Video Summarization using Histogram Analysis

Yash S. Patange,Prashant P. Porwal, Akhil S. Raut, Smita K. Jawale

Student, Student, Student, Assistant Professor.
Department of Computers, Vidyavardhini's College of Engineering and Technology, India.

Abstract :Everybody in our world is so busy with their lives that taking the time out for some activities is nigh difficult. Not everybody can sit through an entire match of their favorite team and support them. Neither all can watch an entire livestream of an important event occurring in history. Through this project, we aim to condense the entire full-length video into a 'highlight stream'. Our project, Video Summarization cuts and summarizes an entire video into a small video clip of relevant information. Thus, allowing all those people who couldn't get time to watch the stream, watch the highlight stream and still enjoy it. We will provide an interface through which the user can specify which video he/she wants to summarize. Once the video is selected summarization process begins and then after its end user has the summarized version of the original clip.

IndexTerms **- k-means, multimedia, summary, video summarization.**

I. INTRODUCTION

In recent years, because of the rapid growth in multimedia information, the advance in internet communication and digital video technologies, multimedia information indexing and retrieval has become more and more important and lots of research efforts have been devoted to the video retrieval and video analysis based on audio or visual features. This analysis shows that, when developing retrieval applications and video indexing, we first have to consider the issue of structuring the huge and rich amount of heterogeneous information related to video content. In addition, to retrieve information from the audio or visual content is a very challenging since it requires the extraction of high-level semantic information from low level audio or visual data.

Video summarization is an important process that facilitates faster browsing of large video collections and also more efficient content indexing and access. There are two main video summarization techniques in the literature: static video summarization (video summary) and dynamic video summarization (video skimming). In order to summarize a video, most of the methods have consists on visual features computed from video frames. Also, there are methods that consider the semantic meaning implied in the video to produce a more informative summary. Due to the increasing volume of video content on the Web, and the human effort taken to process it, new technologies need to be researched in order to develop efficient indexing and search techniques to manage effectively and efficiently the huge amount of video data. One of the most upcoming and advancing research areas is Video summarization. As the name implies, video summarization is a mechanism to produce a short summary of a video to give to the user a synthetic and useful visual abstract of video sequence, it can either be a images (keyframes) or moving images (video skims).

II. EXISTING WORKS

There is a large amount of research being carried out to generate a concise video summary. The most commonly used methods are keyframe extractions. They select key frames based on shot similarities and merge these keyframes as video summary [1]. Yang and Wei proposed a video summarization method based on genetic algorithm. In this method, frames are first generated. Then the least similar frames are selected based on histogram differencing. These frames are the keyframes of that video. Next, a fitness function is defined and based on this fitness function optimal frames are selected [2]. Another method uses Singular Value Decomposition. In here the video is divided into sub shots based on different distance measures. From these sub shots, keyframes are extracted by Singular Value Decomposition [3]. Another popular method is video summarization based on event detection. The moving objects motion is detected and the events are predicted. Based on these predictions the summary is generated based on the most interesting event and small number of frames [4].

III. RESEARCH METHODOLOGY

3.1Background Theory

A. Keyframes

Keyframes are those frames that define a start and end point in a smooth transition. They are called "frames" because they are measured in frames on a strip of film. Keyframes define the timing as well as what the viewer will see in the video. Since only 2 or 3 keyframes are present in a span of a second "inbetweener" frames are used to fill out these gaps between the keyframes.

B. Frame Differencing

In frame differencing method there is no need to identify the shot boundary. Whenever drastic change is identified, the next frame is selected as key frame. The identification of change is depended on threshold value. If threshold value is large, then only drastic change can identified and if it is less, then minor change can identified. The more the threshold value, the less the number of key frames, the less the threshold value, the more the number f key frames. So the number of key frames can be controlled by the threshold value.

C. Clustering

Clustering is a process of grouping data or objects that are more similar to each other than the other data or objects present in a given set. It is a common technique of statistical data analysis and is used widely in machine learning.

3.2 ALGORITHM

For a video sequence V, generate the frames $f_i$where i is from 1 to n. 'n' is the last frame generated. Frames are generated as per the video quality. As our example we took videos with 25 frames per second. Histogram analysis is then performed on these frames.

Algorithm:V**ideoSummarization**.
Inputs: Video *V* to be summarized.

