International Journal of Engineering Research and Development
e-ISSN: 2278-067X, **p**-ISSN: 2278-800X, www.ijerd.com
Volume 7, Issue 4 (May 2013), PP. 19-23

Video Fingerprint Extraction Using TIRI – DCT

Vaishali V. Sarbhukan[1], Prof. V. B. Gaikwad[2]
1Terna Engineering College,PG Student, Nerul,Navi Mumbai
2Terna Engineering College, Associate Professor, Nerul,Navi Mumbai

Abstract:- Nowadays copyright infringement and data piracy are major serious concerns. Thousands of videos are being uploaded to the internet and are shared every day. Out of these videos, considerable numbers of videos are illegal copies or some videos are manipulated versions of existing media. Due to these reasons copyright management on the internet becomes a complicated process. To detect such kind of infringements, there are two approaches. First is based on watermarking and other is based on content-based copy detection (CBCD). Video copy detection system relies on a fingerprint extraction algorithm. The fingerprint extraction algorithm extracts compact content- based signatures from special 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. These images are denoted by temporally informative representative images. Existing video fingerprint extraction algorithms have certain limitations. Therefore temporally informative representative images - discrete cosine transform (TIRI-DCT) is developed. TIRI-DCT is based on temporally informative representative images which contains spatial and temporal information of a short segment of a video sequence.

Keywords:- Fingerprinting System ,Video Copy Retrieval, Fingerprint Extraction Methods, TIRI-DCT, Content based Copy Detection

I. INTRODUCTION

Digital videos can be found on TV Channels, Web-TV, Video Blogs and the public Video Web servers. The massive capacity of these sources makes the tracing of video content into a very hard problem for video professionals. Recently for increasing online video repositories copyright infringements and data piracy have become serious concerns. Copyright infringement occurs when someone other than the copyright holder copies the "expression" of a work. Today's widespread video copyright infringement calls for the development of fast and accurate copy-detection algorithms. To detect infringements, there are two approaches. First is based on watermarking and other is based on content-based copy detection (CBCD).The watermarking is a widely used technique in the photography field. It allows the owner to detect whether the image has been copied or not. The limitations [1] of watermarks are that if the original image is not watermarked, then it is not possible to know if other images are copied or not. The primary aim of content-based copy detection (CBCD) is "the media itself is the watermark".

II. STRUCTURE OF FINGERPRINTING SYSTEM

Figure 1 shows the overall structure of this fingerprinting system [8].Content-based copy detection finds the duplicate by comparing the fingerprint of the query video with the fingerprints of the copyrighted videos. To find a copy of a query video in a video database, one can search for a close match of its fingerprint in the corresponding fingerprint database (extracted from the videos in the database). Closeness of two fingerprints represents a similarity between the corresponding videos; two perceptually different videos should have different fingerprints. Ideally, video fingerprint should be robust, discriminant, easy to compute, compact, secure, low complex.

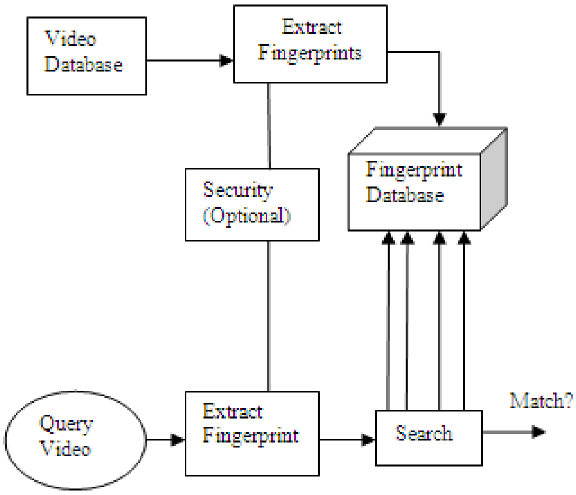

Fig. 1 Overall Structure of Fingerprinting System.

