

# PASSIVE FORENSIC METHOD FOR DETECTING DUPLICATED REGIONS AFFECTED BY REFLECTION, ROTATION AND SCALING

*Sergio Bravo-Solorio and Asoke K. Nandi*

Department of Electrical Eng. & Electronics. The University of Liverpool  
Brownlow Hill, Liverpool, L69 3GZ, UK.  
email: {sbravo, aknandi}@liv.ac.uk

## ABSTRACT

Most of the existing forensic methods aimed at detecting duplicated regions in images are sensitive to geometrical changes in the replicated areas. Consequently, a simple reflection/rotation of the copied region could be used to not only suit the scene of the image, but also hinder its proper detection. In this paper, we propose a novel forensic method to detect duplicated regions, even when the copied portion have experienced reflection, rotation or scaling. To achieve this, overlapping blocks of pixels are re-sampled into log-polar coordinates, and then summed along the angle axis, to obtain a one-dimensional descriptor invariant to reflection and rotation. Moreover, scaling in rectangular coordinates results in a simple translation of the descriptor. This approach allows us to perform an efficient search of similar blocks, by means of the correlation coefficient of its Fourier magnitudes. Results are presented to demonstrate the effectiveness of the proposed method.

## 1. INTRODUCTION

Existing image editing software provides powerful tools that can be used to doctor images which cannot be easily identified, even by trained observers. A general classification of the techniques proposed to address this concern is shown in Table 1. The concept of Trustworthy Digital Camera, introduced by Friedman [4], can be considered as an *active* technique, whereby a digital signature is computed with the camera at the time of capturing, and then appended to the image as metadata. This method is *non-invasive* in the sense that the image is not distorted at all by this process. Fragile watermarking [2] is another *active* technique, whereby certain information is imperceptibly embedded in images for future authentication. Since certain distortion is unavoidable during the embedding process, this technique can be regarded as *invasive*. Active techniques are limited to controlled environments, such as surveillance cameras, where the capturing devices are equipped with a proper system to either compute the digital signature or embed the watermark. Passive image forensics [3, 10], in contrast, are designed to determine whether an image has been manipulated without relying on any information generated *a priori*. Sometimes, these techniques work on the assumption that forgeries may alter the underlying statistics of images, even though no traces can be identified under visual inspection [3]. In practise, because of the difficulty of the problem, forensic evidence will often rely on various methods, rather than one specific algorithm.

Sergio Bravo-Solorio is supported by the National Council of Science and Technology (CONACyT) of Mexico.

Table 1: General classification of techniques intended for integrity verification of images.

	Active	Passive
Non-invasive	Trustworthy digital cameras [4]	Passive image forensics [3, 10]
Invasive	Fragile watermarking [2]	-

A commonly considered manipulation consists in concealing undesired regions of the image by covering them with segments copied from another part of the same image [3, 10]. Forensic methods aimed at detecting duplicated regions must also consider other manipulations which could seriously affect their performance. In this paper, we explore the case where the copied region has undergone geometrical transformations such as reflection, rotation or scaling. This sort of distortions can be employed not only to suit the scene in the image being doctored, but also to hinder its detection with current forensic methods. We propose a novel forensic technique to detect duplicated regions in images by means of 1-D reflection/rotation-invariant descriptors derived from log-polar maps. Results show that the method can effectively detect duplicated regions affected by geometrical transformations. The paper is organised as follows. Related research is described in Section 2. Section 3 is focused on the proposed 1-D descriptors and its properties. In Section 4, the proposed method is presented. Some results are reported in Section 5, and the paper is concluded in Section 6.

## 2. PRIOR ART

There have been various recent forensic methods designed to detect duplicated regions. Fridrich et al. [3] presented an approach where the Discrete Cosine Transform (DCT) is computed on overlapping blocks of pixels. Then the coefficients are JPEG-like quantised and lexicographically sorted. Hence, matching pairs of blocks can be readily identified, as they are expected to have equal coefficients. Popescu and Farid [10] presented a similar approach where a dimension-reduced representation is computed for each block by means of Principal Component Analysis (PCA). The computational cost is significantly reduced with this technique, which is also robust to minor variations induced by additive noise or lossy compression. Langille and Gong [5] improved the efficiency of the previous methods by grouping the blocks with a modified k-dimensional tree. Then the detected duplicated regions are refined by employing a set of colour-based morphology operations. Mahdian and Saic [8] presented a

