ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

Video Retrieval by Detecting Key Frames in CBVR

Poonam V.Patil[1], Prof.S.V.Bodake[2]

P.G. Student, Department of Computer Engineering, PVPIT, Bavdhan, Pune, India[1]

Assistant Professor, Department of Computer Engineering, PVPIT, Bavdhan, Pune, India[2]

ABSTRACT: In this paper a complex and wide area of Content Based Video Retrieval Systems has been presented in a comprehensive and simple way. Processes at different stages in CBVR systems are described in a systematic way. In proposed CBVR system used the behaviour of the Eigen values of covariance matrix over certain frames is analyzed to detect the shot boundary and subsequently extract the key frame means Eigen values is used measuring the dissimilarity between the consecutive frames, by using auto-dual threshold method. BTC and Gabor features algorithms are used in proposed to extract low level features BTC is used to obtain features from colour information of pixels belonging to the small blocks. Gabor features provide good representation of edge and texture features for objects and texts and help to distinguish them effectively from the background. The Kekre's Fast Codebook Generation (KFCG) Algorithm generated is used in the proposed system as a feature vector for video retrieval purpose. Kekre's Fast Codebook Generation (KFCG) Algorithm is basically used for image compression. It requires less time to generate the codebook through vector quantization method. Then use of SVM in automatic classification of videos, similar features extracted from other videos is used by SVM for classification of videos. The videos classified by SVM to form one category show greater similarity among them. Retrieval of most similar video form that category can be obtained by using Single Value Decomposition method , as per query video. The system is evaluated using a database of 1000 videos consisting of 20 different categories by calculating the graph of precision and recall.

KEYWORDS: Content based video retrieval (CBVR), key frames, near-duplicates, dual-threshold method, SVD , Eigen values ,SVM, Kekre's Fast Codebook Generation (KFCG), BTC , Gabor features.

I. INTRODUCTION

With the advent of computer technology, a large amount of data is generated and stored. The data is stored in various ways such as in the form of text, image, videos, etc. As the stored data is increasing enormously over the years, efficient retrieval of the data has become difficult. Video sharing on the web is constantly on the rise due to the fact that mobile phones equipped with cameras has increased over the years and number of people accessing internet has increased manifold and video is one of the most preferred type of data owing to its content richness. Content-based retrieval of video data has been still an active research area. The efficient retrieval of video data is proven a difficult because of the unstructured way the videos are stored. To date, handheld devices and the Internet have become a common method to create and transport video documents. As a result, there has been a huge increase in the utilization of video .So as to keep pace with the growth of video data; enhancing the current solutions for Content Based Video Retrieval (CBVR) is an important task.

Proposed approach consists of various modules for key frame extraction, indexing, features extraction, similarity search etc. We use a dynamic programming approach to compute the similarity between the feature vectors for the query and feature vectors in the feature database. Proposed Video Storage and Retrieval System, stores and manages a large number of video data and allows users to retrieve videos from the database efficiently. It is interactive web based application which takes video frame from users and retrieve the information from the database. Database consists of various video data like still video frames, audio and video. The retrieval is based on the content of the video object. Proposed System provides different functionality for two main clients-which are Administrator and user. Administrator is responsible for controlling the entire database including security and adding, updating and deleting

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

videos to and from database. User can only retrieve videos based on submitted query based on content as well on metadata.

II. LITERATURE SURVEY

Vijeetkumar Benni et.al [1] new technique for key frame extraction and shot boundary detection is proposed. The proposed technique uses Eigen values for measuring the dissimilarity between the consecutive frames. The behavior of the Eigen values of covariance matrix over certain frames is analyzed to detect the shot boundary and subsequently extract the key frame. In this paper a novel method for dynamically updating the covariance matrix is also proposed. This greatly reduces the computational time compared to conventional method of calculating the covariance matrix. The proposed technique is tested on a large variety of videos and also works on live videos. The recall and precision of simulation results reveal that the proposed technique is highly efficient and accurate in detecting the abrupt cuts.

Hong Liu et. al. [2] states segmentation and graph-based video sequence matching method for video copy detection. Specifically, due to the good stability and discriminative ability of local features, use SIFT descriptor for video content description. Then, we propose an SVD-based method to match two video frames with SIFT point set descriptors. To obtain the video sequence matching result, we propose a graph-based method. It can convert the video sequence matching into finding the longest path in the frame matching-result graph with time constraint.

Shimna Balakrishnan et. al. [3].explain features like color, shape and texture are considered for retrieval. In the new approach the frames are selected as multiples of a number and then the feature ext action takes place. The time involved is reduced to a great extent and error rate also is reduced. The main advantage of the system is it compares also the existing algorithms under each feature and analyses the performance. More features can be added so that there is precise and accurate retrieval of the videos.