III. TRADITIONAL VIDEO FINGERPRINT EXTRACTION ALGORITHMS

Existing video fingerprint extraction algorithms are color-space based fingerprints, temporal fingerprints, spatial fingerprints, and spatio-temporal fingerprints.

3.1 Color-space-based fingerprints

They are mostly derived from the histograms of the colors in specific regions in time and/or space within the video. According to Lienhart [2] the color coherence vector (CCV) differentiates between pixels of the same color depending on the size of the color region they belong to. Lienhart et al. and Sanchez et al. have been tested for the domain of TV commercials and are susceptible to color variations. Naphade el al. proposed a technique that was also experimented by Hampapur el al. Naphade[3][4] proposed to use YUV histograms as the signature of each frame in the sequence and the use of histogram intersection as a distance measure between two signatures (frames). Disadvantages of color-space-based fingerprint are that color features change with different video formats and color features are not applicable to black and white videos.

3.2Temporal fingerprints

Temporal fingerprint algorithms can be applied to the luminance (the gray level) value of the frames. According to Indyk and Shivkumar[5], first, a video sequence is segmented into shots. Then, the duration of each shot is taken as a temporal signature, and the sequence of concatenated shot durations form the fingerprint of the video. Temporal signatures are computed on adjacent frames in a video. Temporal fingerprints are extracted from the characteristics of a video sequence over time. These features usually work well with long video sequences, but do not perform well for short video clips since they do not contain sufficient discriminant temporal information.

3.3 Spatial fingerprints

Spatial fingerprints are features derived from each frame or from a key frame. They are widely used for both video and image fingerprinting. Spatial fingerprints can be further subdivided into global and local fingerprints. One shortcoming of spatial fingerprints is their inability to capture the video's temporal information, which is an important discriminating factor. Therefore spatio-temporal fingerprints [9][10] are developed.

3.4 Spatio-temporal fingerprints

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. Some spatio-temporal algorithms consider a video as a three-dimensional (3-D) matrix and extract 3-D transform-based features [6][7].

IV. PROPOSED TIRI- DCT ALGORITHM

Figure 3 shows the block diagram of our proposed approach which is based on temporally informative representative images (TIRIs). TIRI-DCT Algorithm includes following steps

Step 1: Generate TIRIs from each segment of J frames after preprocessing of input video .TIRIs are generated using $w_k = \gamma^k$.
Step 2: Segment each TIRI into overlapping blocks of size $2\omega \times 2\omega$, using
$B^{i,j} = \{l'_{x,y} \mid x \in i\omega \pm \omega , y \in j\omega \pm \omega \}$
Where $i \in \{0,1,2, \ldots , W/\omega - 1\}$ and $j \in \{0,1,2, \ldots , H/\omega - 1\}$
When indexes are outside of boundary then TIRI image is padded with 0's.
Step 3: Extract DCT coefficient from each TIRI block. These are first horizontal and first vertical DCT coefficient. First vertical frequency $\alpha_{i,j}$ can be found for $B^{i,j}$ as
$\alpha_{i,j} = v^T B^{i,j} 1$
Where $v = [\cos(0.5\,\pi/2\omega), \cos(1.5\,\pi/2\omega), \ldots \cos(\pi - (0.5\,\pi/2\omega)]^T$
Step 4: Concatenate all coefficients to form feature vector f.
Step 5: Find median m, using all elements of f.
Step 6: Generate binary hash h, using f

$$h_k = \begin{cases} 1, & f_k \geq m \\ 0, & f_k < m \end{cases}$$

And 1 is column vector of all ones. Similarly first horizontal frequency $\beta_{i,j}$ can be found for $B^{i,j}$ as $\beta_{i,j} = 1^T B^{i,j} v$

