Accurate and Robust Identifying Forged Region Method in Scanned Images

Zeinab F. Elsharkawy, Safey A. Abdelwahab, Sayed M. Elaraby
Engineering Department, Nuclear Research center, Atomic Energy Authority
Cairo, Egypt

Moawad I. Dessouky, Fathi E. Abd El-Samie
Department of Electronics and Electrical Communications, Faculty of Electronic Engineering, Menoufia University
Menouf, Egypt

ABSTRACT
Digital images have been spread all over the world. With rapidly development and ease of use of digital image editing tools like Photoshop TM and paint TM, it is important to authenticate or detect the forged regions in any suspicious digital image. In this paper, new and robust method for authenticating and identifying the forged regions in the scanned images is proposed. The method is based on using dust/scratch and source imperfection pattern of scanned images to identify the forged regions. Each suspicious image is divided into non-overlapping blocks. The correlation features between each block and acquisition scanner's template have been extracted to identify the correct scanned blocks and the tampered ones. Four groups of the tampered images are tested. The experimental results have approved the validity, efficiency, and robustness of the proposed method to identify the tampered images and define their forged regions. The proposed results are then compared to the results of the two previously published methods. The proposed method is simple and easy to apply to all types of the tampered images, regardless the acquisition source noise or the image contents.

General Terms
Pattern Recognition, Digital Forensics.

Keywords
Digital Forensics, Scanner Identification, Forgery Detection, Image Processing.

1. INTRODUCTION
Digital images acquisition devices such as camcorders, cameras, and scanners are became very popular due to the rapid improvement of digital technologies. In today's digital age, the creation and manipulation of digital images is very simple by low-cost or even free software tools (such as Picasa, GIMP, and Photoshop software) that are easily and widely available. Digital images are used in military, medical, broadcasting, certification applications, and also in our daily life. Hence, it is important to authenticate these images.

One of the main problems in authenticating a digital image is identification of source acquisition device. Among various image acquisition devices, scanners are chosen to be experimentally investigated, because it is the commonly used device for transforming images and documents to digital form that can be forged such as fake passports and counterfeit notes [1], [2].

Several methods are introduced to identify the source scanner. One of them is based on extracting various statistical features, such as mean, median, mode, maximum, and minimum from the noise of the scanned images [2-6]. In [7-10] support vector machine (SVM) has been trained with these features to identify the source scanner. In [11] SVM has been trained again by these statistical features to detect forged regions in the scanned images. The forged images had been synthesized using scanned images from different scanners. Image acquisition devices such as scanner, camera, and a computer generated graphic has been classified in [12]. Scanner identification using spectral noise in the frequency domain has been proposed in [13]. In [14], this method has been used to detect forged regions in scanned images and the forged images synthesized using the scanned images and images acquired by digital camera. Dust and scratch positions on flatbed scanners have been used by us for scanner identification in [15], and SVM classifier has been implemented and trained using correlation features of scanned images to classify different scanners' brands and models. The robustness of the used individual source scanner identification scheme on resized and different resolutions scanned images is also proposed.

In this paper- a new and accurate method to identify forged regions in scanned images has been introduced. The traces of dust, dirt, and scratches over different scanners platen on the scanned images are used to create unique patterns. If there are some glass defects and scratches over the platen, their positions do not change, even if they are cleaned manually. The scratches and dusts which are strongly adhered to the scanner platen can create a unique pattern of the scanner. Since the relative scratch and debris positions are fixed and not affected by scanner use, which are used later to formulate the fingerprint of each scanner as explained in detail in our previous work [15]. This fingerprint is finally used to match the scanned images to their sources. This procedure has been used here to detect forged regions in the scanned images. Each given image has been sliced into non-overlapping blocks, then each block in correlated with the scanner template by Normalized Cross Correlation NCC after filtering using high pass filter and the NCC output is applied to detected threshold to authenticate this block. Four groups of tampered images are tested, the first group is synthesized using two scanned images from different scanners, the second group is synthesized using scanned images and images acquired using digital camera, the third group is synthesized using scanned image and images acquired using cellular phone camera, and the fourth group is synthesized using scanned image and computer generated image. The experimental results have approved the validity, efficiency, and robustness of the proposed method to identify forged image blocks in the suspicious images.

