Impact Factor (SJIF): 3.632

International Journal of Advance Research in Engineering, Science & Technology

e-ISSN: 2393-9877, p-ISSN: 2394-2444

Volume 3, Issue 3, March-2016

Video Copy Detection Using TIRI-DCT Algorithm

Khan Faizan[1], Khan Shumaz[2], Ansari Sami Ahmed[3], Ansari Sohrabalam[4], Mujeeb Shaikh[5]

[1]U.G. Student, Electronics & Telecommunication Engineering, Theem college of Engineering, Boisar
[2]U.G. Student, Electronics & Telecommunication Engineering, Theem college of Engineering, Boisar
[3] U.G. Student, Electronics & Telecommunication Engineering, Theem college of Engineering, Boisar
[4]U.G. Student, Electronics & Telecommunication Engineering, Theem college of Engineering, Boisar
[5] Faculty, Department of Electronics & Telecommunication Engineering, Theem college of Engineering, Boisar

Abstract *— The process of extracting a fingerprint from the video content is referred to as fingerprinting the video or video fingerprinting. Proposed fingerprint extraction algorithm Temporally Informative Representative Image -Discrete Cosine Transform (TIRI-DCT) 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. Here we evaluate the Performance analysis of Video Fingerprints on different parameters such as noise, contrast, brightness, time shift, spatial shift, frame shift, TPR,FPR & average etc. and analyzed for video Copy Detection. This system is tested on a database of different videos in the presence of different types of distortions which emphasize the robustness and discrimination properties of the copy detection system. The proposed fast approximate search algorithm facilitates the online application of the system to a large video database of tens of millions of fingerprints, so that a match (if it exists) is found in a few seconds.*

Keywords- *Content based fingerprinting, multimedia fingerprinting, video hashing, video copy detection, video copy retrieval.*

I. INTRODUCTION

A fingerprint is a content-based signature derived from a video that represents the video. To find a copy of a query video in a video database, the close match of its fingerprint in the corresponding fingerprint database is obtained from the videos in the database. Closeness of two fingerprints represents a similarity between the corresponding videos and two different videos will be having different fingerprints. This work, deals with spatio-temporal fingerprints because of their comprehensiveness. The temporally informative representative images (TIRIs) are formed from a video sequence since TIRI contains both spatial and temporal information of a short segment of a video sequence. Two fast search methods are used to matches the piracy of the video. Applying a 3-D transform to a video is a computationally demanding process and may pose problems in online applications. The proposed method for forming temporally informative representative images (TIRIs) from a video sequence. Because videos are available in different formats, it is more efficient to base the copy detection process on the content of the video rather than its name, the fingerprints should also be discriminant, to ensure that two perceptually different videos have distinguishable fingerprints. The TIRI-DCT method along with the search algorithm introduces a fingerprinting system that is robust, discriminant and proves to be the best in providing accurate results when compared to other fingerprinting techniques. In the proposed video copyright detection system consists of the following stages

a. Feature Extraction
b. Searching of fingerprints
c. Decision making

II. PROBLEM STATEMENT

A video fingerprint is a compact bitstream representation of the underlying content that is robust against many signal processing operations on the content and can be used to uniquely identify that content. One of the main challenges of a video fingerprinting method is to provide robustness against geometric attacks such as rotation, aspect ratio change and cropping. These modifications are particularly difficult as there is either a loss of information (e.g cropping) or a loss of registration (e.g rotation,cropping) between modified content and original content. There have been lot of approaches proposed in literature to tackle these challenges. The process of extracting a fingerprint from the video content is referred to as fingerprinting the video or video fingerprinting. Previous fingerprinting extraction methods can be applied to specific videos, some can be applied only to large video sequences, and some contain only spatial information. Therefore spatiotemporal fingerprinting extraction algorithms are designed. Proposed a fingerprint extraction algorithm Temporally Informative Representative Images - Discrete Cosine Transform (TIRI-DCT) 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 videosegment. These images are denoted by temporally informative representative images. 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 matches for the fingerprints of the query video. The proposed fast approximate search algorithm facilitates the online application of the system to a large video database of tens of millions of fingerprints, so that a match (if it exists) is found in a few seconds.

