International Journal of Scientific Research & Engineering Trends
Volume 5, Issue 6, Nov-Dec-2019, ISSN (Online): 2395-566X

A Review: Speech Recognition

Sunil Sunil.sharma8397@gmail.com
Shweta Shweta.koli.8888@gmail.com
Tanya Gulati Gulati.tanya1998@gmail.com
Asst. Prof. Garima Singh
Department of Computer Science Engineering
ADGITM, New Delhi, India

Abstract **- Speech recognition is the capability of a program to identify words and expression in spoken language and convert them to a binary language so computer system can understand. Each spoken word is split into discontinuous division which include many tones. This paper brief about the speech recognition system and how it's work and application in day to day life.**

Keywords- **Speech recognition; speech recognition working, speech recognition application, Future of speech recognition**

I. INTRODUCTION

Speech recognition technology has progressed immensely over the last few years. Research into speech recognition and its applications has been going on since the 70s.Speech recognition technology is concerned with using voice samples to identify a speaker (voice identification) or to confirm.

II. HOW DOES IT WORK

1. How does Speech recognition technology work?
Speech recognition technology works by recording a voice sample of person's speech and digitizing it to create a unique voice print or template. Each spoken word is split into discontinuous division which includes many tones. These tones can be digitized and are hold to create a speaker's different voice sample and covert into text[1].

2. Physiological component
The physiological component of a person's voice is based on the shape of that person's vocal system, i.e. the shape of the larynx, nose and mouth. Biometric technology uses the wave form of the voice sample to digitally recreate the shape of an individual's vocal tract. No two individuals can have the same vocal tract and therefore every person will have a proper voice imprint so machine can hear and translate into text so it can perform given task.

3. Behavioral component
This component shows the physical motion of the individual's jaw, tongue and larynx. Variation in this movement causes changes in the pace, manner and pronunciation of a person's voice, which include the person's accent, tone, pitch, pace of talking etc.the identity of a speaker (voice verification). These are the two different applications of speech recognition technology [1].

4. Factors that can cause a template mismatch
Currently speech recognition systems are not completely infallible, and there are factors that can cause a voice recognition system to return an erroneous result. Differences in the physical existence and sentimental of an individual can also reason a person's speech to change. For example, having a cold can change the person's speech enough to create a mismatch between the stored samples and the person's current voice sample. This is further possible if the person is fussed, depressed, tension or even on medication. Other factors, such as background noise, ambient temperature and the input device can also affect the result of the system. That are the some factor speech recognition work on machine should understand the clear words spoken by the person and respond to correspond query spoken by the user.

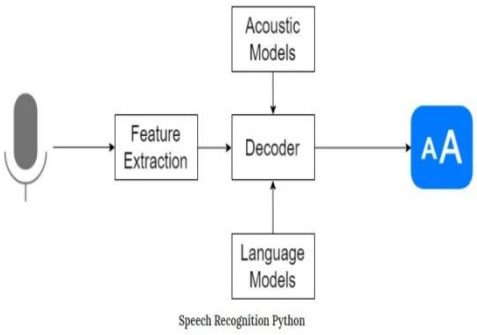

Fig.1 Model in Speech recognition[5]

It is based on these principle of acoustic and language modeling. Acoustic modeling describe the relationship between language related units of speech and audio signals. Language modeling matches sounds with word order to help differentiate between words that sound identical [5].
Speech recognition program is evaluated on these factors:
1. Accuracy of the speech
2. Speed of the speech.

© 2019 IJSRET

III.SPEECH RECOGNITION APPLICATION

These days speech recognition technology is too vast and it is using everywhere. Some of the industry are:
1. In Automobile Industry
2. Health care
3. Military- high performance unmanned vehicle like drone and training air traffic controllers.
4. Telephony and other domain
5. Usage in education and daily life
6. People with disabilities.
7. Artificial intelligence

1. In Automobile Industry

Driver in the car must take hold of the steering wheel all along, but if recognition system is installed in the vehicle, it will be convenient to make phone calls without hand, and can also be select radio stations or play music.[3]

2. Health Care

In health care sector speech recognition can help with the monitoring the patient health by the voice.