The rest of the paper is organized as follows. In Section II, the forged region identification method is explained in details. The efficiency of the forged region identification method is substantially verified by experimental results in Section III. Finally, the conclusion is presented in Section IV.

2. THE FORGED REGION IDENTIFICATION METHOD

In this section, the forged region identification method is explained in detail. The method consists of three steps. First, creation of scanner template by scanning various images on the scanner to detect the dust and scratch positions and create a fixed pattern for each scanner. Second, identify the scanner using the dust and scratch template of each scanner. Third, detection of forged regions in scanned images.

2.1 Creation of Scanner Template

The scanned images are first filtered using high pass filter to detect dust/scratch locations. To detect both scratch and dust locations with the same model, the absolute of the high pass filtered image is calculated. Then, the platen is scanned with a completely black background. Scanning black background helps detecting scratches, since the scratches reflect the incoming light and shine as small light spot. To separate dust/scratch spots from other details such as edges and high frequency components, a dust/scratch pattern is searched all over the high-pass filtered image by Normalized Cross Correlation (NCC) as in equation (1) and (2) [16]. Finally, the NCC output is applied to an empirically detected threshold to select local maxima regions. The regions with high cross-correlations are assumed as candidate of dust / scratch location, which explained in detail in our previous work [15].

To generate a scanner template, solely two different scans of completely black background is sufficient. Once, black scans are obtained, likely dust and scratch positions are detected. To compensate the shifts between the two scanned images, the dust and scratch positions of the two images are matched with each other through cross correlation. The scanner dust and scratch template is finally generated by taking Hadamard product of the scanned images that are correctly aligned.

2.2 Source Identification

Source identification is to match each scanned image to its source. To identify a given image to source, likely dust/scratch positions are detected with our previously proposed model [15]. The image dust/scratch positions are then correlated with the scanner dust/scratch template using NCC. If the dust/scratch pattern extracted from the given image is matched with the scanner template, it is very likely that the given image is created with the suspected scanner. The Dust/Scratch detection and identification process can be summarized as shown in Fig.1.

Basic definition of normalized cross correlation is given as follows:

$$\gamma = \frac{\sum_{x,y}[I(x,y)-\bar{I}_{u,v}][B(x-u,y-v)-\bar{B}]}{\sqrt{\sum_{x,y}(I(x,y)-\bar{I}_{u,v})^2 \sum_{x,y}(B(x-u,y-v)-\bar{B})^2}} \quad (1)$$

Where $I(x,y)$ the intensity value of the scanned image I of size $M_x \times M_y$ at point$(x,y), x \in \{0,\dots,M_{x-1}\}$, $y \in \{0,\dots,M_{y-1}\}$. The pattern is represented by a given template B of size $N_x \times N_y$, $\bar{B}$ is the mean of the template, and $\bar{I}_{u,v}$ is the mean of $I(x,y)$ within the area of the template B shifted to (u,v) which is calculated by:

$$\bar{I}_{u,v} = \frac{1}{N_x N_y}\sum_{x=u}^{u+N_x-1}\sum_{y=v}^{v+N_y-1} I(x,y) \quad (2)$$

2.3 Forged Regions Detection

Each suspicious image is sliced into smaller non-overlapping blocks. Each block separately is tested with the previously proposed model to detect its source scanner. The scanner template here is the whole flatbed scanner. Each block of the suspicious image can be considered as a template B, and the scanner template can be considered as the image I in Eq. (1).