III. LITERATURE SURVEY

M. Malekesmaeili, M. Fatourechi, and R. K.Ward, [2] proposes an approach for generating representative images of a video sequence that carry the temporal as well as the spatial information. These images are denoted as TIRIs, Temporally Informative Representative Images[2]. Performance of the approach is demonstrated by applying a simple image hashing technique on TIRIs of a video database.The most crucial part in CBCD is the feature extraction. Features for fingerprinting should be carefully chosen since they directly affect the performance of the entire video fingerprinting system. What information in a video can be used to build a well-identified and robust fingerprint? In the proposed paper of S. Lee and C.D. Yoo [3], three important qualities have been proposed.

Robustness: The fingerprints extracted from a degraded video should be similar to the fingerprinting of the original video.

Pair-wise independent: Two videos, that are perceptually different, must have different fingerprints.

Database search efficiency: Fingerprints must suitable for fast database search. Many features have been proposed for the video fingerprinting [2,3], e.g. color (luminance) histogram, mean luminance and its variant, dominant color, etc. Feature extraction techniques based on more information of video content have been proposed. For example, in the paper published by A. Hampapur and R. Bolle [7], CBCD with motion direction, ordinal intensity, color histogram are introduced and compared.Previous video fingerprint extraction algorithms are classified into four groups as color-space-based fingerprints, temporal fingerprints, spatial fingerprints and spatio-temporal fingerprints.

IV. BRIEF HISTORY

Research that began a decade ago in video copy detection has developed into a technology known as "video fingerprinting". A video fingerprint is a quantifier that is extracted from a piece of video content. The process of choosing a fingerprint from the original video content is referred to as video fingerprinting. There is an obvious correlation between human fingerprint and video fingerprinting. Just as human fingerprint can uniquely identify a human being, video fingerprint can uniquely identify a piece of video content. The parallelism expands to the process of subject identification by fingerprint: At the start, known fingerprints are stored in a database, and then a subject's fingerprint is queried against the database to match.

3.1. Video Fingerprints used previously

There are four types of Fingerprinting which were used previously.

3.1.1. Color-Space-Based Fingerprints

Color-space-based fingerprints are among the first feature extraction methods used for video fingerprinting. They are normally derived from the histograms of colors in certain regions in time and/or space within the video. Advantages of color histograms are efficiency and insensitivity to small changes in camera viewpoint. Color histograms are frequently used to compare images. Color histograms have some limitations such as it provides no spatial information and it rarely describes what all colors are there in the image and their exact quantities.

3.1.2. Temporal Fingerprints

To overcome drawback of color-space-based fingerprints new video fingerprint extraction algorithm is developed that can be applied to the luminance (the gray level) value of the frames. In order to find pirate videos on the Internet, [2] use temporal fingerprints based on the shot boundaries of a video sequence and the time distance between boundaries is its signature. The feature of a boundary image isn't used since the content of this image is unreliable, usually black or hard to be identified by human eyes. This technique needs a shot-boundary-detection algorithm, and it can be adequate for finding a full movie, but it might not work well for short episodes with a few boundaries. The common way for shot detection is to evaluate difference value between consecutive frames represented by a given feature.

3.1.3. Spatial Fingerprints

Spatial fingerprint algorithm [3] converts a video image into YUV color space in which the luminance (Y) component is kept and the chrominance components (U, V) are discarded. The luminance image is further subdivided into a fixed-sized grid of blocks independent of frame resolutions. Spatial fingerprints are features derived from each frame or a key frame and are mostly used for both image and video fingerprinting. There is a huge research in the area of image fingerprinting and a large no of researchers have extended the concepts developed for image fingerprinting to the video fingerprinting field.

