Design of Image Copy-Paste Forensics System Based on Moment Invariants

Shen Jingyi, Yao Yun, Mei Hao

(School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China）

Abstract

Copy-paste tampering is a common type of digital image tampering, which refers to copying a part of the image area in the same image, and then pasting it into another area of the image to generate a forged image, so as to carry out malicious operations such as fraud and framing. This kind of malicious forgery leads to the security problem of digital image. The research of digital image copy paste forensics has important theoretical significance and practical value. For digital image copy-paste tampering, this paper is based on moment invariant image copy paste tampering detection algorithm, and use Matlab software to design the corresponding tampering forensics system.

Keywords: Copy-paste forgery; image tampering detection; forensics system

0 Introduction

With the rapid development of information technology, digital image is widely used as a common information carrier. Some popular digital image editing tools (such as Photoshop, Meitu, Pitu, etc.) are becoming more and more powerful and intelligent, which makes it easier to edit images without leaving visual traces even if they are not experts in image processing. It is understandable if people are just to beautify the well-intentioned purpose of decorating pictures, but malicious people will take advantage of this convenient environment to forge some images to achieve fraudulent and other bad purposes. In recent years, deliberate image forgery events occur frequently in newspapers and magazines, the Internet, court trials, scientific research achievements, military information and other fields, researchers pay more attention to the authenticity and integrity of digital images because once these forged images are used in illegal situations, it will cause incalculable losses.

In view of this, the security of digital image content has become one of the research hotspots in the field of information content security, which not only has important theoretical research significance, but also has good practical significance.

Image copy-paste tampering is a common kind of image tampering, which is easy to operate and implement. Image copy-paste tampering is to copy a part of the image area in the same image, then paste it into another area of the image and save it as a new image. As the copy-paste tampering operation is processed in the same image, the physical characteristics of the composite image will hardly change, so the forgery marks will not be too obvious, which makes it difficult to distinguish the authenticity of the processed image. In the actual image tampering, in order to make the tampered image closer to the real image, the tamper often carries out some post-processing operations on the image, which makes the forged areas in the image more difficult to detect. Therefore, it is of great significance to analyze and detect this kind of image tampering in the field of image forensics.

Aiming at the copy-paste tampering of digital images, on the basis of previous work, the feature data is constructed by Hu invariant moments, and the corresponding tampering forensics system is designed by using Matlab software. The experimental results verify the effectiveness of the system.

1 Image copy-paste forensics algorithm based on invariant moment

1.1 Color image graying

RGB tricolor mode is a common color mode of color image. Given a suspicious color image, this paper first grays it for further feature data analysis and extraction process. The formula of color image graying processing [1] is as follows:

$$Y = 0.299R + 0.587G + 0.114B \tag{1}$$

where R、G and B represent the red, green and blue components of the color image, respectively, and Y represents the grayed image.

1.2 Hu invariant moment extraction

Hu moment invariants [2] are robust to conventional image processing and image geometric transformation, which can better describe the morphological features in the image.

The order moment of the image $f(x, y)$ is defined as:

$$m_{pq} = \sum_{x}\sum_{y} x^{p} y^{q} f(x, y) \tag{2}$$

where $p, q = 0,1,2,\cdots$, the quotient obtained by dividing each image moment by its zero-order moment is independent of shape-scale change. The corresponding center distance is defined as:

$$u_{pq} = \sum_{x}\sum_{y} (x-\bar{x})^{p} (y-\bar{y})^{q} f(x, y) \tag{3}$$

where $\bar{x} = m_{10} / m_{00}, \bar{y} = m_{01} / m_{00}$.

The central moment of the normalized order is shown in formula (4):

$$\eta_{pq} = \frac{u_{pq}}{u_{pq}^{\gamma}} \tag{4}$$

where $\gamma = \frac{p+q}{2} + 1$.

On the basis of the above, seven Hu invariant moment groups which satisfy the shape scale, rotation and translation invariance can be constructed.

$$\begin{aligned}
h_1 &= \eta_{02} + \eta_{20} \\
h_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
h_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
h_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
h_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{120})^2 - 3(\eta_{21} + \eta_{03})^2] \\
&\quad + (3\eta_{12} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
h_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
&\quad + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
h_7 &= (3\eta_{21} - \eta_{02})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
&\quad + (3\eta_{12} - \eta_{30})(\eta_{21} - \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]
\end{aligned} \tag{5}$$

The image to be detected is divided into a number of image blocks with a size of 8 × 8 pixels, and the adjacent overlapping blocks have only one different column or row. If the size of the image is m × n, then the total number of image blocks should be (m-8+1) × (n-8+1). Each image block is processed, and the h_1 ~ h_7 is calculated. Finally, a 7-dimensional feature vector F is obtained.

