Video Copy Detection using TIRI based Texture Features

S.Divya Almeida
PG Student, CSE
Dr. Sivanthi Aditanar College of Engineering,
Tiruchendur-628215
Tamilnadu, India

G.Wiselin Jiji
Professor, CSE
Dr. Sivanthi Aditanar College of Engineering,
Tiruchendur-628215
Tamilnadu, India

ABSTRACT
Identification and verification of a video clip based on its content fingerprints to find applications in video browsing, large database set and it also take a lesser amount of time to find matches of video clip in the copyright applications. A fingerprint extraction algorithm is used along with a fast approximate search algorithm in order to simplify and speedern with the process. The fingerprint extraction algorithm extracts compact content-based signatures from special images denoted by temporally informative representative images, constructed from the video. Each such image represents a short segment of the video and contains temporal as well as spatial information about the video segment. To find whether a query video or a part of it is copied from a video in a video database, the fingerprints of all the videos in the database are extracted and stored in advance. The search algorithm searches the stored fingerprints to find close enough matches for the fingerprints of the query video. The fast approximate search algorithm facilitates the online application of the system to a large video database of millions of fingerprints and if a match occurs it is identified in a matter of seconds.

General Terms
Video Identification, Video Retrieval

Keywords:
Content-based Video Copy Detection, Illegal Video Copy Detection, Video Copy Retrieval.

1. INTRODUCTION
Huge and increasing amount of videos broadcast through networks has raised the need of automatic video copy detection for copyright protection. In Today's world there is widespread video
violation and it calls for the need of fast and accurate copy-detection algorithms. Recent developments in multimedia technology introduced content-based copy detection as a new research field alternative to the watermarking approach for identification of video sequences. Multimedia fingerprinting also known as robust hashing has been recently proposed for this purpose. A fingerprint is a content-based signature that is derived
from a video. So the evaluation process is done based on the signature of video content.

2. TYPES OF FINGERPRINTS
Video fingerprint extraction algorithms can be classified into five groups based on the features they extract- color-space-based, temporal, spatial, Spatio-temporal and texture. In color-space based method color information from the histogram of the color is extracted. But this feature alone is not suited for video copy detection. In temporal, we get characteristics of the video over the time. In Spatial fingerprints are features derived from each frame or from a key frame. Spatial fingerprints can be further subdivided into global and local fingerprints. Texture fingerprints refers to surface characteristics and appearance of an object given by the size, shape, density, arrangement, proportion of its elementary parts.

3. SUMMARY OF THE PAPER
One drawback of spatial fingerprints is their inability to capture the video's temporal information. So we have extracted Spatio-temporal features. Spatio-temporal fingerprints that contain both spatial and temporal information about the video are thus expected to perform better than fingerprints that use only spatial or temporal fingerprints. In [3], proposed 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. Spatio-temporal features have also been used in the literature [2]

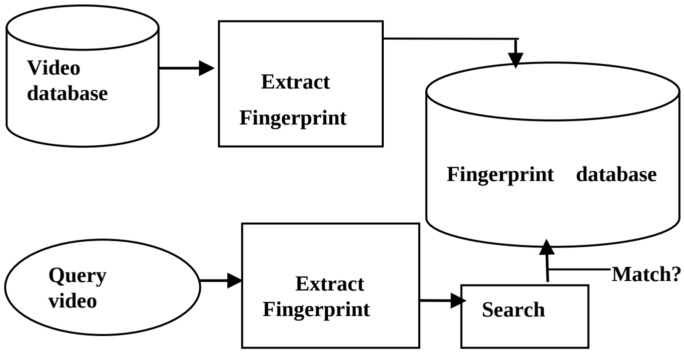


Figure 1 shows the overall structure of the proposed system [5].

Based on TIRIs, in [4], proposed an efficient fingerprinting algorithm (TIRI-DCT).In this paper we adopted GLCM based texture features along with fingerprinting algorithm TIRI-DCT.

