

An Approach towards text messaging to voice message for Smart Android phone

1Mohammed Waseem Ashfaque, 2Sumegh Tharewal, 3Abdul Samad Shaikh, 4Sayyada Sara Banu, 5Shaikh Abdul Hannan

1Department of Computer Science & IT, College of Management and Computer Technology, Aurangabad, India

2Department of Computer Science & IT, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, India

3Greenfo Tech Aurangabad

4Department of Computer Science & IT Jizan University , Saudi Arabia

5Department. of Computer Science & Information Technology, Albaha University, Albaha, Saudi Arabia

Abstract – There are several researches’ results into the implementation of applications for voice messaging into text messages And since last few year developers have designed and developed varieties of applications with versatilities for mobile devices, automatic reading machines etc because speech and voice recognition system. Not been used in the field of computer and electronics. Due to the complexity and variety of speech signals and sounds. Now a day’s mobile phones are widely using and passing through under rapidly changes, and sms is one of the vital applications. In this paper we are attempting vice versa that is from text messaging to voice message. We believe that such type of efforts will be helpful into the interfacing between the communication devices with different aspects. We can process text and recognized into speech signals smoothly. In this paper we are taking samples of various languages to converts text messaging into the voice message. Such type of application can act as a bridge between two different entities for communication that is from text to voice messages. This paper describes three different levels.

First level is to convert the English speech into English text. This type of work can be helpful for those persons who blind by means of if they are handicap.

Second level provides regional language for conversion. It is beneficial for those regional persons they don’t know English language

Third level is that it is associated with mobile device (Android platform) that everyone can be used depend upon his need.

Keywords:- SMS; speech acquisition; Hidden Markov Model (HMM); HMM-based recognition, Android Os, DVM, Speech Recognition, NLP, TTS.

I. INTRODUCTION

Mobiles devices became compulsory part of our day today life, many of us need to make a call or massage at anytime from anywhere. Many of them needs their cell phones when they can’t do so e.g. At the time of driving, cooking accidents may occur because of this activity an speech to text converter for mobile design for this purpose so to avoid accidents. The study of speech to text conversion is from 1970s where the first experiment of phoneme- to-grapheme conversion, this conversion consists of segmentation of phoneme string into word. This work is again extended to stenotype-to-grapheme conversion. Voice massaging is slowly and gradually reducing the importance of text massaging because it is safer to massage at the time of cooking and driving. This paper introduces an idea about the speech to text conversion for SMS application. This software enable user to send the SMS without using keypad with fully spelled word. Now let’s limit our focus towards short massage system it is text messaging service component of phone, using standardized communications protocols that allow the exchange of short text messages between mobile phone devices. SMS text messaging is the most widely used data application in the world, with 2.4 billion active users, or 74% of all mobile phone subscribers [1].Smart phones offer customer enhanced methods to interact with their phones but the most natural way of interaction remains speech. Market for smart mobile phones provides a number of applications with speech recognition implementation. Google's Voice Actions and recently iphone's Siri 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 tasks. Both Siri and Voice Actions require an active connection to a network in order to process requests and most of Android phones can run on a 4G network which is faster than the 3G network that the iPhone runs on. There is also an issue of availability, Voice Actions are available on all Android devices above Android 2.2, but Siri is available only for owners of the iPhone 4S. The Siri's advantage is that it can act on a wide variety of phrases and requests and can understand and learn from natural language, whereas Google's Voice Actions can be operated only by using very specific voice commands. In this work we have developed an application for sending SMS messages which uses Google's speech recognition engine. The main goal of application txt SMS is to allow user to input spoken information and send voice message as desired text message.[2].In this work, TTS conversion is performed on android platform. TTS is one of the major applications of NLP. The NLP module of general TTS conversion system consists of preprocessor, text analyzer, morphological analyzer, contextual analyzer, syntactic prosodic parser, letter to sound module and prosody generator. A text analyzer block is composed of a pre-processing module, which organizes the input sentences into manageable lists of words. It identifies numbers, abbreviations, idiomatic and transforming them into full text when needed. A morphological module performs task to propose all possible part of speech categories for each word taken individually, on the basis of their spelling. Inflected, derived and compound words are decomposed into their elementary graphemes units by simple regular grammars exploiting lexicons of stems and affixes. The contextual analyzer module considers words in their context, which allows it to reduce the list of their possible part of speech of neighboring words. Finally a syntactic praser, which examines the remaining search space and finds the text structure which more closely relates to its expected prosodic realization. The presented research aims at developing a working model of speech synthesizer for English as well as for regional language for android based mobile phones along with creation of a light weight English speech database for android mobiles. The work will create a user friendly environment to present the application effectively. The major requirement about implementing the work, is , we need a library of English text to its phoneme equivalent. There are number of such libraries available online. We can get these libraries by performing an online search. The TTS conversion is implemented for the mobile android environment. It is under the NLP and provides easy communication [3].

