Review on Text to Speech System

[1]Harshada Parab, [2]Nikita Shetye, [3]Seeba Bhatkar, [4]Tejaswini Bhosale, [5]Dr. Sharda Chougule
[1,2,3,4]UG Student, [5]HoD EXTC, FAMT, Ratnagiri (India).
[1]parab272@gmail.com, [2]nikitashetye97@gmail.com, [3]cbabhatkar1998@gmail.com, [4]bhosaletejaswini12@gmail.com,
[5]shardavchougule@gmail.com

Abstract: In today's world, where one cannot live without communication, automation plays a vital role. The world is moving towards digitization, so are the means of communication. But illiteracy and visual impairment are the most important barrier which cease the communication. In such case Text To Speech System plays a vital role in allowing a smooth communication. As the name indicates it converts written text documents into human speech. Text To Speech conversion is the process in which input text is analyzed, process and understood and later on this text is converted into digital audio and then spoken. Text To Speech allows user to see text and hear it read aloud simultaneously. There are many apps available, but typically as text appears on the screen, it's spoken. Some software uses a computer generated voice and other use a recorded human voice.

Keywords — automation, communication, digital audio, digitization, Text To Speech conversion system,

I. INTRODUCTION

Now-a-days, we hear many news related to misuse of illiterate people or visually impaired people, So giving due importance to such people we have designed & implemented a project which will convert given text documents in written form into human voice. i.e. audible form. This is basically process of making PCs to talk. There are various processes involved in this conversion and many more which will be much more cleared later on. We are going to implement Text To Speech system for Marathi language. Along with above mentioned two applications of Text to Speech system there are various other applications like it helps people with learning disabilities who have difficulty in reading large amounts of text due to dyslexia or other problems. Also TTS allows people to enjoy, and also provides an option to the go, taking content away from the computer screen and into any environment that's convenient for the consumer.

II. PROBLEM STATEMENT

Development of Text to Speech Conversion System in Marathi Language.

Text-To-Speech (TTS) is a technology that converts a written text into human understandable voice. A TTS synthesizer is a computer based system that can be able to read any text aloud that is given through standard input devices. Text-To-Speech is a process through which input text is analyzed, processed and "understood", and then the text is rendered as digital audio and then "spoken".

III. REVIEW ON REASEARCH PAPER

Text to speech system convert the written text to a spoken words or voice. In text analysis written documents or text is analyzed into the list of words. In text normalization the main list of words from analysis data is converted into pronounceable form. Then to transform the text to pronounceable form text normalization is used. It is used to identify the pauses and punctuation marks. Then conversion of given text into a sequence of synthesis unit is done by text processing. Then acoustic processing is done. In this processing voice characteristic of a person will be spoken out. There are three types of voice characteristic such as concatenative synthesis, forman ant synthesis, articultory synthesis. To generate user specified sequences of sound. Concatenative synthesis is used in speech synthesis. Articultory synthesis is used for synthesizing speech depending on models of human vocal tract. Sequence of synthesis units is formed by the conversion named as text processing. The main objective is to identify punctuation marks and pauses between the words and then it is converted into the spoken format.

IV. WORKING

The main function of text-to-speech (TTS) system is to convert an arbitrary text to a spoken Waveform. This task generally consists of steps, i.e., text analysis, text normalization, text processing, acoustic processing and speech generation. Text analysis part is preprocessing which analyze the input text and organize into manageable list of words, Text normalization is the transformation of text to pronounceable form. The main objective of this process is to identify punctuation marks and pauses between words,

Text processing is the conversion of the given text into a sequence of synthesis units, Acoustic processing –the speech will be spoken out the voice characteristic of a person like three types such as Concatenative synthesis, Forman ant synthesis, Articulatory synthesis whereas speech generation is generation of an acoustic wave form corresponding to each of these units in the sequence. A synthesizer can integrate a representation of the vocal region and other human voice characteristics to produce a "synthetic" voice output. The quality of a speech synthesizer is judged by its comparison to the human voice, and by its ability to be understood.

This text-to-speech program allows people with visual impairments or interpretation disabilities to listen to written works on a home computer. A text-to-speech system is composed of two parts known as front-end and back-end. The front-end consist of two tasks. First task is that, it converts raw text having symbols like abbreviations and numbers into the equivalent of written-out words. This process is called pre-processing text normalization, or tokenization. The front-end then allots phonetic transcriptions to each word, and breaks up and marks the text into prosodic units, like sentences, clauses, and phrases. The process of assigning phonetic transcriptions to words is known as text-to phoneme or grapheme-to-phoneme conversion. Phonetic transcriptions and Prosody information together make up the symbolic linguistic representation which is output by the frontend. The synthesizer is often known to as the back-end, and then converts the symbolic linguistic representation into necessary sound.

