A Survey on Video Summarization Techniques

Tinumol Sebastian
Dept.Of Computer Science
College Of Engineering Poonjar

Jiby J. Puthiyidam
Dept.Of Computer Science
College Of Engineering Poonjar

ABSTRACT
Video summarization methods attempt to abstract the main occurrences, scenes, or objects in a clip in order to provide an easily interpreted synopsis. The main aim of Video summarization is to provide clear analysis of video by removing duplications and extracting key frames from the video. Video Summarization will divide the frames of the video into blocks and calculating the mean, variance, skew, kurtosis histogram of every block and comparing the same with the corresponding blocks of the next frame. There are many different methods used for Key frame extraction in video Summarization. Some important methods are compared. The frame with highest mean is selected as the key frame. The best method is selected based on the color distribution.

General Terms
Algorithm, Database, Samples.

Keywords
Video Summarization, Frames, video sequence, key frame extraction, mean, variance, Histogram, skew.

1. INTRODUCTION
Nowadays, we get the information from the digital media such as news, movies, T.V shows, internet etc. Due to time limitation, it is difficult for an individual to watch its full content. To reduce this limitation, we use video summarization. Video Summarization is a process of creating and presenting, an abstract view of entire video within a short period of time. This technique will generate the summaries of the videos which contain maximum information that makes individual more easier.

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 an image (key frames) or moving images. Video summarization is a vital process that facilitates well-organized storage, quick browsing, and retrieval of large collection of video data without losing important aspects. Basically, there two different kinds of Video Summarization techniques in literature. One is static Video summarization and the other is dynamic Video Summarization. Static video storyboard summary involves a set of key frames from original video and there is no restriction with time and sequence issue. On the other hand, it selects the most significant, small, dynamic portions of audio and video in order to generate the video summary. Most of the methods consist of visual features, computed from video frames. Also, there are methods that judge the semantic implication shown in the video to produce a more informative summary. In terms of browsing and navigation, a good video abstract will enable the user to get maximum information about the target video sequence in a specified time limitation or adequate information in the minimum time. Manually generated summaries can support users in navigating large video archives and in taking decisions more efficiently regarding selecting, consuming, sharing, or deleting content.

Video Summarization starts with basic steps: Shot boundary detection and Key frame extraction. Shot boundary detection process will divide the video into shots. From each shot a key frame is selected, specify maximum information about the shot. On using Higher Order Color Moments (VSUHCM) for shot boundary detection and key frame extraction based on Higher order color moments image Histogram, skewness and kurtosis which convert a long boring video into a short thrilling video.

Techniques in video summarization touch various domains, such as movies, sports, news, home videos, e-learning, etc. Different techniques used for video summarization such as object base summaries, event based summaries, content based summaries, feature based summaries etc. The methods VSUHCM, DT, STIMO, VSUMM1, VSUMM2 and OV summaries are compared and the comparison analysis is performed based on the color distribution.

2. RELATED WORK
A Video Summarization is a summary representing an abstract view of the original video sequence and can be used for video browsing and retrieval systems. Different methods used to select key frames. These methods are mainly based on low level features such as color Histogram, edge Histogram, frame correlation and it does not consider the colors in the image:-

2.1 Delaunay Triangulation
Padmavathi Mundur,Yong Rao ,Yelena Yesha proposed a method for Video Summarization using Delaunay Triangulation[1], which generates video summaries by capturing the visual content of the original videos in fewer frames Most clustering methods require user-specified parameters or prior knowledge to produce their best results, this demands pre-processing or several trials. Both are extremely expensive and inefficient, because the best-fit parameters are not easy to get. But, video summarization technique by using Delaunay clusters generates good quality summaries with fewer frames with less redundancy Compared to many other clustering techniques. The Delaunay clustering algorithm is fully automatic with no user specified parameters and is well suited for batch processing.

An automatic video summarization technique based on Delaunay clustering is well suited for batch processing of videos without user intervention. Each cluster shows very high-quality clustering performance in terms of similarity of content. The clustering process results in clusters of different sizes based on the content represented in the original video.

