Detection of Frame Duplication Type of Forgery in Digital Video Using Sub-block Based Features

Vivek Kumar Singh(✉), Pallav Pant, and Ramesh Chandra Tripathi

Department of Information Technology, Indian Institute of Information Technology,
Allahabad India
vivekkr.singh@hotmail.com, pallav.pant@gmail.com,
rctripathi@iiita.ac.in

Abstract. With the easy availability and operability of video editing tools, any video could be edited in short span of time. Sometimes, these modifications change the actual meaning of targeted video. Hence, before making any judgment and opinion about such multimedia contents, it is necessary to verify their genuineness. A video can be tampered by various different attempts. Each different attempt derives a new type of forgery in videos. Among various types of attack on video, frame duplication is a common type of attack. Frames are duplicated and pasted into same video in order to either hide or add false information. We propose Sub Blocked based features to detect frame duplication. The experimental results show higher accuracy that not only detects but also localize duplicated frames as well.

Keywords: Frame duplication · Image forensic · Sub-blocking method · Video forgery · Correlation

1 Introduction

In the recent years, the availability of low cost and more interactive digital multimedia hardware such as web cam, digital camera, surveillance camera and mobile phones etc. have made it easy to capture instances at any time. With the proliferation of such digital contents, multimedia editing software (Adobe Photoshop, Avid, Audacity etc.) allow manipulations in it even with little effort. These manipulations/modifications can be performed perfectly such that it looks like original content. As a result, digital contents should not be blindly accepted. Due to potential alteration, it has become sparingly difficult to judge the authenticity of a given multimedia content by naked eye. Thus, there is an urgent need of such a tool that can assure the authenticity and originality of altered contents.

A lot work has already been reported in image forgery detection. In recent years, many cases regarding sting of famous personalities (political, actor or social activist) are reported and telecasted by television media in India. These stings (Video/Audio) affect beliefs of viewers and change their psychology. A long manual procedure in examination of such controversial video or audio tapes cannot return faith of innocent people. So there is an urgent need of such tools which can check the originality and authenticity of such contents.

J.I. James and F. Breitinger (Eds.): ICDF2C 2015, LNICST 157, pp. 29–38, 2015.
DOI: 10.1007/978-3-319-25512-5_3

Therefore, it is necessary to develop tools to detect any type of forgery if exists. In present work, we have focused a particular type of video forgery detection. Before getting into video forgery detection techniques, a small description of digital video is given below.

"Videos refer to pictorial (Visual) information, including still images and time-varying images. A still image is a spatial distribution of intensity that is constant with respect to time. In time varying image (video) visual pattern changes image by image."

Video is a sequence of images (frames). Fraudsters attempt to change the information contained within these sequence of frames thus changing the content of the original video. These changes can either be in spatial information within a frame in temporal information between two frames. Thus, video forgeries can be of various types depending upon the way in which information is tampered in a video sequence.

A most popular type of forgery is frame duplication. In such forgery, some sequence of frames are copied and pasted elsewhere in the same video sequence. This type of forgery is performed either to hide any particular information or to insert any false information in the video sequence. Duplicating a group of frame in a video sequence is an easy task. As a result of frame duplication, particular information can be hide as shown in Fig. 1.

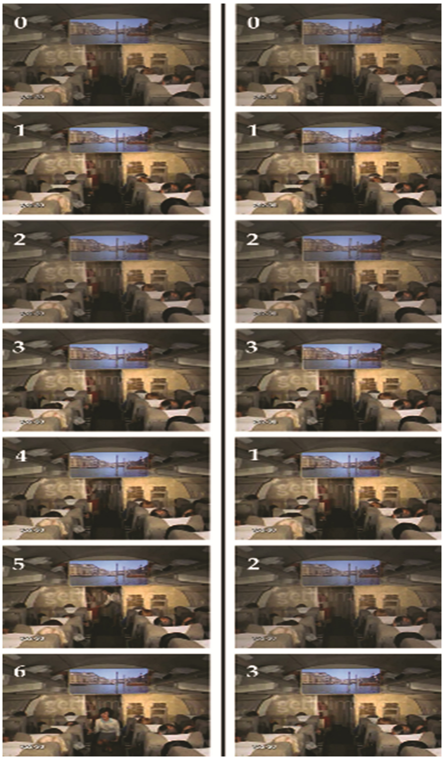


Fig. 1. Example of frame duplication [1]

Another type of duplication is to duplicate a specific region in the video sequence. This type of tampering can duplicate a motion of the same object elsewhere in the same video sequence.

