International Journal of Modern Engineering Research (IJMER)
www.ijmer.com Vol.2, Issue.4, July-Aug. 2012 pp-1747-1753 ISSN: 2249-6645

Enhancing Security in a Video Copy Detection System Using Content Based Fingerprinting

E. Meenachi[1], G. Selva Vinayagam[2], C. Vinothini[3]
1. PG Student, Dept of IT, SNS College Technology, Coimbatore
2. Assistant Professor, Dept of IT, SNS College of Technology, Coimbatore
3. PG Student, Dept of IT, SNS College Technology, Coimbatore

Abstract: *Information Security means protecting information and information systems from unauthorized access, use, disclosure, disruption, modification, inspection, recording or destruction. The "Video Copy Detection" system is based on detecting video copies from a video sample to avoid copyright violations. To identify video sequences, Content Based Copy Detection (CBCD) presents an alternative to the watermarking approach. Content based Video fingerprinting methods extract several unique features of a digital video that can be stored as a fingerprint of the video content. Video copy detection system aims at deciding whether a query video segment is a copy of a video from the indexed dataset or not. The evaluation and identification of video content is then performed by comparing the extracted video fingerprints. The search algorithm searches the fingerprints that are stored in the database to find closest match with the fingerprints of the query video. The Interest point Matching algorithm is implemented to make the fingerprint robust against content changing attacks such as changing the background of the video. The proposed algorithm is tested on videos which are affected by the distortions like noise, changes in brightness/contrast, rotation, time shift and changes in background. The results gives the high average true positive rate and low average false positive rate. The results demonstrate that the fingerprint extracted using this algorithm is robust. Video fingerprinting system is important for the applications like Digital Rights Management (DRM) area, particularly regarding the distribution of unauthorized content on the internet.*

Index Terms: ***Content-Based Fingerprinting, Video Copy Detection, Video Copy Retrieval, Interest points***

I. INTRODUCTION

The widespread availability of video content and services and the rapid diffusion of broadband platforms have brought new challenges to content owners, publishers and distributors. Video copy detection is a recent research domain which has emerged largely in response to this technological trend. The goal of video copy detection is to develop automated video analysis procedure to identify the original and modified copies of a video among the large amount of video data for the purposes of copyright control, monitoring and structuring large video databases.

A. Content Based Video Copy Detection

Content-Based Copy Detection (CBCD) schemes are an alternative to the watermarking approach for persistent identification of images and video clips. The primary concept of content-based copy detection (CBCD) is "the media itself is the watermark," i.e., the media (video, audio, image) contains enough unique information that can be used for detecting copies. Content-based copy detection schemes extract signatures from the original media and store it in a database. The same signatures are extracted from the test media stream and compared to the original media signature which is already extracted and stored in a database to determine if the test stream contains a copy of the original media. The signatures that are extracted from the media are termed as 'fingerprints'.

B. Fingerprint

Fingerprints are compact content-based signature that summarizes a video signal or another media signal. These signatures are feature vectors that uniquely characterize specific signal. Video fingerprinting is a proven and commercially available technique that can be used for content based copy detection. The task of a video-fingerprinting system is to detect whether a particular segment of video is (partly) based on the same original video as video footage in a database of reference videos.

The fingerprint of a video should be *robust* to the content-preserving distortions i.e., the changes made to the videos unintentionally or intentionally by the users of video sharing websites. It determines the tolerance of the system to different encoding processing that give rise to several distortions. Fingerprint should also be *discriminant* to make sure that different videos has distinguishable fingerprints. It should also be *secure* so that it is difficult for an adversary to generate similar fingerprints for different videos and manipulate the video copy detection systems.

C. Type of Fingerprints

Existing video fingerprints can be classified into four groups: spatial, temporal, colour and spatio-temporal fingerprints. A *spatial fingerprint* characterizes spatial features of a video frame and is computed independent of other

frames.These approaches are based on intensity statistics such as mean, variance, centroid, and other higher order moments of the spatial content of the different frames of the video. The frames are divided into subsections and for each section, the features are calculated. This approach allows uniform spatial processing over the entire frame. However it is less robust to geometrical operations like rotation and scaling. Examples of spatial features include luminance patterns, differential luminance or gradient patterns, and edges.