method based on blur moment invariants, where duplicated regions can be identified even after blur degradation, additive noise, or changes in contrast. In the method presented by Myna et al. [9], the Discrete Wavelet Transform (DWT) is firstly computed over the input image. Overlapping blocks of wavelet coefficients in the coarser sub-band of the lower resolution level are mapped to log-polar coordinates. The resulting blocks are lexicographically sorted and analysed to identify similar pairs. The formed pairs are iteratively compared in subsequent resolution levels, and discarded when the similarity criterion is not fulfilled. Results show that the algorithm is capable of detecting duplicated regions even after undergoing certain rotation. However, the lack of detail in the coarser sub-band in the lower resolution level can significantly increase the correlation between blocks, rising thus the number of blocks that must be analysed in subsequent resolution levels. This can significantly affect the computational cost, since the blocks size must be doubled for every resolution level.

### 3. SOLVING REFLECTION, ROTATION AND SCALING

To cope with reflection, rotation and scaling, blocks of pixels are mapped to 1-D descriptors resulting of summing the pixels in their log-polar representation along the angle axis. This approach was inspired by the watermarking technique in [6], which was designed to overcome the synchronisation problems caused by rotation, scaling and translation. In that case, 1-D descriptors were computed by summing a log-polar map along the log-radius axis.

#### 3.1 Computing the 1-D descriptors

Consider the point  $(x, y) \in \mathbb{R}^2$ , which can be written using (natural) log-polar coordinates,

$$\begin{aligned} x &= e^\rho \cos \theta , \\ y &= e^\rho \sin \theta , \end{aligned} \quad (1)$$

where  $\rho \in \mathbb{R}$  and  $0 \leq \theta < 2\pi$ . Let  $(x', y')$  denote the coordinates of a reflected, rotated and scaled point, i.e.,

$$\begin{aligned} x' &= \mu(x \cos \varphi + y \sin \varphi) , \\ y' &= \mu(x \sin \varphi - y \cos \varphi) , \end{aligned} \quad (2)$$

where  $\varphi$  and  $\mu$  are the parameters of rotation and scaling, respectively. Rewriting (2) in log-polar form we get,

$$\begin{aligned} x' &= e^{(\rho + \log \mu)} \cos(\varphi - \theta) , \\ y' &= e^{(\rho + \log \mu)} \sin(\varphi - \theta) . \end{aligned} \quad (3)$$

Observe that scaling in rectangular coordinates results in a simple translation of the log-polar map. Consider a block of pixels<sup>1</sup>  $B_i(x, y)$  and its log-polar representation  $B_i(\rho, \theta)$ . We can define a 1-D descriptor  $\vec{v}_i$  as

$$\vec{v}_i(\rho) = \sum_{\theta} B_i(\rho, \theta) . \quad (4)$$

<sup>1</sup>The index  $i$  is merely included for the sake of consistency in the notation employed in the subsequent section.

A reflected, rotated and scaled version of  $B_i$  can be expressed in log-polar coordinates as  $B'_i(\rho, \theta) = B_i(\rho + \log \mu, \varphi - \theta)$ . Computing a descriptor for this block, we get

$$\vec{v}'_i(\rho) = \sum_{\theta} B'_i(\rho, \theta) , \quad (5)$$

$$= \sum_{\hat{\theta}} B_i(\rho + \log \mu, \hat{\theta}) , \quad (6)$$

where  $\hat{\theta} = (\varphi - \theta)$ . Since cosines and sines are periodic functions, and we consider a whole period, and the change in  $\hat{\theta}$  is linear, it follows that we can re-write (6) as,

$$\vec{v}'_i(\rho) = \sum_{\theta} B(\rho + \log \mu, \theta) . \quad (7)$$

Hence, both descriptors are invariant to both reflection and rotation, i.e.  $\vec{v}_i(\rho) = \vec{v}'_i(\rho)$ , when  $B'_i$  is a mirrored and/or rotated version of  $B_i$ . Furthermore,  $\vec{v}_i(\rho) = \vec{v}'_i(\rho - \log \mu)$ , when  $B'_i$  is a scaled version of  $B_i$  by a factor  $\mu$ . In discrete signals though, we must consider loss due to interpolation and round-off errors. Thus, because of the well-known translation properties of the Fourier transform [1], the Fourier magnitude of both descriptors are expected to be very closely correlated to each other, i.e.,

$$c(\vec{V}_i, \vec{V}'_i) = \frac{\vec{V}_i^T \vec{V}'_i}{\sqrt{(\vec{V}_i^T \vec{V}_i)(\vec{V}'_i^T \vec{V}'_i)}} \approx 1 , \quad (8)$$

where  $c$  is the correlation coefficient,  $\vec{V}_i$  and  $\vec{V}'_i$  are the Fourier magnitudes of  $\vec{v}_i$  and  $\vec{v}'_i$ , respectively.

