International Journal of Digital Crime and Forensics
Volume 13 • Issue 2 • March-April 2021

Multiple Fusion Strategies in Localization of Local Deformation Tampering

Yongzhen Ke, School of Computer Science and Technolgy, Tiangong University, China
Yiping Cui, School of Computer Science and Technolgy, Tiangong University, China

ABSTRACT

Tampering with images may involve the field of crime and also bring problems such as incorrect values to the public. Image local deformation is one of the most common image tampering methods, where the original texture features and the correlation between the pixels of an image are changed. Multiple fusion strategies based on first-order difference images and their texture feature is proposed to locate the tamper in local deformation image. Firstly, texture features using overlapping blocks on one color channel are extracted and fed into fuzzy c-means clustering method to generate a tamper probability map (TPM), and then several TPMs with different block sizes are fused in the first fusion. Secondly, different TPMs with different color channels and different texture features are respectively fused in the second and third fusion. The experimental results show that the proposed method can accurately detect the location of the local deformation of an image.

KEYWORDS

Forensic Science, Image Forensics, Local Deformation, Multiple Fusion, Textural Features

INTRODUCTION

The rapid development of image processing software makes it easy for people to use the retouching software to tamper with the image according to their wishes. Tampering images may be used by criminals to publish false information, and widespread dissemination of false information may cause social panic. Tampering with images can also involve areas of crime such as public opinion and violations of portrait rights. The purpose of news work is to publicize the truth, not to mention the use of tampered images to attract people's attention. For entertainment or beauty, people will modify photos before posting them on social networks. Many magazines and businesses also use unrealistic images to attract consumers. For example, many photos of the stars we saw in magazines have been modified with retouching software. This modified beauty will attract the public's attention and cause the public's pursuit of this beauty, but this beauty may not be in line with the health of the human body. Excessive pursuit of this unhealthy beauty will affect people's physical and mental health. In addition, ads such as fitness and weight loss often use false information to attract consumers to buy their products, which is unfair to consumers. People have long-term seen this over-modified "beautiful" images, which makes person mistakenly believe that this modified figure is "normal". This misunderstanding may affect people's values, and at the same time, the pursuit of such excessive

DOI: 10.4018/IJDCF.2021030107

Figure 1. An example of local deformation

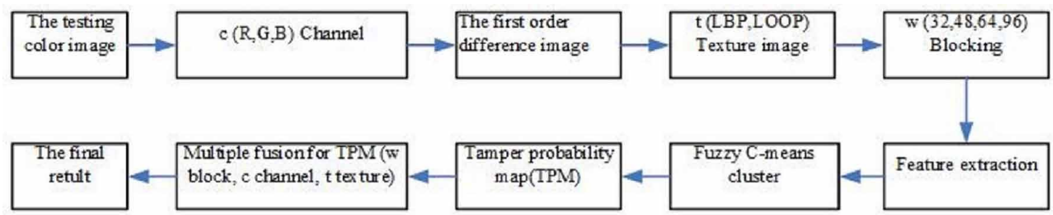
(a) original image (b) use Forward Warp Tool (c)tampered image

"beauty" will affect people's physical and mental health. Therefore, it is necessary to forensic the authenticity of the image.

Image local deformation is one of the common methods of image tampering, including three basic operations of local translation, local scaling, and local rotation (Andreas Gustafsson, 1993). Through the use of Liquefy Tools like Photoshop, people can freely distort any area of the image and change the shape of objects in the image. People also use this tool to make the people whom in the figure look thinner and so on, an example is shown in Figure 1. Figure 1 shows that Forward Warp Tool in the Liquefy Tool is used to achieve translational deformation, Figure 1(a) is the original image without deformation, Figure 1(b) shows the process of using the Forward Warp Tool, and Figure 1(c) shows the tampered image of deformation. It can be seen that the arms of the woman in the image are deformed. The liquefaction tool tampers the image through a circular selection, which makes the pixel values in the middle of the selected area a great change while the pixel values at the edge of the selected area almost unchanged. Multiple basic deformations will be applied in practical applications.

