Content Based Video Retrieval With Motion Vectors and the RGB Color Model

Bandi Ritish Reddy [1], K. Prathyusha [2], G. S. Naveen Kumar [3]
1. Malla Reddy College of Engineering and Technology, Secunderabad, India
2. Malla Reddy College of Engineering and Technology, Secunderabad, India
3.Research Scholar, JNTU Kakinada, India

Abstract:**The surveillance system is widely used in business and government organizations, such as post offices, convenient stores, banks and public streets. The conventional surveillance systems only provide video data that record events. When an event occurs, we can check the video data to figure out what happened. However, the event searching process from the video data is tedious and time consuming. Recently, content based video retrieval technologies have become a popular research topic. The content based video retrieval technologies can be used in two applications: "on line" and "off-line". In "on-line", we can make the surveillance system to track target automatically. They can provide a warning signal for police or security guards when suspects appear or crime happens. This kind of video retrieval systems may provide a possible way to prevent crime happening. In "off-line", we can extract the target events from video data and avoid tedious searching processes. In this paper, we use the RGB color model to retrieve interested moving objects from recorded video data.Experiments verify the effectiveness of the proposed approach**

Key words:
CBVR, Shot Segmentation, Key frame extraction, Feature extraction.

1. INTRODUCTION

With recent advances m multimedia technologies, digital TV and information highways, more and more video data are being captured, produced and stored. However, without appropriate techniques that can make the video content more accessible, all these data are hardly usable. So the research on the management of video data is now a hot field. But the differences between multimedia and textual data in continuity and dimensionality make traditional database technology unavailable for access to handle the multimedia information. Consequently, the content-based access and retrieval become a proper solution.

Video content can be grouped into two levels: low-level visual features and high-level semantic content. Low-level visual content is characterized by visual features such as color, shapes, textures, etc; On the other hand, semantic content contains high-level concepts, such as objects and events [1].

The need for efficient content based video retrieval has increased tremendously in many application areas such as biomedicine, military, commerce, education and web image classification and searching. Currently, rapid and effective searching for desired videos from large scale video database becomes an important and challenging research topic[2]. Computational complexity and retrieval efficiency are the key objectives in the design of CBVR system[3].

The paper is organized as follows. Section 2 gives our method to generate successive frames using key frame extraction and obtain feature vector using feature extraction.[4] The video retrieval system is presented in section 3, followed byexperimental results in section 4. Conclusion with future scope in section 5 concludes the paper

2. PROPOSED ALGORITHM

2.1 Key Frame Extraction:

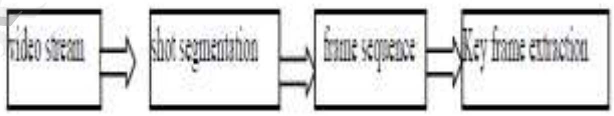

Image database is a storage area which is source of all key frames used in the Video retrieval process. In order to reduce the transfer stress in network and invalid information transmission [5], the transmission, storage and management techniques of video information become more and more important. Video segmentation and key frame extraction are the bases of video analysis and content-based video retrieval [6]. Key frame extraction, is an essential part in video analysis and management, providing a suitable video summarization for video indexing, browsing and retrieval. The use of key frames reduces the amount of data required in video indexing and provides the frame work for dealing with the video content. In recent years, many algorithms of key frame extraction focused on original video stream. It can introduce processing inefficiency and computational complexity when decompression is required before video processing. Key frame is the frame which can represent the salient content and information of the shot [7]. The key frames extracted must summarize the characteristics of the video, and the image characteristics of a video can be tracked by all the key frames in time sequence.

Furthermore, the content of the video can be recognized. A basic rule of key frame extraction is that key frame extraction would rather be wrong than not enough. So it is necessary to discard the frames with repetitive or redundant information during the extraction. A new algorithm of key frame extraction from compressed video data is presented in

this paper. We analyze the features of compressed data and finally obtain the key frames. For video, a common first step is to segment the videos into temporal “shots,” each representing an event or continuous sequence of actions. A shot represents a sequence of frames captured from a unique and continuous record from a camera. Then key frames are to be extracted. Video segmentation is the premise of key frame extraction, and key frames are the salient content of the video (key factors to describe the video contents)[8].
Every Frame is compared with its adjacent frame’s Pixel intensity value using the method “Color histogram two comparison method”. The Correlation Coefficient of the compared frames is obtained from the above method[9].
An appropriate Threshold Value is selected.
The correlation coefficients obtained from comparing all consecutive frames are compared with the threshold value. The frames which are below the threshold value are selected as key frames, the frames above the threshold value are neglected. Hence, the key frames below threshold value are selected and stored in the image database.

