

Key Frame Extraction From Video Sequence : A Survey

¹Kavita Sahu, ²Prof Miss Satya Verma

¹Research Scholar, CSE Dept. MATS University, Raipur

²Asst. Professor, CSE Dept. MATS University, Raipur

Abstract— In the present era of IT and communication technology, application of video based information is increasing day by day. This attract the researcher to work more in the video processing. With decreased cost of digital storage device, availability of higher transmission rate by introducing 4-G technology along with the improved compression method, it is now possible to make the digital video available to all. Each and every video contain some vital information which need to be extracted without losing any information from the video, much attention is now given to the video processing technology. Video sequence basically consist of number of still images called frames. Key frames are the frames of the video which contain some vital information. Extraction of these frames is one of the important task in video processing. Extraction of these key frames which reflects the actual contents of the video in very essential for efficient browsing and retrieval of the video sequence from a large database of video. Such kind of operation is very useful for developing the advanced digital video technology. This paper present a various method of key frame extraction proposed in the past.

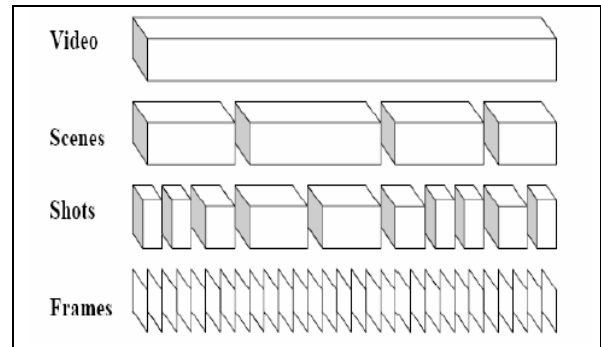


Figure1 Video to Frame Conversion Steps

All the frames in which scene in the video gets changed contain important information and hence called key frames. If these frames are extracted from the video then we can shrink the video or compress the video.

This makes the video smaller and hence require much lesser amount of information and at the same time it also contain all the vital information. The time and resources required for processing these video is also less. Analyzing such kind of video with lesser number of frames can improve the performance of the system manifold. All the frames in the video which contain vital information are called key frames.

These key frames is the summary of the video characteristics. Shot based method is one of the algorithm for extracting the key frames from the video. In this method, shots of the original video sequence are first detected and then one or more key frames are extracted from each shots. In this method, shot transition is detected. There are two methods of shot transition detection i.e. pixel based method and histogram based method.

1. INTRODUCTION

Digital process of extracting the frames of the video which represent the content of the video is known as the key frame extraction.

Now a days everything is shrinking, there is need to shrink the digital data as well. Compression of the data is one of the important steps in today as information technology. Compression of the data can make the communication and file transfer faster. It also require lesser space. It is well known facts that video consume enormous amount of space in digital storage. In any video so many information and hence the frames contain insignificant information which are of no use if information of video is concern.

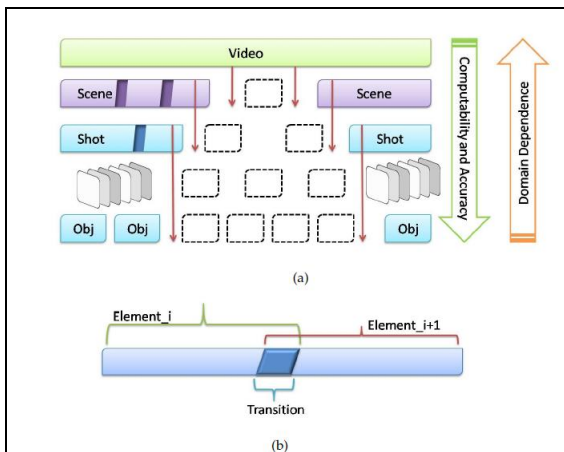


Figure 1.2 (a) Video Structure (b) Video Elements

Both the methods have their pros and cons. Pixel based method are very sensitive to the object motion and hence it is widely used where camera or object movements are required to be detected. Since in this method, pixels are compared therefore it is time consuming. On the other hand, histogram based method does not consider the local information. Two entirely different images may have the similar histogram.

