Multimed Tools Appl
DOI 10.1007/s11042-017-4829-0

A robust forgery detection algorithm for object removal by exemplar-based image inpainting

Dengyong Zhang[1,2] · Zaoshan Liang[1] · Gaobo Yang[1] · Qingguo Li[3] · Leida Li[4] · Xingming Sun[5]

Received: 22 March 2016 / Revised: 2 February 2017 / Accepted: 15 May 2017

Abstract Object removal is a malicious image forgery technique, which is usually achieved by exemplar-based image inpainting in a visually plausible way. Most existing forgery detection approaches utilize similar block pairs between inpainted area and the rest areas, but they invalidate when those inpainted images are further subjected to some post-processing operations such as JPEG compression, Gaussian noise addition and blurring. It is desirable to develop a forensic method which is robust to object removal with post-processing. From some preliminary experiments, we observe that post-processing destroys the similarity of block pairs and simultaneously disturbs the correlations among adjacent pixels to some extent. Inspired by the strong ability of joint probability density matrix (JPDM) in characterizing such correlation, we propose a hybrid forensics strategy. Firstly, our earlier method is employed to detect whether a candidate image is forged or not. Secondly, for those undetected images after the first step, JPDM is computed for each difference array to model the correlations among adjacent DCT coefficients, and the average of these matrixes are computed as feature vectors to further expose tampering traces. Experimental results show that the proposed approach can effectively detect object removal by exemplar-based inpainting either with or without post-processing.

✉ Gaobo Yang
yanggaobo@hnu.edu.cn

[1] School of Information Science and Engineering, Hunan University, Changsha, 410082, China

[2] School of Computer & Communication Engineering, Changsha University of Science & Technology, Changsha, 410005, China

[3] College of Mathematics and Economics, Hunan University, Changsha, 410082, China

[4] School of Information and Electrical Engineering, China University of Mining and Technology, Xuzhou, 2211166, China

[5] School of Computer and Software, Nanjing University of Information Science & Technology, Nanjing, 210044, China

Published online: 22 May 2017

Keywords Passive image forensics · Exemplar-based image inpainting · Post-processing · Joint probability density matrix

1 Introduction

Digital image is the main information source which enriches our daily life everyday and everywhere. Meanwhile, powerful image editing tools lead to large amounts of doctored images without leaving any noticeable traces. This destroys our traditional concept of "*seeing is believing*", which causes serious crises in respect to public confidence. There is an urgent demand for automatic forgery detection to address this issue. Different from active approaches such as digital watermarking or signatures, passive image forensics does not need any prior information about image content and decides the image authenticity by itself. Since original images are usually unavailable, passive image forensics is becoming one of the hottest research topics in the field of image information security [3, 18].

Object removal is widely accepted as a malicious forgery since it might change the semantic content that an image conveys. Existing object removal techniques can be roughly grouped into two categories [18]: copy-move and image inpainting. Copy-move removes undesired object by copying a region from the rest image or another image and then pasting it back to the region of removed object. Due to its simplicity, copy-move is widely used for object removal. In recent years, extensive works have been presented for copy-move forgery detection [1, 11, 14, 17]. Image inpainting was originally presented to restore damaged information and remove scratches for old photographs. It can also be used for object removal forgery by exploiting the patches preserved in surrounding regions to fill the hole left by object removal [2, 4, 9, 21]. Image inpainting can simultaneously maintain both textural and structural consistency. That is, it achieves desirable object removal without leaving any visually annoying artifacts, which makes its passive detection much more challenging. Up to now, there are few works reported about the blind detection of image inpainting. As the first attempt, Wu et al. [21] proposed a forgery detection approach for exemplar-based image inpainting. It exploits zero-connectivity labeling to yield the matching degree of those blocks in suspicious regions, and identifies forged regions by fuzzy membership. However, it requires a manual selection of suspicious regions. Moreover, the full searching of suspicious blocks leads to a high computation complexity. Later, Bacchuwar et al. [2] devised a jump patch-block matching for this kind of forgery detection. It greatly reduces the computational costs, but is still a semi-automatic approach. Chang et al. [4] presented an automatic forgery detection algorithm for exemplar-based image inpainting. Suspicious regions are roughly located by zero-connectivity features, and a multi-region relation technique is exploited to locate forged areas. In our recent work [15], an efficient approach was proposed to detect object removal forgery by exemplar-based image inpainting. It exploits central pixel mapping to speed up suspicious block searching, which maps image blocks into a hash table. Similar block pairs are searched among those blocks with similar hash values. Moreover, fragment splicing detection (FSD) is used to further refine tampered regions.

