Blind Tamper Detection to Copy Move Image Forgery using SURF and MSER

Kelsey Ramirez-Gutierrez, Mariko Nakano-Miyatake, Gabriel Sanchez-Perez, Hector Perez-Meana
Instituto Politecnico Nacional
Mechanical and Electrical Engineering School
Mexico, City, Mexico
hmperezm@ipn.mx

Abstract— **The sharing of digital images has become a common practice in our daily life, with the risk that these images can be accessed and easily modified by malicious people with the intention of causing moral or economic damage; or even to incriminate innocent people in legal issues. This paper proposes an algorithm to authenticate digital images by means of blind tampering detection against one of the principal manipulations that an image is put through, i.e. the *Copy-Move* which intends to erase or replicate a part of the image. The development and evaluation results of this proposal are presented in this paper.**

Keywords-Tamper detection; SURF; MSER; copy-move

I. INTRODUCTION

Nowadays, a huge amount of digital images, with or without commercial value, are easily shared among the general public via Internet or stored using any of the several available digital formats. Such images, which include private pictures or confidential images, have in general high quality and can be easily manipulated using computational tools such as: Photoshop®, Corel Paint Shop®, etc. Such kind of malicious attacks can be divided in copy-move and cut-and paste attacks. The copy-move is one of the most studied forgery techniques which consist in copying a portion of an arbitrary size and shape of a given image and pasting it in another location of the same image. Clearly, this technique is useful when the forger wants either to hide or duplicate something that is already present in the original image [1][2]. On the other hand, in the cut-and-paste attack or splicing, the attacker firstly chooses a region of a given image and pastes it into a second one, usually to alter its content and meaning. Splicing is probably more common than the copy-move attack, because it is far more flexible and allows the creation of images with a very different content with respect to the original image [2].

The image authentication has been a topic of active research during the last several years, because the tampered images may cause moral or economic damages to the persons related to the maliciously modified images, giving as a result the publication of several image authentication techniques, which can be broadly classified into two types: active and passive image authentication methods. The main difference among them is that in the active methods some useful information is extracted from the image to be authenticated and embedded in it or stored separately. This information is then used during the authentication process. On the other hand, in the passive methods, also called forensic methods, the authentication must be carried out without previous information about the processing that the image to be authenticated had passed through [1][2].

The active methods can be classified into two categories: the watermarking-based and the image hashing-based schemes, both of them with advantages as well as some drawbacks. The watermarking-based techniques embed an imperceptible signal into the image to be authenticated to create a watermarked image. The embedded signal can be a random signal or a signal derived from the image to be authenticated. During the authentication process, the watermark is extracted from the watermarked image to be used for authentication purpose [3] or even to restore the tampered image. Several high performance methods for embedding information into digital images have appeared in the literature [3][4][5]. These methods perform fairly well and in several cases have the ability to restore the tampered regions [3]. However, if the parameters are not properly chosen some distortion may be introduced in the image to be protected [3]. On the other hand, the image hashing-based techniques, or multimedia fingerprinting, take out a set of robust features from the image to be authenticated to create a compact code, called perceptual hashing code, which is stored or transmitted separately. During the authentication process, employing the same method, an authentication code is extracted from the suspicious image, which is then compared with the stored code and if the difference between both codes is smaller than a given threshold the suspicious image is considered as authentic; otherwise it is determined as a tampered image. It is necessary to point out that the perceptual hashing technique is different from the cryptographic hashing since in the last one, any change in the image to be authenticated, even if it is perceptually similar to the original one, produces a quite different hash value [6]; while the perceptual hashing technique has the capacity of discriminating between malicious attacks and distortions resulting for standard image processing tools. Because these methods have proved to be very efficient, several image hashing algorithms [6][7][8] have been suggested. These methods do not distort the image, although the authentication code must be stored or transmitted separately.

In many practical situations the investigators have only the image under analysis, such that passive image authentication schemes are required, which carry out the authentication process analyzing the processing artifacts to infer the potential alterations introduced during the image generation process [1]. This paper analyzes image authentication schemes to deal with images tampered using the copy-move scheme. This tampering method creates a forged image by copying a certain portion of an image and moving it to another part of the same image. [1]. The main characteristic of this kind of tampered images is that, because the duplicated region is picked from the image itself, the noise components, texture and color patterns are compatible with the rest of the image. This fact makes it not easy to detect this kind of forgeries.

