symmetry

MDPI

Article

Image Region Duplication Forgery Detection Based on Angular Radial Partitioning and Harris Key-Points

Diaa M. Uliyan, Hamid A. Jalab *, Ainuddin W. Abdul Wahab and Somayeh Sadeghi

Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia; diaa_uliyan@siswa.um.edu.my (D.M.U.); ainuddin@um.edu.my (A.W.A.W.); ssomayeh@siswa.um.edu.my (S.S.)
* Correspondence: hamidjalab@um.edu.my; Tel.: +6-03-7967-2503

Academic Editors: Young-Sik Jeong, Laurence T. Yang and Stefanos Gritzalis
Received: 30 April 2016; Accepted: 6 July 2016; Published: 13 July 2016

Abstract: Region duplication forgery where a part of the image itself is copied and pasted onto a different part of the same image grid is becoming more popular in image manipulation. The forgers often apply geometric transformations such as rotation and scaling operations to make the forgery imperceptible. In this study, an image region duplication forgery detection algorithm is proposed based on the angular radial partitioning and Harris key-points. Two standard databases have been used: image data manipulation and MICC-F220 (Media Integration and Communication Center– of the University of Florence) for experimentation. Experiment results demonstrate that the proposed technique can detect rotated regions in multiples of 30 degrees and can detect region duplication with different scaling factors from 0.8, to 1.2. More experimental results are presented to confirm the effectiveness of detecting region duplication that has undergone other changes, such as Gaussian noise, and JPEG compression.

Keywords: copy-move detection; duplicated region; Harris interest points; angular radial partitioning

1. Introduction

Digital images play an important role in our daily life. However, due to the powerful image editing software, images can be easily tampered with. Therefore, we need to think about the authenticity of the images. This can be done by digital image forensic techniques.

Two main types of authentication methods in digital image forensics have been explored in the literature: active methods and passive methods.

In active methods, mainly watermarking and steganography techniques are used to embed the digital authentication information into the original image. The authentication information may be used as a verification for forensic investigation when the image has been falsified, and even point out if the image has been tampered with or not. These techniques are restricted because authentication information could be embedded either at the time of recording, or later by an authorized person. The limitation of this technique is that it needs special cameras or subsequent processing of the digital image. Furthermore, some watermarks may distort the quality of the original image. Due to these restrictions, the researchers tend to develop passive techniques for digital image forensics.

Passive image forensics techniques inspect the images without embedding authentication information, such as signatures or watermarks. Passive image forensics approaches have been classified into five categories:

1. *Pixel based techniques* are based on detecting the statistical irregularity or pixel level correlations, introduced at the pixel level during the forgery process [1]. Pixel-based approaches are the most popular in image forgery.
2. *Format based techniques* are based on detecting the transformation of image forgery via analysis of JPEG compression artifacts [2].

3. *Camera based techniques* concentrate on detecting the clues of image forgery by exploiting the artifacts introduced by different stages of the image capturing process [3].
4. *Physics-based techniques* are based on estimating the lighting directions and differences in lighting between image regions in the image as a telltale sign of image tampering [4].
5. *Geometric based techniques* are based on estimating principal point of image regions across the image, and the inconsistency between principal points, can be used as evidence of image forgery [5].

In pixel-based techniques, the key idea is exposing image tampering by analyzing pixel level correlations. Based on the operation used to create a tampered image, pixel based image forgery techniques can be categorized into three groups: image splicing, image retouching and copy-move forgery.

1. *Image splicing* adds a part of an image into another image in order to hide or change the content of the second image [6–9].
2. *Image retouching* manipulates an image by enhancing or reducing certain features of the image without making significant changes on image content [10].
3. *Copy–move forgery* is copying a region of an image and pasting it in another location of the same image. The forgers perform duplicate regions with different geometric and post-processing operations to hide traces and make consistency with surrounding area [11–13].

Copy–move forgery is becoming one of the most popular image operations in image tampering especially with many easy to use photo editing software. The key characteristic of the duplicated regions is that they have the same noise components, textures, color patterns, homogeneity conditions and internal structures.

For copy–move forgery detection (CMFD), many methods have been proposed that use algorithms that are either block based or key-point based algorithms.

1.1. Block Based Algorithm

First, an image is divided into overlapping sub-blocks. Then, some features are extracted from each block, and compared with other blocks to find the most similar blocks. Various techniques have been developed by several researchers to deal with copy–move forgery using block based methods. The first block based method for detecting copy–move forgery was introduced by Fridrich et al. [14]. Discrete cosine transformation (DCT) based features have been used. The main drawback of their proposed algorithm is the high computational complexity and inefficiency against some post-processing operations such as rotation, scaling and blurring.

In [15], the ridgelet transform of divided blocks was applied to extract features and compute Hu moments for these features to produce feature vectors. Euclidean distance of feature vectors is computed for similarity measure. However, in [16], a copy–move forgery algorithm based on discrete wavelet transform (DWT) was proposed to extract features from input image to yield a reduction in feature dimensional size.

Common limitations of block-based methods include direct uses of quantized or frequency based features from divided blocks for matching, which makes the size of feature vectors quite high, and makes the dimension reduction step mandatory. To overcome these issues, an alternative approach used for the copy–move tampering detection is the key-point based method, as discussed in the next subsection.

1.2. Key Point Based Algorithm

The key-point based methods depend on the extraction of local interest points (key-points) from local interest regions with high entropy without any image sub-blocking. Good key-points should be efficient for computing, be capable of identifying distinct locations in the image regions, and be robust in detection of geometric transformations, illumination changes, noise and other distortions.

The main advantage of key-point based methods is that they have high detection rates in duplicated regions, which exhibit a rich structure such as image regions, but still struggle to reduce the false matches in the flat regions like sky and ocean, etc.

Huang et al. [17], proposed a Scale-invariant feature transform (SIFT) based detection method to identify duplicate regions with scale and rotation, then used the best bin first search (BBF) method to find the nearest neighbors with high probability that return the matched key points (inliers) as a possible duplicate region. To increase the accuracy of detection methods, a nearest neighbor distance ratio (NNDR) is applied for matched key-points.

Amerini et al. [18] have proposed detecting multiple duplicated regions based on SIFT features, and then employed generalized nearest neighbor (G2NN) to improve the similarity match between key-points. The agglomerative hierarchical linkage (AHL) clustering method has been employed to group the similar key-points into the same cluster and merge closest pair of clusters into single cluster to represent the cloned regions. The estimation of affine transformation parameters is computed between duplicated regions.

Battiato et al. [19] proposed a framework for detection of duplicated region based on SIFT. The hashing method is applied to the feature vectors, and then saved into a hash table, which is used for comparing the hash code of corresponding feature vectors after image alignment. The alignment process of the hash is used to estimate the geometric transformation parameters.

All above techniques have the use of the SIFT key-points in common, which are invariant to changes in rotation, illumination, and scaling. Hence, they commonly inherit the limitations of a lack of key-points in flat or smooth areas in images, where little structures are exhibited. This motivated researchers to utilize other key-points descriptors to overcome these limitations, such as Harris corners, Hessian and Laplacian of Gaussian (LOG) detectors.

Liu et al. [20] have applied Harris detectors to localize key-points, and then adaptive non maximal suppression (ANMS) is employed to identify more key-points in the flat regions in images, which is the main drawback of the SIFT algorithm. The daisy descriptor has been implemented based on an orientation gradient map of image regions around the key-point to perform image matching in a dense way.

Ling et al. [21] proposed a near duplicate image detection framework based on Hessian affine and polar mapping. At the final stage of detection, a nearest neighbor distance ratio (NNDR) here is used for matching between feature vectors.

Kakar et al. [22] proposed an algorithm where local interest points are detected by a modified scale Laplacian of Gaussian and Harris corner detector, which make features invariant to geometric transformations. In addition, the feature descriptor was built based on MPEG-7 signature tools.

It is very obvious from the literature that the CMFD approaches discussed have their pros and cons regarding the geometrical changes in the copied regions. Therefore, an efficient CMFD should be robust to some geometrical changes, such as rotation and scaling. These issues have been under extensive study during the past few years.

Rotation is considered the most difficult geometric transformation to deal with in CMFD. Three key approaches were introduced to achieve a rotation invariant CMFD: polar mapping, circle blocking and image moments such as Zernike moments.

In [23], polar mapping based on log-polar transformation of divided blocks in images was employed. Then, Fast Fourier transformation is applied to build descriptors under different orientations range from 0°–180°. These feature descriptors are then saved into matrices and lexicographical sorting is applied to them. To improve the detection decision, the counting bloom filter has been used to detect descriptors that have the same hashing values.