Sophisticated counterfeiters might further conceal the tampering traces of image inpainting by some post-processing techniques. They include JPEG or double JPEG compression, Gaussian noise and Gaussian blurring. Since almost all existing detection approaches exploit the abnormalities between similar block pairs, post-processing will invalidate them because the similarities between block pairs are likely to be destroyed.

Meanwhile, post-processing operations inevitably disturb the correlations among adjacent pixels to some degree. In this paper, an improved method is proposed to detect object removal by exemplar-based inpainting with/without post-processing. It is actually a hybrid strategy. Firstly, our earlier method [15] is exploited to detect whether a candidate image is inpainted or not. Secondly, for those undetected images, they are further addressed in a steganalytic way to improve the robustness against post-processing. Specifically, joint probability density matrix is exploited to model the correlations among adjacent DCT coefficients, and some statistical features are designed for further blind detection. Since the first step is focus on the the detection of image inpainting without post-processing and the second step aims at the robustness to post-processing, the hybrid approach is expected to achieve desirable detection performance for object removal by exemplar-based inpainting with/without post-processing.

The rest of this paper is organized as follows: Section 2 briefly introduces background knowledge, which includes the preliminary of exemplar-based image inpainting and a brief analysis of correlations among adjacent DCT coefficients. Section 3 presents the proposed approach. Experimental results are provided in Section 4. We conclude this paper in Section 5.

2 Background knowledge

2.1 Preliminary of exemplar-based image inpaining

Criminisi et al. [6] proposed an examplar-based inpainting algorithm by propagating known patches (i.e., examplars) into missing patches gradually. To handle the missing region with composite textures and structures, patch priority is defined to encourage the filling-in of patches on the structure. It can achieve plausible results when inpainting relatively large missing region. Therefore, we briefly introduce Criminisi's algorithm as an example of exemplar-based inpainting. Figure 1 shows the main procedures of exemplar-based image inpainting. Users firstly select a target region to be removed and filled, as shown in Fig. 1a. Given an image with the missing region Ω and the known region Φ, the task of image inpainting is to fill in the target region (i.e., the missing region Ω) using the image information in the source region (i.e., the known region Φ). The boundary of the target region is denoted by $\partial\Omega$. The main steps of image inpainting procedure are summarized as follows.

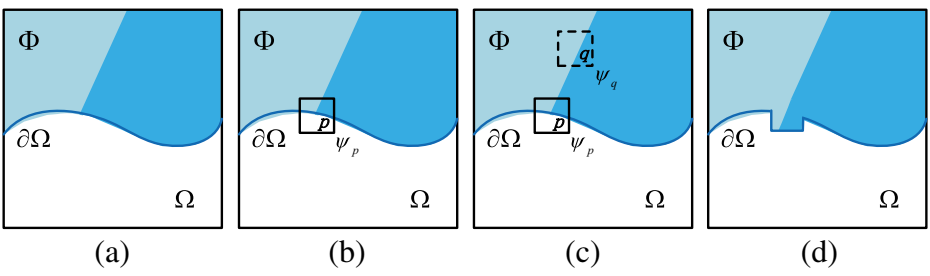


Fig. 1 Exemplar-based image inpainting: (**a**) specify target region Ω; (**b**) select target block Ψ_p; (**c**) search for reference block Ψ_q; (**d**) update priority and contour

Step 1: Compute the priorities of those pixels along $\partial\Omega$ and find the pixel p with the highest priority. Then, a block Ψ_p centered at pixel p is selected as target block, as shown in Fig. 1b.

Step 2: Search for the reference block Ψ_q which is the most similar with Ψ_p in the source region Φ, as shown in Fig. 1c.

Step 3: Fill the area to be inpainted in Ψ_p using the corresponding pixels of Ψ_q, and update the priorities among Ψ_p as well as $\partial\Omega$, as shown in Fig. 1d.

Step 4: Repeat step 1 to step 3 until the target region is entirely filled.

