Automatic Multiple Language Subtitle Generation for Videos

Aiswarya K R

Student,Dept. of Dual Degree Computer Appliications,Sree Narayana Guru Institute of Science and Technology N.Paravur,Kerala,India

---***---

Abstract - *Subtitles are textual content material derived from each a transcript or screenplay of the dialog or commentary in films, tv programs, video games, and as a result the like, commonly displayed at rock bottom of the display , however additionally can be at the best of the display screen if there is already textual content at rock backside of the screen. They can both be a type of written translation of a dialog all through a overseas language, or a written rendering of the dialog inside the equal language, with or besides brought statistics to aid viewers who are deaf or tough of listening to comply with the dialog, or humans that can't recognize the spoken talk or who have accent cognizance problems. In this paper for producing subtitles, frequently three modules are introduced. Audio extraction, speech recognition and multiple language subtitle generation.*

Key Words: Audio extraction, speech recognition, multiple language subtitle generation.

1. INTRODUCTION

Subtitles can appreciably make contributions to the viewing trip of the common viewer, even if they are no difficult of listening to or non-native. Furthermore, a subtitle aids comprehension and reminiscence retention for viewers, particularly when they are gazing fantastically technical standards being explained. Subtitles are a textual content model of the speech and audio information, wished to recognize the content. They are displayed inside the media participant and are synchronized with the audio. Normally, subtitles are downloaded from internet for videos .This is a difficult process. Consequently, To generate subtitles routinely thru the software program itself is a legitimate situation of research. At present, in Video LAN Client (VLC) media participant the subtitles have to be inserted first to the media participant and then it is synchronized with the song. The file inserted wishes to be .srt file containing the time intervals of the textual content spoken. It does now not receive .txt file to synchronize the subtitles as it solely carries the lyrics of the tune and no longer the time intervals of the spoken words. On the different hand, YouTube accepts each the .txt file and the .srt file for synchronizing the text. Nonetheless, software program producing subtitles except intervention of an man or woman the use of speech attention have now not been developed.

In this mission video file taken as enter , and the output of the project will be a video file with Subtitle. The entered video file taken and the audio extracted, and a .wav file generated. Now, the extracted .wav file is used to generate a .srt file the usage of Speech Recognition with the assist of cloud carrier recognized as Google API. Then the video with subtitle generated through fixing the srt file into video [1].

2. EXISTING SYSTEM

The best way to get subtitle is go to a subtitle website, search for your TV exhibit or film and down load the SRT file. The two most famous websites for getting subtitles are Subscene and OpenSubtitles.org. On OpenSubtitles, function a search and then you will see all the handy subtitles for specific languages at the bottom. The authors Abhinav Mathur, Tanya Saxena and Rajalakshmi Krishnamurthi collectively proposed a device for producing subtitle. That research paper resolves the above difficulty by converts an enter file to .wav format. Here 24% reduction rate has been performed in the dimension of the tune after the extraction. Then Speech Recognition of the extracted .wav file is applied and finally, .txt/.srt file is generated which is synchronized with the enter file.

3. PROPOSED SYSTEM

The venture "Automatic a couple of language subtitle technology for videos" generate subtitle for movies in any of the three languages English, Hindi and Malayalam as the desire of user. The task resolves the above difficulty thru three wonderful modules specifically audio extraction, speech recognition, and subtitle generation. Audio Extraction which converts an enter file of any structure supported via MPEG requirements to .wav(Waveform Audio) format. Then in the speech cognizance extracted .wav file is used to generate trendy subtitles file structure acknowledged as "SubRip Subtitle" or SRT. Result of the .srt file which consists of the phrases (lyrics) spoken in the audio file. Speech attention is accomplished the use of Google Translate API. In subtitle technology srt file is embedded in the video the place the lyrics are synchronized with the time and displayed with the video. Thus, this paper gives the customers a fundamental advantage of now not downloading the subtitles from the web as a substitute producing them robotic-ally with their preferred language.

User can also generate subtitle in default English language if the character does now not be aware of the language of the modern-day video. It additionally consists of some of the extra features like extract audio from video, consumer can view the srt file of that video at any time, srt file carries time sequence and corresponding speech that is generated with the assist of cloud carrier acknowledged as Google API[4].

4. METHODOLOGY

Mainly three modules are proposed for producing a couple of language subtitle generation. Audio extraction, speech recognition and multiple language subtitle generation.

4.1 Audio Extraction

This module aims to output an audio file from a video file. It takes as enter a file URL and offers audio layout of the output. The speech recognition requires a .wav audio file as input. Hence, it is fundamental to convert the video into .wav structure .The enter file should be of the codecs supported via FFMPEG standards such as .mp3, .mp4, .avi, .au, and .flac. FFMPEG is a complete, move platform solution to convert and move audio and video. Audio extraction presents a way to extract audio from a media file. It specifically makes use of ffmpeg for audio extraction and conversion to the layout required by using speech focus system. After extracting the audio from video files, it writes the file with .wav extension and then supplied to speech recognition module for in addition processing.