Mohd. Aasif Ansari et. al. [4] states a complex and wide area of CBVR, and CBVR systems has been presented in a comprehensive and simple way. Processes at different stages in CBVR systems are described in a systematic way. Types of features, their combinations and their utilization methods, techniques and algorithms are also shown. For brighter future of CBVR, automatic retrieval systems should be the focus and it requires more attention from researchers for improved retrieval results. A trend to reduce computational cost is needed to project commercialized systems for video indexing, classification and retrieval to facilitate the availability of low cost, fast and efficient CBVR systems

Kalpana Thakre et. al. [5] presented a review on recent developments in content-based video indexing and retrieval. In this work, we have summarized various techniques used in content based video retrieval systems. Some of the most efficient and popular tasks used in CBVR include, video segmentation, video annotation, video indexing and creation of a video database, but the major challenge and research has to be performed in bridging the semantic gap through efficient annotation and indexing techniques.

The paper is organized as follows. Section II includes the general structure of video as well as general background techniques of video retrieval such as Content-Based Video Retrieval with framework structure. Section III includes problem definition. Section IV consists of proposed system architecture with detail explanation of each method which used in proposed system to improve performance of video retrieval .Section V includes a conclusion of proposed system, direction to continue this work in video applications. In the recent years, researchers are continuously focusing on solving the real problems .Section VI includes future scope of proposed system it is demonstrated that video retrieval techniques has wide scope which is utilized in the work of video summarization, video segmentation, video colorization and video analysis.

III. PROPOSED APPROACH

The generic framework of content based video retrieval system is as The framework of content-based video detection. It is composed of two parts: 1) An offline step. 2) An Online step.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

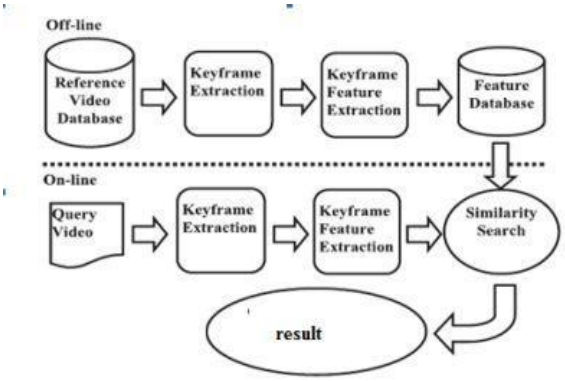

Fig 2.1. A framework on Content Based Video retrieval System

1) An offline step: Key frames are extracted from the reference video database and features are extracted from these key frames. The extracted features should be robust and effective to transformations by which the video may undergo. Also, the features can be stored in an indexing structure to make similarity comparison efficient.
2) An online step: Query videos are analysed. Features are extracted from these videos and compared to those stored in the reference database. The matching results are then analysed and the detection results are returned.

The framework accepts input as a video sequence with the standard structure consisting of shots, key frame and scene.Fig2.2 show the hierarchical structure of video.

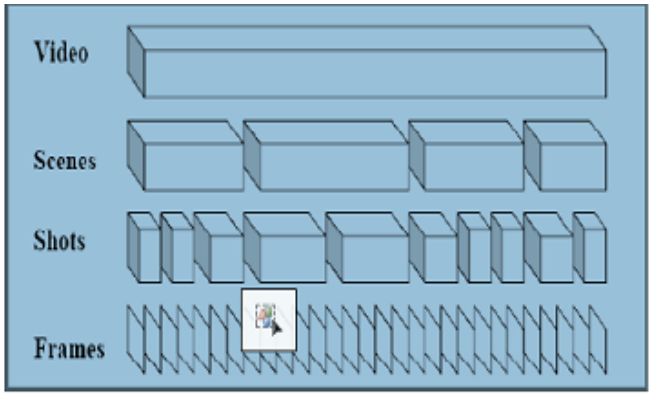

Fig 2.2. Hierarchical Structure Of Video

IV. PROBLEM DEFINITION

To improve performance and speed of CBVR by an optimized Key frame extraction by eliminating redundant video frames in CBVR .For these purpose used efficient algorithm and techniques on each step of CBVR system and calculate graph of precision and recall to showing performance and efficiency of system video retrieval system.