The authentication of this kind of tampered images has many important practical applications giving as a result the proposal of several authentication algorithms during the last several years. Among them, there is the feature matched technique proposed by Pan and Lyu [10], which employs local statistics features together with a verification step which tries to find duplicated regions using normalized correlations maps and thresholding. The main weakness of this method is its lack of accuracy [10]. Jaberi et al. [11] proposed a copy-move detection method in which firstly the Scale Invariant Feature Transform (SIFT) [12][13] is used to extract the key points, then the affine transform of a region around each key point is estimated and finally, to reduce the false detection, the Dense Mirror Invariant Feature Transform (MIFT) is estimated [11]. This scheme performs fairly well although it does not work well if the duplicated region corresponds to a flat surface where not key points are located. Kumar et al. [14] propose an image authentication scheme in which the image under analysis is divided in overlapping sub blocks which are then transformed to the frequency domain using the Discrete Cosine Transform (DCT), keeping only the lowest frequency components. These components are then ordered in a lexicographic way to carry out the evaluation of each sub block. This scheme performs well when the duplicated regions do not presents scaling or rotational distortions. A similar approach was also proposed by Fridrich [15] which presents the same advantages and disadvantages. Popescu [16] proposes to replace the DCT by a Principal Component Analysis (PCA) to reduce the dimensionality of features vector. However this method lacks of robustness against even small rotations of the copy moved regions. Several other methods have been proposed, some of them based on intensity method, which assume that the image may be under any JPEG, rotation or scaling operations [17]. To solve some of the problems still present in the image authentication algorithms describe above, this paper proposes an algorithm that allows the authentication of digital images that have gone under copy-move tampering attacks. Evaluation results show that the proposed scheme performs fairly well when it is required to authenticate tampered images, even when the duplicated region has been rotated and scaled.

The rest of the paper is organized as follows. Section II provides a detailed description of proposed algorithm. In Section III the evaluation results using the CASIA database [21] are given. Finally Section IV provides the conclusions of the paper.

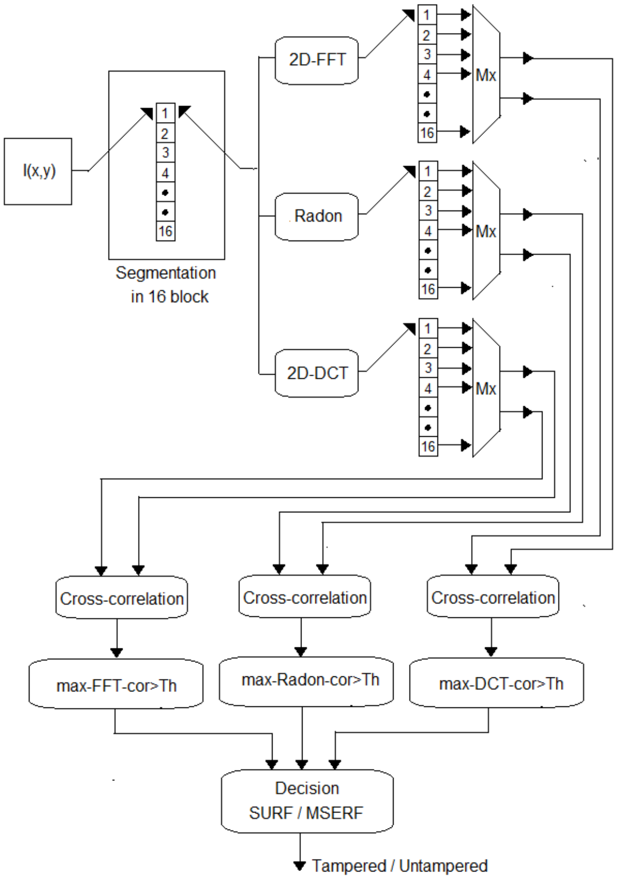
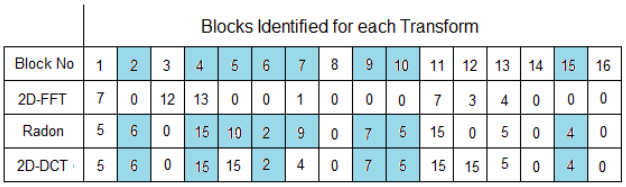