#### 3.2 Advantages of 1-D descriptors

Compared to complete log-polar maps, the main benefits of 1-D descriptors are,

1. *Reflection* - The detection of mirrored duplicated blocks is simplified by 1-D descriptors. Reflecting an image results in a flip of its log-polar map.
2. *Memory* - Since every single block will be compared with several blocks, it is a good idea to form an array with the descriptor corresponding to each block, before the search stage. Thus, instead of storing a complete  $m_\rho \times n_\theta$  log-polar map for every block, we only need to store a vector of length  $m_\rho$ .
3. *Computational cost* - The calculation of the proposed descriptors involves the same computational cost as the complete log-polar map. Nonetheless, the computational cost of whole algorithm significantly drops, as we only need to compute the Fast Fourier Transform (FFT) of vectors, whose complexity is known to be  $O(m_\rho \log m_\rho)$ , instead of computing the FFT of log-polar maps, whose complexity would be  $O((m_\rho n_\theta) \log(m_\rho n_\theta))$ .

## 4. PROPOSED ALGORITHM

Consider a colour image  $X$ , of size  $m_X \times n_X$ , which is tiled as blocks of pixels selected by sliding, pixel by pixel, a window of size  $q \times q$ , from the top-left corner to the bottom-right corner, in a raster-scan order. Let  $A_i$  denote the  $i$ -th block of pixels -for  $i = 1, \dots, (m_X - q + 1)(n_X - q + 1)$ . The proposed

method is comprised of three general stages, namely block sorting, search, and refinement. The algorithm is detailed below:

- 1: The blocks are sorted to reduce the computational cost of the search stage. The centre of each block  $A_i$  will be the centre of a disk of diameter  $q$ . Let  $f_{1_i}$ ,  $f_{2_i}$ , and  $f_{3_i}$  be the average of the red, blue and green colour components<sup>2</sup>, respectively, of the pixels within the disc. Additionally, the luminance of the pixels within the disc is computed as  $Y = 0.2126r + 0.7152g + 0.0722b$ , where  $r$ ,  $g$ , and  $b$  are components of red, green and blue, respectively. Then the entropy is calculated as,

$$f_{4_i} = - \sum_k p_k \log_2 p_k ,$$

where  $p_k$  is the probability of each luminance value in the disc. To reduce the occurrence of false matches, blocks where the computed entropy is lower than a predefined threshold  $e_{\min}$  are discarded. Then a list  $L$  is formed with the tuple of features  $(f_{1_i}, f_{2_i}, f_{3_i}, f_{4_i})$  corresponding to the remaining blocks.

- 2: Let us define  $\hat{L}$  as the result of lexicographically sorting the list  $L$ . Additionally, let  $B_i$  be the  $i$ -th block of pixels in  $\hat{L}$ , and the tuple  $(x_i, y_i)$  be the coordinates (the upper-left corner) of  $B_i$  in the image  $X$ .

A descriptor  $\vec{v}_i$  is computed for every  $B_i$  by (4), and its Fourier magnitude  $\vec{V}_i$  is calculated. Then, the correlation coefficient  $c_{ij} = c(\vec{V}_i, \vec{V}_j)$  is computed, for every  $j > i$  that satisfies the three following conditions:

- a)  $d_{ij} > \tau_d$ ,
- b)  $(f_{k_i} - \tau_a) \leq f_{k_j} \leq (f_{k_i} + \tau_a)$ , for  $k = 1, 2, 3$ ,
- c)  $(f_{4_i} - \tau_e) \leq f_{4_j} \leq (f_{4_i} + \tau_e)$ ,

where  $d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$ , and  $\tau_d$ ,  $\tau_a$ ,  $\tau_e$  are predefined thresholds. Note that, since  $\hat{L}$  is sorted, the comparisons for  $\vec{V}_i$  can stop once a descriptor  $\vec{V}_u$  is reached, such that  $f_{1_u} > (f_{1_i} + \tau_a)$ . Let us assume that  $c_{ir}$  was the higher correlation coefficient computed for  $\vec{V}_i$ . If  $c_{ir} > \tau_{\text{sim}}$ , a tuple  $(x\delta_{ir}, y\delta_{ir}, x_i, y_i, x_r, y_r)$  is appended to a list  $Q$ , where  $x\delta_{ir} = |x_i - x_r|$  and  $y\delta_{ir} = |y_i - y_r|$  are the offsets of the two pairs of coordinates.