Our focus is to detect frame duplication type of forgery. Various researchers have already contributed to detect frame duplication. Most relevant researches are discussed here to detect frame duplication type of forgery.

Wang and Farid [1] have explained a duplication type of video forgery with the example shown in Fig. 1. To hide a particular person from the sequence 4–6 (left), frames 1–3 is copied and then pasted at the place of frame 4–6(right). This kind of forgery is generally known as frame duplication. A common technique is to compare correlation of the two consecutive frames. This technique is computationally high. Wang and Farid [1] have proposed a method in which video is split in group of frames, and correlation coefficients are computed. For similarity check, these coefficients are compared and highlighted where found similar. Results are good up to 90 % of accuracy. The given method was good for stationary camera and robust to MPEG compression.

Kobayashi et al. [2] proposed a powerful method to detect a duplicated region in the video based on noise characteristics. With the help of inherent parameters of a camera, result found was more accurate and reliable. Consistency of noise level function in each frame is sufficient to differentiate between attacked and original region. However, this method is good only for static camera and performance dramatically decreases as the compression is performed on the videos.

Hsu et al. [3] adopt a method using noise residue between two consecutive frames for frame duplications. Parameters of GMM are used to lower down the complexity. The results are accurate for stationary camera but not good for compressed videos.

Lin et al. [4] proposed a method to detect and localize the duplicated frames. Method uses histogram difference of RGB values and takes the correlation of histogram differences of adjacent frames between query and test clip. Based on threshold value candidate clips are selected for further analysis. Once candidate clips are selected, block wise of histogram of luminance values of blocks are calculated and if difference between query and test frame is below threshold the blocks are similar. Process is applied on all blocks of all frames between test and query frame. Finally a frame duplication classification scheme is used that uses the number of matching blocks between frames of query clip and test clip to classify them into duplicated or not duplicated frames.

Milani et al. [5] propose a method to detect spatio-temporal forgeries by analyzing left footprints. Method deals with spatio-temporal region copy paste either, by part of group of frames, or by repeated slice of single image. For the image based attack they take residue of adjacent frames for all frames. A 3D residual matrix is calculated. Small regions with zero residues are eliminated using the morphological erosion operation. A coarse-to-fine or big to small scheme iterations are performed on the morphed residual matrix, to find the duplicated 3D spatio-temporal slices. For image based attack true positive, false negative, true negative and false positives are 62 %, 38 %, 93.8, and 6 %. For video based attack detection rate is 88 %.

Mondaini et al. [6] proposed a method to detect whole frame duplication and for object insertion. Authors used camera sensor pattern noise as a characteristic of a camera. Presented work also claims to be robust to some level of invariance to MPEG compression.

2 Proposed Methodology

The Frame duplication is one of the major attacks performed on videos because it is easy to perform in comparison with other types of attack. Hence a need of efficient, fast and reliable algorithm arises to detect frame duplication. In this type of attack some sequence of frames are replaced with some other sequence of frames from the same video sequence.

A novel method is presented through this paper to detect frame duplication. After performing a number of experiments on any raw video (even with static scene), it is concluded that there is very less probability to have two or more frames exactly same on the basis of intensity values. The reason behind such result is noise caused during acquisition time. We tested it with different quality camcorders and for the different scene with different light intensities even though results are same.

Proposed algorithm is able to identify and localize the frame duplication forgery in targeted video sequences $f(x, y, t_i)$.

If $i = M$, we assume that there are M instances of frames.

Let us assume that a group of frames (i = 1 to k) is duplicated at any other place in the same video sequence. The one very obvious approach is to perform an exact match between all possible pairs of frames to detect duplication. Such type of detection is computationally very high and never should be adapted for detection. Therefore, we present a novel method which is not only computational efficient but also robust to compression artifacts. Figure 2, shows the block diagram of proposed method.