II. CASE STUDY USING ANDROID

Android is free to use, improvable, and designed with multiple hardware implementations, it is open source and the code was released under Apache license. It is a software bunch comprising not only operating system but also middleware and key applications. It is one of the most widely used OS. Android has a large community for developing their own applications, written primarily on customized Java Programming language. Anyone with basic knowledge in java can start developing Android applications. Android OS version varies from 1.0 to 4.2 Jellybean (July 2012). Notice carefully you can see all the OS names are desserts, but still the naming reason is not released by Google and the names are alphabetically ordered, starting from 1.5 Cupcake to 4.2 Jellybean (C-J).

Kernel in android OS is Linux, it makes use of Linux kernel 2.6 and 3.x (Android 4.0 onwards), so Android is a Linux based operating system. Android comes along with open-source API libraries for application development including SQLite, Webkit, OpenGL and a media manager. Android is a software environment for mobile devices that includes an operating system, middleware and key applications [4]. In 2005 Google took over company Android Inc., and two years later, in collaboration with the group the Open Handset Alliance, presented Android operating system.

Main features of Android operating system:-

 Enables free download of development

Environment for application development.

 Free use and adaptation of operating system to manufacturers of mobile devices. Equality of basic core applications and additional applications in access to resources.

 Optimized use of memory and automatic control of applications which are being executed.

 Quick and easy development of applications using development tools and rich database of software libraries.

 High quality of audiovisual content, it is possible to use vector graphics, and most audio and video formats.

 Ability to test applications on most computing platforms, including Windows, Linux…

The Android operating system (OS) architecture is divided into 5 layers (fig. 1.). The application layer of Android OS is visible to end user, and consists of user applications. The application layer includes basic applications which come with the operating system and applications which user subsequently takes. All applications are written in the Java programming language. Framework is extensible set of software components used by all applications in the operating system. The next layer represents the libraries, written in the C and C + + programming languages, and OS accesses them via framework. Dalvik Virtual Machine (DVM), forms the main part of the executive system environment. Virtual machine is used to start the core libraries written in the Java



Fig1: Android Architecture & Linux Kernel & Libraries

III. LITERATURE REVIEW

In year 2012, sheilly padda and nidhi performed a work on ‘A step towards making an effective text to speech conversion system’. In this paper TTS conversion for Punjabi language is discussed [5].Aidan kehoe performed a work on ‘designing help topics for use with text to speech’. This paper proposes a number of guidelines to assist in the creation and testing of help material that may be presented to users via speech synthesis engines. The paper also provides a brief overview of an ongoing project that provides online help using speech technology. In this paper, the initial work involves the creation of a system to present help topics to a user using TTS. Later, speech recognition capabilities will be integrated to allow more dynamic interaction between the help system and the user.[6]. Eyub B. Kaise, publish a paper on ‘concatenative speech synthesis for Amharic using unit selection method’. In this paper algorithms and methods that address critical issues in developing a general Amharic text to speech synthesizer are proposed [7]. Erik blankin ship, performed a work on ‘tools for expressive text to speech markup’. This paper describes handicapped accessible text to speech markup software developed for poetry and performance. The author developed an application, called poet shop, to let a user graphically modify volume and pitch contours [8]. Ankit Asthana performed a work,” IOS 5, Android 4.0 and Windows 8 – A Review”. This paper presents a brief review ofthree top mobile operating systems viz. IOS 5 from Apple, Android 4.0 from Google and Windows 8 from Microsoft. The salient key new features introduced in IOS 5 are also described in some detail [9].