Shao et al. [24] proposed a duplicated region detection based on circular window expansion and phase correlation. The algorithm starts with calculating the Fourier transform of the polar expansion on overlapping circular windows pairs, and then an adaptive process is applied to obtain a correlation matrix. Then, estimating the rotation angle of the forged region, a searching procedure is implemented to display the duplicated regions. The algorithm was robust to rotation, illumination changes, blurring, and JPEG compression.

In [25], the Zernike moments are extracted from overlapping blocks and their magnitudes are used as feature representation. Locality sensitive hashing (LSH) is employed for block matching, and falsely matched block pairs are removed by inspecting phase differences of corresponding Zernike moments.

In this research, we propose developing an efficient copy–move forgery detection algorithm that is able to detect and locate different duplicated regions under various geometric transformations and post processing operations.

The main advantage of key-point based CMFD methods is invariance under rotation, scaling operations, but still struggle to reduce the false matches in flat regions. In order to improve the efficiency and capability of region duplication detection, we propose a rotation robust method for detecting duplicated regions using Harris corner points and angular radial partitioning (ARP). The proposed method depends on statistical region merging segmentation techniques (SRM) of images as a preprocessing step to detect smooth and patterned regions. However, to improve the time complexity of detection algorithm, a linkage clustering, based on Tamura texture features of detected image regions, is employed in the next step. Then, Harris corner points are detected in angular radial partitions (ARPs) of a circle region for each detected image region in order to obtain a scale and rotation invariant feature points. The matching procedure includes two main steps: (1) matching between chain codes to find similar regions based on internal structure; (2) matching the feature point's patches based on Hölder estimation regularity based descriptor (HGP-2) to find similar regions based on texture. Finally, the matched keypoints between similar regions are linked and the duplicated regions are localized in a blue color to show forgery.

The organization of this paper is presented as follows. In Section 2, the proposed algorithm is explained. Section 3 presents the experiment results. Comparisons and discussions are described in Section 4. Finally, conclusions are drawn in Section 5.

2. Proposed Method

The whole proposed detection method generally can be described as follows:

(1) Input the suspicious color image.
(2) Segment the input image of size $M \times N$ pixels into image regions using Statistical Region Merging Segmentation (SRM).
(3) Locate centroids of each image region.
(4) Crop each segmented image region to 11×11 pixel square image blocks around the centroids of detected regions.
(5) Apply Tamura texture on each square image block.
(6) Cluster similar image regions by texture.
(7) Apply angular radial partitioning (ARP) on each image region indexed by its centroid coordinates in the same cluster.
(8) Convert input image into grayscale and extract key points with Harris detectors.
(9) Calculate the total number of Harris corners in each sector of a circle image region in the same cluster.
(10) Represent the total number of Harris corners by a chain code, and search for closely one-to-one matching between two images regions in the same cluster.
(11) Extract regularity-based features the centroid and around each Harris corners of the image regions that have the same chain code by using an HGP-2 descriptor.
(12) Compute median absolute deviation (MAD) for all HGP-2s of the two matched image regions, and save it in feature vector f_v.
(13) Find the Euclidean distance between two corresponding final feature vectors fv and fv'.
(14) Detect and localize the tampered regions.

2.1. Statistical Region Merging Segmentation (SRM)

The duplicated region is copied and moved to another location in the same image. This makes the duplicated regions in forged images have a similar homogenous texture and internal structure.

Segmentation of the input image is carried out using the SRM algorithm [26]. The SRM algorithm has excellent performance with capturing the main structural image features using effective statistical image analysis. The SRM depends on two steps: the merging predicate and the testing order to test the merging of regions in the image.

The advantage to using the SRM method is the ability to detect the smooth and patterned regions like edges and corners.

As a result of the SRM algorithm, the input image is segmented into image regions of similar intensity or color, and saved into mapped images, then the mapped image is scanned horizontally and vertically to find the least frequently occurring values in the mapped image plane. These values are retained to detect regions that will provide a more accurate detection of the tampered regions. Finally, the centroids of the detected segmented image regions are located in the input images, as shown in Figure 1.

(**a**) (**b**) (**c**)

Figure 1. Results of Statistical Region Merging Segmentation (SRM). (**a**) the initial image; (**b**) detected image regions; (**c**) centroids of image regions.

To evaluate the performance of the SRM algorithm, we use two standard datasets, MICC-F220 dataset [18], and image manipulation datasets [27].

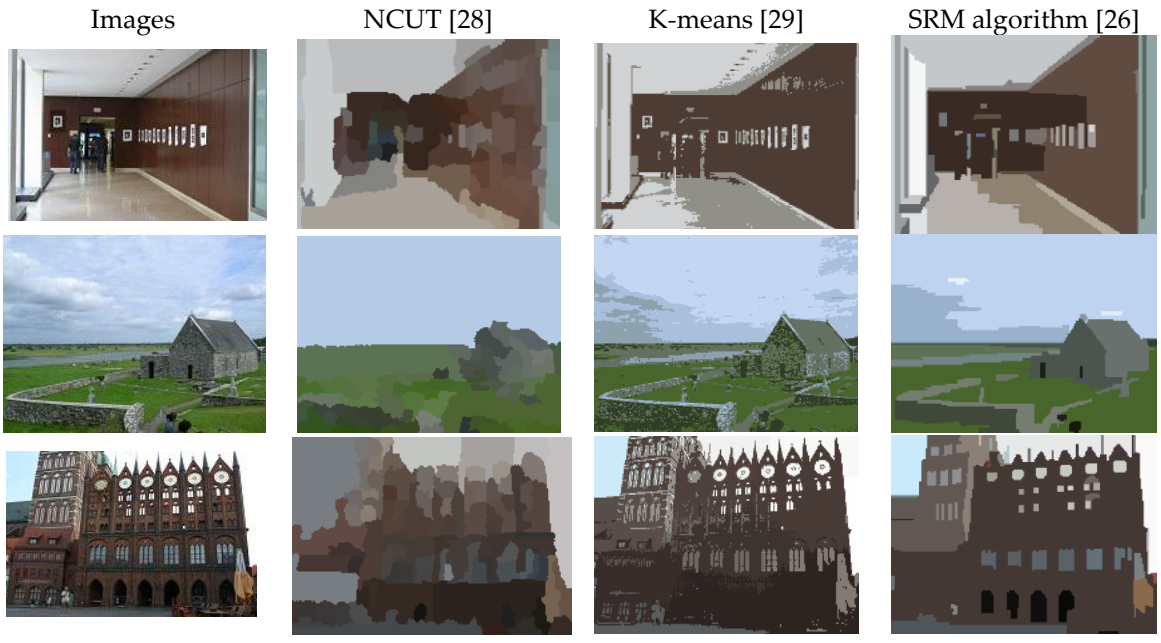


Figure 2. *Cont.*

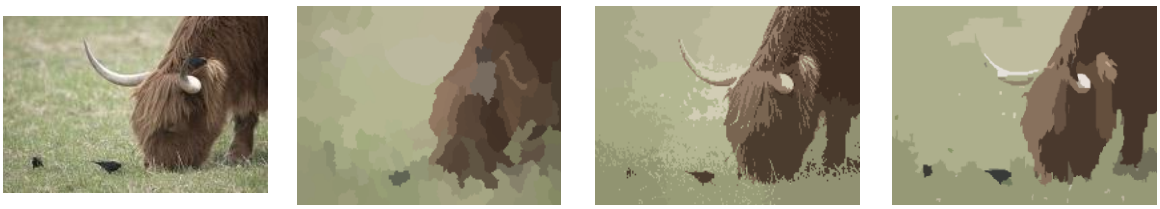


Figure 2. Image segmentation results between different techniques.

We further compare the results of image segmentation with the benchmark segmentation techniques, such as Normalized cuts (NCUT) [28], and K-means [29], as shown in Figure 2. The SRM segmentation that yields the best pixel-wise mean accuracy would be chosen.

2.2. Linkage Clustering of Image Regions Based on Tamura

In order to improve the computational complexity of detection algorithm, the image regions are grouped into clusters. A straightforward linkage image block clustering algorithm has been applied to divide all image blocks into clusters according to their texture features [30].

Improved Tamura texture features are defined as the new criteria for finding similar image regions in block-clustering steps. Improved Tamura features consider only coarseness and contrast as features to create a two-dimensional feature vector [31]. In this study, the Ward's linkage clustering method can be described as follows:

1—Locate centroid of each segmented image region.
2—Crop each segmented image region to 11 × 11 pixels around its centroid.
3—Apply Tamura texture on each square image block.
4—Cluster similar image regions by texture.