Figure 1 Proposed image authentication system

II. PROPOSED IMAGE AUTHENTICATION SYSTEM

The proposed image authentication system is shown in Fig. 1. Here the image under analysis is converted to a gray scale image and divided in 16 blocks. Next, the magnitude of the bi-dimensional Fast Fourier Transform (2D-FFT) [18], the Discrete Radon Transform (DRT) [7][9] and 2D-DCT [18] of each block are estimated. The main idea behind the proposed schema is to take advantage of the translation invariance property of the 2D-FFT, the rotation and scaling properties of the DRT and the compression capability of the DCT. Next, the cross correlation between the 16 blocks in each domain is calculated and the block index with the higher correlation values greater than a given threshold, are kept to form a matrix of 3×16 elements. Here, the threshold is given by the highest correlation value between the 16 blocks. At the end of this process a matrix of 3×16 elements is obtained containing the possibly tampered blocks. The second part of the authentication process can be more easily explained with the example shown in Fig. 2. Here, we look for a block which can be found as tampered by at least two of the three frequency transformations applied and that also correspond to the block that is being compared to.



	Blocks Identified for each Transform															
Block No	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
2D-FFT	7	0	12	13	0	0	1	0	0	0	7	3	4	0	0	0
Radon	5	6	0	15	10	2	9	0	7	5	15	0	5	0	4	0
2D-DCT	5	6	0	15	15	2	4	0	7	5	15	15	5	0	4	0

Figure 2. Block identified as tampered by each transform.

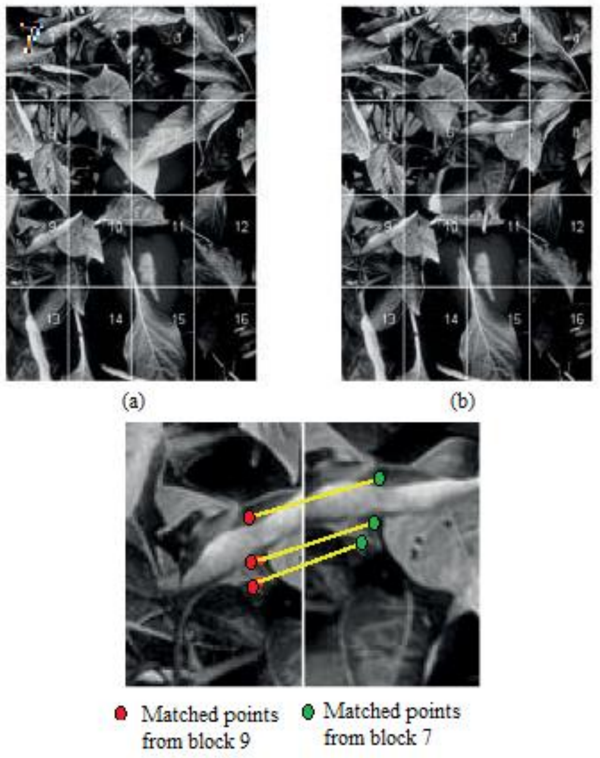


Figure 3. (a) Original image, (b) tampered image, (c) matched points in blocks 7 and 9.

For example, block 2 is highly related with block 6, according to Radon and 2D-DCT transforms. In a similar form block 6 appears to be related with block 2. Thus, blocks 2 and 6 are the first blocks to be compare using a SURF detector [19]. After this evaluation, if at least one matched point is found the block is considered as tampered; otherwise the block is labeled as untampered. The system continues analyzing the next block that is found to be related with other block according to at least two transformations. For example, block 4 is related with block 15. Again, blocks 4 and 15 are compared among them using the SURF (Speeded Up Robust features) [19] and Maximally Stable Extremal Regions, (MSER) [20]; and labeled as tamper or untampered depending if there are matched points or not inside them. This process continues until all blocks of the image related among them according to at least two transforms are labeled as tampered or untampered. Next, if after all blocks are analyzed the decision is that all of them are untampered, the blocks related with other block according to only one transform are analyzed. For example block 4 is related with the block 13 according with the 2D-FFT and block 9 and with block 7. After applying the SURF [19] and MSER [20], it was found that in block 9 and block 7 at least one matched point is found, as shown on Fig. 3 and then the system decides that the image was tampered. This process can be repeated in each one of the 16 regions for a more accurate evaluation to detect region duplications inside each sub-block. This fact reduces the computational complexity avoiding the use of overlapping blocks. Next subsections describe each stage of proposed system.

