International Journal of Advance Research in Science and Engineering
Volume No. 11, Issue No. 06, June 2022
www.ijarse.com
IJARSE
ISSN 2319 - 8354

Dictate Web Application

Ms. Rutuja Annaso Gouraje, Ms. Sonali Dattatray Jadhav,

Ms. Dnyaneshwari Ajitsinh Suryavanshi, Ms.Nafisa Imtiyaj Mulani,

Prof. Swati A Shinde

Department of Computer Engineering, ATS's Sanjay Bhokare Group of Institutes, Sangli India

ABSTRACT

Over the past few years, Cell Phones have become an indispensable source of communication for the modern society. We can make calls and text messages from a source to a destination easily. It is known that verbal communication is the most appropriate modem of passing on and conceiving the correct information, avoiding misquotations. To fulfil the gap over a long distance, verbalcommunication can take place easily on phone calls. A path-breaking innovation has recently come to play in the SMS technology using the speech recognition technology, where voice messages are being converted to text messages. Quite a few applications usedto assist the disabled make use of TTS, STT, and translation. They can also be used for other applications, taking an example: Siri an intelligent automated assistant implemented on an electronic device, to facilitate user interaction with a device, and to helpthe user more effectively engage with local and/or remote services makes use of Nuance Communications voice recognition and text- to-speech (TTS) technology.

1. INTRODUCTION

Over the past few years, Cell Phones have become an indispensable source of communication for the modern society. We canmake calls and text messages from a source to a destination easily. It is known that verbal communication is the most appropriate modem of passing on and conceiving the correct information, avoiding misquotations. To fulfil the gap over a long distance, verbalcommunication can take place easily on phone calls. A path-breaking innovation has recently come to play in the SMS technologyusing the speech recognition technology, where voice messages are being converted to text messages. Quite a few applications usedto assist the disabled make use of TTS, STT

2. OBJECTIVES

The objectives of this project are given below:

- Speech to Text
- Text to Speech

3. PROPOSED SYSTEM

3.1 Introduction of Design Methodology:

A technique essentially makes use of the standards to supply tips and notation for the plan hobby.Even though methodologies are useful, they do no longer minimize the exercising of format to a chain of steps that may be accompanied automatically..

3.2 Working::

SPEECH TO TEXT:

Speech reputation is the capacity of device/program to discover words and phrases in spoken language and convert them into device- readable format. Speech reputation structures can be labeled on foundation of the following parameters :

Speaker: All audio system have a distinct type of voice. The models as a result are both designed for a particular speaker or an unbiased speaker.

Vocal Sound: The manner the speaker speaks also plays a role in speech recognition. Some fashions can recognizeeither unmarried utterances or separate utterance with a pause in between.

Vocabulary: the scale of the vocabulary plays an critical position in figuring out the complexity, overall performance, and precision of the machine

TEXT TO SPEECH CONVERSION

Text-To-Speech is a process in which input text is first analysed, then processed and understood, and then the text is convertedto digital audio and then spoken. Figure 3 shows the block diagram of TTS. The figure shows all the steps involved in the text to speechconversion but the main phases of TTS systems are:

Text Processing: The input text is analysed, normalized (handles acronyms and abbreviation and match the text) and transcribed intophonetic or linguistic representation.

Speech Synthesis: Some of the speech synthesis techniques are :

Articulator Synthesis: Uses mechanical and acoustic model for speech generation. It produces intelligible synthetic speech but it isfar from natural sound and hence not widely used.

Formant Synthesis: In this system, representation of individual speech segments are stored on a parametric basis. There are two basic structures in formant synthesis, parallel and cascade, but for better performance, some kind of combination of these 2 structures is used.A cascade formant synthesizer consists of band-pass reconnected in series. The output of each formant resonator is applied to the input of the successive one. The cascade structure needs only formant frequencies as control information. A parallel formant synthesizer consists of resonators connected in parallel. The excitation signal is applied to all formants simultaneously and their outputsare summed.

