A Review Paper on Content based Video Retrieval

Hetaxi Kamli[1] Mahasweta Joshi[2] U. K. Jaliya[3]
[2,3]Assistant Professor
[1,2,3]Department of Computer Engineering
[1,2,3]B.V.M Engineering College, Anand, Gujarat,388120, India

Abstract— Nowadays visual information is gaining much more importance than textual information as pictures convey more than words. As image has richer content than text, Video database is increasing exponentially over the Internet. Content-based Video retrieval (CBVR) from large video storage has become a challenging and important issue for researchers over worldwide. There is lack of tools so it is difficult to distinguish between low and high level feature. There has been lot of CBVR systems developed but still there is a lack of efficient retrieval system, as many retrieval techniques are not suitable for extensive search and so it deteriorates the overall performance of system. In this paper we have given an outline of the sequential steps of existing CBVR systems.

Key words: CBVR, Feature Extraction, SBD, Key-Frame, Video Indexing, Video Annotation

I. INTRODUCTION

Digital Video over Internet is emerging as a "boom" in the recent years. Digital video is leveraging the internet as it contains richer content than images. Information content in video are as follows: 1)Audio or Sound Track 2)Metadata like title, author, description etc 3)Speech transcripts, text caption 4)Visual information[12]. Due to the increased video database, its indexing and retrieval stage has become difficult to achieve. When we retrieve a video, users search the video in the form of query. Query can be on key-frame, motion, object, texture, color, shape, etc. A shot is defined as a unbroken sequence of frames taken by a single camera in an uninterrupted way. Shots are the fundamental units of a video. The shot consists of two types of transition: 1) Abrupt Transition 2) Gradual Transition [3]. Abrupt Transitions is a sudden transition without any delay. Gradual Transition is a kind of transition where two shots have some added spatial or chromatic effects. Abrupt Transition are known as hard or cut transition whereas Gradual Transition is known as soft transition and can be of various types i.e., dissolve, wipe , fade. Mainly research work on shot boundary detection (SBD) is been carried out on cut transition only as gradual transition is hard to detect. Features are the most important aspect in any video. Feature in video such as text, audio, object, etc. convey rich information for searching a desired video. Feature extracted are useless until we have a good retrieval system.

CBVR systems have a wide scope of applications such as, analysis of visual electronic commerce (such as analysis of interest trends of users' selections and orderings, analysis of correlations between advertisements and their effects), quick browsing of video folders remote instruction, digital museums, news event analysis, intelligent management of web videos (useful video search and harmful video tracing), and video surveillance [12].

The remaining of this paper is structured as follows: Methodology of Content Based Video Retrieval is described in Section 2. In Section 3, we explain future research required. Finally, we give our conclusions in Section 4.

II. METHODOLOGY FOR CBVR

The CBVR system follows the following steps as shown in the block diagram in figure1 [12].The CBVR system works on the following sequence: Shot Identification, Feature extraction, Video Annotation, Video Indexing and Retrieval and Browsing. We will give an overview about the different methods we have been reviewed in each step of the existing CBVR systems.

A. Shot Identification Stage:

It is the first stage carried out to provide necessary elements for indexing and retrieval stage. In this module we are finding shot boundary detection for different transitions like abrupt transitions and gradual transition. Abrupt transition is easy to detect but gradual transitions are not easy to identify because these is not eye catching sequence. There are various shot boundary detection techniques which are stated as follows: 1) Adaptive Threshold 2) Histogram Based Difference 3) Entropy 4) Block Based difference. Jaspreet Kaur at el.(2015),adaptive threshold works well than global threshold but the limitation of adaptive threshold is that it is difficult to estimate the threshold [4]. Junaid Baber at el. (2011) gave us different methods for Shot Boundary Detection (SBD) in which block based difference and histogram difference also have some limitations. Block Based difference solves all problems of histogram techniques except noise, as noise still exists in the output image. Histogram Difference is simple to implement but does not consider pixel position and can result in mistake [6]. B.V.Patel at el.(2010) in his paper used entropy to detect the scene change so can get the better result for cut, fade and dissolve transitions. Entropy method also have certain limitations that its value is nearly zero of a dark image and has problems with fade-in and fade-out transition [7].

B. Feature Extraction Stage:

There are various features such as shape, texture, color, object, motion. It is very time consuming process so we can reduce time complexity by using multicore- architecture. Feature extraction helps to reduce the size of the database frames.

1) Key-Frame Features:

Key Frame features are the essential visual property of any image. It consists of three types of features: 1) Shape Based Features, 2) Color Based Features and 3) Texture Based Features. Shape Based Features is used to extract edges in the image and then show it in the histogram. Edge Histogram (EH) uses Sobel operator to capture the spatial distribution of the edges [8]. Color Based Feature are most effective feature and it include color histogram color space such as RGB, HSV and YCbCr. Mostly Color Histogram is used due to its robustness to scaling, orientation, perspective

All rights reserved by www.ijsrd.com

and occlusion [8] [12]. Kulkarni at el. (2015) used ZigZag DCT to extract shape color and motion features [1]. Texture Based Feature is difficult to achieve. Gabor Filter [8] and Chi-Square [13] is used to detect texture in an image.

