International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2019): 7.583

CBCD Methods in Video Copy Detection

Jan Mary Thomas

Computer Science and Engineering Department, KTU University, Musaliar College of Engineering and Technology, Pathanamthitta, Kerala, India

Abstract: ***Due to the overwhelming addition of videos in the internet today, the authenticity and authority of the video owner is not protected. Several algorithms have been proposed which works on the content based copy detection (CBCD) scheme in detection and extraction phases of the video data. In this paper, the different extraction and searching algorithm in the detection and copy-searching phase are compared.***

Keywords: Content-based fingerprinting, multimedia duplicate detection, multimedia fingerprinting, robust video hashing, video copy detection, video copy retrieval

1. Introduction

Content based copy detection works by identifying copied videos by contrasting it with original copies. The technique is implemented on the basis of a content based signature – extracted component from the video itself. This is known as digital video fingerprinting. Each fingerprint extracted which corresponds to one multimedia document, is compared with fingerprints in the database. The extracted fingerprint should be distinguishable, robust and efficient.

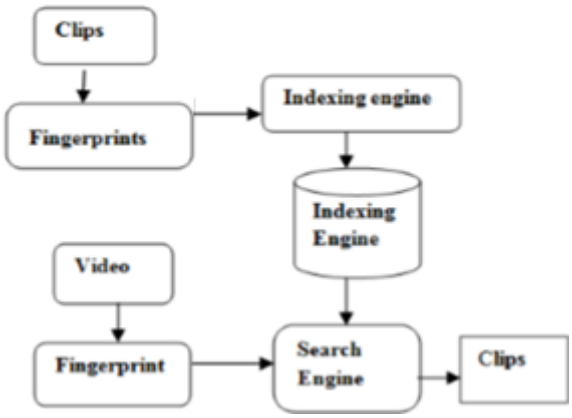


Figure 1: Video Copy detection based on the content

CBCD methods outperforms the watermarking methods outperforms the watermarking methods in the field of copyright infringement.

Image processing is a method to convert an image into digital form and perform some operations on it, in order to get an enhanced image or to extract some useful information from it. It is a type of signal dispensation in which input is image, like video frame or photograph and output may be image or characteristics associated with that image. Usually Image Processing system includes treating images as two dimensional signals while applying already set signal processing methods to them.

The two types of methods used for Image Processing are Analog and Digital Image Processing. Analog or visual techniques of image processing can be used for the hard copies like printouts and photographs. Image analysts use various fundamentals of interpretation while using these visual techniques. The image processing is not just confined to area that has to be studied but on knowledge of analyst. Association is another important tool in image processing through visual techniques. So analysts apply a combination of personal knowledge and collateral data to image processing.

Digital Processing techniques help in manipulation of the digital images by using computers. As raw data from imaging sensors from satellite platform contains deficiencies. To get over such flaws and to get originality of information, it has to undergo various phases of processing. The three general phases that all types of data have to undergo while using digital technique are Pre- processing, enhancement and display, information extraction.

2. General Architecture

2.1 Fingerprint –Types

Fingerprint is a signature having the following properties:

1) Fingerprints must be small in dimension and capture all perceptually important video-related information crucial for identifying it.
2) Fingerprints of two different videos must also be significantly different.
3) Matching two fingerprints should be enough to declare the corresponding videos as being the same. The major set of fingerprints are as follows
   a) ColorSpace Based Fingerprints: These are obtained from the histograms of the colors in distinct regions within a particular time /space in the video and these are not relevant to black and white videos
   b) Temporal Fingerprints: These are acquired from the properties of a video sequence over time. These features perform well for only long sequence of video.
   c) Spatial Fingerprints: They are features derived from each and every frame or from a key frame. More the spatial information, stronger the method will be
   d) Spatio-temporal fingerprinting: Fingerprint features with both spatial and temporal information

2.2 Content-based copy detection (CBCD)

Content-based copy detection (CBCD) The main idea of CBCD is that the media visually contains enough information for detecting copies. The copy detection is

considered as video similarity detection by using the visual similarities of video clips. Multimedia fingerprinting is also known as robust hashing works on the basis of CBCD. Signature is extracted from the trained media and stored in a database. The above procedure is repeated for the extraction of signature for query media which is differentiated with the trained media for similarity, it is stored in a database.

