# An Efficient Forgery Detection Method of Image Inpainting

**Xing Ruofei**

Tianjin Polytechnic University, Tianjin 399 Binshuixi Road, Xiqing District, Tianjin, China

**ABSTRACT**

Image inpainting, which replaces pixels in target region with pixels from multiple image regions, is used for not only repairing the damaged image but also removing the target object. This paper extended a forgery detection algorithm for locating the tampered region of inpainting image. First, the suspicious regions of image are marked. Then, huge amount of false positives caused by smooth area are filtered out. Finally, the final repaired region is located. Experimental results show that the algorithm can effectively filter out a large number of false positives and further identify the real fake area. The proposed method is effective for the scenarios of both the removal of single object and multiple objects.

## 1. INTRODUCTION

Digital images can be seen everywhere and are widely used. Image inpainting can be used to repair the images with a scratch and modify the images in a visually plausible way, while it can be made to be used for malicious purpose of removing the object and fabricating evidences. Hence specific techniques for this kind of images are needed to verify the authenticity of image content.

Objects removal can be achieved by several manipulations, such as copy-move and inpainting in digital images. Many studies have been done on copy-move forgery detection chiefly based on key point match [1-2]. Few researches have been devoted to forensic of image inpainting[3-6]. As the most popular approach, exemplar-based image inpainting[7] which combines structure restoration with texture recovery replaces the image blocks with similar blocks coming from different parts of source region. Q. Wu [3] advocated image inpainting forgery detection, in which zero-connectivity labeling was applied on block pairs to yield matching degree feature and fuzzy membership was computed to measure the level that the block belongs to the forged region. However, the suspicious region was manually selected and the removal of false positives caused by smooth was neglected. In [4], though median comparison of the blocks was employed with a lower computational complexity, they failed to consider the false positives either. D.

Trung [5] adopted distance between the blocks to reduce the false positives, but threshold selection for different images isn't versatile. I. Chang [6] filtered out most of the false positives through the characteristics of the similarity vectors, the length of which is equivalent to distance limit. However, some concentrated regions still do exist.

In order to improve the performance of filtering out false positives caused by smooth region, we present an efficient forgery detection method of image inpainting. The proposed algorithm can be summarized by three main steps. Firstly, similar blocks are labeled. Secondly, false positives in smooth region are filtered. Lastly, the final tampered region is detected.

This paper is organized as follows. Section 2 presents the proposed forgery detection method. Experimental procedure and results are discussed in Section 3. Finally, conclusions are drawn in Section 4.

## 2. FORGERY DETECTION AND LOCALIZATION FOR EXEMPLAR-BASED IMAGE INPAINTING

### 2.1. Labeling similarity block

Considering that image inpainting removes the unwanted objects by replacing blocks with several blocks coming from different parts of the image, the image blocks located at forgery region and the blocks used to override the object

must be highly similar. Accordingly, labeling similarity block is applied to achieve the location of tampered region.

Labeling similarity block is constructed in three steps:

(1) The image is divided into overlapping blocks. Assume that the size of the image is m*n, each overlapping block is N*N.

(2) The matching degree between blocks is calculated. The method in [6] is adopted in this paper. Each overlapped block is chosen as current block (CB) and compared with all the other remained blocks as matching blocks (MB), eventually the most similar block with maximum similarity is selected as best matching block (max_MB).

1. Because the best matching blocks are used to replace the current blocks where the undesired object is located, there must be lots of identical pixels between current blocks and best matching blocks. Subtraction is performed on each two blocks. The difference array ($\sim D_{CB,MB}$) between current block (CB) and matching block (MB) is derived by the following formula.

$$a(i,j) = \begin{cases} 1, if \ R_{CB}(i,j) = R_{MB}(i,j), 1 \le i, j \le N \\ 0, otherwise \end{cases}$$

$$b(i,j) = \begin{cases} 1, if \ G_{CB}(i,j) = G_{MB}(i,j), 1 \le i, j \le N \\ 0, otherwise \end{cases}$$

$$c(i,j) = \begin{cases} 1, if \ B_{CB}(i,j) = B_{MB}(i,j), 1 \le i, j \le N \\ 0, otherwise \end{cases}$$

$$\sim D_{CB,MB}(i,j) = \begin{cases} 0, \ if \ a(i,j) \& b(i,j) \\ \& c(i,j) = 1, 1 \le i, j \le N \\ 1, \ otherwise \end{cases}$$

Where $R_{CB}(i,j)$, $G_{CB}(i,j)$ and $B_{CB}(i,j)$ represent current block for R,G and B color channels, respectively. $R_{MB}(i,j)$, $G_{MB}(i,j)$ and $B_{MB}(i,j)$ denote matching block for R,G and B color channels, respectively.