If all the blocks in an image are matched with the scanner template, it is very likely that the given image is an authentic image coming from that source. Otherwise, different regions are coming from other sources and the image is a forged one. Hence, the forgery blocks are identified accurately and marked. The steps of the proposed forgery detection algorithm are indicated in flow chart shown in Fig. 2. Where the B(K) is the block of suspicious image B, and B1(k), I1 are the absolute of the high pass filtered block and scanner template respectively

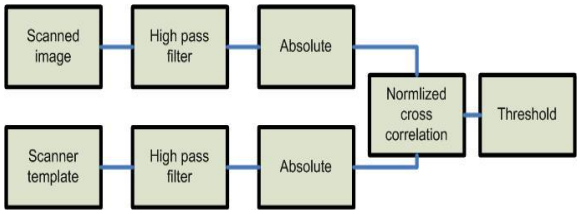


Fig 1: Dust/Scratch detection and scanner identification process.

3. EXPERMINTAL RESULTS

A lot of images are scanned using different scanner brands and have been modified to create the tampered images. The tampered images are classified into four groups, in the first group the images are synthesized using two different scanned images from different scanner brands, in the second group the images are synthesized using scanned images and images acquired using a digital camera, in the third group the images are synthesized using scanned images and images acquired using a cell phone camera, and the fourth group the images are synthesized using scanned images and computer generated images. The tampered images are synthesize by making cropping, scaling, rotation, hue and intensity histogram modification for a part of one image and insert it to the other one. The scanning resolution in these experiments has been fixed to 300 dpi and the scanner templates were also created at the same resolution. Table. 1 shows the scanner brands and the cameras that used in our experiments.

Each tampered image is sliced into a number of non-overlapping identical blocks, and each block is tested using the proposed detection algorithm. Finally, all the original (none tampered) blocks are arranged in the output image and tampered ones are masked with red color before arranging them in the output image. The block size has been scaled down to determine the best size that enable correct tampering detection. In the beginning, three different images from the first synthesized group are tested using the proposed detection algorithm. Figures. 3a, b and c show the scanned image using

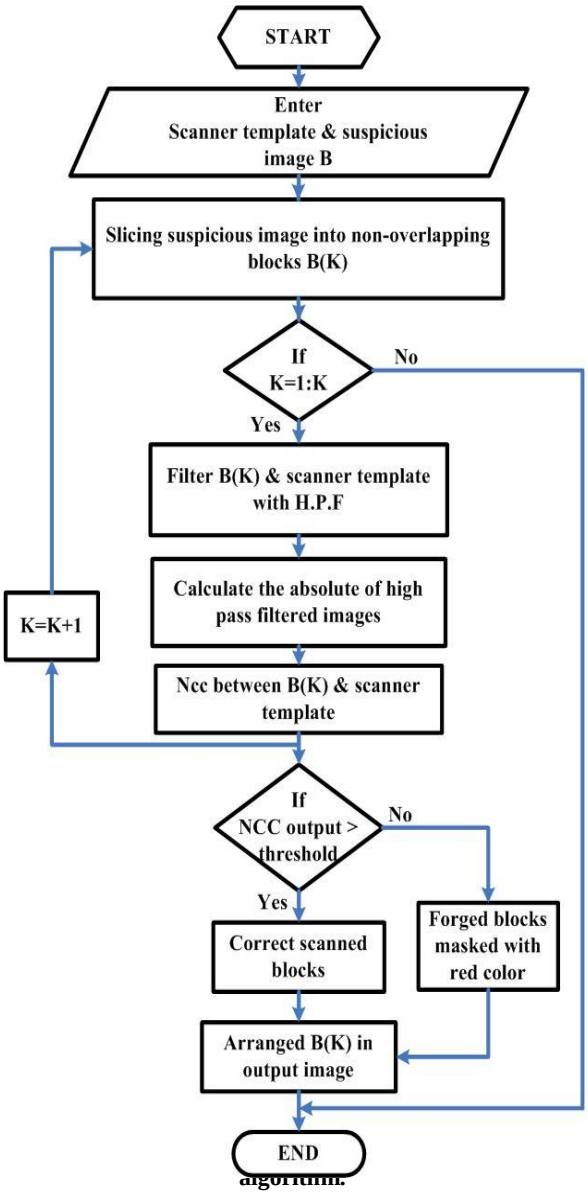


Table 1. Image sources used in experiments.