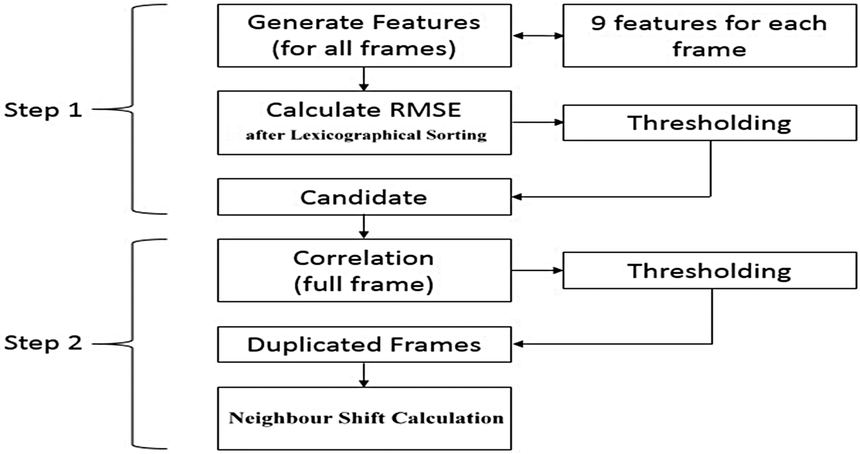

Step 1
Generate Features (for all frames)
9 features for each frame
Calculate RMSE after Lexicographical Sorting
Thresholding
Candidate
Step 2
Correlation (full frame)
Thresholding
Duplicated Frames
Neighbour Shift Calculation

Fig. 2. Block diagram of proposed method. Step 1 candidate frame detection and Step 2 duplicated frame detection.

2.1 Feature Calculation

In spite of comparing intensity values of each pixel as a feature to match two frames, we have chosen nine symmetrical features [7] for each frame. After a lot experiments on our duplicated frame data, we found these features are strong enough to detect duplicated frames. To calculate features each frames are divide into four sub blocks (B_1, B_2, B_3, B_4) as shown in Fig. 3. Features are discussed in following sub-sections.

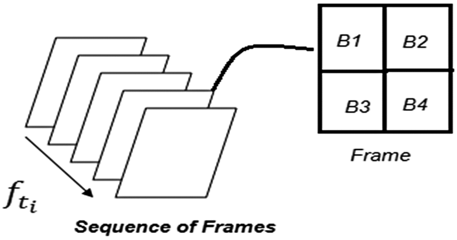


Fig. 3. Sub-blocking of frames

Features are calculated as follows:-

1. **Mean of Block:** *One feature is extracted for each frame.* Mean of a frame is calculated as shown in (1).

$$\mu = 1/MN \sum_{i=0}^{i=M} \sum_{j=0}^{i=N} X\,(i,j) \tag{1}$$

   Where X(i, j) is the set of intensities in each frames. M, N is the dimension of frame. Hence for k number of frames, we will get k means.

2. **Ratio for each Sub Block:** This will result in *four identical features for each frame.* Mean of intensities of each sub block is divided by mean of respective frame one by one as shown in (2). Hence for four different sub blocks of a particular frame, four new features are derived.

$$r_i = \frac{4\sum B_i}{\sum B}\; for\;\; (i = 1, 2, 3, 4) \tag{2}$$

3. **Residue for each Sub Block:** *Four features are extracted for each frames.* Formula given below (3) is used to get residual feature of each of the frame.

$$e_i = \sum B - 4\sum B_i for\,(i = 1, 2, 3, 4) \tag{3}$$

For four sub block of a particular frame, four new features are derived again. Hence, we get nine features for each frame f_i. Maintain a feature vector for each of the frames. $f_i = \{\mu, r_1, r_2, r_3, r_4, e_1, e_2, e_3, e_4\}$, where i is the number of frames in video sequence. Therefore, we get above similar feature set for each frame.

2.2 Lexicographical Sorting

For each frames, we get nine features as f_i. For l no of frames, there will be l feature vectors. Let v1, v2, …, vl be the feature vectors corresponding to these l frames. To perform lexicographical sorting on these vectors of size 9, we regard each of them as a

9-digit number with each digit ranging from 0 to 255. Each vector is assigned with the position of the frame in targeted video. After sorting, likely similar frames are grouped together. It should be ensured that each feature vector (vi) contains its actual position even after sorting.

2.3 Calculation of RMSE and Suspected Frame Detection

RMSE (Root Mean Square Error) is calculated between the feature vectors of adjacent frames after sorting.

$$\boldsymbol{RMSE} = \sqrt{\left(f_i^{\mu,r,e} - f_{i+1}^{\mu,r,e}\right)^2} \quad \boldsymbol{for\ i} = 1\boldsymbol{l} - 1$$

If RMSE is found more than threshold then discards such frames and rest of the frames are marked as suspected frames. We set the threshold value little bigger in order to detect suspected frames in compressed videos.

