

# A Survey on Multi Media based Indexing and Retrieval

Mr. Rahul S Patel<sup>1</sup> Miss. Wable A Aparna<sup>2</sup> Prof.R.M.Mulajkar<sup>3</sup>

<sup>1,2</sup>M.E. Student <sup>3</sup>Assistant Professor & PG coordinator

<sup>1,2,3</sup>Department of Electronics & Telecommunication Engineering

<sup>1,2,3</sup>Jaihind College of Engineering, Pune

**Abstract**— Video classification and retrieval have an increasing wide spectrum of promising applications, motivating the interest of researchers worldwide. The content-based retrieval technique is helpful in video police work and likewise as getting ready to field (CCTV) imagination. The goal of this paper is to retrieve video section supported the visual content. During this paper associate outline of the visual content based video classification and retrieval methodology is conferred. It includes methods for video shot boundary detection and key frame extraction.

**Key words:** Video Surveillance, Shot Boundary Detection, Key Frame Extraction, with the Increasing of Digital Video Contents

## I. INTRODUCTION

Efficient techniques for analysis, classification and retrieval of videos in step with their contents became even further necessary. Multimedia data Retrieval (MIR) square measure planning to be relating to the search of data for all its kind, multimedia data retrieval and classification [1] are required to elucidate store, and conjointly the organize multimedia data and to assist of us to seek out multimedia resources handily and quickly from large dataset. This paper is supported for researching inside the content-based retrieval of multimedia. This paper aim may be a way to improve the fundamental has been thanks to alter or improve multimedia victimization content based methodology. Videos have the following characteristics: 1) far more content than individual images; 2) Brobdignagian amount of raw data; and 3) very little previous structure. Further recently, these dataset became plenty of larger and content-based classification and retrieval square measure planning to be required, supported the machine-controlled analysis of videos with the less human participation.

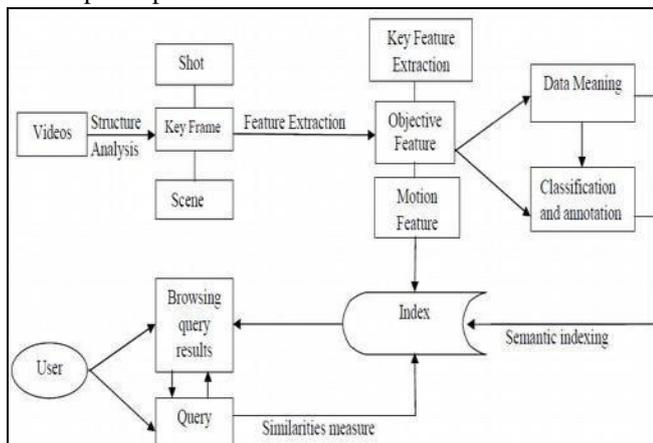


Fig. 1: Generic framework for visual content-based video indexing and retrieval

The video police investigation systems victimization CCTV manufacture pot of data. Looking by video compartmentalization and retrieving the relevant

video from this data-set could be a vital issue. The content - based retrieval technique is helpful throughout this regard. The goal of this paper is to retrieve video section supported the visual content. This can be achieved by a content-based video retrieval and compartmentalization framework.

## II. RELATED WORK

The videos unit structured in step with a down hierarchy of video clips, scenes, shots, and frames. Video structure analysis aims at segmentation of a video into kind of structural elements means linguistics contents, beside shot boundary detection, key frame detection, and scene mesmerizing. The hierarchy of video clips, scenes, shots and frames unit organized throughout a down manner as shown in figure two.

### A. Video Structure Analysis

The videos sq. measure structured in step with a descendent hierarchy of video clips, scenes, shots, and frames. Video structure analysis aims at segmentation of a video into kind of structural elements suggests that linguistics contents, in addition as shot boundary detection, key frame detection, and scene segmental. The hierarchy of video clips, scenes, shots and frames unit organized in Associate in nursing passing descendent manner as shown in figure 2.