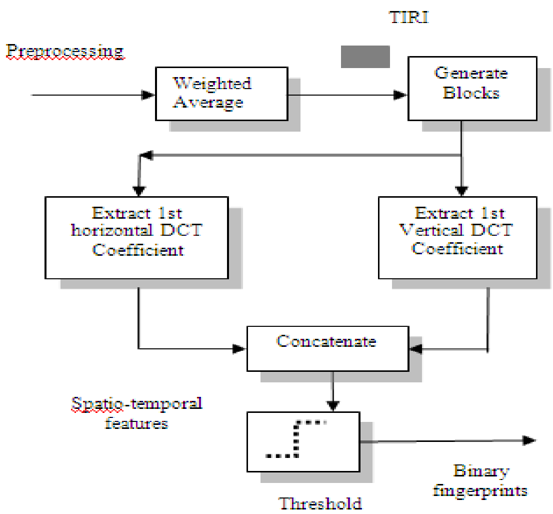

Fig. 2 Schematic of the TIRI-DCT algorithm

Due to increasing online video repositories, copyright infringements and data piracy have become serious concerns .Many videos uploaded on internet are illegal copies or manipulated versions of existing media. So today's widespread video copyright infringement calls for the development of fast and accurate copy-detection algorithms. Temporally Informative Representative Images (TIRI) - Discrete Cosine Transform (DCT) extracts compact content-based signatures from special images constructed from the video. Each image in TIRI represents a short segment of the video and contains temporal as well as spatial information about the video segment. The disadvantages of existing fingerprint extraction systems are overcome in Proposed TIRI-DCT. TIRI-DCT is based on temporally informative representative Images. As a Temporally informative representative Images (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.TIRI-DCT is faster than 3D-DCT while maintaining a very good performance over the range of the considered attacks. Another important property of TIRI-DCT is that it is computationally less demanding than 3D-DCT.

V. EXPERIMENTAL RESULTS

Fig. 3 Frame 50 of the sequence akiyo_qcif.yuv and the resulting TIRIs with different weighting functions for constant, linear, exponential

We have examined different weight factors (constant, linear, and exponential). Figure 3 shows the corresponding TIRIs using three different weighting functions as constant, linear, and exponential. It observed that exponential weighting factor generates images that best capture the motion. So we have chosen the exponential weighting function $w_k = \gamma^k$ for generating TIRIs.

Fig. 4 Overlapping Block Image

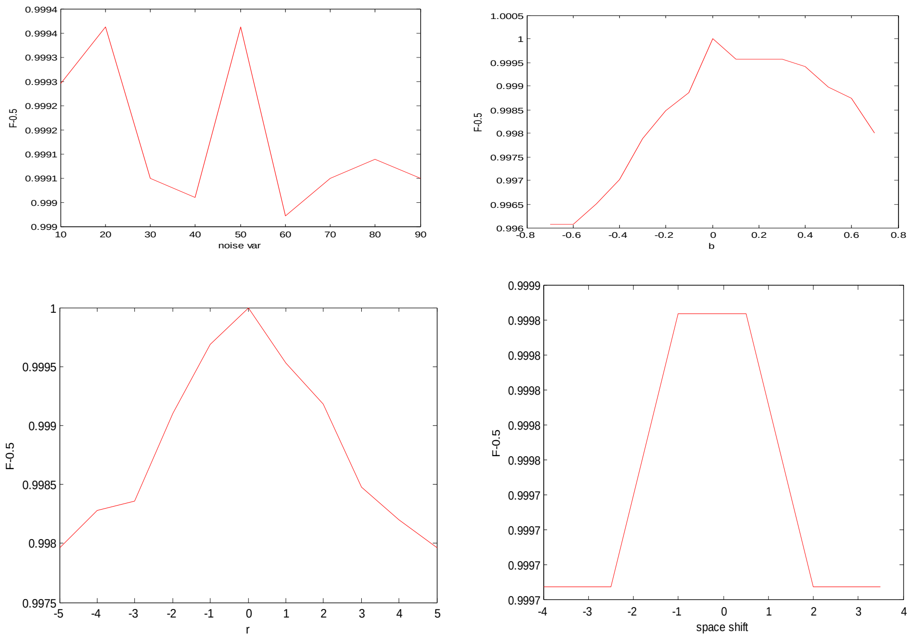

Fig. 5 F-score of TIRI -DCT for different attack parameters like Noise, Brightness, Rotation and Space Shift respectively

