A Hypothesis Test based Robust Technique for Video Sequence Matching

Debabrata Dutta
Tirthapati Institution, Kolkata, India
debabratadutta2u@gmail.com

Sanjoy Kumar Saha*
CSE Dept, Jadavpur University, Kolkata, India,
sks_ju@yahoo.co.in

Bhabatosh Chanda
ECSU, Indian Statistical Institute, Kolkata, India
chanda@isical.ac.in

Abstract

Video sequence matching is the most crucial step to verify whether a video sequence has been copied from another or not. The video sequence in question is represented by a set of visual descriptors. Based on such descriptors the sequence has to be matched with the sequences present in the database. Thus, sequence matching becomes a crucial task. Furthermore, to evade the verification process, one may incorporate visual distortions in the original sequence to generate the copied version. Again, the copied sequence may be generated by taking few sampled frames or a part of the original sequence. Thus, the task of sequence matching becomes more challenging. In this work, rather than concentrating on the visual descriptors, we have focused on the technique of sequence matching that can sustain the challenges posed. We have proposed a hypothesis test based scheme for sequence matching which has inherent strength to withstand the said attacks considerably. Experimental result also indicates its capability of the scheme in doing so.

Keywords: *Video Copy Detection, Video Fingerprinting, Sequence Matching, Hypothesis Test.*

1. Introduction

Technological development has made capturing and storage of video data easier and inexpensive. Moreover, development in the arena of network and communication and increase in bandwidth has encouraged video sharing, broadcasting enormously. Availability of digital videos on various media like TV-channels, web-TV, video blogs, public video web servers has led to huge growth in video data volume. All these have an adverse effect on copyright management. The technology has enabled easy access, editing and duplicating of video data. Such activities result into violation of digital rights. Considering the huge volume of the video database, controlling the copyright of videos generated everyday has become a critical issue. But, it is the basic requirement in protecting the intellectual property right. Driven by the importance of copyright protection, a new area of research called video fingerprinting has come up. Lee et al. [1] has defined fingerprint as perceptual features for short summaries of a multimedia object. The goal of video fingerprinting is to judge whether two video have the same contents even under quality-preserving distortions like resizing, frame rate change, lossy compression [2].Video fingerprinting is also commonly referred to as video copy detection.

There are two basic approaches to address the issue of copyright detection -- watermarking and content-based copy detection. In the first approach, watermark/non-visible information is embedded into the content and later, if required, this embedded information is used for establishing the ownership. On the other hand, in content-based approach, no additional information is inserted. It is said that "Video itself is the watermark" [3]. Unique signatures (features) are derived from the content itself. Such signatures are also extracted from the questioned version of the media and are compared with those of the original media stored in the database [4-7, 3].

Performance of a copy detection scheme relies on suitable signature extraction and a sequence matching scheme. The system must survive in the presence of various distortions adopted by the copier. In this work, we

*** communicating author**

have focused on the aspect of sequence matching. The paper is organized as follows. After this introduction, section 2 presents a brief review of video copy detection techniques. Section 3 describes the proposed methodology. Experimental results are presented in section 4 and finally, concluded in section 5.

2. Past Work

Features of a video copy detection system must satisfy the properties outlined in [2]. It must be *robust* so that fingerprint of a degraded video and the original one should be similar. It should be *pair wise independent* to have different fingerprints for perceptually different fingerprints. Finally, the fingerprint must support fast search i.e. it should be search efficient.

In order to achieve the goal of a content based copy detection system, more effort have been put in designing the global/local descriptors based on spatial and/or temporal information of video data. Global descriptors are derived from the whole video sequence or from a subset of frames. Such descriptors may fail to identify if the distorted sequence under test is much shorter than the original. On the other hand, local descriptors are computed for each individual frame various features like colour histogram [8-10], luminance based descriptors [11-14], dominant colour [3] have been tried. Various gradient based features [2, 15] are also used. Joly et al. [16] considered local descriptors based on Harris detector. Wu et al.[17] have suggested trajectory based visual patterns. Temporal ordinal measurement has been proposed as global descriptors by Chen and Stentiford [18]. DCT based hash algorithm has been used by Coskun et al. [19]. Ordinal measure[20], combination of spatio-temporal information [11] also have been used as signature. Maani et al. [21] have developed local descriptors for identified regions of interest based on angular intensity variation and region geometry. Sarkar et al. [22] were motivated by the concept of summarization and keyframes of the sequence are represented by compact Fourier-Mellin transform and Scale invariant feature transform based descriptors. Su [23] et al. has relied on visual attention region based fingerprinting. Such regions are defined in a manner so that they can be taken as invariant in a content-preserving distorted video. Thus, the success of the scheme relies heavily on the identification of such regions.

