ISSN 2250-3153

Efficient Framework for Video Copy Detection Using Segmentation and Graph-Based Video Sequence Matching

A.PerumalRaja[*], B.Venkadesan[**], R.Rajakumar[***]

[*] M.E student, Department of Computer Science, Arasu Engineering College, kumbakonam
[**] M.E student, Department of Computer Science, Arasu Engineering College, Kumbakonam
[***] Assistant Professor, Department of Computer Science, Annai Group of Institutions, Kumbakonam.

Abstract- A segmentation and graph-based video sequence matching method specifically, due to the good stability and discriminative ability of local features, we use SIFT descriptor for video content description. However, matching based on SIFT descriptor is computationally expensive for large number of points and the high dimension. Thus, to reduce the computational complexity, we first use the dual-threshold method to segment the videos into segments with homogeneous content and extract Keyframes from each segment. SIFT features are extracted from the Keyframes of the segments. 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. Experimental results demonstrate that the segmentation and graph-based video sequence matching method can detect video copies effectively. Also, the proposed method has advantages. Specifically, it can automatically find optimal sequence matching result from the disordered matching results based on spatial feature. It can also reduce the noise caused by spatial feature matching. And it is adaptive to video frame rate changes. Experimental results also demonstrate that the proposed method can obtain a better tradeoff between the effectiveness and the efficiency of video copy detection.

Index Terms- Video copy detection, graph, SIFT feature, dual-threshold method, SVD.

I. Introduction

The rapid development and wide application of multimedia hardware and software technologies, the cost of image and video data collection, creation, and storage is becoming increasingly low. Each day tens of thousands of video data are generated and published. Among these huge volumes of videos, there exist large numbers of copies or near-duplicate videos. According to the statistics of, on average, there are 27 percent redundant videos that are duplicate or nearly duplicate to the most popular version of a video in the search results from Google video, YouTube, and Yahoo! video search engines. As a consequence, an effective and efficient method for video copy detection has become more and more important. A valid video copy detection method is based on the fact that "video itself is watermark" and makes full use of the video content to detect copies. To facilitate the discussion of "video copy" in this paper, we use the definition of video copy in TRECVID 2008 tasks. Definition of copy video: A video V1, by means of various transformations such as addition, deletion, modification (of aspect, color, contrast, encoding, and so on), cam-cording, and so on, is transformed into another video V2, then video V2 is called a copy of video V1. The objective of video copy detection is to decide whether a query video segment is a copy of a video from the video data set. A copy can be obtained by various transformations. If a video copy

Detection system finds a matching video segment, it returns the name of copy video in the video database and the time stamp where the query was copied from. Fig. 1 shows the framework of content-based video copy detection.

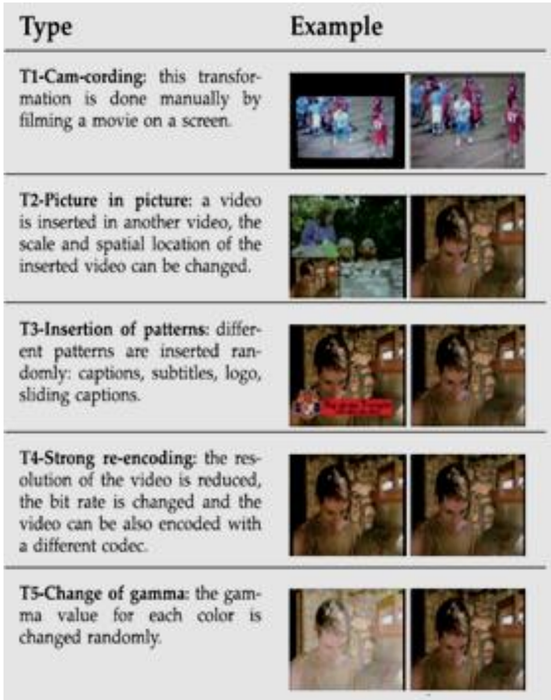
It is composed of two parts:

1) An offline step. Keyframes are extracted from the reference video database and features are extracted from these Keyframes. 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 analyzed. Features are extracted from these videos and compared to those stored in the reference database. The matching results are then analyzed and the detection results are returned.