Overlapping block image for sequence akiyo_qcif.yuv is as shown in figure 4. We evaluate the performance of the proposed TIRI-DCT algorithm .We evaluate F-score of TIRI-DCT for different attack parameters like noise, brightness,rotation space shift etc. as shown in figure 5. Figure 6 shows the receiver operating characteristic (ROC) curve for different values of ω when $\gamma = 0.65$

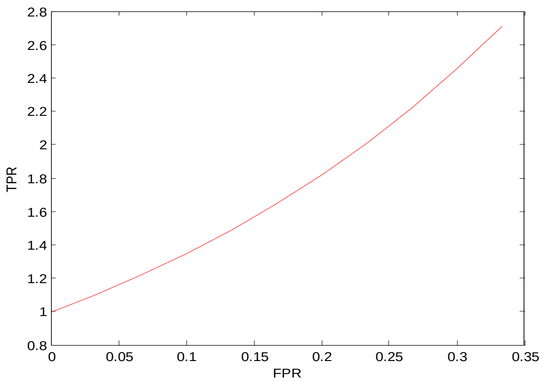

Fig.6 ROC for akiyo_qcif.yuv

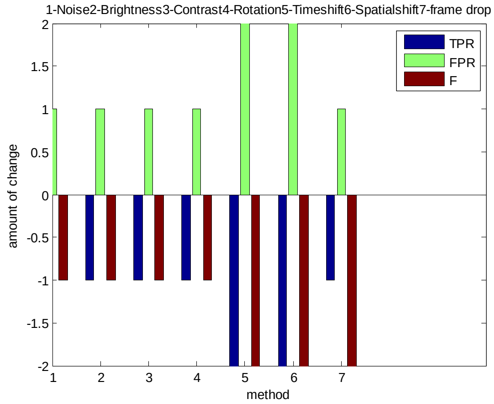

Fig. 7 Amount of change in the average TPR, FPR and F score

Fig 7 shows amount of change in average TPR, FPR and F score for akiyo_qcif.yuv.

REFERENCES

[1] Sunil Lee and Chang D. Yoo, "Robust Video Fingerprinting for Content-Based Video Identification", IEEE Transactions Circuits And Systems For Video Technology, Vol. 18, NO. 7, July 2008

[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 Proc. Conf. Image Video Retrieval (CIVR), 2007

[3] C. K. R. Lienhart and W. Effelsberg, "On the detection and recognition of television commercials", in Proc. of the IEEE Conf. on Multimedia Computing and Systems, 1997.

[4] J. V. J. M. Sanchez, X. Binefa and P. Radeva., "Local color analysis for scene break detection applied to TV commercials recognition" ,in Proceedings of Visual 99, June 1999.

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

[6] A. Hampapur and R. Bolle. "Comparison of sequence matching techniques for video copy detection", In Conference on Storage and Retrieval for Media Databases, 2002

[7] Chen, L. and Stentiford, F. W. M., "Video sequence matching based on temporal ordinal measurement", Pattern Recognition Letters, vol. 29, no. 13, Oct. 2008.

[8] Mani Malek Esmaeili, Mehrdad Fatourechi, and Rabab Kreidieh Ward, "A Robust And Fast Video Copy Detection System Using Content-Based Fingerprinting", IEEE Trans. On Information Forensics and Security", Vol. 6, No. 1, March 2011

[9] B. Coskun, B. Sankur, and N. Memon, "Spatiotemporal transform based video hashing", IEEE Trans. Multimedia, vol. 8, no. 6, Dec. 2006.

[10] G.Willems, T. Tuytelaars, and L. Van Gool, "Spatio-temporal features for robust content-based video copy detection", in Proc. ACMInt. Conf. Multimedia Information Retrieval, New York, NY, 2008, ACM.