- 3: The list  $Q$  is sorted in accordance with the offsets to form a new list  $\hat{Q}$ . Then,  $\hat{Q}$  is scanned to identify clusters with similar -i.e. not necessarily equal- offsets. Finally, a bitmap is encoded with the clusters that contain more blocks than a predefined threshold  $\tau_{\text{num}}$ .

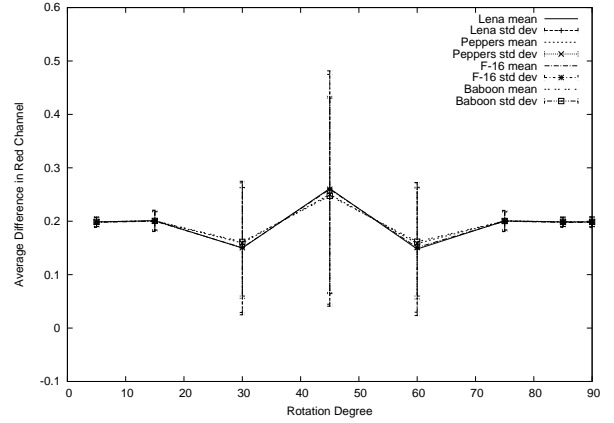
## 5. RESULTS

Before proceeding with the experiments, we will briefly describe the set-up of the system.

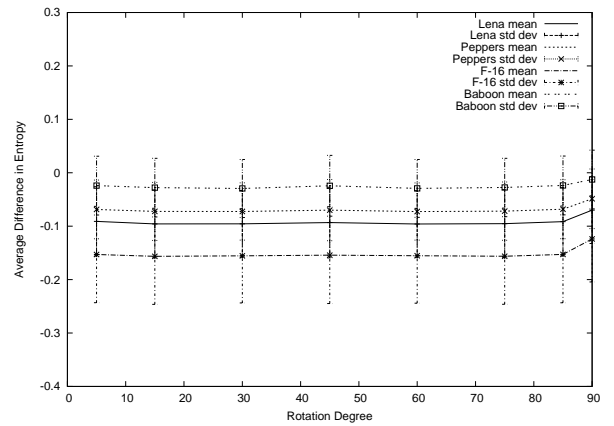
### 5.1 Parameters set-up

In our experiments, the blocks-size was set to  $16 \times 16$ , and the length of the 1-D descriptor computed for each block was set to 8. We have empirically found that such a configuration allows us to maintain a reasonable balance between the appearance of false matches and the capacity of the system

<sup>2</sup>Luo *et al.* [7] used the average of the colour channels and showed that such features are not significantly modified by JPEG compression or Gaussian blurring.



(a)



(b)

Figure 1: Distortion caused by rotation in the average of the red channel and the entropy.

to detect small duplicated regions. Additionally, we set the threshold  $\tau_d$  to 32, to avoid matches of blocks in a very close proximity to each other.

To determine the thresholds  $\tau_a$  and  $\tau_e$ , required in the search stage, we formulated an experiment with the standard images: Lena, Peppers, F-16 and Baboon. As in the proposed algorithm, the images were tiled in blocks by sliding a window of size  $16 \times 16$ , from the upper-left corner to the lower-right corner in raster scan order. Each block was translated<sup>3</sup> a random amount within the range  $[-0.5, 0.5]$ , in a random direction, and subsequently rotated  $5^\circ$ . We applied a mask to extract a disk of diameter 16 from both the original and the rotated blocks, and computed the difference between the average of the three colour channels and the entropy of the luminance channel (see stage 1 in the algorithm in Section 4). The experiment was repeated for different rotations, namely  $15^\circ$ ,  $30^\circ$ ,  $45^\circ$ ,  $60^\circ$ ,  $75^\circ$ ,  $85^\circ$  and  $90^\circ$ . We gathered these results, calculated their mean and their standard deviation. For the sake of brevity, we only present the graphs obtained for the red channel and the entropy in Figs. 1(a) and 1(b), respectively. We found that the distortion induced in the average value of the three channels, in 99.8% of blocks, was in the range  $[-1, 1]$ . Besides, the distortion induced in the entropy, in 96.6% of the blocks, was in the range  $[-0.3, 0.3]$ .

