Scientific Journal of Impact Factor (SJIF): 4.72

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 4, Issue 3, March -2017

Review on Video Summarization Techniques for Surveillance Video Using Keyframe Extraction

Mr. Amitkumar M. Satpute[1],Mr. K. R. Khandarkar[2]

[1](PG Research Student, Department of Computer Science & Technology, Dr. BAM University, Maharashtra Institute of Technology, Aurangabad (MS), India)

[2](Professor, Department of Computer Science & Technology, Dr. BAM University, Maharashtra Institute of Technology, Aurangabad (MS), India)

Abstract *—*

Nowadays Video data creation and its storage play a important role in every industry. We require it in security, entertainment industries, advertising industries and more other industries. In security field data is collected using CCTV camera and stored at data centers of organization. Same video content is referred to check if someone is violating the organizations security. In advertising industry product details are conveyed using seconds of video. If video length is more, then cost of advertisement creation is also more. Video summarization is also a technique in which your long period activity is represented using 5-10 minutes of video. Compare to verbal communication video representation is best way of communication and its trustworthy way to communicate your message.

By considering the importance of video data there is huge scope of improvement in video data processing. Video processing techniques can be used to remove redundant frames from video so that it will solve the problem of data storage. Video without redundant frames is the best input for information retrieval operation. As less time will require for retrieval of information compared to original video. So main object is to develop a system to compress the video i.e. to remove redundant data from video. Also will provide the information retrieval technique in which compressed video is taken as input and image with require object is taken as query so that we will reach to that relevant video content in less time.

Keywords *—Key frame extraction, Faber-Shauder wavelet, Haar Wavelet, Scale Invariant Feature Transform.*

I. INTRODUCTION

In recent years, sudden technical advancement in video data creation and its storage are improved a lot. So it is more about to how efficiently we handle those data with indexing and retrieval methods. As current activity of data creation and its storage is sequential, so it consumes more storage space. Video Storage space increases only due to the fact that video has 60-70% of data which is sufficient to represent all important video content and rest 30-40% data is redundant frames. As to solve this problem we have solution at the same time, by using data management. As for particular video there will be repeated video content which may not be useful and hence we can skip those data and extract whichever we want or one with less repetition.

The popular term video summarization is the same meaning of what we are suggesting here. [1] As global video summarization will be fast and efficient data assessment system. As in video summarization methods, some popular techniques are being researched nowadays like trajectory curve, clustering , faber-shouder wavelet [3] and thepade's sorted n-ary block truncation coding [4].

In current paper among these mentioned methods, their advantages and weakness in particular applications are studied. In current work, review is made in field of key frame extraction and related research going on. Some works those are related to work are also mentioned in this review work.

II. TECHNIQUES

Already lots of techniques are available in the field of video data compression and efficient retrieval of data. These techniques are implemented using Correlation, Histogram measure, moments of inertia, haar transformation and many more.

2.1. Aggregation Mechanism

In [1] the video summarization using key frame extraction is shown in the Fig.1.Video is taken and frames are extracted from video. These extracted frames are taken as input to the algorithm. On each frame preprocessing is performed i.e. color image is converted into grey scale image so that computational time will be less as we have only one plane need to compare but in color image we require to compare RGB plane.

