www.ijecs.in
International Journal Of Engineering And Computer Science ISSN: 2319-7242
Volume 5 Issue 10 Oct. 2016, Page No. 18389-18391

Speech Recognition using e-Speaking: A Review

Anju Bisen[1] , Krishna Hirpara[2], Madhu[3], Roshani Rathore[4]

BE student[123],Dept. Of Computer Science,Govt Engg. College Raipur,C.G
bisen4267@gmail.com
krishnahirpara111@gmail.com
thakurmadhu352@gmial.com
Asst. Professor[4] Dept. Of Computer Science, Govt. Engg. College, Raipur

ABSTRACT

This paper presents the basic idea of speech recognition and its progress till date. Speech recognition is essential for a communication between human and machine.The ultimate goal of the technology is to able to produce a system that can recognize with 90-100% accuracy all words that are spoken by any user in different language .This paper helps in selecting the techniques along with their relative advantage and disadvantage. Under the speech recognition using e-speaking product,we create a small and efficient program to take humans voice as input and convert it into keyboard ,mouse system and program events and even speak to you to let you know what it has performed.

KEYWORDS-Automatic Speech Recognition ,e-speaking , DFT, LPC , MFCC ,VQ

1 . INTRODUCTION

Speech recognition has the potential of being on important mode of interaction between human and computers[1].And is now being implemented in different languages around the world. Speech or voice in human can be said as the most common means of the communication because the information maintains the basic role in conversation. This can be achieved by automatic speaker recognition system(ASR) which capture the spoken words from user and the final result will be the set of the words or it can then apply the synthesis to Pronounce intosounds, which mean speech to speech[2].

1.1**ASR-** Automated Speech recognition is a technology that allows users of information System to speak entries rather than punching number on a keyboard.

1.2 **e-speaking**– It supplements the user's ability to command and control your computer through your

voice for dictation software it using Microsoft's SAPI speech engine. And to allow the computer to read document and email to the user.

How does it work?

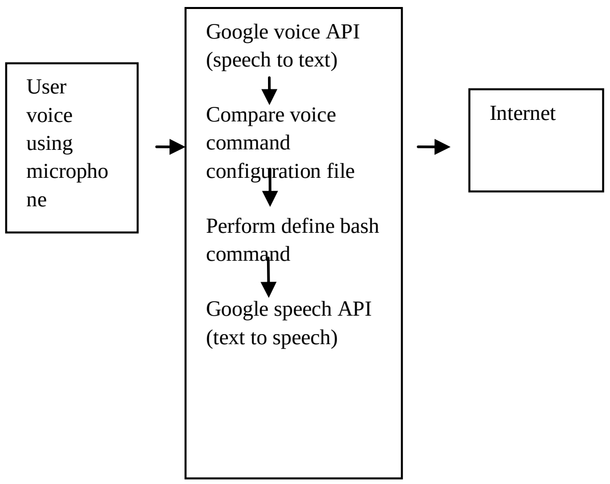

Figure 1 working of voice Recognition software

The software being described here uses Google voice and speech APIs. The voice command from the user is captured by microphone .this is then converted to text by using Google voice API. The text is then compared with the other previously defined commands inside the command configuration file. If it matches with any of them, then bash command associated with it will be executed. You can also use this system as an interactive voice response system by making the raspberry pi response to your command via speech. This is be achieved by using the Google speech API, which convert text into speech. Here's a block diagram(fig. 1) showing you the basic working of the voice recognition software.

1.3 Typology Of Speech Recognition System

- **Speaker dependent- s**ystem may require a user to train the system according to their voice and also for different voice for different users [3].
- **Speaker independent-** In these no system training is required that is they are development to operate of any speaker.
- **Isolated word Recognizer-**the system accepts one word at a time it means user can speak naturally continuous.
- Connected word system[4] allow speaker to speak slowly and distinctly each word with short pause.

2. Literature Survey

Speech recognition is one of the most complex area of speech recognition has in year has become a concept which is now being implemented in different language. In this survey we discussed the some approaches of speech recognition. the broad approaches to speech recognition are acoustic phonetic, pattern recognition,and artificial intelligence approach[5] the goal of speech recognition is to analyze, extract, characterized and recognize the information about the speaker identity[6]the various technique used for determining the speech characteristic.

Raj Raddy was the first person to take on the continuous speech recognition as a graduate student at Stanford university in late 1960.previous system required the user to make a pause after each word .Raddy's system was designed to issue spoken command for the game of chess.also around this time sovient research invented the dynamic time wrapping algorithm and used it to create a recognizer capable of operating 200 word vocabulary [7].

In February 1975, F. Itakura founded a new measure of distance for pole model of speech has been derived on the basis of the probability ratio criteria and is applied to automatic recognition of isolated words. And define algorithm to find to the best match between the input pattern and a reference pattern is derived. The dynamic programming technique is used in conjunction with a sequential decision scheme. The system is being implemented on a DDP-516 computer to recognize 200 isolated words. The validity of the scheme has been confirmed experimentally and Further work is in progress to test the system for a greater number of talkers and for telephone connection switched over greater distances [10].