2.2 Feature Extraction:
In pattern recognition and in image processing, feature extraction is a special form of dimensionality reduction. When the input data to an algorithm is too large to be processed and it is suspected to be notoriously redundant (much data, but not much information) then the input data will be transformed into a reduced representation set of features (also named features vector). Transforming the input data into the set of features is called feature extraction [10]. If the features extracted are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input. Feature extraction involves simplifying the amount of resources required to describe a large set of data accurately. When performing analysis of complex data one of the major problems stems from the number of variables involved. Analysis with a large number of variables generally requires a large amount of memory and computation power or a classification algorithm which over fits the training sample and generalizes poorly to new samples. Feature extraction is a general term for methods of constructing combinations of the variables to get around these problems while still describing the data with sufficient accuracy. Suppose that the key frame set R consists of K frames, R = KFj| j = 1,2,□,k}, while the shot frameset S consists of N frames, S =Fi | i =1,2,□, k} . Let the distance between any two frames KFjand Fi be d(KFj , Fi) . Define di for each frame Fi as: di = min(d(KFj, Fi)) , j =1,2,□, kThen the Semi-Hausdorff distance between S and R is given as: dsh= max(d i) , i =1,2,□, N .

The extraction of visual features of an image is called as feature extraction.The Extraction of color is much Easier and reliable than texture and shape. In the CBVR system, features are extracted on the basis of color. Features of an image should have a strong relationship with semantic meaning of the image[11]. CBVR system retrieves the relevant images from the image data base for the given query image, by comparing the features of the query image and images in the database.In color distribution and quantization is used for color image retrieval. Relevant images are retrieved according to minimum distance or maximum similarity measure calculated between feature of query image and every image in the image database.Only color Layout is used to extract the color features because global color features gives too many false positives.An Image can be divided to three matrices Red, Green and Blue matrices as given in figure:2.2..

3. CBVR SYSTEM

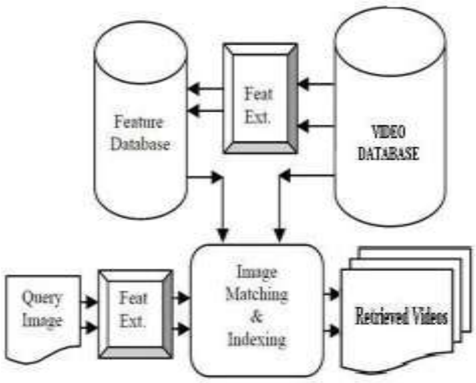

Fig 3.1: Block Diagram for CBVR

3.1 Video Database
Video database is digital memory storage area in which a large content of videos are stored.The user can search this data base in different processes for his/her desired videos[14].Video retrieval is a process through which the user can search the database by giving an image which has similar color or textures of his desired video.In the CBVR system we chose AVI(Audio Video Interleave) format videos to be used.AVI files can contain both audio and video data in a file container that allows synchronous audio-with-video playback. CBVR Video database contains about 20 different videos.A Video is a Stream of continuous sequence of images called as “Frames”. Every second in a video has 24 frames. A frame can be considered as an image, which has visual features of that video.
Query Image
A query image is an image given by the user to identify his/her video from the database. The query image contains the visual features of the video such as color, texture and shape. We keep 20 images which can be used as query images, i.e., those are the images which have maximum number of features present in the video. The feature vector of the query image given is generated online and compared with the feature vectors present in the feature database.

Feature Vector
Color layout method is used for detecting the features of the image, i.e., for both the images in the database and the query image. In an RGB image, there are three intensities for each and every pixel, so there are three separate matrices present for every image. Those matrices can be called as Red matrix, Green matrix and Blue matrix. Each of those

matrices is divided into sub-blocks. Mean of every block is calculated and stored in a matrix called Feature Matrix. There by a feature vector consists of all features of an image, i.e., it represents an image.

Image Matching

Image matching is an important step in the CBVR system, this is the place where main task of Video retrieval is processed. The user gives a query image to the system[15], the feature vector of the query image is found and it is sent to the Image Matching Block where it is compared with all feature vectors related to key frames are stored in the database. Euclidean distance formula is used here to match the feature vectors. The formula is shown below.

Where Elements 'P1,P2,P3…' represents the mean values of query image ,where as elements 'q1,q2,q3….' are the mean values of feature vectors stored in the database. All the feature vectors are matched with the query using above formula, if the result is nearer to zero, then images are said to be matching or if it is nearer to one, then those images are matched. The image which has more features compared to the query is retrieved[16].