V. BLOCK DIAGRAM

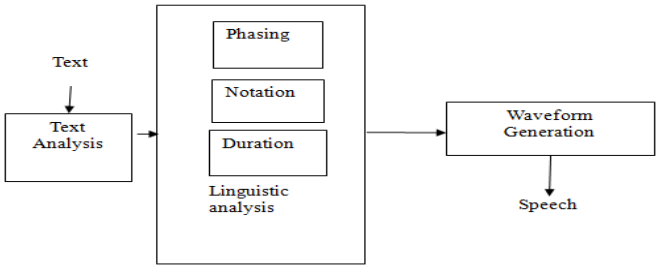

Figure1. Text processing

VI. SOFTWARE USED

Festival -2.4:-

Festival offers a general framework for building speech synthesis systems as well as including examples of various modules. As a whole it offers full text to speech through a number APIs: from shell level, though a Scheme command interpreter, as a C++ library, from Java, and an Emacs interface.

VII. WORK CARRIED OUT

To do the Text to speech system (TTS) project, firstly we search to know about the concept of TTS system, how it is actually work and how to implement it. Then in group of four we individually understand the concept of TTS system, then we discussed about the TTS system.

Then we individually search the review papers of TTS system. After reading all the review papers we get the idea about TTS system. In some review papers TTS system is implemented using hardware like OCR and in some review papers it is implemented using software like MATLAB software.

According to the some review papers we prepared the literature review and give the presentation on the literature review.

VIII. CONCLUSION

In this phase we carried out literature survey by studying the research papers related to our topic and decided software which we are going to use.

ACKNOWLEDGMENT

It gives us an immense pleasure to present the report of our project here. It has been quite experience, facing a number of problems at stages and coming up with appropriate solutions, at time the discussion amongst us or suggestions from our friends and teachers.

We thank our guide Dr. S. V. Chougule, Associate Professor, Electronics and Telecommunication Engineering, in the best possible way. Without her guidance it wouldn't have been possible to reach this stage. We are very grateful for his support and motivation.

We express our gratitude to Dr. S. V. Chougule, Associate Professor and Head of the Department, Electronics and Telecommunication Engineering for her invaluable suggestions and constant encouragement.

Lastly, we would like to put our thanks on record to the teaching and non-teaching staff for rendering their support directly or indirectly.

REFERENCES

[1] R.Shantha Selva Kumari,R.Sangeetha,''Conversion of English Text To Speech using indian language",

[2] Sim Liew Fong, Abdelrahman Osman Elfaki, Md Gapar bin Md johar and kevin Loo Tow Aik, Mobile Language Translator, 5th Malaysim Conference in software Engineering (Misses); 2011.

[3] Kaveri Kamble, Ramesh Kagalkar," A Review: Translation of Text to Speech Conversion for Hindi Language", International Journal of Science and Research (IJSR) ISSN

[4] Rhead, Mke, "Accuracy of automatic number plate recognition (ANPR) and real world UK number plate problems." IEEE International Carnahan Conference on Security Technology (ICCST), 2012.

[5]Mukta Gahlawat,et al, "Natural Speech Synthesizer for Blind Persons Using Hybrid Approach" Procedia Computer Science Volume 41, 2014, Pages 83–88

[6] Archana A, Shinde D. Text preprocessing and text segmentation for OCR. International journal of Computer Science Engineering and Technology. 2012:810-12.

[7] Bhargava A, Nath KV, Sachdeva P, Samel M. Reading assistant for visually impaired. International journal of current Engineering and Technology. 2015 Apr; 5(2).

[8] Helal Uddin Mullah, "A Comparative Study of Different Text-to-Speech Synthesis Techniques", International Journal of Scientific & Engineering Research, Volume 6, Issue 6, June-2015

[9] R.Shantha Selva Kumari, et, al, "Conversion of English Text-to-Speech (TTS) Using Indian Speech Signal" International Journal of Scientific Engineering and Technology, Volume No.4 Issue No.8, pp: 447-450

[10] Prabhsimran Singh et.al, "A Text to Speech (TTS) System with English to Punjabi Conversion", November 2014

[11] Aiswarya TS,et.al , " Malayalam Text To Speech Synthesis with Emotions", International Journal of Advance Research in Computer Science and Management Studies Volume 3, Issue 5, May 2015

[12] O.O. Adeyemo and A. Idowu, "Development and integration of Text to Speech Usability Interface for Visually Impaired Users in Yoruba language", African Journal of Computing & ICT, Vol 8. No. 1 – March, 2012