Image Class	Used Devices	Code
Digital Camera	Sony DSC-W610 14.1 Mega pixel	DC
Cell Phone Camera	Nokia E6 8 Mega pixel	CC
Flatbed Scanner	Hp Scanjet G3110	S1
	Benq 5000 color scanner	S2
	Canon canoscan LiDE25	S3
Computer Generated	www.maxon.net/gallery/vehicles.html	CG

S1 and S2, and the tampered image respectively. To detect the forged regions in the tampered image, the image is sliced into same size blocks. The size of each block defines the accuracy of the tampering detection, where with small block size the correlation between the block in the tampered region and the scanner template may be high. This leads to false tampering detection. So for each test, a large block size is used as a first attempt then smaller size is utilized so as to detect the accurate forged regions. A block of size 362 × 408 has been reached.

The output image of the proposed detection algorithm applied on the synthesized image with the template of S1 is shown in Fig. 3d. Figures 4, and 5 show the detection output of the proposed algorithm of the two tampered images, which are synthesized using the scanned images by (S2, S3), and (S3, S1), respectively.

In the second test, the forged regions in the tampered image from the second group are detected by the proposed algorithm. The scanned image using S1 and the acquired image using a digital camera (DC) are shown in Figs. 6a, and 6b. Figures 6c, and 6d show the tampered image and the detection algorithm output.

In the third test, a tampered image from the third group is tested. The scanned image using S1 and the acquired image using a cell phone camera CC are shown in Figs. 7a, and 7b. Figures 7c, and 7d show the tampered image and the detection result. Finally, a tampered image from the fourth group is tested. The scanned image using S1 and the computer generated CG image are shown in Figs. 8a, and 8b. Figures 8c, and 8d show the tampered image and the detection result. It is clear from results that the proposed detection algorithm can effectively use to identify accurately forged regions even large or small in any type of synthesized images.

The first proposed method used for forged region detection is [11] and uses the statistical feature of sensor pattern noise to identify the forged regions on only one group of tampered images. This group is similar to our first group. Figure 9 shows one of the tested examples in [11]. There are some errors in missing some forged regions or detecting original regions as forged ones. The second method that proposed in [14] uses spectral noise extracted from the image to detect forged regions on one group of tampered images, and these tampered images are similar to our second group. Figure 10 shows one of the tested examples in [14]. Also, there are some errors in missing some forged regions or detecting original regions as forged ones. To compare the results of the proposed method with the published ones, the following equations are defined:

$$\text{FA(forged area \%)} = \frac{\text{the area of the froged region}}{\text{the total area of the suspious image}} \times 100 \quad (3)$$

$$\text{correct detection (\%)} = \frac{\text{the detected area of forged region}}{\text{the area of the forged region}} \times 100 \quad (4)$$

$$\text{false detection (\%)} = \frac{\text{the false detection area}}{\text{the area of the forged region}} \times 100 \quad (5)$$

(a)

(b)

(c)

(d)

Fig 3: The detection result of synthesized image (S1&S2) (a) scanned image1 withS1, (b) scanned image2 With S2, (c) synthesis image, (d) output image with detected forged blocks masked with red color.

The comparison between the three detection methods is summarized in Table 2. It is clear from the table that the highest detection accuracy for large and small forged regions is obtained in the proposed method; however the block size used in the proposed method lies between the other two published methods. The detection accuracy for [11] is higher than that of the method used in [14] for large forged regions but its accuracy is smaller than that of [14] for small forged ones.

(a)

(b)

(c)

(d)

Fig 4: The detection result of synthesized image (S2&S3) (a) scanned image1 withS2, (b) scanned image2 With S3, (c) synthesis image, (d) output image with detected forged blocks masked with red color.

The FA of the proposed method lies between the two other methods, but it has been detected with the highest and small false detection.

The value of FA for one example that shown in [11] is the smallest one, but it has the worst correct detection and false detection value. The false detection occurred in the proposed method is due to the usage of rectangular shape.