The *temporal fingerprint* describes temporal features of a video and is computed between two frames in the temporal direction. This involves comparing each frame against the previous frame. Differences between the two are flagged as motion. This is done by comparing macro-blocks of a pre-determined size. Motion is detected as a change in the color between frames within a macro-block, while the magnitude of the motion is deduced from the difference in color. A measure of the amount of motion of each frame relative to the preceding frame is calculated by summing the differences in color between these frames for all macro blocks. This results in a series of features, each indicative of a measure of motion between two adjacent frames over time. Consequently, the size of the motion fingerprints is dependent on the amount of motion that has been detected. Examples of temporal features include frame difference measures, motion vector patterns, and shot durations.

A *color fingerprint* captures color characteristics of a video frame and is computed in a color space such as RGB or YUV. Color-space-based fingerprints are among the first feature extraction methods used for video fingerprinting. They are mostly derived from the histograms of the colors in specific regions in time and/or space within the video. Since color features changes with different video formats these features have not been popular.Another drawback of the color features is that they are not applicable to black and white videos.

A *spatio-temporal fingerprints* contains both spatial and temporal informations about the video. They perform better than the fingerprints that use only spatial or temporal fingerprints. To capture relevant information from both temporal and spatial domain they apply temporal and spatial differentiation in the feature space. They take the differences of corresponding features extracted from subsequent blocks from one frame, and from subsequent frames in the temporal direction. This feature extraction is robust to global changes in luminance, also robust to luminance and contrast variations because only the sings of the difference are retrained. Their experimental results show that the method is also robust against MPEG compression and median filtering. Different types of fingerprints are combined to form a video fingerprint.

II. LITERATURE REVIEW

Hye-Jeong Cho et al.[6] proposed the hierarchical video copy detection method which estimates similarity using statistical characteristics between original video and its spatial variations. The target video is transformed by various spatial variations such as blurring, constrast change, zoom in, and zoom out. Zheng Cao et al.[15] computed a video signature based on ordinal measurement of video spatio-temporal distribution feature. The video similarity is measured by the computation of the distance of video signature. The duplicate videos are generated by spatio and temporal changes. Juan Chen[7] extracted key frames from the video sequences to save the computing time and storage space. Candidate descriptors are then determined in the local context by applying Harris point detection. The descriptors are invariant to spatial deformations like shifting, cropping and change of ratio.

Coskun et al. proposed a hash function [3] to extract a fingerprint based on signal processing operations. Two robust hash algorithms for video were implemented in this paper. Discrete Cosine Transform (DCT) based on classical basis set and Discrete Cosine Transform (DCT) based on a novel randomized basis set (RBT).The video here is considered as the 3D matrix.DCT transformations are applied to extract the coefficients and hashing was performed. The 3D transformations extract both spatial and temporal coefficients. The fingerprint is robust against signal processing modifications and channel transformations. Does not robust against malicious attacks. Security measures are not solved thoroughly in RBT [9].

Malekesmaeili.M et al generates fingerprints of a video sequence that carries both temporal and spatial information's [5]. This algorithm is applied for a 2D data. Gaussian filtering is applied to the video signals to prevent aliasing. Then the video signals are pre-processed (re-sampling and spatial resizing) to get the fixed frame size and frame rate. Then the video is divided into frames and weighted average of the frames is taken to obtain TIRI (Temporally Informative and Representative Images). DCT based hashing is applied which results in equal number of 0's and 1's. The fingerprint is robust to time shift, frame dropping, added Gaussian noise. But the performance is low for rotation and shift attacks. Hampapur et al uses a combination of feature based matching and inverted index files to detect copies of video clips [3].

CBCD does not modify the video stream and hence can be applied to find copies of media in circulation. The Reference Signature extract a set of signatures for the original media (M).The Test Signature extract the same set of signatures from the test media (T). Then the distance between the test and reference signatures are measured to compare both the signatures. If distance is lesser than a threshold, then it is a copied video. A set of representative frames are selected from each video item in the collection. The corresponding representative frames from each set are represented by an inverted image index table. Once a inverted image index has been created, it can be used to match a query image against the collection [3].

III. EXTRACTION OF FINGERPRINT

A fingerprint is a content-based signature which is derived from a video (or other form of a multimedia asset) so that it specifically represents the video or asset. 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[10].