The test/query video and those in the database are to be matched on the basis of extracted signature. This matching is a crucial part of a video copy detection system. A variety of schemes have been tried by the researchers.

In [11], spatio-temporal measure to compute the similarity between two video sequences has been presented and it relies on a threshold in detecting a copy. Moreover, computing the distance with all the database clips is prohibitive. The scheme presented in [2,24] also suffer from the problem of threshold selection. Wu et al. [17] had to take the burden of computing a huge similarity matrix and in hash function based scheme [19], selection of suitable hash function is difficult. Moreover, a hash function is very sensitive to the changes in the content and making the system robust against distortion is of great challenge.

Various schemes like ordinal measure based technique [20], histogram intersection of DCT frames [25], voting technique [26] has been proposed. Similarity between two sequences also have been measured by calculating the number of frames matched between two shots [27]. Again, such comparisons are to be carried out with all the sequences in the database. Shen et al. [28] proposed to compute similarity based on the volume intersection between two hyper-sphere governed by the video clips. Schemes based on indexes built over the signatures are also found [12, 29, 30].

Several keyframe based schemes have been reported. Jain et al [31] have proposed a set of keyframe (or sub-sampled frame) based sequence matching method. Similar approaches have also been reported in [32,9,33]. Various clustering based scheme [34,27] have also been tried. Frames are clustered and one or more keyframes are extracted from each cluster. Comparison is restricted to keyframes only. Maani et al [21], in their technique, selected a set of matched keyframes from the database corresponding to each keyframe in the query sequence. From the matched set of keyframes, it tried to find out continuous subsequence. If the length of such subsequence exceeds a threshold then considered as a copy. The scheme reduces computation as the final matching is restricted with in a limited set of database sequence. But, selection of threshold poses a problem. Thus, it appears from the past work that sequence matching is an important issue and it demands attention.

3. Proposed Methodology

In a video copy detection method, the task is to verify whether or not a test/query sequence is a copied version of a sequence present in the database. It has already been discussed that such a system consists of two major modules namely *extraction of signature (feature vector)* and *sequence matching*. Signatures must fulfill the diverging criteria such as discriminating capability and robustness against various geometric and signal distortions. Sequence matching module bears the responsibility of devising the match strategy and verifying the test sequence with likely originals in the database. It is evident from the past work that to achieve robustness of the detection system, the emphasis has been put mostly on the development of attack invariant features. For matching the sequences, the researchers have mostly relied on certain threshold based comparison of feature vectors. Thus, a robust matching technique capable of sustaining the attacks can further enhance the capacity of a copy detection system. In this work, we put our effort in developing the sequence matching module which will have its own inherent strength to combat the attacks adhered by copier. We have relied on hypothesis test based strategy as presented in [35] for the purpose. As the video signature is likely to be multi-dimensional, we have considered multivariate Wald-Wolfowitz run test [36] based hypothesis testing.

3.1. Multivariate Wald-Wolfowitz Test

Wald-Wolfowitz runs test is used to solve the similarity problem of non-parametric distribution of two samples. Suppose, there are two samples X and Y of size m and n respectively and the corresponding distributions are F_x and F_y. H_0, the null hypothesis to be tested and H_1, the alternative hypothesis is as follows:

H_0: X and Y are from same population, i.e. $F_x = F_y$

H_1: They are from different population, i.e. $F_x \neq F_y$

In classical Wald-Wolfowitz test, it is assumed that sample points are univariate. N = n + m observations are sorted in ascending order and the labels X or Y are assigned to them depending on the sample to which they belong. Friedman and Rafsky [37] have suggested a multivariate generalization by using the minimal spanning tree (MST) of the sample points. In this approach each sample point is considered as a node and every node is connected to the closest node (based on the similarity between their feature vectors) to form the MST of the sample points. Now if we remove all the edges connecting pair of points coming from two different samples, each subtree formed would consist of samples from one and only one population and is equivalent to a run of a univariate case. Thus number of nodes in each subtree is equivalent to run-length and R, the number of subtrees, is equivalent to number of runs. Test statistic W is defined as

$$W = \frac{R - E[R]}{\sqrt{Var[R]}}$$

$$Var[R] = \frac{2mn}{N \cdot (N-1)} \cdot \left(\frac{2mn - N}{N} + \frac{C - N + 2}{(N-2)(N-3)} \cdot (N(N-1) - 4mn + 2) \right)$$

C is the number of edge pairs in MST sharing a common node and

$$E[R] = \frac{2mn}{N} + 1$$

As W follows standard normal distribution, a critical region may be chosen for a given level of significance which signifies the maximum probability of rejecting a true H_0. If W falls within the critical region, H_0 is rejected. Physically, low value of R expresses that two samples are less interleaved in the ordered list and it leads to the interpretation that they are from different populations.