These characteristics of the deformation method bring great difficulties to forgery detection task. Firstly, this kind of deformation involves interpolation operation. Many studies on image warping focus on the geometric deformation of the rotation and scaling classes. These geometric transformations introduce periodic variations in pixels into the image when interpolating the image. Some papers (Popescu & Farid, 2005; Mahdian & Saic, 2008; Cao, 2012; Birajdar, 2014) use this periodic variation to detect the authenticity of images. Popescu and Farid (2005) analyzed the correlation between resampled signals and adjacent pixels, and proposed an expectation maximization (EM) method, which uses E-step and M-step iterative calculations to determine whether a pixel has been resampled. Mahdian and Saic (2008) studied the periodicity of the covariance structure of the interpolated signal and its derivatives. Cao (2012) used first-order derivative signals to detect images horizontally and vertically. Birajdar (2014) used the second difference zero-crossing auto-covariance sequence to detect the resample of an image. However, this periodic variation is not introduced in local deformation. Secondly, these methods also cannot detect the resampling area after multiple interpolation operations, and the local deformation in actual application is the result of multiple deformations, which brings difficulties to the local deformation method of detecting images. Finally, the local deformation is performed in a circular area. This tampering method does not uniformly affect the entire image. Bharati (2016) proposed a new algorithm using a supervised deep Boltzmann machine, which selected four partial learning features in the face image and used the proposed algorithm to determine whether the face images were retouched or original images. This method requires many pictures for training, and could unable to detect the location of the tampering. Farid and Kee (2011) proposed a metric (range 1-5) to quantify changes in markup digital images. They used four statistic variables to calculate the geometric changes of the subject's face and body. However, their method needs to be compared to the original image, and cannot locate where the deformation occurs.

Figure 2. Flowchart of the proposed method



Hwang MG (2016) tried to detect the local image deformation in 2016. Firstly, they used interpolation for images, then created a detection map in the frequency domain using the re-interpolated image, and finally intuitively identified the tampering region. However, they do not locate the local deformation of the image. It is only a rough view of the location of the tampering by the human eye, and easy to misjudge.

In order to detect the location of local deformation of a tampering image, multiple fusion strategies are proposed in this paper. The authors use the first-order difference image and its texture features to extract the features of the image tampering trajectory, and then send the features into the fuzzy c-means clustering method to obtain the TPM. Since each TPM with different block size, different color channel, different texture feature can roughly detect the tampering of the image and the detection result is not very ideal, a plurality of fusion strategy methods are proposed and the detection result can be improved by fusing different TPMs to detect the deformation position. Our proposed multiple fusion strategies can detect the location of local deformation and it is easy to judge the location of the tampering by the human eye.

MULTIPLE FUSION STRATEGIES FOR LOCAL DEFORMATION DETECTION

The local deformation of the image causes the pixel value of the tampering position in the image to change, which causes the correlation between the pixels to change. The author believes that the first-order difference image and the texture image can detect the change of the correlation between the pixels, and propose a multiple fusion strategies method based on first-order differential image and texture feature, which is shown in Figure 2.

Firstly, a first-order difference image for one color channel of the testing image is obtained, and one texture image is generated from the first-order difference image. Secondly, some features are extracted by using overlapping blocks and fed into a fuzzy c-means clustering method to generate a tamper probability map(TPM). Thirdly, several TPMs with different overlapping block sizes in one color channel is fused in the first fusion. The second fusion is to fuse the TPMs obtained by different color channels with the mean value method to get the binary image, and use morphological operation to remove the incoherent areas. Lastly, Different texture features will lead to different detection binaries, and the third fusion uses the voting method to fuse different detection binaries to get the final result. The multiple fusion strategies with different block sizes, different color channels, different texture features can greatly improve the detection effect.

Differential Image

The local deformation of the image will change the correlation between the pixels of the image. The difference image of the image can provide the relationship between the image pixel and its neighbors. The authors extract the difference image between the horizontal and vertical directions of the image to extract the subtle hidden information between the pixels. Firstly, for the input color image I with the size of M×N, horizontal and vertical difference images (Li Weihai, 2009) of different color channels are extracted, show as Eq.(1) and Eq.(2):

$$d_c^{\rightarrow}(x,y) = | S_c(x,y) - S_c(x,y+1) | \tag{1}$$

$$d_c^{\downarrow}(x,y) = | S_c(x,y) - S_c(x+1,y) | \tag{2}$$

This paper uses the symbol {→, ↓} denote the difference in horizontal and vertical directions. $S_c(x,y)$ is the value at pixel (x,y) in the tampering image of the c channel, where $c \in \{R,G,B\}$, $d_c^{\rightarrow}(x,y)$ is the value of the pixel in the corresponding horizontal absolute first-order difference $D_c^{\rightarrow}$ of the c channel, and $d_c^{\downarrow}(x,y)$ is the value of a pixel in the absolute first-order difference $D_c^{\downarrow}$ in the vertical direction.