IV. SPEECH RECOGNITION

Speech recognition is very important part of this system in this phase speech samples are obtained from speaker at real time and stored for preprocessing. For speech recognition we require microphone to receive voice speech signals, Speech acquisition can be easily done by the microphone present in the mobile phone, In the acquisition phase the different M/C is depends upon the its own configuration, hence there is need to store the sample of different users to make system more compatible to any type of voice. To recognize the speech HMM-based automatic recognition was conducted. For continuous phoneme recognition, an 86% phoneme correct was achieved for the normal-hearing. To achieve speech preprocessing sphinx frame work is used this is the best tool found to acquiesce speech signals. Sphinx is design with high flexibility modularity. Recognition or pattern classification is the process of comparing the unknown test pattern with each sound class reference pattern and computing a measure of similarity (distance) between the test pattern and each reference pattern.. The digit is recognized using a maximum likelihood estimate, such as the Viterbi decoding algorithm, which implies that the digit whose model has the maximum probability is the spoken digit. Preprocessing, feature vector extraction, and codebook generation are same as in HMM training. The input speech sample is preprocessed and the feature vector is extracted. Then, the index of the nearest codebook vector for each frame is sent to all digit models. The model with the maximum probability is chosen as the recognized digit.[1]

V. SPEECH PREPROCESSING

The voice which is taken at the real time will require noise free speech signals now to reduce noise we need to consist of background noise that need to be removed. The preprocessing reduces the amount of efforts in next stages. Input to the speech preprocessing is speech signals which then converted into speech frames and gives unique sample Steps: 1. The system must identify useful or significant samples from the speech signal. To accomplish this goal, the system divides the speech samples into overlapped frames. 2. The system performs checks for the voice activity using endpoint detection and energy threshold calculations. 3. The speech samples are then passed through a pre-emphasis filter. 4. The frames with voice activity are passed through a Hamming window. The system performs autocorrelation analysis on each frame. 6. The system finds linear predictive coding (LPC) coefficients using the Levinson and Durbin algorithm. We apply a Hamming window to each frame to minimize signal discontinuities at the beginning and end of the frame [1].

VI. HMM TRAINING

An important part of speech-to-text conversion using pattern recognition is training. Training involves creating a pattern representative of the features of a class using one or more test patterns that correspond to speech sounds of the same class. A model commonly used for speech recognition is the HMM, which is a statistical model used for modeling an unknown system using an observed output sequence. The system trains the HMM for each digit in the vocabulary using the Baum-Welch algorithm. The codebook index created during preprocessing is the observation vector for the HMM model [1].

Speech recognition for application Voice SMS is done on Google server, using the HMM algorithm. HMM algorithm is briefly described in this part. Process involves the conversion of acoustic speech into a set of words and is performed by software component. Accuracy of speech recognition systems differ in vocabulary size and confusability, speaker dependence vs. independence, modality of speech (isolated, discontinuous, or continuous speech, read or spontaneous speech), task and language constraints [10].

Approach to proposed Work & methodology:-

DESIGN MODULE: - This project converts text message into speech. At the run time text data is acquire in the form of sms from storage and converted into speech frames the speech frames are then pass for preprocessing and after preprocessing of the sample frames HMM-based training is applied on speech frames

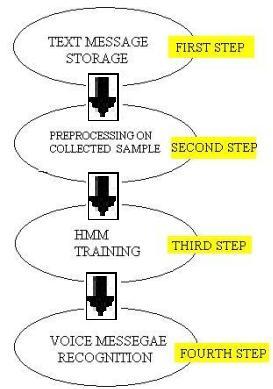


Fig:-2 text to Speech conversion system

Methodology:-

1. Speech synthesis techniques will be used in order to get the naturalness quality in the synthetic speech.

2. The process of the English language can be Used as the basic unit for speech synthesis.

3. Speech database for the English language will be developed using phoneme.

4. Speech database for the Hindi language will be developed using phoneme.

5. Speech database for the Urdu language will be developed using phoneme

4. The input text will be separated into English Phoneme.

5. Phonemes will be searched in the database and corresponding phonemes sounds will be Concatenated to generate synthesized output Speech.

