International Journal of Future Generation Communication and Networking
Vol.14, No. 1, (2021), pp. 2602–2606

Artificial intelligence based Information Retrieval System using Video Montage Technique

Khushal Senghani[1*], Shaikh Mohammad Bilal Naseem[-2]
[1]*Department of Information Technolgy*
Somaiya Vidyavihar University Mumbai, India
[2]*Assistant Professor, Dept of Computer Science/IT*
Somaiya Vidyavihar University Mumbai, India
[1]*senghani.k@somaiya.edu*, [2]*mohammadbilal@somaiya.edu*

Abstract

Information system has proved to have an important role in improving our social ethics. In recent time during pandemic various campaigns run through social media, print media, television media and various offline form informing us to wear mask, wash our hands regularly. Instead, creating new forms of information we can use available data to create a new form of data. In this paper, we will understand more with the speech building example that uses available data and create a montage of it. In this system user will enter the written speech from the input then AI will search and collect all segments of video and audio from dataset, and then compiles it to form a single video.

Keywords: *Segmentation; NLP; Video Montage; Information system; Speech; Montage.*

1. Introduction

Information plays an important role in a person's life as it significantly influences the decision-making and lifestyle in general. As we can see in this recent time Plant-for-the-Planet for the climate change, GoVote[1] during the election campaigns, RallyforRiver[2] for the environment. Even campaigns in crucial times in during wartime, recession, etc plays a vital role. Historically, information was passed physically through print, speeches, banners/posters, but in recent technological advancements and social media revolution this information is passed digitally to greater mass at a single time. There has been many inspiring videos and audio. This has created a great potential to use this data so to form a new form of data that we can use for such information systems. We are going to discuss such type of system that uses previously generated data to create new one. In our system we are going to input a user generated written speech and the system analyze and collects all the videos needed then forms a new video which is the compilation of all the video that it has fragmented.

2. Problem Statement

This paper is mainly focused on generating informative videos using available videos by passing written speech into the system input through video montage technique described in propose solution. It help in creating speech.

ISSN: 2233-7857IJFGCN

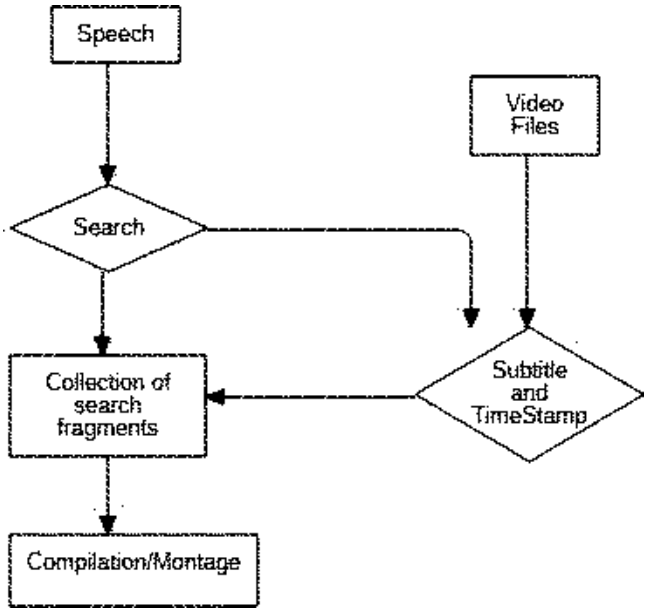
3. General Idea

First Fragment the video and audio file – When the user enters the speech AI system should search for the word which has been spoken in the video file and should fragment and store it for compilation

- Read and store the user speech – The user input typed speech, the system reads and analyze the speech. Then it looks it into the data for that terms available in our system fragments the data and stores the data.
- Compile – The stored fragments then needs to be stitched to form a single video for this video is sequenced same as the speech word sequence.

