Robust Detection of Copy-Move Forgery in Color Images

Nathalie Diane Wandji [1], **Sun Xingming** [2]
[1] School of Information Science and Engineering, Hunan University, Changsha, Hunan, P.R China
[2] Jiangsu Engineering Center of Network Monitoring, Nanjing University of Information Science and Technology, Nanjing, Jiangsu, P.R China

Abstract - *With rapid advances in digital image processing technology, there is a massive development of sophisticated tools and techniques for digital image forgery. Copy-move forgery is one of the techniques most commonly used. A part of an image is copied and pasted into the same image, in order to maliciously hide or clone an object or a region. In this paper, we propose a method that addresses this kind of fakery. Each component of the suspicious RGB image is used for feature extraction. Let Y be the Y-component of the corresponding YUV image. Then R, G, B and Y are splitted into fixed-size overlapping blocks, and characteristics are computed from each component. The 10-dimensional feature vectors made by concatenating the characteristic values of each component are then stored row wise in a matrix that will be lexicographically sorted to make similar image blocks neighbors. Duplicated image blocks are identified using Euclidean distance as similarity criterion. Experimental results showed that the proposed method can detect the duplicated regions even in case of slight JPEG compression, blur and noise addition.*

Keywords: Digital image forgery, Image forensics, Copy-move forgery.

1 Introduction

Powerful and easy to use digital cameras and image-editing software packages are becoming rampant. As a result, it has become relatively easy to manipulate digital images without leaving any visual clue, even for a novice. Recently, a lot of effort has been made to be able to decide on the authenticity and integrity of digital images, resulting into many approaches.

Generally, these approaches are divided into two main categories: the active methods on one hand and the passive/blind methods on the other hand. The use of active techniques such as digital signature and watermarking is limited, mainly by the requirement that certain information is embedded into the image during its creation since only few capturing devices on the market possess that feature. Unlike the watermark-based and signature-based methods; the passive approaches are concerned with determining the source and potential authenticity of an image without using any prior information. In this latter category, we will put our focus on one of the most commonly used tampering methods, namely the copy-move forgery. It proceeds by copying a part of an image and pasting it into another part of the same image, most of the time with the intention of deceiving people, or hiding/cloning an object. An example for this type of forgery can be seen in Fig.1, where the yellow hat is cloned.

(a)

(b)

Figure 1. Example of Copy-Move forgery (a) original image (b) tampered image

Several methods have been developed to provide solutions for detecting such forgery.

Jessica Fridrich et al. proposed a method using discrete cosine transform (DCT) of overlapping blocks and the lexicographical representation of the quantized DCT coefficients. Alin C. Popescu and Hany Farid [2] proposed a method based on the use of principal components analysis (PCA) to represent image blocks. A. N. Myna et al. [3] proposed a method using the idea of log-polar mapping and wavelet transforms. Hailing Huang et al. [4] used the SIFT algorithm to detect duplicated regions in the image. SIFT features were proven to be stable with respect to changes in illumination, rotation and scaling. Li Kang et al. [5] suggested to apply improved singular value decomposition to each image block to yield a reduced dimension representation and then lexicographically sort the feature matrix formed by the singular values. Their method was proven to be robust against noise distortion. Weiqi Luo et al. [6] presented a technique

robust to various forms of post region duplication processing, including blurring, noise contamination and lossy compression. They represented each block by 7 intensity-based characteristics extracted from both the RGB color image and the YCbCr corresponding image. Sevinc Bayram et al. [7] proposed to first apply Fourier Mellin Transform (FMT) on the image blocks and then use the projection of the obtained log-polar values onto 1-D as feature vectors. W. Li et al. [8] analyzed the block artifact grids (BAG), caused by the blocking processing during JPEG compression assuming they usually mismatch when tampering with objects by copy–paste operations. G. Li et al. [9] proposed to decompose the image into four frequency sub-bands using DWT and represent each block by the singular vector obtained by applying Singular Value Decomposition (SVD) only on the low-frequency component to yield a dimension reduction. H.T Sencar et al. [10] presented a method based on the assumption average sharpness/blurriness value of the tampered area is expected to be different as compared to the non-tampered parts of the image. They estimated the sharpness/blurriness value of an image based on the regularity of its wavelet transform. J. Zhang et al. [11] started by applying DWT to the input image and then, computed the phase correlation to estimate the spatial offset between the copied region and the pasted region. Finally, they used the idea of pixel matching to locate the forged region. S. Khan et al. [12] proposed the use of DWT followed by the comparison of blocks extracted from the low frequency subband using Phase Correlation as similarity criterion. Their technique fails to detect duplicated regions with rotation or scaling.