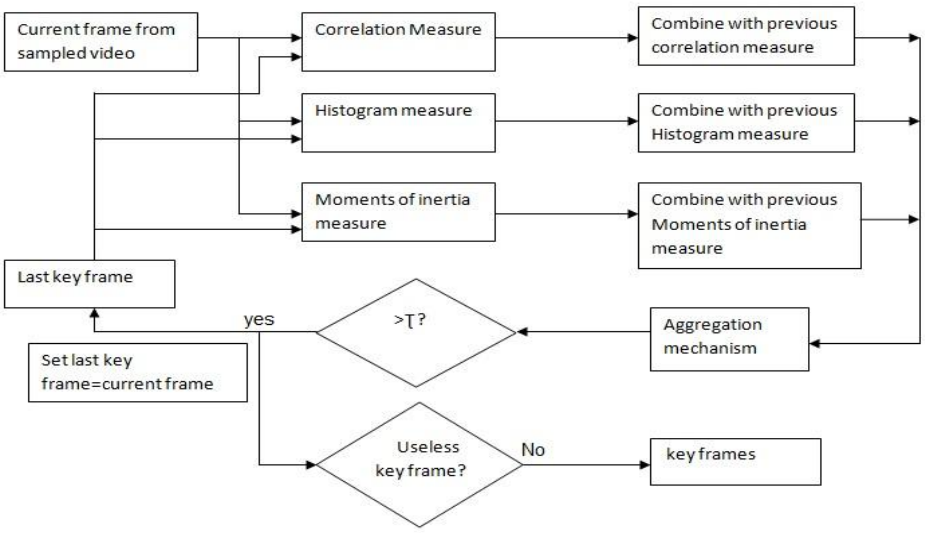

Fig.1. Video Summarization using key frame extraction

Correlation measure, Histogram measure and moments of inertia is performed on each frame and these results are combined with previous frame measures. Aggregation mechanism is used to get the aggregate value of combined result. Based on this aggregate value difference between current frame and key frame is calculated. If current frame is different to key frame more than certain threshold value then current frame is considered as a key frame else current frame is skipped. At the end all key frames are combined to form a summarize video.

2.2. Block truncation with Haar Wavelet

In [2] along with the Haar transformation, mathematical techniques such as mean, standard deviation and threshold are used to calculate the difference between two frames. This mechanism is efficient compared to previous algorithm because it took less time for comparison of two frames while previous technique compare the frames pixel by pixel. The flow chart of proposed key frame extraction method is shown in the Fig.2.

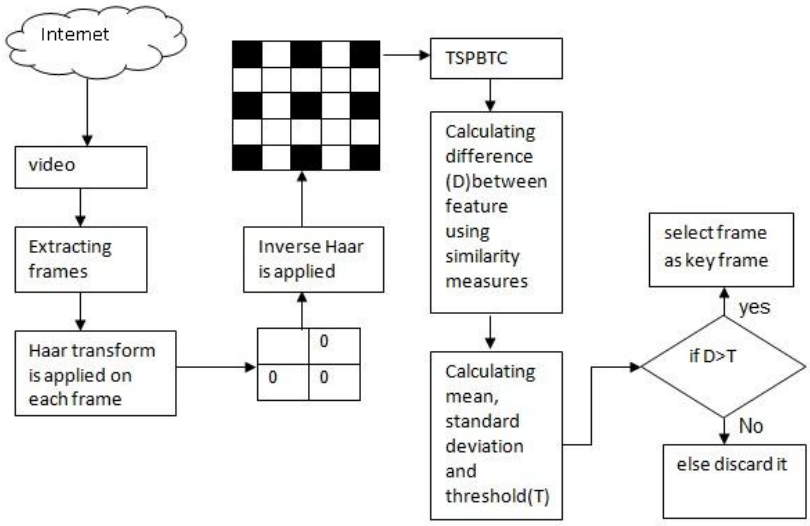

Fig.2. key frame extraction method

In the proposed strategy, video is taken from standard database. Video constitutes of nonlinear edges. Casings are extricated from video. On each casing from video haar change is connected. After this procedure the changed edge is isolated into four sections
The left upper part is held as it is and other three are made invalid. At that point on that Inverse Haar is connected. After utilization of Inverse Haar, just interchange lines and section qualities are chosen to get Haar wavelet level 1.

At that point, on the edge acquired after level 1, again entire of this procedure is rehashed to get Haar wavelet level 2 and same is for Haar wavelet level 3 and Haar wavelet level 4, 5 and 6 is taken after. At that point on each of the casing got from level 1, level 2, level 3 and level 4. Thepade's sorted pentnary piece truncation coding (TSPBTC) is connected.
This will give fifteen values for each frame. This is the feature vector of that frame. On these consecutive feature vectors of frame similarity measures are applied to calculate the difference between two frames. Canberra distance, Sorensen distance, Wave hedge distance, Euclidean distance and mean square error are taken here as a similarity measures. Then difference between each consecutive frame is taken which can be referred as diff.

2.3. Faber-Shauder wavelet

In [3] key frames are extracted and compressed video is formed using artificial intelligence techniques such as precision and recall. Video is taken as input to the algorithm then framing is performed to separate out frames present inside the video. Each frame is converted into grey to minimize the processing time and computational effort. This is block selection technique in which dominant region of frame is considered as a block.