Then, the Eq. (3) is used to combine the difference images in the horizontal and vertical directions to get the final difference image:

$$d_c(x,y) = d_c^{\rightarrow}(x,y) + d_c^{\downarrow}(x,y) \tag{3}$$

where $d_c(x,y)$ is value at pixel (x,y) in the first-order difference image D_c.

Texture Features

In this paper, texture features are extracted from the first order difference images D_c. Two texture features are used in this paper as described below.

LBP Texture Feature

Local Binary Pattern (LBP) (Ojala, 1994) is the earliest proposed texture feature. Let i_c be the intensity of an image D_c at the pixel (x_c, y_c) in c channel of the color image, where $c \in \{R,G,B\}$, and $i_n (n = 0,...,7)$ be the intensity of a pixel in the 3×3 neighborhood of (x_c, y_c) excluding the center pixel i_c. The LBP value for the pixel (x_c, y_c) is given by:

$$LBP_c(x_c, y_c) = \sum_{n=0}^{7} F_c(i_n - i_c).2^n \tag{4}$$

where:

$$F_c(x) = \begin{cases} 1 & if\ x \geq 0 \\ 0 & otherwise \end{cases} \tag{5}$$

The LBP image of c channel is obtained by finding the LBP_c value of the pixels of the entire image.

LOOP Texture Feature

The main disadvantage of LBP and LDP (Jabid, 2010) is the arbitrary sequence of binarization weights that add dependence to orientation. LDP suffers from the empirical assignment of value to the threshold variable, which puts an ad hoc restriction on the number of bits allowed to be 1, thus reducing the number of possible words. Chakraborti (2017) proposed a Local Optimal Oriented Pattern(LOOP). LOOP is a nonlinear fusion of LBP and LDP to preserve the advantages while removing the disadvantages of both.

Similar to the equation for calculating the LBP value, the equation for LOOP is:

$$LOOP_c(x_c, y_c) = \sum_{n=0}^{7} F_c(i_n - i_c).2^{w_n} \tag{6}$$

where:

$$F_c(x) = \begin{cases} 1 & if\ x \geq 0 \\ 0 & otherwise \end{cases} \tag{7}$$

The eight responses of the Kirsch masks are m_n corresponding to pixels with intensity $i_n, n = 0,...,7$. Each of these pixels is assigned an exponential w_n (a digit between 0 and 7) according to the rank of the magnitude of m_n among the 8 Kirsch mask outputs.

Image Tampering Probability Map

After extracting the texture image, three features of mean, standard deviation, variance (Eq. 8-10) and histogram features extracted by using sliders of different sizes are used in this paper:

$$f_1 = \frac{1}{w_l^2} \sum_{i=1}^{w_l} \sum_{j=1}^{w_l} T_c(i,j) \tag{8}$$

$$f_2 = \sqrt{\frac{1}{w_l^2} \sum_{i=1}^{w_l} \sum_{j=1}^{w_l} \left(T_c(i,j) - f_1\right)^2} \tag{9}$$

$$f_3 = (f_2)^2 \tag{10}$$

where w_l represents the sliding window, where $l = \{1,..,l\}$ represent the lth sliding window, $T_c(i,j)$ is the value at the point (i,j) in the texture image of c channel in color space, where $c \in \{R,G,B\}$. f_1 defines the average of all texture feature values in the sliding window. f_2 and f_3 define the standard deviation and variance in the sliding window, respectively.

Then, three features and histogram features are fed into the fuzzy c-means clustering method to obtained TPMs P_c^l, where P_c^l represents the TPM corresponding to the lth window of c color channel.

Multiple Fusion Strategies

Fusion of Different Size Blocks

For sliding window analysis, large windows are expected to provide reliable classification accuracy, yet with limited localization resolution, while small windows could potentially yield much better localization capability, but are expected to be impractically prone to errors (P. Korus & J. Huang, 2016). In order to make better use of the advantages of these two windows, a multi-map fusion method (P. Korus & J. Huang, 2017) is used in the first fusion.