1. Capture the video *V*.
2. Generate frames *fi* for the video.
3. Calculate the Histogram for each frame.
4. Calculate the histogram difference between i^{th} and $(i+25)^{th}$ frame.
5. Form 3 clusters of these difference values using K-Means Clustering Algorithm.
6. Select clusters 1 and 3, as these contain the key-frames of our video.
7. Form buckets of time *t* seconds and sort the frames from clusters 1 and 3 corresponding to the seconds they belong to in the video.
8. The buckets with the maximum number of corresponding frames, above a certain threshold, are selected.
9. The *t* second clips are extracted from the original video and merged together.
10. After merging summarized video is the output.

In three clusters formed the 1^{st} and the 3^{rd} clusters represent scene changes. The 2^{nd} represents scene transitions that are largely inconsequential and thus aren't key-frames. Also the 2^{nd} cluster is on the elbow, thus they are not relevant.

IV. RESULTS AND DISCUSSION

4.1 Experimental Result

- Extensive testing was performed on a number of videos.
- Each video depending upon its frame rate, video quality and length differed in total time taken.
- The algorithm applied by us ensures no video will take an exceedingly long time. Each video will take a reasonable time to generate summary.

table 1: experimental results

Sr. No.	Video Size (in MB)	Video length (in Minutes)	Summary Size (in MB)	Summary length (in Minutes)
1	56.3	13.04	2.56	1.51
2	122	12.10	2.50	3.09
3	11	4.05	5.10	2.31
4	139	17.41	6.46	2.48
5	107	12.54	8.1	1.55

Table 1 shows the performance of the algorithm when applied. The time taken by the algorithm is dependent on the video quality of size. Higher the frame rate of the video, time taken is more. Even though it depends upon quality, experimenting showed us

that the algorithm takes a reasonable amount time. Average time taken for a video of size 7-10 minutes with a framerate of 25 to 29 frames/second to be summarized is 12 mins.

Figures

figure 1

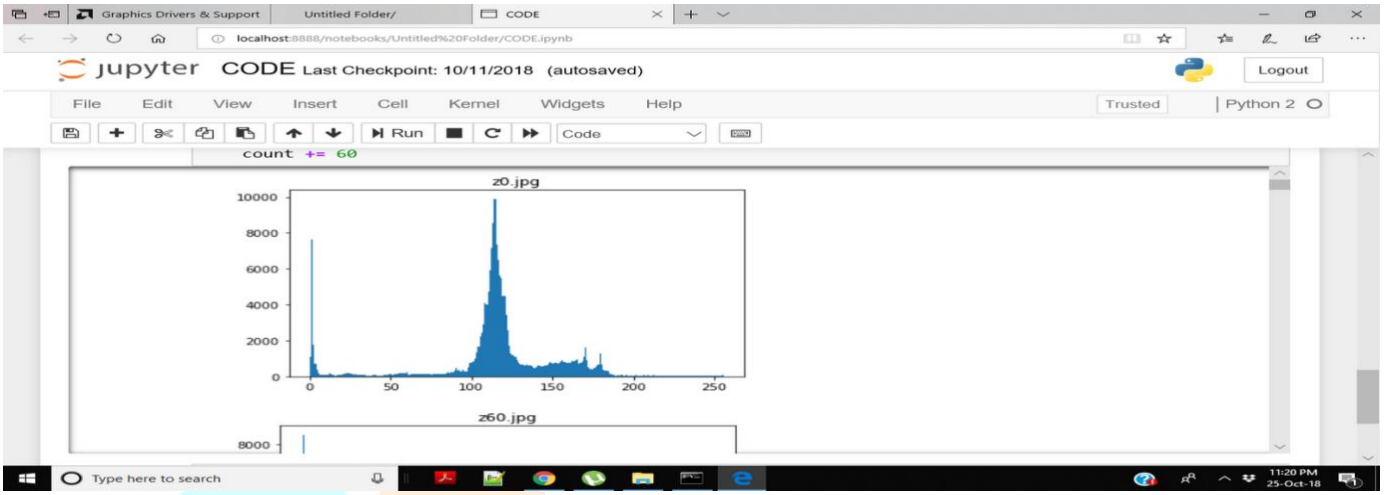
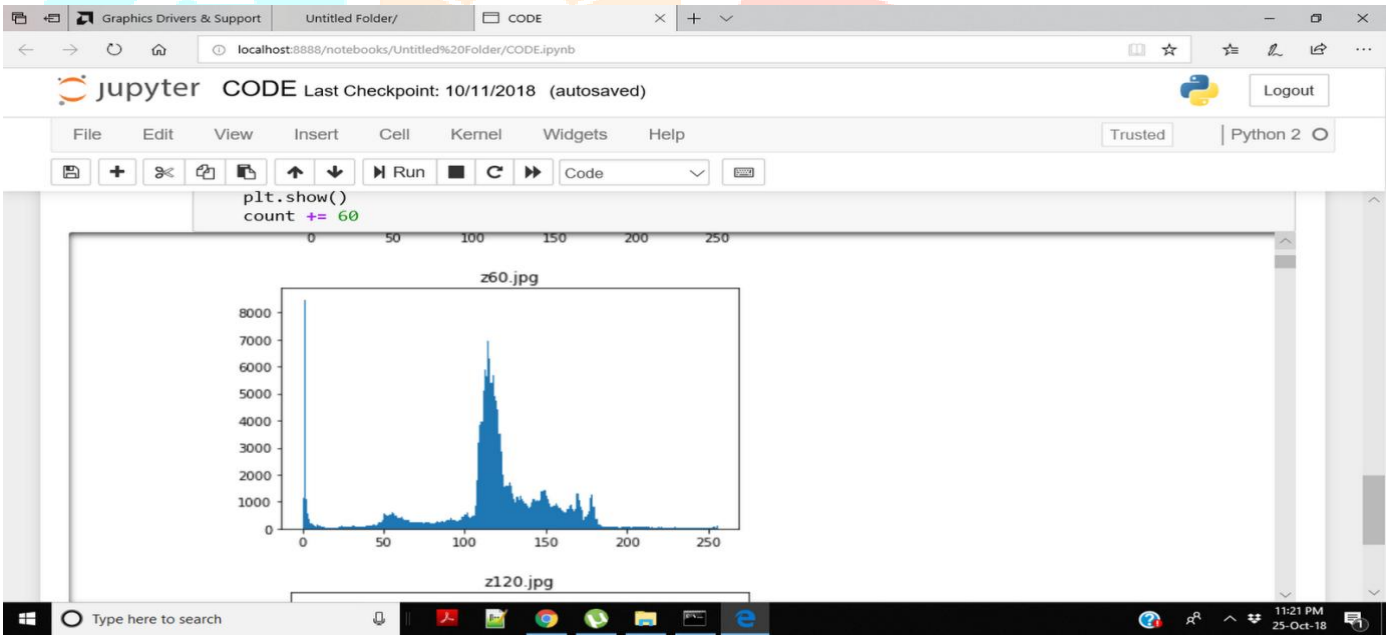