By default, it is assumed that the block size of 11 × 11 pixels is smaller than segmented image regions. However, if the block of size 11 × 11 pixels is larger than the size of forged region, the block will cover more neighborhood pixels outside the forged region. As a result, the clustering process for blocks becomes futile.

For every stage in the clustering process, there are N numbers of clusters. The pairs of clusters are formed using $N(N-1)/2$. Every pair of clusters is merged together if they form the least increment to the sum of squares for errors. Once all pairs clusters are merged, the process is considered finished [32]. It is noteworthy that Ward's linkage clustering method has no threshold.

2.3. Angular Radial Partitioning (ARP)

For efficient CMFD systems with rotation and scale invariance, a rotation robust region description method is employed using ARP [33]. The ARP was originally developed as an edge based descriptor. We proposed a simple modification of the original ARP by selecting corner points in the image regions as features.

The algorithm applies ARP of a circular region that is invariant to rotation and scale transformations on each image region indexed by its centroid coordinates in the same cluster.

The image regions are divided into circular sectors of the surrounding circle. The algorithm defines N as the number of angular divisions equal to 12, and $\theta = 2\pi/N$ is the angle between adjacent sectors as shown in Figure 3. The radial partitions are labeled as S_1, S_2, ... , S_{12}. All 12 sectors are saved to find Harris corner points in the next subsection. In our method, we experimentally set the radius of circle region r = 8.

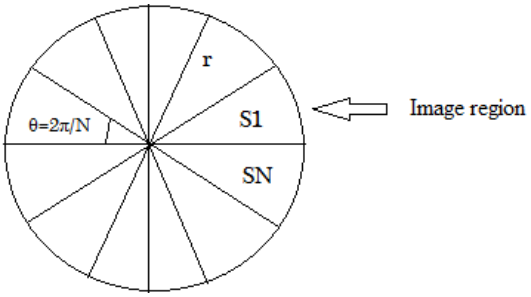


Figure 3. Angular radial partitioning of an image region with angular sectors.

2.4. Harris Corner Detection

The proposed algorithm employs Harris corners as a key-points based method for region duplication forgery detection after converting colored images into grayscale.

A Harris corner detector renders features invariant to geometric transformations. It is used as a key-points method to represent the internal structure of the matched image regions.

The Harris corner detector is an invariant to rotation technique that identifies corners of an image [34]. Harris corner detection is based on the second moment auto correlation matrix. This matrix describes the Gradient distribution of input images at point x, weighted by Gaussian $G(x,\sigma)$ as the following:

$$M = G(x,\sigma)\begin{bmatrix} I_x^2(x,\sigma) & I_xI_y(x,\sigma) \\ I_xI_y(x,\sigma) & I_y^2(x,\sigma) \end{bmatrix}, \quad (1)$$

where I_x^2, I_y^2, I_xI_y are square derivatives of input image I.

Figure 4 shows detection of centroids of the image region, and the detection of Harris corner points around these centroids.

(**a**) (**b**)

Figure 4. Centroids and Harris corners of image regions. (**a**) Centroids of image regions represented in **red**; (**b**) Harris corner points around centroids represented in **green**.

2.5. Region Descriptor Based on Chain Code

In the proposed algorithm, the clustering of image regions based on Tamura texture is followed by matching procedures. However, the matching procedure is divided into two main steps:

In the first step, initial matching is done based on the internal structure of the image region by using the chain codes. While in the second step, the feature based on the HGP-2 descriptor is used as texture-based matching. The second matching step will be used only if the first matching is found, otherwise the matching procedure is repeated between two other image regions in the same cluster.

To achieve good performance with rotation attacks in the first step matching, we developed a rotation invariant method that can detect duplicate regions with rotation by encoding the number of circular shifts expressed in multiples of 30 degrees (counter clockwise) between two chain codes of image regions.

Chain codes are saved into feature vectors to represent the sequence of the total number of Harris corner points in each sector around the centroid of the image region. The significance of using chain codes is to provide a rotation invariant method that can detect duplicate regions with rotation.

The chain code of detected region is defined as follows:

$$chain_{Seg} = [sum(H, S_i), sum(H, S_{i+1}), \ldots, sum(H, S_{i=12})]$$

where sum calculates the total number of Harris corners in each sector.

The chain code and its circularly shifted version are looking for closely one-to-one matching between two image regions in the same cluster.

The rotation duplication forgery, as well as the corresponding detection performance, is illustrated in Table 1. For example, the chain code C2 of the duplicated image region with an angle of rotation of 90° should be circularly shifted anticlockwise three times to be matched with C1. As a result, the proposed algorithm estimates the rotation angle according to the total number of circular shifts of the chain codes of the forged image. Each one time of circular shift is equal to 30°.

Table 1. Threshold value analysis for our method.

Chain Code (C1) for Region 1	Chain Code (C2) for Region 2	Estimated Rotation Angle
[2 2 1 2 1 1 2 2 2 2 2 2]	[2 1 2 1 1 2 2 2 2 2 2 2]	30°
[1 1 2 1 1 1 1 1 1 2 1 1]	[2 1 1 1 1 1 1 2 1 1 1 1]	60°
[1 1 0 0 1 1 1 1 0 0 1 1]	[0 1 1 1 1 0 0 1 1 1 1 0]	90°

2.6. Regularity Based Descriptor

For each cluster of image regions, the first step of the matching process is to find the similar chain codes that represent the internal structure of the matched regions. However, this matching process alone struggles to detect a non-uniform duplicated regions. Thus, in order to reduce the false positives detection, we employed a local descriptor using HGP-2 to measure the amount of irregularities around the centroid and around each Harris corner of the image regions that have the same chain code as shown in Figure 5. The HGP-2 be employed only if the first matching is found.

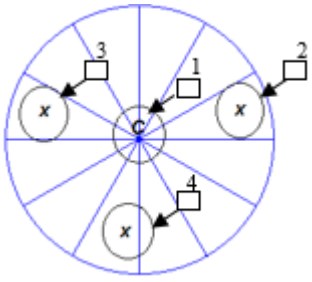


Figure 5. Hölder estimation regularity based descriptor (HGP-2) for image region. (**1**) Represents HGP-2 descriptor for centroid; (**2–4**) represent HGP-2 descriptors for Harris points.

The HGP-2 descriptor is based on formal concept of image regularity. The symbolic expression for HGP-2 estimator is defined by a Gaussian smooth filter, which gives a description of texture and is described as follows [35]:

$$\mathrm{HGP}-2 = G_1 \times |\log|G_1 \times (k \times (R - G_1 \times R))|\,| \quad (2)$$

where R is selected as a 9×9 pixel block, G_1 is a Gaussian kernel at scale 1; k is the optimal operator experimentally fixed as equal to 11 to perform a good representation of local image patterns.

2.7. Region Duplication Localization

The common assumption is that the forged duplicate regions have been subjected to geometric transformation such as rotation. Therefore, we need to represent the image region by robust features that are insensitive to rotation.

To achieve this goal, a median absolute deviation (MAD) has been used as a similarity measure for second matching step between HGP-2 descriptors of centroids and HGP-2 descriptors of all Harris corners of the image regions in the same clusters [36].

MAD as a measure of similarity for the second matching step between HGP-2 sets is proposed to remedy the problem of texture-based matching. MAD is computed separately for $(\text{HGP-2})_{\text{c}}$, and for $(\text{HGP-2})_{\text{H1}}$, $(\text{HGP-2})_{\text{H2}}$,.., $(\text{HGP-2})_{\text{Hn}}$, where n is the number of Harris corners that surround the centroid as shown in Figure 5. As a result, the final M_{c} and M_{H} are denoted as follows:

$$M_{\text{c}} = \text{MAD}(\text{HGP} - 2_{\text{c}}); M_{\text{H}} = \text{MAD}(\text{HGP} - 2_{\text{Hi}}, \text{HGP} - 2_{\text{H2}}, \ldots, \text{HGP} - 2_{\text{Hn}}) \quad (3)$$

M_{H} is created as an array of MAD for each Harris corner that surrounds the centroid. However, the size of $M_{\text{c}} = 1$, while the size of $M_{\text{H}} = n$.

Both M_{c} and M_{H} are defined as a measure of similarity for detecting duplicated regions and are saved into feature vector f_v. This vector is used to detect duplicated image regions. The final feature vector f_v for two image regions in the same cluster is given by:

$$f_v = [M_{\text{c}}, M_{\text{H}}]; f_v{}' = [M_{\text{c}}{}', M'_{\text{H}}] \quad (4)$$

The dimension of both f_v, and $f_v{}'$ will be $n + 1$.