<sup>3</sup>The translation was necessary to reproduce a typical situation when the centre of rotation is not the exact centre of the block being analysed.

A similar experiment was formulated for different scaling factors, from 0.8 to 1.15 with increments of 0.05. We found that the distortion experienced by the average of the three colour channels is in the range  $[-1, 1]$  in 18% of the blocks. Moreover, in 94% of the blocks, the distortion induced in the entropy was in the range  $[-0.3, 0.3]$ . We decided to set the thresholds  $\tau_a$  and  $\tau_e$  to 1.0 and 0.3, respectively, to maintain a low computational cost, at the expense of reducing the capacity of detecting changes in scaling.

To define the threshold  $\tau_{\text{sim}}$ , we performed a similar experiment, but, instead of computing the features, we calculated the correlation coefficient between the Fourier magnitude of the descriptors extracted from the original and the distorted blocks. The threshold  $\tau_{\text{sim}}$  was set to 0.99985, as the correlation coefficient calculated from 98.2% of the blocks fell above this value.

## 5.2 Example forgeries

To show the capability of our method to detect duplicated regions affected reflection, rotation or scaling, we used the test images, shown in Figs. 2(a)-2(d), to create four different forgeries.

- A segment of the image in Fig. 2(a) was copied, without applying any further distortion, to conceal the woman near the centre, and get the forgery in Fig. 2(e).
- An horizontally flipped region of the Leaves image, in Fig. 2(b), was used to conceal the tree by the upper-right corner. The forgery is shown in Fig. 2(f).
- The person on the cliff, in Fig. 2(c), was covered by a 30° rotated portion of the rocks in the opposite side, to get the forgery in Fig. 2(g).
- A copy of the cow, in Fig. 2(d), was scaled 95% of its original size and superimposed in the left part to cover the other animals in the scene to get Fig. 2(h).

Let  $E$  be the bitmap map encoded by our system, and  $G$  be the ground truth, which shows the exact duplicated regions. The accuracy  $\alpha$  of our results, and the false positives  $\varepsilon$  (regions mistakenly detected as duplicates), will be given by,

$$\alpha = \frac{E \cap G}{G}, \text{ and, } \varepsilon = \frac{E \cup G}{G} - 1,$$

where the ideal detection would give  $\alpha = 1$  and  $\varepsilon = 0$ .

The results can be visually inspected in Figs. 2(i)-2(l), where the retrieved bitmaps are compared with their respective ground truth. The darker grey areas depict the duplicated regions that were successfully detected, the black areas illustrate undetected duplicated regions, and the white areas show the parts that were mistakenly deemed duplicates. The measures of accuracy and the false positive are summarised in Table 2. The system managed to detect the four forgeries successfully. As expected though, the best results were obtained from the Beach and the Leaves forgeries, as the copied blocks suffered no loss due to interpolation and round-off errors.

## 5.3 Further tests

We selected three images with particular characteristics. Figure 3(a) depicts a building with many symmetrical structures, which are expected to increase the false positives. An image with large textured regions is shown in Fig. 3(b). And Fig.

Table 2: Results obtained from the example forgeries.

Image	Accuracy ( $\alpha$ )	False positives ( $\varepsilon$ )
Beach	1.0	0.21
Leaves	0.86	0.0
Clif	0.63	0.21
Cow	0.49	0.04

3(c) depicts an image with large obscure regions, which are also expected to pose a challenge to the proposed method. A randomly chosen block of size  $52 \times 52$  was subjected to a geometrical distortion before being superimposed in another random location within the same image. From each test image, we formed four forgeries, each one involving a different distortion, namely, reflection, random rotation, random scaling, or no distortion at all (for comparison purposes). Results are summarised in Table 3. In all the cases, the duplicated regions were successfully detected in the Building image, in Fig. 3(a), and the symmetrical structures did not increase the false positives. The copied region was properly detected in the House image, in Fig. 3(b), except for the scaling test. Unexpectedly though, some regions by the building in the rear part were mistakenly deemed duplicates. The duplicated areas were successfully detected in the River image, in Fig. 3(c), in all the cases. Nonetheless, as expected, some regions in the bottom-left part were mistakenly deemed duplicates.

Table 3: Results obtained after applying different distortions to the duplicated region.