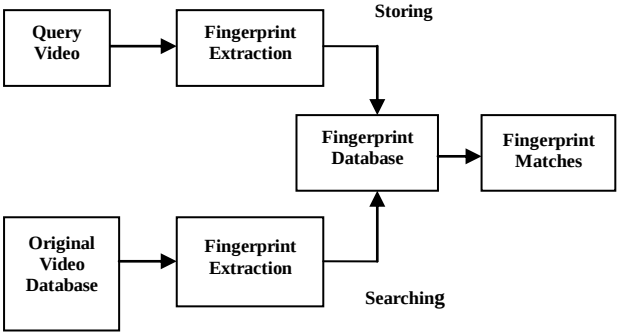

Fig.1. A complete Fingerprinting system

A. Tiri-Dct Algorithm

The existing algorithm for extracting the fingerprint is TIRI-DCT (Temporally Informative Representative Images). This method calculates a weighted average of the frames to generate representative images. This sequence will carry the temporal as well as spatial informations. The image is then divided into blocks. Then the first horizontal and the first vertical DCT coefficients (features) are then extracted from each block. The value of the features from all the blocks is concatenated to form the feature vector.[10] Each feature is then compared to a threshold (which is the median value of the feature vector) and then a binary fingerprint is generated.

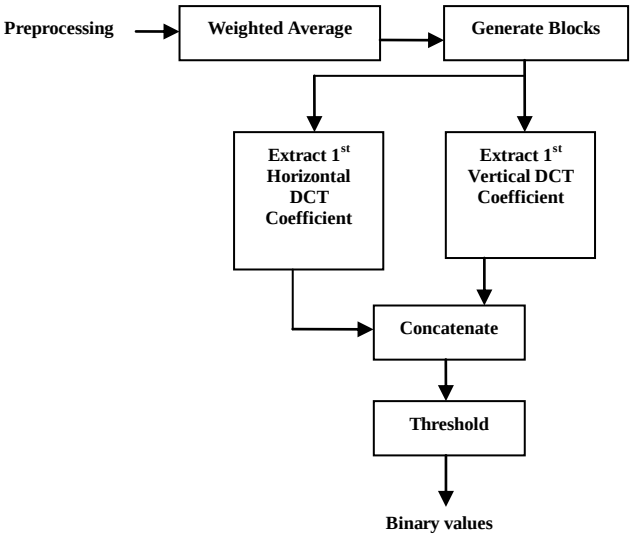

Fig.2Finger Print Algorithm Preprocessing Steps

In order to determine whether a query video is an attacked version of a video in a database or not, its fingerprint is first extracted. The fingerprint database (previously created from the videos in the video database) is then searched for the closest fingerprint to the extracted query fingerprint.

Two searching techniques are implemented:

- Inverted-File-Based Similarity Search
- Cluster-Based Similarity Search

1). Inverted-File-Based Similarity Search

The 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. Divide each fingerprint into small non overlapping blocks of bits. [10]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. The query is then compared to all the fingerprints that start with the same word. The Hamming distance between these fingerprints and the query is then

calculated. If a fingerprint has a Hamming distance of less than some predefined threshold, it will be announced as the match. When no match is found in the end, it is stated that the query does not belong to the database.

2). Cluster-Based Similarity Search

Clustering is used to reduce the number of queries that are examined within the database. By assigning each fingerprint to one and only one cluster (out of K clusters), the fingerprints in the database will be clustered into K non overlapping groups. To do so, a centroid is chosen for each cluster, termed the cluster head.[10]

A fingerprint will be assigned to a 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 (of less than a certain threshold) from the query. This process continues until a match is found or the farthest cluster is examined. If no match is found, the query is declared to be out of the database.