3.2. Sequence Matching

Video sequence is a collection of frames. Each frame is described by a n-dimensional feature vector. Thus, a sequence may be thought of as $\{V_i\}$, the set of feature vectors where V_i is the feature vector corresponding to i-th frame in the sequence. Let, S_t and S_d are the test sequence and a sequence from database which are to be compared. Signatures are extracted for S_t and S_d to obtain the set of feature vectors $\{V_t\}$ and $\{V_d\}$ respectively. Thus, $\{V_t\}$and $\{V_d\}$ may be thought of as two samples and hypothesis testing described in section 3.1 can be applied to verify whether they are from same population or not. If V_t and V_d belong to the same population, it is declared that the sequences are similar. Thus, even if the feature values of the elements vary to an extent, the scheme itself can handle the situation and thereby it becomes inherently robust.

As the database consists of large number of sequences, it is prohibitive to compare test sequence with each and every sequence in the database. In order to address this issue and to reduce the number of sequences, we outline the proposed scheme as follows.

- *Obtain K_d, the collection of keyframes extracted from all the in video sequences in the database.*

- *Obtain K_t, the collection of keyframes extracted from the test sequence.*
- *For each keyframe in K_t, Find the most similar one (in terms feature vector) from K_d to obtain matched keyframe set, K_m.*
- *Form a candidate sequence set, S_c by taking the sequences corresponding to the keyframes in K_m.*
- *Verify S_t only with S_d (where, $S_d \in S_c$) using multivariate Wald-Wolfowitz test.*

Further refinement of the scheme may be done based on the following consideration. In order to avoid the possible exclusion of any possible candidate sequence, instead of best matched one, few top order matched keyframes may be considered to generate K_m. It may increase the size of S_c. But the growth of size of S_c can be controlled by putting higher priority on the candidate sequence with higher number of matched keyframes. Indexing scheme may be employed to search for similar keyframes in the database to reduce the cost of matching scheme.

3.3 Extraction of Signature

It has been outlined earlier that our focused effort is on the sequence matching technique. But, in order to verify the effectiveness of the said scheme, we need to extract the signatures of the frames of a video sequence. The feature should have the discriminative power to distinguish between images. On the other hand they must not be too sensitive to minor variation. For this purpose we have relied on edge and wavelet based features described as follows.

Edge based feature is computed using the grayscale version of the images. Applying Sobel operator on the intensity image, gradient image is obtained first. As noise may give rise to additional edge pixels, to minimize such effect, the pixels with gradient magnitude higher than the average gradient are considered as the strong edge points and only such pixels are retained to compute the feature (see Figure 1). It also may be noted that edge pixels can sustain the attacks like change in brightness and contrast. The image is divided into 16 x 16 grids. Normalized count of strong edge points in the grids forms the k-dimensional feature where, k is the number of grids in which the image has been divided. Thus, it reflects the distribution of the dominant edge pixels in the image. Same could have been achieved by computing the moments. But, in that case because of the motion of the camera and/or objects, edge pixels may get shifted in the subsequent frames leading to the change in the moments. But, in our case, as long as the new position of the pixel is within the same grid, there will be no impact on the feature vector. Thus, it is less affected by such variation.

To compute the wavelet based feature we have considered the grayscale version of the image and it has been decomposed into four sub-bands (LL, LH, HL and HH) as shown in Figure 2 using 2-dimensional Haar wavelet transformation. Thus, average intensity or the low frequency component is retained in the LL sub-band and other three show the high frequency components i.e. the details as shown in Figure 3. Energy of the values in each sub-band is considered as features. Iteratively, decomposition is continued considering LL sub-image as the image. Normally, along with the energy, average intensity is also considered as the feature. But, as average intensity gets more affected by the common attacks like change in brightness and contrast, we have relied only on energy. In successive iteration, as we deal with the average image in LL band, the impact of the noise also gets reduced and enables us to cope up with some specific attacks. In our experiment, the intensity values of the gray scale image are normalized within [0, 1]. We have considered 5 levels of decompositions to obtain 20-dimensional feature vector. Finally, edge based and wavelet based features represent the signature of each frame.