TABLE 1

Based on the study, in these transformations, picture in picture is especially difficult to be detected. And for detecting this kind of video copies, local feature of SIFT is normally valid. However, matching based on local features of each frames in two videos is in high computational complexity.



Examples of Five Single Transformations

T1. Cam-cording; T2. Picture in picture; T3. Insertions of pattern: Different patterns are inserted randomly: captions,subtitles, logo, sliding captions; T4. Strong re-encoding; T5.Change of gamma; T6, T7. Decrease in quality: Blur, change of gamma (T5), frame dropping, contrast, compression (T4), ratio, white noise; T8, T9. Post production: Crop, Shift, Contrast,caption (text insertion), flip (vertical mirroring), Insertion of pattern (T3), Picture in picture (the original video is in the background); T10. Combination of random five transformations among all the transformations described above.

II. RELATED WORKS

As reviewed in many content-based video copy detection methods have been proposed. Furthermore, copy is a subset of near duplicate. Copies have an origin, while near-duplicates may not. Specifically, two news videos on the same event from two broadcasting corporations are not copies, but near duplicates since they deliver the same information to audience, although some variations on the scenes may exist. Also, there are many methods proposed on near-duplicate detection. The methods on copy and near duplicate detection can be grouped into two types.
One type of copy detection methods uses global descriptor. Specifically, Hampapur et al. compared distance measures and video sequence matching methods for video copy detection. They employed convolution for motion direction feature, L1 distance for ordinal intensity signature (OIS), and histogram intersection for color histogram feature. The results show that the method using OIS performs better. Yuan et al. combined OIS with color histogram feature as a tool for describing video sequence.

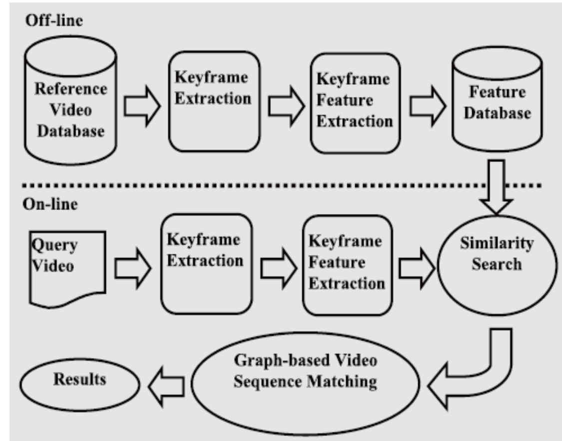
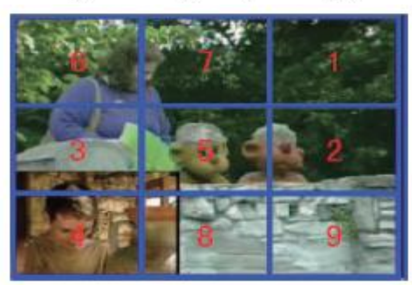


Fig. 1. A framework of video copy detection system

As the basis, designed region intensity rank signature along time sequence. Specifically, they divided each video frames along the time sequence into several blocks and proposed average gray values for each block. Then, they linked gray values of these divided blocks separately along the time direction before they use those sequence information to describe the video content. Shen et al. Introduced a real-time near-duplicate video detection system, UQLIPS, which globally summarized each video to a single vector.

Huang et al. used global image feature such as color histogram and texture to represent each video frame. Wu et al. adopted the color histogram in HSV color space to detect and remove the majority of duplicates of web videos.

(a) OIS (query image)



(b) OIS (PiP image)

As SIFT descriptor has good stability and discriminative ability, we choose SIFT descriptor to describe video characteristics. Meanwhile, we suggest two solutions to the lack of high computational cost in the process of copy detection: 1) dual-threshold method to eliminate video redundant frames; 2) using singular value decomposition (SVD) for matching two feature sets of SIFT features on key points.

