International Journal of Advances in Electronics and Computer Science, ISSN: 2393-2835 Volume-4, Issue-11, Nov.-2017
http://iraj.in

A SURVEY: SPEECH RECOGNITION APPLICATION

[1]ASHA SHAJEE, [2]DHRUV PATEL, [3]RAHUL MISHRA, [4]HIMANGSHU SAIKIA, [5]NARENDRAN.M

[1,2,3,4,5]Department of Computer Science Engineering, SRM University, Chennai, India
E-mail: [1]ashashajee@hotmail.com, [2]dhruvp3012@gmail.com, [3]rahulbhai2496@gmail.com, [4]dimpusaikia6@gmail.com

Abstract -
Speech is a very essential element of communication.
This project of ours will research word-based and phonological feature-based audio for vocabulary speech recognizers. Many clients who have dysarthria would find this very helpful, as their greatest handicap is their relative inability to control PCs.
Hence, we thus propose the idea of developing a speech recognition application to decode the acoustic signals and convert it to a syllable or set of words.
As of yet, there isn't any commercial or open-source product available that would enable users to enter unrestricted text into a PC via automatic speech recognition.
This paper gives an overview of the speech recognition application and its recent trends. The main objective of this paper is to compare and summarize the well known methods used in various stages of speech recognition system.

Keywords - Speech recognition, voice recognition, text to speech, speech to text, speech translate.

I. INTRODUCTION

Speech is the most basic and important form of communication for interaction with anyone. Thus to interact with computers via speech, rather than using the general devices is what one looks for. This can be accomplished by developing a Speech Recognition Application which allows a computer to identify the words that a person speaks into a microphone and convert it into written text. As a result, it has the potential of being an important mode of interaction between human and computers.
This has proved to be a very challenging task of the conversion. Communication among the human being is dominated by spoken language. Hence, it is natural for people to expect speech interfaces with computers which can speak and recognize speech in native language.

Speech Recognition is the methodology through which the computer can identify words or syllables spoken by the person and convert that into text accordingly.
Hence, this gives an interaction between human and computer. Here, accuracy plays the major role for developing and enhancing it.
As a result, it has the potential of being an important mode of interaction between human and computers. Since the 1960s computer scientists have been researching ways and means to make computers able to record interpret and understand human speech.
Even the toughest and complex problem such as digitalizing voice was a huge challenge in the early years.
It took until the 1980s before the first systems arrived which could actually decipher speech.
Communication among the human beings is dominated by spoken language, therefore it is natural for people to expect speech interfaces with computer. Especially with those computers which can speak and recognize speech in the native language. Machine reorganisation of speech involves generating a sequence of words best matches the given speech signal.

II. TYPES OF SPEECH

Speech has various different classifications based on the way they are spoken and said. Here, the modulation of each person's voice is different and altered in every speech type.

The following are the types of speech which are commonly being used. Thus, these are mentioned in brief below.
Thus, we now look at the types of speech.

Basically, speech can be sub-divided into:-

1) Continuous speech:- Here, the person speaks almost naturally and the SRS should identify all of the words/statement spoken correctly of accurately.
2) Non-continuous speech:- Here, the person utters single words and SRS must recognize each and every single word accurately. It's also called as Spontaneous Speech.
3) Isolated Speech:- It recognizes and usually requires utterance to have quiet on both the sides of the sample windows. It also accepts single words or single utterances at a time. This is called Isolated Speech.

Thus these are the different types of speech which is present. They also give the variations and modulations which are presented in its types.

III. LITERATURE SURVEY

S.NO	TITLE	YEAR	METHODOLOGY	DRAWBACK
1	Laughing voice recognition	2016	Senses periodic waveforms of laughing	Different waveforms can mix up easily
2	Wireless room automation system	2016	Uses voice commands for automation processes	Cant access if there is a system failure
3	Smart mobile attendance system	2016	Uses voice and fingerprint recognition	Fingerprint is less efficient as compared to VRS
4	Intelligent automatic starting engine	2016	Uses voice commands for engine functioning	Cant access if there is a system failure
5	Driving awareness detection	2016	Evaluation of awareness based on answers of questions	Answers can be manipulated accordingly
6	Smart wheelchair	2014	Uses touchpad and voice commands for the movement of wheelchair	Touch-screen is less efficient as compared to VRS
7	VRS for visually impaired	2008	Uses voice commands to access internet	Doesn't gives the required result accurately
8	Silencing VRT	2004	Uses stick-on sensors to grab the word before speaking	Doesn't senses exact words
9	Acoustic interference cancellation	2013	Reduces the acoustic interference for smart TV'S	Some minute noise waveform are not cleared totally
10	VERA-Voice Encrypted recognition authentication	2015	Uses voice to encrypt and decrypt phone data	Cant access data due to system failure

Thus, we see that in this literature survey, it shows the related wok of the other authors and their papers for the designing of their systems. The papers of these scholars that we have thus taken are all dated of the years in which they were published. Hence, these 10 papers that we have taken is unique and specific from the authors.