1.3 Image block matching

By integrating the eigenvalues of each block into a matrix, a total eigenmatrix can be obtained, in which each row represents the eigenvector of an image block, and the number of rows is equal to the number of image blocks. If the similarity of two image blocks in the image to be detected is high enough, then the similarity of the corresponding feature vectors in the feature matrix must also be very high. If the feature matrix is sorted in dictionary, the suspected tamper blocks and the original image blocks will be arranged together. Since the tampering block is generally not adjacent to the original image block, which means that there must be a certain distance between them, a distance threshold T_d [3] is defined to judge the Euclidean distance between the suspected tampering block and the original image block. As there may be natural small block pairs with high similarity in the original image, these disturbances need to be removed. Since there cannot be only one pair of corresponding blocks in the tampered region and the original region and both have the same displacement vector, a frequency shift threshold T_{shift} [3] is defined. It is used to determine the number of times of occurrence of the corresponding regional displacement vector between the image block in the suspected tampering region and the image of the original region. Whenever there is a pair of continuous rows in the total feature sorting matrix, the algorithm will automatically calculate the displacement vector T_r between them. Then it is normalized and stored in the displacement vector counter C for statistics. Whenever the same displacement vector appears, the counter C is added with 1. When the value $C\left(T_r\right)$ of the counter is greater than the threshold T_{shift}, it can be judged that the area is tampered with and the original region is marked. Through the double threshold judgment, we should be able to eliminate most of the misjudgments because the algorithm will display the two block marks only when the above two thresholds are met at the same time.

1.4 Post processing

The threshold processing in Section 1.3 can match copy-paste image blocks to some extent, but there will still be omissions or misjudgments, so post-processing is needed. The morphological operation is carried out on the preliminary detection results of the algorithm, so as to eliminate isolated blocks and holes, so as to obtain the final detection results.

2 Design of forensics system based on Matlab GUI

Graphical user interface (GUI) is object-oriented programming. GUI is an interface for people to communicate with computers. Users use the mouse to click on graphical objects to execute commands. GUI adopts a graphical operation mode, and what users see and operate are graphical objects, which is easier to operate than the command line interface, and it is also easier for users to understand the functions of the program visually. Matlab GUI [4] has strong graphics processing ability, supports the visualization of algorithms, and has a friendly interface and is easy to operate. This paper uses 3 exit text, 3 push button, 2 axes in Matlab GUI to build a copy-paste tamper forensics system.

3 Experimental results

In order to verify the effectiveness of the system built in this paper, this paper downloads some uncompressed color digital images from the Internet, and then uses Photoshop software to copy-paste and tamper with these images to get the corresponding forged images. Fig. 1 shows an example of copy-paste tamper detection based on this system. Among them, the safflower in the test image on the left of Fig. 1 is a copy-paste tampering area, and the result detected and marked by the system is shown on the right side of Fig. 1. According to the detection results in Fig. 1, we can see that the system can accurately detect and locate the copy-paste tampering area, indicating that the image copy-paste forensics system based on invariant moments designed in this paper is effective.

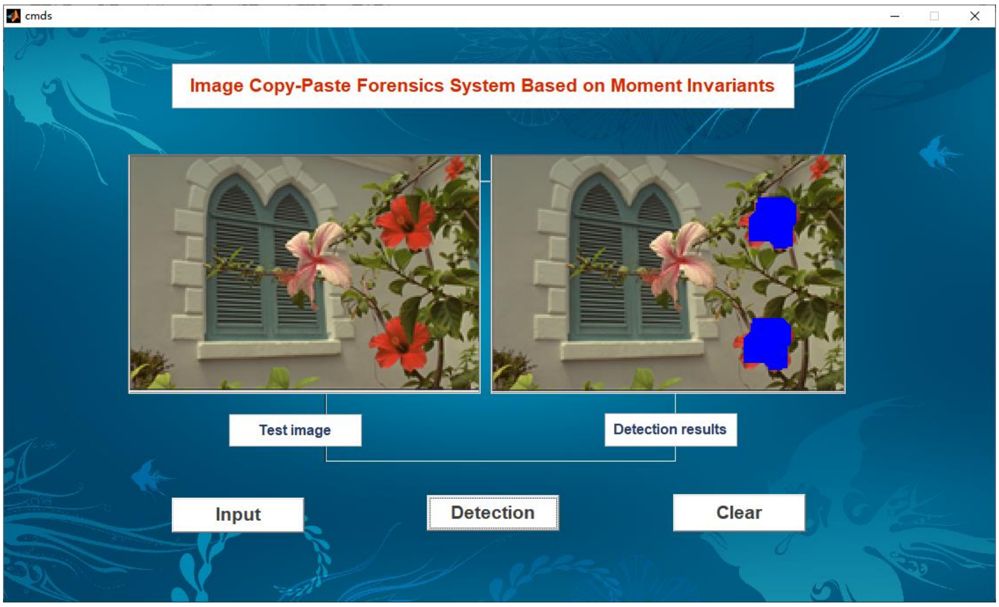
cmds
Image Copy-Paste Forensics System Based on Moment Invariants
Test image
Detection results
Input
Detection
Clear

Fig.1 An example of copy-paste forgery detection based on the proposed system

4 Conclusion

Aiming at the copy-paste tampering of digital images, taking the Hu moment invariant feature as an example, this paper designs an image copy-paste tampering forensics system based on invariant moments. The experimental results verify the effectiveness of the detection system. In the future, we will comprehensively analyze a variety of existing image tampering operations and forensics methods, and build a more perfect image tampering forensics system to meet the actual needs of image forensics.

References

[1] Gonzalez Rafael C, Woods Richard E, and Eddins Steven L. Digital Image Processing (MATLAB Edition). Electronic Industry Press, 2005.

[2] Tejas K, Swathi C, and Rajesh K M. Copy Move Forgery using Hus Invariant Moments and Log Polar Transformations. The 3rd IEEE International Conference on Recent Trends in Electronics, Information and Communication Technology, Bangalore, India, 18-19 May 2018.

[3] Zhao J and Guo J. Passive Forensics for Copy-Move Image Forgery using a method Based on DCT and SVD. Forensic Science International, 2013, (233):158-166.