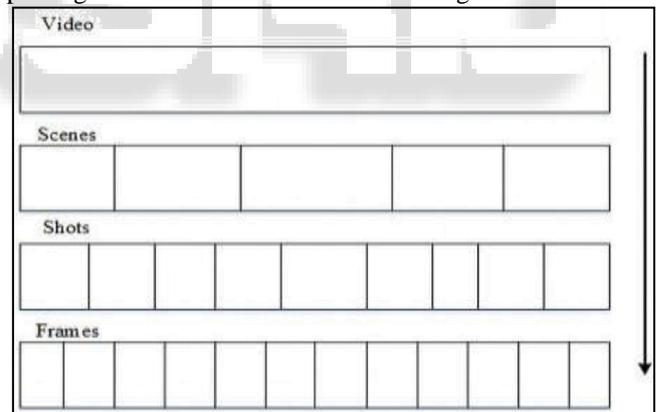


Fig. 2: General Hierarchy of Video Parsing

### 1) Shot Boundary Detection

The shot is made public because the consecutive frames from the start to the tip of recording in an exceedingly} very camera. It shows a nonstop action in an exceedingly image sequence [11]. There are 2 differing kinds of transitions which will occur between shots, abrupt means (discontinuous) to boot referred as cut, or gradual means (continuous) like fades, dissolves and wipes. The cut boundaries show the abrupt modification in image intensity or color, whereas those of fades or dissolves show gradual changes between frames. A video is countermined in scene, shot and frames. a shot could be a sequence of frames is captured by one camera in an exceedingly} very single continuous action. An endeavor boundary is that the

Transition between 2 shots. A scene may be a logical grouping of shots into a linguistics unit [11].

a) Shot Transition Types

There unit of measurement a pair of sorts of shot transition. Abrupt: Abrupt transition happens in an exceedingly} very single frame. Gradual: it's once more classified into 3 varieties. they are dissolve, fade, and wipe. A dissolve may be a gradual transition from one scene to a unique scene at intervals which the first scene fade-out and second scene fade-in. it is a mixture of fade-in and fade-out.

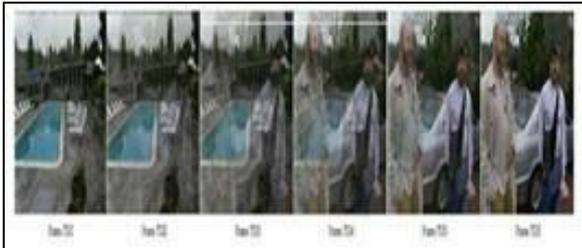


Fig. 3: Dissolve a fade – out is a slow decrease in brightness resulting in a black frame.



Fig. 4: Example of fade-out

B. Shot Boundary Detection Techniques

There are many methods to detect shot boundary we have discussed very few techniques.

1) Pixel based mostly Shot Boundary Detection

It's the one amongst the simplest methodology for crucial shot boundaries. The excellence between corresponding pixels of two frames is computed. If the excellence is bigger than some threshold, then its shot boundary is assumed [12].

2) Threshold-Based Approach

The threshold-based approach detects shot boundaries by comparison the measured pair-wise similarities between frames with a predefined threshold [6]. once a similarity is a smaller amount than the edge, a boundary is detected.

3) Learning-Based Approach

Threshold-based approach detects shot boundaries by comparison the measured pair-wise similarities between frames with a predefined threshold [6]. Once a similarity is a Smaller amount than the edge, a boundary is detected.

C. Key Frame Extraction

The frame that represents the content of an endeavor or scene in key frame. This content of shot and scene should be the foremost representative as realizable. Within the good quantity of video data, we tend to initial cut back each video to a bunch of representative key frames.