2.4 Detection of Duplicated Frames

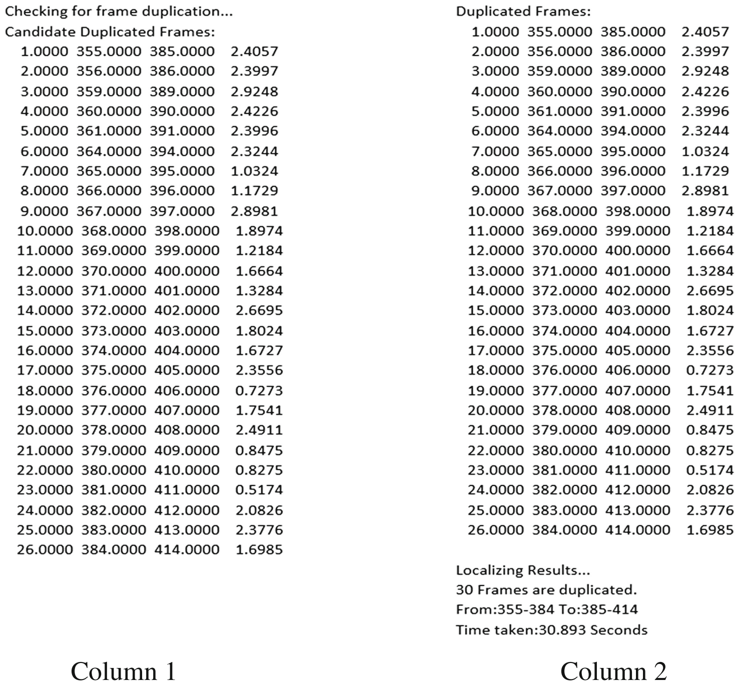
Finally, we get a list of frames that are suspected frames according to their RMSE. These frames are regarded as candidate frames. For exact match, we used correlation between each frames listed in the candidate list.

$$r_{xy} = \frac{\sum_{i=1}^{n}\left(x_i - x'\right)\left(y_i - y'\right)}{(n-1)\,s_x s_y} = \frac{\sum_{i=1}^{n}\left(x_i - x'\right)\left(y_i - y'\right)}{\sqrt{\sum_{i=1}^{n}\left(x_i - x'\right)^2 \sum_{i=1}^{n}\left(y_i - y'\right)^2}}$$

If the correlation between pair of frames is very high (nearly 1) that frames are marked as duplicated. This is assumed that a fraudster duplicates more than five consecutive frames in order to cheat human eyes because single frame duplication can not affect original meaning of video. Therefore, if two frames are found duplicated after correlation matching, very next four consecutive frames of these two frames are also compared as above. Frames are declared duplicated only if such group of frames is found duplicated with another group of frames. Alternatively, we can discard isolated frame matches. Single frame duplication does not have any significance for the fraudsters.

3 Results

The snapshot of the implementation result on self made video is given in Fig. 4. Column 1 is the result of the candidate selection and Column 2 is the result of the duplicated frames. First, second, third and fourth value in detected results are indicating the serial number, first frame sequence number, second frame sequence number and mean square error between their features respectively. At the bottom of column 2 localization results are given. This is clear in the snap shot that thirty frames are copied and pasted elsewhere in the same video sequence. Results are also depicting copied frames as well as the result.



```
Checking for frame duplication...
Candidate Duplicated Frames:
   1.0000  355.0000  385.0000    2.4057
   2.0000  356.0000  386.0000    2.3997
   3.0000  359.0000  389.0000    2.9248
   4.0000  360.0000  390.0000    2.4226
   5.0000  361.0000  391.0000    2.3996
   6.0000  364.0000  394.0000    2.3244
   7.0000  365.0000  395.0000    1.0324
   8.0000  366.0000  396.0000    1.1729
   9.0000  367.0000  397.0000    2.8981
  10.0000  368.0000  398.0000    1.8974
  11.0000  369.0000  399.0000    1.2184
  12.0000  370.0000  400.0000    1.6664
  13.0000  371.0000  401.0000    1.3284
  14.0000  372.0000  402.0000    2.6695
  15.0000  373.0000  403.0000    1.8024
  16.0000  374.0000  404.0000    1.6727
  17.0000  375.0000  405.0000    2.3556
  18.0000  376.0000  406.0000    0.7273
  19.0000  377.0000  407.0000    1.7541
  20.0000  378.0000  408.0000    2.4911
  21.0000  379.0000  409.0000    0.8475
  22.0000  380.0000  410.0000    0.8275
  23.0000  381.0000  411.0000    0.5174
  24.0000  382.0000  412.0000    2.0826
  25.0000  383.0000  413.0000    2.3776
  26.0000  384.0000  414.0000    1.6985
```

