Voice Recognized System

Anju Pawar
Department of Computer Science Engineering
Sagar Institute of Research and Technology, India

Prof. Dilip Solanki
Department of Computer Science Engineering
Sagar Institute of Research and Technology, India

Abstract

Speech is the first important primary need, and the most convenient means of communication between people. The communication among human computer interaction is called human computer interface. This paper basically gives an overview of major technological perspective and appreciation of the fundamental progress of speech to text conversion and also gives complete set of speech to text conversion and Speech and text Hindi to English & vice versa translation. The language which is used as a communication to interact between others. But not everyone has the same language to communicate; there are many languages from different countries of the world. Speech-to-speech translation system able to converts source language into the target language. A comparative study of different technique is done as per stages. This paper concludes with the decision on future direction for developing technique in human computer interface system in different mother tongue and it also discusses the various techniques used in each step of a language translation process and attempts to analyses an approach for designing an efficient system for speech recognition. However, with modern processes, algorithms, and methods we can process speech signals easily and recognize the text and translate them in another language. In this system, we are going to develop an on-line language translation engine.

Keywords: Speech to Text conversion, Automatic Speech Recognition, Speech Synthesis, Text to Text Translation, Speech to Speech Translation

I. INTRODUCTION

Voice recognition (VR) and electronic health records (EHRs) have both entered mainstream medicine in the past decade. Currently the increased time burden of data entry into EHRs is one of the reasons that the EHR adoption rate is low. With voice recognition software continuing to improve in speed and accuracy, it could potentially improve the process of inputting data into electronic health records and thereby decrease one of the key barriers to EHR adoption. Similar to the introduction of many new technologies, VR may succeed or fail based on personal experience, training, or technical or logistical reasons. We sought to explore the factors that influence the continuation or discontinuation of voice recognition as an inputting method for an electronic health record by surveying all clinicians who volunteered to receive the software.

Speech & then it's converted into text form which is display on desktop. Speech processing is that the study of speech signals and therefore the various methods which are won't to process them. Speaking different languages. These technologies play an important role in multi-lingual societies like India which has about 1652 dialects/native languages. Speech to Text conversion take input from microphone within the sort of speech & then it's converted into text form which is display on desktop. Speech processing is that the study of speech signals and therefore the various methods which are won't to process them. During this process various applications like speech coding, speech synthesis, speech recognition and speaker recognition technologies; speech processing is used.

II. LITERATURE REVIEW

1) Yee-Ling Lu, Man-Wai and Wan-Chi Siu explain about text-to-phoneme conversion by using recurrent neural networks trained with the important time recurrent learning (RTRL) algorithm.
2) Penagarikano, M.; Border, G explains a way to perform the speech to text conversion also as an investigational test administered over a task oriented Spanish corpus are reported & analytical results also.
3) Quoc Truong Do, Sakriani Sakti, et al., in 2018 proposed a work on emphasis speech translation .Emphasis is used to distinguish between focused and unfocused part of an utterance and it is useful in misheard situations in which speakers must repeat the most important words or phrases accuracy and assure that the theme of this study also will be helpful for other languages for Speech-to-Text conversion and similar tasks.
4) Moulines, E., in his paper "Text-to-speech algorithms supported FFT synthesis, "present FFT synthesis algorithms for a system supported diaphone concatenation. FFT synthesis techniques are capable of manufacturing top quality prosodic adjustments of natural speech. Several different approaches are formulated to reduce the distortions thanks to diaphone concatenation.
5) Hariz Zakka Muhammad, Muhammad Nasrun, et al., in 2018 proposed a work on speech recognition by using hidden markov model. They consider the translation language from English to Indonesian. The classification method used is the Mel frequency Cepstral coefficients (MFCC) and Hidden markov model (HMM). The proposed system converts speech-to-text and uses the

existing Google translation or Microsoft translation for translation. The speech signal will be processed by using MFCC. The algorithm used here is kmeans algorithm.