The TPMs $P_c^l \in [0,1]$ is obtained by sliding windows w_l of different sizes. The process of fusion is to generate a probability map from the different P_c^l. Use the following Eq (11) to indicate:

$$t_c = \frac{1}{l}\sum_{i=1}^{l} P_c^{(i)} \tag{11}$$

A simple average fusion method is used to fuse the TPMs obtained by different sliding windows in c channel, whereTPMs $P_c^{(l)} \in [0,1]$ identify its unreliable regions, where close to 1 represents the highest probability of tampering, and close to 0 is the higher probability of tampering. And t_c is the probability map obtained by fusing different windows of c channel.

Color Channel Fusion

The existed studies(Johnson & Farid, 2006; Zhao, 2010; Muhammad, 2013) have found that color channels rather than gray brightness enhance the performance of the assay. The average fusion is also used, as shown in Eq.(12), where different color spaces are merged into a binary image in the second fusion:

$$C = \frac{1}{3}\sum_{i=1}^{c} t_c \tag{12}$$

where $c \in \{R, G, B\}$, represents the different color space channel. C is the result of color space fusion.

The choice of color space will be explained in the experimental section. After using the average fusion, a threshold value of 0.6 is set to transform the probabilistic map into a binary image, and incoherent areas are removed using first dilation and then erosion (Serra, 1983) in this paper.

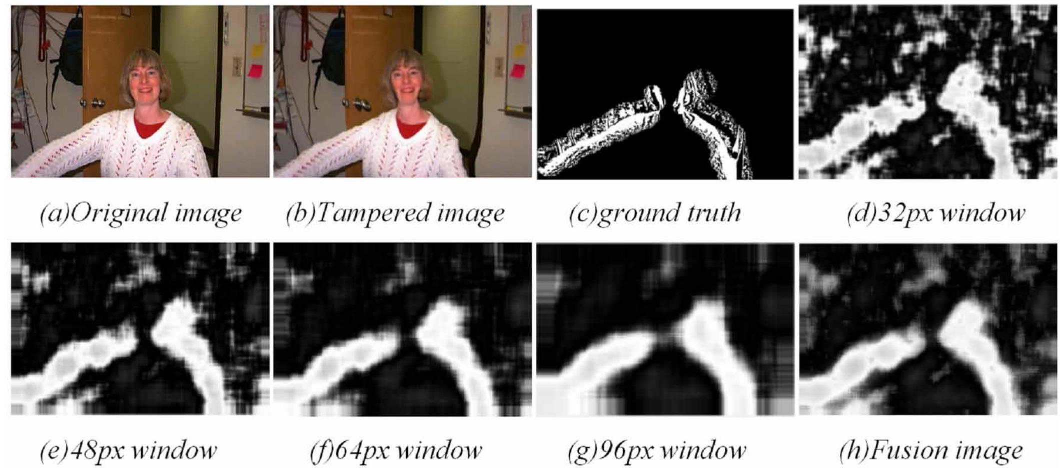
Fusion of Different Texture Features

Different texture features can bring different detection results, and using multiple texture features can improve the accuracy of local deformation detection. In the third fusion, the detection results of different texture features are fused by the voting mechanism, AND operator [18] is used for voting, shown as Eq.(13):

$$R = C_{LBP} \wedge C_{LOOP} \tag{13}$$

where C_{LBP} and C_{LOOP} are the second fusion results, the difference is that LBP and LOOP texture features are used in the first fusion respectively.

Figure 3. Tampering probability map of different size detection window



EXPERIMENTS

This paper randomly downloaded some images from the image library (Linda G. Shapiro) for experimentation. The authors use the Forward Warp Tool in the Liquefy Tool of Photoshop to perform local deformation operations on the downloaded image. The tampering position in the tampered image is arbitrary to prove that our method can be free from the influence of the object texture in the image. The program is built in matlab2016. Texture features are extracted using different sizes sliding windows with $w_l \in \{16, 32, 64, 96\}$, central-pixel attribution is used for different sliding windows, and repeated padding is used for unclassified pixels at image boundaries.