These include the current knowledge, substantive findings, as well as theoretical and methodological contributions of each of the authors to Speech and Voice Recognition.

Thus, it explains the methods which were used in them along with all their outcomes as well as their drawbacks, which are focused on each paper.

In our Literature Survey, we have referred around 10 different papers from the year 2008-2016.

One of the example:

- Laughing voice recognition system –

Here, the developers who developed this system sensed different waveforms of sound such as laughing, crying, etc.

They then separated laughing voice waveform from the others so that only that particular waveform can be detected.

Hence, we intend to take these references from the methodology used in different projects of speech recognition system.

IV. ANALYSIS ON SPEECH RECOGNITION

Speech recognition (SR) has made great strides with the development of digital signal processing hardware and software. But despite of all these advances, machines cannot match the performance of the human counterparts in terms of accuracy and speed. Before recognition, speech processing has to

be carried out to get a feature vectors of the signal. So, front end analysis plays an important role.
The reasons are its wide range of applications, and limitations of the available techniques of speech recognition.
So, in this analysis, we briefly discuss the different aspects of the front end analysis of speech recognition including sound characteristics, feature extraction techniques, spectral representations of speech signal, etc. Thus, it shows the suitability of each method to the particular application.

The objective of the speech recognition application is:-

- To understand speech recognition and the way it works.
- To see its applications in different areas.
- To understand its implementation as a single application.
- To check how it's faster than writing by hand and check its hands-free capability.
- To make it allow better spelling and grammar.
- To make it useful for people having mental and physical disabilities.
- To be used for voice dialing, dictation, navigation and industrial applications.

Hence, thus our solution to this would be a proposed system wherein this is to develop a translation feature in the near future to spread availability of product to all user types and to make system platform independent.
It will be faster than typing and allows for better spelling along with being hands-free capable.
Thus, speech recognition will revolutionize the way people conduct meetings over the Web.
Hence, in the near future, people would be using their home and business computers by speech and not by keyboard or mouse.

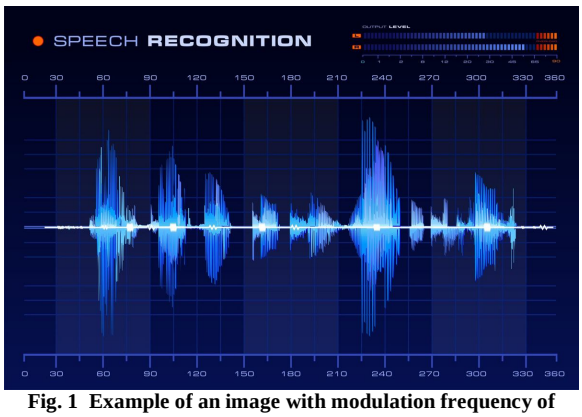

Fig. 1 Example of an image with modulation frequency of speech

Thus, we see that in this image of the frequency modulation, comparing the audio visualization and the speed in which the voice or speech is heard, the image is thus processed for the speech recognition.

There are already a few existing systems as in ,any cell-phone handsets have basic dial-by-voice features built in. That basic feature can be used in all type of communication devices.

Some applications of voice recognition system are as follows:

- Alexa – Amazon Echo
- Siri Personal Assistant
- Microsoft Cortana
- Google Assistant.ai
- Google Now
- Jeannie
- Indigo Virtual Assistant

V. EXPECTED OUTCOME

Thus, we expect our application will have a translation feature so as to spread availability of product to all user types and to make system platform independent.
Many clients who have dysarthria would find this very helpful, as their greatest handicap is their relative inability to control PCs. This would also help dyslexic people to help them interpret it correctly though speech.