3.1.4. Spatio-Temporal Fingerprints

One shortcoming of spatial fingerprints is their inability to capture the video's temporal information, which is an powerful discriminating factor. Hence the new fingerprinting approach is designed to call as spatio-temporal fingerprints. A Spatio-temporal fingerprints that consists of both spatial and temporal information about the video are thus expected to perform better than fingerprints that use only temporal or spatial fingerprints. Temporal information is related to time. Spatial information describes the physical location of objects and metric relationship between objects.

V. PROPOSED METHOD

In TIRI – DCT, first step is generation of temporally informative representative images TIRI.So proposed TIRI-DCT method along with fast search algorithm outperforms than 3D-DCT since it is more robust, discriminant, and fast. Figure 5.1 shows the overall structure of fingerprinting system. When an identifier or the signature is extracted from the content without changing the content, it is fingerprinting. Video fingerprinting has been used to refer to the technology encompassing algorithms, systems, and workflows that use video fingerprint
for video identification

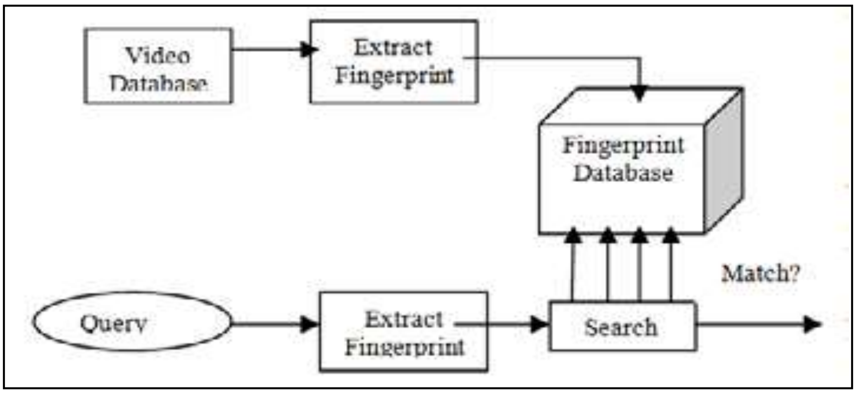
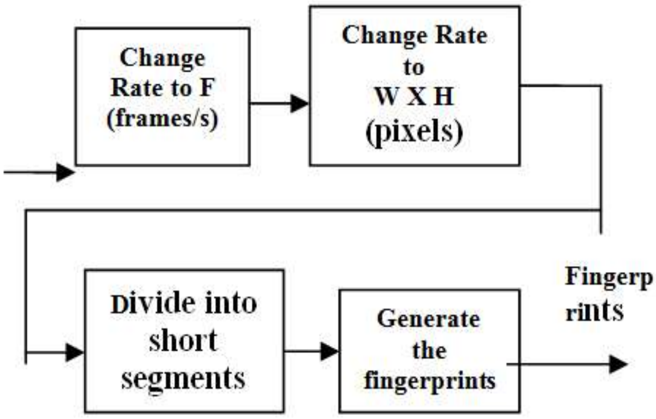


Fig. 5.1 : Structure of a Fingerprinting

System Advantages of proposed system are as follows.

- A Spatio-temporal fingerprint is adopted because of their comprehensiveness.
- The TIRI-DCT method introduces a fingerprinting system that is robust, discriminant, and fast.
- The TIRI-DCT outperforms the well-established (3D-DCT) algorithm and maintains a good performance for different attacks like noise, time shift, spatial shift, brightness/contrast, rotation, frame loss that normally occurs on video signals.
- TIRI –DCT algorithm is applied to color videos as well as black and white videos.