III. SYSTEM IMPLEMENTATION:

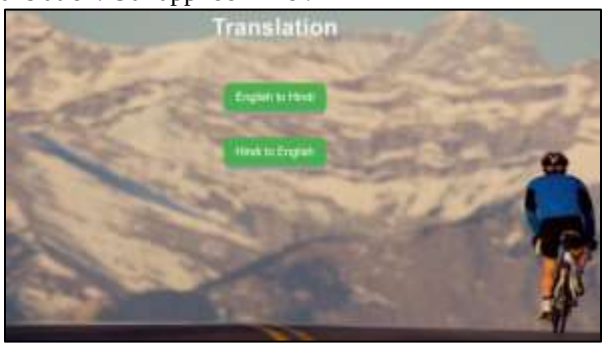
Implementation: The Implementation process achieves text to text and speech to text and speech to speech translation of hindi to english and vice versa, we have used google trans api for implementation of translation and we have used gtts for text to speech conversion, speech_recognition api for recognition of speech and use of flask api for web application of our translation app.

IV. AUTOMATIC SPEECH RECOGNITION:

Basic Principle: ASR systems operate in two phases. First, a training phase, during which the system learns the reference patterns representing the various speech sounds (e.g. phrases, words, phones) that constitute the vocabulary of the appliance. Each reference is learned from spoken examples and stored either within the sort of templates obtained by some averaging method or models that characterize the statistical properties of pattern. Second, a recognizing phase, during which an unknown input pattern, is identified by considering the set of references.

V. LANGUAGE TRANSLATION

Language Translation: Language Translation is done by Google Translator. Google translate Google's translation system works primarily on text, but has a built-in feature that allows you to pick up a microphone input and then play back the sound from the speakers. Google Translate uses the classic speech-to-speech translation style with the use of a speech identifier for text speech, text translation, and speech synthesis to generate the audio associated with the text. It should be noted that Google's translation service is a difficult candidate to beat in terms of correct translations due to its well-designed implementation with huge amounts of input data from many different sources. A basic machine translation system makes a simple word substitution to convert text from one language to another. The translation process involves decoding the meaning of the source text and re-encoding the meaning in the target language. Word meaning refers to the concept or sense of the source text. Machine Translation Systems can be built using several approaches like Rule-based, Statistical-based, and Examplebased systems. In Rule-based translation to translate from English-to-Telugu and Example-based translation to translate from Telugu-to-English. In Rule-based translation, we perform source language text reordering, to match target language syntax, and word substitution. On the other hand, in Example-based translation, we map examples of text in source language to the corresponding representations in destination language, to achieve translation. Our app look like :



VI. CONCLUSION

This section summarizes and concludes the contributions made by our project. A system is developed for the real time speech to speech translation based on the input speech given by user in any of the four languages they are Kannada, English, Hindi and Telugu. After the input speech is given, language translation is done via the Google API. Language barrier reduction is done via the speech to speech translation system. The system can defeat the constant challenges of unskilled individuals and improve their way of life The outcomes show sensibly great achievement in perceiving Continuous speech from different speakers, for an enormous vocabulary. The various modules were examined in their separate areas and were effectively checked for various speech input.

REFERENCES

[1] Sanjib Das, "Speech Recognition Technique: A Review",International Journal of Engineering Research and Applications (IJERA) ISSN: 22489622 Vol. 2, Issue 3, May-Jun 2012.
[2] NEY, H., Speech translation: Coupling of recognition and translation, Acoustics, Speech, and Signal Processing, IEEE International Conference on., Vol. 1, 1999, pp. 517–520.
[3] Mrinalini Ket al: Hindi-English Speech- 2014.
[4] Santosh K.Gaikwad, BhartiW.Gawali,PravinYannawar, "A Review on Speech Recognition Technique", International Journal of Computer Applications (0975 – 8887)Volume 10– No.3, November 2010.
[5] https://emerj.com/ai-sector-overviews/machine trans lation -14-current-applications-and- services/
[6] B. RaghavendharReddy, M.; Quilis, A.; Bernstein, J., "Real timetext-to-speech conversion system for spanish," Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '84. , vol.9, no., pp.85,87, Mar 1984.
[7] Olabe, J. C.; Santos, A.; Martinez, R.; Munoz, E.; Martinez, to-Speech Translation System for Travel Expressions, 2015 International Conference On Computation Of Power, Energy, Information And Communication.
[8] Kavaler, R. et al., "A Dynamic Time Warp Integrated Circuitfor a 1000-Word Recognition System", IEEE Journal of Solid-State Circuits, vol SC22, NO 1, February 1987, pp 3-14.