Perspectives in Communication, Embedded-Systems and Signal-Processing (PiCES) – An International Journal
ISSN: 2566-932X, Vol. 4, Issue 4, July 2020

A Case Study of Key Frame Extraction in Video Processing

K Ragavan
Department of Electronics and Communication Engineering, Ramco Institute of Technology, Rajapalayam, India, ragavme@gmail.com

K Venkatalakshmi
Department of Electronics and Communication Engineering, University College of Engineering, Tindivanam, India, venkata_krish@yahoo.com

K Vijayalakshmi
Department of Computer Science and Engineering, Ramco Institute of Technology, Rajapalayam, India, vijayalakshmik@ritrjpm.ac.in

Abstract: ***Video is an integral part of our everyday lives and in too many fields such as content-based video browsing, compression, video analysing, etc., Video has a complex structure that includes scene, shot, and frame. One of the fundamental techniques in content-based video browsing is key frame extraction. In general, to minimize redundancy the key frame should be representative of the video content. A video can be more than one keyframes. The utilization of key frame extraction method speeds up the framework by choosing fundamental frames and thereby removing additional computation on redundant frames. Key frame extraction significantly reduces the video processing overhead time and increase the throughput. In this paper different types of key frame extraction methods are compared with their advantages and disadvantages.***

Keywords: Key frame extraction; Content-based video; Video processing

I. INTRODUCTION

Video is the most powerful multimedia with which to capture the world around us. But the application of videos such as storing, retrieving, indexing and summarizing became very difficult nowadays. To resolve these problems key frame extraction technique is mainly used. Key frames are those frames that have the video's rich content. Video may describe as a visual representation of the data. Videos can be arranged as scene sequence, scenes as shots series, and shots as frame sequence.. Key frames are the frames which can represent the salient content and information of the shot. The key frames extracted will summarize a video's characteristics, and all of the key frames in the time series will control the image characteristics of a video. A simple rule for extracting key frames is that key frame extraction will rather be incorrect than not appropriate. So it is necessary to eliminate the repetitive information during key frame extraction.

Video:	A video is a series of scenes
Scene:	A scene is a logical collection of shots that take place through a semantic unit
Shot:	A shot is a series of frames in one continuous movement taken by a single camera
Frames:	A frame is created by pixels

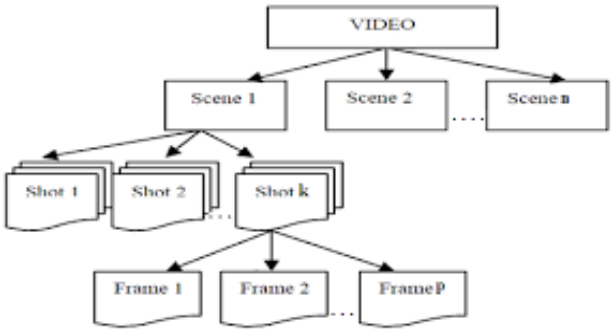

Fig 1. Video to frame convert

II. KEY FRAME EXTRACTION

Key Frame Extraction plays an important role in many applications for video processing, such as video compression, recovery, skimming, editing etc. each described as an unbroken sequence of frames taken from a single camera that forms the video building block. Key Frame Extraction typically involves selecting one Frame to properly reflect in its cluster should obey two key principles.

First it should be fairly close to the frames in its cluster and second it should vary tolerably from frames in other clusters. First, a series of features are extracted from each frame in a standard key Frame Extraction process to form a feature vector. Next, to analyze the similarity / dissimilarity between frames, an effective measure of distance is applied to the feature vectors.

Eventually, shot and cluster boundaries are identified, and the main frames are extracted based on the distance measurement in the selected function space. These frames are very useful for object identification and for keeping track of the particular object.

III. CLASSIFICATION

Key frame extraction methods Broadly classified in two categories.

a) Object Based Video Segmentation

b) Shot Based Video Segmentation.

A. *Object Based Video Segmentation*

Video segmentation based on the target is to decompose one video shot into objects and context which are typically dependent on application. The segmentation based on artifacts will provide objects representing a raw video at a higher semantine level. This approach classifies a video into many objects, each of which can be seen as a pattern described in a video by temporal and/or spatial characteristics.

It can be done either using spatial priority or temporal priority or using both spatial and temporal priority. Joint spatial and temporal segmentation of video is used in recent years to achieve the construction of a multidimensional feature space for object segmentation.

