International Journal of Applied Mathematics and Computing Science. Volume 3, Issue 4, (2014) 61-69
http://www.cayley-nielson.com/ijamcs

Blind Digital Image Forensics based on Correlation Detection

Douglas Armstrong
Department of Mathematics and Computer Science Albion College, 611 E. Porter St., Albion, Michigan 49224, USA
E-mail: *DRA11@albion.edu*

Abstract

For the most common fuzzy operation in digital image tampering, a digital image forensics algorithm based on the correlation detection is proposed. The algorithm assumes that linear correlation of pixel and neighboring pixels of the original image after fuzzy operation and the image has been tampered is enhanced, which uses the shortest distance clustering method to detect linear correlation of pixels, and reveals the trace of tampering, thus achieves forensics of digital image tampering. Through a lot of experiments, the results show that, the algorithm can effectively detect fuzzy operation trace of the tampered image, at the same time, which can be able to accurately locate the tampered.

Keywords: MIMO relay broadcasting systems; block diagonalization; low complexity; power loading.

1. Introduction

Currently, the function of image editing software is diversified, especially digital image won't leave any trace after being tampered, and this brought challenges for public security detection and judicial forensics, thus promoting the development of the digital image identification technology research. The document [1] has Summarize the digital image identification technology into three types: digital signature, blind digital watermarking and digital image forensics. In the above three kinds of technology, digital image blind forensics does not rely on any signatures and typing information to check the authenticity and source of the image; therefore, it is becoming a hot topic in the field of multimedia security, attracts more and more attention in all fields, and has wide applications[2-5].

Digital images are usually tampered by fuzzy and gradient operation, such as polishing, to eliminate the forgery traces left by tampering which can be detected by the human eyes. The detection of retouching in digital image tempering is one of the highlights of blind digital image forensics research. According to the different features of fuzzy images, a lot of image blind forensics based on fuzzy detection method has been put forward. In 2005, people like D. Hsiao [6] have proposed the fuzzy estimation method based on image frequency domain, and regarded the estimated fuzzy area as the most likely to be tampered. In 2006, Wang Bo and others [7] proposed a forensics based on the knowledge that fuzzy operation would affect the rate of abnormal hue. In 2007, Y. Sutcu and others [8] proposed to estimate the sharp and fuzzy part of image edge based on wavelet coefficients, so as to detect and locate image tampering area. In 2008, Zhou Linna and others [9] proposed blind forensics algorithm based on image morphology filtering edge feature.

Aiming at the most commonly used fuzzy operation in digital image tampering, the algorithm assumes that linear correlation of pixel and neighboring pixels of the original image after fuzzy operation and the image has been tampered is enhanced, which uses the shortest distance clustering method to detect linear correlation of pixels, and reveals the trace of tampering, thus achieves forensics of digital image tampering.

Through a lot of experiments, the results show that, the algorithm can effectively detect fuzzy operation trace of the tampered image, at the same time, which can be able to accurately locate the tampered.

2. Blind Digital Image Forensics Based on Correlation Detection Algorithm

2.1 Fuzzy Operation on Images

The mathematical expression of fuzzy operation on images:

$$g(x,y)=\sum_{s=-m}^{m}\sum_{t=-q}^{q}h(s,t)f(x+s,y+t) \tag{1}$$

In which, the original image is marked as *f*, and the image which has been fuzzily operated is marked as *g*, the Spread Function as *h*, (2*m*+1)×(2*q*+1) as the size of fuzzy template.

Supposing image which has been fuzzily operated is marked as *g*, in which, *g*(*x, y*) means pixel values when it is located at the position of (*x, y*); *f*(*x, y*) means pixel values when the original image *f* is located at the position of (*x, y*). Thus, *g*(*x, y*) can be expressed through the linear combination of *f*(*x, y*) and its adjacent pixel values with the same coefficient.

This paper assumes each *g*(*x, y*) can be expressed through the linear combination of *g*(*x, y*) and its adjacent pixel values with the same coefficient. It is:

$$g(x,y)=\sum_{\substack{s=-m\\ s\neq 0}}^{m}\sum_{\substack{t=-q\\ t\neq 0}}^{q}a(s,t)g(x+s,y+t)+\delta(x,y) \tag{2}$$

In which, *a* is a linear combination coefficient (a constant value), δ is the error, the neighboring size is (2*m'*+1)×(2*q'*+1).