In this paper, we propose an efficient and robust algorithm for detecting and locating Copy-Move forged regions within a color image.

2 Proposed Method

The detailed procedure proceeds as follows:

1. Read the suspicious RGB color image $\boldsymbol{f}$ of size $\boldsymbol{M \times N \times 3}$ and convert it into YUV color space.
2. For each R, G, B and Y color component :
   2.1. Divide it into overlapping *b*x*b* blocks
   2.2. For each block of the Y component, compute its arithmetic average value *Ave_Y*.
   2.3. For each block of the R, G and B components, calculate 3 characteristics, namely the Average gray level (*Ave_g*); the average contrast (*Ave_c*) and the third moment (*Skew*) [13]. Let

$$\mu_n = \sum_{i=0}^{L-1} (z_i - m)^n \, p(z_i) \qquad (1)$$

where z_i is a random variable representing intensity, $p(z_i)$ is the intensity-level histogram and L the number of possible levels. Then, the above characteristics are obtained by the following equations :

$$Ave\_g = \sum_{i=0}^{L-1} z_i \, p(z_i) \qquad (2)$$

$$Ave\_c = \sqrt{\mu_2(z)} \qquad (3)$$

$$Skew = \mu_3 = \sum_{i=0}^{L-1} (z_i - m)^3 \, p(z_i) \qquad (4)$$

3. For each block of the color image, we therefore obtain a feature vector *V* of length *10*.

$$V = [Ave_Y, {Ave_g}^R, {Ave_c}^R, Skew^R, {Ave_g}^G, {Ave_c}^G, Skew^G, {Ave_g}^B, {Ave_c}^B, Skew^B] \quad (5)$$

4. Form a matrix **A** of dimensions ***(M-b+1)(N-b+1)*** rows and ***10*** columns and store the obtained features into rows of **A**.
5. Sort the matrix A lexicographically.
6. Set a threshold $\boldsymbol{T_n}$ controlling the amount of neighboring feature vectors to use for similarity check. For a given feature vector $A_p = (A_1^p, A_2^p, \dots, A_{10}^p)$, its neighbors are defined by $A_q = (A_1^q, A_2^q, \dots, A_{10}^q)$ where $q \in [k, l]$

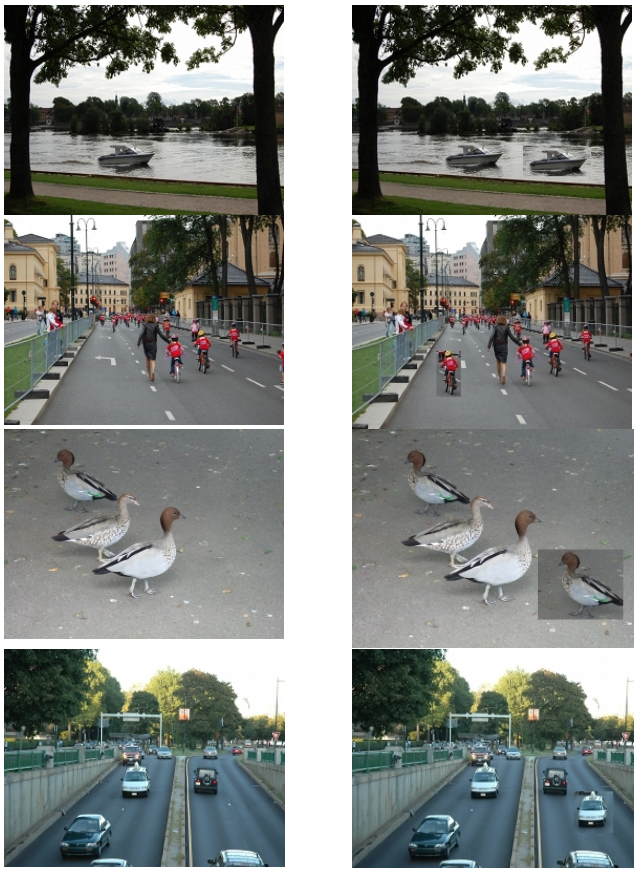
$$k = \begin{cases} 1, & p < T_n \\ p - T_n, & otherwise \end{cases} \qquad l = \begin{cases} b^2, & p > b^2 - T_n \\ p + T_n, & otherwise \end{cases}$$