Both f_v and $f_v{}'$ should have similar feature dimension because both image regions have the same number of HGP-2 descriptors.

The final set of matched key points be obtained by using Euclidean distance between two regions R_1, R_2 that share the chain code with corresponding feature vectors f_v and $f_v{}'$:

$$D(R_1, R_2) = \sqrt{\left(f_v - f'_v\right)^2} \quad (5)$$

The two regions R_1, R_2 are considered as matched, only if D satisfies the following condition:

$$D(R_1, R_2) \leqslant \theta \quad (6)$$

where θ is the threshold value. We have experimentally chosen the value of θ as 0.18.

Finally, lines will connect the Harris corner points and centroids of the two matched regions in order to locate the tampered areas in the image. Finally, matched keypoints between similar regions are linked and the duplicated regions are localized in blue to show forgery.

2.8. Threshold Selection Value

The proposed algorithm is analysed to determine the threshold value for achieving the highest true positive rate (TPR) and lowest false positive rate (FPR) scores. Different threshold values are tested to measure their influences on the detection of the forged and original image regions.

Table 2 illustrates how decision thresholds affect the true positive rate (TPR) and false positive rate (FPR). The best threshold (θ) value is empirically found to be 0.18 which identified the best TPR and FPR.

Table 2. Threshold value analysis for our method.

Threshold Value (θ)	True Positive Rate (TPR)	False Positive Rate (FPR)
0.14	0.92%	3%
0.16	0.94%	6%
0.18	0.96%	2.8%
0.20	0.96%	7%
0.22	0.96%	8%
0.25	0.96%	12%

2.9. The Algorithmic Complexity

The goal of computational complexity is to assess the proposed algorithm according to its performances. The main problem in image region duplication forgery detection is the computational complexity related to block matching. The technical viewpoint of this study is based on three components: SRM segmentation, linkage clustering and Harris corner points.

If we consider the input image having ($M \times N$) pixels, then the SRM segmentation time will be defined as [26]:

$$T1 = O(M \times N) \times \log Q \tag{7}$$

where Q is independent random variable taking positive values ranging from 1 to 256 as defined in [26]. The Ward's linkage time used for clustering the image regions can be represented as:

$$T2 = O\left(L^2\right) \tag{8}$$

where L is the total number of detected image regions to obtain clustered regions based on their texture features.

The time used to find Harris corner points for each image region in the same cluster followed by the searching algorithm with sliding search window of size (b × b) would be equal to:

$$T3 = O\left(K \times b^2\right) \tag{9}$$

where ($b \times b$) << ($M \times N$), and K = 12 which is the total number of divided sectors in the image region. Therefore, the computational complexity of the proposed algorithm mainly depends on the image size ($M \times N$) and the total number of detected image regions L.

3. Experiment Results

The experiments were carried out using the MATLAB R2015a software (The Math Works, Natick, MA, USA) on Windows 8.1 Pro 64-bit (Microsoft, Redmond, WA, USA). The experimental images were tested, based on the following two image datasets:

1. MICC-F220 which contains 220 JPG images, 110 are forged images and 110 are originals [18].
2. Image manipulation dataset which includes 48 PNG true color images [27].

The detection of forgery is accessed by the following measures:

i TP (True Positive): forged images detected correctly as forged images.
ii FP (False Positive): original images detected wrongly as forged images.
iii FN (False Negative): forged images falsely missed as forged images.

From these measures, we define four evaluation metrics: Precision (P), Recall (R), True Positive Rate (T_{PR}) and False Positive Rate (F_{PR}) as follows:

$$P = \frac{T_P}{T_P + F_P},\ R = \frac{T_P}{T_P + F_N} \tag{10}$$

Precision (P) indicates the probability that a detected forgery is truly a forgery, while Recall (R) denotes the probability that a forged image is detected:

$$T_{PR} = \frac{No.\ of\ forged\ images\ detected\ as\ forged}{No.\ of\ forged\ images} \tag{11}$$

$$F_{PR} = \frac{No.\ of\ original\ images\ detected\ as\ forged}{No.\ of\ original\ images} \tag{12}$$

where T_{PR} represents the performance of the detection algorithm in correctly locating the pixels in forged regions, while F_{PR} is the fraction of original images that was not correctly detected. The higher the T_{PR} (towards one) and the lower the F_{PR} (towards 0), the better the detection performance.

3.1. Performance Test

The performance of the proposed method is evaluated on a set of images having duplicate regions that are non-regular shapes and meaningful image regions with different size of 32 × 32, 64 × 64, 96 × 96 and 128 × 128 pixels. The images shown in Figure 6 are the detection results of simple copy–move forgery without any post processing operation.

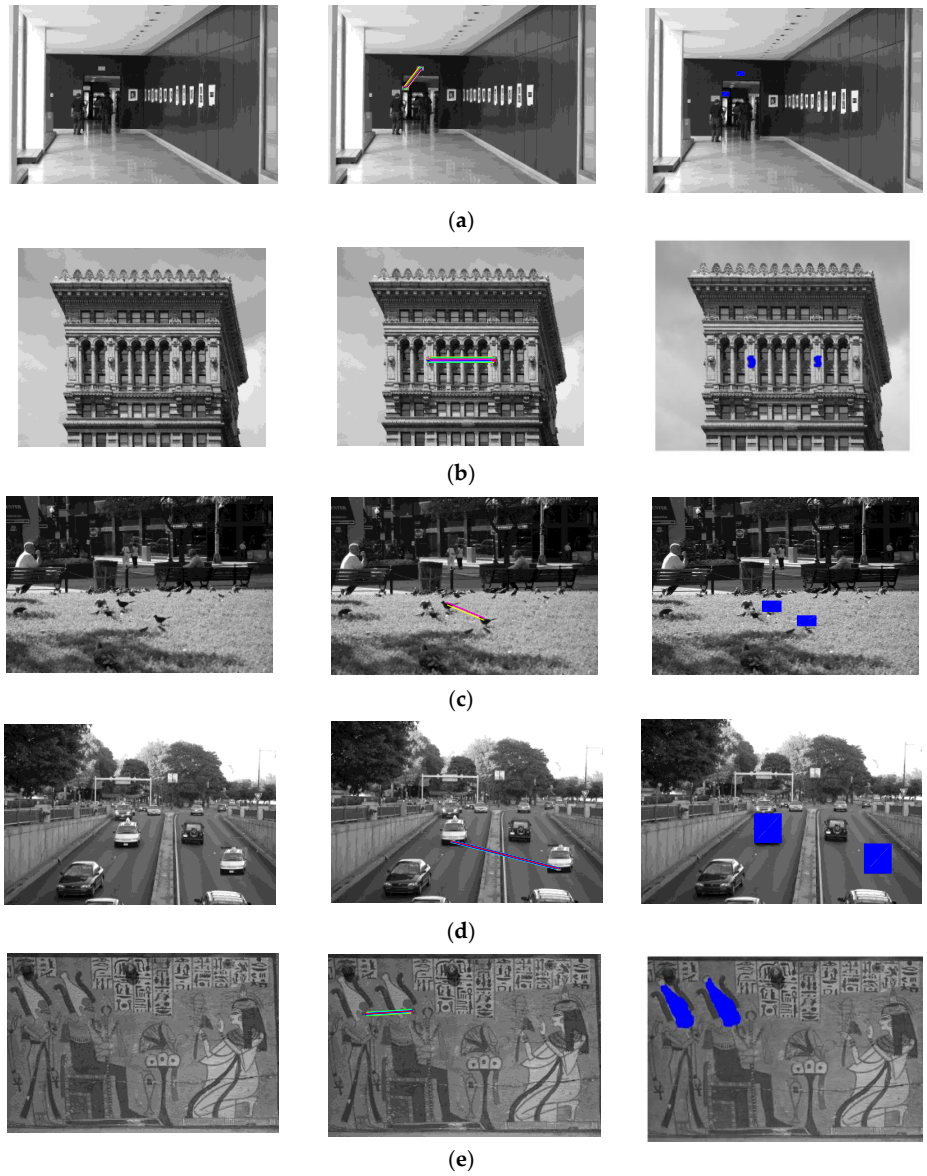


Figure 6. Detection results of various region sizes. (**a**) 20 × 20; (**b**) 32 × 32; (**c**) 64 × 64; (**d**) 96 × 96; and (**e**) 128 × 128 pixels.

3.2. Robustness Test

The robustness performance of the proposed method is evaluated against sets of various kinds of post-processing operations including rotation, scaling, Additive white Gaussian noise (AWGN) and JPEG compression.