II. CLASSIFICATION OF KEY FRAME EXTRACTION METHODS

Conservative approach of extracting the key frames from the video sequence preferred to extract more key frames in order to get the vital information from the video. In these conventional method the irrelevance between two adjacent key frames was considered as the principle of the key frame extraction. This section present the some noteworthy and common method of key frame exatrtion from the video sequence.

A. Video shot based Method

Video shot based method of key frame extraction utilized the frame average method and histogram average method for key frame exatrtion. frame average method compute the average value of the pixel in a specific location of the video shots. All the frames whose average value of the pixel in that specific location are very close to the average value computed earlier are taken as the key frame. In histogram average method first of all the average histogram of all the frames are computed. All the frames whose histogram is close to the average histogram is selected as the key frame.

Both the above mentioned method has the benefit of having low computational complexity and extracted

frame has average representative meaning. However this method has the drawback of not including the content complexity and hence fixed number of extracted video frames. Moreover the extracted frames does not represent the many changes in the video shots.

B. Content Analysis based Method

In this method, key frames are extracted on the basis of the color, texture and other valuable visual information of each frame. All the frames of the video in which these information are changing significantly are considered as the key frames.

In this approach, first frame is selected as the new frame and hence reference frame, next subsequent frames are then compared with the reference frame in order. Kth frames become the new frame if and only if the distance between kth frame and the reference frame exceeds some predefined threshold. So from this description it is clear that this method select the key frame on the basis of degree of change in the content of the frame. This method is insensitive to the camera movement and hence cant use the movement information to select the key frames. This is why this method produce unstable key frames and poor efficiency of the key frames.

C. Motion-based Analysis Method

Movement in the video shots can easily be detcted or analyzed by analyzing the optical flow of the video sequence. In this method, local minimum in the movement is considered as the key frame. One of the drawback of this method is its low robustness as this method depends on the local information and does not count the global information for the key frame extraction.

III. CLUSTER-BASED METHOD

Cluster based approach of extracting the key frames from the video shots in one of the advanced technique of extracting the key frames from the video sequence. In this method cluster analysis is implemented for key frame exatrtion. In this approach , first of all the cluster center in initialized and then the current frame is classified as the class of this kind or as a new class by computing the distance between the current frame and the cluster center. At the final stage all the frames which are close to the cluster center is considered as the key frame. One of the drawback of this approach is that it is very difficult for any cluster algorithm to get general cluster parameters.

IV. REVIEW OF NOTEWORTHY KEY FRAME EXTRACTION METHOD

This section presents a brief summary of the works carried out in the field of the key frame extraction. Different methods of key frame extraction are reviewed here.

Ajay Diwakaran et al. [1] proposed a method of key frame extraction which is based on the facts that more motion in the video requires more key frames for summarization.

In their method, they divided the whole video sequence into equal motion activity and then by locating the half way point of each segment. Frames that lie in the half way are the key frames as per their method. They also established an empirical relationship between the motion activity in the segment and the number of key frames for that segment.

Mr. Thomas Sikora et al. [2] presented a video summarization approach for video available in the video sharing web site like "U-TUBE". Main aim of their work is to link the existing image search engine to video sequence or video data. User generated video i.e. video available at the u tube are basically unstructured and do not follow the rules. Camera work in these videos is very poor, coding quality and the resolution is also poor due to the higher compression. First step in their method is to detect the gradual and abrupt cut and hence segmenting the video shots.

Moreover, longer video shots are then again segmented into shorter video shots by using a location and motion features of the video sequence. One key frame which is representative of these shorter video shots are then extracted by using visual attention features. Lighting, camera motion, face, appearance of text etc are some of the common visual attention features. These key frames are very useful for indexing and searching similar video in database of video.