ISSN 2250-3153

c) The Matching Result Using SIFT

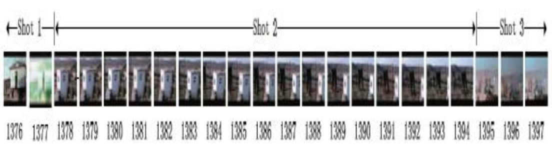
III. USING AUTO DUAL-THRESHOLD METHOD TO ELIMINATE REDUNDANT VIDEO FRAMES

Normally, visual information of video frames is temporally redundant. So, video sequence matching is not necessarily to be carried out using all the video frames. An effective way of reducing non necessary matching is to extract certain Keyframes to represent the video content. And the matching of two video sequences can be first performed by matching the Keyframes.

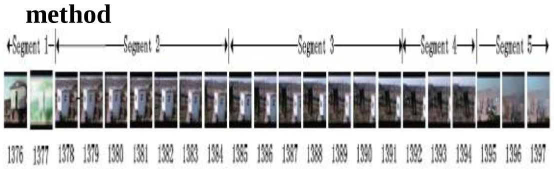
Specifically, Guil et al. proposed to cluster video frames by computing the similarity between neighboring frames and choose a keyframe from each cluster to represent it. However, the extracted Keyframes cannot represent the temporal information among frames. On the other hand, some methods were proposed to detect video shots and extract Keyframes from each shot to represent the video content.

Since there are some camera motion and object motion, the content within one shot will still has much variance. Then, we propose to detect video segments, which is an intermediate representation between video frames and video shots. Furthermore, matching two video sequences based on extracted Keyframes from the segments can meet the requirement of two videos being in different frame rates.

a) The Segmentation result using shot segmentation method



a) The Segmentation result using auto dual - threshold method



IV. MATCHING SIFT FEATURE POINT SETS BASED ON SVD

To better represent the local content of video frames, we choose SIFT descriptors to present the video sequences. On the other hand, since the number of SIFT feature points in video sequences is large, it thus exists high computational cost for video copy detection. Matching the SIFT feature points in two frames with the BBF-Tree method needs about several seconds. And the computational cost for matching the whole video sequences is high.

Thus, many methods, such as bag of features (BoFs) or visual word for video copy detection, locality sensitive hashing (LSH), and hierarchical indexing structure for efficient video retrieval, and so on, have been proposed for efficient video search. However, by using these indexing methods, the temporal information of the SIFT feature points in different frames will be lost. Thus, our motivation is to match the two SIFT feature sets in two video frames and make use of the temporal information of video frames.

The matrix singular value has the following characteristics: Characteristic 1: transposition and replacement invariance. That is to say, after transposition or row-column replacement operation of the matrix, its singular value remains unchanged.
This characteristic can be directly proved according to the definition of singular value and the characteristic of elementary matrix. Characteristic 2: energy concentricity. The matrix A can be approximately restructured by the first k largest singular values of A. It can be proved that the matrix corresponding to the first k largest singular values of A is the closest to matrix A under the Frobenius norm.

V. GRAPH-BASED VIDEO SEQUENCE MATCHING METHOD FOR VIDEO COPY DETECTION

Step 1: Segment the video frames and extract features of the Keyframes.

Step 2: Match the query video and target video.

Step 3: Generate the matching result graph according to the matching results.

Step 4: Search the longest path in the matching result graph.

Step 5: Output the result of detection.

VI. ADVANTAGES OF THE GRAPH-BASED VIDEO SEQUENCE MATCHING METHOD

Since the matching results based on visual features of the video frames do not incorporate the videos' temporal characteristics, the goal of the proposed graph-based video sequence matching method is to refine and order the segment matching results by incorporating the temporal information. The proposed methods demonstrate the following advantages:

1. It can automatically find optimal sequence matching result.
2. It can automatically remove the noise caused by visual feature matching.
3. It is adaptive to video frame rate change.
4. It can detect multiple copies existed in the detected video.