4. PREPROCESSING
Before generate TIRI, we need to perform some preprocessing steps. Because copies of the same video with different frame sizes as well as frame rates are usually exist in the same video database. As a result, a fingerprinting algorithm should be robust to changes in the frame size as well as the frame rate. So down sampling is applied to form frame with some fixed size ($\mathbf{W}\times\boldsymbol{H}$) pixels. And then segment each frame into 50% overlapping of segments.

4.1 TIRI-Generation
In [4], adopted a method that involves the calculation of weighted average of total number of frames. The final output is normally a blurred image and it contain possible existing video

motion.TIRI is generated as follows. let be $l_{m,n,k}$ be the luminance value of the (m,n) th pixel of the k^{th} frame in a set of J frames. The pixels of TIRI are then obtained as a weighted sum of the frames

$$l'_{m,n} = \sum_{k=1}^{J} w_k \, l_{m,n,k}. \qquad \text{--------- (1)}$$

Figure 2 shows Frames (a) 1, (b) 3, and (c) 5 of the sequence of this baby's video and the resulting TIRIs with different weighting functions:(d)$w_{k=1}$(constant),(e)$w_{k=}$ k (linear), (f) w_k =0.65^k (exponential).

Compare to different weighting function, Figure 2 shows exponential weighting function produces perceptually better quality. Our experiments with other videos led to the same conclusion, thus we have chosen the exponential weighting function for generating TIRIs.

5. FEATURE EXTRACTION BASED ON GLCM

In [6], proposed a texture feature extraction based on GLCM.It creates a matrix with the directions and distance between pixels, and then extracts meaningful statistics from the matrix as texture features.GLCM express the texture feature according the correlation of the couple pixels gray level at different positions.

It provides the texture features include energy, contrast, correlation, homogeneity are extracted. These features are use along with Spatio-temporal fingerprints.

6. PROPOSED TIRI-BASED BINARY FINGERPRINTS

To find whether a query video is copied or not, first we extract binary fingerprints of the video from special image TIRI. The searching can be made efficient by combining the binary fingerprints.

7. FAST MATCHING OF FINGERPRINTS WITHIN A LARGE VIDEO DATABASE

As shown in Figure 1, to determine if a specific query video is a pirated version of a video in the database, its fingerprint is first extracted and searched for the closest fingerprint to the extracted query fingerprint. Fingerprints of two different copies of the same video content are similar but not necessarily equal. This is why we try to find a close match of the query in the fingerprint database and not an exact match should be deployed. In real-world applications, the size of online video databases can reach millions of videos, which translates into a very large fingerprint database size. So we need an efficient method to find closeness in a quick time. As an example of a fast search algorithm, Oostveen et al. proposed a search algorithm for their video fingerprinting algorithm, based on the inverted file technique [1]. In [5] proposed a modified version of [1] and cluster based search can be used in any fingerprinting search. They also showed that the second method yields superior results. Here we use cluster based similarity search technique for fingerprint processing.

7.1 Cluster based similarity search

The main idea [5] 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. So, a centroid is chosen for each cluster, termed as the cluster head. To
determine if a query fingerprint matches a fingerprint, the cluster head closest to the query is found. All the fingerprints belonging to this cluster is 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.

The various steps involved in the generation of fingerprints are expressed in below Figure 3.

1: Generate TIRIs from each segment of *J* frames using $w_k=\gamma^k$.

2: Segment each TIRI into overlapping blocks of size $2w\times 2w$, as follows,

$$B^{i,j} = \{ l'_{x,u} \mid x \in iw \pm w, y \in jw \pm w \}$$

3: Extract two DCT coefficients from each block. These are the first horizontal and vertical coefficients adjacent to the DC coefficient. The first vertical frequency, $\alpha_{i,j}$ as follows:

$$\boldsymbol{\alpha}_{i,j} = \mathbf{v}^T \mathbf{B}^{i,j} \mathbf{1}$$

Similarly, the first horizontal frequency $\beta_{i,j}$ can be found as follows:

$$\boldsymbol{\beta}_{i,j} = \mathbf{1}^T \mathbf{B}^{i,j} \mathbf{v}$$

4: Concatenate all the coefficients to form vector *f*.