2.2 Linear Correlativity

To analyze the formula (2), this paper introduces the following definition. That is: the image *g* is known,

$$g(x,y)=\sum_{\substack{s=-m'\\ s\neq 0}}^{m}\sum_{\substack{t=-q'\\ t\neq 0}}^{q}a(s,t)g(x+s,y+t)+\delta(x,y) \tag{3}$$

and meets the following requirements:

① The model parameter *a*(*s, t*) remain the same;

② The model error $\delta(x,y)$ has a small variance;

Then, it can be said there is a linear correlation between the image pixel of *g* and its neighborhood pixel.

Note:

① The linear correlation is relative.

② The closer the variance of *δ* (x, y) approaches to zero, the stronger the linear correlation is; conversely, the weaker it is. When the variance of (x, y) reaches zero, the linear correlation will be the strongest.

2.3 Enhancement of Linear Correlation

(1) Experiment 1 (ideal)

According to the above definition, it can be inferred: as the original image is f, the fuzzily operated one is g, and then the linear correlation of pixel of g is stronger than f. That is to say the fuzzy operation enhances the linear correlation between the pixel and its neighborhood pixel. To prove it, the Experiment 1 and 2 are introduced.

Randomly generated the two-dimensioned matrices $I_0$256×256, then it is fuzzily operated for 8 times, $I_n(n=1,\ldots,8)$ can be achieved. The fuzzy parameters are mentioned in Table 1. The least square method is used to respectively calculate the model error $\delta_n(n=0,\ldots,8)$ of 9 images, and then count its variance. Making 100 random experiments, its result can be shown in Figure 1.

Table 1. Fuzzy Parameter

Fuzzy Images	Fuzzy Types		Fuzzy Parameters
I_1	Defocusing Blurring	Gauss	Model Size：3×3 Variance σ ： 1.0
I_2			Model Size ： 3×3 ； Variance σ ： 2.0
I_3			Model Size ： 9×9 ； Variance σ ： 1.0
I_4			Model Size：9×9 ； Variance σ ： 2.0
I_5		Disk	Radius：3
I_6			Radius：7
I_7	Motion Blurring		Movement：5， Direction：45°

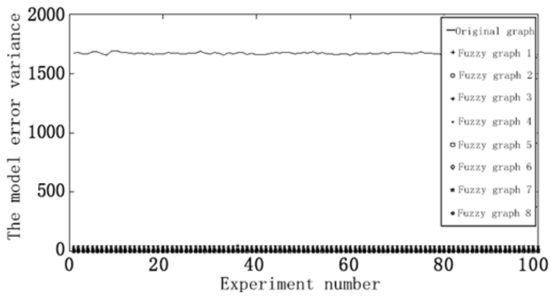


Figure 1. The Fuzzy Operation Influence on the Variance of δ

(2) Experiment 2(Reality)

Randomly select 79 commonly used images reduce them to 256 x 256, carry the operation taken in experiment 1 and experiment result is shown in Figure 2.

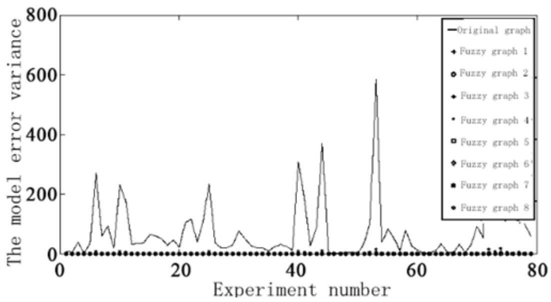


Figure 2. The Fuzzy Operation Influence on the Variance of δ

Figure 1 and Figure 2 show that the variance of δ is reduced after fuzzy operation.

To sum up, experiment 1 and experiment 2 indicate that the fuzzy operation enhances the linear correlation between the pixel and its neighborhood pixel. Therefore, the forensics mentioned in this paper is proved.