The Criminisi's algorithm performs well in terms of both perceptual quality and efficiency. Later, there are a few improved algorithms, which mostly concentrate on improving priority calculation and optimizing patch searching strategies. However, they introduce similar abnormity when filling the target block with its best-match block. It is well-known that there exists large amount of irregular pieces in natural images, which seemingly share the same textured surface but are actually different. Therefore, the filling scheme of exemplar-based inpainting can not represent well the subtle differences in natural scene. Instead, it leads to abnormal similarity between blocks, which provides useful clues for image forensics. This is the idea behind our earlier work [15], which is effective to detect this kind of forgery images in uncompressed format without post-processing.

2.2 The influences of post-processing towards the blind detection of exemplar-based image inpainting

After object removal by exemplar-based image inpainting, sophisticated counterfeiters might conceal the forgery traces by post-processing operations including JPEG compression, Gaussian noise addition and/or Gaussian blurring. Thus, there are at least two operations for those inpainted images with post-processing. The forgery traces left by them might be mixed together, which leads to ambiguous processing artifacts. Some experiments have been conducted to illustrate the influences of post-processing towards blind detection of object removal forgery by exemplar-based inpainting on two popular image databases UCID and yokoya. Table 1 reports the average ratios of pixel changes and PSNR (Peak Signal-to-Noise Ratio) between the forged images and their original ones. From it, more than 20% pixels are changed, and the qualities of the resultant images also degrade significantly due to these operations. Consequently, we infer that exemplar-based image inpainting and post-processing greatly modify many pixels and thus destroy the inherent correlations among adjacent pixels to some degree. This motivates us to devise more discriminative features to distinguish forged images from original ones. Moreover, since the correlations among adjacent pixels have been exploited to detect image forgery [5, 10, 12, 19, 25], we are inspired to adopt joint probability density matrix (JPDM) [16] to model the correlations among adjacent pixels for the detection of exemplar-based inpainted images with/without post-processing.

3 The proposed forgery detection algorithm

Figure 2 is the framework of the proposed passive forensics approach. It is a hybrid detection scheme which includes two stages by our recent work [15] and the JPDM-based approach, respectively. Specifically, the first stage is to detect whether a candidate image is forged

or not by our recent work [15]. Since it has weakness in detecting those inpainted images with post-processing, those undetected images or original images are further detected with the JPDM-based approach in the second stage, which is the key content of this paper. Please note that the JPDM-based approach is also made up of two stages: training stage and testing stage. In the training stage, sufficient natural images and their forged ones after exemplar-based image inpainting are collected for training. That is, the JPDM-based statistical features are extracted from them and input into an ensemble classifier [13] for blind forensics. In the testing stage, the same JPDM features are extracted from these undetected images by our recent work [15]. Then, they are input into the trained classifier for binary decision about their authenticities.

3.1 Extraction and analysis of JPDM-based features

It is well-known that adjacent pixels or DCT coefficients are highly correlated for natural images. When exemplar-based image inpainting is used for object removal, the patches in the known regions are copied to fill in the hole left by object removal. Since the original pixels are replaced with new ones during hole-filling, this inevitably changes the DCT coefficients within the region of original object. Apparently, this might also destroy the correlations of those coefficients within the holes and their neighbors. Thus, it is expected that the block-based correlations among DCT coefficients are promising clues for image forensics. Meanwhile, post-processing might further distort the afore-mentioned correlations, as summarized in Table 1. In this paper, we are motivated to design some steganalytic features to expose the traces of object removal left by exemplar-based image inpainting. Specifically, the intra-block and inter-block correlations among DCT coefficients are modeled with JPDM to extract statistical features for blind forensics. Moreover, their effectiveness and robustness against post-processing as analyzed by some preliminary experiments.

3.1.1 JPDM-based feature extraction

There are intra-block correlations among adjacent DCT coefficients within an 8×8 block after block-based discrete cosine transform (BDCT) [16]. It is claimed that the quantified DCT coefficients and the differences between adjacent DCT coefficients are subject to Generalized Gaussian Distribution (GGD) [13]. Moreover, since patch matching is involved in the process of expemplar-based image inpainting, there are also inter-block dependencies among those coefficients located in the same position of neighboring blocks. That is, the inter-block correlations among neighboring blocks are also important for object removal forgery detection. The extraction procedures of intra-block and inter-block correlations are summarized as follows.

