An Efficient Rebroadcast Video Detection Algorithm Using Local Binary Patterns and Colour Moments

Rudrani Bhadra
Department of Computer Science and Engineering
National Institute of Technology, Delhi
Delhi, India-110040
rudranistar@gmail.com

Aayushi Pradeep Kumar
Department of Computer Science and Engineering
National Institute of Technology, Delhi
Delhi, India-110040
aayushi.p.kumar@gmail,com

Preeti Mehta
Department of Electronics and Communication Engineering
National Institute of Technology, Delhi
Delhi, India-110040
preetimehta@nitdelhi,ac.in

Dr. Sushila Maheshwar
Department of Computer science and Engineering
National Institute of Technology, Delhi
Delhi, India-110040
sushila@nitdelhi.ac.in

***Abstract*—Nowadays, it is extremely easy for anyone to recapture a high quality video from LCD screens and from projectors. In the modern age, there are many ways of video recapturing and video tampering. Video piracy is widespread which result in huge revenue loss for businesses and manufacturers. This is one reason why, digital video forensics has become very important. Recaptured video detection is one part of that. It can applied for detecting illegal video copies in professional cinema and home entertainment. In the paper, we have used features such as local binary patterns, colour moments and other statistical measures in order to differentiate between an originally captured video from a recaptured video. This experiment has been carried out on a dataset that was created by us. The experimental results show that the proposed method is very effective in video recapture detection for our database.**

Keywords—Local binary patterns, colour moments

I. Introduction

The increasing popularity of social media platforms has caused videos to be uploaded, shared and circulated on the internet. Videos are gaining more popularity than photos. As internet is at the ease of people, the abundance of internet has increased the popularity of videos on all social media. Nowadays, movies and entertainment play a very important role in our society. It helps us understand other cultures and also helps us contribute in globalising. Videos are also a common source of information. Newspapers are drastically being replaced by video as a source of information. Movies are a main contributor to the video production. Bollywood being a vast film industry based in India, produces approximately 1500 movies per year. The Indian film industry looses approximately 1.2 billion dollars each year due to online video piracy. The US film industry also is a victim to video piracy. It is estimated that 13.7% of losses are incurred by the US distributors of theoretical films due to piracy worldwide [7]. Hence piracy is one of the biggest threats to motion picture industries. Numerous film industries has suffered worldwide losses due to film piracy. A lot of financial wastage occurs due to the same. A rebroadcast attack is an attack where changes are made in an image. It is against the forensic methods which are made to differentiate between original images and edited images [6]. Mobile phones are widely used to capture films in movie theatres and to acquire good quality recaptured videos and they are sold on a black market and illegally distributed over the Internet. Mobile phones are used for piracy as it is more preferred than digital camcorders due to the ease in availability, low cost and high resolution. Illegal video copies of movies which are newly released in theatres are a huge source of fake and unauthorised copies of movies that are distributed all over the internet. General public is also capable of producing fine quality videos by using video editing softwares and making changes in the videos such as correcting colour and reducing noise and blurriness. Video forensics, such as the applications of video recapture detection is widely used in piracy. In such applications it can be used to check the authenticity of a video. In figure I, we can see a basic implementation of the original video being recaptured.

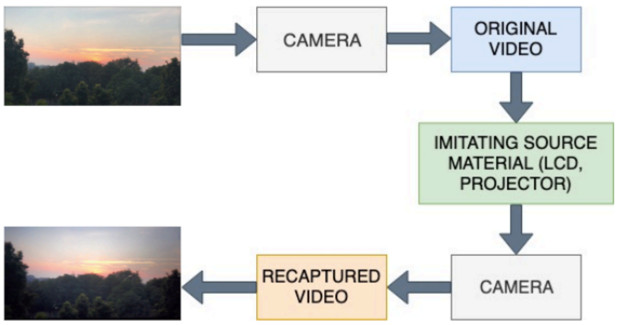

Figure 1
Block diagram for the implementation of Rebroadcast Attack.