In 1975,J. K. Baker state that termination that the hidden articulatory Markov model as an alternative or companion to standard phone-based HMM models for speech recognition. In this paper Found that either in noisyconditions, or when used in tandem with a traditional HMM, The WER results improved by hidden articulatory model. Also shown that the HMM is able to reasonably estimate articulator motion from speech. There are a number of avenues to improve this work. In the future, the plan to add more articulatory knowledge, with rules for phoneme modification that arise as a result of physical limitations and shortcuts in speech production, as was done in (Erler 1996) (for example, vowel nasalization)[11]. this rules may help the speech recognition systems in the presence of strong co articulation, such as in conversational speech

.In 1976 F.Jelinek presented a new approaches , In this paper describe visual speech recognition based on a data driven lip model and HMMs. Under Fred jelinek's lead , IBM created a voice typewriter called Tangora which could handle 20,000 word vocabulary by the mid 1980's [12]. This was controversial with linguistic since HMM are too simplistic to account for many common feature of human language[13]. However HMM proved to be highly useful way for modelling speech and replace dynamic time wraping to become a dominate speech recognition algorithm in 1980 [14].

In 1980 , using statistical method known as HMM in which vocabulary size are increases into several thousand word. It is capable to recognize unlimited no. of words.

In 2000, stiphan & Leila concluded fast speaker adaption technique dedicated automatic speech recognition system by using artificial neural network for HMMs state probability. The result show that decrease the no. of word error rate nearly 25% over the system independent system

In 2006, ranbir and junag defined that the recognizer consist of three processing step namely feature analysis ,pattern matching and confidence scoring

In 2009 , Ladan & Dougal investigated a non parametric classification in speaker independent continuous speech that employ powerful classifier. The result of this recognition shows increases in the percentage of correctness over the conventional HMM based phenomenon recognition.

Santosh et al(2010) proposed that the goal of speech recognition is the ability of machine has capable to understand , hear and act upon spoken information.

3. TECHNIQUES USED IN SPEECH RECOGNITION

DOI: 10.18535/ijecs/v5i10.28

Techniques	Functions	Advantage
Acoustic Phonetic Approach • DFT • LPC • MFCC	Spectral analysis with feature detection Phonemes ,nasality(nasal resonance),frication(random excitation),labelling.	Formant transitions, Silence detection,Voicing Detection, Zero-crossing rate.
Pattern Recognition Approach • Template • DTW • VQ	Correlation distance measure Clustering function Dynamic warping Optimal algorithm Pattern matching theory, feature analysis	Recognition pattern quickly, with ease, with atomicity
Artificial intelligence Approach	Knowledge based representation	The focus in this approach has been mostly in the representation of knowledge and integration of knowledge sources
Neural Network	Network function	Neural Networks are capable of solving more complicated recognition task, achieve more accuracy than HMM model.

4. CONCLUSION

In this research we study about speech recognition with e-speaking product that will results us better communication with system i.e. System can also give reply us with speaking tone. like human beings. And basically the main strength of speech verification technology is that relies on a signal that is natural and unobtrusive to produce and can be obtained easily from almost anywhere using the familiar telephone network or internet with no special user equipment or training .In our research we faces some problem related to language .we wants to involved many languages for batter communication but there is some noisy problem with many voices in different languages .

In future there will be a focus on development of the system automatically guess what the user intended to say , rather than what was actually said , to avoid mistakes and use many languages with translation according to user needs and microphone and sound system that will be designed to adapt more quickly to changing background noise level , different environment.

REFERENCES-

[1]. Sanjib das "speech recognition technique", university of north Bengal, India , vol 2,2012.

[2]. Kavita S.Yadav, M.M.Mukhedkar "Review on speech recognition" Electronics and Telecommunication Dept. Dr.D.Y.Patil COE, Ambi,Talegoan, Dist: Pune. , vol 1 No. 2,2013.

[3]. Preeti Saini , Parneet Kaur ,"Automatic speech recognition: A Review ",CSE Department, Kurukshetra University ACE , Haryana, India , vol 4, issue 2,2013.

[4]. M.A.Anusuya and S.K.Katti, "Speech Recognition by Machine: A Review", (IJCSIS) International Journal of Computer Science and Information Security, vol. 6, no. 3, pp. 181-205, 2009.

[5]. L.R.Rabiner, B.H.Juang. Fundamentals of Speech Recognition, Prentice-Hall, Inc.,Upper Saddle River, NJ. 1993.

[6]. Shikha Gupta, Mr.Amit Pathak , Mr.Achal Saraf, "A STUDY ON SPEECH RECOGNITION SYSTEM: A LITERATURE REVIEW ",International Journal of Science, Engineering and Technology Research (IJSETR), Volume 3, Issue 8, August 2014

[7]. Benesty, Jacob; Sondhi, M. M.; Huang, Yiteng , Springer Handbook of Speech Processing. Springer Science & Business Media. ISBN 3540491252,2008.

[8]. Blechman, R. O.; Blechman, Nicholas (June 23, 2008). "Hello, Hal". The New Yorker. Retrieved 17 January 2015.

[9]. Blechman, R. O.; Blechman, Nicholas (June 23, 2008). "Hello, Hal". The New Yorker. Retrieved 17 January 2015.

[10]. Er. Shweta Doda , Er. Rajni Mehta"Speech Recognition Techniques: A Review", JCD Vidyapeeth, Sirsa, India , Volume 4, Issue 8,August 2014.

[11]. J. K. Baker, "The DRAGON system - An overview" , IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-, pp. 24-29,23, February 1975.

[12] "Pioneering Speech Recognition". Retrieved 18 January 2015.

[13]. Xuedong Huang, James Baker, Raj Reddy (Retrieved 20 January 2015). "A Historical Perspective of Speech Recognition". Communications of the ACM..

[14]. juang, B. H.; Rabiner, Lawrence R (Retrieved 17 January 2015.). "Automatic speech recognition–a brief history of the technology development" Georgia Institute of Technology, Atlanta ,Rutgers University and the University of California, Santa Barbara.