A. 2D- Discrete Fourier Transform

The bi-dimensional Discrete Fourier Transform (2D-DFT) has found a large amount of applications in several fields, because it and its inverse allow analyzing the frequency spectral characteristics of images [18]. The 2D-DFT pair is given by

$$F(u,v)=\frac{1}{NM}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)e^{-j2\pi\left(ux/M+vy/N\right)} \quad (1)$$

$$f(x,y)=\frac{1}{NM}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} F(u,u)e^{j2\pi\left(ux/M+vy/N\right)} \quad (2)$$

Some general statements can be made about the relationship between the frequency components of the Fourier transform and spatial features of an image. For instance, because the frequency is directly related to the spatial changes rate, it is not difficult intuitively to associate frequencies in the 2D-DFT with intensity variations patterns in an image, because the low frequencies correspond to the slowly varying intensity components of an image and the higher frequency components correspond to the faster intensity changes in the image [18]. Other important feature is the fact that the magnitude of the 2D-DFT is translation invariant, i.e.

$$\begin{aligned}&IF \quad f(t)\leftrightarrow F(\omega)\\ &and \;\; f_1(t)=f(t-T)\leftrightarrow F_1(\omega)=F(\omega)e^{-j\omega T}\\ &Then\; |F(\omega)|=|F_1(\omega)|\end{aligned} \quad (3)$$

B. *Radon Transform*

The Radon Transform [7] is used in this proposal because it is robust against rotation, scaling and translation. The Radon transform for a set of parameters (ρ, θ) is the line integral through the image $f(x,y)$, where the line corresponding to the value of (ρ,θ) is given by (4)

$$g(\rho,\theta) = \sum_{m-s_{\max}}^{s_{\max}} \sum_{n=-s_{\max}}^{s_{\max}} f(m,n)\delta(\rho - m\cos\theta - n\sin\theta) \qquad (4)$$

where $\delta(\eta)$ is the Dirac delta function which is equal to one when η=0 and zero for all other arguments [7][9]. The definition of the Radon transform forces the summation of $f(m,n)$ along the line ρ=$mcos\theta$+$nsin\theta$ and consequently the value $g(\rho,\theta)$ for any (ρ,θ) is the sum of the value of $f(m,n)$ along this line [7]. The Radon transform has the following useful properties for the affine transformations of an image [7][9].

1. The translation of an image by (xo,yo) causes the Radon transform to be translated in the direction of s

$$f(m - m_0, n - n_0) \leftrightarrow g(s - m_0\cos\theta - n_0\sin\theta, \theta) \qquad (5)$$

2. Scaling (retaining aspect ratio) of an image by a factor ρ (ρ>0) causes the Radon transform to be scaled through the same factor

$$f(\rho m, \rho n) \leftrightarrow \frac{1}{|\rho|} g(\rho s, \theta) \qquad (6)$$

3. Rotation of an image by an angle θ causes the Radon transform to be shifted by the same amount

$$\begin{aligned} &f(m\cos\theta_r - n\sin\theta_r, m\sin\theta_r + n\cos\theta_r) \\ &\qquad \leftrightarrow g(s, \theta - \theta_r) \end{aligned} \qquad (7)$$

C. *2D-Discrete Cosine Transform*

The 2D Discrete Cosine Transform is widely used in image compression applications, because it is able to represent a given image with a reduced number of coefficients, besides that the DCT of a real a valued signal is also real valued. The general equation of a 2D-DCTof an image of N×M pixels, *f(m,n),* is given by [18]

$$\begin{aligned} F(u,v) = (4/MN)^{1/2} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} A(i)A(j)f(m,n) \times \\ \cos(\pi u(2i+1)/2N)\cos(\pi v(2j+1)/2M) \end{aligned} \qquad (8)$$

where

$$A(\xi) = \begin{cases} \dfrac{1}{\sqrt{2}}, & for\,\xi = 0 \\ \xi = 0 & otherwise \end{cases} \qquad (9)$$

