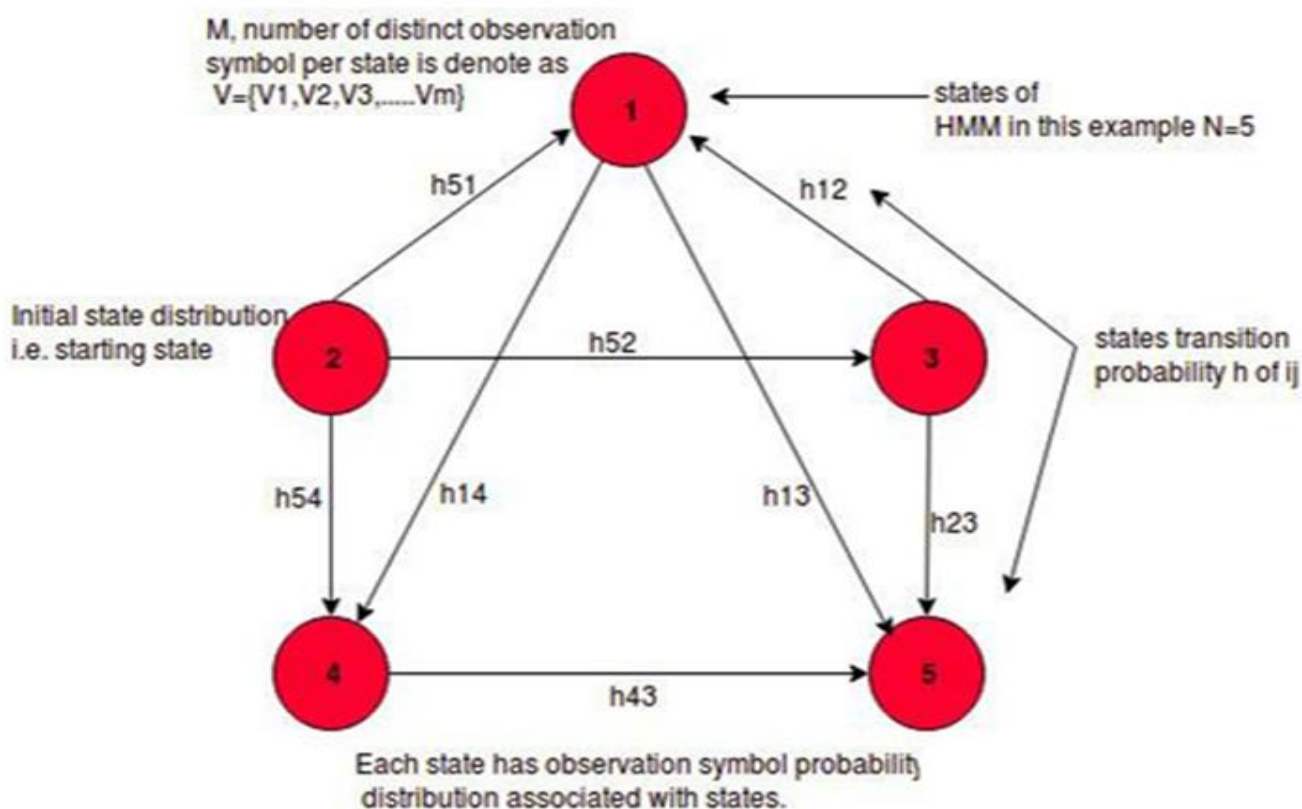
Fig 4: Working of HMM

E. Mel Frequency Cepstral Coefficients (MFCC)

MFCC is a method used in speech recognition for feature extraction. Before MFCC LPC were available,but it has lots of drawbacks which overcome by MFCC. It uses a frequency domain, which I s more accurate than a time domain. MFCC can be derived from FFT.

