Key Frame Extraction Techniques

Mie Mie Khin, Zin Mar Win,

Phyo Phyo Wai, Khaing Thazin Min

mie_mie_khin@miit.edu.mm, zin_mar_win@miit.edu.mm,
phyo_phyo_wai@miit.edu.mm, khaing_thazin_min@edu.mm

Myanmar Institute of Information Technology

Abstract. Video is the most challenging filed because it combines all the form of multimedia information. Key Frame extraction plays an important role in many video processing applications. Key Frame extraction reduces the useless information of video. Key frames are essential to analysis on large amount of video frame sequences. This paper describes different key frame extraction techniques and helps to choose the best key frame method for key frame extraction.

Keywords: Key Frame, Key Frame Extraction Techniques.

1 Introduction

Nowadays, video data is increasing because of the development of digital video capture and technology. Video is a collection of key frames. To discover the content of the video, the main point is to eliminate the redundant information from this video. Key frame extraction is a powerful tool that implements video content by selecting a set of summary key frames to represent video sequences. A key component of the video is the Image; the standard term for the image is the frame.

Key frame extraction is useful to summarize video data. Video summarization plays a major role where the resources like storage, communication bandwidth and power are limited. It has several applications in security, military, data hiding and even in entertainment domains [7].

Key-frames provide a suitable framework for video browsing and retrieval. The use of key-frames significantly reduces the amount of information required in video indexing [8]. The static key-frames are those frames that are extracted from the video which hold the important content of the video. Thus they are representative of video. Dynamic key-frames preserve dynamic nature of video in the sense that they are temporal ordered sequence of key-frames extracted [3].

Key frame extraction techniques can be roughly categorized into four types, based on shot boundary, visual information, movement analysis, and cluster method. And then sometimes it could be completed in compressed domain. Nowadays, cluster-

based methods are mostly applied in video content analysis research. In these methods, key frame extraction is usually modeled as a typical clustering process that divides one video shot into several clusters and then one or several frames are extracted based on low or high level features. The methods in compressed domain usually are not suitable for diverse formats of videos from the Internet. Transcoding may increase time complexity and inaccuracy.

The organization of the paper is as follows. Key frame extraction methods are presented in Section 2 and the conclusions are drawn in Section 3.

2 Key Frame Extraction Techniques

N. K. Narayanan extracted key frame by using the absolute value of histogram of consecutive frames. The experiment is tested on the KTH action database which contains six types of human actions (walking, jogging, running, boxing, hand waving and hand clapping. The performance of their technique is evaluated based on fidelity measure and compression ratio [2].

Sanjoy Ghatak used the threshold technique for extracting key frame. This technique consists of three modules: Module 1 inputs the video of avi format and outputs a set of frames from the input video file. After that, module 2 finds the frame difference. To find the difference, firstly the images are converted to gray scale and gray scale images are plotted to histogram. Finally, the difference matrix is calculated. In module 3, threshold value is calculated. The input frame is considered as key frame or not based on the threshold value [1].

Markos Mentzelopoulos et al. proposed key frame extraction algorithm by using Entropy Difference [4]. Entropy gives the information that can be coded for by a compression algorithm. In this algorithm author used the entropy as local operator instead of global feature for the total image. Entropy is one of the good way of representing the impurity or unpredictability of a set of data because it is dependent on the context in which the measurement is taken. Here author said that first distribute the entropy among the image, then the higher distribution of entropy will give us the regions containing the salient objects of a video sequence. Therefore any change in the object appearance of salient regions will affect its relevant semantic information in the entire story sequence.

Zeeshan Rasheed and Mubarak Shah [5] have proposed the color histogram based method of UCF. This algorithm uses the color histogram for measuring the intersection similarity to extract key frames. This algorithm proposes a method to select the first key frame and then based on first frame and next frame difference, key frame will be decided. There will be comparison with the next frames and new frame will be selected based on difference between frames if there is significant difference, the current frame will be the key frames.

Yueting Zhuang *et al* proposed a GGD model for key frame extraction. GGD model, $\alpha \; and \; \beta$, estimates the density function of coefficients. GGD estimates give a good approximation to the 2D wavelet subband histograms. In content analysis process, appropriate features extraction like motion information, color histograms or features has to be employed from a 2D-transform of the video frame. However, none of the existing methods are structurally matched with the Human Visual System (HVS), which is the main sources of the error in locating the key frames. To clear this error, features are extracted from wavelet transformed sub-bands of each frame, leading to a better match with the HVS in GGD. GGD parameters are helpful to construct the feature vectors and examine similarity between two frames in a cluster and discriminate between two frames from different clusters. These results are more accurate and more subjective to the human section as compared with the previous methods [6].

Dolly et al. presented a paper that A various number of Scene Change Detection (SCD) methods are proposed in recent years. Today's research topic on video is video summarization or abstraction, video classification, video annotation and content based video retrieval. Applications of scene change detection are video indexing, semantic features extraction, multimedia information systems, Video on demand (VOD), digital TV, online processing (networking) neural network mobile application services and technologies, cryptography, and in watermarking. Therefore, in this paper we provide block processing scene change detection techniques & algorithms [9].