4.2 Speech Recognition

It is the key phase of the device which influences at once to overall performance results evaluation. When the audio extraction is completed, the speech cognizance section is carried out. Now, the extracted .wav file is used to generate a subtitles file format regarded as "SubRip Subtitle" or SRT file the use of Speech Recognition. Result is the .srt file which carries the phrases (lyrics) spoken in the audio file. Speech recognition is finished the usage of Google Translate API. Here performs voice recreation detection to find speech regions, makes parallel requests to Google Web Speech API to generate transcriptions for these regions, and interprets them to chosen language.

4.3 Multiple Language Subtitle Generation

This module is predicted to get a listing of phrases and their respective speech time-frames from the speech attention module and then produce a .srt subtitle file. In this module the srt file is then in addition embedded in the video the place the lyrics are synchronized with the time and displayed with the video. This is executed with assist of Moviepy video enhancing library in python. MoviePy is an open supply Python module for video enhancing that

can be used for primary operations (like cuts, concatenations, and title insertions), video composting, video processing, or to create superior effects. It can examine and write the most frequent video formats. Subtitles are textual content translations of the dialogues in a video displayed in actual time for the duration of video playback on the pinnacle of the screen. A normal .srt file has three sections first the Subtitle range indicating which subtitle it is in the sequence. Secondly, time that the subtitle should appear on the screen. Last, time for disappear the subtitle. And at ultimate the textual content accompanied via a blank line to point out give up of present day subtitle.

5. ALGORITHM

An algorithm is a step by means of step system to remedy logical and mathematical problems. A recipe is a correct instance of an algorithm due to the fact says what should be done, step by using step.

Step 1 : Start

Step 2 : Choose the video file to generate subtitle

Step 3 : Specify the supply language of the video

Step 4 : Extract the audio file from video

Step 5 : Select the choice if the person desires to convert to default language otherwise video generated with supply language.

Step 6 : SRT file generated and constant in to video

Step 7 : Play the video with Subtitle

Step 8 : Stop

6. IMPLEMENTATION

This paper is applied using high-level programming language Python the place Tkinter and PyQt4 acts as front-end with Sublime as platform. Python facets a dynamic kind gadget automated reminiscence management. It helps a couple of programming paradigms such as object-oriented, imperative, functional and procedural, and has a giant and complete general library. In this assignment, the speech attention of movies performed via Speech Recognition in Python the usage of Google Speech API. This utility can be used via humans to generate subtitles robotic-ally for movies via deciding on any of the three languages (English,Hindi,Malayalam) as the preference of user. User can additionally generate subtitle in default English language if the character does no longer comprehend the language of the contemporary video. It additionally consists of some of the extra features like

extract audio from video, consumer can view the srt file of that video at any time, srt file incorporates time sequence and corresponding speech that is generated with the assist of cloud carrier acknowledged as Google API.

6.1 Development Tools

Major equipment that are help for the implementation of software program utility.

6.1.1 Python

Python is an interpreted, object-oriented, highlevel programing language with dynamic semantics. Its support a lot of packages with variety of features, blended with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as proper as to be used as a scripting or glue language to join existing elements together .python's simple, convenient to study syntax emphasizes readability and consequently reduces the value of software maintenance. Python helps modules and packages, which encourages application modularity and code reuse. The Python interpreter and the big preferred library are available in supply or binary structure besides cost for all primary platforms, and can be freely distributed.

6.1.2 PyQt4

It is a Python interface for generating graphical user interface, one of the most powerful, and famous cross-platform GUI library. PyQt is a combo of Python programming language and the Qt library.PyQt4 helps Python v2 and v3.PyQt is well suited with all the famous working structures including Windows, Linux, and Mac OS. It is twin licensed, on hand underneath GPL as nicely as industrial license.

6.1.3 PIL

Python Imaging Library (abbreviated as PIL)(new variations recognized as Pillow) is a free library for the python programming language that provides help for opening, manipulating, and saving many special picture file formats. Pillow offers several trendy methods for photo manipulation. These include:per-pixel manipulations, masking and transparency handling, adding textual content to pictures and much more.

7. RESULT ANALYSIS

User can add their video and pick a language for producing subtitle as proven in figure1.Subtitle for video generated in any of the three languages English, Hindi, Malayalam as the preference of user.
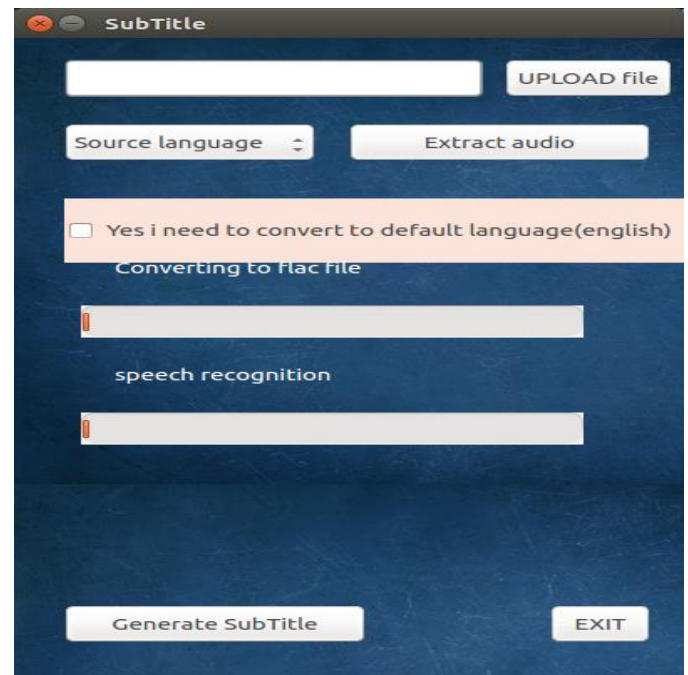