VII. CONCLUSIONS

This paper first analyzes different video copy types and the features used for copy detection. Based on the analysis, we use local feature of SIFT to describe video frames. Since the number of SIFT points extracted from a video is large, so the copy detection using SIFT features has high computational cost. Then, we use 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. Experimental results show that this method can obtain a better tradeoff between the detection effectiveness and time cost. Furthermore, for video sequence matching, we propose a graph-based video sequence matching method. It skilfully converts the video sequence matching result to a matching result graph. Thus, detecting the copy video becomes finding the longest path in the matching result graph. Experimental results show that the proposed graph-based video sequence matching method has several advantages:

1. The graph-based method can find the best matching sequence in many messy match results, which effectively excludes false "high similarity" noise and compensate the limited description of image low-level visual features.

2. The graph-based method takes fully into account the spatiotemporal characteristic of video sequence, and has high copy location accuracy.

3. The graph-based sequence matching method can automatically detect the discrete paths in the matching result graph. Thus, it can detect more than one copies.

4. Compared to exhaustive search method, graph based method can also reduce detection time.

REFERENCES

[1] X. Wu, C.-W. Ngo, A. Hauptmann, and H.-K. Tan, "Real-Time Near-Duplicate Elimination for Web Video Search with Content and Context," IEEE Trans. Multimedia, vol. 11, no. 2, pp. 196-207, Feb. 2009.

[2] A. Hampapur and R. Bolle, "Comparison of Distance Measures for Video Copy Detection," Proc. IEEE Int'l Conf. Multimedia and Expo (ICME), pp. 188-192, 2001.

[3] TRECVID 2008 Final List of Transformations, http://www-nlpir. nist.gov/projects/tv2008/active/copy.detection/final.cbcd. video.transformations.pdf, 2008.

[4] Final CBCD Evaluation Plan TRECVID 2008 (v1.3), http:// www-nlpir.nist.gov/projects/tv2008/Evaluation-cbcd-v1.3.htm, 2008.

[5] O. Ku¨cu¨ ktunc, M. Bastan, U. Gu¨du¨ kbay, and O ¨ . Ulusoy, "Video Copy Detection Using Multiple Visual Cues and MPEG-7 Descriptors," J. Visual Comm. Image Representation, vol. 21, pp. 838-849, 2010.

[6] M. Douze, H. Je´gou, and C. Schmid, "An Image-Based Approach to Video Copy Detection with Spatio-Temporal Post-Filtering," IEEE Trans. Multimedia, vol. 12, no. 4, pp. 257-266, June 2010.

[7] M. Douze, A. Gaidon, H. Jegou, M. Marszalek, and C. Schmid, TREC Video Retrieval Evaluation Notebook Papers and Slides: INRIALEAR's Video Copy Detection System,http://www.slpir.nist.gov/projects/tvpubs/tv8.papers/inria-lear.pdf, 2008.

[8] J. Law-To, C. Li, and A. Joly, "Video Copy Detection: A Comparative Study," Proc. ACM Int'l Conf. Image and Video Retrieval, pp. 371-378, July 2007.

[9] A. Hampapur, K. Hyun, and R. Bolle, "Comparison of Sequence Matching Techniques for Video Copy Detection," Proc. SPIE, Storage and Retrieval for Media Databases, vol. 4676, pp. 194-201, Jan. 2002.

[10] J. Yuan, L.-Y. Duan, Q. Tian, S. Ranganath, and C. Xu, "Fast and Robust Short Video Clip Search for Copy Detection," Proc. Pacific Rim Conf. Multimedia (PCM), 2004.

AUTHORS

First Author – A.PerumalRaja, received his B.E degree in computer science and engineering from Info institute of Engineering and also currently pursuing in M.E computer science and Engineering degree from Arasu Engineering college,kumbakonam, Email id: ashokperumalraja@gmail.com

Second Author – B.venkadesan, received his B.E degree in computer science and engineering from Anjalai Ammal Mahalingam Engineering college, kovilvenni-614403, Thiruvarur district, Tamilnadu. He is currently pursuing his M.E degree in computer science at Arasu engineering college kumbakonam.
Email id: coolboy.venkat@gmail.com

Third Author – R.Rajakumar, received his M.sc., M.Phil. M.Tech Degree in computer science and computer science and engineering from various recognized universities. He is currently pursuing PHD degree in computer science at Bharathiar University., Email: raja2012mtech@gmail.com