VI. DESIGN STEPS



As shown in figure 6.1, each video is down-sampled both in time and space. Prior to down-sampling, a Gaussian smoothing filter is applied in both domains to prevent aliasing. This down-sampling process provides the fingerprinting algorithm with inputs of fixed size of pixels and fixed rate (F frames/second. After pre-processing, the video frames are divided into overlapping segments of fixed length, each containing J frames. The fingerprinting algorithms are applied to these segments. Overlapping reduces the sensitivity of the fingerprints to the "synchronization problem" which is called as "time shift". Features are derived by applying a 2D-DCT on overlapping blocks of size from each TIRI. As shown in figure 6.2 the first horizontal and the first vertical Discrete Cosine Transform (DCT) coefficients (features) are extracted from each block. The value of the features from all the blocks is 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
TIRI-DCT Algorithm:

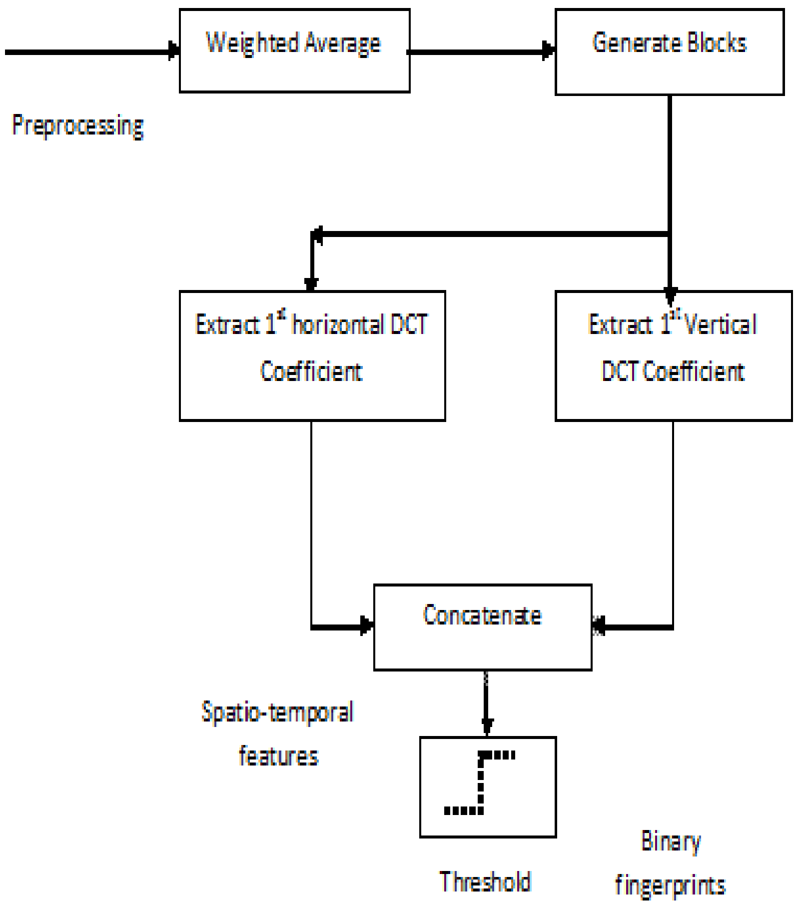


Fig.6.2 : Schematic of the TIRI-DCT algorithm

Figure 6.2 shows the block diagram of proposed approach which is based on temporally informative representative images (TIRIs). Pre-processing is often used to improve visual quality and coding efficiency of video compression systems requirement and then cause the aliasing by being below Nyquist minimum.

The steps involved in generating binary fingerprint are:

a) TIRI-DCT is segmented into number of blocks.
b) Extract 1st horizontal coefficients of DCT.
c) Extract 1st vertical coefficients of DCT.
d) Concatenate horizontal and vertical coefficients.
e) Compute median m from the concatenated values.
f) Compare coefficients with median.
g) If the values of coefficients are greater than or equal to median value, then the value 1 is assigned to it.
h) If the values of coefficients are greater than or equal to median value, then the value 1 is assigned to it.

TIRI-DCT algorithm includes the 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 $2w \times 2w$, using

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

Where $i \in \{0,1,2,\ldots,W/w-1\}$ and $j \in \{0,1,2,\ldots,H/w-1\}$

When indexes are outside of the 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 $\propto_{i,j}$ can be found for $B^{i,j}$ as

$$\propto_{i,j} = v^T B^{i,j} 1$$

Where

$$v = [\cos(0.5\pi/2w), \cos(1.5\pi/2w), \ldots\ldots\ldots\ldots \cos(1-0.5\pi/2w)]^T$$

And 1 is a column vector of all ones.