II. Related Work

A. Literature Review

In [1], the problem of facial spoofing detection against replay attacks based on the analysis of aliasing in spoof face videos is addressed. The aliasing patterns that commonly appears during the recapture of video or photo replays on a screen in different channels and regions is analysed. Multi-scale LBP and DSIFT features are used to represent the characteristics of aliasing patterns that differentiate a replayed fake face from a real face. In [2], they analyse the differences between genuine face images and spoof images, and propose to extract three types of features: specular reflection ratio, Hue channel distribution and blurriness, to determine whether a face image is captured from a real face or not. In [3], a face spoofing detection scheme based on colour texture Markov feature (CTMF) and support vector machine recursive feature elimination (SVM-RFE) is proposed. In this paper, the adjacent facial pixels discrepancy between the real and the fake face is analysed, and texture information between the colour channels is fully considered. In addition, SVM-RFE is used to reduce the feature dimension and makes it suitable and effective for real-time detection. In [8], the noise features produced by a recaptured video is used to differentiate between a fake and a real face. To capture the noise and generate a compact representation, Fourier spectrum is used which is followed by the calculation of of the visual rhythm and extraction of the grey-level co-occurrence matrices, which are used as feature descriptors in this method. The compact representation achieved by using visual rhythms as texture maps has positive impacts on the implementation of the method.

III. Methodology

The features used in this paper are multi-scale local binary patterns (MLBP), colour local binary patterns (CLBP), colour moments, correlation and energy coefficients.

A. Dataset Creation

The data set consists of 80 originally captured videos and 320 recaptured videos. In figure II, we can see still frames from original and recaptured videos.

The following cameras were used to recapture the videos:

TABLE I
Cameras USED TO CREATE DATABASE

Camera	Model	Frame/	Resolution	Sensor
C1	Lenovo P2	30fps	1080p	13 mp
C2	Redmi note 4	30fps	1080p	13 mp
C3	Redmi note 5	30fps	1080p	13 mp
C4	iPhone 7	30fps	1080p	12 mp
C5	Samsung S6	30fps	1080p	16 mp
C6	Oppo realme	30fps	1080p	12 mp
C7	Moto G 2015	30fps	1080p	13 mp
C8	Mi Max 2	30fps	720p	12mp

The following cameras and screens have been used for recapturing the videos.

TABLE II
CAMERAS USED TO RECAPTURE THE ORIGINAL VIDEOS

Camera	Model	Frame/Second	Resolutio	Sensor
C9	Nikon	30fps	1080p	24.2 mp
C10	Sony Lens	30fps	1080p	20.4 mp

TABLE III
Screens USED TO CREATE DATABASE

Screen (S.No)	Model	Frame/Second	Resolution
S1	Acer Predator	60fps	1920 x 1080 px
S2	iPad Pro 17	120fps	2224 x 1668 px

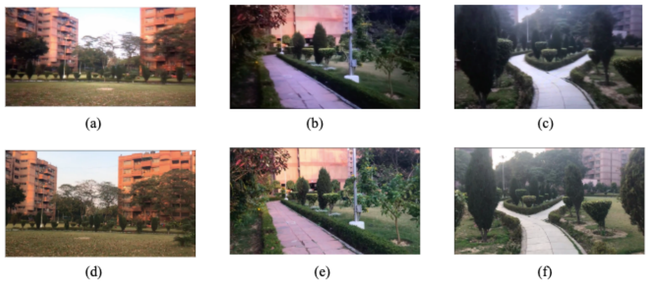

FIGURE 2
(A), (B), (C): IMAGE FRAMES OF RECAPTURED VIDEOS. FIG (D), (E), (F): CORRESPONDING IMAGE FRAMES OF THE ORIGINAL VIDEOS

B. Algorithms

Algorithm 1: Texture Feature Extraction

Input: *V* is a set of original and recaptured videos, *F* is a frame of a video, *P* is the number of neighbour elements, *R* is the radius about centre element.

Output: $f_{8,1}$,$f_{16,2}$,$f_{24,3}$,$f_{24,4}$ are the multi-scale LBP parameters, *r1,g1,b1* are the RGB LBP parameters, *r',g',b'* are the normalised LBP parameters and *o1,o2* and *o3* are the opponent LBP parameters.

for *each Video V* **do**

extract F frames

crop frame using imcrop(F);

compute r, g, b frames;

for each *r, g, b frame* **do**

compute RGB_LBP(r), RGB_LBP(g), RGB_LBP(b); //RGB_LBP is function for rgb lbp feature extraction

compute nRGB_LBP(r), nRGB_LBP(g), nRGB_LBP(b); //nRGB_LBP is function for normalized rgb lbp feature extraction

compute O_LBP(r), O_LBP(g), O_LBP(b); //O_LBP is function for opponent lbp feature extraction

end

compute rgb2gray(F)

for each *(P,R)* **do**

compute extractLBPfeatures(F,P,R);

end

end

Algorithm 2: Colour Moments

Input: *V* is a set of original and recaptured videos, *F* is a frame of a video.