(c) Images with strong edge points

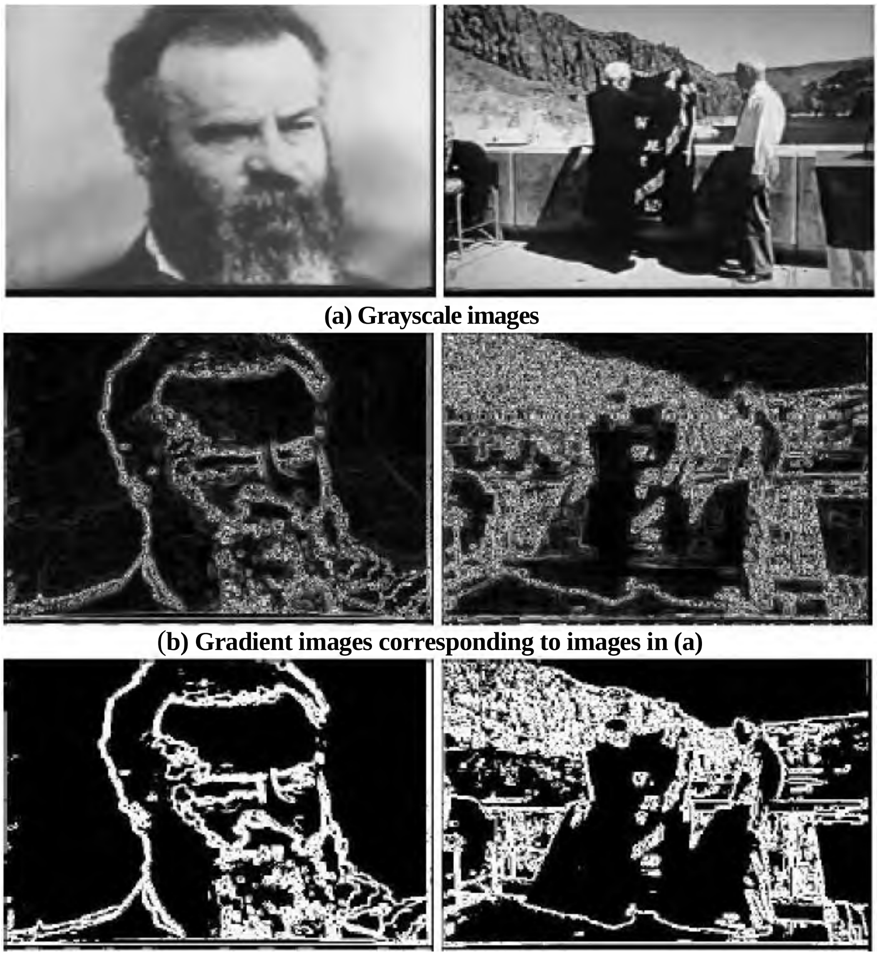

Figure 1. Detection of Strong Edge Points

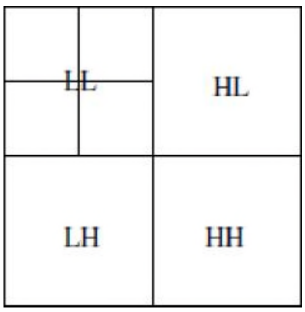

Figure 2. Wavelet Transformation.

Figure 3. Wavelet Decomposed Images: Image on the right is decomposed version of the image on the left.

4. Experimental Results and Discussion

In order to carry out the experiment we have worked with a database of approximately eight hour duration. It contains 4150 sequences which have collected mostly from TRECVID 2001, 2002 and TRECVID 2005 databases. Few other news, sports and documentary sequences also have been incorporated.

For each sequence in the database, keyframe(s) are obtained based on the methodology presented In [38] and using the features described in section 3.3. Moreover, distance between the feature vectors are computed as $dist = d_w + d_e$. Euclidean distance between wavelet features is taken as d_w. sim_e, similarity between the edge histograms is measured using Bhattacharya distance between the two and d_e is taken as $1\text{-}sim_e$. Keyframes thus obtained are stored in the database. Whenever a test sequence has to be verified, its keyframe(s) are also detected following the same strategy. In the present scope of work, the set of candidate sequences are obtained by comparing the keyframes of test sequences with those in the database in an exhaustive manner. Corresponding to each keyframe in the test sequence, the database sequence containing the best matched keyframe is included in the candidate set. Finally, test sequence is matched with those in candidate set using the hypothesis test based scheme and level of significance for the test has been considered as 0.1.

Extensive experiment has been performed to judge the performance of the proposed sequence matching scheme. Initially, no attack has been considered and each of the database sequence has been treated as the sequence. It was found that the proposed scheme was successful in detecting all the cases. In order to verify whether the scheme can successfully identify the cases of *true-false* (the cases where the sequence are truly not a copy) or not, we have considered 500 test sequences which are different from those in the databases. It has been observed that all such test sequences are correctly detected as the original.