Similarly first horizontal frequency $\beta_{i,j}$ can be found for $B^{i,j}$ as

$$\beta_{i,j} = 1^T B^{i,j} v$$

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

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

Generation of Temporally Informative Representative Images (TIRIs):

Some temporal information about a video segment can be embedded in the representative images. For a generation of temporally informative representative images, a weighted average of the frames is calculated. The resulting image is basically a blurred image that contains information about possible existing motions in a video sequence. The TIRI is thus generated as follows

$$l'_{m,n} = \sum_{k=1}^{J} w_k \, l_{m,n,k}$$

Where $l_{m,n,k}$ be the luminance value of the (m, n) th pixel of the k th frame in a set of J frames. $l'_{m,n}$ represents the pixels of the temporally informative image and w_k is the weight associated with each frame. Different weighting functions are as follows

$w_k = 1$ is called as simple averaging weight

$w_k = k$ is called as linearly changing weights

$w_k = 1 - e^{(k-\mu)^2}/\sigma$ is called as Gaussian weighting

$w_k = \gamma^k$ is called as exponential weighting

A very large, or one close to 0, generates a TIRI with detailed spatial information and low temporal information, resulting in a more discriminant representative. By changing from 0 to 1, it is possible to move from a single frame selection (high spatial information) to selecting all frames with equal weights (high temporal information).

VII. RESULTS

Here Performance analysis on Video Fingerprint on different parameters such as noise, contrast, brightness, time shift, spatial shift, frame shift, TPR,FPR & average etc. is been observed and analyzed for video Copy Detection.

Performance Measures:

There is a database of n videos. It is used to evaluate the performance of the proposed TIRI-DCT method followed by inverted file based similarity search and cluster based similarity search algorithms. Here true positive rate (TPR), false positive rate (FPR) and F-Score are used as performance measures. For these performance measures, noise, brightness, contrast, rotation, time shift, spatial shift and frame loss are used as attack parameters.