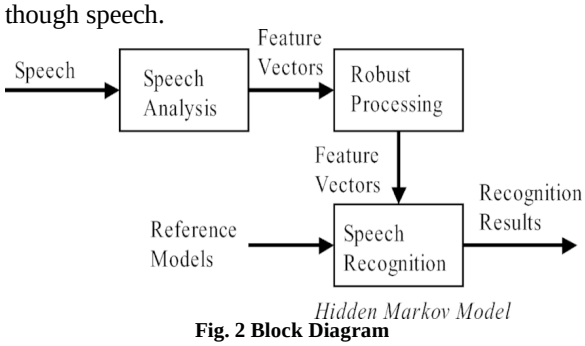

Fig. 2 Block Diagram

This is used for decoding the acoustic signals of the client that can be captured by the microphone to be converted to a syllable or set of words.
The hardware which would be used is microphone and soundcards and the software would deal with Windows OS and Audio Driver Software.
It is faster than typing and allows for better spelling along with being hands-free capable. It can also allow better spellings and grammar.

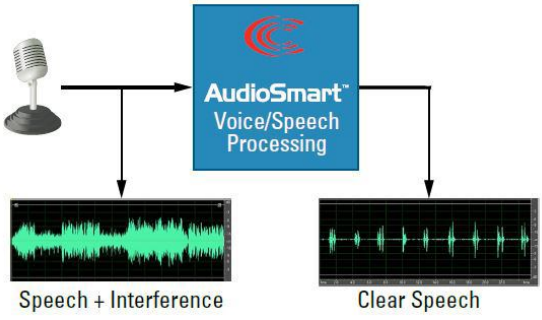

Fig. 3 Image showing how speech with inference is converted to clear Speech

Thus, it would prove to be very helpful in In-Car Systems, Health Care in medical documentation, therapeutic use, Military in high-performance fighter aircrafts, helicopters, training air traffic controllers, Telephony, Education, People with disabilities, Aerospace, Robotics, Court Reporting, Pronunciation, Home Automation, Telematics, Navigation and many others.

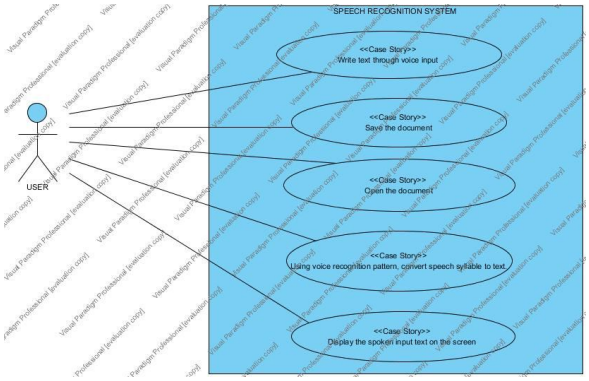

Fig. 4 Image showing use case model of Speech Recognition

CONCLUSIONS

Speech recognition has been in development for more than 50 years, and has been entertained as an alternative access method for individuals with disabilities for almost as long. In this paper, the fundamentals of speech recognition are discussed and its recent progress is investigated. The various approaches available for developing a SR system are clearly explained. The performance of the SR system based on the adopted feature extraction technique and the speech recognition approach for the particular language is compared in this paper. In recent years, the need for speech recognition research based on large vocabulary speaker independent continuous speech has highly increased. Based on the review, the potent advantage of HMM approach along with MFCC features is more suitable for these requirements and offers good recognition result. These techniques will enable us to create increasingly powerful systems, deployable on a worldwide basis in future, along with higher level of performance and robustness.

Thus, this will represent the next wave of the Web

REFERENCES

[1] http://ieeexplore.ieee.org/Xplore/home.jsp
[2] https://www.slideshare.net/sarangafle/speech-recognition-project-report
[3] https://www.codeproject.com/Articles/5820/Speech-Recognition
[4] Hakan Erdogan, Ruhi Sarikaya, Yuqing Gao “Using semantic analysis to improve speech recognition” performance in Elsevier 2005 .
[5] https://www.slideshare.net/charujoshi/speech-recognition
[6] Chunyi Guo, Runzhi Li, Lei Shi “Research on the Application of Biomimetic Computing in Speech Recognition” in IEEE 2008
[7] Kavita Sharma, Prateek Haksar “ Speech Denoising Using Different Types of Filters” in International Journal of Engineering Research and Applications Vol. 2, Issue 1,Jan-Feb 2012, pp.809811.
[8] Bhupinder Singh, Neha Kapur, Puneet Kaur “Speech Recognition with Hidden Markov Model: A Review” in International Journal of Advanced Research in Computer Science and Software Engineering Volume 2, Issue 3, March 2012.

★★★