2.2 K-means clustering algorithm
An Automatic Video Summarization and Quantitative Evaluation[2] proposed by Sandra E.F, de Avila, Antonio da Luz Jr., Arnaldo de A, and Matthieu Cord. The K-means clustering algorithm is a simple method for estimating the mean (vectors) of a set of K-groups. It is an unsupervised clustering algorithm. This is a simple and easy way to classify a given data set through a certain number of clusters (assume

k clusters) fixed a priori. The main plan is to describe k centroids, one for each cluster. These centroids should be placed in a cunning way because of different location causes different result. So, it is better to place them as far as possible from each other. The next step is to take each point belonging to a given data set and associate it to the nearest centroid. When no point is pending, the first step is completed. At this point we need to re-calculate k new centroids as bar centers of the clusters resulting from the previous step. After we have these k new centroids, a new binding has to be done between the same data set points and the nearest new centroid. A loop has been generated. This loop may notice that the k centroids change their location step by step until no more changes are done.

In this technique, video summaries are generated based on visual features and uses K-means clustering algorithm to cluster the frames. Quantitative Evaluation is based on quality of summary involving different users. This is a simple and efficient approach for video summarization and a method to quantitatively calculate the video summaries quality. The approach was applied to a sample of 20 videos selected from the Open Video Project [10].

2.3 Algorithm On Shot Boundary Detection And Key Frame Extraction

A method called shot boundary detection and key frame extraction is used in video summarization using higher order color moments by Mrs. Poonam S. Jadhava, Prof. Dipti S. Jadhav. Shot boundary detection is an early step for the majority of the video applications involving the understanding, indexing, characterization, or categorization of video, temporal video segmentation has been an active topic in content based video analysis .The key frame extraction is main task in producing summaries and skims of video. Shot boundary detection is the process of identifying the significant content change in the video. The key frame is extracted by using a reference frame approach per shot. The method has been tested on various videos and is able to detect the abrupt cuts with high precision and accuracy.

In this approach, a shot is defined as the consecutive frames from the start to the end of recording in a camera. It shows a continuous action in an image sequence. Deleting redundant information is achieved by segmenting the video into shots. A shot is a continuous recording of video content without breaks in a scene. Then, key frames may be extracted from each shot with different techniques based on pixel-to-pixel comparison, histogram-based comparisons, motion flow vectors, etc. This process is called Shot Boundary Detection. Shot Boundary detection is based on automatic threshold and by using reference key frame extraction videos are extracted.

2.4 Still and moving

Marco Geraci and Montenegro has proposed a technique STILL and MOVING (STIMO) is a fast clustering algorithm used to select visual contents using HSV color distribution. STIMO produces still and moving storyboards and allows advanced users customization (e.g., users can select the storyboard length and the maximum time they are willing to wait to get the storyboard). STIMO is based on a fast clustering algorithm that selects the most representative video contents using HSV frame color distribution. Experimental results show that STIMO produces storyboards with good quality and in a time that makes on-the-fly usage possible.

The summarization technique designed to produce on-the-fly video storyboards. STIMO produces still and moving storyboards and allows advanced users customization (e.g., users can select the storyboard length and the maximum time they are willing to wait to get the storyboard). STIMO [5] is based on a fast clustering algorithm that selects the most representative video contents using HSV frame color distribution. Experimental results show that STIMO produces storyboards with good quality and in a time that makes on-the-fly usage possible.

2.5 Edge Change Ratio

Dr. Shobha G and Azra Nasreen uses Edge Change Ratio[8] to detect the shots of video and from every shot key frame is selected. Edge change ratio algorithm is used for detecting shots of the video and key frames are extracted. The extracted key frames represent the video as a whole and can be used in variety of applications such as content based video retrieval, video summarization etc.

In order to obtain useful information from video, reduce redundancy in a video by extracting key frames. Key frames represent the well-known content of video. This paper presents algorithm for key frame extraction. Edge change ratio method extracts the key frames representing the content of whole video. The extracted key frames can be used in various applications like content based video retrieval, video surveillance, video summarization etc. In this ratio of entering and exiting edges of consecutive frames are calculated. The frames having maximum difference between the edge change ratio of is considered as key frame. The extracted key frames provide the summary of the whole video. The method is highly efficient and can identify the fast moving objects easily.

2.6 Vscan

VSCAN is an enhanced Video Summarization using Density based Spatial clustering by Karim M. Mohamed, Mohamed A.

Ismail, and Nagia M. Ghanem. A novel approach for generating static video summaries. This approach is based on a modified DBSCAN clustering algorithm to summarize the video content utilizing both color and texture features of the video frames. The paper also introduces an enhanced evaluation method that depends on color and texture features. Video Summaries generated by VSCAN are compared with summaries generated by other approaches found in the literature and those created by users.