3.2.1. JPEG Compression

The robustness of the proposed algorithm against JPEG compression has been evaluated under different JPEG compressions (Q = 90, 80 70, 60, and 50), which are shown in Figure 7.

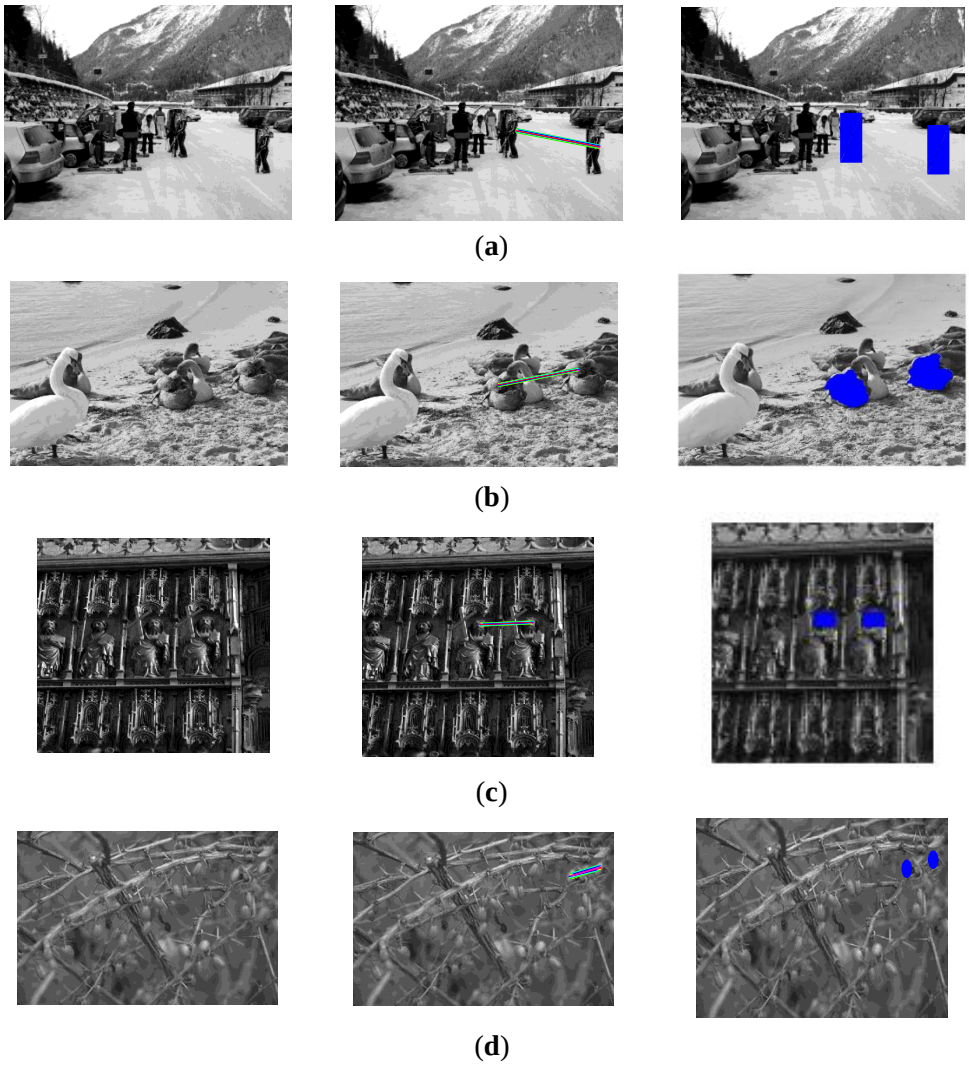


Figure 7. The detection results after JPEG compression. (**a**) $Q = 90$; (**b**) $Q = 70$; (**c**) $Q = 50$; and (**d**) $Q = 30$.

The forged image was compressed with different quality factors and subjected to the rotation and normal copy–move operations. The detection performance for JPEG compression of different qualities is shown in Table 3. Even when the quality factor of the distorted image is quite low, such as JPEG compression ($Q = 50$), the detection performance of the proposed method is still reliable. However, the detection performance tends to decrease with the decrease of the quality factor of JPEG compression. The main issue is that with low quality JPEG compression images, it becomes more difficult to identify the exact Harris corner points. The matching and linking between these points might be strongly confused. As a result, there may be less or even no matched Harris corner points at all. However, with such low quality factors, the compressed images are not usually visually good enough anyway. Thus, these low factors are not usually commonly used.

Table 3. The average detection rate of copy–move forgery for JPEG compression based on MICC-F220 (Media Integration and Communication Center—of the University of Florence).

Operations	Quality Factors	90	80	70	60	50
Normal copy move	TPR	0.96	0.94	0.92	0.90	0.90
	FPR	0.06	0.08	0.08	0.10	0.10
Rotation	TPR	0.96	0.92	0.90	0.86	0.80
	FPR	0.08	0.09	0.09	0.1	0.2

3.2.2. Additive White Gaussian Noise (AWGN)

The proposed method is tested for robustness to AWGN in forged images. The experimental results are shown in Table 4. The forged image was corrupted by White Gaussian Noise with different SNRs of 15, 20, 25, 30 and 35 dB.

Table 4. The average detection rate of copy move forgery for Additive white Gaussian noise (AWGN) on MICC-F220.

Signal-to-noise ratio (SNR)decibel (dB)	**35**	**30**	**25**	**20**	**15**
True Positive Rate (TPR)	0.94	0.94	0.96	0.96	0.96
False Positive Rate (FPR)	0.10	0.07	0.07	0.06	0.06

3.2.3. Rotation Copy–Move Forgery

It is often necessary to apply a rotation operation to an image region before being pasted in order to create convincing forgeries. Figure 8 indicates that our algorithm can identify duplicated regions in the cases of different angles of rotation θ = 30°, 90°, and 180°.

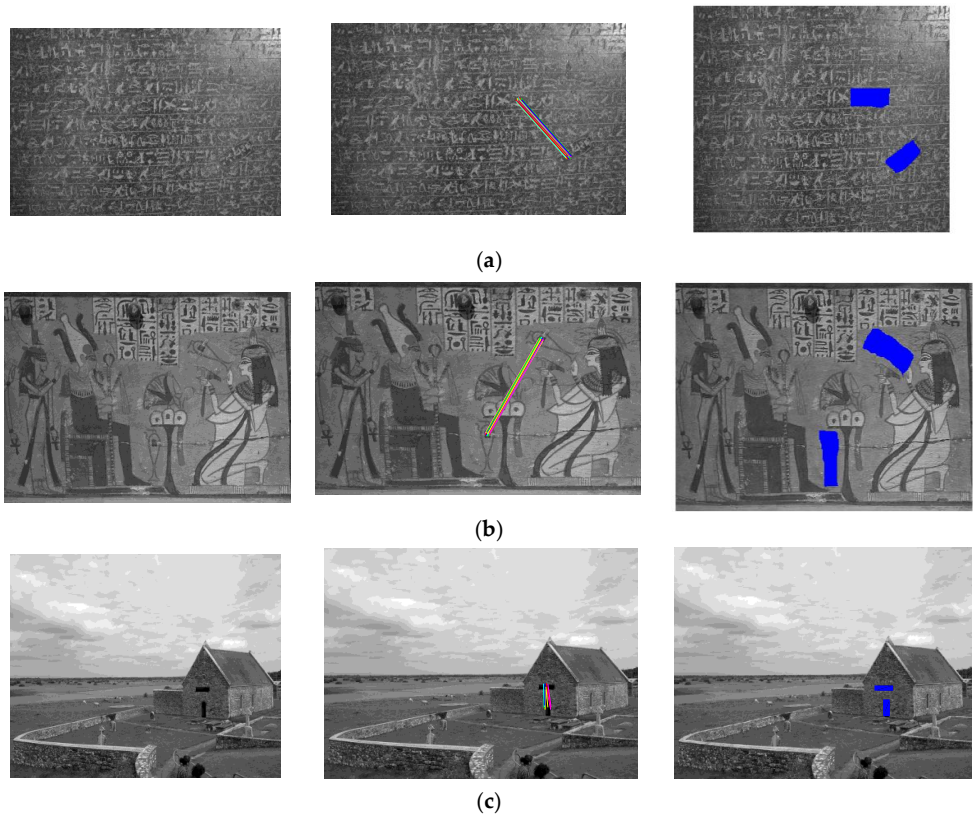
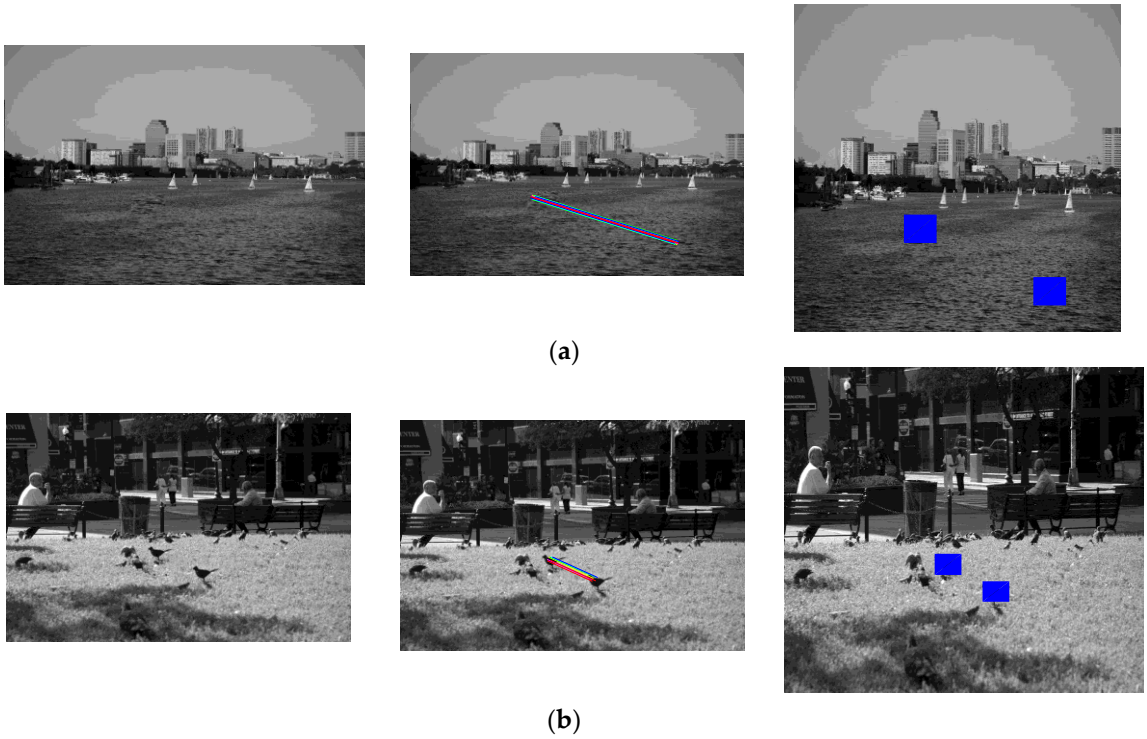