V .IMPLEMENTATION DETAIL

Proposed CBVR system show the combination of algorithm those gives best result on each step of CBVR and help to improve the performance ,accuracy ,and efficiency of video retrieval system. CBVR system include the six steps 1) Video Segmentation 2) Key Frame Extraction 3) Feature Extraction 4) Feature classification and indexing 5) Similarity Measures 6)Video Retrieval

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

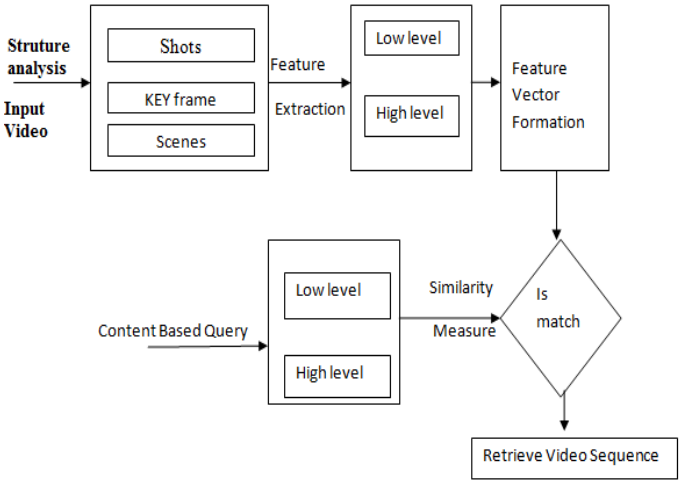

Figure 4.1: System architecture

step 1: Video Segmentation:-

In video segmentation divide the video into frames, for video segmented is done commonly by using the shot boundary detection with the different distance measurement algorithm. The video shots can be obtained by the shot boundary detection method. To detect near duplicates frames in video data set, used Auto- dual threshold method. A video retrieval by removing near duplicates in CBVR. These method aims to eliminate the near-duplicate frames along the video time direction; it does not take into account the concept of the shot, also does not require post processing to obtain the actual shots.

step 2: Shot detection and Key frame Extraction by eliminate Redundant Video Frames

The proposed method for shot segmentation and key frame extraction has four stages: generation of data matrix for k consecutive frames, computation of co-variance matrix, computation of Eigen values and shot boundary detection/key frame extraction. Initially the covariance matrix of k consecutive frames of the video stream is calculated. The Eigen values of this covariance matrix is calculated which are used to find the similarities/dissimilarities between the consecutive frames. Based on the similarities/ dissimilarities the key frames are extracted. These stages are explained in detail

- Generation of Data Matrix

Let $\{f_1, f_2, f_3, f_4, \ldots f_k, f_{k+1}, f_N\}$ be the number of frames in a video V. Generating the data matrix D.

$$D = P_1, P_2, P_3, \ldots, P_k$$

The first step is to move the origin to mean of the data. This is achieved by finding the mean image by averaging the columns of D, then subtract the mean image from each image in the data matrix (i.e., each column of D) to create the mean centered data vector which is U.

- Computing the Co-variance matrix

The purpose of computing the co-variance matrix is to find the variations between the intensity levels of consecutive frames, the computed co-variance matrix is used to derive the Eigen values. Generally, the covariance matrix is computed by using the below formula

$$C = U^T * U/n - 1$$

- computation of Eigen values

The k Eigen values of the covariance matrix C is obtained the minimum Eigen value chosen is used to find the dissimilarities between the frames.

- shot boundary detection/key frame extraction

If the minimum Eigen value exceeds the predefined threshold value then the $(k-1)^{th}$ frame will be the shot boundary and k^{th} frame is said to be a key frame of the next shot indicating the beginning of the new shot. If the k^{th} frame is not a key frame then the data matrix is modified by eliminating the first frame and appending the next

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

consecutive frame $(k+1)^{th}$ to the data matrix, else the data matrix is newly created with the consecutive frames starting from k^{th} frame to next k frames.

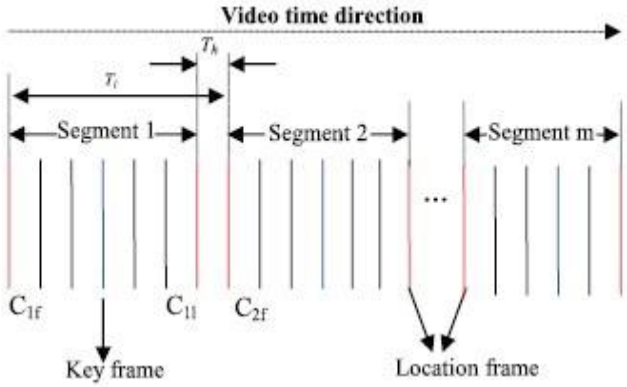

Figure 4.2: Eliminate redundant video frames, select key frames.

step 3: Feature Extraction:-

In video analysis the features applicable to images are considered same for video signal. Features are extracted for the key frame and stored into feature vector. Features are of two types that are spatial (low level) and temporal(high level). The fundamental techniques for content based video retrieval are: spatial analysis, and temporal analysis:

- **Spatial Analysis:**

  Spatial Analysis uses parameters like color, shape and texture.