figure 2



Figures 1 and 2 show the histogram of the frames generated. This histogram is used further for selecting the key-frames.

IV. CONCLUSION

The methodology proposed here will certainly help laymen who are not computer savvy enough to watch highlight streams of their own choice of video. It can also be used to obtain key scenes or sequences of any type of video. Our proposed Algorithm is efficient. The throughput of the program depends upon the size and framerate of the video. The larger and better the video more time is required for output. During the testing phase we took many different samples and found that time taken is not too much.

V. FUTURE SCOPE

Our proposed method has an incredible amount of scope in the field of video processing. This can be used in video hosting websites to show us what the video is about in its thumbnail.

GoPro cameras are lot popular in the market right now. When people go on an adventure, they tend to record the whole trip in this camera. After returning from the journey it gets tedious for people to see the hours and hours of footage and edit it down to their fancy. By using video summarization, they can input the video and the software will do the work. Output is a highlight reel of their vacation.

People love to watch sports. They will sometimes spend lot of time in watching their favorite teams play the match. Sometimes it may not be possible for people to spare such large amount of time. By applying machine learning and event detection, video summarization can perfectly form the highlight stream of the match for the viewers pleasure.
Another great use of this project is for security purposes. CCTV cameras record footage 24/7. This is a lot of data to view in case of any mishap. With the help of artificial intelligence and object detection and tracking, data to be viewed can be reduced drastically and help catch any incident easily in the summarized video.

REFERENCES

[1] Nada Jasim Habeeb, Rana Saad Mohammed, MuntahaKhudairAbbass, "Surveillance Video Summarization Based on Histogram Differencing and Sum Conditional Variance", International Journal of Computer and Information Engineering Vol:10, No:9, 2016.

[2] Gao, Yue and Qiong-Hai Dai, "Shot-based similarity measure for content-based video summarization." In Image Processing, 2008, ICIP 2008. 15th IEEE International Conference on, pp 2512-2515. IEEE, 2008.

[3] Lei, Shoashuai, Gang Xie, and Gaowei Yen "A Novel Key-Frame Extraction Approach for both Video Summarization and Video Index". The Scientific World Journal 2014(2014).