Figure 8. Detection results of duplicated regions in the cases of different angles of rotation. (**a**) 30°; (**b**) 90°; (**c**) 180°.

3.2.4. Scaled Copy–Move Forgery

To test the robustness of our proposed method for detecting region duplication in the case of scaled duplicated regions, forged regions are scaled up or down with various scaling parameters (s = 0.8,0.9,1.1,1.2). One example of a visual result is shown in Figure 9. Furthermore, in order to quantitatively evaluate the robustness of our algorithm under different scaling factors, we randomly selected 50 original images from MICC-F220 image datasets [18]. For each original image and each

duplicate region with square image blocks of sizes 64 × 64, 96 × 96 and 128 × 128 pixels. The detection performances of scaled duplicated regions for each image region size are presented in Table 5.



(**a**)

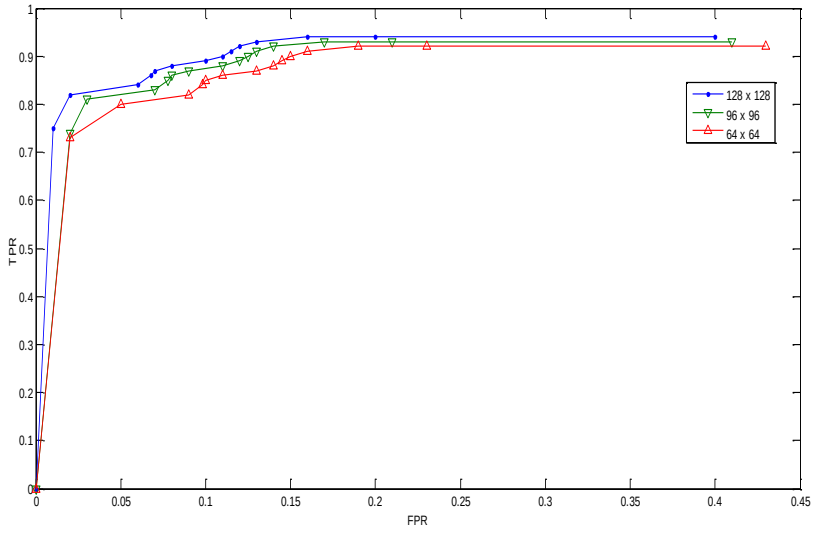
(**b**)

Figure 9. The detection results on sample images. (**a**) scale down with scaling factor = 0.9; and (**b**) scale up with Scaling factor = 1.1.

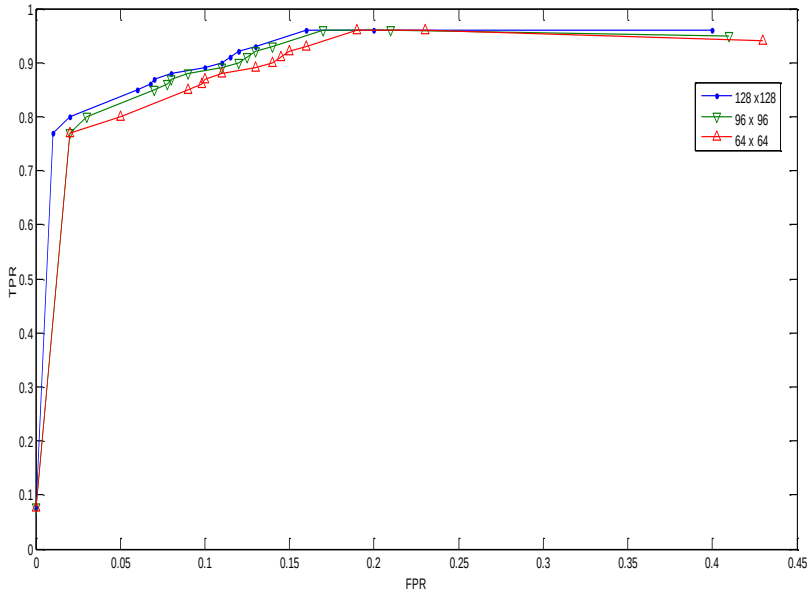
Table 5. The detection performance of scaling duplication for 50 forged images with different image region sizes.

Title	**64 × 64**		**96 × 96**		**128 × 128**	
Scale	TPR	FPR	TPR	FPR	TPR	FPR
0.8	0.90	0.06	0.88	0.07	0.96	0.08
0.9	0.94	0.08	0.96	0.08	0.96	0.08
1.1	0.92	0.08	0.94	0.10	0.92	0.10
1.2	0.92	0.10	0.94	0.10	0.91	0.07

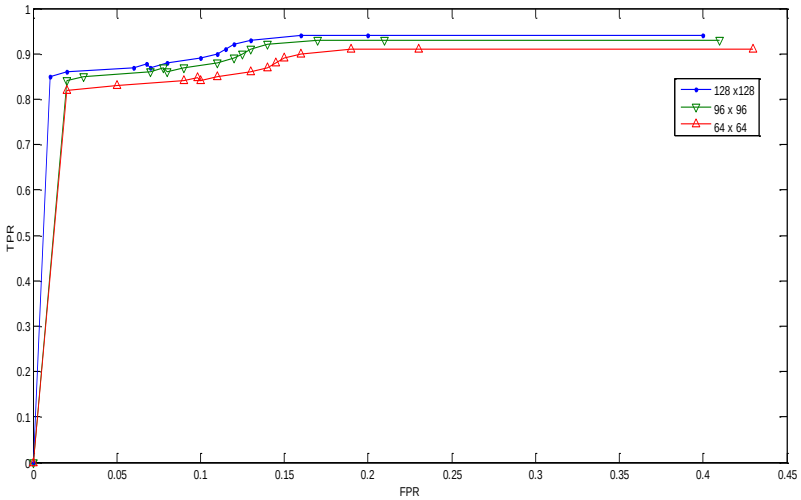
Figure 10 illustrates the TPR and FPR rates that are combined in the receiver operator characteristics (ROC) curve to show the performance of our method that is subjected to various kinds of post processing operations and sizes of duplicated regions. Under different image region sizes, our method performs well with high accuracies at low positive rates. For instance, in the case of normal copy–move forgery, we allowed for random increments in image region sizes ranging from 64 × 64 to 128 × 128 pixels. It can be seen that our method provides TPR greater than 90% and FPR less than 0.08. Furthermore, in the case of forgeries with rotation, our method gives better performance with TPR greater than 85%. In Figure 10c, it can be seen that the performance tends to give high FPR with increasing scaling factor. This is probably due to the greater loss of Harris corner points being caused by scaling operation, which makes it difficult for matching processes to work. Nevertheless, TPR and FPR are acceptable even with high scaling factors. The reason for the overall improved performance is partly due to the fact that the features that we use are also intended to be rotation invariant.