7. To evaluate similarity between image blocks, the Euclidian distance is exploited. The more similar the examined blocks are, the smaller the Euclidian distance $\boldsymbol{d}$ between their corresponding feature vectors is. If $\boldsymbol{d}$ is smaller than a pre-defined threshold $\boldsymbol{T_{max}}$, the two blocks are considered as candidates for forgery and their respective positions $(\boldsymbol{x_i}, \boldsymbol{y_i}), (\boldsymbol{x_j}, \boldsymbol{y_j})$ together with and the shift vector between them $[|\boldsymbol{x_i} - \boldsymbol{x_j}|, |\boldsymbol{y_i} - \boldsymbol{y_j}|]$ are stored.
8. Set a third threshold $\boldsymbol{T_s}$. If the accumulative number of the corresponding shift vector is greater than $\boldsymbol{T_s}$, the corresponding blocks are marked as suspicious.
9. Considering the fact that blocks close to each other in the input image might with a high probability have similar feature vectors, we conclude the regions are duplicated only if the actual euclidian distance between both regions is greater than a predefined threshold $\boldsymbol{T_d}$.
10. Plot the blocks as copied and pasted regions on the gray-scaled image corresponding to the input image.

3 Experimental Results

Experiments are carried out on a computer with a configuration of CPU 2.7 GHz, RAM 2 G, Windows 7 32-bit Operating System, GIMP 2.6.12, and Matlab 7.12.0.635 (R2011a).

Original images are made available by authors in [14]. We scaled them to the size 267 × 200 and tampered with them as shown in Fig. 2. We tested different kinds of post processing such as noise addition, Gaussian blur and JPEG compression. The experimental results presented were all done with the same parameters, namely *b=16* (block size)*;* T_{max}=0.002; T_d=*2*b=32*; and T_n= T_s=b=16.



(a) (b)

Figure 2: (a) Original image (b) Tampered image

3.1 JPEG Compression

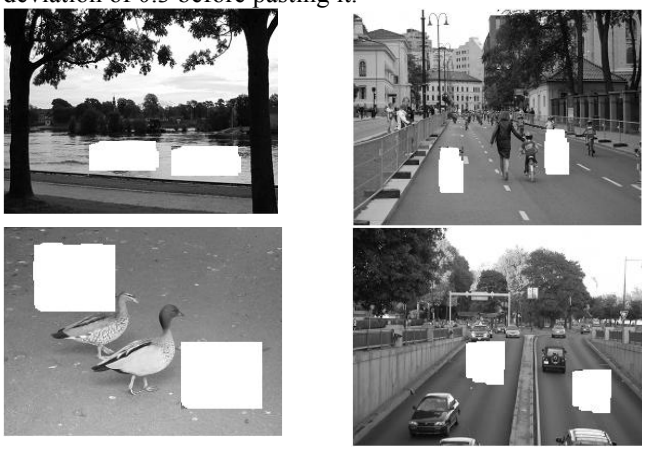
The whole forged image is compressed with a quality factor of 90.

3.2 Gaussian Blur

Gaussian blur with radius 3 is applied to to the whole forged image.

3.3 Additive Gaussian noise

Gaussian noise is added to the copied region with a standard deviation of 0.3 before pasting it.