Experiment Results With Different Size Blocks

In this section, the authors show the tamper probability map(TPM) extracted from the tamper image of Figure 1 using different sized windows to verify the first fusion method. The experimental results are shown in Figure 3.

Figure 3(a)-(c) shows an original image, tampered image, and its ground truth. Figure 3(d)-(g) shows the TPMs obtained by using the window of 32×32px, 48×48px, 64×64px, and 96×96px respectively. Figure 3(h) shows the first fusion result with different size of windows. The pixel in the light part means that the probability of tampering is close to 1, while the dark part means that the probability of unaltered pixel is close to 1. It should be noted that the use of a larger window can roughly detect the location where tampering occurs, while the smaller detection window can more accurately locate the tampering location, but small windows contain more noisy and uncertain displays. It could be able to achieve better performance by combining the result of different scales.

Experimental Results with Different Color Space

Three common color Spaces are compared in this section, including RGB, HSV, and YCbCr. The authors show the results of using the first fusion method to detect the tampering position of the tampering image in Figure 1 in different color channels,and use the second fusion method to fuse the results of different color spaces. The experimental results are shown in Figure 4- Figure 6.

Figure 4(a)-(c) show the results of the first fusion of each channel in RGB space. Figure 5(a)-(c) shows the result of the first fusion of each channel in HSV space. Figure 6(a)-(c) shows the result of the first fusion of each channel in YCbCr space. Figure 4(d), Figure 5(d) and Figure 6(d) show the second fusion result of the different color channel. The light part in Figure 4-6 represent the location

Figure 4. Fusion result of one channel in RGB space and fusion result of RGB color space

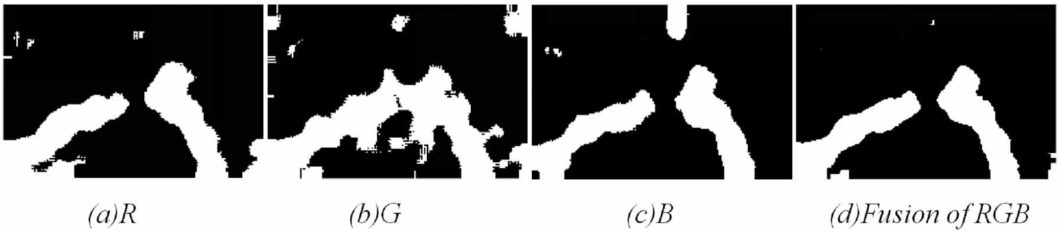


Figure 5. Fusion result of one channel in HSV space and fusion result of HSV color space

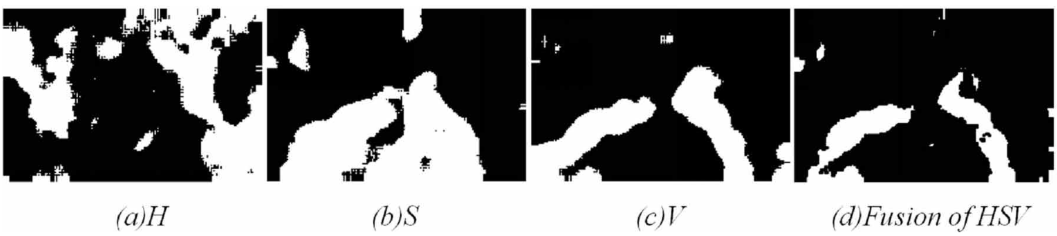
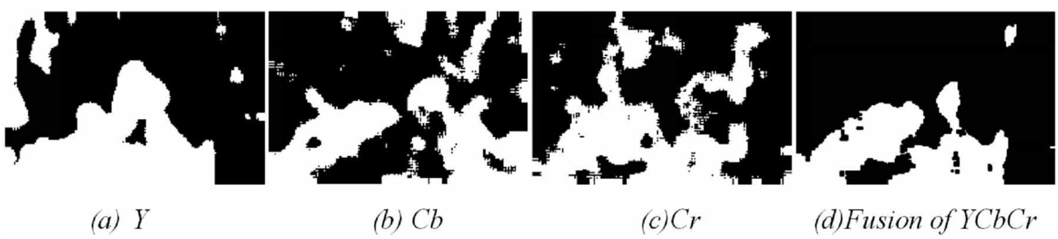


Figure 6. Fusion result of one channel in YCbCr space and fusion result of YCbCr color space



where the detection is tamper, and the dark part represents the untampered area in the image. It is clearly seen that the tampering area can be detected in every channel of the RGB space, while the misdetection area is different. In the other two color spaces, not all three channels can detect tampering areas. By comparing the results of the fusion of these three different color Spaces, it can be clearly seen that RGB space provides better detection results.