Output: *M* is the mean, *SD* is the standard deviation, *Corr* is the correlation and *E* is the energy ratio.

for each *Video V* **do**

extract F frames

crop frame using imcrop(F);

compute r, g, b frames;

for each *r, g, b frame* **do**

compute M, SD values;

compute Corr(r,g), Corr(g,b), Corr(b,r), E1(g,b), E2(g,r), E3(b,r);

end

end

C. Multi-Scale LBP

The LBP operator first thresholds a 3×3 neighbourhood. It compares the value of the centre pixel to its immediate 8 neighbours and assigns a value of '0' if it is greater and a value of '1' if it is lesser than its neighbouring pixel. A local binary pattern is then formed which is interpreted as a binary number. In the same manner, LBP codes are calculated for each pixel and the LBP operator is formed by computing a histogram based on these codes. Then it is used for analysing the texture of an image or a region. The fact that the LBP operator covers a small neighbourhood area is its biggest limitation. Features that are calculated in a local 3×3 neighbourhood cannot capture much local information. The operator is not very robust against local changes in the texture of an image which is caused by changing viewpoints or illumination directions. Therefore, MLBP is used to increase the power of the original LBP operator. A multi-scale LBP is computed by extracting a number of LBP codes for each pixel with different P and R values. The MLBP features are calculated as [4]:

$$f_{MLBP}(I)=\{LBP_{P,R}\}_{(P,R)\in\{(8,1),(16,2),(24,3),(24,4)\}} \quad (1)$$

where P and R are the sampling points and the radii respectively. $LBP_{P,R}$ has the formula of a single scale LBP ,

$$LBP_{P,R} = \sum_{p=0}^{P-1} sig(I(p) - I(c)) \quad (2)$$

where *I(c)* and *I(p)* are the intensities of the current pixel c and the sampling point p, respectively. The MLBP features from individual regions are concatenated together to form a histogram.

D. Colour LBP

Here, colour information is introduced into the original LBP operator to increase its photometric invariance properties of dealing with different kinds of illumination changes. Two colour LBP operators are used as follows [5]:

RGB-LBP- This LBP operator is obtained by calculating LBP over all three channels of the RGB colour space independently and then concatenated.

nRGB-LBP- This is a normalised operator which is obtained by calculating LBP for both R' and G' channels of the normalised RGB colour space (Since R'+G'+B'=1, B' channel is not calculated):

$$\begin{aligned} R' &= R/(R+G+B) \\ G' &= G/(R+G+B) \\ B' &= B/(R+G+B) \end{aligned} \quad (3)$$

Identify applicable funding agency here. If none, delete this text box.

R' and G' channels are scale invariant due to normalisation, which make this operator invariant to light intensity change.

E. Opponent LBP

This LBP operator is obtained by calculating LBP over all three channels of the opponent colour space [5].

$$\begin{aligned} A &= (R- G)/\sqrt{2} \\ B &= (R+G-2B)/\sqrt{6} \\ C &= (R+G+B)/\sqrt{3} \end{aligned} \quad (4)$$

A and B channels are invariant to light intensity shift due to the subtraction. C channel represses the intensity information and does not have invariance properties.

D. Colour Moments

Even if videos are recaptured by using high quality cameras and equipments, colours of a finely recaptured video can become slightly different from its original video. Since colour moments encode both shape and colour information, they are a good feature to use. In this case, the mean and standard deviation of each video frame are calculated.

E. Other statistical measures

We have also calculated the correlation coefficients and energy ratios of the image.

The correlation coefficients between colour channels are calculated by the following formula:

$$c = \frac{\sum_{m=1}^{M}\sum_{n=1}^{N}(A-A')(B-B')}{\sqrt{(\sum_{m=1}^{M}\sum_{n=1}^{N}(A-A')^2)(\sum_{m=1}^{M}\sum_{n=1}^{N}(B-B')^2)}} \quad (5)$$

where, M = total number of rows of image

N = total number of columns of image

A = first input array

B = second input array

A' = mean of array A

B' = mean of array B

The correlation coefficients are calculated between the R and G channels, the G and B channels and the B and R channels.

The energy ratios for the RGB channel are calculated as:

$$\begin{aligned} E_1 &= \Sigma G^2/\Sigma B^2 \\ E_2 &= \Sigma G^2/\Sigma R^2 \\ E_3 &= \Sigma B^2/\Sigma R^2 \end{aligned} \quad (6)$$

F. Classification

After all the types of features are computed, we train a SVM classifier with a RBF kernel.