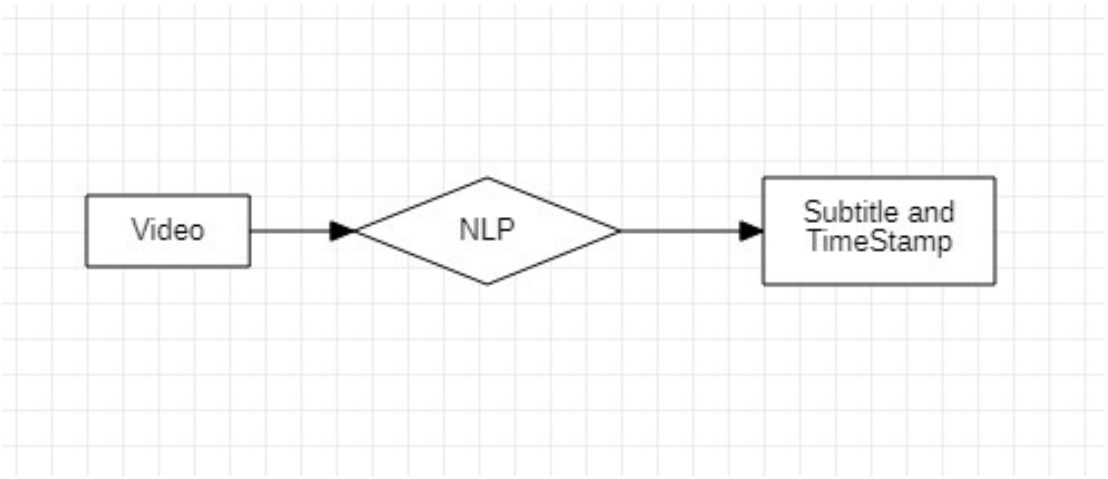
4. Proposed Model

For the AI to search in the video event in the video need to be tagged. This is then included in the metadata of the video this metadata is then searched while the user generate the video request. This metadata is the subtitle/caption of the video input. This caption is created using the NLP. This caption is then store in the metadata file each metadata.

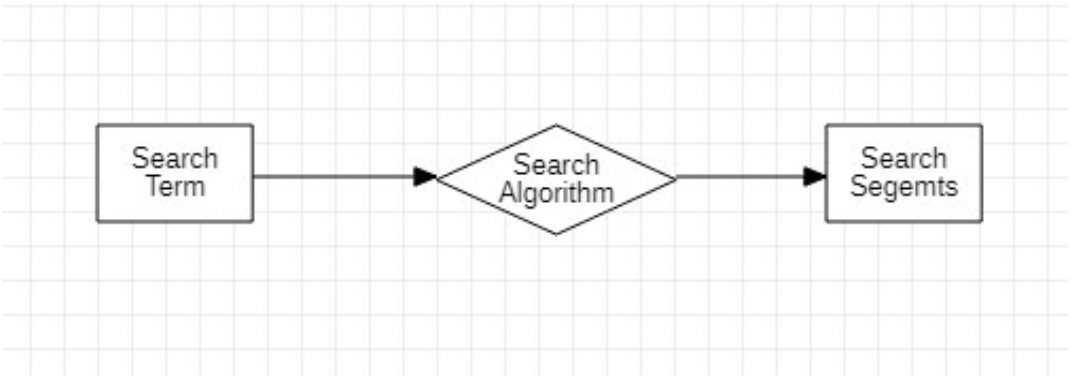


A metadata crawler be used to search for the video option available for the part of the sentence. A video composition method selects segments from storage medium and denotes serially connected sequences of the segments to thereby form a program sequence. The tactic employs metadata associated with each segment for easy manipulating the segments to form a program sequence. In this metadata associate to recall the segments whenever required. This enables us for further manipulation of segments but this requires greater processing power and time to label all the segments.

ISSN: 2233-7857IJFGCN



When the user enters the speech the AI system searches for the video segment of the speech that is available in the video database. A metadata crawler be used to search for the video option available for the part of the sentence. The video this stored into the final store where the system makes the compilation of the video segments form a single video
A strategy for organizing video by probabilistic converging of video sections incorporate the means of acquiring a majority of casings of unstructured video; creating video fragments from the unstructured video by distinguishing shot limits dependent on shading difference between back to back edges. This then stitches the video segments into one to form a single video.



6. Conclusion

While compiling all the segment it is possible that the frequency, volume and the pitch between two audios do not match which is undesirable in our system. If there are many segments of the same word it makes it difficult to make decision to choose. Different videos have different accent while hearing two different accent may cause some difficulties to the listener. Pauses need to be regulated. The new video can be created using existing data, but when in real world when the user input some words are not available in on our data then it excludes that part but this can be overcome by using more diverse data set.

7. Acknowledgement

ISSN: 2233-7857IJFGCN

We can use this system for creating speeches using the speeches of historic leaders, informative videos, for social cause. This system is an ideal fit to create movie trailers[6].

ISSN: 2233-7857IJFGCN

We can put movie into the data set then it generates trailer with emotional factor which can be added to this in the future scope.

8. Future Scope

As we have seen all possible challenges that our system faces iteration can be made to solve this problem

9. References

11.1. Journal Article

[1] https://govote.org/

[2] https://en.wikipedia.org/wiki/Isha_Foundation

[3] https://www.researchgate.net/publication/332376333_Inter-Sentence_Segmentation_of_YouTube_Subtitles_Using_Long-Short_Term_Memory_LSTM.

[6] https://dl.acm.org/doi/abs/10.1145/3123266.3127906.

11.2. Patent

[4] https://patents.google.com/patent/US4979050A/en.

[5] https://patents.google.com/patent/US7296231B2/en

ISSN: 2233-7857IJFGCN