B. Problem Description

In existing system the fingerprint which is extracted using TIRI-DCT algorithm is not robust against the content changing attacks such as changing the background of the video or replacing picture in picture and the performance of the system is low in the presence of some other attacks, such as cropping, and logo insertion. The security of the fingerprint can be achieved only with the fingerprints of smaller length. Larger fingerprints results in a decrease in detection speed as they require more computation in calculating the hamming distance between the fingerprints. Also performance of the detection speed gets reduced when there are a large number of fingerprints in the database

IV. INTEREST POINT MATCHING ALGORITHM

Corner detection is an approach used to extract certain kinds of features and infer the contents of an image. Corner detection is frequently used in motion detection, image registration and video tracking. Corner detection overlaps with the topic of Interest point detection. An interest point is a point in an image which has a well defined position in an image and can be robustly detected. A good interest point detector has the following three properties:(1)The interest points are repeatable, (2)the descriptors produced from them are unique, (3) they are well-distributed spatially. An Interest point in an image has a clear, mathematically well-founded definition and has an well defined position in an image. It is stable under local and global distortions in an image domain. To increase the robustness of the fingerprint against the content changing attacks such as changing the background of the video an Interest point Matching algorithm is proposed. The conceptual basis of this algorithm is the detection of "super points," those points which have the greatest interest strength (i.e., which represent the most prominent features) and the subsequent construction of a control network. Sufficient spatial information is then available to reduce the ambiguity and avoid false matches.

The algorithm proposed in this paper includes two parts: A) Interest-point detection;B) interest-point matching.

A. Interest-Point Detection

The Harris detector is a well-known interest-point detection algorithm[17] to detect and extract the interest points. The Harris algorithm determines whether a point is a corner based on the Harris matrix A at the point $P(x, y)$.The interest strength is determined based on the magnitudes of the eigenvalues ($\gamma 1$ and $\gamma 2$) of A. The following function Mc was suggested as the interest strength:

$Mc = \det(A) - \kappa trace2(A)$. The value of κ has to be determined empirically, and in the literature, values in the range of 0.04–0.06. If $Mc > 0$, it is a corner; otherwise, it is not a corner. Obviously, the corner should be the point with the local maximum value of Mc. By calculating the interest strength Mc over the whole image, an image which shows the interest strength can be obtained[17]

B .Interest-Point Matching

After the detection of interest points, a control network is constructed with the correspondences. Correspondences are defined as those interest points with the minimum difference in position and angle. Then the control network of each frame is compared.

If the location of the points are same then it is marked as 1 else it is marked as 0. Thus the binary sequences that are obtained are stored as the fingerprint.

V. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed algorithm the videos are collected and stored in the database. TIRI-DCT and Interest-point matching algorithm were separately applied to each video in the database. A fingerprint database is formed for each algorithm and the extracted fingerprints are stored. Then, videos in the database were attacked(disorted) to generate query videos. The attacks include added Gaussian noise, changes in brightness/contrast, time shift, rotation and changes in background of the video. True Positive Rate and False Positive Rate are calculated for each algorithm.

Let true positives (*TP*) be positive examples correctly labelled as positives

$$TPR=TP/P$$

False positives (*FP*) be negative examples incorrectly labelled as positives.

$$FPR=FP/N$$

F-Score is calculated to measure the accuracy of the system. If the value of the F-Score is low then it represents the poor system in terms of both robustness and discrimination. It considers both the precision *p* and the recall *r* of the test to compute the score

p is the number of correct results divided by the number of all returned results.

$$P = TP /(TP + FN)$$

and *r* is the number of correct results divided by the number of results that should have been returned.

$$R = TP /(TP + FP)$$

The F_1 score can be interpreted as a weighted average of the precision and recall, where an F_1 score reaches its best value at 1 and worst score at 0.

$$F=(1+\beta^2)\ (precision.\ Recall)\ /\ (\beta^2\ precision+recall)$$

β value is chosen as 0.5 to give precision twice the importance of recall. Following table shows the F-score for different attack parameters: noise addition, change in brightness, contrast, rotation, time shift and changes in background.

TABLE I
COMPARING TIRI-DCT AND INTEREST-POINT-MATCHING ALGORITHM

	TIRI-DCT			INTEREST-POINT-MATCHING		
Attacks	TPR(%)	FPR(%)	F-SCORE	TPR(%)	FPR(%)	F-SCORE
Noise	91.9724	8.0276	0.8542	97.7549	2.2451	0.9049
Brightness	91.9500	8.0500	0.8615	97.7523	2.2477	0.9023
Contrast	91.2927	8.7073	0.8623	97.6701	2.3299	0.9021
Time shift	91.8374	8.1626	0.8587	97.6856	2.3144	0.9058
Rotate	92.1669	7.8331	0.8583	97.7106	2.2894	0.9023
Background Changed	91.5541	8.4459	0.8582	97.4372	2.5628	0.9057
Average	91.7955	8.204	0.8585	97.6684	2.3315	0.9038

The proposed Interest-Point-matching algorithm shows the higher F-score value than the existing algorithms. It maintains higher performance for all the attacks such as noise addition, change in brightness, contrast, rotation, time shift and changes in background. The algorithm maintains a high True Positive Rate of 97.66% and low False Positive Rate of 2.33%.