B. *Shot Based Video Segmentation*

Video segmentation based on shot can be considered as a data abstraction process, in which two major steps, i.e., Temporary segmentation and key frame extraction are usually involved. Temporary segmentation classifies one video series into one or more video frames. Most frame-wise functions, such as color layout, entropy etc., Key-frame extraction implements data abstraction by choosing a set of key-frames. It is typically modelled as a standard processing of clusters that divides one video shot into multiple clusters and selects centroid clusters as main frames. Key-frames are extracted from each shot using K-means process. In the Gaussian mixture model, the temporal color histogram variance in the RGB color space is modelled. Frames present in the shot can be grouped into multiple clusters based on representation of the function. The frames closest to the centroid are chosen as a key-frame for each cluster. Bayesian Knowledge Criterion can be used to determine the number of clusters. The main downside to this approach is that it cannot automatically calculate the number to clusters and thus, it does not automatically adjust the clustering to the content of the images.

IV. LITERATURE SURVEY

Liu. H & Hao. H proposed an algorithm of using hierarchical clustering algorithm based on calculation of information entropy. They have used it for news, ad, cartoon and movie videos and produced good precision rate and recall rate. But they experienced high computational complexity.[1]

Liu, G., & Zhao, J. proposed an algorithm of using similarities in histogram of frames for aircrash video and obtained low error. But this method is only suitable for small size of videos. [2]

Dr. Aziz Makandar, Daneshwari Mulimani used histogram values of gray images to extract key frames. They used this for complex videos like kabbadi videos but they can't use this method for online videos.[3]

AssmaAzeroual, KarimAfdel, Mohamed EL Hajji, Hassan Douzi used Faber-Shauder Discrete Wavelet Transform for key frame extraction process and achieved high robustness for sports videos but failed to fix the threshold value automatically.[4]

Miss.A.V. kumthekar, Prof.Mrs.J.K.Patil combined the color histogram and edge detection to extract the key frames. They got 98% of compression ratio but they didn't clearly mentioned about threshold conditions.[5]

Ganesh. I. Rathod, Dipali. A. Nikam propsed an algorithm of using square histogram model for extracting key frames and reached efficiency of 95% to 98% but with some memory constrains.[6]

Yujie Li, Benying Tan, Shuxue Ding, Incheon Paik and Atsunori Kanemura computed matrix determinant value to find sparse measurement which helps to extract the key frames. This method produces low compressive ratio but complex to compute algorithm.[7]

Sanjoy Ghatak calculated histogram difference to get key frames which requires less time to execute. But he got redundant ley frames.[8]

Yunyu Shi,Haisheng Yang,Ming Gong,Xiang Liu,and Yongxiang Xia used both histogram difference and structure feature extraction for extraction of key frames which have low computational complexity but can be used only for specific application.[9]

Kevin Thomas Abraham, Ashwin M, Darshak Sundar, Tharic Ashoor and Jeyakumar, G used differential evolution method to achieve high accuracy for key frame extraction process but it can't be used for online videos.[10]

Vignesh S, V. Vaidehi, M. Kannan, proposed an algorithm based on MBType Difference metric information and Unimodality of the GOP structure of the frames post L2-Norm Optimization which produces precision rate and recall rate above 90% but can't be applied for real time videos.[11]

Zhikai Zong, Qing Gong proposed an algorithm based on dynamic color space algorithm and fast wavelet histogram which has no redundancy.[12]

