# Video Summarization

**Pallavi Taru[1], Snehal Hiray[2], Shashank Gurnalkar[3], Ashlesha Gokhale[4]**

[1]*Computer Engineering (B.E.), P.I.C.T, Pune, Maharashtra, India*

[2]*Computer Engineering (B.E.), P.I.C.T, Pune, Maharashtra, India*

[3]*Computer Engineering (B.E.), P.I.C.T, Pune, Maharashtra, India*

[4]*Computer Engineering (B.E.), P.I.C.T, Pune, Maharashtra, India*

---

**Abstract -** *This paper proposes an innovative summarization of video in text format.Natural Language Processing is the vast area which has great importance when people started to interpret human language from one form to another form. A summarization is a system that produces a condensed representation of its inputs for user consumption.With the explosion of abundant data present on social media, it has become important to analyze this text for seeking information and use it for an advantage of various application and people.Generating the summary of video in text format will be useful to quickly get the relevant information of the video content as well as to get the abstract idea of video content of long duration.*

***Key Words***: NLP,video analysis,text analysis, summarization,speech recognition and synthesis.

## 1.INTRODUCTION

In today's world the abundant information available online but time is limited . With the advent of personal mobile computing devices, we are being presented with a barrage of information every minute. This makes it increasingly important to consume as much information as possible in the least amount of time while eliminating irrelevant and redundant data. The Video is another domain which falls prey to this information overload.

Text Summarization is an approach that can be used to get text format summary of the video. This basically uses Natural Language Processing principles and algorithms to generate efficient Summaries.

## 2. MOTIVATION

There are some videos which is of very long duration and viewers of that video not have much time to go through the whole video.The viewer just wants to know an overview of that video.Many times it also happens that viewer goes through video on the particular topic and at the end of viewer come to know that video is not a relevant topic for which viewer is searching. It is a dire need of the day to save time and grasp just summary of video which is in text format. This problem can be solved using Intelligent Summarization of Videos which will be useful for educational purpose where the time of students can be saved and they will have like notes of that video.

## 3. ARCHITECTURE

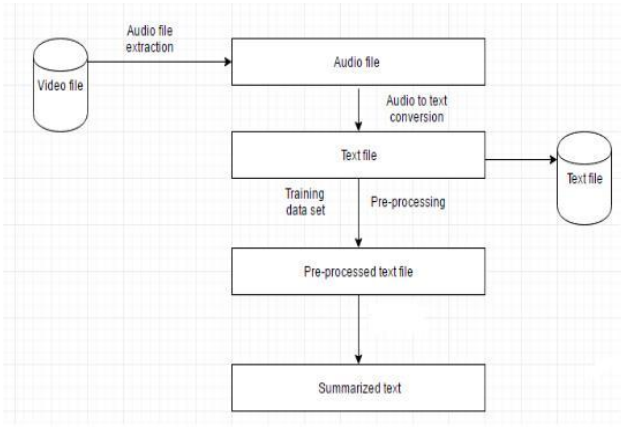


**Fig -1**: Architecture diagram

In our system,the main task is to get the text file of video content from video.By using various API's text file of that video can be achieved.Following steps explains the architecture in detail:

A) *Conversion of Video file into an Audio file:*

Firstly we will extract audio from video using API.

B)*Conversion of Audio file into a Text file:*

Then next task is to convert that audio to text using API.So,we will get the text file that video.

C)*Pre-processing of Text:*

Pre-processing method plays a very important role in text mining techniques and applications. It is the first step in the text mining process.Pre-processing involves stop

words removal, stemming etc.

D)*Summarization of Text file:*

This text file is given to the summarization system and generates summary of the video.This summarization module uses extractive summarization approach to get summary.From that generated summary user can get know whether that video satisfies his requirements or not.Ultimately it saves users time to go through the whole video.

## 4. IMPLEMENTATION METHODOLOGY

The main task in this system is to generate an effective summary of the video which depends on summarization technique used .
There are two types of summarization one is Abstractive summarization and second is Extractive summarization.Abstraction based summarization involves paraphrasing sections of the source document.In general, abstraction can condense a text more strongly than extraction, but the programs that can do this are harder to develop as they require the use of natural language generation technology, which itself is a growing field.Extraction techniques merely copy the information deemed most important by the system to the summary (for example, key clauses,sentences or paragraphs),deemed to contain a document's essential information, and of assembling these units in an adequate way.[2]

The text extraction based method is used to summarize documents because of less computationally intensive nature, ease of scalability and availability of various techniques for analysis.[3]In 1958, Luhn does the research about automation summary and introduces the computer to the text extraction for the first time. In the next more than half a century,keywords extraction methods based on statistics have been widely developed, commonly used statistical information includes word frequency, co-occurrence frequency, TF-IDF and so on[1].In recent years,keywords extraction algorithms based on graph model have a rapid development.[1]

The algorithm basically involves preprocessing the words on a corpus followed by graph based text ranking based on relevance. It is an unsupervised graph based ranking algorithm. It takes into account the keywords, frequency, relationship between sentences and the distortion measure. The graph is connected, undirected/directed and represents the text. Each sentence is represented by a vertex.[3]

## 5.CONCLUSION

In this paper we used extractive summarization approach instead of abstractive.The extractive approach is more simple as compare to abstractive approach.In this approach it does not require deep linguistic knowledge.
we are just using one video for summarization but we can do the summary of more than one video of same type in the text format.For example,if we are having three videos on one topic then we can summary of that topic by summarizing the that three video.
The goal of this summarization system is to save the time of viewer and give an effective summary of the video.

## REFERENCES

Wengen Li,Jiabao Zhao,"TextRank algorithm by exploiting Wikipedia for short text keywords extraction",2016 IEEE 3rd International Conference on Information Science and Control Engineering,DOI 10.1109/ICISCE.2016.151.

Ajinkya Zadbuke,Sahil Pimenta,Deepen Padwal, Varsha Wangikar, "Automatic Summarization of News Articles using TextRank," volume 6,Issue 3, Mar. 2016.

https://en.wikipedia.org/wiki/Automatic_summarization

Nitin Agrawal,Shikhar Sharma, Prashant Sinha, Shobha Bagai,"A Graph Based Ranking Strategy for Automated Text Summarization",DU Journal of Undergraduate Research and Innovation

https://en.wikipedia.org

Rada Mihalcea,"Graph-based Ranking Algorithms for Sentence Extraction Applied to Text Summarization",University of North Texas