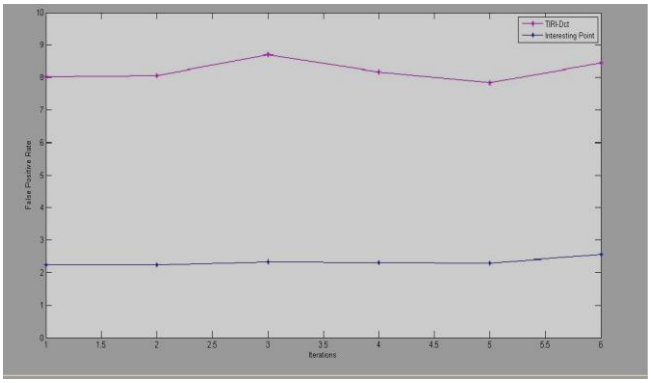

Fig .3. False Positive Rate of TIRI-DCT and Interest Point- Matching algorithm for different attack parameters from Table I

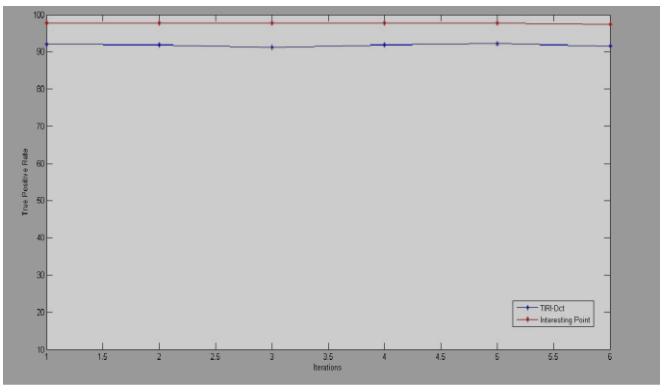

Fig .4. True Positive Rate of TIRI-DCT and Interest-Point-Matching algorithm for different attack parameters from Table I

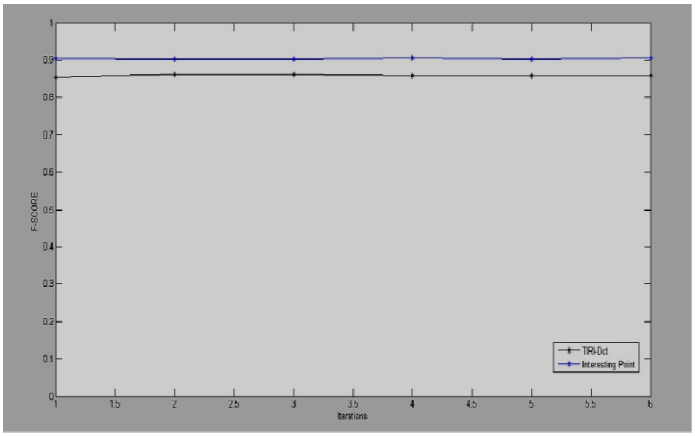

Fig.5. F-Score value of TIRI-DCT and Insert-Point-Matching algorithm for different attack parameters from Table I

VI. CONCLUSION AND FUTURE WORK

The fingerprinting system is proposed for video copy detection system.It can be used for copyright management and indexing applications.The system consists of a fingerprint extraction algorithm followed by an approximate search method. The proposed fingerprinting algorithm which is Interest point Matching Algorithm extracts robust, discriminate and compact fingerprints from videos in a fast and reliable fashion. The fingerprint extracted using this algorithm maintains a good performance for attacks such as noise addition, changes in brightness or contrast rotation, time shift, and changes in background. Two fast searching methods: Inverted file based searching and cluster based similarity search are implemented for efficient searching in the fingerprint database.