Column 1

```
Duplicated Frames:
   1.0000  355.0000  385.0000    2.4057
   2.0000  356.0000  386.0000    2.3997
   3.0000  359.0000  389.0000    2.9248
   4.0000  360.0000  390.0000    2.4226
   5.0000  361.0000  391.0000    2.3996
   6.0000  364.0000  394.0000    2.3244
   7.0000  365.0000  395.0000    1.0324
   8.0000  366.0000  396.0000    1.1729
   9.0000  367.0000  397.0000    2.8981
  10.0000  368.0000  398.0000    1.8974
  11.0000  369.0000  399.0000    1.2184
  12.0000  370.0000  400.0000    1.6664
  13.0000  371.0000  401.0000    1.3284
  14.0000  372.0000  402.0000    2.6695
  15.0000  373.0000  403.0000    1.8024
  16.0000  374.0000  404.0000    1.6727
  17.0000  375.0000  405.0000    2.3556
  18.0000  376.0000  406.0000    0.7273
  19.0000  377.0000  407.0000    1.7541
  20.0000  378.0000  408.0000    2.4911
  21.0000  379.0000  409.0000    0.8475
  22.0000  380.0000  410.0000    0.8275
  23.0000  381.0000  411.0000    0.5174
  24.0000  382.0000  412.0000    2.0826
  25.0000  383.0000  413.0000    2.3776
  26.0000  384.0000  414.0000    1.6985

Localizing Results...
30 Frames are duplicated.
From:355-384 To:385-414
Time taken:30.893 Seconds
```

Column 2

Fig. 4. Sub-blocking of frames

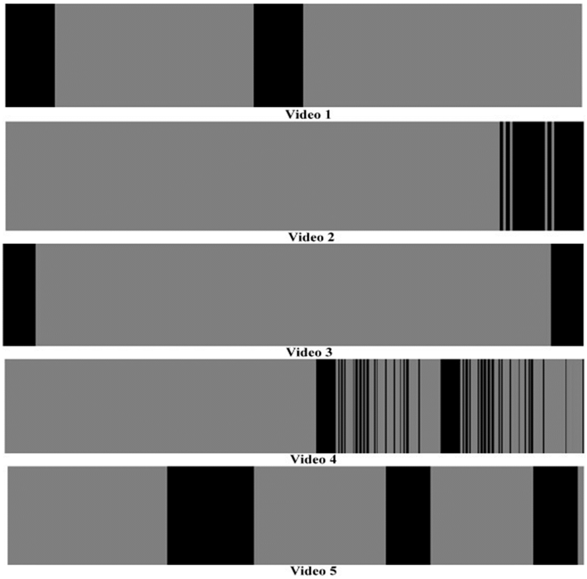


Fig. 5. Duplicate frames detected in Videos 1, 2, 3, 4 and 5

Results of method applied on five videos (as shown in Table 2) are represented in the Fig. 5. Black regions indicate duplicated frame(s) and the gray regions indicate the non-duplicated frames. Results shown in Fig. 5 are in time domain. X axis of each Video Sequence is plotted with respect to time. Hence, a black patch is representing a group of adjacent frames which are detected as a duplicate group of frames.

4 Discussions and Performance Measurement

We have performed experiment on both compressed and uncompressed forged videos with different frame rates. Experiments on our database show accuracy approximately 98.1 % in detection of duplicated frames. False positives found in such exercise were approximately 1 %. Interestingly, In the case of moving camera 99.5 % of duplicated frames were detected. The reason behind this significant improved accuracy is least probability of occurrence of static frames in the moving camera. Accuracy is summarized in following Table 1.

