



Copy Frame Detection in Video by Using Data Mining Techniques

K.Girija¹, R.Maheswaran², P.Sabarinathan³

PG scholar, Pavendar Bharathidasan College of Engg and Tech, Tiruchirappalli, Tamilnadu, India¹

PG scholar, SKP Engineering College, Thiruvannamalai, Tamilnadu, India²

Research scholar, Pavendar Bharathidasan College of Engg and Tech, Tiruchirappalli, Tamilnadu, India³

ABSTRACT: There are a number of methods available for video copy detection. Some of the methods were employing the application of local and global descriptors which were found to be ineffective in detections involving complex transformations. In order to overcome the above specified inefficiency, SIFT descriptors came into picture but was found to have a high computational cost. The method proposed in this paper involving five different types of MPEG-7 descriptors namely CEDD, FCTH, SCD, EHD, CLD descriptors for extracting the features of the frames in the selected video is found to be cost effective and efficient even in case of high level of transformations. This paper also throws light on certain improvements in graph-based video sequence matching method which is used to overcome the level of noise, to detect videos with different frame rates and optimal sequence matching is found automatically from the disordered video sequences by applying spatial features during copy detection. Experimental results have showed that the proposed method is far effective than the previously existing video detection scenarios.

KEYWORDS: Video copy detection; graph; SIFT feature; SVD; CEDD, FCTH, SCD, EHD, CLD descriptors; graph-based matching

I. INTRODUCTION

Now-a-days, server space is becoming a major issue for high level organizations to maintain enormous amount of data. For instance, organizations like YouTube, Google, Metacafe and others dealing with enormous video storage are found to have acres of racks consisting of hard disks each could be holding a capacity of around 1 to 2TB. But we are aware of the fact that most of the videos available in these video storage websites are redundant. According to the recent statistics [13], there are about 27 percent of redundant videos in You Tube, Google videos. Redundant videos are of two types: Copy videos and near duplicate videos.

A copy can be defined as a segment of video derived from another video, usually undergone a lot of transformations, such as Cam-cording, PiP (Picture in Picture), Insertions of patterns such as captions, subtitles, logo; Strong re-encoding; Change of gamma; Decrease in quality: Blur, frame dropping, compression, ratio and white noise, Post production is shown in Table 1. All of the above specified transformations can be done to a specific original video or else some of these transformations can also be done based on the needs of the one who is implementing these transformations. Both of these above specified cases are considered to be a copy of the original version of that respective video which is available in the dataset maintained by us.

The framework of video copy detection is done as two parts:

- (1) Client side: Key frames are extracted from the reference video dataset and features are extracted from those key frames. The features extracted should not be vulnerable to video transformation. The features can be stored in a feature database to make similarity comparison effective.
- (2) Server side: Query videos are verified. Features are extracted from the videos and are compared with the features which are already stored in the database.
- (3) The sequence matching results are returned along with the detection results [5].

II. THE PROBLEM

Watermarking is one of the classical techniques handed down for a long time for the purpose of representing the ownership of a particular video by means of inserting the logo or any other digital signature of the owner [1]. This information can be either visible to human eye or even can be hidden. This seems to be a better way for representing the ownership but it is vulnerable to the slightest changes. For instance, this watermarking information can be removed easily by anyone by making any small transformation to the watermarked video which includes re-encoding and change of bit rate. It is important to note that watermarking is not destined to retrieve a query video clip. To overcome this demerit, CBCD (Content Based Copy Detection) came into existence. Methods which are employing the use of global descriptors were found to take the low level features into account always for the purpose of video copy detection. But this technique is found to work well only for those videos which have undergone some small amount of transformations. In order to overcome the demerits of global descriptor, local descriptor were employed for video copy detection which makes use of the local spatiotemporal features and then making use of the features around the local features which helps them to be effective even in the case of high level transformations [3]. One of the efficient local descriptor that came into existence was SIFT (Scale Invariant Feature Transform).