In order to verify the robustness of the proposed detection scheme, we have considered different types of attacks that a copier may adopt to evade the possible detection. In the present scope of the work, we have considered the attacks in the form of sampling, contrast variation, brightness variation, and corruption by noise and application of colour filters. Corresponding to each sequence in the database, a set of test sequences are generated by incorporating the attacks.

For each sequence, frames are sampled randomly to generate the test sequences. Moreover, a sequence may have considerable variation so that visually it may be thought of as a collection of different sub-sequences. Sampled frames of one such sequence is shown in Figure 4(a). Test sequences are also generated focusing on one part (see Figure 4(b)) of such varying sequences and may be referred as burst dropping of frames. The proposed matching scheme relies on the interleaving pattern of the frames of a database sequence and those of a test sequence in the feature space. Hence, the absolute count of frames in the test sequence bears insignificant impact in decision making and as shown in Table 1, the proposed scheme has successfully handled the attack in the form of sampling, be it random or burst.

(a)

(b)

Figure 4. (a) Sample frames from different sub-parts of a database sequence (b) Sample frames from test sequence

While copying a sequence, copier may change the contrast or brightness. Corresponding to each sequence in the database, test sequence has been generated by changing the contrast (brightness). Such sequences have also been generated varying the contrast or brightness by 10%, 20% and 30%. Sample frames from such test sequences have been shown in Figure 5 and 6. Edge features are not much influenced the variation in contrast or brightness. But, variation in contrast leads to considerable change in the energy of wavelet sub-bands and the scheme can not withstand such variation beyond a limit. For brightness variation, the change in sub-band energy is less significant than that due to contrast variation. But, due to brightness shift runs get affected. Thus, it seems that features under consideration may be influenced by such attack. Still the performance of the scheme as shown in Table 1 is good enough. Furthermore, it may be noted that wide variation affects the quality of the copy heavily and degradation makes it of no use. Thus, such attacks can only be adhered to an extent.

(a)

(b)

(c)

(t

Figure 5. (a) Sample frames from a database sequence (b), (c) and (d) Frames after 10%, 20% and 30% contrast enhancement respectively.