5: Find *M*, which is the median element of *f*.

6: Generate binary hash value, the value below the threshold value is represented by 0, and the value that is greater is represented by 1.

7. GLCM method is applied in the TIRI for extracting texture features *t*.

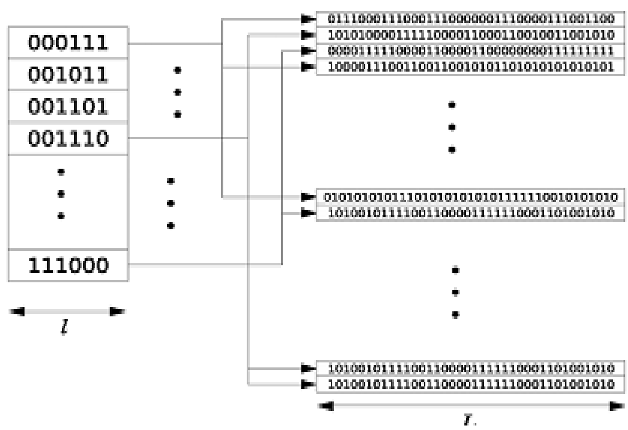


Figure 4 shows clustering the fingerprint for TIRI-DCT

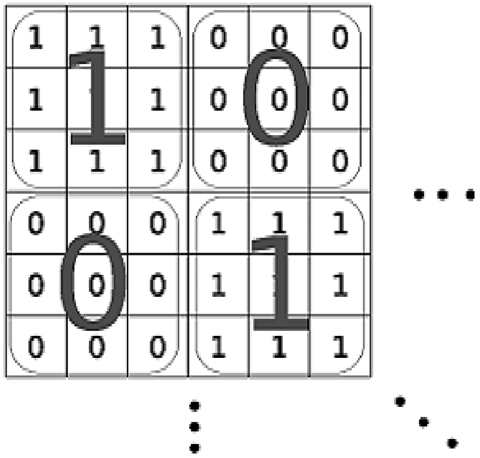


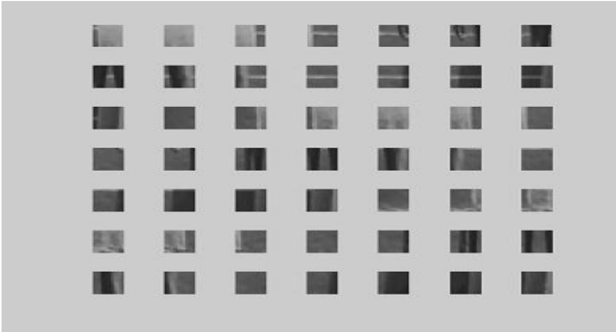
Figure 5 shows expanding a cluster head to compare it with a fingerprint.

8. RESULTS

In this section we have studied the generation of fingerprints using Spatio-temporal and texture features. To evaluate the performance of TIRI-DCT, a database with 130 videos are created. Before extracting the fingerprints every video is filtered and down-sampled to a frame rate of 4 frames/s and a frame size is 144 ×176.After preprocessing, videos are then segmented to 50% overlapping frame sequences. And this method employs 'J' as a total frames that is available in the particular video. Now segment each frame into a 50% overlapping sequence. In this fixed frame size, 357 overlapping blocks are expected.

Figure 6 shows sample video frames numbered 5,9,11 of total 40 frames

Figure 7 shows sample overlapping segments. Next step is the process of the generation of TIRI. This method calculates a weighted average of the frames to generate a representative image. The resulting image is basically a blurred image that contains information about possible existing motions in a video sequence. By using equation (1) an average image is derived. Since exponential weight is helpful in the faster perception of the image's motion and brings desired results. Hence here Exponential (γ=0.65) weight is used to generate TIRI and it is also shown in Figure 8 of the selected video.



