Video Copyright Detection Using High Level Objects in Video Clip

Abdul Rasheed Balouch[1], Ubaidullah alias Kashif[1], Kashif Gul Chachar[2], Maqsood Ali Solangi[1]

Abstract:

Latest advancements in online video databases have caused a huge violation of copyright material misuse. Usually a video clip having a proper copyright is available in online video databases like YouTube without permission of the owner. It remains available until the owner takes a notice and requests to the website manager to remove copyright material. The problem with this approach is that usually the copyright material is downloaded and watched illegally during the period of upload and subsequent removal on request of the owner. This study aims at presenting an automatic content based system to detect any copyright violation in online video clips. In this technique, a video clip is needed from original video that is used to query from online video to find out shot similarity based on high level objects like shapes.

Keywords: Digital Image Processing; Video Copyright Detection; Video Processing; Edge and Object Detection.

1. Introduction

Increasing number of videos day by day on web servers are violating the copyright rules. They are uploading pirated videos on their web server. It is extremely difficult to keep a watch on a huge number of videos which are being uploaded on different web servers on daily basis. It is practically impossible for video producers to check each and every video manually to find out any copyright violation. In few cases, these pirated videos are uploaded even before the original version released by legitimate producers. Therefore, many systems are proposed for detecting pirated video. Like watermark and copyright detection by content based retrieval system. In watermark approach, some extra information is added to detect the video like signature [1]. But still there is no robust approach of watermark [1, 2].Content Based Copy Detection(CBCD) is an alternative approach of watermark. In Content Base Copy Detection (CBCD) Frames are taken as image to extract the features from image, motion of frames, spatial temporal features of video, events of video like fighting scene and love scene [3]. But extracting of video features and then converting these features in histogram, or calculation of features entropy are also time cost tasks specially when there are huge low level features calculations in a query. In many cases, videos are not copied from original version; they are copied from a copy. It means that when a second copy is compared with the original, simple histogram matching may not produce accurate results [1].

The approach followed in this study exploits high level features, which include objects. A user poses a query to video database, then it will extract key frame from query clip and apply image analysis method to highlight high level objects, which are highlighted by Canny Edge detection algorithm shown in figure.2. In this research authors calculate position and size of each object for comparing it with video database objects.

---

[1] *Department of Computer Science, Sukkur Institute of Business Administration University Sukkur, Pakistan*

[2] *Department of Computer Science, IQRA University Karachi, Pakistan*

Corresponding email: abdulrasheed.ms@iba-suk.edu.pk

Figure 1: Image taken from a Video

- Man
- Books
- Flowers
- Telephone Set
- Shelf
- Papers
- Table

For accuracy of objects retrieved by our system, authors will use Precision recall methods. These methods are used in object retrieval to measure accuracy of retrieval system[4].

Our system is extension of Canny Edge Detection in which system will extract objects highlighted by canny algorithm. Figure 1 is taken from video shows many object in one frame. Figure.2 shows the edge detection result of Canny Edge Detection algorithm which is clearly depicting the edge of each object in frame. Figure.1 is taken from a video, in this figure there are so many objects, authors have enlisted these objects beside the picture. Through Canny Edge Detection, authors have highlighted all objects in figure 2.

Figure 2: Objects detection Using Canny algorithm

Authors calculate the position and size of each object individually. It is rare to have two scenes in a video where same objects appear at the same location and size; thus, proposal can be used to differentiate scenes and extract a query scene.