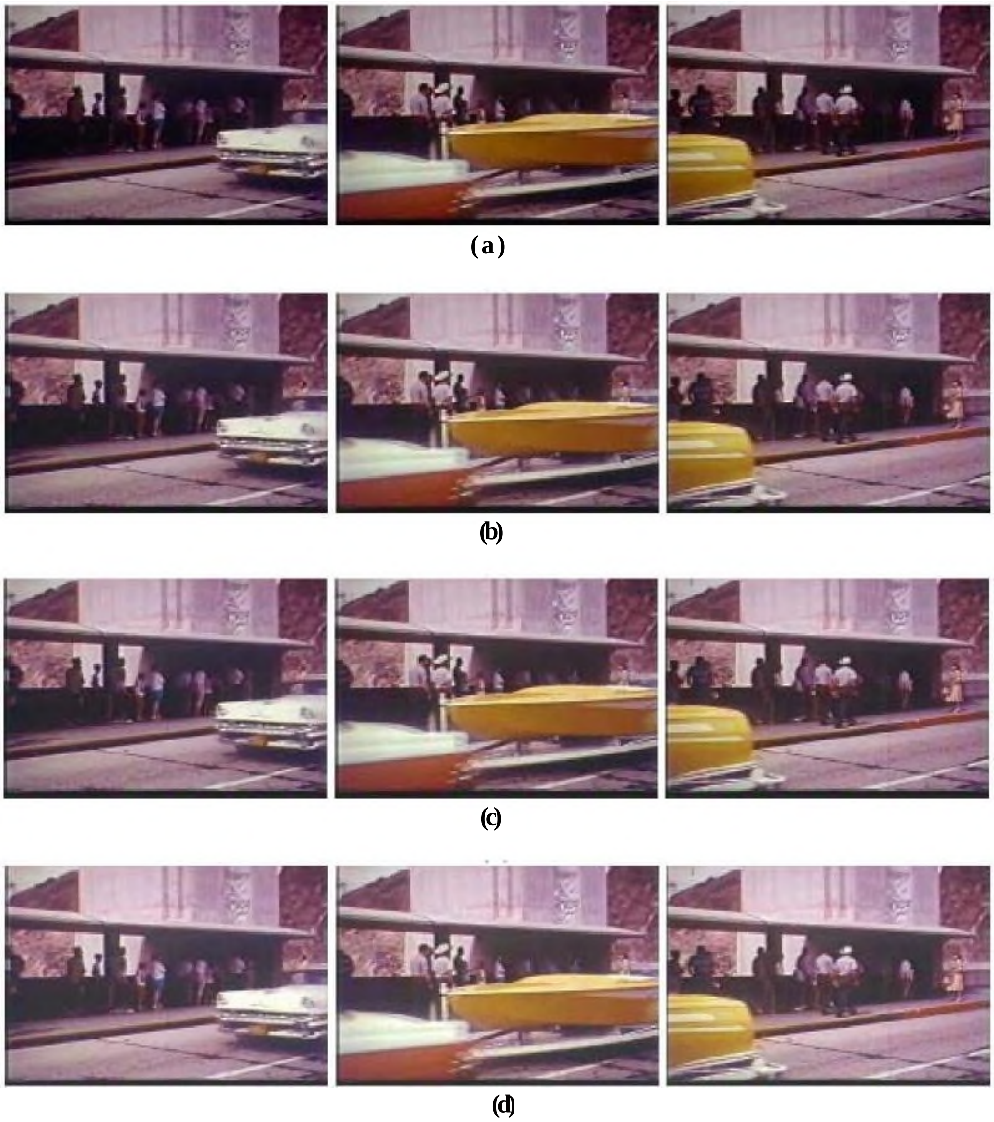

Figure 6. (a) Sample frames from a database sequence (b), (c) and (d) Frames with brightness increased by 10%, 20% and 30% respectively

(a)

(d)

Figure 7. (a) A sample frame from a database sequence (b) Noise corrupted frames.

Table 1. Copy Detection Performance

Type Of Attack	Description	Correct detection (in %)
No Attack		100.00
Sampling		100.00
Variation in Contrast	10% enhanced	87.73
	20% enhanced	86.62
	30% enhanced	84.54
Variation in Brightness	10% enhanced	93.07
	20% enhanced	82.32
	30% enhanced	81.19
Noise	10% pixel affected	89.89
	20% pixel affected	88.30
	30% pixel affected	83.10
Colour Filtering	Red	83.16
	Green	86.56
	Blue	77.82

(a)

(b)

(c)

(t

Figure 8. (a) Sample frames from a database sequence (b), (c), (d) Frames obtained applying red, green and blue filter respectively.

As shown in Figure 7, test sequences have been generated by adding random noise in the sampled frame of original sequence. The test sequences have been obtained by corrupting 10%, 20%, 30% pixels in each frame. Intensity values of the affected pixels are modified within a wide range of ±10 to ±30. Table 1 show that the performance of the proposed scheme in presence of noise is also considerably well. Addition of noise increases the spread of feature value but runs in hypothesis test are less affected. As a result, proposed scheme can identify the copy in most of the cases.

Test sequences are generated by applying the red, green and blue filters for each sequence in the database. Sample frames for a sequence has been shown in Figure 8. For such filtered version, a mixed result has been achieved. Details of a particular colour is lost leading to significance change in energy of the sub-bands. Thus, if the presence of a colour is insignificant in a sequence then the corresponding filtered version does not provide much clue about the original one and the scheme may fail.

Table 1 shows the performance of the proposed scheme under various attacks. The performance is quite satisfactory but has certain cases of failure also. More devoted effort is required to judge the performance against such attacks and it has to be dependent on the features/signatures used to represent the frames. In this work, we have concentrated our effort in designing the sequence matching methodology without looking for attack invariant features. In general, the experiment has established that the proposed methodology has the strong potential in addressing the issue of sequence matching even under the possible attacks.

5. Conclusion

In this work, we have presented a novel scheme for video copy detection. By comparing the keyframes of test sequence and database sequences, a subset of database sequence is taken as candidate set. Finally, we have proposed a multivariate Wald-Wolfowitz run based hypothesis testing scheme to verify whether the test sequence and any sequence of the candidate set are same or not. The sequence matching scheme has inherent strength to cope up with the commonly adopted visual attacks like corruption by noise, contrast and brightness variation, colour filtering and also the attack in the form of random/burst dropping of frames in a sequence. The hypothesis test arranges the elements in the test sequence and a database sequence based on their neighborhood in the feature space. Thus, as long as the attack does not change the features to an extent to change the arrangement heavily, the scheme survives. Experimental result also justifies the robustness of the proposed sequence matching scheme. In future, further work may be carried out to develop attack invariant features which are also capable of handling other attacks like cropping, logo insertion. A suitable indexing scheme may also be used to speed up the detection process.

Acknowledgement

This work is partially supported by the facilities created under DST-PURSE program in Computer Science and Engineering Department of Jadavpur University, India.