Sl.No.	Title	Author, Publication, Pages, Volume No, Issue No Year	Methodology Used	Advantages	Limitations
1	Key Frame Extraction Based on Improved Hierarchical Clustering Algorithm	Liu, H., & Hao, H. (2014). "Key frame extraction based on improved hierarchical clustering algorithm," 2014 IEEE 11th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD) .	K-means hierarchical Clustering algorithm	Precision Ratio and recall Ratio is above 85%	Calculation of information entropy is expensive
2	Key Frame Extraction From MPEG Video Stream	Liu, G., & Zhao, J. (2010), "Key Frame Extraction from MPEG Video Stream", 2010 IEEE Third International Symposium on Information Processing.	Histogram matching method is used	Good feasibility, High efficiency, Low error, High robustness	Longer video can not be used
3	Key frame extraction and Object Detection in the Sports Video	Dr. Aziz Makandar, Daneshwari Mulimani, " Key Frame Extraction from Video Based on Determinant-Type of Sparse Measure and DC Programming," 2017 IEEE 11th International Symposium on Embedded Multicore/Many-core Systems-on-Chip.	Gray histogram method	More suitable for Kabaddi videos	Effective only for self-developed videos
4	Key Frame Extraction Using Faber-Shauder Wavelet	Azeroual, A., Afdel, K., El Hajji, M., & Douzi, H. (2014). "Key frame extraction using Faber-Shauder wavelet," 2014 IEEE Second World Conference on Complex Systems.	Based on Faber-Shauder Discrete Wavelet Transform and Singular Value Decomposition	High robustness for sports videos	Threshold value cannot be fixed automatically
5	Key frame extraction using color histogram method	Miss.A.V.Kumthekar, Prof.Mrs.J.K.Patil , "Key frame extraction using color histogram method," International Journal of Scientific Research Engineering & Technology (IJSRET) , Volume 2, Issue 4, pp 207-214, July 2013.	Color histogram and edge detection method	Less computational complexity The compression ratio is 98%	Absence of threshold conditions
6	An Algorithm for Shot Boundary Detection and Key Frame Extraction Using Histogram Difference	Ganesh. I. Rathod, Dipali. A. Nikam "An Algorithm for Shot Boundary Detection and Key Frame Extraction Using Histogram Difference," International Journal of Emerging Technology and Advanced Engineering, Volume 3, Issue 8, August 2013.	Based on square histogram model	Efficiency is about 95% to 98%	Suitable only for ".avi" type with memory range of 3MB to 4.5MB
7	Key Frame Extraction from Video Based on Determinant-Type of Sparse Measure and DC Programming	Li, Y., Tan, B., Ding, S., Paik, I., & Kanemura, A. (2017), "Key Frame Extraction from Video Based on Determinant-Type of Sparse Measure and DC Programming.," 2017 IEEE 11th International Symposium on Embedded Multicore/Many-Core Systems-on-Chip (MCSoC).	Based on sparsity measurement using matrix determinant	Compressive ratio is less	Computational complexity is high
8	Key-frame extraction using threshold technique	Sanjoy Ghatak, "Key-frame extraction using threshold technique," International Journal of Engineering Applied Sciences and Technology, 2016 Vol. 1, Issue 8, ISSN No. 2455-2143, Pages 51-56.	Histogram difference based method	Execution time is less	Redundancy of key frames
9	A Fast and Robust Key Frame Extraction Method for Video Copy rightProtection	Yunyu_Shi,Haisheng_Yang,Ming_Gong,Xiang_Liu,and Yongxiang_Xia , "A Fast and Robust Key Frame Extraction Method for Video Copyright Protection," Journal of Electrical and Computer Engineering Volume 2017, Article ID 1231794, 7 pages.	Based on histogram difference and structure feature extraction	Lower computational complexity	Specific for video copyright protection
10	An Evolutionary Computing Approach for Solving Key Frame Extraction Problem in Video Analytics	Kevin Thomas Abraham, Ashwin M, Darshak Sundar, Tharic Ashoor and Jeyakumar, G	Differential Evolution method is used	Accuracy is high	Not suitable for online videos
11	Unimodality of MBType frame different metrics for efficient Key Frame Extraction in Video	Vignesh, S., Vaidehi, V., & Kannan, M. (2017), " Unimodality of MBType frame different metrics for efficient Key Frame Extraction in Video," 2017 Fourth International Conference on Signal Processing, Communication and Networking (ICSCN).	Based on the MBType Difference metric information and Unimodality of the GOP structure of the frames post L2-Norm Optimization	Precision rate and recall rate is above 90%	Not suitable for online real time videos
12	Key Frame Extraction Based on Dynamic Color Histogram and Fast Wavelet Histogram	Zong, Z., & Gong, Q. (2017), " Key frame extraction based on dynamic color histogram and fast wavelet histogram," 2017 IEEE International Conference on Information and Automation (ICIA).	Based on dynamic color space algorithm and fast wavelet histogram algorithm	Absence of redundancy	———
13	Key frame extraction based on entropy	Zhang, M., Tian, L., & Li, C. (2017), " Key Frame Extraction Based on Entropy Difference	Based on entropy difference and perceptual hash.	Low miss rate	Redundancy for background

	difference and perceptual hash	and Perceptual Hash," 2017 IEEE International Symposium on Multimedia (ISM).			moving videos
14	Key Frame Extraction for Video Content Summarization Using Orthogonal Transforms and Fractional Energy Coefficients	Tonge, A. A., & Thepade, S. D. (2015), "Key frame extraction for video content summarization using orthogonal transforms and fractional energy coefficients," 2015 International Conference on Information Processing (ICIP).	Based on DCT, Walsh transform, Haar transform, Hartley transform, Slant transform and Kekre transform	Less computational complexity	Completeness is less
15	3D Human Motion Key-frames Extraction Based on Asynchronous Learning Factor PSO	Zhang, Y., & Cao, J. (2015)., "3D Human Motion Key-Frames Extraction Based on Asynchronous Learning Factor PSO," 2015 Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC).	Based on Asynchronous Learning Factor PSO	Low error rate	———