D. *SURF Detector*

The *SURF* [19] employs integral images and efficient scale space construction to generate key points and descriptors very efficiently. SURF uses two stages namely key point detection and key point description. The detector is based on the Hessian matrix with the Laplacian-based detector. It relies on integral images to reduce the computation time and therefore call it the “Fast-Hessian” detector. The descriptor, on the other hand, describes a distribution of Haar-wavelet responses within the interest point neighborhood. In the first stage, integral images allow the fast computation of approximate Laplacian of Gaussian images using a box filter. The computational cost of applying the box filter is independent of the size of the filter because of the integral image representation. The determinants of the Hessian matrix are then used to detect the key points, because of its good performance in computation time and accuracy. It relies on the determinant of the Hessian for both location and the scale. Given a point p=(x,y) in an image *I(x,y)*, the Hessian matrix section $H(x,\sigma)$ in p at scale σ is defined as follows

$$H(p,\sigma) = \begin{bmatrix} L_{xx}(p,\sigma) & L_{xy}(p,\sigma) \\ L_{yx}(p,\sigma) & L_{yy}(p,\sigma) \end{bmatrix}, \qquad (10)$$

where $L_{xx}(p,\sigma)$, $L_{xy}(p,\sigma)$, $L_{yx}(p,\sigma)$, $L_{yy}(p,\sigma)$ are the convolution of the Gaussian second order derivative with respect to x and y, respectively, with the image *I(x,y)* in the point p [19]. The SURF builds its scale space by keeping the same image size while varying only the filter size. In the final stage, to each detected key point is firstly assigned a reproducible orientation. For orientation, the Haar wavelet responses in x and y directions are calculated for a set of pixels within a radius of 6σ where σ refers to the detected key point scale. The SURF descriptor is then computed by constructing a square window centered on the key point and oriented along the orientation obtained before. This window is divided into 4 x 4 regular sub-regions and Haar wavelets of size 2σ are calculated within each sub-region. Each sub-region contributes 4 values thus resulting in 64D descriptor vectors which are then normalized to unit length. The resulting SURF descriptor is invariant to rotation, scale and contrast; besides that it is also partially invariant to some other transformations [19].

E. *MSER Detector*

The Maximally Stable Extremal Regions (MSER), proposed by Matas et al. [20], estimates a set of distinguished regions that are detected in a gray scale image and defined by an extremal property of the intensity function in the region and on its outer boundary. The MSER has properties that allow it to achieve a superior performance as stable local detector compared with other local point detectors. Two of the main properties of the set of MSER are that it is closed under continuous geometric transformations and invariant to affine intensity changes. Furthermore the MSER regions are detected at different scales. Some other important properties of the *MSER* detector are:

a) Invariance to affine transformation of image intensities.
b) Covariance to adjacency preserving (continuous) transformation *T*: *D* on the image matched point domain.
c) Stability of the detected regions which means that only the regions whose support is nearly the same over a range of thresholds is selected.
d) Multi-scale detection without any smoothing involved, thus both fine and large structure is detected.

The set of all extremal regions that can be enumerated in worst-case is of $O(n)$, where n is the number of pixels in the image.

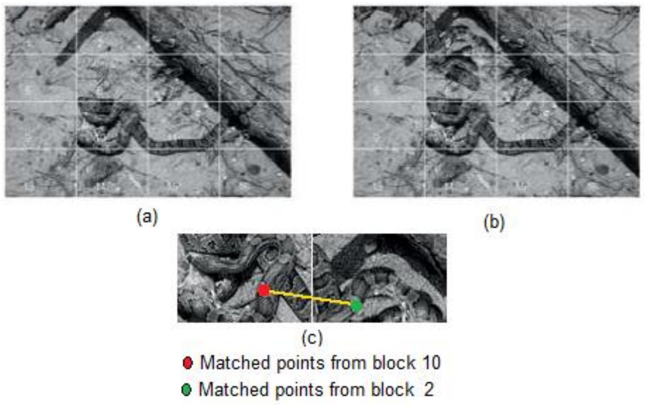


Figure 4. Performance of proposed scheme (a) Original image, (b) tampered image and c) matched points.

