TO PROVIDE A SYSTEM FOR THE DEAF FOR PROPER COMMUNICATION WITH NORMAL PEOPLE

[1]Akshay Ithape, [2]Khushal Adlakha, [3]Tanvi Shinde, [4]T.J. Parvat
[1]BE Student, [2]BE Student, [3]BE Student, [4] Asst. Professor
[1]Department of Computer Engineering, [2]Department of Computer Engineering, [3]Department of Computer Engineering, [4]Department of Computer Engineering
Sinhgad Institute of Technology, Lonavala, Pune, India

Abstract: The objective of this project is to design and develop an application that can fulfill the communication gap between the DEAF and the normal people i.e. an application that can convert speech to ASL (American Sign Language) and basic text. The Final project is envisioned to be a windows phone application which helps the user with hearing disabilities to communicate in real time. This application has been conceived to provide quick information exchange between people with hearing impairments and those around them who may not know sign language. The proposed system would recognize and convert Speech to Text and further to American Sign Language. This application can be implemented through all the platforms. Ultimately, this application can be viewed as a translator to sign language. This application will encompass a speech recognizer for speech to text conversion, transitional models and an inventory of American sign language signs. This application will function in a three-phase implementation. This application will be user friendly from both the ends i.e. the deaf as well as the normal people. The speech converted into text uttered by the normal person will be then translated into images for the deaf to understand in the form of the standard sign language which will be standardized.

Keywords-American Sign Language

I. INTRODUCTION

Hearing loss could result from genetic causes, complications throughout birth, few infectious diseases, chronic ear infections. Current production of hearing aids meets less than 10% of global need. hearing disorder is that the impairment that worst have an effect on the communication between individuals. Deaf individuals (especially for those who became deaf before language acquisition) have serious issues once expressing themselves or understanding written texts. they face issues with verb tenses, gender and variety etc., and that they have difficulties once making a internal representation of abstract ideas. These deficiencies became apparent as a result of the shortage of feedback in speak-listen procedure. This act causes deaf individuals to face various issues once accessing information. Some other places where deaf people face problems are while accessing education, culture, social relationship. Deaf individuals use a sign language (their mother tongue) for act. sign languages square measure fully- fledged languages that have synchronic linguistics, similar to any auditory communication, contrary to what the general public assume. Sign languages aren't disseminated enough among hearing individuals inflicting communication barriers between Deaf and hearing individuals. These barriers square measure even a lot of problematic after they seem between a deaf and a traditional person. Hence, we glance forward to create a system which might cut back this gap between the deaf and normal individuals for better communication.

A. Motivation to Topic

As it is mentioned earlier about the statistical scenario of the deaf people worldwide, something needs to be implemented on those lines. There isn't any system developed for the deaf to make their life much easier when it comes to communicating with the normal people around. Some deaf people may choose not to talk because it is difficult for them to regulate the volume pitch or sound of their voices in a way that most people can understand. This drive one through the thought process of building an application which would be efficient enough for the speech-to-text-to-image

B. Problem Definition and Scope

Tackling the communication problem between the deaf and the normal people with suitable speech recognition algorithms and converting the same into hand gesture images in a windows phone application.

C. Goals and Objectives

To design and develop an application that can fulfill the communication gap between the DEAF and the normal people.

- Accurate recognition of speech
- Conversion from speech to text
- Conversion from text to images

II. COGNITIVE SERVICES

A. Bing Speech API

Bing Speech API is a cloud-based API using which speech driven actions are added to the software, this also includes real time interaction with the user. It provides advanced algorithms to process speech. The API consists on two components Speech to Text API and text to speech API. Speech to text API is used to convert spoken words into text. Text to speech PI is used to convert text into audio which can be played to the user. This API supports almost all platforms including Windows, Android, iOS and Third party IoT devices. Developers have a choice of using

the Client Library, Service Library or the REST API. REST API provides just one recognition back and no partial results. Client library supports real-time streaming, which means partial recognition results are returned as the audio which is spoken is sent. Service library also supports real-time streaming. This library is supported for Window

B. Custom Recognition Intelligent Service (CRIS)
Overcome speech recognition barriers like speaking style, background noise, and vocabulary.

C. Speaker Recognition API
The cloud-based APIs for Speech recognition provides the foremost advanced algorithms for speaker verification and identification. Speaker Recognition is divided into 2 categories: speaker verification and identification.