2.4 Tampering Fuzzy Detection Based on Pixel Linear Correlation

2.4.1 Algorithm Establishment

To analyze the pixel Linear Correlation of the image being detected by using the principle that fuzzy operation can enhance the Linear Correlation, if in a connected area, the Linear Correlation of the pixel is stronger than that in any other areas, then it can be said the area has been fuzzily operated and finally concluded that the image has been tampered.

(1) Rewrite the Model in Formula 3

As the image g is known, put the influence of error δ(x,y) into the model parameter $a(s, t)$, then the formula (3) can be rewritten as the following:

$$g(x,y)=\sum_{\substack{s=-m'\\ s\neq 0}}^{m}\sum_{\substack{t=-q'\\ t\neq 0}}^{q}a'(s,t;x,y)g(x+s,y+t) \tag{4}$$

In which, the model coefficient $a'(s, t; x, y)$ is changeable in areas.

When $a'(s, t; x, y)$ tends to stay the same, Linear Correlation will be stronger; conversely, it will be weaker. That is to say, when the neighboring pixel Linear Combination is used to represent each pixel, when the linear combination coefficient tends to keep the same, the Linear Correlation will be the strongest.

Therefore, it can be concluded that when the Linear Combination coefficient $a'(s, t; x, y)$ becomes more similar and concentrated in an area, then it is fuzzily operated and other areas are not.

(2) Algorithm Idea

To realize it, this algorithm will calculate the linear combination coefficients in divided blocks. a'_I means the linear combination coefficient of the *ith* block; the set $\{a'_i\}$, which means all linear combination coefficients of all blocks, can be divided into two categories by using the shortest distance clustering method based on threshold value.

If $\exists a' \in \omega_1, d(a'_i, a'_j) < t$;

Then $a'_j \in \omega_1$;

Otherwise $a'_j \in \omega_2$,

In which, $d(a'_i, a'_j)$ is distance operation, t is the given threshold, ω_1 is the fuzzily operated pixel class, while ω_2 is the pixel class that hasn't been fuzzily operated.

After the first classing, ω_1 and ω_2 can be achieved. Since the fuzzy operation is carried in the connected area, there must be errors if it is divided according to the distance, morphologic method shall be adopted further so that the fuzzily operated area can be connected with fewest errors.

2.4.2 Algorithm Implementation

This paper proposes an algorithm which can effectively detect and locate the tampered area in digital images. That is to obtain the neighboring matrix (8 neighboring areas) of the images being tested, then divide them into blocks and calculate the linear combination coefficient of each block by using the least square method, later use shortest distance clustering method based on the threshold value, finally operate them morphologically so as to detect the tampered area.

Algorithm steps are described in detail in the following:

(1) Dividing in Blocks: to evaluate the 8 neighboring areas of image I (regardless boundaries) in blocks (block size: 3×3);

(2) Evaluating the Linear Combination Coefficient: to get the linear combination coefficient a'_I by using the least square method and get a vector with starting address, and store in matrix A;

(3) Clustering: to cluster A by using the shortest distance clustering method based on the threshold value. The distance mentioned in this experiment is Euclidian distance. The distance $a'I$ that is less than the given threshold value t will be clustered in ω_1 class, T_i=1; otherwise, the distance will be clustered in ω_2 class, T_i=0. All results are expressed in the matrix T;

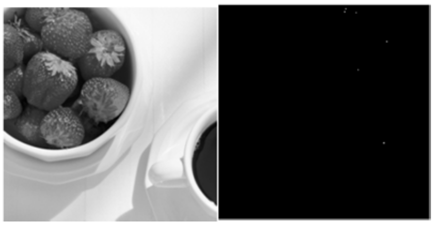
(4) Locating and Morphologically Operating: to locate the clustered results with the help of T and A, and then get the result figure I_d. Making morphological operation like corrosion, expansion and inversion, and then multiply the original image, I_r will be got.

3. Experimental Results and Analysis

3.1 Experimental Results

(1) Experiment 3 (Simulation Data)

In order to verify the proposed algorithm, *Photoshop* is used to tamper the selected image (Figure 3): copy a part of the image to the original image and fuzzily operate it, then tampered image will be got, which is shown in Figure 3c. The proposed algorithm is used to calculate the pixel correlation of the original image and tampered one. Experiment Parameters: block size: 3×3, threshold value t =0.054， structure element is disc, corrosion radius =1，expansion radius =15。The results are shown in the Figure 3.