III. Evaluation Results

To evaluate the proposed images authentication system, the CASIA Image Tampering Detection Evaluation Database [21] was used, which consists of 102 Images. To obtain a correct operation, the parameters of proposed system are set as follows: The threshold value used to determine if two blocks in each transformation domain are similar using the cross correlation among them is the mean value of the highest correlation among the 16 blocks. For the Radon Transform the projection of the image intensity along the 180 angles equally spaced in the interval $0 \le \theta < \pi$ are analyzed. For the SURF detector the number of octaves is set equal to 3, which will give a filter size of 27x27. Finally for the MSER detector the step size between intensity threshold levels is set equal to 0.8.

Table I shows the detection performance of proposed system when it is required to evaluate both tampered and original images, where a false positive is an error in which the test result indicates that an image is tampered when it is an original one, while a false negative is an error produced when the test result indicates that the image is original, although it is in fact tampered.

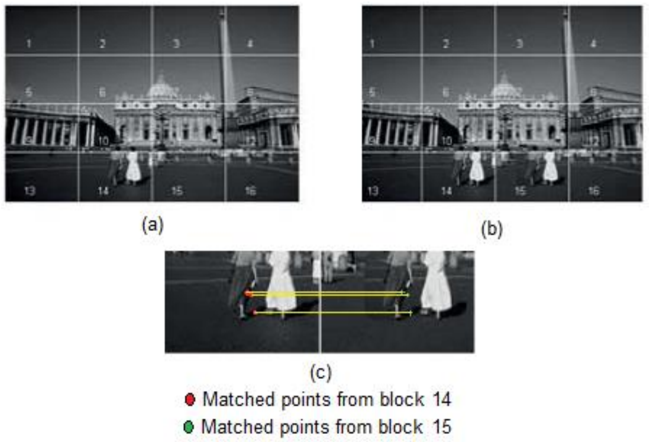

1
2
3
4
5
6
10
13
14
15
16
(a)
(b)
(c)
● Matched points from block 14
● Matched points from block 15

Figure 5. Performance of proposed scheme (a) Original image, (b) tampered image and c) matched points.

Figures 4 and 5 show the evaluation results in which the proposed scheme correctly detects a tampered image. In both cases the original and tampered images are shown in (a) and (d), while in (c) the matched points obtained by the SURF and MSER are shown which confirm that the image under analysis was tampered is shown in (c). In some cases, depending on the alteration introduced on the original image, for example if the pasted object suffer some affine transformation, the SURF features are not robust enough to detect these changes, so in this case to use the MSER features may allow to detect that the image under analysis was tampered, as shown in Figs. 6 and 7. Finally Table II shows the evaluation results obtained using the Mean Opinion Scoring (MOS) criterion in which several images were presented to 100 peoples who were asked to determine if the images were tampered or untampered.

The total time of calculation evaluating the three transformation techniques and the SURF and MSER detector is of 3.53 minutes.