3. Existing System

Many of the existing algorithms for video fingerprinting has focused on the extraction of robust and discriminant features from a video. They also address the search time of the proposed fingerprinting algorithm. They propose two fast search methods. They divide each fingerprint into small non-overlapping blocks of bits. They call these small blocks words. Words are then used to create an inverted file from the fingerprints of database. All fingerprints have equal lengths, so the inverted file can be represented as a table of size, where is the number of words in a fingerprint of length. They propose another similarity search algorithm for binary fingerprints. The main idea is to use clustering to reduce the number of queries that are examined within the database. Their final proposed fingerprinting system, i.e., the TIRI-DCT method along with the search algorithm specifically developed for it introduces a fingerprinting system. By using a secret key, they show that the security of the fingerprinting system can be increased. They also articulate the trade-off between security and robustness. This system also intend to improve the robustness of the global features to large geometric attacks, and further reduce the required search time, while maintaining acceptable performance. The optimization of the system is not efficient. The overall system performance is decreased for large databases.

4. Proposed System

The proposed system introduces a product quantization-based approach for approximate nearest neighbor search. The idea is to decompose the space into a Cartesian product of low-dimensional subspaces and to quantize each subspace separately. A vector is represented by a short code composed of its subspace quantization indices. The Euclidean distance between two vectors can beefficiently estimated from their codes. An asymmetric version increases precision, as it computes the approximate distance between avector and a code. The goal is to estimate distances using vector-to-centroid distances, i.e., the query vector is not quantized; codes are assigned to the database vectors only. This reduces the quantization noise and subsequently improves the search quality. To obtain precise distances, the quantization error must be limited. Nearest neighbor search depends on the distances between the query vector and the database vectors or, equivalently, the squared distances. The method introduced in this section compares the vectors based on their quantization indices. First how the product quantizer properties are used to compute the distances. Then provide a statistical bound on the distance estimation error, and an estimator for the squared Euclidean distance is proposed.

4.1 Advantages of Proposed System

This system introduce the robustness of the global features to large geometric attacks, and the required search time is efficient, while maintaining acceptable performance;

- The compact coding scheme provides an accurate approximation of the Euclidean distance.
- The optimization of the process is efficient

5. Problem Definition

The spatio-temporal interest-point descriptors used to generate the fingerprints, a method for forming temporally informative representative images (TIRIs) from a video sequence. As a TIRI contains spatial and temporal information of a short segment of a video sequence, the spatial feature extracted from a TIRI would also contain temporal information. Propose two fast search methods. The first is a generalization of and the second is based on a novel approach involving clustering of the fingerprints in the database. The TIRI-DCT method specifically developed for it introduces a fingerprinting system that is robust, discriminant, and fast.

Initial module of the proposed approach, describes the pre-processing of the input video. First the input video is processed to change the rate and then change the pixel size of the video. These resized video is again divided into short segments. Final step in the pre-processing approach of this method is to generate the fingerprints from the segments obtained from the previous steps. extraction process using the TIRI-DCT transform is used for the capture of the temporal information in a video. In this TIRI transform method there are three different weight factors such as constant, linear and exponential and generate the images that best capture the motion. In inverted file based similarity search, to find a query fingerprint in the database, first the fingerprint is divided into words. The query is then compared to all the fingerprints that start with the same word. The indices of these fingerprints are found from the corresponding entry in the first column of the inverted file table. The Hamming distance between these fingerprints and the query is then calculated.

6. Overview of the System

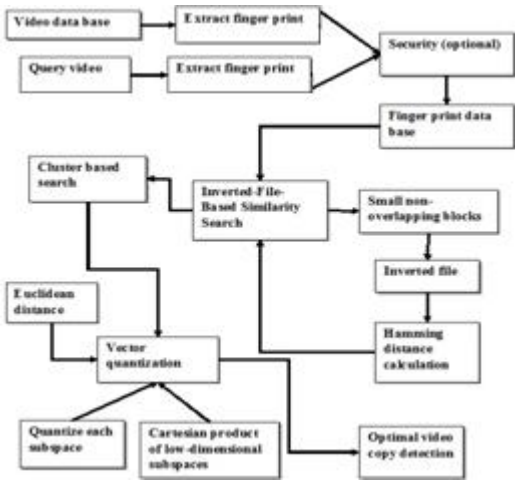


Figure 5.2.1: Overview of System

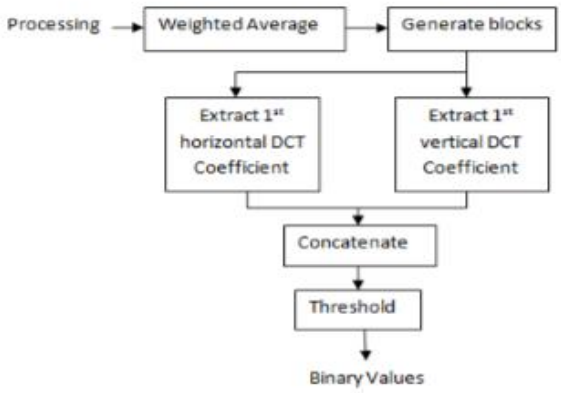
6.1 Modules Description

6.1.1Preprocessing

In pre-processing stage before extracting the fingerprints, we pre-process the video signals. The input video is divided into frames then these frames are converted into blocks. These blocks are converted into short segments. From the segments we generate the finger prints. After pre-processing, the video frames are divided into overlapping segments of fixed-length, each containing frames. The fingerprinting algorithms are applied to these segments. The amount of overlapping is experimentally chosen to be 50%. Overlapping reduces the sensitivity of the fingerprints to the "synchronization problem".