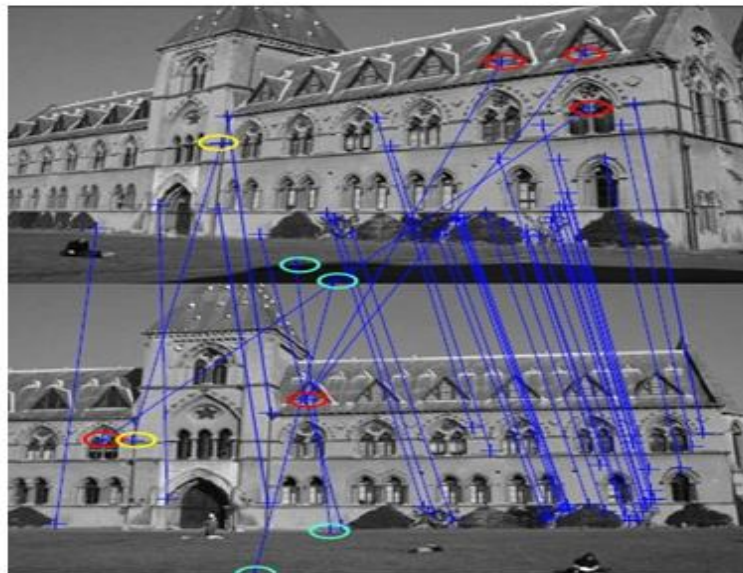


Fig. 1. The matching result using SIFT

In existing system, SIFT descriptor [6] seems to have a good stability to describe video characteristics. The SIFT descriptor (Local Descriptor) is found to give a better performance in identifying the objects is shown in Fig.1. The methods based on local descriptors on points, lines and also shape found to play an important role in video copy detection. Of the above mentioned properties, spatiotemporal interest points are found to play a vital role in classifying human actions and to detect periodic motions It has a good tolerance to scale changes, illumination variations and image rotations and also it has a greater strength against affine distortion and additive noise. Performance of methods based on local descriptor is found to be far better than the methods based on global descriptor in terms of logo insertion, shifting or cropping. In spite of the above specified merits, high computational cost is found to be highest demerit.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

III. CONTRIBUTION

In order to increase the efficiency of video copy detection and make it more effective, we have employed the use of five MPEG7 type descriptors such as CEDD, FCTH, SCD, EHD, CLD descriptors are used to extract the feature for Sequence Matching and also used to reduce the computational complexity. They describe elementary characters such as shape, color, texture or motion among others.

CEDD – Color & Edge Directivity Descriptor

Low computational power for extraction compared to others. It incorporates color and texture information in histogram. CEDD low level features that are extracted from the images can be used for indexing and also for retrieval.

FCTH – Fuzzy Color & Texture Histogram

FCTH is used for extraction of features that combines color and texture information in one histogram. This feature is appropriate for accurately retrieving images even in case of distortion such as deformation.

SCD –Scalable Color Descriptor

SCD derived from a color histogram defined in a Hue Saturation value color space with fixed color space quantization.

EHD – Edge Histogram Descriptor

EHD use global, semi-global, local histogram generated from local histogram bins to increase performance. MPEG-7 is used for the case of video or still images containing texture, color, shape and motion.

CLD – Color Layout Descriptor

CLD is designed to capture spatial distribution of color in image. It is used to describe color relation between sequence and group of images.

Graph-Based Method

Graph method can find the longest path in the frame matching-result graph considering time constraint

- It can automatically remove the noise caused by visual feature matching.
- It is adaptive to video frame rate change.

It can find the optimal sequence matching result automatically. The above specified five types of descriptors are used to extract exact individual features from the frames which are more accurate for comparison of the results. In addition to that graph based method is used to make this method an optimal one for video copy detection.

IV. SYSTEM DESIGN

The system design involves the different steps involved in the proposed video copy detection method. The entire process is summarized in the Fig.2 which gives a clear cut idea about the proposed method.

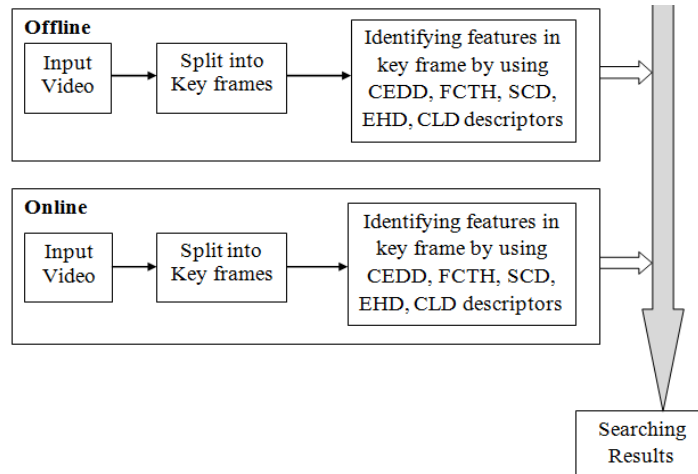


Fig. 2. System Architecture showing Client and Server side

Client side:

It is important to note that the client side and offline process are one and the same which can be represented in any of the descriptions. The procedure followed at the client side involves getting the user input in the form of video. First step would be extracting the key frames from the user given input video. The different features are extracted from the extracted key frames using the different type of descriptors and are stored into the reference database.