References

[1] Seo, J.S., Jin, M.,Lee, S., Jang, D., Lee, S.J., D.Yoo, C.: Audio Fingurepriting based on normalized spectral subband centroids. In:Proc. ICASSP. (2005) 213 -216.

[2] Lee, S., Yoo,C.D. :Video fingerprinting based on centroids of gradient orientations. In Proc. ICASSP. (2006) 401-404.

[3] Hampapur, A., Bolle, R. : Comparison of sequence matching techniques for video copy detection. In Proc. Intl. Conf. on Multimedia and Expo. (2001) 188-192.

[4] Chang, E.Y., Woang, J..Z., Li, C., Wienderhold,G.: Rime: A replicated image detactor for the world-wide-web. In: Proc.SPIE Media Storage and Archiving Systems III (1998) 68-71.

[5] Kim, C. : Ordinal Measure of DCT Coefficients for Image Correspondence and Its Application to Copy Detection. In Proc. for SPIE Storage and Retrieval for Media Databases, (2003) 199-210.

[6] Kim. ,C., content-based Image Copy Detection. In Signal Process. Image Comm. 18, No. 3, (2003169-184.

[7] Mohon., R.: Video Sequence Matching. In Proc. ICASSP,(2001) 3697-3700.

[8] Ferman, A.M., Tekalp,A.M., Mehrotra, R. : Robust color histogram descriptor for video segment retrival and identification. IEEE Trans. On IP 11(5) (2002) 497 – 508.

[9] Cheung, S-C. S., and Zakhor., A.,: Efficient Video Similarity Measurement with Video Signature. In IEEE Trans. CSVT,13 No. 1,(2003) 59-74.

[10] Li, Y., and Jin, L., and Zhou, X.,: Video Matching Using Binary Signature. In Proc. Intl. Symp. on Intelligent Signal Processing and Comm. Systems, (2005) 317-320.

[11] Kim, C., and Vasudev, B.,.: Spatiotemporal Sequence Matching for Efficient Video Copy Detection. In IEEE Trans. on CSVT 15, No. 1,(2005) 127-132.

[12] Oostveen, J., and Kalker, T., and Haitsma, J.,: Feature Extraction and a Database Strategy for Video Fingerprinting. In Proc. VISUAL, (2002)117-128.

[13] Hua, X-S., and Chen, X., and Zhang., H-J. ,.: Robust Video Signature based on Ordinal Measure. In Proc. ICIP,(2004)685-688.

[14] Radhakrishnan, R., Bauer, C.: Robust video fingureprints based on subspace embedding. In Proc. ICASSP (2008) 2245- 2248.

[15] Lowe., D. G.,: Object Recognition From Local Scale Invariant Features. In Proc. ICCV, (1999) 1150-1157.

[16] Joly, A., and Buisson, O., and Frelicot, C.,: Content-Based Copy Retrieval Using Distortion-Based Probabilistic Similarity Search. In IEEE Trans. Multimedia 9 No.2, (2007) 293-306.

[17] Wu, X., and Zhang, Y., and Wu, Y. and Guo, J., and Li., J.: Invariat Visual Patterns for Video Copy Detection. In Proc. ICPR, (2008)1-4.

[18] Chen., L., and Chua., T.S.,: A match and tiling approach to content-based video retrieval. In Proc. Intl. Conf. on Multimedia and Expo.(2001).

[19] Coskun, B., and Sankur, B., and Memon., N. ,: Spatio-Temporal Transform Based Video Hashing. In IEEE Trans. Multimedia,8 No. 6 (2006)1190-1208.

[20] Bhat, D. N. , and Nayar., S. K. ,: Ordinal Measures for Visual Correspondence. IEEE Trans. PAMI, 20 No. 4, (1996) 415-423.

[21] Maani, E., and Tsaftaris, S. A., and Katsaggelos, A. K.,: Local Feature Extraction for Video Copy Detection. In Proc. ICIP, (2008) 1716-1719.

[22] Sarkar, A., Ghosh, P., Moxley, E.,Manjunath, B.S., : video fingureprinting: Features For duplicate and similar video detection and query base video retrieval. SPIE – Multimedia Content Access: Algorithms and Systems II (2008)

[23] Su, X., Huang, T., Gao, W.,: Robust video fingureprinting based on visual attention regions. In Proc. ICASSP.(2009).

[24] Yan, Y., and Ooi., B. C. , and Zhou, A.,: Continuous Content-Based Copy Detection over Streaming Videos. In Proc. Intl. Conf. on Data Engg., (2008)853-862.

[25] Naphade, M., and Yeung, M., and Yeo, B.,: A Novel Scheme for Fast and Efficient Video Sequence Matching using Compact Signatures. In Proc. SPIE Conf. Storage and Retrieval for Media Databases, volume 3972, (2000) 564-572.

