International Journal of Computer Engineering in Research Trends
Multidisciplinary, Open Access, Peer-Reviewed and fully refereed
Research Paper | Volume-5, Issue-7 ,2018 Regular Edition | E-ISSN: 2349-7084

Piracy Detection of Video Contents by Signature Matching Method

Aishwarya. M. Chavan[1*]

[1*]*Dept. of Computer Science and Engineering, DKTE's TEI (An Autonomous Institute), Ichalkaranji, India.*

e-mail: chavanaishu@gmail.com

Available online at: http://www.ijcert.org

Received: 11/July/2018, Revised: 12/July/2018, Accepted: 14/July/2018, Published: 18/July/2018

Abstract:- Security has become the primary concern to protect the critical and sensitive information such as multimedia. A problem faced by nowadays is that multimedia contents are getting pirated on a large scale. These contents need to be protected from getting duplicated. The primary goal is to detect the duplication of both 2D and 3D video contents. Essential components in identifying the piracy are the generation of unique signatures and a matching engine to match them. The system detects pirated multimedia contents.

Keywords: 3D SBS, SUFR, video signatures, KD Tree.

1. Introduction

Due to progress in the technological field,, people communicate through the Internet and share their contents. Due to advanced recording and processing of multimedia contents and due to the availability of free online hosting sites it became easy to duplicate the copyright material. There are the illegal copying and distribution of substances over the social media, so the content owners need to protect their contents from getting pirated. Finding unauthorized copies over the Internet is a complicated and computationally expensive operation, because of a large volume of multimedia contents available on the Internet. Two fundamental approaches to detect video copies are Watermarking and Content-Based Copy Detection.

With the exponential growth of multimedia content, there need to protect the content. On video sharing sites a large number of videos are uploaded due to which issue like copyright arises. Digital media should be safeguarded from getting into wrong hands. So different options are present in front of software publishers to protect contents from thieves.

Watermarking approach used before were not suitable for protecting piracy. For online videos watermarking approach was not practical as they can be easily removed and replaced with another watermark. User downloads the video from different online hosting sites and makes transformations such as blurring, cropping, and text insertion. For comparing original video with newly uploaded video, it requires more memory and needs more processing time in the existing system. So a digital signature is to be created for 3D videos. A digital name is a scheme is used in ensuring the authenticity of digital message or documents.

Content-Based Copy Detection (CBCD) is another approach for detecting multimedia content [1]. To create and match signatures different methods were proposed. These methods are classified into four categories: spatial, temporal, color and transform-domain. Spatial signatures are mainly the block-based and are the most widely used. The drawback of the spatial signature is the lack of capacity to recover quickly against great geometric transformations. As compared to spatial signatures, Temporal and color signatures are less robust and they are used to enhance spatial signatures. Transform-domain signatures are difficult to carry out and not widely used in practice. Youtube Content ID [2], Mark Monitor [3] used a signature method for media protection, and these are modern examples.

The system can help us to protect multimedia contents like 2D and 3D videos, images, audio clips, songs. This system is to be developed on public/ or private cloud. This system has two types of components (i) method to create signatures of 3D videos (ii) Matching engine to check these signatures. The signature method creates a robust, and representative signature of 3D videos that capture the depth signals in these videos and the matching engine achieves high scalability, and it is designed to support different multimedia objects. By using 3D signature methods, we handle the 3D videos.

The remainder of this paper is organized as follows. In Section 2, we present the related work about Piracy protection system. Section 3 explains the methodology; Section 4 is about Result and discussion. Conclusion and Future work are drawn in Section 5.

2. Related Work

Earlier methods [4, 5, 6, 7] extracted the local features like SIFT to deal with copping problem and occlusions. Here dynamic behaviour on the local descriptor of interest points is extracted under the method named video content indexing method. Video copies can be found by considering local salient points and local fingerprints [4, 6, 7, 8] by using complex indexing scheme.