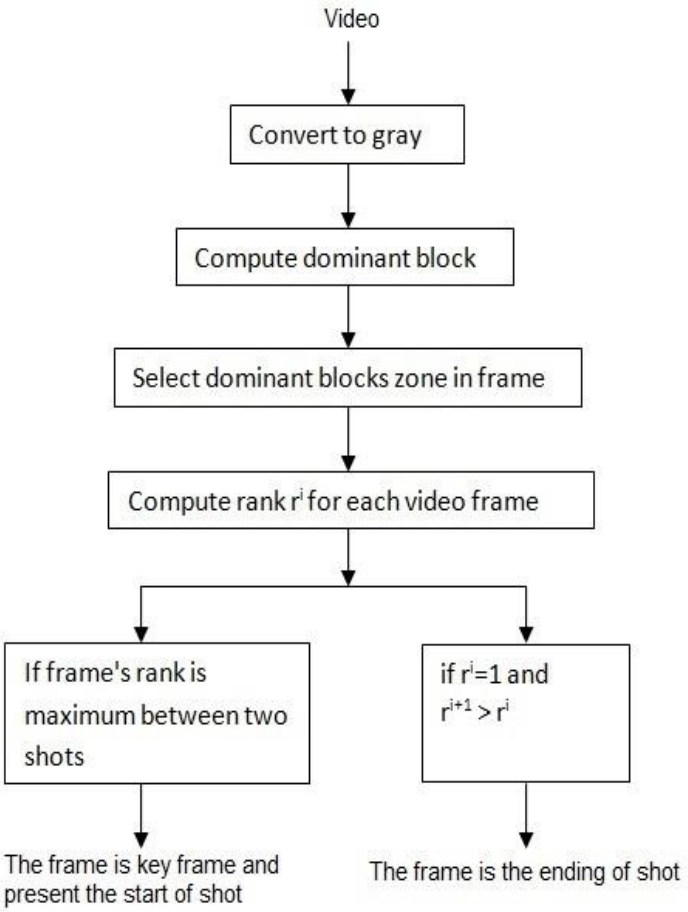

Fig.3. Key frame extraction using block selection.

After that rank of each dominant block is computed. This technique is somewhat similar to Google's page ranking algorithm which is used to extract the accurate results for given input query. After calculating rank's for each block, if frames rank is maximum between two shots then frame is considered as key frame and representing the start of next shot else the frame is ending of the given shot.

2.4. Scale Invariant Feature Transform(SIFT)

In [4] their describes a framework using SIFT for shot boundary detection and keyframe extraction , as shown in Fig.4. In that there SIFT keypoints are extracted from each frame according their sequence of temporal information.
.

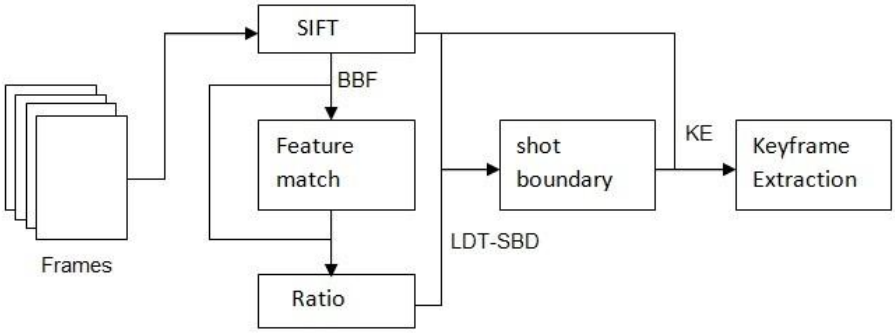

Fig.4. The framework using SIFT for shot boundary detection and keyframe extraction

Each keypoint represented by a descriptor of 128 dimensions. After that Best-Bin-First (BBF) algorithm is used for comparing two adjacent frames and then for detection of extract keyframes and shot boundaries they calculate the ratio of matched keypoints number and total number.

2.5.Clustering

In[5] their first calculate visual similarity between every pair of frames according to selected feature space and it stored in a matrix. All calculation is done based on low-level features such as color, shape and texture. After that dominant-set clustering algorithm used with similarity matrix. If all frames are grouped in cluster then clustering process will stops. Shown in Fig.5.