III. FUNCTIONAL MODEL AND DESCRIPTION

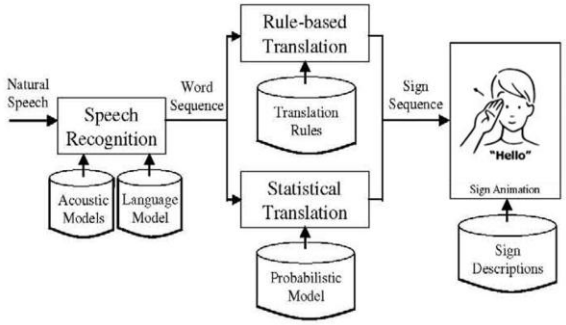


Figure 1. Speech Functional Model

This section introduces the system summary and also the new tools and utilities for adapting the system to a brand-new domain. Fig. 1 shows the module diagram sketched for converting spoken language into ASL. the primary module, the automated speech recognizer, converts natural speech into a sequence of words (text). It uses a wordbook, a language model and acoustic models for each phoneme. The linguistic communication translation module converts a word sequence into a sign sequence. For this module, a combination of 2 totally different ways are used. The first one consists of associate example-based strategy: the interpretation method is applied based on the similarity between the sentence to be translated and also the samples of a parallel corpus. The other relies on a applied math translation approach wherever parallel corpora area unit are used for training language and translation models. At the ultimate step, the sign animation is created by employing a extremely correct illustration of the movements (hands, arms and facial expressions) during a Sign list information and a Non Linear Animation composition module, both needed to get clear output. This illustration is freelance of the virtual character and also the final illustration part. In this manner, the virtual character may be simply modified and also the results can be custom-made for the employment in several devices. At the lowest a part of Fig. 1, the new tools for increasing the system flexibility area unit bestowed. The speech recognizer includes associate acoustic adaptation module, or adapting the acoustic models to a new specific atmosphere, a new speaker, or a brand-new Spanish accent. The speech recognizer uses a vocabulary and a language model generated from the parallel corpus. during this method, a new module has been enclosed for introducing language variants, increasing the speech recognizer flexibility. The language translation module presents better configuration compared to the previous version (San-Segundo et al., 2012) wherever all the interpretation strategies area unit data-oriented ones.

IV. FLOW CHART FOR SPEECH RECOGNITION

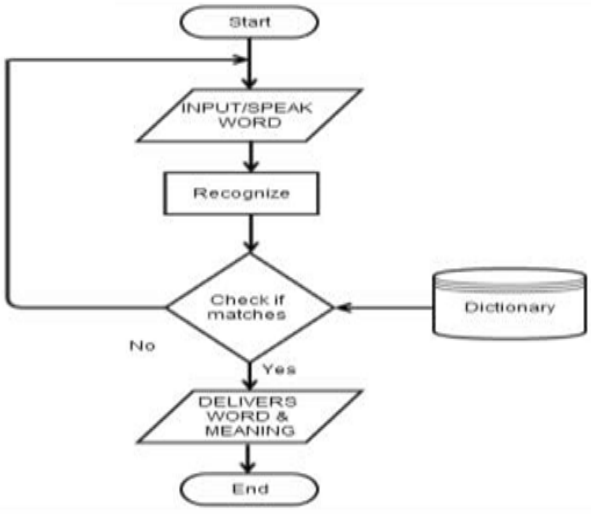
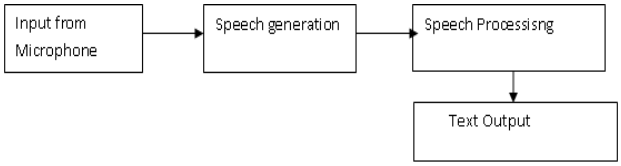


Figure 2. Flow chart for speech recognition

Fig.2:-The flow chart begins with the input i.e. the word spoken by the user1.This can be again recognized irrespective of the sounds in the background leading to an accurate input. When the output is given into the microphone, Cortana is enabled and the speech recognition takes

place. When this speech recognition is happening it first checks the matches from the cache memory which includes the recently used words else scans the dictionary at the whole. If the required output is acquired by the recognizer, then it moves further else turns back to step 1. When the words are matched with it,it delivers the words and its meaning in the form of text.

V. STATE TRANSITION DIAGRAM



This is our state transition diagram wherein we take the input from the microphone irrespective of the sound in the background and this can be enabled with the help of cognitive services. When the input is taken from the microphone and an acoustic signal is generated. To term it into simple words a speech is generated with the help of Cortana integration in it. The device recognizes the language and the words. This part is usually carried out in the speech processing section. After the speech processing method, it further drives us to the output for the given input. The speech is now converted to a text which is readable and creating to language barrier as multiple language inputs can be generated.

VI. WORKING

This application is built in a IDE named visual studio which consist of its own packages which helps in the enhancement of the same. This application is written in C#. Extensible Application Markup Language (XAML) page is added for a better graphical appearance which would attract the users and make it all the way more and more user friendly. The cognitive services which we have mentioned earlier, while building this project will be implemented through their API. The solution explorer of our project will lead us to manage the nudgets which include the packages for running this application. the required packages can be installed from there. The application consists of two modules where one will be convenient for the ones who are familiar with the English language. These are basically the people who have turned out to be deaf due to some unfortunate mishap in their life. The other module will consist of only sign language and its sub parts which could be the basic objects which are often used by people. This will make it all the way more convenient to the people for a clearer conversation. One of the prominent features of this application would be that its feasible enough to work on all screen sizes and devices. Interestingly it will also cover the embedded systems which would cover the IoT hub.