4. EXPERIMENT &ANALYSIS

In this paper, experimental data set contains 100 videos, using key frame extraction each video has 10 successive frames, the video data base is divided into 109 categories, and each category has 10 videos.

KEY FRAME EXTRACTION

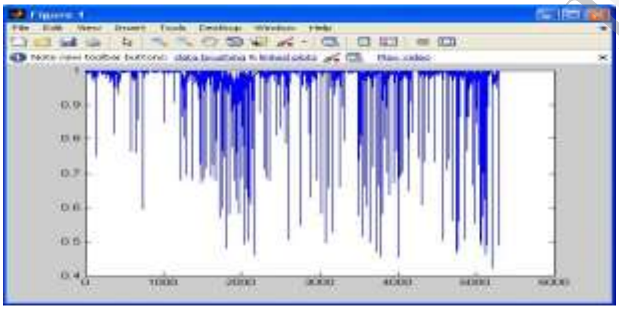

Fig 4.1: Histogram Based Correlation coefficients comparison plot

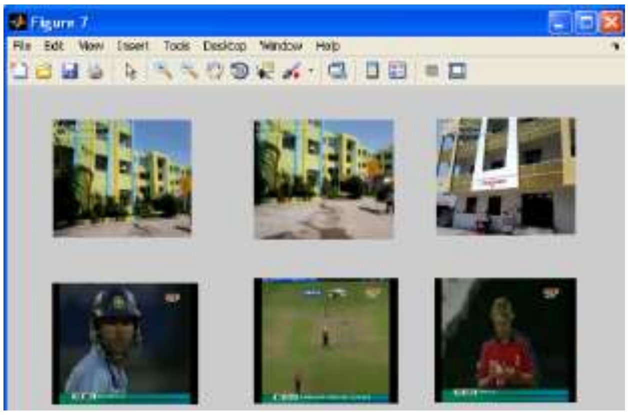

Fig 4.2: Extracted key frames

QUERY BY EXAMPLE

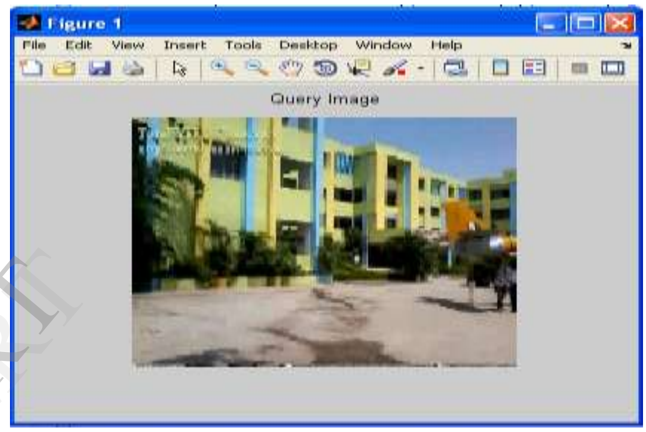

Fig 4.3: Query Image

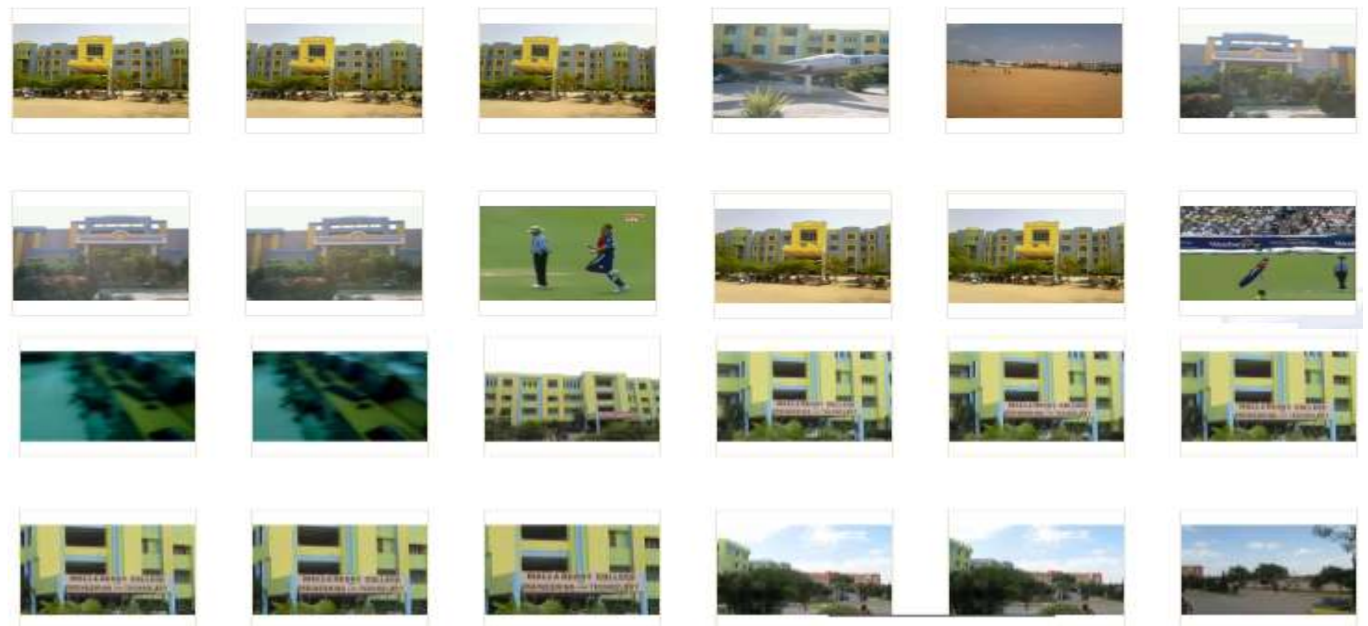

Fig 4.4: Retrieved Videos

CONCLUSION& FUTURE SCOPE:

In this paper we present our current research about developing a video retrieval system based on color lay out. This system works in a query by example manner. It has been successful at all the stages i.e., at key frames extraction, feature extraction and finally video retrieval.Experimental results have shown good performance. Our future work involves including features such as shape and texture to retrieve similar video objects more robustly.

REFERENCES

1. M Petkovic. Content-based Video Retrieval. Conf. onExtending Database Technology (EDBT), Ph.D.Workshop, Konstanz, Germany, March 2000
2. Costas Cotsaces, Nikos Nikolaidis, and Ioannis Pitas,“Video shot detection and condensed representation: a review,”IEEE Signal Processing, vol. 23, no. 2, pp. 28-37, 2006.
3. T. Liu, H. Zhang, and F. Qi, “A novel videokey-frame-extraction algorithm based on perceived motionenergy model,” IEEE Transactions On Circuits And SystemsFor Video Technology, vol. 13, no. 10, pp. 1006-1013, 2003.