(**a**)



(**b**)



(**c**)

Figure 10. The Receiver Operator Characteristics (ROC) curves for different post-processing operations and image regions sizes of duplicate regions. (**a**) Duplicated region forgery; (**b**) rotation attack; (**c**) scaling attack.

4. Comparison and Discussion

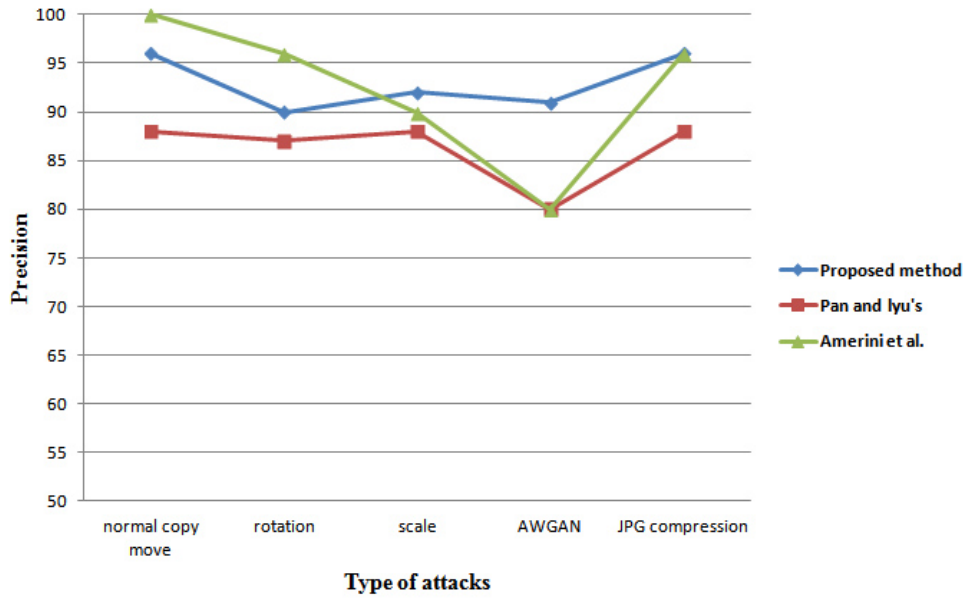
The performance of the proposed method is compared with other methods. Key-points-based techniques [18,22,37,38] and image regions based methods [14,39]. The true positive rates, false positive rates, Precision and Recall are evaluated on the tampered and authentic images from the MICC-F22 and Image data manipulation datasets. As seen from Table 6, our method achieved a TPR around 96% and FPR of 2.8%. Our method reduces the false positive rate while still maintaining a high true positive rate.

Table 6. Average True Positive Rate (TPR) and False Positive Rate (FPR) values in (%) and processing time (sec) for each method on MICC-F220.

Methods	*TPR*	*FPR*	*Time*
Pan and Lyu's method [38]	89.96	1.25	10
Amerini et al. [18]	100	8	4.94
Fridirch et al. [14]	89	84	294.96
Popescu and Farid [39]	87	86	70.97
Kakar et al. [22]	90	3	NA
Mishra et al. [37]	73.64	3.64	2.85
The proposed method	96	2.89	4

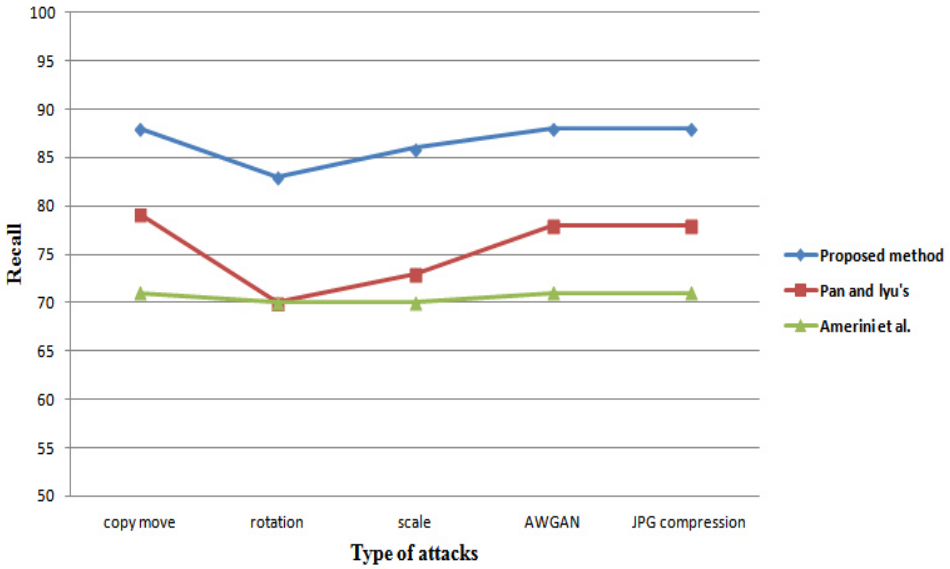
As seen from Table 4, the execution time for block based methods: Fridirch et al. [14] and Popescu and Farid [39] are high compared with keypoint based methods: [18,37,38] and our method. The time required to detect forgery in our method is faster than Pan and Lyu's method [38] and Amerini et al. [18]. However, Mishra et al. [37] method is quite faster than our method due to SURF features.

The Precision and Recall performance of the proposed method were compared to two standard methods: the Pan and Lyu method [38] and Amerini et al. [18], as illustrated in Figure 11. The graph indicates that our method has its best performance with precision of more than 90% under most conditions: normal copy–move, scale, AWGAN and JPEG compression. Furthermore, for different amounts of rotation, our method performs better than Pan and Lyu's method. In the recall rate, our method solved the problem of detecting region duplication forgery with acceptable recall rate of 85% or higher. Among these, the proposed method provides a good balance between precision and recall followed by the Amerini et al. method [18].



(**a**)

Figure 11. *Cont.*



(**b**)

Figure 11. Comparison of detection performance in Precision and Recall of the proposed method, Pan and Lyu's method [38], and the Amerini et al. method [18].

5. Conclusions

Forged images created with duplicated and distorted regions are visually difficult to detect. An effective and robust forensic method based on angular radial partitioning and Harris key-points is proposed. We demonstrated the effectiveness and robustness of the proposed method with a series of experiments on realistic forged images with high resolutions from two image databases: MICC-F220 and image data manipulation. The experiment results showed that the proposed method can detect duplicated and multiple regions effectively, and with high accuracy, in the presence of several geometric transformation operations including (rotation and scaling), image degradations including JPEG compression and Additive White Gaussian Noise. The proposed method can detect rotated regions in multiples of 30 degrees and different rotation angles up to 360 degrees with estimation of rotation angles between duplicated regions. A current limitation of the proposed technique is that it cannot detect the duplicate regions when distorted with blurring and illumination changes. Therefore, involving blur invariant features and multiresolution local binary pattern descriptors is part of the future work that we are exploring to improve the technique.

Acknowledgments: The authors would like to thank the reviewers for their comments and suggestions for improving the paper. This research is funded by the Ministry of Higher Education Malaysia under the Fundamental Research Grant Scheme (FRGS). Project No.: FP073-2015A.

Author Contributions: Diaa M. Uliyan developed the algorithm; Hamid A. Jalab performed the experiments; Ainuddin W. Abdul Wahab analyzed the data; Somayeh Sadeghi contributed the Matlab software; Diaa M. Uliyan and Hamid A. Jalab wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Silva, E.; Carvalho, T.; Ferreira, A.; Rocha, A. Going deeper into copy-move forgery detection: Exploring image telltales via multi-scale analysis and voting processes. *J. Vis. Commun. Image Represent.* **2015**, *29*, 16–32. [CrossRef]
2. Chen, Y.-L.; Hsu, C.-T. Detecting recompression of JPEG images via periodicity analysis of compression artifacts for tampering detection. *IEEE Trans. Inf. Forensics Secur.* **2011**, *6*, 396–406. [CrossRef]
3. Li, C.-T. Source camera identification using enhanced sensor pattern noise. *IEEE Trans. Inf. Forensics Secur.* **2010**, *5*, 280–287.