(a)

(b)

(c)

(d)

Fig 6: The detection result of synthesized image (S1&DC) (a) scanned image1 withS1, (b) photographed image captured with digital camera DC, (c) synthesis image, (d) output image with detected forged blocks masked with red color.

(a)

(b)

(c)

(d)

Fig 7: The detection result of synthesized image (S1&CC) (a) scanned image1 withS1, (b) photographed image captured with cell phone camera CC, (c) synthesis image, (d) output image with detected forged blocks masked with red color.

Finally we can say that the accuracy of the forged regions detection methods depend greatly on the block size which affect the result of the comparison between the suspicious image and the source pattern (i.e., sensor noise, spectral noise, or dust/scratch and source imperfection patterns). The proposed method is simple and easy to apply to all types of the tampered images, regardless the acquisition source noise or the image contents. The two previously stated methods and our proposed method succeeded in forged image detection but all the methods have errors in determining the exact forged regions, depending on the shape and size of the forged regions.

(a)

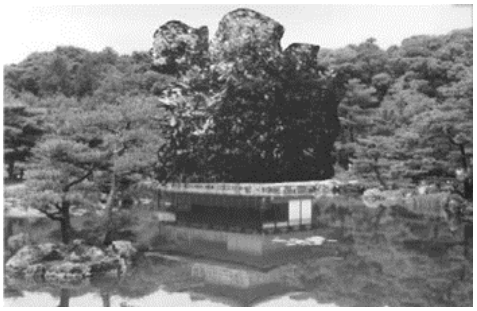
(b)

(c)

(d)

Fig 8: The detection result of synthesized image (S1&CG) (a) scanned image1 withS1, (b) Computer Generated image CG, (c) synthesis image, (d) output image with detected forged blocks masked with red color

(a)



(b)

(c)

Fig 9: Tested examples of forged region detection in [11]

4. CONCLUSION

In this paper, a novel method to authenticate and identify the forged regions in the scanned images is presented. The method is based on using dust/scratch and device imperfection pattern of the scanned images to identify the forged regions. Four groups of tampered images are tested. Each tampered image is sliced into a number of non-overlapping identical blocks, and each block is tested using the proposed detection algorithm. The block size has been scaled down to determine the best size that enables correct tampering detection. The experimental results have approved the validity, efficiency, and robustness of the proposed method to identify the forged image blocks in the suspicious images.

A comparison between the two previously published forged regions detection methods and the proposed method has been done. The accuracy of these methods depends greatly on the block size which affects the result of the comparison between the suspicious image and the source pattern. The proposed method is simple and easy to apply to all types of the tampered images, regardless the acquisition source noise or the image contents. The two previously stated methods and our proposed method succeeded in forged image detection but all the methods have errors in determining the exact forged regions, depending on the shape and size of the forged

regions. The errors include missing some of the forged regions or detecting some original regions as forged ones.

(a)

(b)

(c)

(d)

Fig. 10 Tested examples of forged region detection in [14]

5. REFERENCES

[1] H. Sencar, N. Memon, "Overview of state-of-the art in digital image forensics," WSPC - Proceedings Trim Size: 9.75in x 6.5in sencar-memon-chapter. 1, September 25, 2007.

[2] N. Khanna, A. Mikkilineni, A. Martone,G. Ali, G.Chiu, J. Allebach, E. Delp, "A survey of forensic characterization methods for physical devices," Digit Investig, vol. 3, pp.17–28, September2006.

[3] T. Gloe, E. Franz, and A. Winkler, "Forensics for flatbed scanners," in Proc. SPIE Security, Steganography, and Watermarking of Multimedia Contents IX, San Jose, CA, USA, vol. 6505, p. 65051I, Feb. 2007.

[4] A. Kot, H. Cao,"Image and Video Source Class Identification," in proc. Digital Image Forensics, Springer, pp 157-178, 2013.