2. Related Work

Content based video copy detection by motion vector is proposed by Kasm Tademir in [1]. It calculates the histogram of motion vector. The author has computed mean value of magnitude of motion and mean value of phase motion vectors of each macro blocks. An advantage of this study is that it does not consider the features of video; it just calculates motion of video. So, it reduces computation cost [1]. Another approach of copyright detection is Watermark approach. It inserts some extra bits in a video. By extracting this extra information, authors could detect pirated video [3]. However, watermark approach is not still robust [1][2][3]. Watermark approach has two significant limitations. First, since

watermarks must be introduced into the original content before copies/duplicates are made, it cannot be applied to content which is already in circulation. Second, the degree of robustness is not adequate for some of the attacks that authors encounter frequently. In CBCD approach, media itself is watermark approach in which features are extracted from image are key frame of video to match video database [3] Another drawback of CBCD is that when you compute the video similarities from a copied video, the results may not be the same as they may be when compared with original video. Another approach used by Idnyk in[5] exploits temporal fingerprints based on shot boundaries of video sequence. The drawback of this technique is that it does not work well with the short video [5]. Oostveen presents hash function for identification of video in [6]. B. Coskun et al present two methods for video copyright detection both based on discrete cosine transform for video copy detection [6]. Hampapur and Bolle have enhanced the work of Bhat and Nyar in which they compare the video on the basis of motion, color and spatial temporal distribution of intensity [7]. Y. Li ET also proposed a method to detect video copy. The author used a binary signature involving color histogram. The main advantage was that the system was robust regarding to inserting signature which frequently cropped in TV production [8]. T.N. Shanmugham and Priya Rajendran presented a good approach of content based retrieval based on query clip in [4]. The approach could be used to detect pirated video in which authors have proposed that spilt video in a sequence of elementary shots and represents each shot as elementary frame, and calculates the frame description including motion, color, edge, feature and, texture. Mani Malekesmaeili has proposed an approach namely video copy detection using temporally informative representative images [8]. Basically, author enhanced another work of video image hashing technique which is applied on either each frame or selected key frame of a video sequence. But approach

never used temporal information in a video [4]. Maini has enhanced this approach and adds temporal as well as spatial information. Performance measured by a simple image hashing function on video database [5]. Earlier video matching method was reducing video sequence in small key frames [1][2][3]. Image sequence method used to match the key frame. In this approach when a shot is missed, the entire process of matching is also missed. They also ignore the temporal information [9]. A fast method of color histogram was discovered by many researchers. Yeh and Cheng have discovered a fast method that is 18 times faster than other sequence matching algorithm. They used and extended color method namely HSV color histogram [5]. Jolly, Frelicot and Buisson proposed a statistical similarity based approach, in which some interesting points are detected and then compared with Nearest Neighbor method. Interested points are extracted by Harris detector [6]. Law et al. proposed a video indexing method by using temporal contextual information which is extracted from interested point by voting function [7]. Another approach proposed by Deepak CR et al Query by video clip. They have proposed an algorithm which retrieves video from database by a query clip. Their results show that retrieval has high precision and comparatively low search time cost. They have used clustered approach for key frame matching [9]. Lienhart et al. used a mechanism to characterize key frame of the video by using color vector [8]. Another approach proposed by Deepak CR et al Query by video clip [9]. They have proposed an algorithm which retrieves video from database using query by clip. Their result shows that retrieval has high precision and searching query has low search time cost. They have used clustered approach for key frame matching [9]. Another efficient method is proposed by Kimiaki Shirahama and Kuniaki Uehara namely Query by shot. In this approach, they have classified features in high level objects and low level objects. In this study, they mainly focus on low level feature

color, region, and motion, high level object is just identified by them but not used further [10].

3. Methodlogy

The presented system extracts objects from the image. As shown in figure.3, each object in a frame is indicated by red circle. Here system extracts these objects from image. Figure.4 shows the result of extracted objects from image.