1) Likelihood Ratio

The chance quantitative relation [4][5] could also be a region-based technique. It's a typical math distinction technique, which might be thought to be extension to component distinction. It are usually solve the matter of false detection due to very little camera motions. Instead of

scrutiny individual component, it compares the applied mathematics characteristic, the alleged chance magnitude relation, of the corresponding regions (i.e. blocks) in a pair of consecutive frames. If the chance magnitude relation is larger than a planned threshold, the region is believed to be being changed.

III. PROPOSED METHODOLOGY

Recent visual content-based video retrieval and categorization methodology is found to be economical and effective. This paper work aims to introducing a plenty of smart visual content–primarily based video retrieval and categorization methodology victimization advanced algorithms. The diagram are projected content–based video retrieval and categorization is illustrated in Figure one. This paper is search video accelerated by content-based video retrieval ways in which. supported the segmentation videos results we've got a bent to are able to modify the feature extraction for the ascertain information which means and classification and annotation the future scope The implementation of the projected methodology for shot boundary and Key frame is given below:

A. Shot Boundary Analysis

In the structural analysis we've to perform 3 tasks that square measure 1) shot boundary 2) key Frame 3) scenes half. The Shot transition detection is employed to separate up a show into basic temporal units known as shots; a trial may be a series of interconnected footage taken contiguously by one camera and representing endless action in time and space.

1) Planned Algorithm

The best due to discover if two frames unit of measurement significantly completely totally different is to count the amount of pixels that modification in value quite some threshold. This methodology is sensitive to camera motion. Zhang, Kankanhalli, and Smoliar1 implemented this methodology with the additional step of using 333 averaging alter before the comparison to reduce camera motion and noise effects together. The component variations for each region were sorted, and therefore the weighted add of the sorted region variations provided the image distinction live.

## 2) Key Frame Extraction

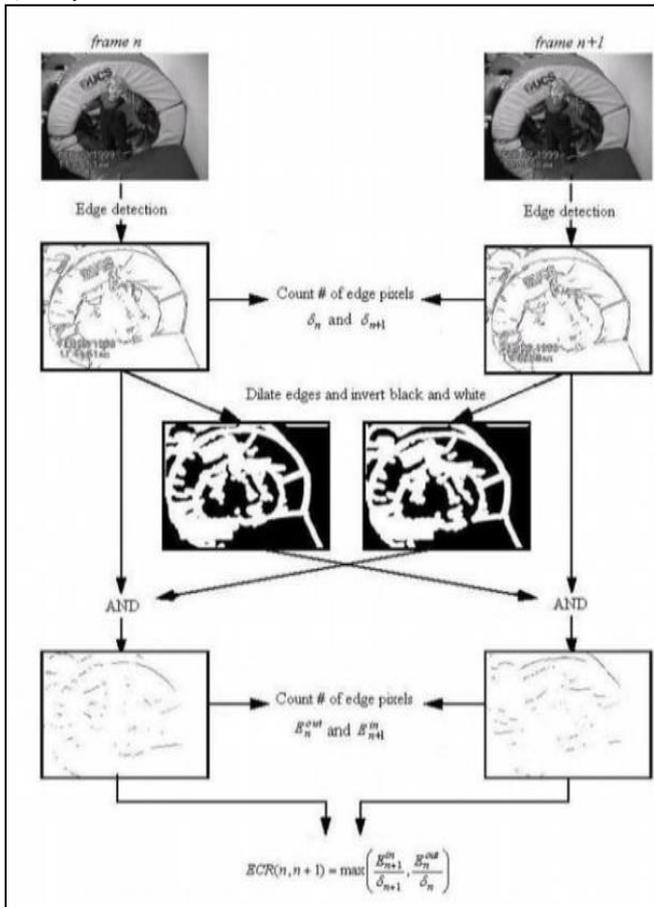


Fig. 5: Calculation graph of the edge change ratio (ECR)

Key border animation and filmmaking is drawing that defines the start and ending points of any sleek transition video. The drawings are said as "frames" as a result of their position in time is measured in frames on a strip of film. A sequence of key frames defines that movement the viewer can see. Throughout this paper, the subsequent methods to note the exhausting cuts in a passing given video is given. Histogram comparison: this system compares a feature of the entire frames, rather than the individual pixels or regions. It has been verified to be a lot of durable against the camera motions and object movements. The calculation of edge change magnitude relation (ECR) of two consecutive frames is given in Figure5.