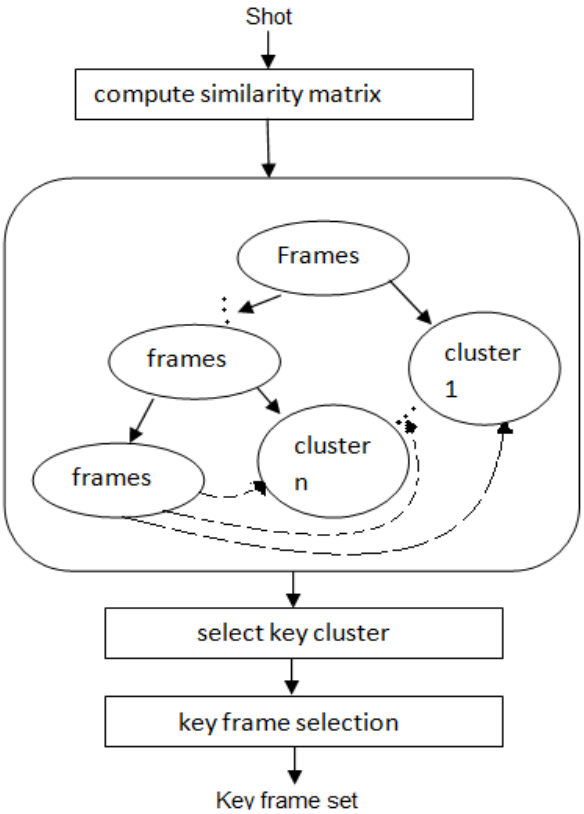

Fig.5. Key frame extraction using clustering

After that discard that cluster who's size is smaller than threshold. And select remaining cluster as a key cluster. Key cluster contains similar sequence of frames, select a frame as a key frame from key cluster which is closest to the centroid of the key cluster.

III. COMPARISON TABLE

Sr.no.	Paper	Description	Research Gap
1.	Adaptive key frame extraction for video summarization using an aggregation mechanism.	In this technique key frames are extracted using correlation, moment of inertia and histogram measure. The computations are pixel based so require more time but gives accurate result.	As computations are pixel based and video frames are huge in size so time complexity is more.
2.	Summarization with key frame extraction using thepade's sorted n-ary block truncation coding applied on Haar Wavelet of Video Frame.	In this technique key frames are extracted using Haar wavelet features. The computational time required for this approach is less compared to previous approach.	Though computational time is less but Haar features are environmental condition dependent. So result may vary based on conditions.
3.	Key Frame extraction using Faber-Shauder Wavelet	In this technique key frames are extracted using artificial intelligence techniques such as precision and recall. Dominant blocks are formed in each frame.	This approach is not good if we consider a frame in which details are distributed on entire image. So in such situation calculation of dominant block is extra overhead.
4.	Shot Boundary detection and keyframe extraction based on scale invariant feature transform.	In this approach key frames are extracted using SHIFT and BBF. BBF is similar to CPU scheduling algorithm. In which priority of two adjacent frames are compared using BBF and based on the result key frame is decided.	In this approach for some of the cases it causes miss or false shot boundary detection.
5.	Key frame extraction using dominant set of clustering	In this technique key frames are extracted using dominant set of clustering for calculate similarity matrix.	In this method it did not consider temporal information. Many times temporal information is important aspect in video analysis..

IV. CONCLUSION

Due to the increasing requirement of video summarization and information retrieval from video , key frame extraction methods are used such as clustering, aggregation mechanism, Scale invariant feature transform, faber-shauder wavelet and block truncation with Haar wavelet. According to survey of the above methods, here conclude that there are some positive points as well as some drawbacks are there. As According to study of those methods, time complexity is more of aggregation mechanism , computation time is less of block truncation with Haar wavelet method but it depends on the environmental condition, in the scale invariant feature transform there is chances of missing shot boundary detection and the clustering method did not consider temporal information but in some cases temporal information is important thing in video analysis.

By consider the positive and negative points of above mentioned techniques which I found during their review, I will try to implement a system which will cover most of the positive points of above mentioned techniques. At the concluding ,we have to focus on correlation method for key frame extraction. Correlation method will be better option for future key frame exaction techniques.

