Automated Audio to Text Converter

Dhyani T. Joshi[1] Dhruvi P. Joshi[2] Prof. Ajaykumar T. Shah[3]
[1,2]Student [3]Head of Department
[1,2,3]Department of Computer Engineering
[1,2,3]Alpha College of Engineering & Technology, India

Abstract— Language translation is the process of translating the information from one/more language/s or medium of expression to another language/s or medium of expression. It comprises of various stages such as interpretation, processing, and translation. There is no objection that translation is an ancient activity still being continued by the modern world for different purpose. A cumulative study of different techniques is done as per stages. This paper terminates with the decision on future direction for developing techniques in human computer interface system in different languages given in translation. The main aim of this project is to make a web-application capable of generating a text file from the given uploaded audio/video file in English language. This is done to enhance the horizon for the deaf and dumb person. This web-application is not for communication within themselves or another person. It is for personal use like self-learning or communicating to virtual world in which they face difficulty. By this it will become effortless for deaf and dumb people to gain extra knowledge and can access information early, and without facing any complication.
Key words: Audio To Text, Video to Text, Translation, Speech to Text

I. INTRODUCTION

Automated audio to text converter is crafted specifically for the differently-abled people. The intended purpose is to create an easy-to-use and associable translation system in order to help automate the process of translating the audio file into other languages. It has been deliberately kept as minimal as to help the differently-abled person to access the information of any language. One can translate the audio file/s of different languages and can read it from the generated text file generated. These text file/s can then be downloaded for future access. Although, this type of system is already used around the globe, the purpose of the research is to create a socially acceptable and easy-to-use platform which can become a handy source for self-learning.

II. LITERATURE REVIEW

The inspection of idea was taken place and it was known that, though differently abled people get the education in regular terms, but to cope with the outer world it is not in satisfactory amount. For that they either have to rely on a mediator to teach them and provide extra coaching, or refer to some books which may not be regularly gets updated. So we come up with the solution of converting an audio file to simply upload whichever audio they contemplate to convert and can translate it respectively.

Google translate is one of the huge achievement in terms of translation in more than 100 languages, in addition to that, 39 languages via photos and 32 via voice in conversation mode. Furthermore, 27 via real time video in augmented reality mode. Also those translated content can be saved or even redirected to ones email address as well. So we are quite fascinated by the process and were thinking in same direction to serve the society. Google translate administrates many functionalities such as written word translation which translates written word or text to a foreign language. Website translation that translates a whole webpage into selected language. Document translation that translates a document uploaded by the user to selected language in given extensions. In addition to that speech translation and mobile application translation, image translation and handwritten translation facilities are implemented.

Another inspection was of different public websites who provides translation for professional purpose for instance meetings or for presentation purpose. They equip multiple levels of service based on client needs, ranging from draft/summary translation up to certified translation, featuring to multistep translation and review the process accordingly to ensure the highest level of accuracy and meet the evolving needs of global business. Also there are plenty of websites who serves this purpose so as to communicate globally and shattering language as a barrier in business world.

III. METHODOLOGY

Automated audio to text converter is a web-application used for converting audio file to text file. It can also used for translating text file from one language to another language. This web-application is for deaf and dumb people for educating themselves. Using this web-application they can convert audio file of any webinar or any speech or lecture into text file. During research we found that live speech or chat can be converted into text file but there is no proper way to convert saved audio file into text file.