Fig - 1 : Home Page

Audio extraction works flawlessly by way of extracting audio file from video and save in task folder as .wav format. The WAV file is a loss-less format, which lets in the sound of the WAV file to be a shut duplicate of the unique audio source.

Fig - 2: wav file

Then recognize speech from audio via removing noise. Here performs voice undertaking detection to locate speech regions, makes parallel requests to Google Web Speech API to generate transcriptions for these regions, and translates them to chosen language and a srt file will be generated.

Fig - 3 : .srt file

Fig - 4 : .srt file of English Subtitle

Fig - 5 : .srt file of Malayalam Subtitle

Fig - 6 : .srt file of Hindi Subtitle

Finally, .srt file embedded in the video the place the lyrics are synchronized with the time and displayed with the video. The subtitle will be present on the top side of videos as shown in the below figures.
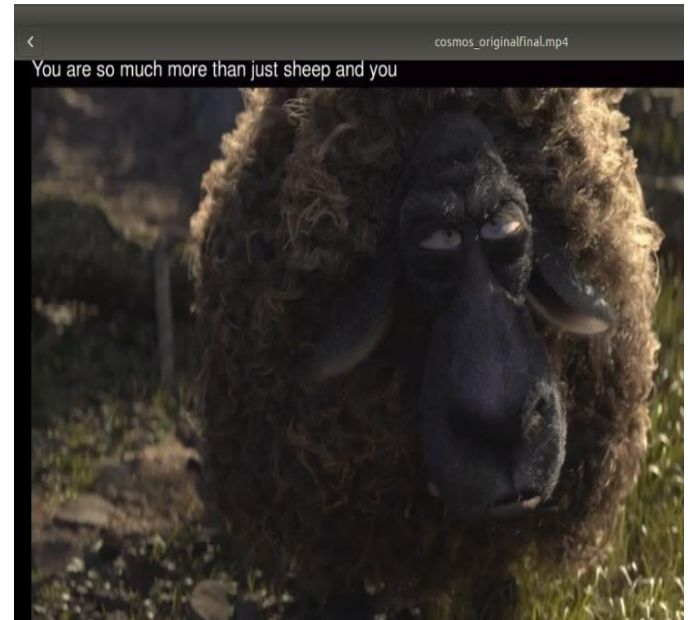

Fig - 7 : Video with English subtitle

Fig -8 : Video with Malayalam subtitle

Fig - 9: Video with Hindi subtitle

8. CONCLUSION

This paper presents the users a predominant advantage of now not downloading the subtitles from the net alternatively producing them automatically. Various research have been finished to accomplish every module Audio extraction, Speech recognition, Subtitle era of the project. The fundamental goal of this challenge is to strengthen an software that reduces the language problem. Currently subtitle of enter video file generated with any three languages(English, Malayalam, Hindi) as the preference of user.

9. FUTURE SCOPE

• It could be enhanced to a media player in future

• In this paper we can generate subtitle with three languages(English, Malayalam,Hindi). In future we can add more languages.

10. ACKNOWLEDGMENT

I would like to lengthen my heartfelt thanks to our HOD, Mrs. Kavitha C R, Department Of Computer Applications for the helps prolonged to me in the course of my direction of study. I am deeply grateful to my information Mrs. Dhanya P Varghese, Assistant Professor of Department of Master Of Computer Applications for the valuable guidance. My honest gratitude towards my household and pals for all support. And additionally thank to all Assistant Professors of Computer Application Department in Sree Narayana Guru Institute of Science and Technology for being a part of this long journey.

11. REFERENCES

[1] Aditya Ramani,Asmita Rao,V Vidya,VR Badri Prasad "Automatic Subtitle Generation for Videos",IEEE,23 April 2020.

[2] Abhinav Mathur; Tanya Saxena; Rajalakshmi Krishnamurthi,"Generating Subtitles Automatically Using Audio Extraction and Speech Recognition",IEEE, 02 April 2015.

[3] Kanika Mishra, Prerak Bhagat, Prof. Atiya Kazi,"Automatic Subtitle Generation for Sound in Videos",IRJET ,Volume: 03Issue: 02 | Feb-2016.

[4] Aditi Sankhe, Vishakha Patil, Subodh Patel, Siddhesh Bhagat4,"Video based Subtitle Generation",IJRASET,Volume 6 Issue III, March 2018.

12. BIOGRAPHIES

Aiswarya K R
D/o K V Ravi
DDMCA student at
SNGIST.N.Paravur,Ernakulam,
Kerala,India