Distortion	Measure	Images		
		Building	House	River
Move	Accuracy ( $\alpha$ )	0.99	0.99	0.85
	F. positive ( $\varepsilon$ )	0.20	0.08	2.15
Reflection	Accuracy ( $\alpha$ )	0.53	0.92	0.78
	F. positive ( $\varepsilon$ )	0.01	0.04	2.10
Rotation	Accuracy ( $\alpha$ )	0.29	0.44	0.51
	F. positive ( $\varepsilon$ )	0.01	0.53	1.94
Scaling	Accuracy ( $\alpha$ )	0.22	0.0	0.95
	F. positive ( $\varepsilon$ )	0.20	0.50	2.57

## 6. CONCLUSIVE REMARKS

We have presented a forensic method to detect duplicated regions, even when the copied portion has experienced geometrical distortions such as reflection, rotation or scaling. The log-polar representation of a block of pixel is summed along its angle axis, to get a 1-D descriptor invariant to reflection and rotation. Reported results show the effectiveness of the proposed approach. Further work involves improving the last refinement stage to increase the accuracy and decrease the false positives. Moreover, we plan to exhaustively test the method in different images with several distortions, including combinations of the previous geometrical transformations, JPEG compression, additive noise and Gaussian blurring. Results are expected to be reported soon in a future publication.



Figure 2: Example forgeries. Original test images: (a) Beach. (b) Leaves. (c) Cliff. (d) Cow. Doctored images where the copied region has undergone: (e) no further distortion, (f) horizontal reflection, (g) 30° rotation, (h) 95% scaling. (i)-(l) Comparison between the obtained results and the ground truth.

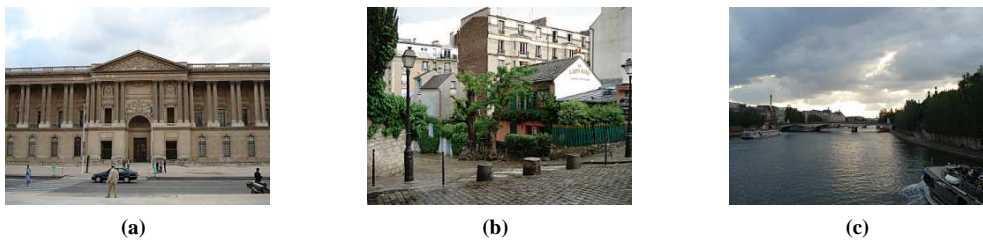


Figure 3: Test images. (a) Building. (b) House. (c) River.

## REFERENCES

- [1] R. N. Bracewell. *The Fourier transform and its application*. McGraw-Hill, 1986.
- [2] I. J. Cox, M. L. Miller, J. A. Bloom, J. Fridrich, and T. Kalker. *Digital Watermarking and Steganography*. Morgan Kaufman, 2nd edition, 2008.
- [3] J. Fridrich, D. Soukal, and J. Lukáš. Detection of copy-move forgery in digital images. In *Proc. of Digital Forensic Research Workshop*, August 2003.
- [4] G. Friedman. The trustworthy digital camera: restoring credibility to the photographic image. *IEEE Transactions on Consumer Electronics*, 39(4):905 – 10, 1993.
- [5] A. Langille and M. Gong. An efficient match-based duplication detection algorithm. In *Proc. of the 3rd Canadian Conference on Computer and Robot Vision*, page 64, 2006.
- [6] C.-Y. Lin, M. Wu, J. Bloom, I. Cox, M. Miller, and Y. Lui. Rotation, scale, and translation resilient watermarking for images. *IEEE Transactions on Image Processing*, 10(5):767 – 782, 2001.
- [7] W. Luo, J. Huang, and G. Qui. Robust detection of region-duplication forgery in digital images. In *Proc. of the IEEE International Conference on Pattern Recognition*, volume 4, pages 746–749, 2006.
- [8] B. Mahdian and S. Saic. Detection of copy-move forgery using a method based on blur moment invariants. *Forensic Science International*, 171(2-3):180–189, 2007.
- [9] A. N. Myna, M. G. Venkateshmurthy, and C. G. Patil. Detection of region duplication forgery in digital images using wavelets and log-polar mapping. In *Proc. of the International Conference on Computational Intelligence and Multimedia Applications*, pages 371–377, Washington, DC, USA, 2007.
- [10] A. C. Popescu and H. Farid. Exposing digital forgeries by detecting duplicated image regions. Technical report, Department of Computer Science, Dartmouth College, 2004.