  1. Color :

Color is the most extensively used visual content for video retrieval. Its three dimensional values makes its 53 discrimination potentiality superior to the single dimensional gray values of videos. Retrieving the videos using color is done by computing a color histogram which identifies the proportion of pixels within an image holding specific values. BTC features are calculated for small blocks formed by dividing an image instead of calculating for each pixel . BTC is used to obtain features from color information of pixels belonging to the small blocks. m1 and m2 are the mean values found for the three components . m1 = {mR1, mG1, mB1} and m2 = {mR2, mG2, mB2} m1 and m2 represent the entire block. Mean values of all the blocks considered together represent the entire image.

2. Shape and Texture

The usage of shape features for video retrieval is restricted be-cause robust and accurate video segmentation is difficult to achieve. Shape, moreover, does not refer to the shape of a video but to the shape of a particular region that is being sought out. Shapes will be determined by applying segmentation or edge detection techniques. Texture representations can be classified as structural and statistical. The are the visual patterns in the videos and how they are spatially defined. Texture represented by texels gives the relative brightness of consecutive pixels and finds the degree of contrast, regularity, coarseness and directionality which classifies textures as 'smooth', 'rough' etc.

Gabor features provide good representation of edge and texture features for objects and texts and help to distinguish them effectively from the background. Gabor filters are capable of extracting features from edges or regions of different objects inside an image directed towards desired orientations with different frequencies

- **Temporal Analysis:**

  Temporal Analysis is done based on parameters like motion and audio.

  1. .Motion:

Motion is an intangible factor which is used for querying. Though tapping the motion vectors for the videos is a task which is why still motion based querying is in its infancy, once this concept is perfected it will help in video retrieval to the best accuracy level and is more dependable than visual and textual query.

2. Audio:

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

Unlike popular belief, from content-based video retrieval point of view, the audio information can be even more important than the visual part since it is mostly unique and significantly stable within the entire duration of the content. However, audio-based studies lag far behind the visual counterpart and the development of robust and generic systems for audio content management is still in its infancy.

Step 4: Stored features in Feature vector table:

Codebook vectors are represented by a set of code words which are used to encode and decode the images .Kekre's Fast Codebook Generation (KFCG) Algorithm is basically used for image compression It requires less time to generate the codebook through vector quantization method. The codebook generated is used in the proposed system as a feature vector for video retrieval purpose

Step5: Classification of Features

Support Vector Machine (SVM) improves performance of content based image retrieval (CBIR) significantly. It is the inspiration to use SVM for CBVR too. Similar features extracted from other videos are used by SVM for classification of videos. Use of SVM is a milestone in automatic classification of videos with better efficiency.

step 6: Similarity matching :-

In singular value decomposition method, match feature point sets based on SVD. The SVD method has been widely used in pattern recognition, data compression, signal processing, and other fields. Specially, the goal of this method is to use the SVD method to reduce and correct the wrong match between the two points in twofeature sets. The method focused on the one point-to-one point correspondence. However, SVD method to measure the similarity between two feature point sets, and emphasize the similarity of frame-to-frame

VI. RESULT

Performance of the system is evaluated based on the Precision and Recall values.

Precision = No. of retrieved videos that are relevant to the query clip / Total no. of retrieved videos

Recall = No. of retrieved videos that are relevant to the query clip / Total no. of relevant videos available in database

After final retrieval of the video from the database, various results are shown such as value of precision and recall then show that value of precision and recall using bar chat graph. Below Table-I shows the how much Precision and Recall is calculated for a given set of query clips.

Table I: Precision and Recall for a given set of query clips

Sl. No.	Query clip	Precision	Recall
1	Q1	0.10	0.90
2	Q2	0.26	0.88
3	Q3	0.31	0.86
4	Q4	0.70	0.32
5	Q5	0.89	0.27

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

After final retrieval of the video from the database value of total no. of frames of retrieve video, and total no. Of key frames of that video also stored in database then user can seen that values with the help of bar chat graph. Below graph show total no. of frames and total no. of key frames of retrieve result video for query Q1.