Experimental Results With Different Texture Features

In this section, the authors compared three texture features of LBP, LOOP and LDP. The experimental results are shown in Figure 7- Figure 9, where the tampering image is shown in Figure 1.

The images (a) in Figure 7- Figure 9 respectively represent the results of different texture images extracted from first-order difference images. The images (b) in Figure 7- Figure 9 show the result obtained by feeding three features of mean, standard deviation, and variance into fuzzy c-means clustering method. The images (c) in Figure 7- Figure 9 show the result which used histogram feature. Moreover, the images (d) in Figure 7- Figure 9 show the results which combine three features and histogram feature together.

It can be found from (a) in figure 7 - figure 9 that after the image is processed by the LBP method and the LOOP method, a slight change at the tampering position can be seen from the untampered

Figure 7. The experiment of LBP texture feature

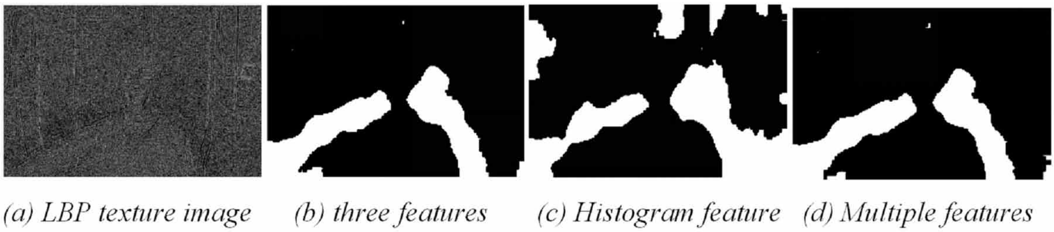


Figure 8. The experiment of LOOP texture feature

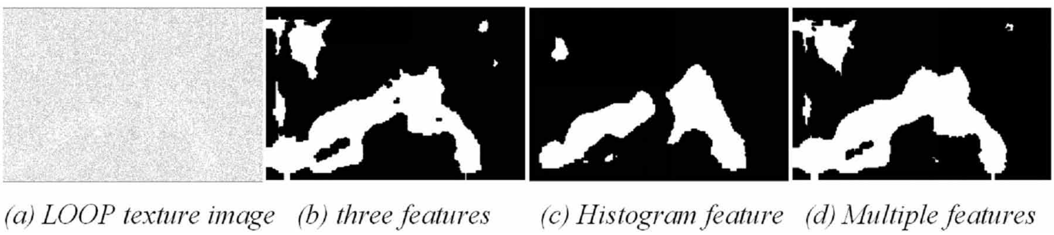
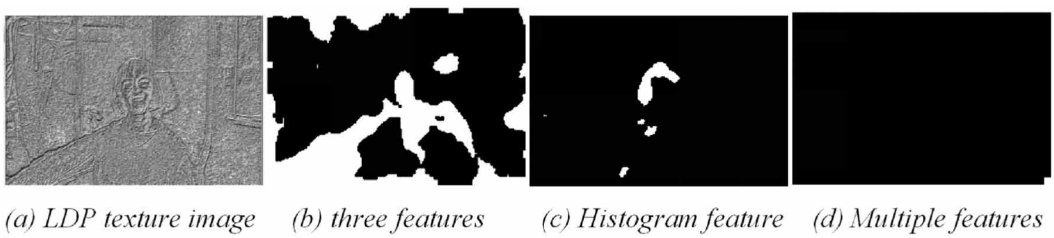