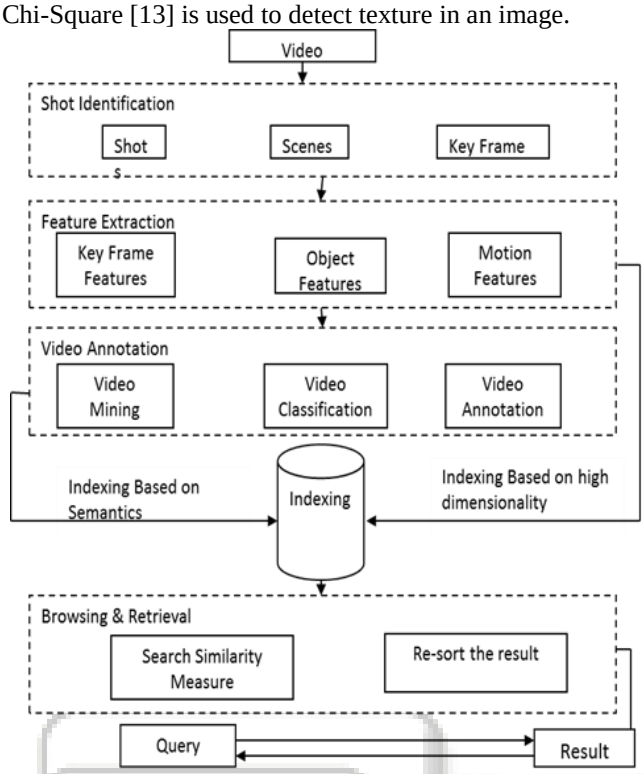

Fig. 1: Block Diagram of CBVR

2) Object Features:

Object Feature is used to detect person, vehicle or similar object in the image. Using Gaussian Mixture Model (GMM) we can detect object in the video. [9]. D. Hari Hara Santosh in their paper used GMM to detect moving object. It produces better result in less number of iterations. New algorithms have been developed which identifies different types of objects in the various scenes.

3) Motion Features:

Motion represents the two dimensional temporal change of video content [8]. It is used to distinguish dynamic videos from still images. It includes background motion by camera movement and foreground motion by moving objects. Camera motion can be done by as "zoom in or out," "panning left or right" and "tilting
up or down" [12].

C. Video Annotation Stage:

Video Mining, Classification and Annotation depend completely on the video pattern. And the extracted features. Video Annotation means to share video segments or video shots to different predefined semantic concepts such as vehicle, sky, mountain, river, person, etc. It is somewhat similar to video classification but has two major differences.1) Video classification has a different concept ontology compared with video annotation, although some of the concepts could be applied to both; and 2) video classification applies to complete videos, while video annotation applies to video shots or video segments [12]. Both classification and annotation used same methods where first low level features are extracted and then data are train to. Weiming Hu at el. (2011) has given a brief idea about different techniques used for video annotation i.e. Support vector Machines (SVM), K-Nearest Neighbour (KNN). KNN [12] has certain limitations that it is highly suspectible to noise in the training data and computationally expensive. The merits of SVM [10] are following: a) Training information and maintenance of the good generalization are fully utilized by them; b) By using the kernel function the large numbers of features are easily handled; and c) Already plenty of SVM codes are obtainable [14]

D. Video Indexing Stage:

Video Indexing are of two types: 1) Syntactic Indexing and 2) Semantic Indexing. Syntactic Indexing use syntactic features as the basis for matching and employ either Query-by-example or Query-through-dialog box to interface with the user [15]. Semantic indexing is confined to data modeling in specific domains in an unstructured manner. It is useful input for knowledge creation. For obtaining video indices the method widely used nowadays is K-Means Cluster. It works well with large variable and computationally faster than hierarchical cluster but it also have certain limitations that it doesn't work well with clusters of different size and different density [1] [12]. Madhav Gitre at el. (2014) in Content Based Video Retrieval System has used B+ Tree for indexing technique. B+Tree automatically reorganize itself with small local changes during insertion and deletion but its performance degrades if the database of video is large [10]. Kulkarni at el. (2015) used Fast pattern index (FPI) tree to traverse along the large video database. For FPI [1] we have to first use K-Means cluster [10] to generate cluster of video shots and then will implement FPI to generate temporal patterns.