Server side:

As specified above, server side is similar to that of online process which can be said in either ways. The process done at the server side is similar to that of the process done at the client side which is the following. Frames are extracted from all the videos available in the server side and the features are already extracted using the proposed five types of descriptors and are stored individually in the server side database for future reference. It is to be noted that each time when the user is uploading a new video into the server, key frames are extracted along with the features and are updated into the database maintained by the server.

V. SYSTEM IMPLEMENTATION

The above specified concepts are implemented in various steps which are described as follows:

a. Splitting of video

In order to analyze the video given as input by the user, the different desired properties of the video is found out such as color, size and quality. The input video should have high quality and brightness to achieve better results. Key frame extraction is done at this stage.

b. Performance of ROI

After key frame extraction is done, the main object features are extracted from the key frames by using ROI (Region Of Interest) and these extracted information is made used to detect the copied videos is shown in Fig.3.

c. RGB color histogram:

This module can used to find out various RGB properties of the frames such as mean, standard deviation, level, count at different scenarios by applying changes to the image using different types of filters such as HSL color contrast to increase the image brightness, Edge detectors, morphology, binarization and lot more. A histogram is shown in Fig.4

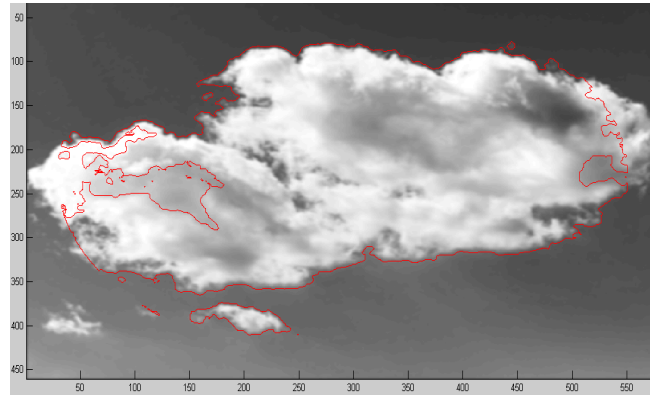


Fig. 3. ROI for extracting main object

d. Sequence Matching:

The extracted key frames are stored in the database and the random frame is chosen here to be compared with all the other frames in the dataset which is done effectively with the help of CEDD, FCTH, SCD, EHD, CLD descriptors is shown

e. Similarity matching detection:

The random frame from the user input is matched with all the frames available in the server for each and every video available in the server and based on the similarity matching results, server can allow the user to upload the video if it is an original one (i.e. it is not available in the server) whereas it would not allow the user to proceed with the upload if the video is a copied one (i.e. the video is already available in the server).

VI. RESULTS AND DISCUSSIONS

The results of the compared videos are formulated in the form of graph with the help of the proposed graph based method. Results show that it is very easy to compare the results in the form of graph which shows the values formulated as a result of the comparison done between the randomly chosen frames along with each and every individual frame in the dataset. The graph also gives a clear comparative study between the values computed by all the five different type of descriptors.

Similarity Matching Results/Descriptors	CEDD	FCTH	SCD	EHD	CLD
FRAME 1	0	0	12	22	0
FRAME 2	5.3	7	17	26	13
FRAME 3	5.9	7.8	21	32	16
FRAME 4	6.5	8.7	25	36	20

Table 2 Similarity Matching Result of Video A

Table 2 shows the similarity matching result of video A done by five different type of descriptors which is represented in the form of graph in Fig.6 whereas Table 3 shows the similarity matching result computed for video B by the descriptors which is graphically represented in Fig.7

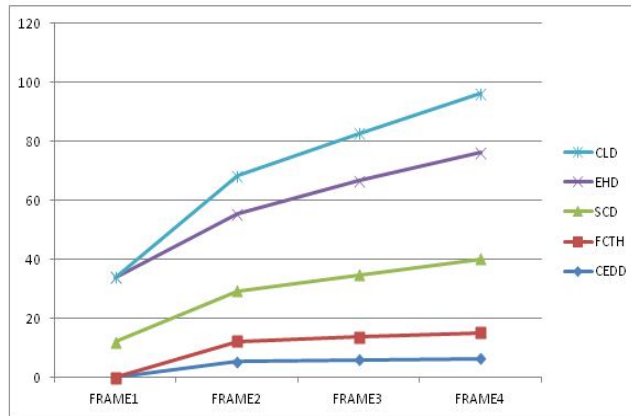


Fig. 6. Graph-Based Method representing Table 2

On comparing the two graphs, we can come to a conclusion that the values computed by the descriptors are entirely different so that both the videos are not the same or we can say that video A is not a copy of video B. Analyzing the above results, the server would allow the user to upload video A into the server as it is not available within the server space.