(a)Original Image (b) Test Result of the Original Image

(c)Tampered Image (d) Test Result of the Tampered Image

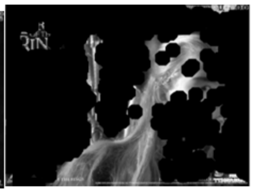
(e) Morphologic Operation (f) Relocation of the Tampered Area

Figure 3. Tampered Test Experimental Result (Simulation Data)

(2) Experiment 4 (Practical Data)

Download images from website (key words: Little Fatty, the Lord of the Rings) , then the Figure 4a and Figure 4c can be got, Figure 4b and Figure 4d are the results calculated according to the proposed algorithm. The experimental results are the same with that of the document [6].

(a)Original Poster (b)Test Result (t=0.054)



(c)Tampered Poster (b)Test Result (t=0.054)

Figure 4. Differential Chart of Adjacent Threshold Value Test Result

3.2 Experimental Analysis

(1) Threshold Value

To analyze whether the selected threshold values is right or not, the following experiment is carried out. *Photoshop* is used to tamper the Figure 5a. that is to fuzzily operate the face and then Figure 5d can be got. Then test the pixel correlation of the original and tampered images by dividing them into 3×3 blocks according to the proposed algorithm. The experimental threshold value is determined according to the shortest distance *dm* of linear combination coefficient.

$\forall a'_i, a'_j \in \{a'\}$

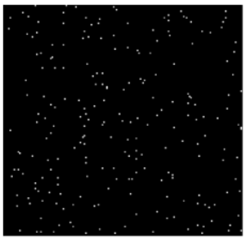
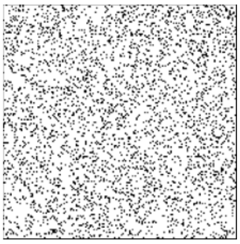
If $\min(d(a'_i, a'_j)) = 0$, then $dm = \varepsilon$;

Otherwise, $dm = \min(d(a'_i, a'_j))$, ε

In which, ε is the minimum but not zero, and in this experiment $\varepsilon = 1\times10^{-9}$.

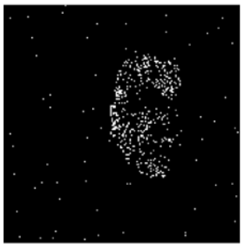
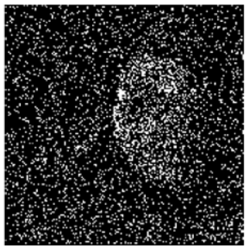
Experimental threshold value $t=2^k \times dm$ (k=1,…,40), some test results are shown in Figure 5 and Figure 6. (dm_1 is the *dm* of original image, while dm_2 means the *dm* of the tampered one)

(a)Original Image

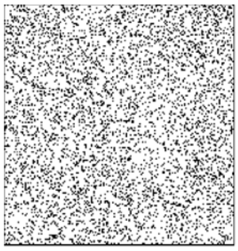
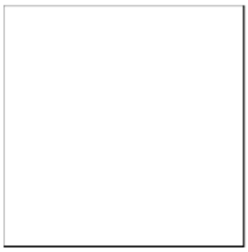



(b) $t_{11}=2\times dm_1$ (c) $t_{14}=2^4\times dm_1$


(d) Tampered Image (e) $t_{21}=2\times dm_2$

(f) $t_{227}=2^{27}\times dm_2$ (g) $t_{228}=2^{28}\times dm_2$

(h) $t_{230}=2^{30}\times dm_2$ (i) $t_{240}=2^{40}\times dm_2$

Figure 5. Some Test Results of Different Threshold Value

Figure 5 shows that when the threshold value is changing from small to large, the fuzzily operated part if it has been will be regional in the result. Otherwise, the experimental results will be approximately distributed on the whole image. Therefore, the threshold value can be increased times in the experiment. If there is obvious regional part, it can be initially concluded that the region has been fuzzily operated.