[5] T. Gloe, E. Franz, and A. Winkler, "Forensics for flatbed scanners," in Proc. SPIE Security, Steganography, and Watermarking of Multimedia Contents IX, San Jose, CA, USA, vol. 6505, pp. 65051I, Feb. 2007.

[6] H. Gou, A. Swaminathan, and M. Wu, "Robust scanner identification based on noise features," in Proc. SPIE Security, Steganography, and Watermarking of Multimedia Contents IX, San Jose, CA, USA, vol. 6505, pp. 65050S, Jan. 2007.

[7] N. Khanna and E. Delp, "Source Scanner Identification for Scanned Documents," IEEE International Workshop on Information Forensics and Security - WIFS , 2009

[8] P. Chiang, N. Khanna, A. Mikkilineni, M. Segovia, J. Allebach, G. Chiu, E. Delp ,"Printer and Scanner Forensics: Models and Methods," Intelligent Multimedia Analysis for Security Applications, Studies in Computational Intelligence Volume 282, pp 145-187, 2010.

[9] N. Khanna, A. Mikkilineni, G. Chiu, J. Allebach, and E. Delp, "Scanner identification using sensor pattern noise," in Proc. SPIE Security, Steganography, and Watermarking of Multimedia Contents IX, San Jose, CA, USA, vol. 6505, p. 65051K, Feb. 2007.

[10] N. Khanna, A. Mikkilineni, and E. Delp,"Scanner Identification Using Feature-Based Processing and Analysis," Information Forensics and Security, IEEE Transactions on, Vol. 4, No. 1, pp.123-139, MARCH 2009.

[11] N. Khanna, T. George, J. Allebach; Delp, J. Edward ,"Scanner Identification with Extension to Forgery Detection," Security, Forensics, Steganography, and Watermarking of Multimedia Contents X, Volume 6819, pp. 68190G-68190G-10, 2008.

[12] N. Khanna, G. Chiu, J. Allebach, E. Delp," Forensic techniques for classifying scanner, computer generated and digital camera images," In: Proceedings of the IEEE international conference on acoustics, speech and signal processing, pp 1653–1656. Las Vegas, NV, 2008.

[13] C. Choi, M. Lee, H. Lee, "Scanner identification using spectral noise in the frequency domain," in Proc. IEEE International Conference on Image Processing (ICIP), Hong Kong, pp. 2121 – 2124, 26-29 Sept. 2010.

[14] C. Choi, M. Lee, D. Hyun, H. Lee, "Forged Region Detection for Scanned Images," Computer Science and Convergence Lecture Notes in Electrical Engineering Volume 114, pp 687-694, 2012.

[15] Z. Elsharkawy, S. Abdelwahab, M. Dessouky, S. Elaraby, F.Abd El-Samie, "Identifying Unique Flatbed Scanner Characteristics for Matching a Scanned Image to

its source," in Proc. CiiT International Journal of Digital Image Processing, Vol 5, No 9, September 2013.

[16] J. Yoo, T. Han," Fast Normalized Cross-Correlation", in Proc. Circuits, Systems and Signal Processing, Springer, vol. 28, Issue. 6, pp.819-843, December 2009.

Table 2. Comparison between the three detection methods of forged regions in scanned images.

	Sensor pattern noise [11]		Spectral noise [14]	The proposed method
Correct detection (eq 4) applied on output images	**First group (S&S)**	98.75% for large region and 39.69% for small region	——	100% for big regions and 100% for small region
	Second group (S&DC)	——	87.87% for large region and 66.8% for small region	100% for large regions and 100% for small region
	Third group (S&CC)	——	——	100% for large regions and 100% for small region
	Fourth group (S&CG)	——	——	100% for large regions and 100% for small region
Block size	384×512		256×256	362×408
The largest FA (Eq 3)	About 53% of		About 38.4%	About 38.3%
The smallest FA (Eq 3)	About 3.7%		About 12.85%	About 6.5%
False detection (Eq 5) applied on output mages	10% for large region and 17.95% for small region		6% for large region and 4.64% for small region	0.021% for large region and 7.5% for small region