6.1.2 TIRI –DCT process

A method for forming temporally informative representative images (TIRIs) from a video sequence is implemented here. As a TIRI contains spatial and temporal information of a short segment of a video sequence, the spatial feature extracted from a TIRI would also contain temporal information. Based on TIRIs, we have proposed an efficient fingerprinting algorithm (TIRI-DCT) and compared it. After preprocessing first we calculate the weighted average then generate TIRI from each segment. Segment each TIRI into overlapping blocks. The first horizontal and the first vertical DCT coefficients (features) are then extracted from each block. The values of the features from all the blocks are concatenated to form the feature vector. Each feature is then compared to a threshold (which is the median value of the feature vector) and a binary fingerprint is generated.



6.1.3 Inverted-File-Based Similarity Search

This search method is based on the idea that for two fingerprints which are similar enough to be considered as matches, the probability of an exact match between smaller sub-blocks of those fingerprints is high. We divide each fingerprint into small non-overlapping blocks of bits. We call these small blocks words. Words are then used to create an inverted file from the fingerprints of database. All fingerprints have equal lengths, so the inverted file can be represented as a table of size.. The horizontal dimension of this table refers to the position of a word inside a fingerprint, and the vertical direction corresponds to possible values of the word. To generate this table, we start with the first word of each fingerprint, and add the index of the fingerprint to the entry in the first column corresponding to the value of this word. We continue this process for all the words in each fingerprint and all the columns in the inverted file table.

6.1.4 Cluster-Based Similarity Search

This is a similarity search algorithm for binary fingerprints. The main idea is to use clustering to reduce the number of queries that are examined within the database. By assigning each fingerprint to one and only one cluster, the fingerprints in the database will be clustered into non-overlapping groups. To do so, a centroid is chosen for each cluster, termed the cluster head. A fingerprint will be assigned to cluster if it is closest to this cluster's head. To determine if a query fingerprint matches a fingerprint in the database, the cluster head closest to the query is found. All the fingerprints (of the videos in the database) belonging to this cluster are then searched to find a match, i.e., the one which has the minimum Hamming distance from the query. If a match is not found, the cluster that is the second closest to the query is examined. This process continues until a match is found or the farthest cluster is examined.

6.1.5 Vector quantization for searching process

Here short codes using quantization are constructed. This reduces the quantization noise and subsequently improves the search quality. To obtain precise distances, the quantization error must be limited. Quantization is a destructive process which has been extensively studied in information theory. Its purpose is to reduce the cardinality of the representation space, in particular when the input data is real-valued. Formally, a quantizer is a function q mapping a D-dimensional vector. Nearest neighbor search depends on the distances between the query vector and the database vectors or, equivalently, the squared distances. The method introduced in this section compares the vectors based on their quantization indices. First e how the product quantizer properties are used to compute the distances is explained. Then the statistical bound on the distance estimation error is provided, and are fined estimator for the squared Euclidean distance is proposed.

6.1.6 Performance evaluation

The proposed approaches are illustrated and evaluated to compare the performance of all the approaches. Here the a the proposed scheme in terms of error in matching and brightness is analyzed. Based on the measurements and the results from the experiment show the proposed approach works better than the other existing systems.

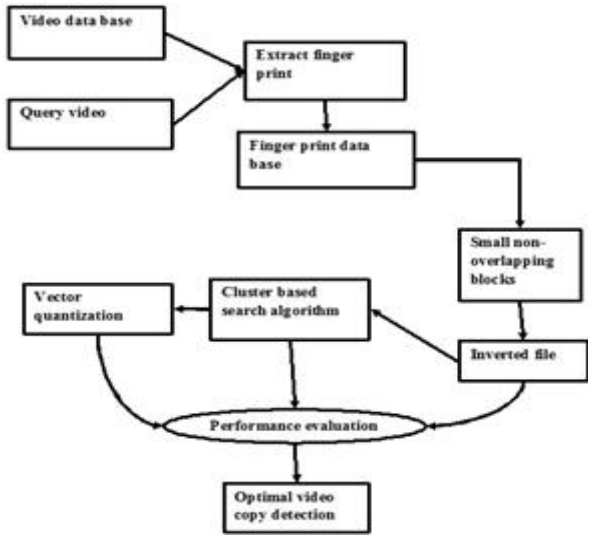


Figure 5.4.1: Data Flow Diagram

7. Input Output Design

Copies of the same video with different frame sizes and frame rates usually exist in the same video database. A fingerprinting algorithm should be robust to changes in the frame size as well as the frame rate. Down-sampling can increase the robustness of a fingerprinting algorithm to these changes. Each video is down-sampled both in time and space. Prior to down-sampling, a Gaussian smoothing filter is applied in both domains to preventaliasing.