E. Retrieval Stage:

Once we achieve video index, we can perform content based video retrieval. When a user gives a query, a similarity check is been performed with the help of following methods:1) Kullback Leibler Distance (KLD) and 2) Euclidean Distance Measure, T.N. Shangmagam in enhanced content based video retrieval used KullBack Leibler Distance(KLD) method to calculate similarity measure. KLD is used to get better result but requires large number of iterations. So we can use Euclidean Distance instead of it [8]. Euclidean Distance is simpler and faster and can work well with large iterations [7]. After performing similarity check we re-sort the result so retrieval becomes easier. Finally we get the result of the matched video in the retrieval stage.

III. FUTURE DEVELOPMENTS

We can eliminate some of the issues by making improvements in the following field:

- Using various good algorithms for gradual transition so we can get quite impressive results in shot boundary detection step.
- Feature extraction of low and high features should be done separately so we can get fine result.
- Indexing can be done using temporal patterns.
- Video Similarity measure check can be made dynamic and fast.
- To reduce time complexity and getting accurate and precise result.

All rights reserved by *www.ijsrd.com*

IV. CONCLUSION

For better overall performance we can combine different methods for shot boundary detection nd feature extraction stage so we can get more précised result. In this paper we have review different techniques of Content Based Video Retrieval process to enhance efficiency and robustness in the system. Therefore, we can conclude that it is rather difficult to get desired result for CBVR system in terms of effectiveness, efficiency and flexibility.

REFERENCES

[1] Dr. Parag Kulkarni, Bhagyashri Patil, Bela Joglekar –"An Effective Content Based Video analysis and Retrieval Using Pattern Indexing Techniques", 2015 International conference on Industrial Instrumentation and Control , 2015 IEEE.
[2] Hamdy.K.Elminir, Mohammed Abu ElSoud, Sahar F. Sabbeh, Aya Gamal –"Multi feature content based video retrieval using high level semantic concept", International Journal of Computer Science Issues , Vol9 , Issue 4, No.2, July 2012.
[3] Zhounglan Wu, Pin Xu –"Shot Boundary Detection In Video Retrieval", 2013 IEEE.
[4] Jaspreet Kaur, Navjot Kaur –"Key Frame Extraction from a Video Using Edge Change Ratio", International Journal of Advanced Research In computer Science and Software Engineering, vol5, Issue 5, May 2015.
[5] Hong Jiang Zhang, Jianhua Wu , Di Zhong, Stephen W. Smoliar –"An Integrated System For Content Based Video Retrieval and Browsing", Pattern Recognition, vol 30, Elsevier Science Ltd.
[6] Junaid Baber, Nitin Afzulpurkar, Matthew N. Dailey, And Maheen Bakhtyar –"Shot boundary detection from videos using entropy and local descriptor ", 2011 IEEE.
[7] B. V. Patel,A. V. Deorankar,B. B. Meshram –"Content Based Video Retrieval using Entropy, Edge Detection, Black and White Color Features", 978-1-4244-6349-7/10, 2010 IEEE.
[8] T.N.Shangmagam, Priya Rajendran –"An Enhanced Content Based Video Retrieval System Based on Query Clip", International journal of Research and Reviewa in Applied Sciences, vol 1,issue 3, december 2009.
[9] D. Hari Hara Santosh, P. Venkatesh, P. Poornesh, L. Narayana Rao, N. Arun Kumar –"Tracking Multiple Moving Objects Using Gaussian Mixture Model", International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-3, Issue-2, May 2013
[10] Madhav Gitre, Harshal Bawaskar, Sourabh Sethi, Ajinkya Shinde –"Content Based Video Retrieval System", International Journal of Research in Engineering and Technology, vol 3, Issue 6, Jun-2014.
[11] Nitya Raviprakash, Meggha Suresh, Asmitha Rathis, Divija Devarla, Aakanksha Yadav, Nagaraja G.S –"Shot Segmentation for Content Based Video Retrieval" ,Proceedings of International Conference ,"Computational Systems for Health & Sustainability", 17-18 April 2015
[12] Weiming Hu, Nianhua Xie, Li Li, Xianglin Zeng, and Stephen Maybank –"A survey on Visual Content Based Video Indexing and Retrieval", Senior Member, IEEE, , IEEE Transactions on systems, Man ans Cybernetics, Vol 41, no 6, Nov -2011.
[13] Prajesh V. Kathiriya, Dhaval S. Pipalia, Gaurav B. Vasani, Alpesh J. Thesiya, Devendra J. Varanva-"X2(Chi-Square) Based Shot Boundary Detection and Key Frame Extraction for Video" , Research Inventory: International Journal of Engineering and Science, January-2013
[14] Muhammad Nabeel Asghar, Fiaz Hussain, Rob Manton –" Video Indexing: A Survey" , International Journal of Computer and Information Technology (ISSN: 2279 – 0764) Volume 03 – Issue 01, January 2014.
[15] p.Geetha, Vasumathi Naryanam – "A Survey of Content Based Video Retrieval", Journal of Computer Science 4 (6), 474-486,2008 Science Publications.

All rights reserved by *www.ijsrd.com*