



VIDEO FOOTAGE RETRIEVAL SYSTEM USING FEATURE PROFILE ANALYSIS ON TIME AXIS

Er. Amarjit Kaur ^{1*}, Er. Parminder singh ²

^{1,2}DIET, Kharar, Mohali, Punjab, India.

*Correspondence Author: Amarjeet2013@yahoo.com

Keywords: CBVR, wavelet coefficient, GLCM, Entropy

Abstract

The video footage retrieval is normally done by using the contents like shape, color or personal personality etc in most of the cases considered so far. However, the existing video retrieval system suffers from the drawback of retrieving the video based on the undefined activity observed in video sequence. This can also be stated in a way that the existing system cannot judge by themselves whether the activity in a video is abnormal or normal. This kind of self detecting mechanism is proposed in the presented work and can be termed as adaptive system of frame feature detection based on video footage retrieval system. The proposed concept may be a time consuming from the computational point of view. The reduction of computation time and image itself as a content for image retrieval has been proposed in the presented work to work on.

Introduction

Content-based video retrieval (CBVR) is the application of computer vision techniques to the video footage image retrieval problem i.e. searching for digital video content in long time recording of a spot using close circuit TV camera e.g. in ATMs, shopping malls etc. Content-based method means that the search analyzes the contents of the video recording rather than the metadata search keywords, tags, or descriptions associated with the recording. The term "content" in this context refers to any unusual activity in an event for long time, like staying of a person for a long time or more than expected time of a person in an ATM.

Till date, most of the work done on content-based video retrieval deal with problems like video positioning and key frame extraction. The compact representation of video content for shot similarity measurement and shot retrieval remains one of the most challenging issues. In the recent literatures, video shots are mostly represented by key frames. Low-level features such as color, texture, and shape are extracted directly from key frames for indexing and retrieval. For efficiency reason, video retrieval is usually processed and analyzed in a similar way as image retrieval. Such strategy, however, is ineffective since spatio-temporal information existing in videos is not fully exploited. Content-based video indexing relies on the processing of a set of features extracted from a video sequence.

Related work

Based on the work done in content based video retrieval speeded up robust feature descriptor a system is presented that retrieves the similar videos based on local features. The video indexing information is gathered but due to the high dimensionality of SURF features large consumption of storage memory required during indexing of the video information. To reduce the dimensionality on speeded up robust feature descriptor. It presents a model data for the videos. Then for the video retrieval the model data of the test clip is classified to its similar videos using minimum distance classifier. [1]

Video retrieval using information theory system is divided into three main parts that is shot boundary detection, hierarchical video summarization, retrieval and index the target video. The result is evaluated on the basis of entropy feature and the databases related to a video. The system is composed of two parts first is database and second is video retrieval. These parts are used for the detection of efficient features to analyze the video. After that the boundary detection algorithm is applied to the segmented video to divide into number of frames. For this there are three main methods are used that are frame blocking, which is mainly based on the entropy factor and another one is block weighting which uses the weighted information for differencing estimation between the frames. The performance evaluation is determine by using the measure like precision and recall values.[2]

Another method named frame based decision pooling method for video classification classifies videos into different classes. In this the model extracts frames from videos and maps the low level features. The various features used here such as color structure and texture. These features extracted from a standard image database. The extracted frames are classified according to five different classes such as mountains, forests, buildings deserts and seas. Here the method applied use the color feature which will detect or retrieved the frame by using histogram, energy and entropy. Second method uses the algorithm h detection that is the structural feature another feature used is texture feature which represent the spatial arrangement of the grey level of the pixel. The entropy of grey scale image is statistical measure of randomness. Based on all these features support vector machine classifiers classify all the frames.[3]



One of the methods named Content based Video Retrieval using Particle Swarm Optimization introduce an efficient methodology that lead to incremental improvement in the video search results against a users query image. [2]

An Effective Content Based Video Retrieval Utilizing Texture, Color and Optimal Key Frame Features presents the intention of retrieving video for a given query, the raw video data is represented by two different representation schemes, video segment representation (VSR) and Optimal key frame representation (OFR) based on the visual contents. [3]

Semantic Web for Content Based Video Retrieval a semantic web based video search engine. A scalable integration platforms to represent extracted features from videos, so that they could be indexed and searched.[4]

Algorithm

The steps of the work done in proposed method are described in steps as below:

1. Video Footage Acquisition: Convert the video footage into 2-d frames i.e. image format. Present the video frames as images for frames retrieval based on query features like color, entropy, texture and time based activities.
2. Video Footage Retrieval based on Color: When the video footage is to be retrieved based on some color, the query color is first indexed using the CAT command in matlab. Say, the query color is C, then, C may be decomposed into its RGB component using the index color coding. Say the RGB components are given by:
 - R-Component = $R/256$;
 - G-Component = $G/256$;
 - B-Component = $B/256$;
 The color components are divided by 256 to make the color components in the range of 0-1. The query color is obtained by the followings:
 - Query Color C = cat (R-Comp, G-Comp., B-Comp)
 A RGB histogram is extracted from the data base images by decomposing the images into index images. A Color Check Margin (CCM) or threshold is defined and if the query color components (R-, G- and C-) are greater than the CCM in data base images, those images are presented to the user.
3. Video Footage Retrieval based on Texture: The texture based features are extracted from the query image and the query image features are compared with the database image features using the distance measure. Images having the least distance with the query image are displayed as the result. Image Texture features are extracted using Gray Level Co-occurrence matrix (GLCM).
4. Video Footage Retrieval based on Shape: The shape based features are extracted from the database images and stored in a feature database. The shape based features are extracted from the query image and the query image features are compared with the database image features using the distance measure. Images having the least distance with the query image are displayed as the result.
5. Video Footage Retrieval based on Entropy: The entropy is extracted from the database images and stored in a feature database. To extract entropy from the query image, the query image features are compared with the database image features using the distance measure. Images having the least distance with the query image are displayed as the result. Entropy is a statistical measure of randomness that can be used to characterize the texture of the input image.
6. Finally all these features that are extracted as stated above are plotted on the time axis.
 - Color/Gray Histogram
 - Mean Red Color Component (μ_R)
 - Mean Green Color Component (μ_G)
 - Mean Blue Color Component (μ_B)
 - Standard Deviation in R-Component Image (σ_R)
 - Standard Deviation in G-Component Image (σ_G)
 - Standard Deviation in B-Component Image (σ_B)
 - Entropy
 - Contrast
 - Power
 - Energy
 - Homogeneity
 - Correlation
 - Wavelet LL-sub-bands image Coefficients
 - Time based profile of each of the above features



Performance measures

Precision and recall are used to evaluate the performance of a content based image/video retrieval system.

Precision: Precision is defined as the ratio of no. of relevant images retrieved to the total no. of image retrieved.

$$\text{Precision (P)} = \frac{\text{Total No.of Relevant Image Frames Retrieved}}{\text{Total No.of Images Frames Retrieve}}$$

Recall: Recall is defined as the ratio of no. of relevant images retrieved to the total no. of relevant images present in the data base.

$$\text{Recall (R)} = \frac{\text{Total No.of Relevant Image Frames Retrieved}}{\text{Total No.of Relevant Image Frames in Dtabase}}$$

Based on these two parameters the results so obtained are discussed in next section.

Results

The time based features are important from the scanning of video clips primarily from the security/forensic application. For example, shopping mall, crowded place bank premises and ATM etc. Here, the continuity of the features are observed and related using the time axis. Below image in fig.1 No. shows the video frames retrieval based on time based analysis.

The query color is obtained by the followings:

Query Color C = cat (R-Comp, G-Comp., B-Comp);

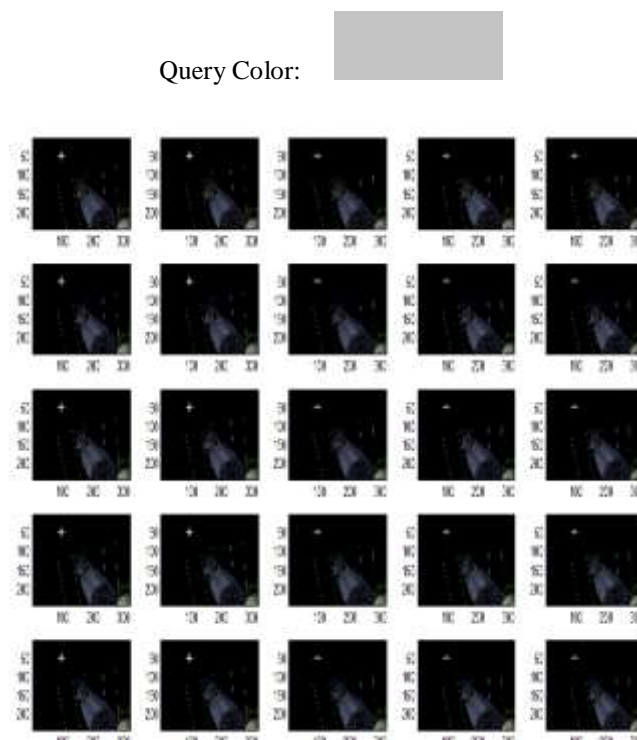


Fig. 1. (Retrieved Images based on Time based Features)

The accuracy is always higher in case of image/frames retrieval when the parameters are taken over the time axis i.e. time based features. The precision and recall value are the performance measures for any retrieval algorithm.

The recall and precision are computed for different video footage clips and summarized in the below table:



Video No.	Total Image Retrieved based on query frame image	Total Relevant Image Frames Retrieved	Total Relevant Images in Database i.e. Video	Precision (P)	Recall (R)	Time(secs.)
1.	125	120	122	0.96	0.98	5.52
2.	130	126	129	0.97	0.97	6.43
3.	105	103	105	0.98	0.98	7.58
4.	75	68	73	0.91	0.93	4.52
5.	63	60	62	0.95	0.97	4.25

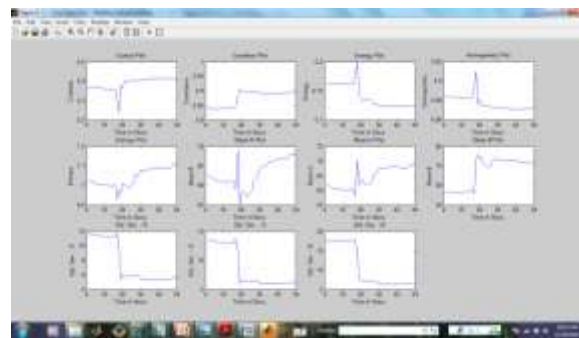


Fig. 2 Props. Vs Time Plot

Fig. 2 shows the different properties trends with respect to time. As the props plot remains flat for the respective time period, it can be said that the same activity is taking place. Prolonged flat region may be enquired after retrieval and analyzed. The retrieval of this region is very important when analyzing the video retrieval.