2. The length of connected zeros in each difference matrix are sorted and the longest length is chosen as similarity value of current block and matched block. The longest length $length(\sim D_{CB,MB}(i,j))$ is computed by the following formula.

$$length(\sim D_{CB,MB}(i,j))$$
$$= \max \{length(connectedzero_p \ in$$
$$\sim D_{CB,MB}(i,j)), 0 < p < P\}$$

Each current block is compared with all the matched blocks to get the corresponding similarity value, and the largest value is regarded as the similarity value of current block. max_length($\sim D_{CB,MB}(i,j)$) is calculated through the formula as follows.

$$\max\_length(\sim D_{CB,MB}(i,j))$$
$$= \max \{length(\sim D_{CB,MB}(i,j))\}$$

Only if the value of S is high enough to the threshold, the corresponding block can be chosen as best matching block. Where S is defined as similarity degree of current block and is calculated as follows.

$$S = \max\_length(\sim D_{CB,MB}(i,j))/(N*N)$$

(3) The largest connected area is marked and the tampered region is determined. According to the position where the largest connected zeros in difference matrix of similarity pairs is located, the corresponding position of current block and matched block are labeled. All the labeled positions are taken as connected areas.

## 2.2. Removal of smooth area

Image blocks with similar color will be mistakenly regarded as similarity blocks and be labeled further, causing lots of false positives. How to filter out such false positives is clarified in this section.

### 2.2.1. Counting up the total numbers

The total number of current blocks and matched blocks belonging to a certain connected area are counted up as follows:

(1) Acquiring the position of current block and matched block in image.

All coordinate position for each current block and matched block are stored in two groups.

(2) Calculating the longest connected component between current block and matched block.

Each current block is subtracted from each corresponding matched block to obtain the difference array, in which the length of each zero-connected component is acquired and the largest will be selected as maximum connected component between current block and matched block.

(3) Checking location relationship between connected component and connected area.

Coordinate position of maximum connected component of current blocks and matched blocks are acquired separately. The position with the location of connected area in section 2.1 will be compared to check whether each maximum connected component is contained in connected area or not.

### 2.2.2. Computing the proportion of matched block pairs

The positions of current block and matched block should be checked and only if both the current block and corresponding matched block are contained in connected area, this block and matched block are regarded as a matched block pair and the number of matched block pairs (NMBP) in connected area is counted. Count the number of all blocks in connected area (NAB). The ratio is obtained by the following formula.

$$ratio = (NMBP * 2)/NAB$$

Those blocks with ratio values greater than threshold (h2) are filtered out.

Since a large range of region with similar color exists in the smooth area, a current block can easily find lots of candidate matched blocks close to it, producing a great many similarity block pairs called false positives. According to

the characteristic of centrality, the block pairs in the connected area are marked and cleared away, achieving the goal of eliminating the false positives caused by smooth area.

### 2.3. Locating the final repaired region

In addition to the false positives caused by smooth area, those blocks used to override the forged region also need to be filtered out. Seeing the tampered region is filled by several parts from the same image, the blocks in real forgery region can be distinguished by judging on whether the area is connected to multiple regions [6].

(1) All the connected areas are categorized into two groups. The first group is the connected area named as selected_area, and the second group is the rest of the areas called remain_area.

(2) All the blocks that share the same position with connected components are stored and marked, the blocks are divided into two groups, the first group is composed of current blocks which are labeled as current_block, the second group consists of matched blocks which are labeled as matched_block.

(3) Each connected area is selected as current selected connected area (CSCA) in turn. The number of current block(current_block) in each current selected connected area(CSCA) is counted. Then the corresponding matched block (matched_block) can be required by similarity computation.

(4) The coordinate position of the matched block (matched_block) is acquired and is used to check which remaining connected area (remain_area) the matched block belongs to.

(5) Matched blocks belong to several remaining connected areas. The number of matched blocks in each remaining connected area is calculated.

(6) Number of matched blocks (NMB) in each remaining connected area and the number of all the matched blocks (NAMB) in selected area are counted, respectively. ratio1 is determined by the

following formula.

$$ratio1 = NMB/NAMB$$

If ratio1 is higher than threshold (h3), the present selected area and the remaining connected areas are labeled at the same time. The tag numbers of each area is increased by one at the same time.

(7) All the values of tag number are verified. The final result are regions left behind, each value of which is larger than threshold (h4 is usually taken as three).

## 3. EXPERIMENTAL RESULTS

The experiments in three scenarios are applied and described as follows. Firstly, an original image with lots of false positives is tested. Secondly, when a single object is eliminated, an image with a great deal of false positives. Thirdly, the tampered image with multiple objects removed is used for forgery detection.

To quantitatively evaluate the performance, recall rates and accuracy rates are defined in the following way [6]:

$$R_{recall} = R_{correct}/(R_{correct} + R_{missing})$$
$$R_{accuracy} = R_{correct}/(R_{correct} + R_{false})$$

Where $R_{correct}$ is the number of pixels correctly identified, $R_{missing}$ is the number of pixels missed, $R_{correct} + R_{missing}$ is the total number of forged pixels, $R_{false}$ is the number of pixels mistakenly regarded as correct pixels, and $R_{correct}+R_{false}$ is the number of pixels totally detected.

To assess the performance for removal of false positives, false positives' percentage is defined as follows.

$$P_{false} = \frac{n_{white}}{n_{all}}$$

Where $n_{white}$ is the number of white pixels (false positives), $n_{black}$ is the number of the whole pixels in the image.