4 Conclusions

With the rapid development of the image processing technology, there is a great need of a method that is able to detect digital image forgeries in general and copy-move forgeries in particular which is the most common forgery. In this paper, we consider the detection of such forgery in color images even when some post-processing operations such has Gaussian blur, Gaussian noise addition or JPEG compressed have been applied. Experimental results show that the proposed method was appropriate to some extent to identify and localize the forged region.

5 Acknowlegement

This work is supported by the NSFC (61232016, 61173141, 61173142, 61173136, 61103215, 61070196, 61070195, and 61073191), National Basic Research Program 973 (2011CB311808), 2011GK2009, GYHY201206033, 201301030, 2013DFG12860 and PAPD fund

6 References

[1] Jessica Fridrich, David Soukal and Jan Lukas. “Detection of copy–move forgery in digital images”, in: Proceedings of Digital Forensic Research Work- shop, IEEE Computer Society, Cleveland, OH, USA, pp. 55–61, August 2003.

[2] Alin C. Popescu and Hany Farid. “Exposing digital forgeries by detecting duplicated image regions”, Technical Report TR2004-515, Depart- ment of Computer Science, Dartmouth College, 2004.

[3] 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 Proceedings of the International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007), IEEE Computer Society, Washington, DC, USA, pp. 371–377, 2007.

[4] Hailing Huang, Weiqiang Guo, and Yu Zhang. "Detection of copy–move forgery in digital images using sift algorithm", in Proceedings of the 2008 IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application (PACIIA '08), IEEE Computer Society, Washington, DC, USA, 2008, pp. 272–276.

[5] Weiqi Luo, Jiwu Huang and Guoping Qiu. "Robust Detection of Region-Duplication Forgery in Digital Images", In proceedings of the International Conference on Pattern Recognition, Washington, DC, pp. 746-749, 2006.

[6] Li Kang and Xiao-pin Cheng. "Copy-move forgery detection in digital image", 3rd International Congress on Image and Signal Processing (CISP), vol. 5, pp. 2419 – 2421, 2010.

[7] Sevinc Bayram, Taha Sencar and Nasir Memon. "An efficient and robust method for detecting copy-move forgery," In Procceedings of ICASSP, IEEE International Conference on Acoustics, Speech, and Signal Processing, 2009.

[8] Weihai Li, Yuan Yuan and Nenghai Yu. "Passive Detection of Doctored JPEG Image via Block Artifact Grid Extraction". Signal Processing, pp. 1821-1829, 2009.

[9] Guohui Li, Qiong Wu, Dan Tu, and ShaoJie Sun. "A Sorted Neighborhood Approach for Detecting Duplicated Regions in Image Forgeries based on DWT and SVD". in Proceedings of IEEE International Conference on Multimedia and Expo, Beijing China, pp. 1750-1753, July 2-5, 2007.

[10] Yagiz Sutcu, Baris Coskun, Husrev T. Sencar and Nasir Memon. "Tamper detection based on regularity of wavelet transform coefficients," In Proceedings of ICIP, International Conference on Image Processing, 2007.

[11] Qiumin Wu, Shuozhong Wang, and Xinpeng Zhang. "Detection of image region-duplication with rotation and scaling tolerance". In Proceedings of the Second International Conference on Computational Collective Intelligence (ICCCI) Part I, pp. 100 -108, November 2010.

[12] Jing Zhang, Zhanlei Feng and Yuting Su. "A New Approach for Detecting Copy-Move Forgery in Digital", in Procceedings of the 11th IEEE Singapore International Conference on Communication Systems, pp. 362–366, 2008.

[13] Rafael C. Gonzalez, Richard E. Woods and Steven L. Eddins. "Digital Image Processing using MATLAB", Second Edition, Pearson Publications, 2004.

[14] Amerini, L. Ballan, R. Caldelli, A. Del Bimbo, and G. Serra. "A sift-based forensic method for copy-move attack detection and transformation recovery", IEEE Transactions on Information Forensics and Security, vol. 6(3), pp. 1099 - 1110, 2011.