REFERENCES

[1] N. Ejaz, T. Tariq and S. Baik, "Adaptive Key Frame Extraction using an Aggregation Mechanism", Elsevier Inc J. Vis commun Image R. 23(2012), pp. 1031-1040

[2] Dr. Sudeep D. Thepade, Shalakha R. Badre " Summarization with Key Frame Extraction using Thepade's Sorted n-ary Block Truncation Coding Applied on Haar Wavelet of Video Frame " 2016 Conference on Advances in Signal Processing (CASP) Cummins College of Engineering for Women, Pune. Jun 9-11, 2016, 978-1-5090-0849-0/16/$31.00 ©2016 IEEE

[3] Assma Azeroual*, Karim Afde!, Mohamed El Hajji and Hassan Douzi " Key Frame Extraction using Faber-Shauder Wavelet", IRF-SIC Laboratory, Faculty of Science, 978-1-4799-4647-1114/$31.00 2014 IEEE Agadir, Morocco.

[4] Gentao Liu, Xiangming Wen ,Wei Zheng and Peizhou He " Shot Boundary Detection and Keyframe Extraction based on Scale Invariant Feature Transform", School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, China, 2009 Eigth IEEE/ACIS International Conference on Computer and Information Science

[5] Xianglin Zeng, WeimingHu, Wanqing Li†, Xiaoqin Zhang, Bo Xu "KEY-FRAME EXTRACTION USING DOMINANT-SET CLUSTERING",National Laboratory of Pattern Recognition (NLPR) Institute of Automation, Chinese Academy of Sciences SCSSE, University of Wollongong, Australia, 978-1-4244-2571-6/08/$25.00 ©2008 IEEE.

[6] X. Zeng, W. Hu, W. Li, X. Zhang and B. Xu, "Key-Frame Extraction using Dominant-Set Clustering", pp. 1285-1288.

[7] J. Yuan, W. Wang, W. Yang and M. Zhang , "Keyframe Extraction using AdaBoost", 2014 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC), Oct. 2014, pp. 91 - 94 .

[8] A. C. Hernandez, M. C. Hernandez, F. G. Ugalde, M. N. Miyatake and H. P. Meana," A Fast and Effective Method for Static Video Summarization on Compressed Domain", IEEE LATIN AMERICA TRANSACTIONS, VOL. 14, NO. 11, NOV. 2016

[9] Vasileios T. Chasanis, Antonis I. Ioannidis and Aristidis C. Likas" Efficient Key-frame Extraction Based on Unimodality of Frame Sequences ", Department of Computer Science and Engineering University of Ioannina Ioannina, Greece fvchasani, aioannid, arlyg@cs.uoi.gr, 978-1-4799-2186-7/14/$31.00 ©2014 IEEE

[10] Werachard Wattanarachothai, Karn Patanukhom " Key Frame Extraction for Text Based Video Retrieval Using Maximally Stable Extremal Regions", Visual Intelligent and Pattern Understanding Laboratory Department of Computer Engineering Chiang Mai University, Thailand karn@eng. cmu. ac. th, 2015 1st International Conference on Industrial Networks and Intelligent Systems (INISCom).

[11] Ashvini A. Tonge, S.D. Thepade " Key Frame Extraction for Video Content Summarization Using Orthogonal Transforms and Fractional Energy Coefficients", Pimpri Chinchwad College of Engineering Pune, INDIA, 2015 International Conference on Information Processing (ICIP) Vishwakarma Institute of Technology. Dec 16-19, 2015.

[12] Antonis I. Ioannidis, Vasileios T. Chasanis and Aristidis C. Likas " Key-frame Extraction using Weighted Multi-View Convex Mixture Models and Spectral Clustering", Department of Computer Science and Engineering University of Ioannina Ioannina, Greece, 2014 22nd International Conference on Pattern Recognition. 1051-4651/14 $31.00 © 2014 IEEE DOI 10.1109/ICPR.2014.596.

[13] Ijya Chugh, Ridhima Gupta2 , Rishi Kumar, Prashast Sahay " Techniques for Key Frame Extraction: Shot segmentation and feature trajectory computation", Computer Science and Engineering Department Amity University, Uttar Pradesh, Noida, India, 978-1-4673-8203-8/16/$31.00_c 2016 IEEE.

[14] Wisnu Widiarto, Eko Mulyanto Yuniarno, Mochamad Hariadi " Video Summarization Using a Key Frame Selection Based on Shot Segmentation", 2015 International Conference on Science in Information Technology (ICSITech), 978-1-4799-8386-5/15/$31.00 ©2015 IEEE.