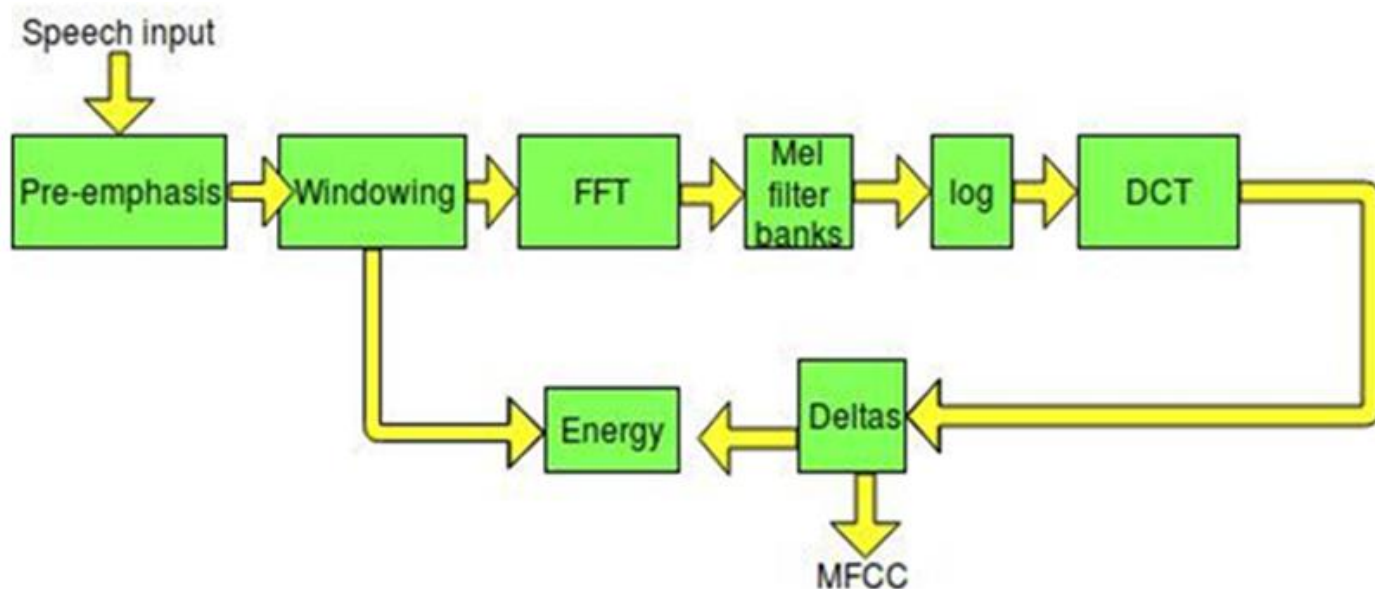Fig 5 shows Steps involved in MFCC feature extraction.

Fig 5: Working of MFCC

F. Technologies, Method and Algorithms:

- HMM[Hidden Markov Model]:

Most successful and most flexible approach to speech recognition. IT is used to send SMS. HMM feature is that it's state are independent i.e current state,past state and future state are calculated independently.

- MFCC[Mel Frequency Cepstrum coefficient]:

It extracts features and also select parametric representation

- N-gram DataBase:

It performs error correction mechanism and data can be collected from world web pages and Internet document.

- Android:

It is complete ,open and free platform.

- Forward Algorithm:

It is used for efficient output sequence.

- Viterbi Algorithm:

To get better observe state.

- Baum Welch Algorithm:

To choose computing parameters.

- AI(Artificial Intelligence) :

It is used for check validity of input speech for phone.

- SMS manager class:

It is provided by android to handle SMS default activity.

- JavaScript:

for recording panel we use javascript

- JavaApplet:

Pure JavaApplet is used button to record.

- Eclipse WorkBench:

It is used for text reconsecration.

G. Advantage

- It uses special technologies do it must be very fast and almost 100% correct to be understandable used SoundX

algorithm which makes use of NLP,SoundX selects best possible match words
- User can select a multiple contacts of same person to reduce multiple reduction
- It recognizes the speech to a more than 90% accuracy,delay form recognition is less than 100ns it gives a voice guidance for direction and destination of moving,
- It gives alarm services and calling services phone number can be selected manually or by using a voice.
- Timer for unread message, notification and alerts are provided when new message arrives,the timer will remind after a time period to read unread message.
- User can monitor their voice signal level by a red signal bar.

H. Disadvantages

It cannot perform a multiple language selection. User should know just one standard language.

This application is only based on English language.

IV. ACKNOWLEDGEMENT

REFERENCES

[1]. Intelligence Hands-Free Speech based system on android. Institute of Electrical and Electronics Engineers (issued on : 11 April 2016).

[2]. Android Speech to Text Converter for SMS Application (IOSR Journal of Engineering Mar. 2012, Vol. 2(3) pp: 420-423).

[3]. Android text messaging application for visually impaired people (IRACST – Engineering Science and Technology: An International Journal (ESTIJ), ISSN: 2250-3498,Vol.3, No.1,February 2013.

[4]. A REVIEW ON SPEECH TO TEXT CONVERSION METHODS International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 4 Issue 7, July 2015.

[5]. International Journal of Innovative Research in Science, Engineering and Technology (An ISO 3297: 2007 Certified Organization) VOL. 4, ISSUE 7, JULY 2015.

[6]. Intelligent Hands Free Speech based SMS System on Android(The Master of IEEE Projects Copyright © 2015 LeMeniz Infotech).

[7]. International Journal of Innovative Research in Science,Engineering and Technology(An ISO 3297: 2007 Certified Organization) Vol. 4, Issue 7, July 2015.

[8]. Brahim Patel, Dr. Y. SrinivasRao ,"Speech recognition using HMM with MFCC- an analysis using frequency spectral decomposion technique" , SIPIJ Dec 2010.

[9]. B. Raghavendhar Reddy, E. Mahender, "Speech to text conversion using android platform ", IJERA Feb-2013.

[10]. International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056 Volume: 04 Issue: 01 | Jan -2017.

[11]. Intelligent Hands Free Speech based SMS System on Android International Conference on Emerging Trends in Applications of Computing (ICETAC 2K17).

[12]. Text to Speech Conversion System using OCR International Journal of Emerging Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 5, Issue 1, January 2015)