Figure 9. The experiment of LDP texture feature

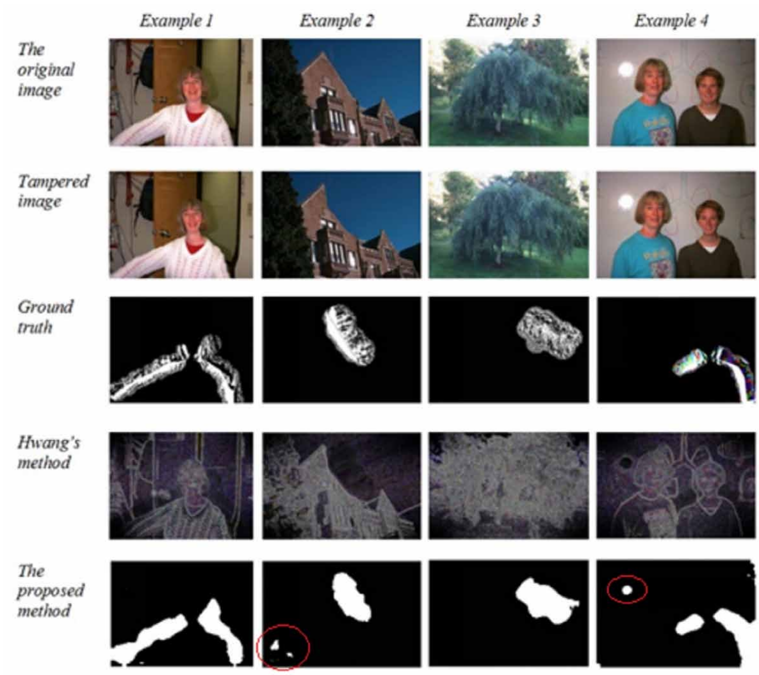


position, and the change is difficult to locate by the human eye, so the authors extract the features of the tampering to locate the tampering location. From the observation of the images (c) in Figure 7- Figure 9, the texture image using the LBP method can see changes in the tamper region that are significantly different from the untampered region. The LOOP method is not well observed because its texture image is bright. Such changes are unably seen with the LDP method. As can be seen from (b)(d) in Figure 7-Figure 9, for the LBP and LOOP methods, the tampering region can be roughly detected by using the three features and the histogram features, while the LDP method could be unable to detect the tampering area. Similarly, in Figure (d), the LBP and LOOP methods are capable of detecting tampering by combining the use of three features and histogram features, while LDP cannot detect. So, the authors chose LBP and LOOP to detect the tampering. Regarding features, we use mean value, variance, standard deviation and histogram features.

Experimental Results Compared With the Existed Method

In order to verify our method is easier to judge the location of the tampering by the human eye compared with the existed method(Hwang MG, 2016), the experiments are done, and the partial results are shown in Figure 10.

Figure 10. The experimental results compared with the existed method



The first row shows the original image, the second row shows the tampered image, the third row shows the ground truth image, the fourth row shows the result of Hwang's method, and the fifth row shows the detection results of our method.

It can be seen from the experimental results that the Hwang's method could unably mark the specific location of local deformation and requires the human eye to observe the deformation position. Our method could better mark the location of local deformation and is easier to judge where is the tampering area. Our results are almost the same as ground truth, but Hwang's results are different from the ground truth. However, our method still has certain defects, and it is easily affected by brightness. High brightness or low brightness will cause false detection, as shown in the red circle of example 2 and 4 in figure 10.

CONCLUSION

This paper mainly studies the blind detection method for the location of local deformation of the image. The authors extracted the trajectory of the local deformation by using the first-order difference image and the corresponding texture features, and locate the position of the local deformation by using the proposed multiple fusion strategies. The experimental results show that our proposed method can not only detect whether an image has been tampered, but also locate the specific location of the local deformation. However, it can be seen that our method will cause false detection where the brightness of the image changes drastically, which is a problem that needs to be solved in future work. The direction of local deformation of the image is also the focus of future research.

ACKNOWLEDGMENT

This research was supported by the National Natural Science Foundation (61771340, 61602344), Tianjin Natural Science Foundation (18JCYBJC15300).

REFERENCES

Gustafsson. (1993). *Interactive Image Warping*. Helsinki University of Technology.

Bharati, A., Singh, R., Vatsa, M., & Bowyer, K. W. (2016). Detecting Facial Retouching Using Supervised Deep Learning. *IEEE Transactions on Information Forensics and Security*, *11*(9), 1903–1913. doi:10.1109/TIFS.2016.2561898

Birajdar, G. K., & Mankar, V. H. (2014). Blind method for rescaling detection and rescale factor estimation in digital images using periodic properties of interpolation. *AEÜ. International Journal of Electronics and Communications*, *68*(7), 644–652. doi:10.1016/j.aeue.2014.01.013

Cao, H., Zhao, Y., & Ni, R. (2012). Detection of interpolation using correlation coefficients. *International Conference on Signal Processing*.