Xianglin Zeng et al. [3] proposed a dominant set clustering based approach of key frame extraction. As it is well known fact that key frames play a very important role in video abstraction. Clustering based approach is one of the popular schemes for extracting the key frame from the video sequence. As compared to the then existing clustering based approach, extract the key frames by dynamically deciding the key frame in a given segmented video shots by considering the complexity of the video. This method works in a progressive way and requires less computation. Experimental results of this method confirm its effectiveness of extracting the key frame from the different types of video shots.

Ashvini A Tonge et al. [4] presented a method to summarize the video content by key frame extraction. Video content management system has become very popular now a days for it being widely used in multimedia retrieval. With the ease of availability of the internet and increased bandwidth, video, movies, TV shows are easily available in the internet which makes it easy for the user to facilitate it. In this era where people have very less time they prefer highlight of the information. Video summarization is one of the tools which fulfill this requirement by providing only useful content to the user. In this paper the authors presented a key frame extraction based video content summarization. Orthogonal transform such as cosine, haar, Walsh, slant etc are used for extraction of the key frames. Coefficients of the transform having fractional energy 25%, 6.25% and 1.5625% are used for efficient extraction of the key frame. Experimental results show that the proposed method is capable of producing completeness up to 69%.

Thomas Breuel et al. [5] proposed key frame extraction method which is based on the shot boundary detection for segmenting the video. K-means clustering is used then to find out the key frames within the video segment. In their method they also perform the additional clustering on the extracted key frames to provide better video summarization.

Maria et al. [6] proposed a content based video retrieval by key frame extraction. In this paper they proposed a sequential search algorithm for key frame extraction. This method bypasses the process of video segmentation based on temporal characteristics. The main aim of their research is to find out an efficient, real time and fully automatic way of extracting key frames from the video. This helps to reduce the laborious task of offline video database indexing while performing the query video processing. They found out that significant reduction can be achieved by exploiting the DCT coefficient in feature extraction. The efficiency of the proposed system is tested in terms of quality and speed on the TREC Vid 2007 video dataset.

Li Zhao et al. [7] presented an effective way to search the video on the large video dataset by key frames. According to them, key frame of the video is an effective and convenient way to search the video in video database. This paper supports such type of search. Main contribution of this approach is to use both feature extraction and distance computation.

Key frames of the video shots represent the feature points in the feature space, they proposed a new metric for measuring the distance between query

video and a shot with the help of nearest feature line(NFL). Experimental results shows that the proposed method outperform the then existing method.

Dianting Liu et.al. [8] proposed a new key frame extraction method which is based on the clustering in the feature extraction phase but also effectively reduces the redundant frames by integrating both local and global information in the video.

Proposed method was tested on the TREC Vid 2007 video dataset and the results obtained clearly reveals the efficiency of the proposed key frame extraction method in term of precision retrieval and compression rate. The main aim of key frame extraction is to extract the frames of the video which effectively represent or summarize the video content and can be used in many video related application. The effectiveness of the key frames extraction method can be evaluated on the basis of the fact that how effectively it extract the frames of the video which contain major object or major event in the video with little redundancy or overlapped video.

Sang Hyun Kim et.al. [9] presented an effective approach for matching the video sequence. In their approach, they used

modified Hausdorff distance and the directed divergence of histograms between the successive frames of the video.in order to get the low computational cost they used cumulative directed divergence for extracting the key frames which later on compared by using modified Hausdorff distance. Effective video indexing and retrieval is required for manipulating large video dataset. Experimental results carried out on the color video reveals that the proposed algorithm for video sequence matching gives better performance than the histogram difference method, histogram intersection method and chi square test method.

Tong-yelee et.al.[10] proposed a animating mesh representation based on the key frame extraction. Three dimensional animating meshes has wide application in the computer graphics and video game industries. For overcoming the network bandwidth, reduction in animating complexity is essential.in the proposed method, deformation analysis of animating mesh is used for preserving the geometrics features and motion characteristics. This approach is able to produce very compact representation of the animation in spatial as well as in the temporal domain. Due to above reason it is beneficial in many application such as animation segmentation and transfer.