## IV. CONCLUSION AND FUTURE SCOPE

A review on recent developments in Shot boundary detection and key frame extraction on visual content based video classification and retrieval is bestowed during this paper. The state of the art of existing approaches in each major issue has been delineating with the most specialize in the following tasks: video structure analysis at the side of shot boundary detection, key frame extraction.

A large amount of labor has been drained visual content-based video categorization and retrieval.

There are many issues are still open and be any analysis, significantly inside the area of feature extraction, video segmentation, question interface, and similarity live.

## REFERENCES

- [1] M. S. Lew, N. Sebe, C. Djeraba, and R. Jain, "Content-based multimedia information retrieval: State of the art and challenges," *ACM Trans. Multimedia Comput., Commun. Appl.*, vol. 2, no. 1, pp. 1–19, Feb. 2006.
- [2] Y. X. Peng and C.-W. Ngo, "Hot event detection and summarization by graph modeling and Matching," in *Proc. Int. Conf. Image Video Retrieval*, Singapore, pp. 257–266, Jul. 2005.
- [3] P. Over, G. Awad, J. Fiscus, and A. F. Smeaton. (2010). "TRECVID 2009–Goals,
- [4] Choi, K.-C. Ko, Y.- M. Cheon, G.-Y. Kim, H.-I, S.-Y. Shin, and Y.-W. Rhee, "Video shot boundary detection algorithm," *Comput. Vis., Graph. Image Process. (Lect. Notes Comput. Sci.)*, 4338, pp. 388–396, 2006.
- [5] A Survey on Shot Boundary Detection Algorithms and Techniques Hattarge A.M.1, Bandgar P.A.2, Patil V. M.3, February 2013
- [6] K. W. Sze, K. M. Lam, and G. P. Qiu, "A new key frame representation for video segment retrieval," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 9, pp. 1148–1155, Sep. 2005.
- [7] B. T. Truong and S. Venkatesh, "Video abstraction: A systematic review and classification," *ACM Trans. Multimedia Comput., Commun. Appl.*, vol. 3, no. 1, art. 3, pp. 1–37, Feb. 2007.
- [8] C.-Y. Chen, J.-C. Wang, and J.-F. Wang, "Efficient news video querying and browsing based on distributed news video servers," *IEEE Trans. Multimedia*, vol. 8, no. 2, pp. 257–269, Apr. 2006.
- [9] Swati D. Bendale, Bijal.J.Talati, "Analysis of Popular Video Shot Boundary Detection Techniques in Uncompressed Domain", *International Journal of Computer Applications (0975 – 8887) Vol. 60– No.3*, December 2012.
- [10] Abdelati Malek Amel, Ben Abdelali Abdessalem and Mtibaa Abdellatif, "Video shot boundary Detection using motion activity descriptor", *journal of telecommunications*, volume 2, issue 1, April 2010.
- [11] Baak A, Müller M, Seidel HP. An efficient algorithm for keyframe-based motion retrieval in the presence of temporal deformations. *New York: 1<sup>st</sup> ACM international Conference on Multimedia Information Retrieval*; pp. 451–458, 2008.
- [12] John, S. Boreczky and D. Lynn, 1998. A hidden markov model framework for video segmentation using audio an image features.
- [13] R. Kasturi and R. Jain, "Dynamic vision," in *Computer Vision: Principles*, R. Kasturi and R. Jain, Eds., IEEE Computer Society Press, Washington 1991.