ALGORITHM FOR ENGLISH TEXT TO ENGLISH SPEECH CONVERSION:-

The above flow chart shows the basic flow of work for English text to English speech conversion. By starting the work, get the input text in the English language. After getting the input text, perform the separation of the English words from the text. Once the text have been separated, than perform the library lookup to get the phonetic equivalent of the text .Later, arrange these entire phonetic equivalents in a series respective to text. After this arrangement of phonetic equivalents, the speech synthesis is performed and the speech quality is maintained.

Lastly, the application is implemented in android Environment. Last step is the checking the accuracy of the implementation. The whole work is done on android environment using android 4.2 versions. The layout of research proposal is defined in terms of given flowchart.

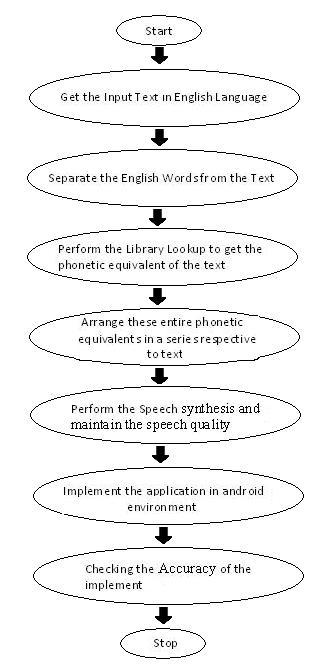
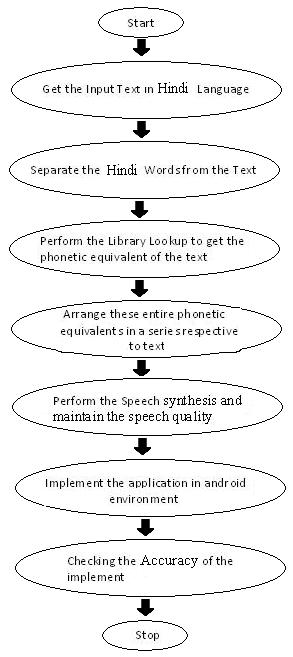




Figure: 3 English Text to English speech conversion

ALGORITHM FOR HINDI TEXT TO HINDI SPEECH CONVERSION:-

The above flow chart shows the basic flow of work for Hindi text to Hindi speech conversion. By starting the work, get the input text in the Hindi language. After getting the input text, perform the separation of the Hindi words from the text. Once the text have been separated, than perform the library lookup to get the phonetic equivalent of the text .Later, arrange these entire phonetic equivalents in a series respective to text. After this arrangement of phonetic equivalents, the speech synthesis is performed and the speech quality is maintained. Lastly, the application is implemented in android Environment. Last step is the checking the accuracy of the implementation. The whole work is done on android environment using android 4.2 versions. The layout of research proposal is defined in terms of given flowchart.



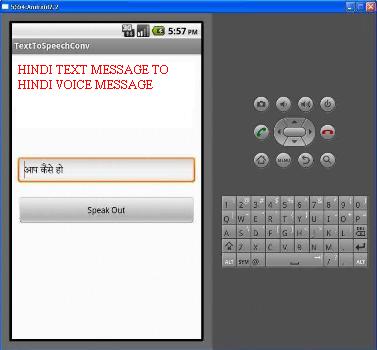


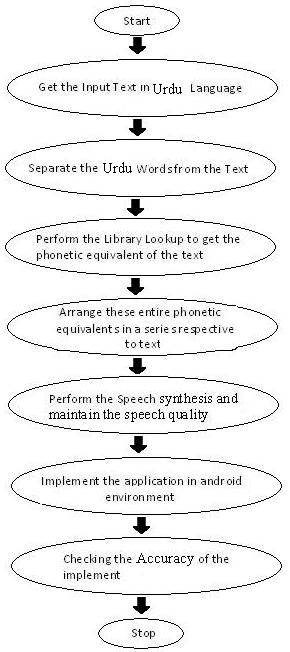
Figure: 4 Hindi Text to Hindi speech conversion