Xiaomusong,et.al. [11] proposed a coherent framework for combined approach of key frame extraction and

object based video segmentation. Due to being a different sementic levels, key frame extraction and object segmentation are generally implemented separately and independently in a conventional method. This conventional method hence overlooked the inherent relationship between key frames and the objects. In the proposed approach, small number of key frames are extracted within a video shot for maximizing the divergence between the video object in a feature space. Which result in robust and efficient object segmentation. This method hence utilize the temporal based as well as the object6 based video segmentation which is very helpful for content based video analysis and structured video representation. Theoretical analysis and the experimental results on the standard test video reveals the effectiveness of the proposed method.

Guozhu Liu et.al. [12] presented a new method for extracting key frames.in his approach ,they first used the improved histogram matching algorithm to segment the video. Next they extracted the key frames from the video sequence by utilizing the features of the I-frames, P-frames and B-frames. They used fidelity and compression ratio for chacking the validity of the method. Experimental results shows that extracted key frames are able to summarize the salient content of the video sequence. High efficiency, high fidelity and the good feasibility along with high robustness are some of the advantage of this method.efficient segmentation of the video sequence and the key frame extraction significantly reduce the video processing overhead and increase the throughput.

V.CONCLUSION

This paper is an attempt to present different method of key frame extraction by explaining its advantages and disadvantages in order to get the best and reliable approach to key frame extraction. Each methods has its limitations and advantages. Which method is best depends on the application on which it is intended to use. Cluster based approach is advanced technique which gives better results for key frame extraction.

References:

- [1]Ajaydivakaran,Regunathan, Radhakrishnan, Kadir Parker, "Motion activity based key frame exatrction from video shots", pg-116, IEEE2002.
- [2]Kelm, Pascal, Sebastian Schmiedeke, and Thomas Sikora. "Feature-based video key frame extraction for low quality video sequences." *Image Analysis for Multimedia Interactive Services, 2009. WIAMIS'09. 10th Workshop on.* IEEE, 2009.

- [3] Zeng, Xianglin, et al. "Key-frame extraction using dominant-set clustering." *Multimedia and Expo, 2008 IEEE International Conference on*. IEEE, 2008.
- [4] Tonge, Ashvini A., and Sudeep D. Thepade. "Key frame extraction for video content summarization using orthogonal transforms and fractional energy coefficients." *Information Processing (ICIP), 2015 International Conference on*. IEEE, 2015.
- [5] Borth, Damian, et al. "Navidgator-similarity based browsing for image and video databases." *Annual Conference on Artificial Intelligence*. Springer Berlin Heidelberg, 2008.
- [6] Chatzigorgaki, Maria, and Athanassios N. Skodras. "Real-time keyframe extraction towards video content identification." *Digital Signal Processing, 2009 16th International Conference on*. IEEE, 2009.
- [7] Zhao, Li, et al. "Key-frame extraction and shot retrieval using nearest feature line (NFL)." *Proceedings of the 2000 ACM workshops on Multimedia*. ACM, 2000.
- [8] Liu, Dianting, et al. "Integration of global and local information in videos for key frame extraction." *Information Reuse and Integration (IRI), 2010 IEEE International Conference on*. IEEE, 2010.
- [9] Kim, Sang Hyun, and Rae-Hong Park. "An efficient algorithm for video sequence matching using the modified Hausdorff distance and the directed divergence." *IEEE Transactions on Circuits and Systems for Video Technology* 12.7 (2002): 592-596.
- [10] Lee, Tong-Yee, et al. "Animation key-frame extraction and simplification using deformation analysis." *IEEE Transactions on Circuits and Systems for Video Technology* 18.4 (2008): 478-486. [11] Song, Xiaomu, and Guoliang Fan. "Joint key-frame extraction and object-based video segmentation." *Application of Computer Vision, 2005. WACV/MOTIONS'05 Volume 1. Seventh IEEE Workshops on*. Vol. 2. IEEE, 2005.
- [12] Liu, Guozhu, and Junming Zhao. "Key frame extraction from MPEG video stream." *Information Processing (ISIP), 2010 Third International Symposium on*. IEEE, 2010.