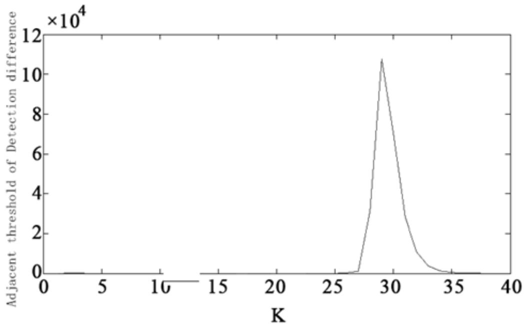


Figure 6. Differential Chart of Adjacent Threshold Value Test Result

Figure 6 shows that if $t \in [2^1 \times dm, 2^{27} \times dm]$, the test results remain unchanged, which proves the robustness of the threshold value used in the algorithm proposed by this paper.

(2) Robustness

In order to prove the robustness of the proposed algorithm, the Figure 4c will undergo further processing: compression (Figure 7a), fuzzy operation (Figure 7c), the parameters are in Table 2, which are proved to be the same with that of experiment 4. The experimental results are in Figure 7.

Table 2. Further Processing Parameters

Type	*JPEG* Compression	Fuzzy Operation (Gaussian)
Parameter	Quality Factor: 85	Fuzzy Size: 3×3 Fuzzy Radius: 1.0

(a)JPEG Compression (b) Test Result

(a)JPEG Compression (b) Test Result

Figure 7. Test Results of Different Tampering for Further Processing

Figure 7 shows that the proposed algorithm in this paper has some robustness on compression and fuzzy operation.

4. Conclusion

The experiments show that fuzzy operation can enhance the linear correlation of pixel and neighboring pixels, based on which as well as fuzzy operation, a digital image forensics algorithm is proposed. This can be used to decide whether the image has been tampered or not according to its fuzzy operation possibility and then locate the tampered position. Lots of experiments prove that the algorithm proposed in this paper can be used to effectively detect whether the image has been tampered and then locate the exact tampered position.

References

[1] Ng, Tian-Tsong, et al. "Passive-blind image forensics." Multimedia Security Technologies for Digital Rights 15 (2006): 383-412.

[2] Hobbs, Brian, and Mohamed Tchoketch Kebir. "Non-destructive testing techniques for the forensic engineering investigation of reinforced concrete buildings." Forensic science international 167.2 (2007): 167-172.

[3] Prasad, S., and K. R. Ramakrishnan. "On resampling detection and its application to detect image tampering." Multimedia and Expo, 2006 IEEE International Conference on. IEEE, 2006.

[4] Nicholson, Jeremy K., John C. Lindon, and Elaine Holmes. "'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data." Xenobiotica 29.11 (1999): 1181-1189.

[5] Alles, Erwin J., Zeno JMH Geradts, and Cor J. Veenman. "Source Camera Identification for Heavily JPEG Compressed Low Resolution Still Images*." Journal of forensic sciences 54.3 (2009): 628-638.

[6] Hsiao, Dun-Yu, and Soo-Chang Pei. "Detecting digital tampering by blur estimation." Systematic Approaches to Digital Forensic Engineering, 2005. First International Workshop on. IEEE, 2005.

[7] Wang, Bo, et al. "Image forensics technology using abnormity of local hue for blur detection." Acta Electronica Sinica 34.12A (2006): 2451-2454.

[8] Sutcu, Yagiz, et al. "Tamper detection based on regularity of wavelet transform coefficients." Image Processing, 2007. ICIP 2007. IEEE International Conference on. Vol. 1. IEEE, 2007.

[9] Alvarez, Luis, Pierre-Louis Lions, and Jean-Michel Morel. "Image selective smoothing and edge detection by nonlinear diffusion. II." SIAM Journal on numerical analysis 29.3 (1992): 845-866.

[10] Gusmão, Leonor, et al. "DNA Commission of the International Society of Forensic Genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis." Forensic science international 157.2 (2006): 187-197.

[11] Albert, A. Midori, Karl Ricanek, and Eric Patterson. "A review of the literature on the aging adult skull and face: Implications for forensic science research and applications." Forensic Science International 172.1 (2007): 1-9.

[12] Bhanu, Bir, and Yingqiang Lin. "Genetic algorithm based feature selection for target detection in SAR images." Image and Vision Computing 21.7 (2003): 591-608.