Figure 3: Objected detected in video

Figure 4: Detected objects and extracted objects from a video

Figure 5 depicts all the steps involved in our methodology. Rest of the paper defines each building block step by step.
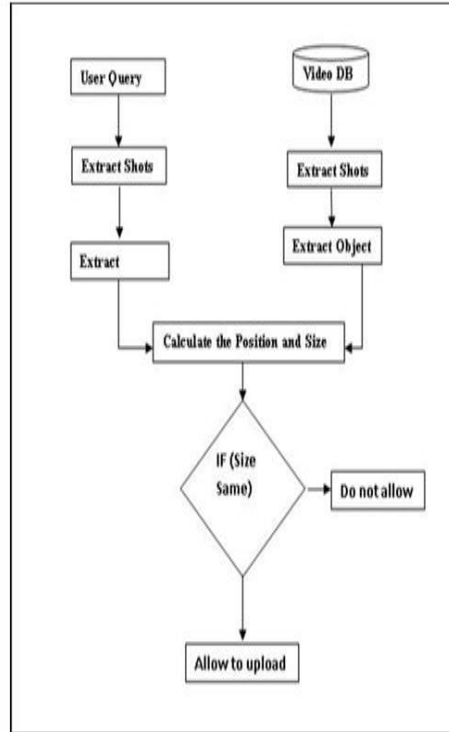

Figure 5: Showing all steps of Methodology

3.1. User Query

Authors define online video database V as,
$$V= \{v_1, v_2, v_3\ldots\ldots v_n\}$$
Where each v_i represents a video file in an online database V.

Authors further define a set of query Q as,
$$Q= \{q_1, q_2, q_3, q_4\ldots\ldots\ldots q_n\}$$
Where each qi represents a clip to be checked in online video database V.

In our approach, the set V represents all online video databases whereas each q in Q is submitted by video producer who want to check its video for piracy. Authors use following algorithm to determine if a video is illegally uploaded by a non-legitimated user.

Algorithm:
1. A set S containing all the shots in query clip q
2. A set V of online video database
3. For each vi in set V

4. For each si in S
   a. Determine shots in vi
   b. From each shot of si and vi, extract objects size and position
   c. if size and position of objects are found same in both clips
      i. Report video to be pirated
5. Continue to check next video

3.2. Extract Shot

Shot detection is primary step to analyze video [10].

Figure 6: Shots extracted from a video

Different algorithms are used to detect shots from video based on visual discontinuity along the time domain between two scenes of video [11]. Hard cut is most advanced algorithm to detect the shots from sequence of video frames [11, 12] .

In this research work authors have also used an existing approach of hard cut, an abrupt video shot detection based on color histogram used in [11] for finding the color histogram difference in frames authors used the expression to find the shot boundary of frames.

$$HistDif[i] = \sum_{j=1}^{M} |h_i(j) - h_{i-1}(j)|$$

Where h_i color histogram with M bins of frames *i* [11]. Figure 7 shows the step of shot detection. It compares current frame color histogram with the previous frame color histogram. If it finds a difference between these frames, it will detect as shot.
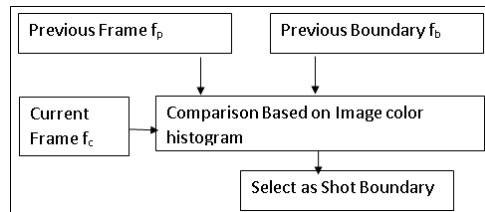

Figure 7: Frame comparison diagram

3.3. Extract Objects

Edge based object detection is a collection of homogenous edge which shows the continuity of image in same region when edge dissimilarity occurred in region shows that object edge is complete here and new object starts. For highlighting each object in image, authors have used existing approach of Canny Edge detection algorithm. This detects objects based on edge. Through canny expression of finding gradient for pixel which shows the color intensity of object [14], authors find the position of object. For the calculation of edge based intensity authors have applied a method 'Crop' which crops the segment of image highlighted by canny edge detection.
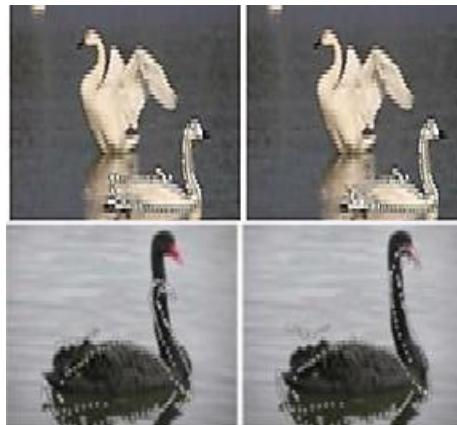
Figure 8: Objects highlighted by our System