4. Kee, E.; O'Brien, J.F.; Farid, H. Exposing photo manipulation with inconsistent shadows. *ACM Trans. Graph.* **2013**, *32*. [CrossRef]
5. Johnson, M.K.; Farid, H. Detecting photographic composites of people. In *Digital Watermarking*; Springer: Berlin, Germany, 2007; pp. 19–33.
6. Moghaddasi, Z.; Jalab, H.A.; Md Noor, R.; Aghabozorgi, S. Improving RLRN image splicing detection with the use of PCA and kernel PCA. *Sci. World J.* **2014**, *2014*. [CrossRef] [PubMed]
7. Ibrahim, R.W.; Moghaddasi, Z.; Jalab, H.A.; Noor, R.M. Fractional differential texture descriptors based on the machado entropy for image splicing detection. *Entropy* **2015**, *17*, 4775–4785. [CrossRef]
8. Moghaddasi, Z.; Jalab, H.A.; Noor, R.M. SVD-Based Image Splicing Detection. In Proceedings of the 2014 International Conference onInformation Technology and Multimedia (ICIMU), Putrajaya, Malaysia, 18–20 November 2014; pp. 27–30.
9. Moghaddasi, Z.; Jalab, H.A.; Noor, R.M. A Comparison Study on Svd-Based Features in Different Transforms for Image Splicing Detection. In Proceedings of the 2015 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW), Taipei, Taiwan, 6–8 June 2015; pp. 13–14.
10. Granty, R.E.J.; Aditya, T.; Madhu, S.S. Survey on passive methods of image tampering detection. In Proceedings of the 2010 International Conference on Communication and Computational Intelligence (INCOCCI), Erode, India, 27–29 December 2010; pp. 431–436.
11. Al-Qershi, O.M.; Khoo, B.E. Passive detection of copy-move forgery in digital images: State-of-the-art. *Forensic Sci. Int.* **2013**, *231*, 284–295. [CrossRef] [PubMed]
12. Uliyan, D.M.; Jalab, H.A.; Wahab, A.W.A. Copy move image forgery detection using hessian and center symmetric local binary pattern. In Proceedings of the 2015 IEEE Confernece on Open Systems (ICOS), Melaka, Malaysia, 24–26 August 2015; pp. 7–11.
13. Sadeghi, S.H.A.J.; Wong, K.S.; Uliyan, D.; Dadkhah, S. Keypoint based authentication and localization of copy-move forgery in digital image. *Malays. J. Comput. Sci.* **2016**. accepted.
14. Fridrich, A.J.; Soukal, B.D.; Lukáš, A.J. Detection of Copy-Move Forgery in Digital Images, Proceedings of the Digital Forensic Research Workshop, Cleveland, OH, USA, 5–8 August 2003; pp. 55–61.
15. Sheng, G.; Gao, T.; Cao, Y.; Gao, L.; Fan, L. Robust algorithm for detection of copy-move forgery in digital images based on ridgelet transform. In *Artificial Intelligence and Computational Intelligence*; Springer: Berlin, Germany, 2012; pp. 317–323.
16. Zimba, M.; Xingming, S. DWT-PCA(EVD) based copy-move image forgery detection. *Int. J. Dig. Content Technol. Its Appl.* **2011**, *5*, 251–258.
17. Huang, H.; Guo, W.; Zhang, Y. Detection of copy-move forgery in digital images using sift algorithm. In Proceedings of the Pacific-Asia Workshop on Computational Intelligence and Industrial Application, 2008 (PACIIA'08), Wuhan, China, 19–20 December 2008; pp. 272–276.
18. Amerini, I.; Ballan, L.; Caldelli, R.; Del Bimbo, A.; Serra, G. A sift-based forensic method for copy–move attack detection and transformation recovery. *IEEE Ttans. Inf. Forensics Secur.* **2011**, *6*, 1099–1110. [CrossRef]
19. Battiato, S.; Farinella, G.M.; Messina, E.; Puglisi, G. Robust image alignment for tampering detection. *IEEE Ttans. Inf. Forensics Secur.* **2012**, *7*, 1105–1117. [CrossRef]
20. Liu, G.; Wang, J.; Lian, S.; Wang, Z. A passive image authentication scheme for detecting region-duplication forgery with rotation. *J. Netw. Comput. Appl.* **2011**, *34*, 1557–1565. [CrossRef]
21. Ling, H.; Wang, L.; Zou, F.; Yan, W. Fine-search for image copy detection based on local affine-invariant descriptor and spatial dependent matching. *Multimed. Tools Appl.* **2011**, *52*, 551–568. [CrossRef]
22. Kakar, P.; Sudha, N. Exposing postprocessed copy–paste forgeries through transform-invariant features. *IEEE Ttans. Inf. Forensics Secur.* **2012**, *7*, 1018–1028. [CrossRef]
23. Bayram, S.; Sencar, H.T.; Memon, N. An efficient and robust method for detecting copy-move forgery. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2009), Taipei, Taiwan, 19–24 April 2009; pp. 1053–1056.
24. Shao, H.; Yu, T.; Xu, M.; Cui, W. Image region duplication detection based on circular window expansion and phase correlation. *Forensic Sci. Int.* **2012**, *222*, 71–82. [CrossRef] [PubMed]
25. Ryu, S.-J.; Kirchner, M.; Lee, M.-J.; Lee, H.-K. Rotation invariant localization of duplicated image regions based on zernike moments. *IEEE Ttans. Inf. Forensics Secur.* **2013**, *8*, 1355–1370.
26. Nock, R.; Nielsen, F. Statistical region merging. *IEEE Trans. Pattern Anal. Mach. Intell.* **2004**, *26*, 1452–1458. [CrossRef] [PubMed]

27. Christlein, V.; Riess, C.; Jordan, J.; Riess, C.; Angelopoulou, E. An evaluation of popular copy-move forgery detection approaches. *IEEE Ttans. Inf. Forensics Secur.* **2012**, *7*, 1841–1854. [CrossRef]
28. Shi, J.; Malik, J. Normalized cuts and image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 888–905.
29. Chen, T.-W.; Chen, Y.-L.; Chien, S.-Y. Fast image segmentation based on k-means clustering with histograms in HSV color space. In Proceedings of the 2008 IEEE 10th Workshop on Multimedia Signal Processing, Cairns, Australia, 8–10 October 2008; pp. 322–325.
30. Sekeh, M.A.; Maarof, M.A.; Rohani, M.F.; Mahdian, B. Efficient image duplicated region detection model using sequential block clustering. *Dig. Investig.* **2013**, *10*, 73–84. [CrossRef]
31. Tamura, H.; Mori, S.; Yamawaki, T. Textural features corresponding to visual perception. *IEEE Trans. Syst. Man Cybern.* **1978**, *8*, 460–473. [CrossRef]
32. Timm, N.H. *Applied Multivariate Analysis*; Springer: Berlin, Germany, 2007; SPIN 10848751.
33. Chalechale, A.; Mertins, A.; Naghdy, G. Edge image description using angular radial partitioning. *IEE Proc. Vis. Image Signal Process.* **2004**, *151*, 93–101. [CrossRef]
34. Harris, C.; Stephens, M. A Combined Corner and Edge Detector. In Proceedings of the Alvey Vision Conference, Manchester, UK, 2 September 1988.
35. Trujillo, L.; Legrand, P.; Olague, G.; Lévy-Vehel, J. Evolving estimators of the pointwise hölder exponent with genetic programming. *Inf. Sci.* **2012**, *209*, 61–79. [CrossRef]
36. Leys, C.; Ley, C.; Klein, O.; Bernard, P.; Licata, L. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *J. Exp. Soc. Psychol.* **2013**, *49*, 764–766. [CrossRef]
37. Mishra, P.; Mishra, N.; Sharma, S.; Patel, R. Region duplication forgery detection technique based on SURF and HAC. *Sci. World J.* **2013**, *2013*. [CrossRef] [PubMed]
38. Pan, X.; Lyu, S. Region duplication detection using image feature matching. *IEEE Trans. Inf. Forensics Secur.* **2010**, *5*, 857–867. [CrossRef]
39. Popescu, A.C.; Farid, H. Exposing digital forgeries in color filter array interpolated images. *IEEE Trans. Signal Process.* **2005**, *53*, 3948–3959. [CrossRef]