TABLE I. TAMPER DETECTION PERFORMANCE OF PROPOSED ALGORITHM

Success rate	False positive	False negative
75%	10%	15%

TABLE II. TAMPER DETECTION PERFORMANCE OF PROPOSED ALGORITHM USING THE MOS CRITERION

Success rate	False positive	False negative
53%	20%	27%

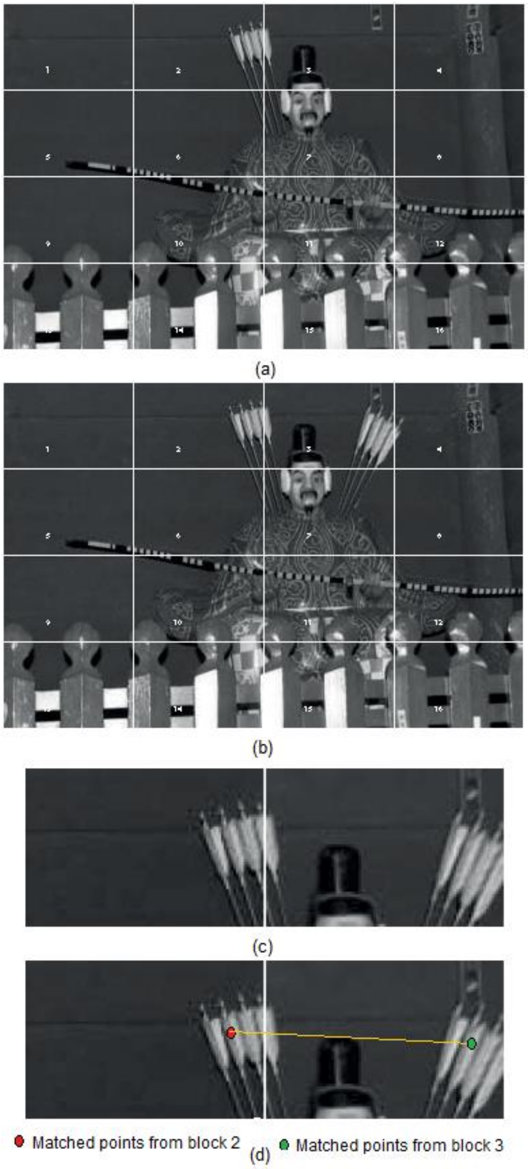


Figure 6. Performance of proposed scheme (a) Original image, (b) tampered image and c) matched points detected using SURF, (d) matched points detected using MSER.

IV. CONCLUSIONS

This paper proposes a copy-move tamper detection algorithm in which firstly the image under analysis is divided in 16 blocks and the 2D-DCT, 2D-FFT and DRT of each block is estimated. Then, the similitude between such blocks, in each domain, is estimated using the maximum of the cross correlation value together with the SURF detector and MSER features to determine if the image was tampered. From the evaluation results presented in this paper we can observe that the proposed scheme is able to identify the copy-move regions of the image under analysis. We must add that this method is not trying to identify any particular type of copy-move forgery mechanism, like rotation, or scaling, or JPEG compression. Instead it is intended to be a more general method able to operate in almost any situation and that, combined with other methods can lead to an accurate detection of a specific type forgery attack.

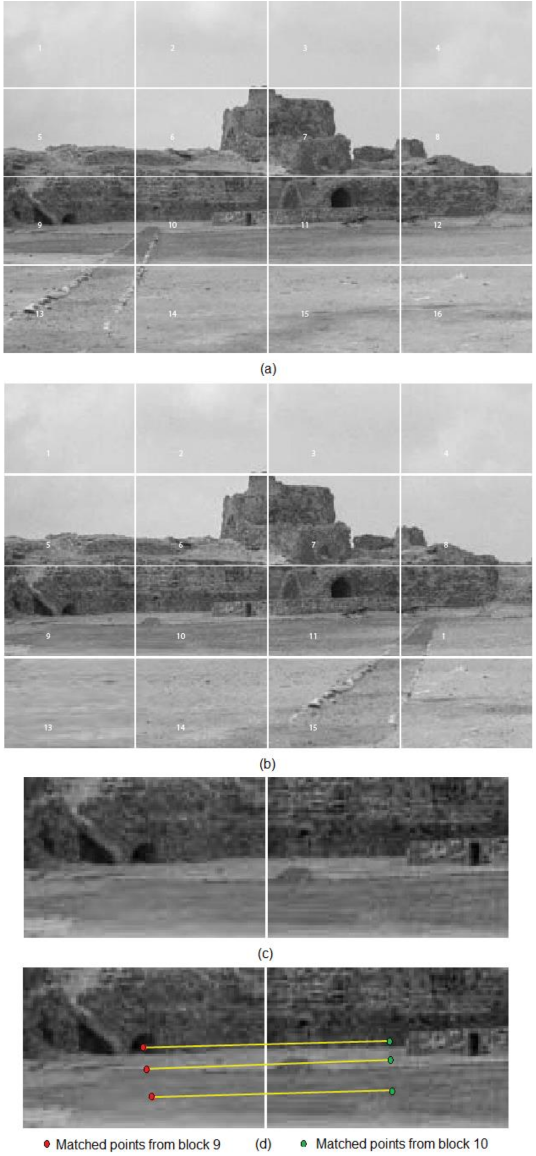


Figure 7. Performance of proposed scheme (a) Original image, (b) tampered image and c) matched points detected using SURF, (d) matched points detected using MSER.

ACKNOWLEDGMENT

We thank the National Science and Technology Council (CONACYT) and to the National Polytechnic Institute of Mexico (IPN) for the financial support provided during the realization of this research.