Future work includes enhancing the security to the fingerprint. It also includes the study of the performance of the system in the presence of some other attacks, such as large geometric attacks like cropping or inserting some logo. Comparing to the other fingerprinting methods our system will reduce the searching time so it improves to faster searching methods.

REFERENCES

[1] Christoph Strecha, Alexander M. Bronstein, Member, IEEE, Michael M. Bronstein, Senior Member, IEEE, and Pascal Fua, Senior Member, IEEE(jan 2012) "LDAHash: Improved Matching with Smaller Descriptors" IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 34, No. 1

[2] Cordelia Schmid, Roger Mohr and Christian Bauckhage (2000) "Evaluation of Interest Point Detectors" International Journal of Computer Vision 37(2), 151–172

[3] Coskun. B, B. Sankur, and N. Memon,(Dec. 2006) "Spatiotemporal transform based video hashing," IEEE Trans. Multimedia, vol. 8, no. 6, pp. 1190–1208

[4] Gerhard Roth, Robert Lagani`ere, Patrick Lambert, Ilias Lakhmiri, and Tarik Janati (2010) "A simple but effective approach to video copy detection" Candian conference computer and robot vision

[5] Hampapur and R. M. Bolle, Videogrep(2001) " Video copy detection using inverted file indices" IBM Research Division Thomas. J. Watson Research Center, Tech. Rep.

[6] Hye-Jeong Cho, Yeo-Song Lee, Chae-Bong Sohn, Kwang-Sue Chung, and Seoung-Jun Oh(2009) "A novel video copy detection method based on statistical analysis"

[7 Juan Chen(2010) " Detection Of Video Copies Based On Robust Descriptors" 978-1-4244-8026-5/10/$26.00 ©2010 IEEE

[8] Law-To.J, L. Chen, A. Joly, I. Laptev, O. Buisson, V. Gouet-Brunet, N. Boujemaa, and F. Stentiford,(2007) "Video copy detection: A comparative study," in Proc. ACM Int. Conf. Image and Video Retrieval, New York, NY, pp. 371–378, ACM.

[9] Malekesmaeili.M, M. Fatourechi, and R. K.Ward,(Dec. 2009) "Video copy detection using temporally informative representative images," in Proc. Int. Conf. Machine Learning and Applications , pp. 69–74.

[10] Mani Malek Esmaeili,Mehrdad Fatourechi and Rabab Kreidieh Ward(March 2011) "A Robust and Fast Video Copy Detection System Using Content-Based Fingerprinting" IEEE Transactions on Information Forensics and Security, vol 6,No1

[11] Radhakrishnan.R and C. Bauer, (Oct. 2007)"Content-based video signatures based on projections of difference images," in Proc. MMSP, pp. 341–344.

[12] Roopalakshmi.R, Ram Mohana Reddy.G(2010) "Recent Trends in Content-Based Video Copy Detection" 978-1-4244-5967-4/10/$26.00 ©2010 IEEE

[13] Shikui Wei, Yao Zhao, Member, IEEE, Ce Zhu, Senior Member, IEEE,(Jan 2011) "Frame Fusion for Video Copy Detection" IEEE Transactions On Circuits And Systems For Video Technology, Vol. 21, No. 1

[14] Swaminathan, Y. Mao, and M. Wu,(Jun. 2006.) "Robust and secure image hashing," IEEE Trans. Inf. Forensics Security, vol. 1, no. 2, pp.215–230,

[15] Zheng Cao, and Ming Zhu (Aug 2009) "An efficient video copy detection method based on video signature" Proceedings of the IEEEInternational Conference on Automation and Logistics Shenyang, China

[16] Yoshiaki Itoh , Masahiro Erokuumae, Kazunori Kojima, Masaaki Ishigame, Kazuyo Tanaka(2010) "Time-space Acoustical Feature for Fast Video Copy Detection" 978-1-4244-8112-5/10/$26.00 ©2010 IEEE

[17] Zhen Xiong and Yun Zhang ,(December. 2009) "A Novel Interest-Point-Matching Algorithm for High-Resolution Satellite Images" IEEE Trans. on Geoscience and Remote Sensing, Vol. 47, No. 12,