The video summaries generated by VSCAN have a higher quality than those generated by other approaches. Here we

DBSCAN clustering algorithm to summarize the video content using both color and texture features of the video frames. We combine both color and texture features to overcome the negative aspect of using color features only as in other approaches. Also, the gain of using a density-based clustering algorithm is that, VSCAN [11] could overcome the drawback of determining a priori number of clusters; thus, the additional step needed for estimating the number of clusters is avoided. An advantage of using a modified DBSCAN algorithm is that, VSCAN also be able to notice noise frames automatically without extra calculation.

3. COMPARISON

Techniques	Algorithm	Features	Aim	Frames	Key frames
VSCAN(An Enhanced Video Summarization using Density based Spatial Clustering)	DBSCAN clustering algorithm.	Extract both color and texture features of the video frames.	Generating static video summaries.	3266	9
VSUMM1	K-means clustering algorithm.	Extract color attributes and visual features.	Generate Dynamic video summaries	3266	15
STILL AND MOVING	k me K- medoids clustering algorithm k	Extract color and motion vector.	Generate dynamic video	1916	19
DT(Delaunay Triangulation)	Delaunay clustering algorithm	By capturing the visual content of the original videos in fewer frames.	Generate Dynamic video summaries	1917	7
VSUHCM(VS using Higher Order Color Moments)	Algorithm for Shot Boundary Detection and Key Frame Extraction.	Majority of key features are extracted.	Generate Dynamic video summaries	3213	**10**

4. CONCLUSION

Key frame extraction is an important part of many video applications, like video indexing, browsing, and video retrieval. This survey concludes that there are several techniques used for video summarization. VSUHCM method becomes more effective and relevent in terms of the number of key frames extracted and color distribution. Thus the quality of the video is improved. This method allows the user to generate a short relevant video from a long video. It improves the quality and also increases the efficiency of the summarized video.

5. REFERENCES

[1] Padmavathi Mundur, Yong Rao, Yelena Yesha,(2006)'Key frame-based video summarization using Delaunay clustering' International Journal on Digital Libraries April 2006, Volume 6, Issue 2, pp 219-232.

[2] Sandra E. F. de Avila, Antonio da Luz Jr., Arnaldo de A. Araujo, and Matthieu Cord, (2008). `VSUMM: An Approach for Automatic Video Summarization and Quantitative Evaluation', XXI Brazilian Symposium on Computer Graphics and Image Processing IEEE

[3] Zhao Guang-sheng, (2008). `A Novel Approach for Shot Boundary Detection and Key Frames Extraction', IEEE.

[4] Miss.A.V.Kumthekar, Prof.Mrs.J.K.Patil, (2013). `Key frame extraction using color histogram method', International Journal of Scientific Research Engineering and Technology (IJSRET), Vol 2,pp 207-214

[5] Marco, Geraci and Montenegro, (2010)' STIMO: Still and Moving video storyboard for the Web Scenario' Journal Multimedia Tools and Applications, Volumes =46, issue1,January 2010,pages 47-69.

[6] Xiaohua He1 and Jian Ling, (2012). `A Video Summarization Method Based on Key Frames Extracted by TMOF', IEEE.

[7] Mr. Sandip T. Dhagdi, Dr. P.R. Deshmukh, (2012). `Key frame Based Video Summarization Using Automatic Threshold and Edge Matching Rate', international Journal of Scientific and Research Publications, Vol 2, ISSN 2250-3153.

[8] Azra Nasreen, Dr Shobha G, (2013). `Key Frame Extraction using Edge Change Ratio for Shot Segmentation', International Journal of Advanced Research in Computer and Communication Engineering, Vol 2.

[9] Proceedings, published by Springer Berlin Heidelberg, (2008). `Video Summarization: Techniques and Classification ', International Conference, ICCVG 2012, Warsaw,Poland,pp.24-26

[10] The Open-Video Project: http://www.open-video.org

[11] Karim M. Mohamed, Mohamed A. Ismail, and Nagia M. Ghanem (2014) VSCAN: An Enhanced Video Summarization using Density-based Spatial Clustering. Computer and Systems Engineering Department Faculty of Engineering, Alexandria University Alexandria, Egypt.