This web-application will work in following way-
– Users can upload audio or video file from their device.
– After uploading file users can convert the audio file into text file.
– Users can also convert live speech or chat into text file using microphone.
– The converted text file can be translated into different languages.
– Users can download the converted text file for future use.
– For queries users can refer to FAQ section of web-application or talk to us by using contact us option.
– Users can also refer to Help section for any kind of help they need while using the web-application.
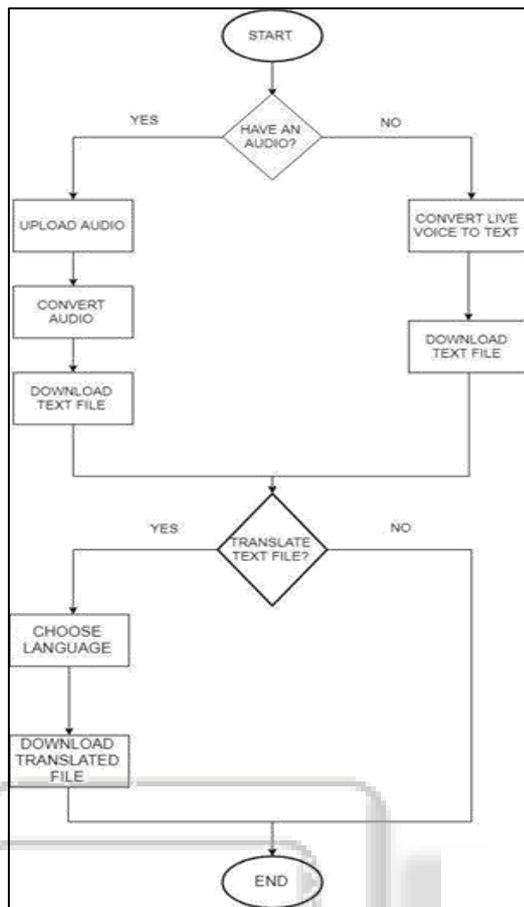
Fig. 1:

IV. SPECIAL FEATURES

The idea to this web-application was inspired by google translate. Likewise in google only live voice can be translated, in our application audio file can be translated and also the purpose behind it is whole at different level. One can translate the file in different languages available in application programming interface and all the possible efforts are made to improve the accuracy of project.

V. LIMITATIONS

- Accuracy of speech recognition vary with the following.
- Dictionary size and conceivability.
- Speaker dependency vs independency.
- Task and language constrain. • Accent of different origin.
- Sometimes machine cannot interpret the high frequency voice and can misprint the word spoken.
- As in present, we do not have higher processor in our system so it takes quite long to complete the process.

VI. FUTURE ENHANCEMENT

(Speed)Accuracy can be improved with better processing system. Greater use can be made of "Intelligent Systems" which will attempt to guess what the speaker intended to say, rather than only relying on what was spoken, as people can often make pronunciation and unintentional mistakes. Furthermore, video can be provided with subtitles. Though it is present in youtube, but if any video is other than youtube, this enhancement can be proved helpful.

VII. CONCLUSION

Automated audio to text converter will be helpful to deaf and dumb people and can also use by normal people. It will become easy for especially abled people to access the knowledge from multiple resources, which are available in the form of audio file. Furthermore, for normal people if they intend to translate the voice using microphone they can do so.

REFERENCES

[1] http://www.ijstr.org/final-print/june2015/Speech-to-text-Conversion-stt-System-Using-Hidden-Markov-Model-hmm.pdf
[2] http://ijarcet.org/wp-content/uploads/IJARCET-VOL-4-ISSUE-7-3067-3072.pdf
[3] https://patents.google.com/patent/US6820055B2/en?q=audio&q=text&q=conversion&oq=audio+to+text+conversion
[4] https://patents.google.com/patent/US7136462B2/en?q=audio&q=text&q=conversion&oq=audio+to+text+conversion
[5] https://patents.google.com/patent/US7174295B1/en?q=audo&q=text&q=conversion&oq=audio+to+text+conversion
[6] https://www.alexkras.com/transcribing-audio-file-to-text-with-google-cloud-speech-api-and-python/
[7] http://www.iosrjournals.org/iosr-jce/papers/Vol19-issue4/Version-3/C1904031418.pdf
[8] https://www.youtube.com/watch?v=CleLvQ Wy8-c
[9] https://www.youtube.com/watch?v=l-eD12KDmRc