REFERENCES

[1] M. Kirchner, “Notes on digital image forensics and counter-forensics,”http://dud.inf.tu-dresden.de/~kirchner/Documents /image_forensics_and_counter_forensics.pdf, Oct. 2011.

[2] A. Piva, *An Overview on Image Forensics*, ASRN Signal Processing, Hindawi, http://dx.doi.org/10.1155/2013/496701, 2012.

[3] L. Rosales-Roldan, M. Cedillo-Hernandez, M. Nakano-Miyatake, H. Perez-Meana, B. Kurkoski, "Watermarking-basewd image authentication with recovery capability using halftoning technique", Signal Processing: Image Communications, vol. 28, pp. 69-83, Jan 2013.

[4] M. Cedillo-Hernández · F. García-Ugalde · M. Nakano-Miyatake, H. Manuel Perez-Meana, "Robust hybrid color image watermarking method based on DFT domain and 2D histogram modification", Journal of Signal Image and Video Processing, April, 2013, doi 10.1007/s11760-013-0459-9.

[5] I. Ismali, S. El-Zoghdy, and A. Abdo, "A novel techinque for data hiding," International Journal of Computers and Applications, vol. 32, pp. 119–124, Jan. 2010.

[6] A. Swaminathan, Y. Mao, and M. Wu, "Robust and secure image hashing," *IEEE Transactions on Information Forensics and Security*,, vol. 1, pp. 215–230, June 2006.

[7] J. S. Seo, J. Haitsmab, T. Kalkerb, and Y. C. D., "A robust image fingerprinting system using the radon transform," *Signal Processing: Image Communication*, vol. 19, pp. 325–339, April, 2004.

[8] Y. Li, Z. Lu, C. Zhu, and X. Niu, "Robust image hashing based on random gabor filtering and dithered lattice vector quantization," IEEE Transactions On Image Processing, vol. 99, pp. 1–14, Jan. 2011.

[9] D. Q. Nguyen, L. Weng, and B. Preneel, "Radon transform-based secure image hashing," International Conference on Communications and Multimedia Security, Springer-Verlag, pp. 186–193, Oct. 2011.

[10] X. Pan and S. Lyu, "Region duplication detection using image feature matching", IEEE Trans. On Forensic Security, pp. 857-867, Dec. 2010.

[11] M. Jaberi, G. Bebis, M. Hussain, G. Muhammad, "Improving the detection and localization of duplicated regions in copy-move image forgery", Proc. Int. Conf. on Digital Signal Processing, pp. 1-6, 2013.

[12] H. Bay, T. Tuytelaars, , and L. V. Gool, "Surf: Speeded up robust features," ETH Zurich and Katholieke Universiteit Leuven, Tech. Rep., 2006.

[13] N. Y. Khan, B. McCane, and G. Wyvill, "SIFT and SURF performance evaluation against various image deformations on benchmark dataset," International Conference on Digital Image Computing: Techniques and Applications, pp. 501–506, Dec. 2011.

[14] S. Kumar, J. Desai, S. Mukherjee, "A fast DCT based method for copy move forgery detection", Proc. Int. Conf. on Image Information Processing, (CIIP'13), pp. 1-6, 2013

[15] J. Fridrich, D. Soukalm, J. Luka, "Detection of copy move forgery in digital images," Proc. Digital Forensic Research Workshop, pp. 19-29, 2003.

[16] A. Popescu, H. Farid, "Exposing Digital Forgeries by detecting duplicated regions," Tech. Report TR2004-515, Dartmouth Collage, Hanover, 2004.

[17] M. A. Qureshi, M. Deriche, "A review on copy move image forgery detection techniques," Proc Multi-Conference Signal, Systems and Devices, pp. 1-5, 2014.

[18] R. Gonzalez and R. Woods, *Digital Image Processing*, Prentice Hall, Englewood Cliffs, NJ, 2008.

[19] H. Bay, T. Tuytelaars, L. Van Gool, "SURF: Seed Up Robust Features", Proc. Int. Conf. on Computer Vision (ECCV´06), May 2006

[20] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide baseline stereo from maximally stable extremal regions," Image Vision and Computing, vol. 22, pp. 761-767, 2004.

[21] CASIA Image Tampering Detection, Evaluation Database http://forensics.idealtest.org:8080/index_v1.htm. June 2014.