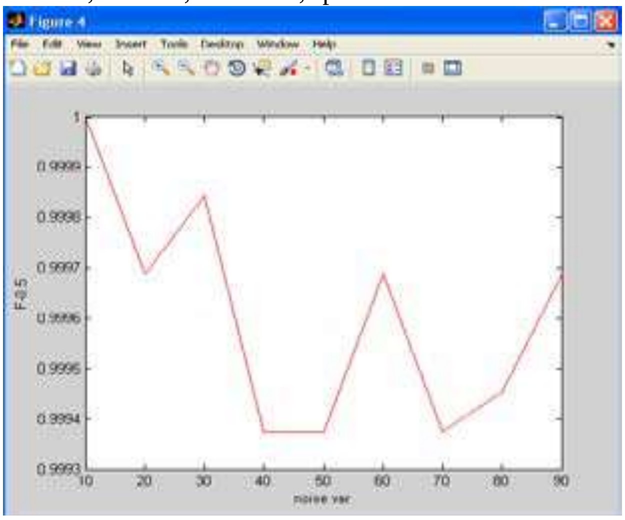


Fig 6.1 (a) : F-Score vs Noise

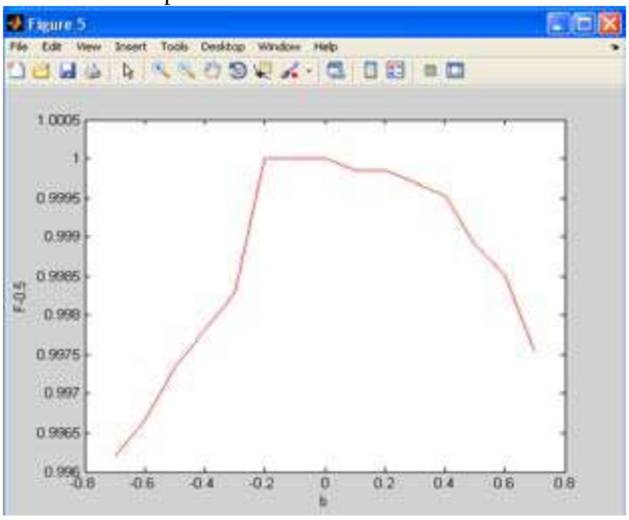


Fig 6.1(b) : F-Score vs Brightness

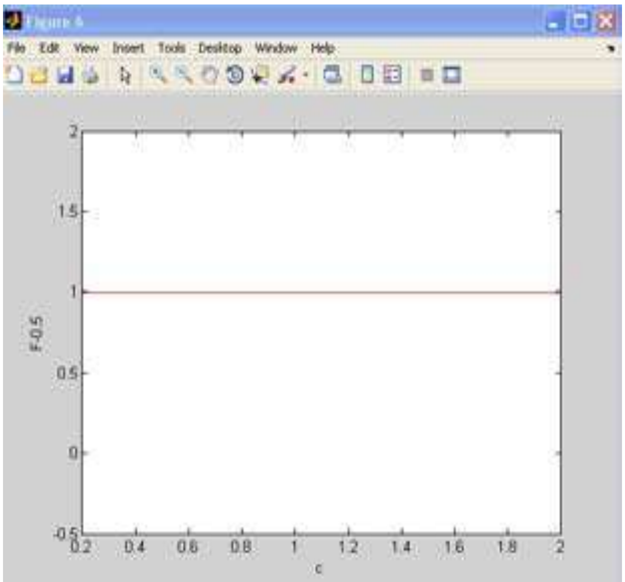


Fig 6.1(c) : F-Score vs Contrast

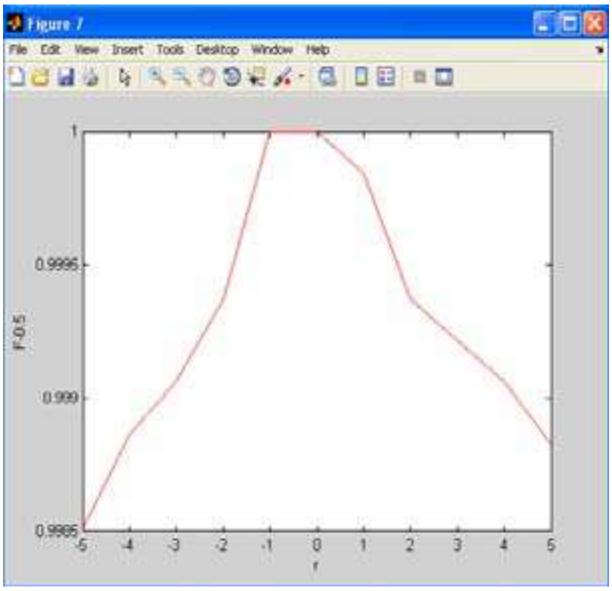


Fig 6.1(d) : F-Score vs Rotation

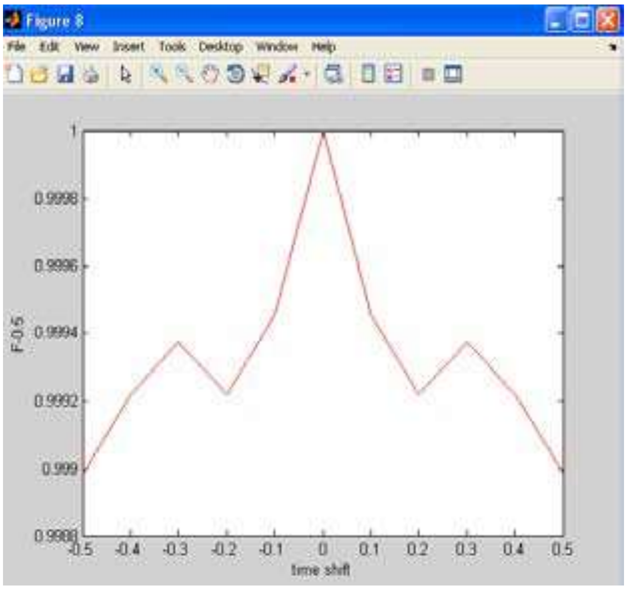


Fig 6.1(e) : F-Score vs Time Shift

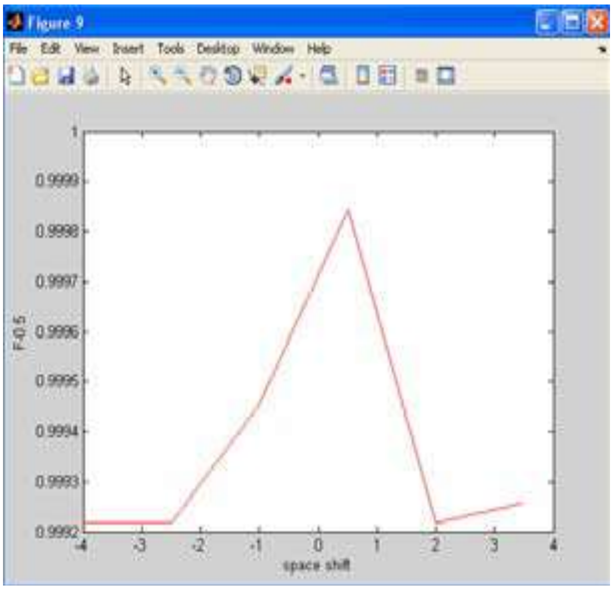


Fig 6.1(f) : F-Score vs Space Shift

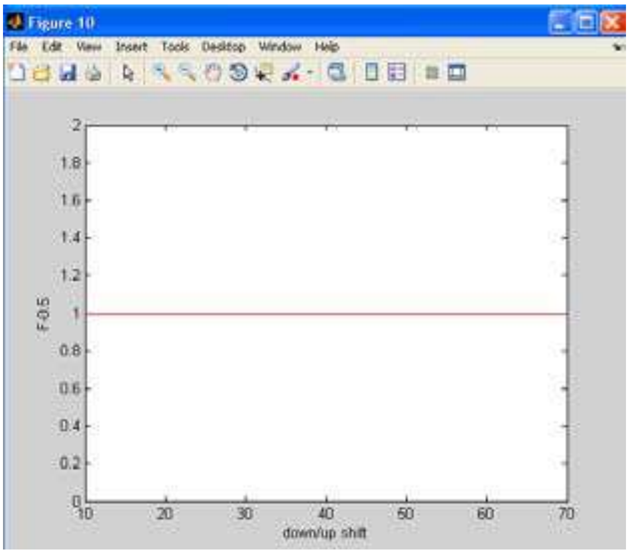


Fig. 6.1(g) : F-Score vs Frame Loss

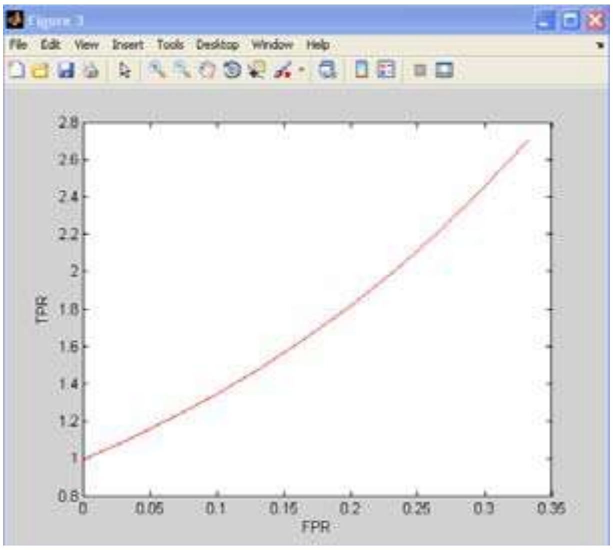


Fig. 6.2 : TPR vs FPR

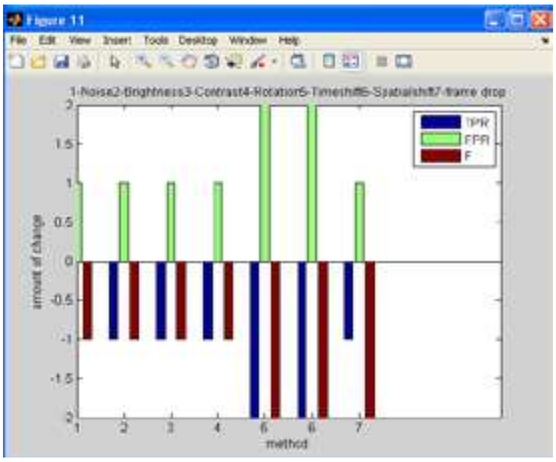


Fig 6.3 Amount of change in average TPR, FPR & F when security is added to the system

VIII. CONCLUSION

This paper proposes a fingerprinting system for video copy detection. It is used to evaluate the performance of the proposed TIRI-DCT method followed by inverted file based similarity search and cluster based similarity search algorithms. Here true positive rate (TPR), false positive rate (FPR) and F-Score are used as performance measures. For these performance measures, noise, brightness, contrast, rotation, time shift, spatial shift and frame loss are used as attack parameters .The system consists of a fingerprint extraction algorithm followed by an approximate