Table 1. Comparison of various key frame extraction techniques

V. CONCLUSION

Video abstraction is an integral part of many video applications including video indexing, browsing, and retrieval. Everybody around the world has "lack of time". Key frame extraction is one of the important process to reduce the time consumption. Since analyzing key frames takes less time than analyzing a whole video. In the above survey Block based method produces good results for shot segmentation and square histogram method produces good efficiency of 95% to 98% . Even though all these methods have a merit but they have some drawbacks in some other way like memory constraint, high computational complexity. Hence we looking for a more better method which should have following constrains threshold free, application range is wide, low error, high robustness, good feasibility, less wipe effect, efficiency > 98%. and compression ratio >98% which makes the key frame extraction process more suitable for all applications.

ACKNOWLEDGMENTS

I would like to thank Ramco Institute of technology, Rajapalayam for providing the excellent facilities and constant support extended to carry out the research work.

REFERENCES

[1] Liu, H., & Hao, H. "Key frame extraction based on improved hierarchical clustering algorithm," 2014 IEEE 11th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), 2014 .

[2] Liu, G., & Zhao, J., "Key Frame Extraction from MPEG Video Stream," 2010 IEEE Third International Symposium on Information Processing, 2010.

[3] Dr. Aziz Makandar, Daneshwari Mulimani, " Key Frame Extraction from Video Based on Determinant-Type of Sparse Measure and DC Programming," 2017 IEEE 11th International Symposium on Embedded Multicore/Many-core Systems-on-Chip,2017.

[4] Azeroual, A., Afdel, K., El Hajji, M., & Douzi, H.. "Key frame extraction using Faber-Shauder wavelet," 2014 IEEE Second World Conference on Complex Systems, 2014.

[5] Miss.A.V.Kumthekar, Prof.Mrs.J.K.Patil , "Key frame extraction using color histogram method," International Journal of Scientific Research Engineering & Technology (IJSRET) , Volume 2, Issue 4, pp 207-214, July 2013.

[6] Ganesh. I. Rathod, Dipali. A. Nikam "An Algorithm for Shot Boundary Detection and Key Frame Extraction Using Histogram Difference," International Journal of Emerging Technology and Advanced Engineering, Volume 3, Issue 8, August 2013.

[7] Li, Y., Tan, B., Ding, S., Paik, I., & Kanemura, A., "Key Frame Extraction from Video Based on Determinant-Type of Sparse Measure and DC Programming.," 2017 IEEE 11th International Symposium on Embedded Multicore/Many-Core Systems-on-Chip (MCSoC), 2017.

[8] Sanjoy Ghatak, "Key-frame extraction using threshold technique," International Journal of Engineering Applied Sciences and Technology, Vol. 1, Issue 8, ISSN No. 2455-2143, Pages 51-56, 2016.

[9] Yunyu Shi,Haisheng Yang,Ming Gong,Xiang Liu,and Yongxiang Xia , "A Fast and Robust Key Frame Extraction Method for Video Copyright Protection," Journal of Electrical and Computer Engineering, Article ID 1231794, pp-15-22, 2017.

[10] Abraham, K. T., Ashwin, M., Sundar, D., Ashoor, T., & Jeyakumar, G. "An evolutionary computing approach for solving key frame extraction problem in video analytics," 2017 International Conference on Communication and Signal Processing (ICCSP), 2017.

[11] Vignesh, S., Vaidehi, V., & Kannan, M., " Unimodality of MBType frame different metrics for efficient Key Frame Extraction in Video," 2017 Fourth International Conference on Signal Processing, Communication and Networking (ICSCN), 2017.

[12] Zong, Z., & Gong, Q., " Key frame extraction based on dynamic color histogram and fast wavelet histogram," 2017 IEEE International Conference on Information and Automation (ICIA), 2017.

[13] Zhang, M., Tian, L., & Li, C., " Key Frame Extraction Based on Entropy Difference and Perceptual Hash," 2017 IEEE International Symposium on Multimedia (ISM),2017.

[14] Tonge, A. A., & Thepade, S. D. , "Key frame extraction for video content summarization using orthogonal transforms and fractional energy coefficients," 2015 International Conference on Information Processing (ICIP), 2015.

[15] Zhang, Y., & Cao, J., "3D Human Motion Key-Frames Extraction Based on Asynchronous Learning Factor PSO," 2015 Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC), 2015.