Initial module of the proposed approach, describes the pre-processing of the input video. First the input video is processed to change the rate and then change the pixel size of the video. These resized video is again divided into short segments. Final step in the pre-processing approach of this method is to generate the fingerprints from the segments obtained from the previous steps. Thus by pre-processing approaches in this module results the fingerprints for each segments in the video data.

Features are derived by applying a 2D-DCT on overlapping blocks of size from each TIRI. The first horizontal and the first vertical DCT coefficients are then extracted from each block. The value of the features from all the blocks are concatenated to form the feature vector. Each feature is then compared to a threshold and a binary fingerprint is generated. For TIRI-DCT, all features are in the same frequency range, and binarization based on a common threshold.

Finally the performance is evaluated for the proposed approach for this problem in this module. For evaluation process a set of videos in which the approaches are applied and the performance metrics used here for the comparison are precision, recall and F-score are used. The experimental results show that the enhanced approach is performing better than the existing approaches.

8. Conclusion and Future Enhancements

Conclusion of this project focuses on implement of construct short codes using quantization. This reduces the quantization noise and subsequently improves the search quality. To obtain precise distances, the quantization error must be limited. Quantization is a destructive process which has been extensively studied in information theory. Its purpose is to reduce the cardinality of the representation space, in particular when the input data is real-valued. The idea is to decompose the space into a Cartesian product of low-dimensional subspaces and to quantize each subspace separately. A vector is represented by a short code composed of its subspace quantization indices. The Euclidean distance between two vectors can be efficiently estimated from their codes. An asymmetric version increases precision, as it computes the approximate distance between a vector and a code. As part of the future work, I intend to improve the robustness of the global features to large geometric attacks, and further reduce the required search time, while maintaining acceptable performance. Further the TIRI-DWT algorithm can be implemented and improved instead of the TIRI-DCT.

References

[1] J. Oostveen, T. Kalker, and J. Haitsma, “Feature extraction and a database strategy for video fingerprinting,” in Proc. Int. Conf. Recent Advances in Visual Information Systems (VISUAL), London, U.K., 2002, pp. 117–128, Springer-Verlag.
[2] J. Lu, “Video fingerprinting for copy identification: From research to industry applications,” E. J. Delp, J. Dittmann, N. D. Memon, and P. W. Wong, Eds., SPIE vol. 7254, no. 1, 2009, p.725402[Online].Available: http://link.aip.org/link/?PSI/7254/725402/1
[3] 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 Proc. ACM Int. Conf. Image and Video Retrieval, New York, NY, 2007, pp. 371–378, ACM.
[4] A. Hampapur and R. M. Bolle, Videogrep: Video copy detection using inverted file indices IBM Research Division Thomas. J. Watson Research Center, Tech. Rep., 2001.
[5] L. Chen and F. W. M. Stentiford, “Video sequence matching based on temporal ordinal measurement,” Pattern Recogn. Lett., vol. 29, no. 13, pp. 1824–1831, 2008.
[6] R. Radhakrishnan and C. Bauer, “Content-based video signatures based on projections of difference images,” in Proc. MMSP, Oct. 2007, pp. 341–344.
[7] C. De Roover, C. De Vleeschouwer, F. Lefebvre, and B. Macq, “Robust video hashing based on radial projections of key frames,” IEEE Trans. Signal Process., vol. 53, no. 10, pp. 4020–4037, Oct. 2005.
[8] S. Lee and C. Yoo, “Robust video fingerprinting for content-based video identification,” IEEE Trans. Circuits Syst. Video Technol., vol. 18, no. 7, pp. 983–988, Jul. 2008.
[9] X. Su, T. Huang, and W. Gao, “Robust video fingerprinting based on visual attention regions,” in Proc. ICASSP, Washington, DC, 2009, pp. 1525–1528, IEEE Computer Society.
[10] A. Joly, O. Buisson, and C. Frelicot, “Content-based copy retrieval using distortion-based probabilistic similarity search,” IEEE Trans. Multimedia, vol. 9, no. 2, pp. 293–306, Feb. 2007.