4. S.F.Chang and et.al, “A fully automated content-based videosearch engine supporting spatiotemporal queries,” IEEETransactions on Circuits and Systems for Video Technology,vol. 8, no. 5, pp. 602–615, September 1998.
5. Y. Deng and B. S. Manjunath, “Netra-v: Toward an objectbasedvideo representation,” IEEE Transactions on Circuitsand Systems for Video Technology, vol. 8, no. 5, pp. 616–627, September 1998.
6. S. Antani, R. Kasturi, R. Jain, "A survey on the use of pattern recognition methods for abstraction, indexing and retrieval of imagesandvideo", Pattern recognition 35 (2002) 945-965.
7. M.K. Mandal, F. Idris, S. Panchanathan, "A critical evaluation of image and video indexing techniques in the compressed domain",Imageand Vision Computing 17 (1999), 513-529.
8. N. Dimitrova, H.J. Zhanhg, B. Shahrary, I. Sezan, T. Huang, A.Zakhor "Applications of Video-Content Analysis and Retrieval", IEEEMultimedia, Vol. 9, No. 4, 2002.
9. A paper on “Keyframe-based Video Summarization using Delaunay Clustering” by PADMAVATHI MUNDUR, YONG RAO, YELENA YESHA, Department of Computer Science and Electrical Engineering,University of Maryland Baltimore County1000 Hilltop CircleBaltimore, MD 21250
10. a paper on”Key Frame Extraction from MPEG Video Stream” by Guozhu Liu, and Junming Zhao College of Information Science & Technology, Qingdao Univ. of Science & Technology,266061,P. R. China.
11. a paper on “Introduction to Digital Video Processing” by Alan C. BovikLaboratory for Image and Video Engineering (LIVE),Department of Electrical and Computer Engineering,The University of Texas at Austin.
12. A paper on “Automated Content Based Video Retrieval” by Sr.an Zagorac, AinhoaLlorente, Suzanne Little, Haiming Liu, Stefan RÄuger Knowledge Media Institute, The Open University Walton Hall, Milton Keynes, MK7 6AA, UK.