3.Military -Every country wants to modernize their army or defense personal with the latest technology. So in military country are using unmanned vehicle like drone are using for Surveillance. Drone operated with the speech or command which is given by their specific user.

4.Telephony and other domain-Voice dialing of telephone uses the voice chip which is based on speech recognition technology. There are a lots of function like voice search, voice control over electronic devices[4].

5.Education and daily life-Speech recognition technology help to practice language for betterment and be more convenient. Before speech recognition it is difficult to correct pronunciation by contrast. And in daily life it is been widely used by peoples such as television, radio communication device and many more application[3].

6. People with disabilities -People with disability is useful to the people with limited mobility in their arms and hand or leg or for those who has problem with the sight.

7. Artificial Intelligence- In robotics it take input in the speech wave form or spoken words[2].

8. Translation- It is an advanced application which translate from one language to another. It take input as speech and give output in spoken words[2].

9. Dictation- Dictation system on the market accepts continuous speech input which replaces menu system[2].

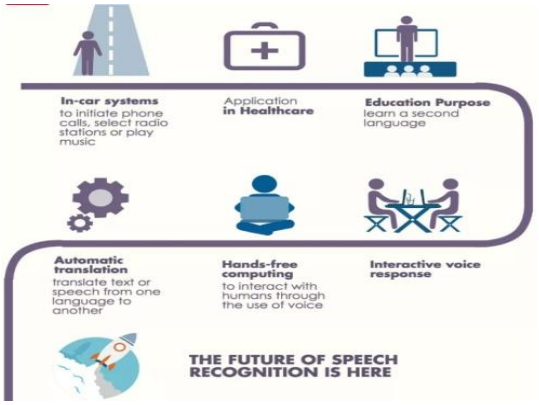

Fig.2Application of speech recognition[5]

IV. FUTURE OF SPEECH RECOGNITION

It is predicted that speech recognition will be the important part of our everyday lives sooner than we think. Thanks to such devices like as Amazon 'Alexa' and Google's 'Home' hub voice search assistance are making their way into our homes and changing the way peoples live. Speech is a more natural means of interacting, so it's no wonder speech recognition is looking bright [6].

Voice search will be increase 50% by 2020 and every 2 guys out of 5 prefer voice searching over typing. Voice assistance is popular amongst the new generation who likes to lay with technology. 13% American have voice assistance in their home for convenience. Now a days people are adapting new technology very freely.

Fig.3 voice search stats[7]

© 2019 IJSRET

REFERENCES

[1] Chris kikel 2019, How voice Recognition technology works <https://www.totalvoicetech.com/how-voice-recognition-technology-works/>

[2]. M.S.Anusuya, S.K. Katti,"Speech Recognition by machine: A review" proc. International Journal of Computer Science and Information Security, Vol. 6, No. 3, 2009 pp-182-183

[3] Youhao YU "Research on Speech Recognition Technology and Its Application" proc. International Conference on Computer Science and Electronics Engineering, 2012, pp-307-309

[4] Povey D, and Woodland P.C,"Minimum phone error and I-smoothing for improved discriminative training, " Proc. IEEE International Conference on Acoustics, Speech, and Singnal Processing, IEEE Press,May. 2002, pp. 105-108

[5] Gulsanober Saba 2018, image1 and image 2 speech recognition python and application of speech recognition<https://www.simplifiedpython.net/speech-recognition-python/>

[6] Paul Matthews, April 2019, "The future of voice recognition1" <https://www.getelastic.com/the-future-of-voice-recognition-infographic>

[7]. Sharoon Emmanuel, April 2019, image 3 voice search stats <https://bester capital media.com/blog/why-voice-is-the-future-of-digital-marketing>

© 2019 IJSRET