Out of the 357 segments, 70 segments are shown of selected videos are shown in Figure 7.

Figure 8 shows TIRI of selected video frame

The steps shown in above are used to extract fingerprint of both Spatio-temporal and texture features. These are applied to the above received TIRI and the expected result is 2 DCT coefficients from each block.

	1	2	3	4
1	0.2347	0.9060	0.9159	0.1924
2	0.2240	0.9111	0.9198	0.1972
3	0.2225	0.9108	0.9210	0.1966
4	0.2259	0.9098	0.9201	0.1932
5	0.2259	0.9086	0.9198	0.1935
6	0.2239	0.9088	0.9204	0.1943
7	0.2192	0.9102	0.9214	0.1954
8	0.2265	0.9075	0.9189	0.1936
9	0.2180	0.9097	0.9207	0.1976
10	0.2171	0.9098	0.9200	0.1988
11	0.2166	0.9104	0.9202	0.1993
12	0.2195	0.9081	0.9184	0.1983
13	0.2211	0.9072	0.9175	0.1983
14	0.2261	0.9051	0.9153	0.1966
15	0.2386	0.9030	0.9110	0.1938
16	0.2314	0.9052	0.9127	0.1975
17	0.2415	0.9030	0.9077	0.1973
18	0.2473	0.9019	0.9057	0.1970
19	0.2449	0.9034	0.9073	0.1973
20	0.2498	0.9035	0.9049	0.1982
21	0.2533	0.9017	0.9030	0.1962
22	0.2679	0.8977	0.8993	0.1925
23	0.2620	0.8982	0.9011	0.1941
24	0.2503	0.9007	0.9080	0.1926
25	0.2528	0.8984	0.9080	0.1892
26	0.2526	0.8977	0.9096	0.1871
27	0.2527	0.8975	0.9092	0.1870
28	0.2488	0.8994	0.9113	0.1882

Figure 9 shows extracted texture features

As cited in [5] we concatenate all the coefficients to form feature vector. Now the median value of the vector can be used as threshold value. The value below the threshold value is represented by 0, and the value that is greater is represented by 1. We have to apply GLCM method to extract texture features from the resulting TIRI by adopting the steps numbered above as 7, 8, 9.Finally when the two fingerprints of the spatio temporal and texture features are combined, the desired binary fingerprints will be obtained. The afore said search algorithm make use of binary fingerprints and it enables one to find faster the closeness of the fingerprints.

9. CONCLUSION

The detection of illegal video copy is important for copyright protection. In this paper Spatio-temporal and texture binary fingerprints are used along with fast search algorithm. And also, this method yields efficient retrieval of copied video. We also aim to evaluate the performance by making use of multiple features.

10. 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]. C. Kim and B. Vasudev, “Spatiotemporal sequence matching for efficient video copy detection,” IEEE Trans. Circuits Syst. Video Technol.,vol. 15, no. 1, pp. 127–132, Jan. 2005.

[3].M. Malekesmaeili, M. Fatourechi, and R. K. Ward**,** “Video copy detection using temporally informative representative images,” in Proc. Int. Conf.Machine Learning and Applications, Dec.2009,pp.69-74.

[4] M. Malekesmaeili and R. K. Ward, “Robust video hashing based on temporally informative representative images,” in Proc. IEEE Int. Conf. Consumer Electronics, Jan. 2010, pp. 179-180.

[5] Mani Malek Esmaeili, Mehrdad Fatourechi, and Rabab Kreidieh Ward, Fellow, “A Robust and Fast Video Copy Detection System Using Content-Based Fingerprinting*”*, in IEEE trans,information forensics and security,vol 6, No 1 ,March 2011.

[6] FAN-HUI KONG,” Image retrieval using both color and texturefeatures”,July2009.