Concatenative Synthesis: This technique synthesizes sound by concatenating short samples of sound called units. It is used in speechsynthesis to generate user specific sequence of sound from a database built from the recording of other sequences. Units for Concatenative synthesis are : Phone- a single unit of sound; Diphone- is defined as the signal from either midpoint of a phone or point of least change within the phone to the similar point in the next phone; Triphone- is a section of the signal taking in a sequence going from middle of a phone completely through the next one to the middle of a third.

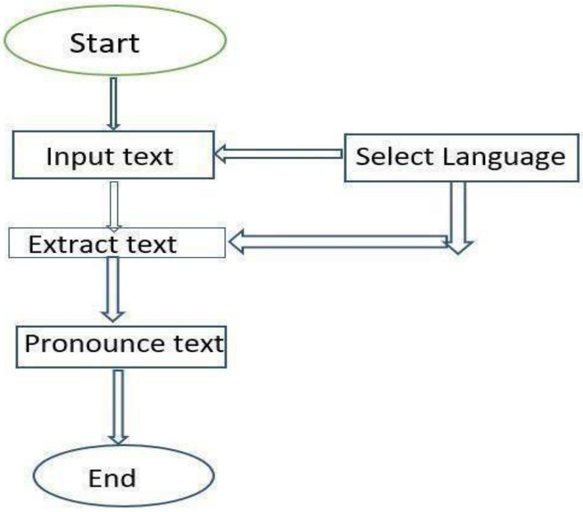
4. Algorithm

Text to speech :

- Take a text as a input.

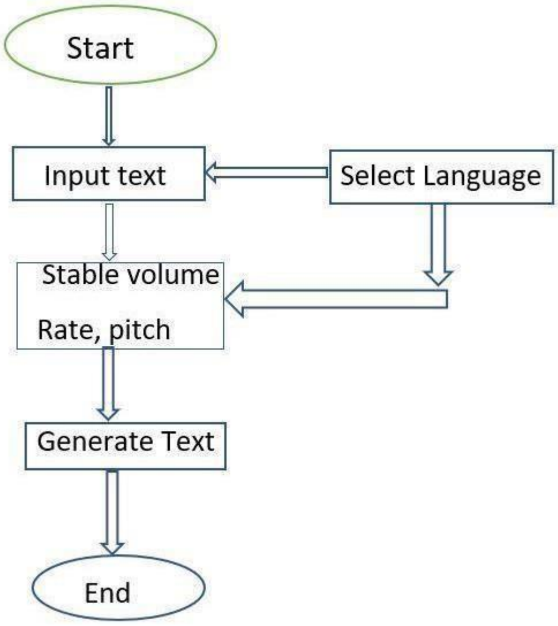
International Journal of Advance Research in Science and Engineering
Volume No. 11, Issue No. 06, June 2022
www.ijarse.com
IJARSE
ISSN 2319 - 8354

- Extract voice by firstly select thevoice which are available in library.
- Convert the text itno speech.



Speech to text:

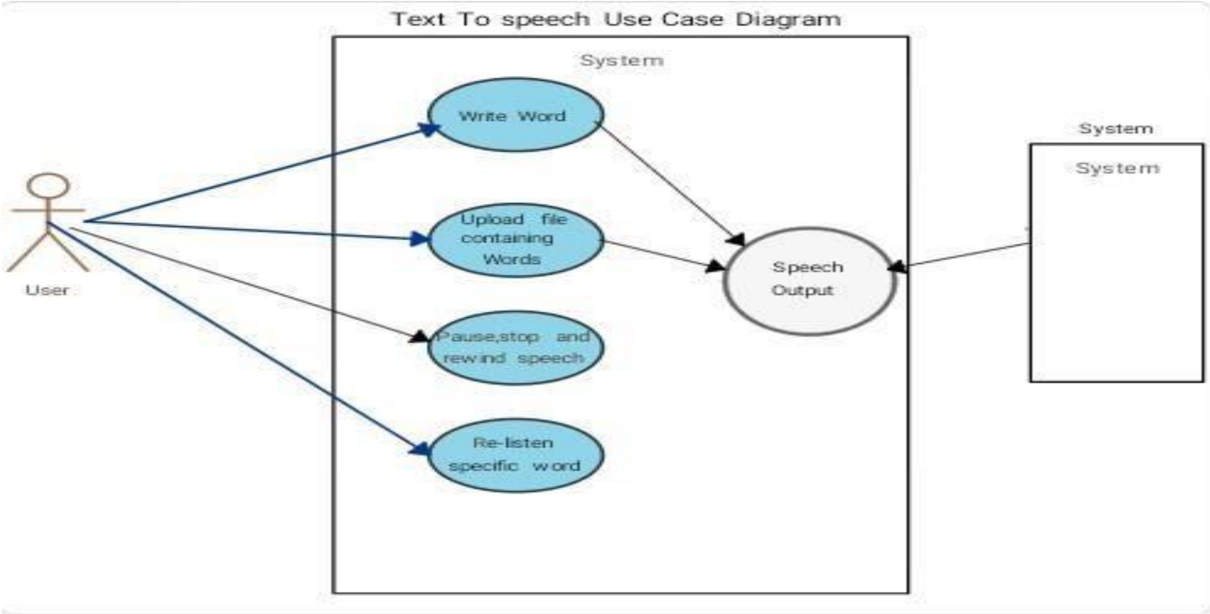
- Take a voice as a input.
- Adjust and stable,rate
- Generate text..



5. FIGURES AND TABLES

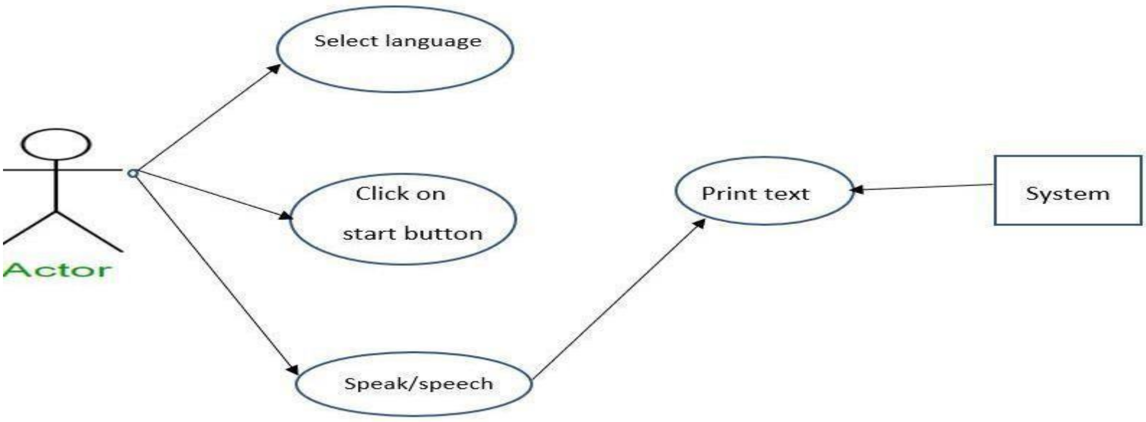
5.1. Use case Diagram –

Text to speech:

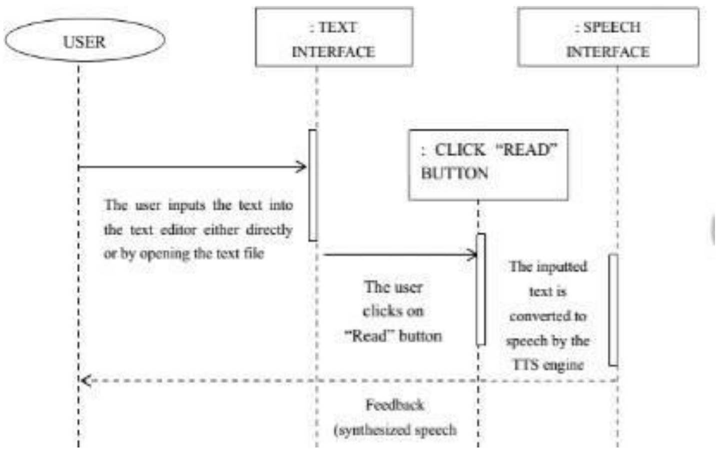


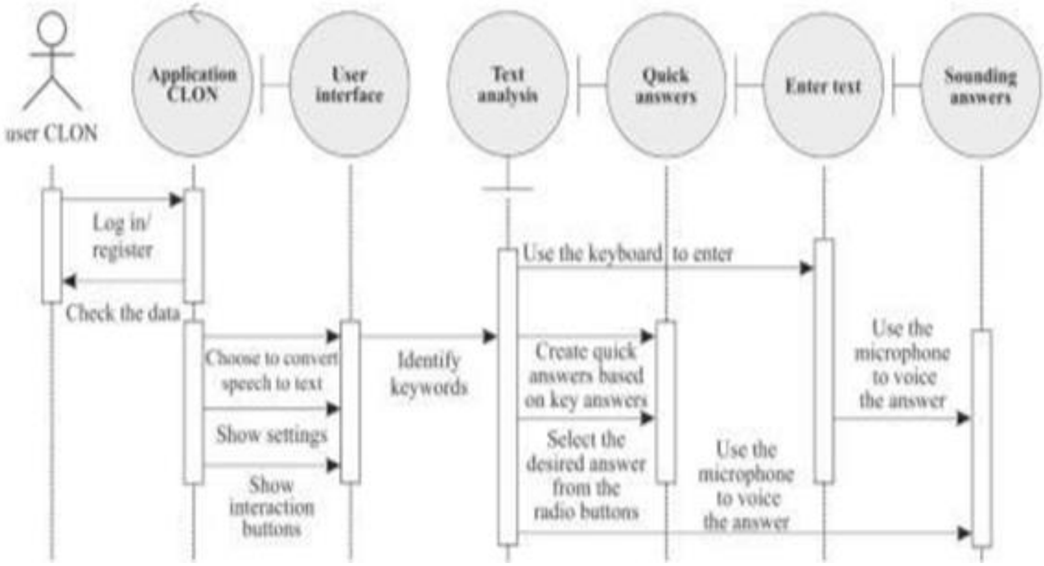
Speech to text:

5.2 Sequence diagram:Text to speech:



Speech to text:

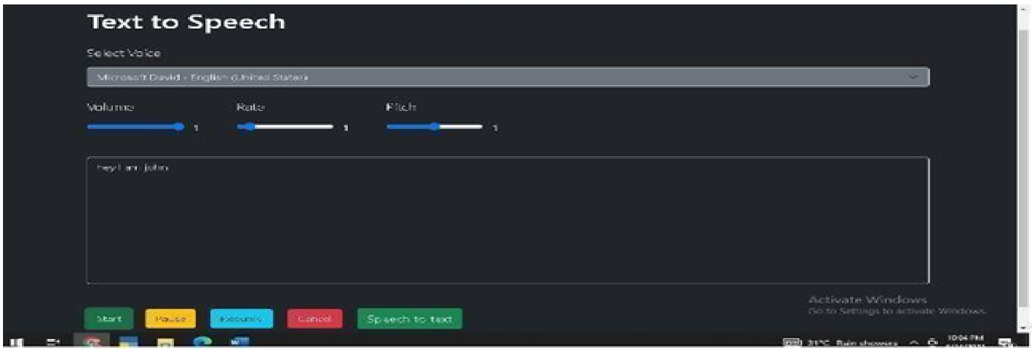




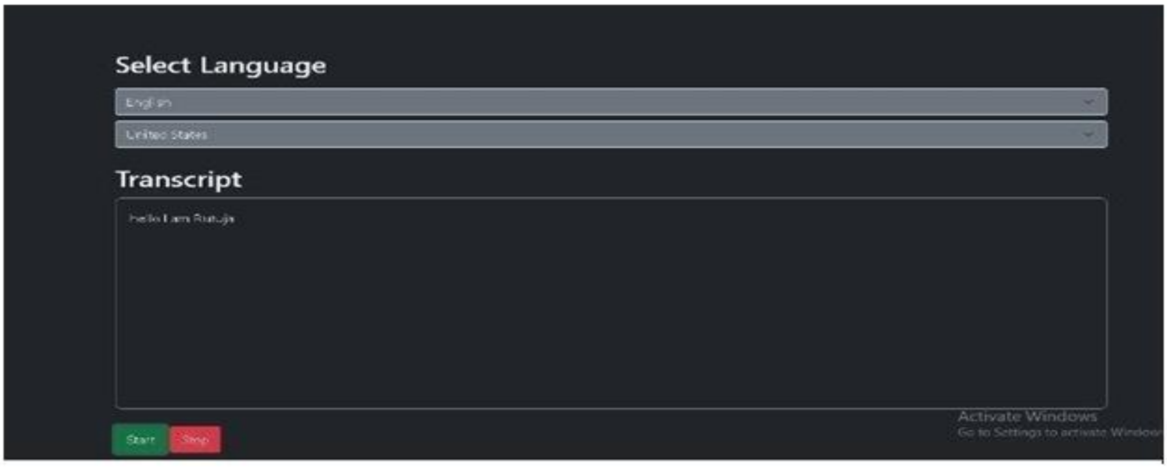
6. Result

6.1 Main Page:

Text to speech:



6.2 Speech to text



CONCLUSION

Hence have learned about various techniques that fall under STT and TTS, and have also read about their applications and usage. After having looked upon closely, at the different types of speech, speech recognition,speech to text conversion, text to speech conversion and speech translation systems.

REFERENCES

1. Jingdong Chen, Member, Yiteng (Arden) Huang, Qi Li, Kuldip K. Paliwal, “Recognition of Noisy Speech using Dynamic Spectral Subband Centroids” IEEE SSIGNAL PROCESSING LETTERS, Vol. 11, Number 2,February2004.
2. Hakan Erdogan, Ruhi Sarikaya, Yuqing Gao, “Using semantic analysis to improve speech recognition performance” Computer Speech and Language, ELSEVIER 2005.
3. Chadawan Ittichaichareon, Patiyuth Pramkeaw, “Improving MFCC-based Speech Classification with FIR Filter”International Conference on Computer Graphics, Simulation and Modelling (ICGSM“2012) July 28-29, 2012 Pattaya(Thailand).
4. Bhupinder Singh, Neha Kapur, Puneet Kaur “Sppech Recognition with Hidden Markov Model:A Review” International Journal of Advanced Research in Computer and Software Engineering, Vol. 2, Issue 3, March2012.
5. Shivanker Dev Dhingra, Geeta Nijhawan, Poonam Pandit, “Isolated Speech Recognition using MFCC and DTW” International Journal of Advance Research in Electrical, Electronics and Instrumentation Engineering,Vol.2, Issue8, August 2013.
6. Ibrahim Patel, Dr. Y. Srinivas Rao, “Speech Recognition using HMM with MFCC-an analysis using FrequencySpectral Decomposition Technique” Signal and Image Processing:An International Journal(SIPIJ),Vol.1, Number.2, December 2010.
7. Om Prakash Prabhakar, Navneet Kumar Sahu,”A Survey on Voice Command Recognition Technique” International Journal of Advanced Research in Computer and Software Engineering, Vol 3,Issue 5,May 2013.
8. M A Anusuya, “Speech recognition by Machine”, International Journal of Computer Science and Informationsecurity, Vol. 6, number 3,2009.