Chakraborti, T., Mccane, B., Mills, S., & Pal, U. (2018). LOOP Descriptor: Local Optimal-Oriented Pattern. *IEEE Signal Processing Letters*, *25*(5), 635–639. doi:10.1109/LSP.2018.2817176

Cozzolino, D., Gragnaniello, D., & Verdoliva, L. (2013). *A novel framework for image forgery localization*. Computer Vision and Pattern Recognition.

Hwang, M. G., Kim, S. M., & Har, D. H. (2017). A method of identifying digital images with geometric distortion. *The Australian Journal of Forensic Sciences*, *49*(1), 93–105. doi:10.1080/00450618.2015.1128970

Jabid, T., Kabir, M. H., & Chae, O. (2010). Gender Classification Using Local Directional Pattern (LDP). *International Conference on Pattern Recognition*.

Johnson, M. K., & Farid, H. (2006). Exposing digital forgeries through chromatic aberration. *ACM Workshop on Multimedia and Security*.

Kee, E., & Farid, H. (2011). A perceptual metric for photo retouching. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(50), 19907–19912. doi:10.1073/pnas.1110747108 PMID:22123980

Korus, P., & Huang, J. (2016). Multi-Scale Fusion for Improved Localization of Malicious Tampering in Digital Images. *IEEE Transactions on Image Processing*, *25*(3), 1312–1326. doi:10.1109/TIP.2016.2518870 PMID:26800540

Korus, P., & Huang, J. (2017). Multi-Scale Analysis Strategies in PRNU-Based Tampering Localization. *IEEE Transactions on Information Forensics and Security*, *12*(4), 809–824. doi:10.1109/TIFS.2016.2636089

Li, W., Yuan, Y., & Yu, N. (2009). Passive detection of doctored JPEG image via block artifact grid extraction. *Signal Processing*, *89*(9), 1821–1829. doi:10.1016/j.sigpro.2009.03.025

Shapiro, L. G. (n.d.). *Object and Concept Recognition for Content-Based Image Retrieval*. http://imagedatabase.cs.washington.edu/groundtruth/

Mahdian, B., & Saic, S. (2008). Blind Authentication Using Periodic Properties of Interpolation. *IEEE Transactions on Information Forensics and Security*, *3*(3), 529–538. doi:10.1109/TIFS.2004.924603

Muhammad, G., Alhammadi, M. H., Hussain, M., Mirza, A. M., & Bebis, G. (2013). Copy move image forgery detection method using steerable pyramid transform and texture descriptor. *Conference on Computer as a Tool*. doi:10.1109/EUROCON.2013.6625188

Ojala, T., Pietikainen, M., & Harwood, D. (1994). Performance evaluation of texture measures with classification based on Kullback discrimination of distributions. *International Conference on Pattern Recognition*. doi:10.1109/ICPR.1994.576366

Popescu, A. C., & Farid, H. (2005). Exposing digital forgeries by detecting traces of resampling. *IEEE Transactions on Signal Processing*, *53*(2), 758–767. doi:10.1109/TSP.2004.839932

Serra, J. (1983). Image Analysis and Mathematical Morphology. *Biometrics*, *39*(2).

Zhao, X., Li, J., Li, S., & Wang, S. (2010). Detecting digital image splicing in chroma spaces. *International Workshop on Digital Watermarking*.

Yongzhen Ke obtained his BE of computer application at Tianjin Polytechnic University in China in 1997, and received his ME and PhD of computer application at Tianjin University in China in 2000 and 2008, respectively, on the research subjects of information security, image processing and data visualization. From 1997 to the present he worked in Tiangong University, China. Now he is a professor at School of Computer Science and Technolgy, Tiangong University, China. His current research interests include image processing, digital image forensic, and information visualization.

Yiping Cui is a master degree candidate of Tiangong University,China. She got the B.E. degrees at School of Computer Science and Technolgy in Tiangong University. Her research interests during period of postgraduate is digital image forensic.