Table 1. Accuracy for stationary camera and moving camera videos

Compressed video (source)	Detection accuracy	Detection time (fixed length video)
Stationary camera	98.1 %	7.96 s (For 9 s Clip)
Moving camera	99.5 %	20 s

For performance measure, we compute the following measures on the entire database:

T-P (true positive): forged frames declared forged.; F-P (false positive): genuine frames declared forged.; T-N (true negative): genuine frames declared genuine.; F-N (false positive): forged frames declared genuine. Thereafter calculating these quantities, results can be given in terms of sensitivity (Sn), specificity (Sp) and accuracy (Ac).

Sensitivity or true positive rate (TPR)

$$TPR = \frac{TP}{P} = \frac{TP}{TP + FN}$$

Specificity (SPC) or True Negative Rate (TNR)

$$SPC = \frac{TN}{N} = \frac{TN}{TN + FP}$$

Fall-out or false positive rate (FPR)

$$FPR = \frac{FP}{N} = \frac{FP}{TN + FP} = 1 - SPC$$

Accuracy *(ACC)*

$$ACC = \frac{TP + TN}{P + N}$$

Performance thus calculated shows a prominent result. Video 1, 2 and 3 from Table 2 shows a highest sensitivity as shown in Table 3. To get all the duplicated frames in forged video is still a point to concern about. Overall presented method is perfect in deciding a forged video with frame duplication. Due to lack of any global database, a comparison with other techniques could not be presented. With our database, better results are found in comparison to other techniques in reference.

Table 2. Different properties of the videos.

Properties → Test Videos ↓	Total frames (count)	Time length (s)	Resolution (Pixels)	Bitrate (kbps)	Detection time (s)
Video 1	465	15	1080 × 1920	37428	43
Video 2	414	13	1080 × 1920	34052	36
Video 3	353	11	1080 × 1920	44048	30
Video 4	469	15	240 × 320	3330	17
Video 5	274	9	240 × 320	3905	8

Table 3. Different detection measures.

	TPR or sensitivity	TNR or specificity	FPR or fallout	Accuracy
Video 1	1.0	1.0	0.0	1.0
Video 2	0.86	1.0	0.0	0.990
Video 3	1.0	1.0	0.0	1.0
Video 4	0.35	0.994	0.005	0.838
Video 5	1	0.995	0.004	0.996

5 Conclusions

A new approach to detect duplicated frames is proposed in this paper. Few frames are copied from the video clip and pasted within the same video in order to manipulate the story in video. Similarity between frames is computed to decide a duplicated frame present in the video sequence. To find out duplicated frames, nine features for each frame are calculated and compared with each other. These features make our proposed algorithm fast and more accurate than earlier such methods. Based on this analysis, we can determine and localize the duplicated frames. The result is prominent even for compressed videos. The proposed scheme exhibits simpler design and implementation. The experimental results have validated the efficiency of our video forgery detection technique.

Acknowledgments. This work is supported by Tata Consultancy Service under RSP scheme.

References

1. Wang, W., Farid, H.: Exposing digital forgeries in video by detecting duplication. In: Proceedings of ACM 9th Workshop on Multimedia and Security, pp. 35–42 (2007)
2. Kobayashi, M., Okabe, T., Sato, Y.: Detecting video forgeries based on noise characteristics. In: Wada, T., Huang, F., Lin, S. (eds.) PSIVT 2009. LNCS, vol. 5414, pp. 306–317. Springer, Heidelberg (2009)
3. Hsu, C.-C., Hung, T.-Y., Lin, C.-W., Hsu, C.-T.: Video forgery detection using correlation of noise residue. In: Proceedings of IEEE 10th Workshop on Multimedia Signal Processing, pp. 170–174 (2008)
4. Lin, G.S., Chang, J.F., Chuang, C.H.: Detecting frame duplication based on spatial and temporal analyses. In: Proceedings of IEEE Conference on Computer Science and Education, pp. 1396–1399 (2011)
5. Bestagini, P., Milani, S., Tagliasacchi, M., Tubaro, S.: Local tampering detection in video sequences. In: IEEE International Workshop on Multimedia Signal Processing, pp. 488–493 (2013)
6. Mondaini, N., Caldelli, R., Piva, A., Barni, M., Cappellini, V.: Detection of malevolent changes in digital video for forensic applications. In: Proceedings of SPIE, Security, Steganography, and Watermarking of Multimedia Contents IX, vol. 6505 (2007)
7. Tripathi, R.C., Singh, V.K.: Fast and efficient region duplication detection in digital images using sub-blocking method. Int. J. Adv. Sci. Technol. **35**, 93–102 (2011)