Similarity Matching Results/Descriptors	CEDD	FCTH	SCD	EHD	CLD
FRAME 1	0	0	14	24	1.3
FRAME 2	6.4	9	16	27	15
FRAME 3	7.1	9.8	20	34	18
FRAME 4	8	10.7	26	38	22

Table 3 Similarity Matching Result of Video B

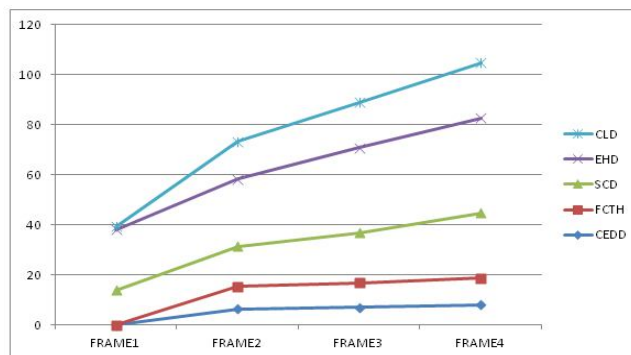


Fig. 7. Graph-Based Method representing Table

VII. CONCLUSION AND FUTURE ENHANCEMENT

Video copies are analyzed and the features are used for copy detection. Based on the analysis, local feature of SIFT was used for describing video frames. Since large number of SIFT points are extracted from a video, the copy detection involving SIFT features has high computational cost. And so the new descriptors have been introduced: EHD, CLD, SCD, CEDD, and FCTH. Edge Histogram Descriptor (EHD) captures the edge spatial distribution around the image in five



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

orientations: horizontal, vertical, 45° degrees, 135° degrees and non-directional. Color Layout Descriptor (CLD) is used for capturing the spatial distribution of color in an image. Scalable Color Descriptor (SCD) can obtain good performance in case of retrieval accuracy and fast matching of the image. Color and Edge Directivity Descriptor (CEDD) captures and relates shape, texture and color from an image. Fuzzy Color and Texture Histogram (FCTH) aims at capturing the image texture, shape and color.

By using these descriptors, the computational cost will be decreased and it can be used for obtaining the effectiveness in key frame feature extraction. Experimental results prove that the proposed method is found to have an optimal tradeoff between the effectiveness and efficiency of video copy detection.

Futuristic approach can be done to improve the efficiency of the descriptors in terms of scalability and to make the result far more accurate than the proposed method.

REFERENCES

- [1] Arun Hampapur, Ki-Ho Hyun and Ruud B “Comparison of Sequence Matching Techniques for Video Copy Detection,” Proc. SPIE, Storage and Retrieval for Media Databases, vol. 4676, pp. 194-201.
- [2] Delponte E., Isgro F., Odone F. and Verri A. (2006) “SVD-Matching Using Sift Features,” Graphical Models, vol. 68, no. 5, pp. 415-431.
- [3] Geert Willems, Tinne Tuytelaars and Luc Van G. (2008) “Spatio-Temporal Features for Robust Content-Based Video Copy Detection,” Proc. ACM Int’l Conf Multimedia Information Retrieval (MIR), pp. 283-290.
- [4] Heng Tao Shen, Jie Shao, Zi Huang and Xiao fang Z. (2009) “Effective and Efficient Query Processing for Video Subsequence identification,” IEEE Trans. Knowledge and Data Eng., vol. 21, no. 3, pp. 321-334.
- [5] Hong Liu, Hong Lu and Xiangyang Xue (2013). “A Segmentation and Graph-Based Video Sequence Matching Method for Video Copy Detection,” IEEE Transaction on Knowledge and Data Engineering, vol. 25.
- [6] Hong Liu, Hong Lu and Xiang yang X. (2010) “SVD- SIFT for Web near duplicate image detection,” Proc. IEEE Int’l Conf. Image Processing (ICIP ’10), pp. 1445-1448.
- [7] Julien Law-To, Li Chen and Alexis J. (2007) “Video Copy Detection: a Comparative Study,” Proc. ACM Int’l Conf. Retrieval, pp. 371-378.
- [8] Li Chen and Stentifordb. F.W.M. (2008) “Video Sequence Matching based on Temporal Ordinal Measurement,” Pattern Recognition Letters, vol. 29, no.13, pp. 1824-1831.