[26] Law-To, J., Buisson, O., Gouet-Brunet, V., Boujemma, N., : Robust voting algorithm based on labels of behavior for video copy detection. In Proc. ACM Multimedia.(2006)

[27] Chen, L., and.Chua, T.S,.: A match and tiling approach to content-based video retrieval. In Proc. Intl. Conf. on Multimedia and Expo. (2001)

[28] Shen, H., and Ooi, B. C.. and Zhou, X.,: Towards Effective Indexing for Very Large Video Sequence Database. In Proc. SIGMOD, (2005) 730-741.

[29] Schmid, C., and Mohr, R.,: Local Grayvalue Invariants for Image Retrieval. In IEEE Trans. PAMI 19 No.5, (1997) 530-535.

[30] Zhao, H. V., and Wu, M. , and Wang, Z. J., and Liu, K. J. R.,: Forensic Analysis of Nonlinear Collusion Attacks for Multimedia Fingerprinting. In IEEE Trans. IP vol 14 No.5, (2005) 646-661.

[31] Jain, A. K., and Vailaya, A. , and Xiong., W., : Query by Clip. In Multimedia System Journal 7,No. 5, (1999) 369-384.

[32] Chang, S-F. S., and Chen, W. , and Meng, H. J., and Sundaram, H. , and Zhong, D.,: Videoq: An Automated Content Based Video Search System Using Visual Cues. ACM Multimedia. (1997) 313-324.

[33] Sze, K. W., and Lam, K. M., and Qiu, G.,: A New Keyframe Representation for Video Segment Retrieval. In IEEE Trans. CSVT vol 15 No. 9, (2005) 1148-1155.

[34] Guil, N., and Gonzalez-Linares, J. M., and Cozar, J. R. , and Zapata, E. L.,: A Clustering Technique for Video Copy Detection. In Proc. Iberian Conf. on Pattern Recog. and Image Analysis, (2007) 415-458.

[35] Dutta, D., Saha, S.K., Chandra, B.,: Video copy detection : Sequence matching using Hypothesis test. In Proc. Intl. Conf. on Ubiquitous Computing And Multimedia Applications. (2010) 499-508).

[36] Wald, A., and Wolfowitz., J. ,: On A Test Whether Two Samples are from the Same Population. In Annals of Mathematical Statistics,volume 11, (1940) 147-162.

[37] Friedman, J. H. , and Rafsky, L. C.,: Multivariate Generalizations of the Wald-Wolfowitz and Smirnov Two-Sample Tests. In The Annals of Statistics 7(4), (1979) 697-717.

[38] Mohanta, P. P., and Saha, S. K., and Chanda, B.,: Detection of Representative Frames of a Shot using Multivariate Wald-Wolfowitz Test. In Proc. ICPR, Florida, USA (2008).

Authors

Debabrata Dutta: Received his B.Sc and M.Sc. degree in Computer Science from Calcutta University and Vidyasagar University, in 2004 and 2006 respectively. Currently, he is doing research project in the Computer Science Department,Jadavpur University in Video Processing. His areas of interests are Image Processing, Pattern Recognision, Astronomical Image Processing, Video Processing and Classification.

Sanjoy Kumar Saha: Received his B.E. and M.E. Degree in Electronics and Tele-communication Engineering from Jadavpur University, West Bengal, India in 1990 and 1992 respectively and obtained his PhD from Bengal Engineering and Science University, West Bengal, India in 2006. Currently, he is working as a Reader in Computer Science and Engineering Department of Jadavpur University. His research interests are in the area of Image Processing, Video Processing, Multimedia Data Retrieval and Pattern Recognition.

Bhabatosh Chanda: Received B.E. in Electronics and Telecommunication Engineering and PhD in Electrical Engineering from University of Calcutta in 1979 and 1988 respectively. He has received `Young Scientist Medal' of Indian National Science Academy in 1989, `Computer Engineering Division Medal' of the Institution of Engineers (India) in 1998, 'Vikram Sarabhai Research Award in 2002, and IETE-Ram Lal Wadhwa Gold medal in 2007. He is also recipient of UN fellowship, UNESCO-INRIA fellowship and Diamond Jubilee fellowship of National Academy of Science, India. He is fellow of

Institute of Electronics and Telecommunication Engineers (FIETE), of National Academy of Science, India (FNASc.), of Indian National Academy of Engineering (FNAE) and of International Association of Pattern Recognition (FIAPR). His research interest includes Image and video Processing, Pattern Recognition, Computer Vision and Mathematical Morphology. He is a Professor in Indian Statistical Institute, Kolkata, India.