Joly et al. [6] proposed a technique in which feature regions were selected by considering the probability and was based on the distribution of feature distortions. After that compatible geometrical matches were used to retrieve the video copy. Law-To el al. took video sequence and built the trajectories of local points and carrying out its indexing task. A generalised Radon transform [7] is used to extract local descriptors of the corner points.

Douze et al. [8] Local feature indexing method are used which is robust to video transformations. In [9], to retrieve the near about duplicate videos a multiple feature hashing methods is used. To evaluate the performance of video copy detection many test beds are built.

Juan et al. [10] Here Video Copy Detection System (VCD) are developed which detects subsystems each of which identifies the duplicate videos by consisting different features like visual, audio etc. and these results are combined to get the final result. Jiang et al. [11] two different techniques are employed namely multimodal feature representation which exploits audio, local and global features, and other is Temporal Pyramid Matching (TPM) that fuses the frame-level similarity results into sequence-level matching results.

Mohamed Hefeeda et al. [12] the main idea of the project is taken from this paper. The focus of the article is on the other approach for protecting multimedia content which is content-based copy detection (CBCD). Algorithms such as SURF is used to detect the features which are unique and to illumination changes and rotation.

3. Methodology

The image of a particular scene can be taken from different viewpoints, even these images can be subjected to transformations like rotation, noise etc. So these images appear different. So to find the correspondence and similarity between these images is a challenging task.

To avoid duplication of images, there are different ways to detect them. Various algorithms are used to detect the copied images. Algorithms like SIFT, SURF, HOG can be used to detect features in an image which in turn are essential to identify duplicated copies. Object recognition, visual localization, image indexing includes efficient image feature detection and matching. Even if the image is taken at different instances the features extracted remains the same. To eliminate the computational complexity along with time-consuming parameters the SURF algorithm is used. Two different views of an image can be matched by extracting features from an image for reliable matching those images.

System Architecture:

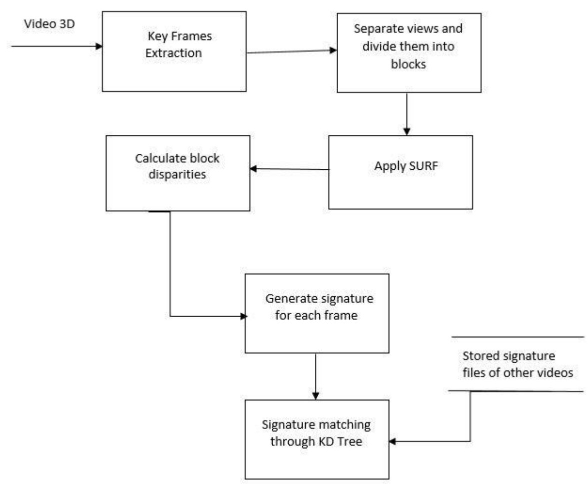

Figure: 3.1. System architecture for video piracy detection system

Figure 3.1. Shows the system architecture which consists of two main components, the process to generate the signatures and method to match those signatures.

A. Frame Extraction / Separation of Right And Left views of 3D frames :

3D videos are of three types Side-by-Side (SBS), Bottom/Up, and Anaglyph. Among that our experiment is conducted on SBS 3D videos only. Standard video is having the frame rate of 25 to 30 fps. Separate all scene change frames from a given video clip. For 3D SBS type of videos separate their left and right views of each frame. Both aspects are having the 15-degree difference in them. That is left view is one or two seconds forward than the correct perspective.

B. Creating Blocks of views and SURF (Speed Up Robust Features) Feature Extraction:

After separating both the opinions we divide each of the left and right view of each extracted frame into blocks (chunks) of 3X3 or 5X5. Then apply the surf algorithm on each corresponding blocks of both views. Similar blocks in left and right image are compared. Take the single descriptor in the left image and find the closest descriptor in the right image. Then compute the Euclidean distance between the descriptor in the left and right image. And the descriptors with minimum distance are taken the as good match.