$$|G| = \sum \; (Gx + Gy) \ldots\ldots\ldots\ldots\ldots\ldots(2)$$

Calculation of the total size of the object through following expression

$$S = X*Y \ldots\ldots\ldots\ldots\ldots\ldots\ldots.. \; (3)$$

Where x and y show the horizontal and vertical pixel of highlighted objects. The Same process will apply on test database to test the position and size of objects.

4. Comparision

At the stage of comparison, the system compares the size and position of each object against each object of video database. If the size and position are found to be equal, the system will not allow the uploading of the video on video database server. If the size and position are not matched from both sides the system will allow to upload the video on database server.

5. Conclusion and Future Work

In this paper, authors presented a robust way to detect the pirated videos based on Content Based Copy Detection (CPCD). The presented system extracts objects from image and calculates the position and size of objects and apply same technique on video database server. Compare the size and position of objects to test the video either it is pirated or not. In future work, authors aim to fully implement the system on a large scale video database server and perform further analysis on the system.

REFERENCES

[1]   K. Tasdemir, "Content based video copy detection using motion vectors," bilkent university, 2009.

[2]   J. Law-To, L. Chen, A. Joly, I. Laptev, O. Buisson, V. Gouet-Brunet, N. Boujemaa, and F. Stentiford, "Video copy detection: a comparative study," in Proceedings of the 6th ACM international conference on Image and video retrieval, 2007, pp. 371-378.

[3]   D. Zhang and S.-F. Chang, "Event detection in baseball video using superimposed caption recognition," in Proceedings of the tenth ACM international conference on Multimedia, 2002, pp. 315-318.

[4]   B. Taneva, M. Kacimi, and G. Weikum, "Gathering and ranking photos of named entities with high precision, high recall, and diversity," in Proceedings of the third ACM international conference on Web search and data mining, 2010, pp. 431-440.

[5]   P. Indyk, G. Iyengar, and N. Shivakumar, "Finding pirated video sequences on the internet," Tecnical Report, Stanford University, 1999.

[6]   J. C. Oostveen, T. Kalker, and J. Haitsma, "Visual hashing of digital video: applications and techniques," in International Symposium on Optical Science and Technology, 2001, pp. 121-131.

[7]   A. Hampapur and R. Bolle, "Feature based indexing for media tracking," in Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference on, 2000, pp. 1709-1712.

[8]   S. Poullot, O. Buisson, and M. Crucianu, "Z-grid-based probabilistic retrieval for scaling up content-based copy detection," in Proceedings of the 6th ACM international conference on Image and video retrieval, 2007, pp. 348-355.

[9]   L. Teodosio and W. Bender, "Salient video stills: Content and context

preserved," in Proceedings of the first ACM international conference on Multimedia, 1993, pp. 39-46.

[10]   A. F. Smeaton, P. Over, and A. R. Doherty, "Video shot boundary detection: Seven years of TRECVid activity," Computer Vision and Image Understanding, vol. 114, pp. 411-418, 2010.

[11]   J. Mas and G. Fernandez, "Video shot boundary detection based on color histogram," Notebook Papers TRECVID2003, Gaithersburg, Maryland, NIST, 2003.

[12]   S. M. Doudpota, S. Guha, and J. Baber, "Shot-Based Genre Identification in Musicals," in Wireless Networks and Computational Intelligence, ed: Springer, 2012, pp. 129-138.

[13]   J. Canny, "A computational approach to edge detection," Pattern Analysis and Machine Intelligence, IEEE Transactions on, pp. 679-698, 1986.

.