(1) Detection result with removal of smooth area

Figure 1(a) is the original image without modification. A great many similarity blocks are labeled shown in the Figure 1(b). That's because if similarity value between two blocks is high enough, then the two blocks will be marked by similarity block detection. In spite of no tampered region in the image, regions virtually identical such as sky produce lots of false positives, which is the major interference for our final detection. After filtering out similarity blocks, the false positives are generally eliminated in Figure 1(c). $P_{false}$ is 34.83% in Figure 1(b) while $P_{false}$ decreased to 2.87% in Figure 1(c), which demonstrates the effectiveness of our algorithm.
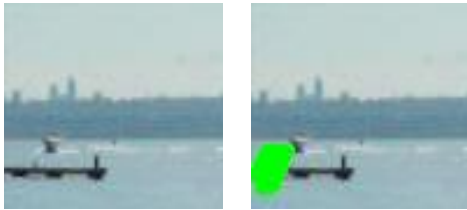


(a)Original Image



(b)Similarity Block Labeling (c)Removal of Smooth Area

Figure 1: Detection for Original Image

(2) Forgery detection of single forged region

In order to demonstrate the robustness of the proposed algorithm, images with lots of false positives are used for detection. Figure 2(a) is the original image. Figure 2(b) is modified by removing the object. Figure 2(c) is forged by the inpainting algorithm [7] Criminisi proposed. Figure 2(d) conducts similarity block labeling. Figure 2(e) is made filtration of false positives in smooth area. Figure 2(f) is the final result after blocks used to cover the object being cleared away. The recall rate is 0.7530 and the accuracy rate is 0.7718. Moreover, $P_{false}$ in Figure 2(d) is

0.5368 and P$_{false}$ decreases to 0.0950 in Figure 2(e).



(a)Original Image(b)Removal of Target Object



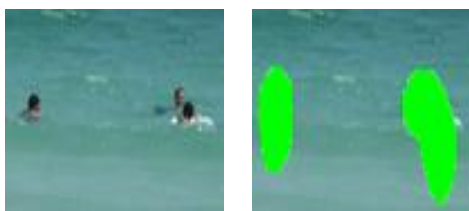(c)Forged Image (d)Similarity Block Detection



(e)Removal of Smooth Area (f)Final Result
Figure 2: Forgery Detection of Single Forged Region

(3) Forgery detection of multiple forged region

To evaluate the performance for multiple forged region. Two regions of the image are removed and used as testing image. A large number of false positives can be found in Figure 2. An image with few false positives is tested in Figure 3. The recall rate is 0.8140 and the accuracy rate is 0.8759, which shows the efficiency of the algorithm.



(a)Original Image (b)Removal of Target Object



(c)Forged Image (d)Similarity Block Detection



(e)Removal of Smooth Area (f)Final Result
Figure 3: Forgery Detection of Multiple Forged Region

## 4. CONCLUSIONS

An efficient forgery detection method is introduced to locate tampered region of image inpainting in this paper. The first step involves similarity block labeling. The second step is related to filtration of false positives in smooth area, which plays an important role in the whole algorithm. The third step deals with the issue that real forgery region is identified from the blocks selected to cover it. Experiments show that the proposed algorithm is effective for original images with lots of false positives itself, forged images with single forged region and forged images with multiple forged regions.

**REFERENCES:**

[1]  F.Liu and H.Feng, "An Efficient Algorithm for Image Copy-move Forgery Detection Based on DWT and SVD", International Journal of Security and Its Applications, vol. 8, no. 5, (2014), pp. 377-390.

[2]  J.C.Lee, "Copy-move Image Forgery Detection Based on Gabor Magnitude", Journal of Visual Communication and Image Representation, vol. 31, (2015), pp. 320-334.

[3]  Q.Wu, S.Sun, W.Zhu, G.Li, D.Tu and C.He, "A Blind Forensic Algorithm for Detecting Doctored Image Region by Application of Exemplar-based Image Completion", ACTA AUTOMATICA SINICA, vol. 35, no. 3, (2009), pp. 239-243.

[4]  K.S. Bacchuwar and K.R. Ramakrishnan, "A Jump Patch-block Match Algorithm for Multiple Forgery Detection", Proceedings of IEEE International Multi-Conference on Automation, Computing, Communication, Control and Compressed Sensing(iMac4s), Kerala, India, (2013) February 22-23.

[5]  D.Trung, A.Beghdadi and M.C.Larabi, "Blind Inpainting Forgery Detection", Proceedings of IEEE Global Conference on Signal and Information Processing(GlobalSIP), Atlanta, GA, United states, (2014) December 3-5.

[6] I.  Chang, J. Yu and C.C. Chang, "A Forgery Detection Algorithm for Exemplar-based Inpainting Images Using Multi-region Relation", Image and Vision Computing, vol. 31, no. 1, (2013), pp. 57–71.

[7]  A.Criminisi, P.Perez and K.Toyama, "Region Filling and Object Removal by Exemplar-Based Image Inpainting", IEEE TRANSACTIONS ON IMAGE PROCESSING, vol. 13, no. 9, (2004), pp. 1200-1212.