Feature detection process is to be carried out to find interesting points to make correspondence between images. Visual Descriptors or Image descriptors are the descriptors of visual features of the contents in images, videos, or algorithms that produce such descriptors. They describe elementary characteristics such as the shape, color, texture or motion among others. Feature detection process focuses only on those parts on an image that provides some unique information. Matching feature descriptor can be based on Euclidean distance.

Distinctive features from an image are searched in two phases. First features are extracted from different locations in an image such as edges, blobs etc. These detected features are repeatable. Neighbourhood region of each element is considered and descriptor are calculated for each region. Extracting robust features and their descriptor is a computationally expensive process. To reduce the number of comparisons only scene change images are taken.

C. Disparity calculation and signature generation:

Block disparity is calculated between each block in left image and its corresponding block in the right image. The disparity of a single descriptor the following formula gives me:

$$\sqrt{((x_i - x_j)/W_b)^2 + ((y_i + y_j)/H_b)^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Equation (1) calculates disparity of blocks [12]. Where the position of descriptor i is (x_i, y_i) in left image and (x_j , y_j) is the position of the corresponding descriptor in the right image. Both the left and right images are divided into an equal number of blocks. Blocks can be of any regular size and shape. Divide each image into an N*M number of blocks. In our experiment, we divide the image into 3*3 or 5*5 rectangular blocks. And the disparity of the block is normalized by Width W and Height H of the image. If the block and the corresponding block in the right image do not have any descriptors, the disparity is 0. Signature of two corresponding images is (SB1, SB2,…, SBn*m). Where B is a block and S is its disparity. This signature is stored. Compute signatures for different videos and store them.

D. KD Tree construction:

KD Tree implements are matching engine. KD Tree is a binary tree in which every node is a k dimensional point. It is space partitioning and directing tree. Matching technique consists of two steps creating an index for each data point and matching query against query data points. Every non-leaf node is a hyper splitting plane that divides the plane into two parts. Points to the left of the tree represent the left sub-tree and aspects to the right of the tree represents right sub-tree[12]. We find the perfect nearest neighbour of the signature in the query in the tree.

4. Results and Discussion

Experiments are carried out on videos which are downloaded from YouTube and other sites. Videos are transformed by using basic video editing software's. Transformations like dropping the clip and attaching clips of 1 and 1.30 minutes. Another different video clip is attached to the original video to transform the video. Similarly, a small clip is dropped from the original video and is cross-checked against the original video. Transformations such as a change in the format were also carried out. Experiments are carried out on total 49 videos among them 45 are detected aspirated by the system. The overall accuracy for different types of transformations is 89.64% for 3D videos.

Experiments are carried out on videos which are downloaded from YouTube and other sites. Videos are transformed by using basic video editing software's. Transformations like dropping the clip and attaching clips of 1 and 1.30 minutes. Another different video clip is attached to the original video to transform the video. Similarly, a small clip is dropped from the original video and is cross-checked against the original video. Transformations such as change in format were also carried out. Experiments are carried out on total 49 videos among them 45 are detected aspirated by the system. The overall accuracy for different types of transformations is 89.64% for 3D videos.

Transformation		**Number of Videos**	**Detected by PDVCSM**	**Precision**
Drop clip	1 min.	14	14	**100%**
Drop clip	1:30 min	16	14	87.5%
Attach another Clip	1 min	7	7	**100%**
Attach another Clip	1:30 min	8	7	87.5%
File format change	Mp4 to WMV	4	3	75%
Average		**49**	**45**	**89.64%**

Figure 4.1. Precision table for 3D videos

Transformation	**Number of Videos**	**Detected by PDVCSM**	**Precision**
Drop clip	3	3	**100%**
Attach another Clip	8	8	**100%**
Average	**11**	**11**	**100%**

Figure 4.2. Precision table for 2D videos

The comparison is done through signature matching method. The accuracy of the signature is improved by increasing the length of the signature. So even small parts of the frames are detected and matched correctly. For 3D videos, results vary for attaching and dropping the clip. The overall average is taken of the videos detected as pirated. Attaching 1 min clip hardly extracts the 15 to 20 keyframes from the video clip attached. So it's easy to identify the duplication, as 60% of the signature match is considered. If the timing of attaching the clip increases it becomes difficult to detect the duplication as a number of frames extracted are more and different.