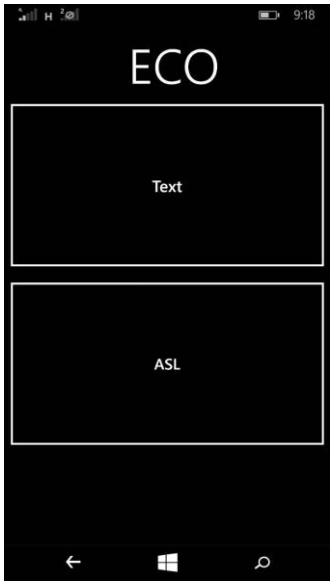


Figure3. Main Page

Figure4. Speech Recognition

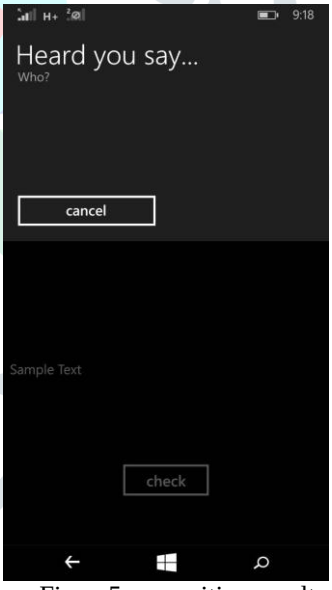


Figure5.recognition result

Figure 6. Graphical Results

As mentioned earlier figure 3 consists of the two modules on which our app draws all the focus on. It consists of on a speech to text converter and a speech to text to image converted.

Figure 4 pops up when we select the first module which is the speech to text converter where the ones who are known to the language will be able to communicate through this medium

Figure 5 is the result of the figure 4 so after getting the input from the user through a microphone we get the result in the form of a text which is displayed on the screen in seconds

Figure 6 is the screen which pops up when we go for the second module that is the speech to ASL converted and after searching and matching the input in the database we come up with a gif.

VII. SUMMARY

Our project is just a virtual implementation of the real-life activity as shown above where is the text and speech is generated with the help of each other and in the output the image is formed for the better learning and understanding Summarizing this project into an application which would be developed for speech recognition turning in text for the readers. It will further convert the text to image for the ones facing the language barrier and are well acquainted with the sign language.

VIII. FUTURE PLANS

After the implementation of the project further changes would be enhancing the application with face and gesture recognition using cognitive services as the normal people are unaware of the sign language which will be used by the deaf to reply them. This will convert the images captured by the camera in the phone into text or speech for the normal people to understand.

IX. CONCLUSION

After looking onto all the possible scenarios and applications leading us to this domain, we conclude to make an application which will consist of gif images to fill in this gap between the deaf people and the normal people. A handle device was designated to be the one initially with a layer of raspberry pi a display and power supply. But as it wasn't user friendly nor was it as cost efficient as one could think to have. Multiple languages could be incorporated in this app with an additional feature of the font size and backlight transformation for clear visual results. Hearing disability was a major issue initially but as the time has passed by technology has eradicated all the constraint which would be an obstacle in any kind of a communication. Many applications run worldwide but this application will bring the wide world all together. The algorithm will run in such a manner that speech –text—image. This can also be implemented otherwise by face and gesture recognition for positive results. Multiple language implementation will generate a foundation for everyday purpose. no external devices or carry on devices will be required to establish a communication between two people. Even if the application turns out to be larger in size its satisfying the needs of the society. It's our whole and soul attempt of eradicating this barrier and yes our thought towards this project is the cutting edge to contemplate it has a whole.

X. REFERENCES

[1] William C. Stokoe. Sign Language Structure [microform] / William C. Stokoe. Distributed by ERIC Clearinghouse, [Washington, D.C.], 1978.

[2] William C. Stokoe, Dorothy C Casterline, and Carl G Croneberg. "A Dictionary of American Sign Language on Linguistic Principles" Linstok Press, [Silver Spring, Md.], New Edition, 1976.

[3] Nathaniel Blanchard, Michael Brady, Andrew M. Olney, Marci Glaus, Xiaoyi Sun, et al. "A Study of Automatic Speech Recognition in Noisy Classroom Environments for Automated Dialog Analysis"

[4] William C. Stokoe, Jr. "Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf" Oxfordjournals, 1960

[5] Clayton Valli, Ceil Laucas. "Linguistics of American sign language: An introduction" Gallaudet University press.

[6] Cognitive services. https://blogs.windows.com/ buildingapps/2017/02/14/cognitive-services-apis-speech

[7] Brown, R. D. (2000). "Automated generalization of translation examples." In Proceedings of the eighteenth international conference on computational linguistics (COLING-2000) (pp. 125–131). Saarbrücken, Germany, August 2000.

[8] Jaballah, K., & Jemni, M. (2010). "Toward automatic sign language recognition from Web3D based scenes." Lecture Notes in Computer Science, 6190, pp. 205–212: Springer.