IV. Experimental Results

To calculate the effectiveness and accuracy of the method, we have made an original dataset and a recaptured dataset. For each video, the first hundred frames are extracted and feature extractions are perfumed on them. The algorithm is implemented for each of the frame sizes- 32, 64, 128 and 512. The accuracies are calculated by taking each feature independently for each frame size and by taking all features together. In table IV,V,VI,VII and VIII we can see the accuracy, recall and precision compared for different features.

TABLE V
PERFORMANCE COMPARISON (%) OF MLBP

Frame size	Accuracy	Recall	Precision
32	87.7	94.6	87.5
64	96.1	99	95.3
128	98.9	99.4	98.9
512	99.6	99.8	99.6

TABLE VI
PERFORMANCE COMPARISON (%) OF COLOUR LBP

Frame size	Accuracy	Recall	Precision
32	95.5	96.1	94.9
64	97.4	95.2	95.1
128	98.2	97.0	94.3
512	99.1	98.1	98.6

TABLE VII
PERFORMANCE COMPARISON (%) OF OLBP

Frame size	Accuracy	Recall	Precision
32	94.6	90.4	98.6
64	95.2	92.3	98.7
128	96.8	94.9	98.6
512	97.5	95	100

TABLE VIII
PERFORMANCE COMPARISON (%) OF COLOUR MOMENTS

Frame size	Accuracy	Recall	Precision
32	90	95.9	89.5
64	94.5	97.9	94
128	98	99	97.9
512	99.4	99.8	99.3

TABLE IV
PERFORMANCE COMPARISON (%) OF ALL FUNCTIONS

Frame size	Accuracy	Recall	Precision
32	93.1	94.8	95.5
64	97.7	97.9	97.4
128	98.6	98.3	97.7
512	99.85	100	99.7

TABLE IX
PERFORMANCE COMPARISON (%) ON DATASET

Frame Size	Accuracy (%)				
	MLBP	***CLBP***	***OLBP***	***Mean, SD, Correlation, Energy***	***All Combined***
32	87.7	95.5	94.6	90.1	95.1
64	96.1	97.4	95.2	94.5	97.7
128	98.9	98.2	96.8	98.0	98.6
512	99.6	99.1	97.5	99.4	99.5

V. CONCLUSION

In this paper, we have extracted several features of both original and recaptured videos to address the problem of video recapture detection. Multi-scale LBP, colour LBP, opponent LBP, colour moments (mean and standard deviation) and statistical measures such as correlation coefficients and energy ratios are used to classify real and fake videos. Experimental results show the effectiveness of the proposed method. We find that accuracy increases with increase in frame size. It can be seen from Table IX that accuracy increases with the increase in frame size. Out of all the features taken individually, MLBP gives the highest accuracy of 99.6%. The highest accuracy achieved for all functions combined is 99.5%.

ACKNOWLEDGMENT

We would like to thank Dr. Sushila Maheshkar and Preeti Mehta Ma'am for their constant support and motivation in this project.

REFERENCES

1. Patel, K., Hu Han, H, Anil. K. Jain, A.K., Greg Ott, G., "Live Face Video vs. Spoof Face Video: Use of Moir´e Patterns to Detect Replay Video Attacks, " in IEEE, ICB 2015, pp. 98-105.
2. Luan, X., Wang, H., Ou, W., Liu, L., "Face Liveness Detection with Recaptured Feature Extraction, " International Conference on Security, Pattern Analysis, and Cybernetics (SPAC), 2017, pp. 429-432.
3. Zhanga, L.B., Penga,⋆, F., Qina, L., Long, M., "Face spoofing detection based on color texture Markov feature and support vector machine recursive feature elimination, " Journal of Visual Communication and Image Representation 51 (2018) , Jan. 2018, pp. 56-69.
4. Maenpaa,T., Pietikainen,M., "Multi-Scale Binary Patterns for Texture Analysis," unpublished.
5. ZHU,C., BICHOT, C.E., CHEN,L., " Multi-scale Color Local Binary Patterns for Visual Object Classes Recognition," 2010 International Conference on Pattern Recognition, 2010 IEEE, pp. 3065-3068.
6. Fan W., Agarwal S., and Farid H., "Rebroadcast Attacks: Defenses, Reattacks, and Redefenses" 2018 26th European Signal Processing Conference (EUSIPCO), pp. 1
7. Waterman D., Wook Ji S., Rochet L.,"Enforcement and control of piracy, coping and sharing in the movie industry." August 2007 International Conference on Pattern Recognition, IEEE
8. Pinto, A.S., Pedrini, H., Schwartz, W.R., Rocha, A., "Video-Based Face Spoofing Detection through Visual Rhythm Analysis," 2012 IEEE, pp. 221-228.