For 2D videos 1 to many frames are compared to high-level comparison is achieved. Therefore the accuracy of matching is increased. 100% precision is achieved for 2D videos. Experiments are carried out on total 11 animated type of videos downloaded from YouTube. The maximum length of downloaded 2D videos is 2 minute. Original videos and transformed videos are of the same dimensions. Here diagonal matrix is created to record the matching frames. The most extended matched sequence on diagonal to the length of the diagonal ratio is considered to detect the duplication of 2D videos. The threshold of 0.6 is taken into account. If the count of frame match exceeds the threshold the video is declared is duplicate.

5. Conclusion and Future Work

This system demonstrates how the duplication of multimedia (like videos and images) contents can be detected. The approach used to identify the substances of images provides better performance regarding computational time and accuracy. The system takes the input of videos which are of the same dimensions. For a 2D video detection full frame to frame, matching is carried out between the query and the reference video. Whereas for 3D videos a different approach of signature matching is used to match the frames. The experimental results show high accuracy is achieved for 2D and 3D videos. Future scope of the system lies in considering the audio part along with video part. The system generated signatures for SBS type of 3D videos. So it can be enhanced to generate signatures for advanced Multiview plus depth type of 3D videos.

6. References

[1] J. Lu. Video signatures for copy identification: From research to industry applications. In IS&T/SPIE Electronic Imaging, volume 7254 of SPIE Proceedings, pages 725402–725402. International Society for Optics and Photonics, SPIE, 2009.

[2] S. Ioffe. Full-length video signature. Google Inc., July 24 2012. US Patent 8229219.

[3] E. Metois, M. Shull, and J. Wolosewicz. Detecting online abuse in images. Markmonitor Inc., Apr. 12 2011. US Patent7925044.

[4] J. Law-To, O. Buisson, V. Gouet-Brunet, and N. Boujemaa. Robust voting algorithm based on labels of behaviour for video copy detection. In ACM Multimedia, 2006.

[5] W. L. Zhao, C.-W. Ngo, H.-K. Tan, and X. Wu. Near-duplicate keyframe identification with interest point matching and pattern learning. Multimedia, IEEE Trans. on, 2007.

[6] A. Joly, O. Buisson, and C. Frelicot. Content-based copy retrieval using distortion-based probabilistic similarity search. Multimedia, IEEE Transactions on, 2007.

[7] E. Maani, S. Tsaftaris, and A. Katsaggelos. Local feature extraction for video copy detection in a database. In ICIP, 2008.

[8] M. Douze, H. Jegou, and C. Schmid. An image-based approach to video copy detection with spatio-temporal postfiltering. Multimedia, IEEE Transactions on, 2010.

[9] J. Song, Y. Yang, Z. Huang, H. T. Shen, and R. Hong. Multiple feature hashing for real-time large scale near-duplicate video retrieval. In ACM Multimedia, 2011.

[10] J. M. Barrios, B. Bustos, and X. Anguera. Combining features at search time: Prisma at video copy detection task. In Proc. TRECVID, 2011.

[11] M. Jiang, S. Fang, Y. Tian, T. Huang, and W. Gao. Pkuidm @ trecvid 2011 cbcd: Content-based copy detection with cascade of multimodal features and temporal pyramid matching. In Proc. TRECVID, 2011.

[12] Mohamed Hefeeda, Senior Member, IEEE, Tarek ElGamal, Kiana Calagari, and Ahmed Abdelsadek Cloud-based Multimedia Content Protection System IEEE Transactions on Multimedia,2015.

Authors Profile

Miss. Aishwarya M. Chavan pursued Bachelor of Engineering from Shivaji University, Kolhapur in the year 2016. She is currently pursuing Master of Technology form DKTE's TEI (An Autonomous Institute), Ichalkaranji, India. Her work focuses on image processing and multimedia.