search method. The proposed fingerprinting algorithm (TIRI-DCT) extracts robust, discriminant, and compact fingerprints from videos in a fast and reliable fashion. These fingerprints are extracted from TIRIs containing both spatial and temporal information about a video segment. We demonstrate that TIRI-DCT generally outperforms the well-established (3D-DCT) algorithm and maintains a good performance for different attacks on video signals, including noise addition, changes in brightness/contrast, rotation, spatial/temporal shift and frame loss.

IX. FUTURE WORK

As part of our future work, we will conduct a detailed analytical study of the security of fingerprinting algorithms including the one proposed in this paper. As another part of our future work, we will carry an extensive comparison study to compare our fingerprinting algorithms to other state-of-the-art algorithms. We will also evaluate our proposed fast search methods when applied to other fingerprinting methods. We also plan to study the performance of the system in the presence of some other attacks, such as cropping, and logo insertion.

REFERENCES

[1] JOostveen, T. Kalker, and J. Haitsma, "Feature extraction and a database strategy for video fingerprinting," in Proc. Int. Conf. Recent Advancesin Visual Information Systems (VISUAL), London, U.K., 2002, pp. 117–128, Springer-Verlag.

[2] IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 6, NO. 1,MARCH 2011 213 A Robust and Fast Video Copy Detection System Using Content-Based Fingerprinting by-Mani MalekEsmaeili, Mehrdad Fatourechi, and Rabab Kreidieh Ward, Fellow, IEEE

[3] Sunil Lee and Chang D Yoo. Video fingerprinting based on centroids of gradients orientations. In ICASSP '06 Proceedings of the 2006 IEEE International Conference on Acoustics, Speech and Signal Processing, pages 401_404, Washington, DC, USA, 2006. IEEE Computer Society.

[4] J. A. Haitsma, A. A. C. M. Kalker, C. P. M. J. Baggen, and J. C. Oostveen, Generating and matching hashes of multimedia content US 2002/0178410, Nov. 2002.

[5] M. L. Miller, "Audio fingerprinting: Nearest neighbour search in high dimensional binary spaces," in IEEE Workshop on 2002, MultimediaSignal Processing, 2002, 2002, pp. 182– 185

[6] 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 CIVR '07: Proceedings of the6thACM international conference on Image and video retrieval, pages 371_378, New York,USA,2007. ACM.

[7] A. Hampapur, K. Hyun, and R. M. Bolle. Comparison of sequence matching techniques for video copy detection. volume 4676, pages 194_201. SPIE, 2001.

[8] Wei-LunChao,The Core of Video Fingerprinting: Examples of Feature Extraction.