Conclusion

Long videos most importantly suffers from a drawback of storage space because as number of frames increases this also increases the requirement of memory, as the whole video sequence is saved in form of contents like time parameters. So if we reduce the time lag between the frames, some sort of memory conservation has been achieved. This is also one of the benefits of adding time parameter to my proposed work. Since time parameter has been added, so when time starts counting, the role of color, entropy and texture comes into play. When a particular frame has been retrieved it shows parameter values according to time. Hence results are obtained. This may improve the learning from the video analysis in the domain of time based activities, entropy, texture and color. Further, as most of the time, there are normal or routine activities in any video footage of ATM or shopping complex or any other public place. Therefore, an average scene based frames could be removed from the entire video and only the frames having unusual activities could be inspected. This reduces the search space and optimizes the query and in turn reduces the time consumption. As seen from the results, the time lag has been reduced successfully between the frames. This work can be further extended by reducing the time further hence improving the video retrieval time more.

References

1. Asha and Sreeraj, (2013) “Content Based Video Retrieval using SURF Descriptor”, 978-0-7695-5033-6/13 \$26.00 © 2013 IEEE
2. Ayesha Salahuddin, Alina Naqvi, Kainat Mujtaba and Junaid Akhtar (2012)” Content based Video Retrieval using Particle Swarm Optimization, 978-0-7695-4927-9/12 \$26.00 © 2012 IEEE



3. S.Padmakala, Dr.G.S.AnandhaMala and M.Shalini (2011) “ An Effective Content Based Video Retrieval Utilizing Texture, Color and Optimal Key Frame Features”, 978-1-61284-861-7/11/\$26.00 ©2011 IEEE.
4. 4.Sancho C Sebastine, Bhavani Thuraisingham and Balakrishnan Prabhakaran (2009) , “Semantic Web for Content Based Video Retrieval”, 978-0-7695-3800-6/09 \$26.00 © 2009 IEEE
5. Yan Yang, Brian C Lovell and Brian C Lovell (2009), “Content-Based Video Retrieval (CBVR) System for CCTV Surveillance Videos”, 978-0-7695-3866-2/09 \$26.00 © 2009 IEEE
6. J. P. Collomosse, G. McNeill and Y. Qian (2009), “Storyboard sketches for content based video retrieval presents a novel Content Based Video Retrieval (CBVR) system”, 978-1-4244-4419-9/09/\$25.00 ©2009 IEEE
7. Tong Lin,Chong- Wah Ngo, Hong-Jiang Zhang and Qing- Yun Shi (2001), INTEGRATING COLOR AND SPATIAL FEATURES FOR CONTENT-BASED VIDEO RETRIEVAL”, 0-7803-6725- UO 1/\$10.00 0200 1 IEEE
8. Markos Zampoglou, Theophilos Papadimitriou and Konstantinos Diamantaras, Member (2003),CONTENT-BASED VIDEO RETRIEVAL USING THE SHOT CLUSTER TREE”, 7803-7865-2/03/\$17.00 02003 IEEE
9. Markos Zampoglou, Theophilos Papadimitriou, Member and Konstantinos Diamantaras (2007) “Support Vector Machines Content-Based Video Retrieval based solely on Motion information”, 1-4244-1566-7/07/\$25.00 ©2007 IEEE.
10. Samita Chavan,Aurangabad and Shubhangi Sapkal (2013), “Color Content based Video Retrieval”, International Journal of Computer Applications (0975 – 8887) Volume 84 – No 11, December 2013
11. Ashwini S. Mane, P. M. Kamde (2013) , “Video Classification using SVM”, International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-2, Issue-3, July 201
12. Kimiaki Shirahama,Kazuyuki Otaka and Kuniaki Uehara.(2007), “content Based Video Retrieval Using Video Ontology”, . 0-7695-3084-2107 \$25.00 O 2007 IEEE
13. Bei Xiong ,Haojie Wen,Xue Liu,Xinxing Li,Zetian Fu and Lingxian zhang (2013), “Design and development of crop video retrieval system”, Vol.11 (3&4): 1863-1866. 2013
14. Zimian Li,Ming Zhu (2013), “A Light –Weight Relevance Feedback Solution for large scale content-based retrieval”, IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 1, No 3, January 2013
15. Lili N.A.(2009),”Hidden Markov Model for Content-Based Video Retrieval”, 978-0-7695-3648-4/09 \$25.00 © 2009 IEEE