Z. Dong, G. Zhang, J. Jia and H. Bao et. al. present a novel key frame selection and recognition method for robust markerless real-time camera tracking. This paper describes an offline module to select features from a group of reference images and an online module to match them to the input live video in order to quickly estimate the camera pose. Their workflow is like that. For offline space abstraction, they extract SIFT (Scale Invariant Feature Transform) features from the reference images, and recover their 3D positions by SFM (Structure-From-Motion). And, select optimal keyframes and construct a vocabulary tree for online keyframe recognition. For online real-time camera tracking, they extract SIFT features for each input frame of the incoming live video. Then, they quickly select candidate keyframes. They match the feature with candidate keyframes. And, they estimate camera pose with the matched features [10].

This proposed approach is based on correlation method for key frame extraction and parallel processing. This approach uses a correlation technique to summarize video. In this paper, they use a simple technique for comparing frames and key frame extraction. They use correlation technique to find how much two frames are similar or different from each other. When frames are similar, they will assume that these frames are duplicate frames and will remove them. Otherwise, different frames will be stored as key frames for using it in summarizing the video. They use mapping for map key frames to retrieve proper video sequence from original video. In their proposed system, they can extract frames from video, and at the same time they compare those extracted frames simultaneously and choose unique frames from them. They use key frame extraction algorithm for choosing unique frames. Although they can apply key frame extraction algorithms directly on this frame, it will work in linear way

which will increase computational time. Therefore, they use Task function to decrease computational time. Task is mainly used for parallel processing. Task is done parallel the operations of extracting frames, creating, applying. By using parallel processing they can decrease processing speed. Parallel processing is also used to find frame difference correlation [11].

Rudinei Goularte et. al. propose a new key frame extraction method based on SIFT local features. Initially, they group key frame candidate sets (KCS) to select the best frames in the key frames of each shot. The first frame in KCS is defined as shot first frame. The next frame in the KCS is the window rule. Widow size 25 is a good one because most of movies have 25 fps as capturing rate and there is no significant variation on consecutive frames content. Then, SIFT features are extracted from the frames in KCS. Each frame represents 128 dimensions of feature vectors. They experimented on five video segments on movies domain [12]. They evaluated their method in the scene segmentation context, with videos from movies domain, developing a comparative study with three state of the art approaches based on local features. The results show that their method overcomes those approaches. Their method is not dependent on the scene segmentation technique.

3 Conclusion

This paper describes about the different key frame extraction techniques. Each technique has its own advantages and disadvantages. Key frame will provide more compact and meaningful video summary. So with the help of key frames it will become easy to browse video and it will give the appropriate search result. Each technique is explained briefly along with their performance results and feedbacks by comparing with each other. This paper will definitely help the researchers to choose the best key frame extraction technique.

References

1. Sanjoy Ghatak, “Key Frame Extraction using threshold techniques”, International Journal of Engineering Applied Sciences and Technology, 2016 Vol. 1, Issue 8, ISSN No. 2455-2143, Pages 51-56
2. M. Omidyeganeh et, “Video Key frame Analysis Using A Segment Based Statistical Metric in A Visually Sensitive Parametric Space”, IEEE
3. M. Chandra Shekhar Mithlesh, D.Shukala, “A Case Study of Key frame Extraction Techniques”, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol. 5, Issue 3, March 2016
4. Markos Mentzelopoulos, Alexandra Psarrou, “KeyFrame Extraction Algorithm using Entropy Difference,” Proceedings of the 6th ACM SIGMM International Workshop on Multimedia Information Retrieval, MIR 2004, October 15-16, 2004, New York, NY, USA
5. Zeeshan Rasheed and Mubarak Shah, “Detection and Representation of Scenes Videos”, IEEE Transactions on Multimedia, vol.7, no.6, December 2005.
6. Azra Nasreen, Dr. Shobha G, “Key Frame Extraction From Video- A Survey”, International Journal of Computer Science & Communication Networks, Vol 3(3),194-198
7. Ebrahim Asadi, Nasrolla Moghadam Charkari, “Video Summarization Using Fuzzy C-Means Clustering”. 20th Iranian Conference on Electrical Engineering, 2012: 690-694.
8. Sheena C V, N. K. Narayanan, “Key-frame extraction by analysis of histograms of videoframes using statistical methods”, 4thInternational Conference on Eco-friendly Computing and Communication Systems, 2015, p-36-40
9. Dolly Shukla, Chandra Shekhar Mithlesh and Manisha Sharma " Design, Implementation & Analysis of Scene Change Detector using Block Processing Method "The Journal of Applied Sciences Research (TJASR), Volume- 2, No 2, 2015, pp103-110.
10. Zilong Dong, Guofeng Zhang, Jiaya Jia, and Hujun Bao, "Keyframe-Based Real-Time Camera Tracking", Computer Vision, 2009 IEEE 12th International Conference, 29 Sept.-2 Oct. 2009.
11. Amitkumar M. Satpute and K. R. Khandarkar, "Video Summarization by Removing Duplicate Frames from Surveillance Video Using Keyframe Extraction", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 5, Issue 4, April 2017
12. Tamires Tessarolli de Souza Barbieri and Rudinei Goularte, "KS-SIFT: a keyframe extraction method based on local features", 2014 IEEE International Symposium on Multimedia, 2014