ALGORITHM FOR URDU TEXT TO URDU SPEECH CONVERSION:-

The above flow chart shows the basic flow of work for Urdu text to Urdu speech conversion by starting the work, get the input text in the Urdu language. After getting the input text, perform the separation of the Urdu words from the text. Once the text have been separated, than perform the library lookup to get the phonetic equivalent of the text .Later, arrange these entire phonetic equivalents in a series respective to text. After this arrangement of phonetic equivalents implemented in android Environment. Last step is the checking the accuracy of the implementation. The whole work is done on android environment using android 4.2 versions. The layout of research proposal is defined in terms of given flowchart



Figure: 5 Urdu Text to Urdu speech conversion



VII. CONCLUSION

In the proposed work we have suggest and design the phenomenon of converting mobile text message into voice messaging services. And the speech synthesis is performed, speech quality is also maintained. Outcome of result we have concluded that English text to English voice messaging, Regional language (Hindi) text to Hindi voice messaging and Urdu text messaging to Urdu voice messaging And Urdu and Hindi conversation gives satisfactorily result and voice is audible with great punctuation accuracy using natural language processing (NLP) we proposed the algorithmic steps to converting for each language. We took TTS for solving the problem of sound equivalent for half words because its quit useful for blind peoples to read the documents but one drawback we have found in our proposed model is that its system voice dependent because Application are only capable to recognized single word only and the output speech may same for every person the current work has been done for the English, Hindi, and Urdu as a sample, but it can also be done for other regional language

REFERENCES

[1] Ms. Anuja Jadhav, Prof. Arvind Patil International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com Vol. 2, Issue 2, Mar-Apr 2012, pp.1126-1128 1126| P a g e A Smart Texting System For Android Mobile Users Ms. Anuja Jadhav \* Prof. Arvind Patil\*\* \* (M.Tech. Student Y.C.C.E. Nagpur India) \*\* (Y.C.C.E., Nagpur, India).

[2] B. Raghavendhar Reddy, E Mahender / International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com Vol. 3, Issue 1, January -February 2013, pp.253-258 253 | P a g e Speech toText Conversion using Android Platform B. Raghavendhar Reddy#1, E. Mahender\*2 #1Department of Electronics Communication and Engineering Aurora’s Technological and Research Institute Parvathapur, Uppal, Hyderabad, India. 2Asst. Professor Aurora’s Technological and Research Institute Parvathapur, Uppal, Hyderabad, India.

[3] Swati Ahlawat ,Rajiv Dahiya Electronics and Communication Engineering PDM College of Engineering, bahadurgarh, IJCSMS International Journal of Computer Science & Management Studies, Vol. 13, Issue 05, July 2013 ISSN (Online): 2231 –5268, A NOVEL APPROACH OF TEXT TO SPEECH CONVERSION UNDER ANDROID ENVIRONMENT.

[4] “Android developers”, http://developer.android.com

[5] Er. SheillyPadda, Er. Nidhi “A Step towards Making an Effective Text to Speech Conversion”,IJERA’ ISSN:2245-9622.

[6] Aidan Kehoe,”Designing Help Topics for use with text to speech”, SIGDIC’06, October 18- 20,2006,myrtle beach, south carolina,USA.

[7] Eyob B. Kaise,” Concatenative speech synthesis for Amharic using unit selection method”, MEDES’12, october 25-31,2012,addis ababa,ethiopia.

[8] Erik Blankinship, “Tools for Expressive Text to Speech Markup” UIST’01 orlando FLO, november 11-14, 2001

[9] Ankit Asthana,” IOS 5, Android 4.0 And Windows 8 – A Review.

[10] J. Tebelskis, Speech Recognition using Neural Networks, Pittsburgh: School of Computer Science, Carnegie Mellon University, 1995.

🞜🞜🞜