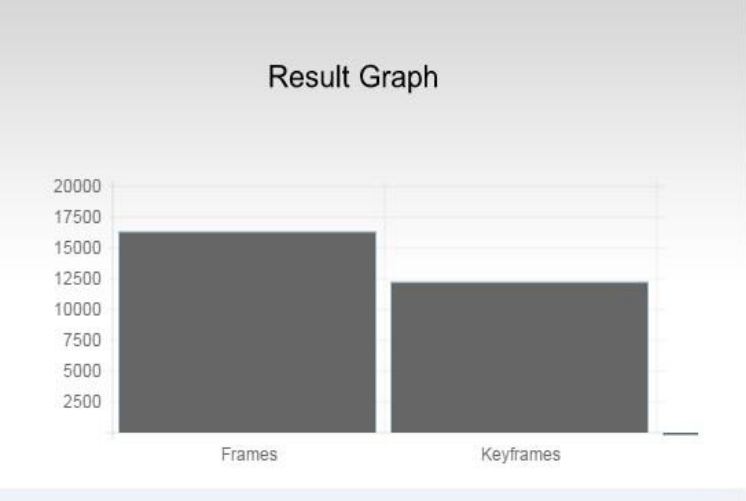

Figure 6.1: Results

VII. CONCLUSION AND FUTURE WORK

In CBVR key frame extraction and elimination of near duplicates will improve the video storage, search and retrieval speed. Using the Auto-dual threshold method for elimination of near duplicate frame extraction the number of frames can be reduced to 30-40 percent. A dual-threshold method to eliminate redundant video frames and use the SVD based method to compute the similarity of two SIFT feature point sets. These will definitely help to achieve better accuracy up to 35-36 percent.

In this paper this method of key frame extraction is applied only in one pass. It can be implemented in multiple passes and improved the performance of video retrieval process. As key frames are the collections of only important content data, once extracted from the videos can be utilized anywhere for video retrieval such as video summarization, video colorization, video annotation etc.

REFERENCES

1. Vijeetkumar Benni, R. Dinesh, Punitha P., and Vasudeva Rao" Keyframe Extraction and Shot Boundary Detection Using Eigen Values", International Journal of Information and Electronics Engineering, Vol. 5, No. 1, January 2015.
2. Hong Liu, Hong Lu, Member, IEEE, and Xiangyang Xue, Member, IEEE "A Segmentation and Graph-Based Video Sequence Matching Method for Video Copy Detection", IEEE transactions on knowledge and data engineering, vol. 25, no. 8, august 2013.
3. Shimna Balakrishnan, Kalpana S . Thakre "VIDEO MATCH ANALYSIS:A Comprehensive Content based Video Retr ieval System", International Journal of Computer Science and Application Issue 2010.
4. Mohd.Aasif Ansari, HemlataVasishtha " Content based Video Retrieval Systems Performance based on Multiple Features and Multiple Frames using SVM",(IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 7, No. 8, 2016
5. Kalpana Thakre, Archana Rajurkar and Ramchandra Manthalkar, "An Effective CBVR System Based On Motion, Quantized Color" Science & information Technology (IJCSiT), Vol 3, No 2, April 2011.
6. Sudeep D. Thepade , Ashvini A. Tonge "An Optimized Key Frame Extraction for Detection of Near Duplicates in Content Based Video Retrieval" International Conference on Communication and Signal Processing, April 3-5,2014, India
7. Sanjoy Ghatak, Debotosh Bhattacharjee "Extraction of Key Frames from News Video Using EDF, MDF AND HI Method for News Video Summarization" In- ternational Journal of Engineering and Innovative Technology (IJEIT) Volume 2,Issue 12, June 2013
8. Weiming Hu, Senior Member, IEEE, Nianhua Xie, Xianglin Zeng, and Stephen Maybank , "A Survey on Visual Content-Based Video Indexing and Retrieval", IEEE Transactions On Systems, Man, And CyberneticsPart C: Applications And Reviews, Vol. 41, No. 6, November 20i i, pp no 797-819.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

9. Xiao Wu, Wan-Lei Zhao, Chong-Wah Ngo, "Near-Duplicate key frame Retrieval with Visual Keywords and Semantic Context", CiVR'07, July 9-11,2007, Amsterdam, The Netherlands.
10. R.T. Minu, K.K. Thyagharajan, "Scrutinizing the Video And Video Retrieval Concept", international Journal oj Soft Computing and Engineering (JJSCE), ISSN: 2231-2307, Volume-I, issue-5, November 2011.

BIOGRAPHY

Poonam V. Patil is a Student of M.E. Computer Engineering in the Padmabhushan Vasantdada Patil Institute of Technology, Pune, India

Prof. Sarika .V.Bodake is a Faculty of Computer Engineering Department, Padmabhushan Vasantdada Patil Institute of Technology, Pune, India