Table 1 Average modification ratio, and PSNR with/without post-processing operations

Operations	Modification Ratio(%)	PSNR(dB)
JPEG compression	35.63	43.90
Gaussian noise addition	37.82	38.30
Gaussian blurring	36.55	36.07
Exemplar-based inpainting	20.33	24.18

Step1. Define four pixel difference arrays in four directions including horizontal, vertical, diagonal and minor diagonal, which are notated as f_h, f_v, f_d and f_{md}, respectively. That is,

$$f_h(x, y) = f(x, y) - f(x, y + 1) \tag{1}$$

$$f_v(x, y) = f(x, y) - f(x + 1, y) \tag{2}$$

$$f_d(x, y) = f(x, y) - f(x + 1, y + 1) \tag{3}$$

$$f_{md}(x, y) = f(x, y + 1) - f(x + 1, y) \tag{4}$$

where $f(x, y)$ denotes a given image, $x \in [1, M - 1]$, $y \in [1, N - 1]$, and M and N are the width and height, respectively. The different values of 2D arrays are constrained to be in the range $[-T, T]$. That is, if the values in the array are larger than T or less than -T , those values are truncated to $[-T, T]$. Therefore, T is the threshold which determines the dimension of f_h, f_v, f_d and f_{md}. Apparently, their dimensions are $(2T + 1)^2$.

Step2. JPDM is exploited to model the correlation among adjacent pixels. For four pixel difference arrays f_h, f_v, f_d and f_{md}, their JPDMs J_h, J_v, J_d and J_{md} are computed as follows:

$$J_h(i, j) = \frac{\sum_{u=1}^{M} \sum_{v=1}^{N-2} \delta(f_h(u, v) = i, f_h(u, v + 1) = j)}{M \times (N - 2)} \tag{5}$$

$$J_v(i, j) = \frac{\sum_{u=1}^{M-2} \sum_{v=1}^{N} \delta(f_v(u, v) = i, f_v(u + 1, v) = j)}{(M - 2) \times N} \tag{6}$$

$$J_d(i, j) = \frac{\sum_{u=1}^{M-2} \sum_{v=1}^{N-2} \delta(f_d(u, v) = i, f_d(u + 1, v + 1) = j)}{(M - 2) \times (N - 2)} \tag{7}$$

$$J_{md}(i, j) = \frac{\sum_{u=1}^{M-2} \sum_{v=1}^{N-2} \delta(f_{md}(u, v + 1) = i, f_{md}(u + 1, v) = j)}{(M - 2) \times (N - 2)} \tag{8}$$

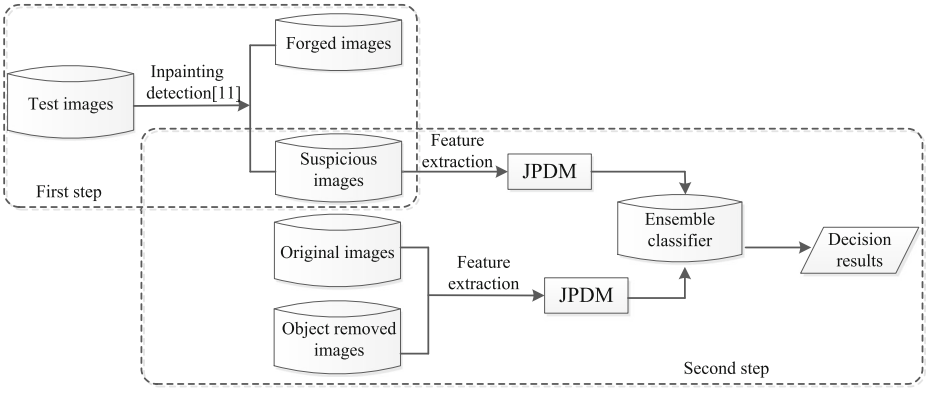


Fig. 2 The proposed two-stage forensic framework

where i and j are integers, $\delta(m, n)$ is a condition function defined in (9). That is, when its arguments are satisfied, the condition function $\delta(m, n)$ will be equal to 1, or else be 0.

$$\delta(m = A, n = B) = \begin{cases} 1 & \text{if } m = A \text{ and } n = B \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

Step3. Compute the average of four JPDMs J_h, J_v, J_d and J_{md} as J_{ave}, which is the intra-block correlation feature. J_{ave} has the same dimension with J_h, J_v, J_d and J_{md}. To tradeoff between computational complexity and feature discriminability [6], T is set with 4 here. Thus, the dimension of J_{ave} is $(2T + 1)^2 = 81$.

$$J_{\text{ave}}(i, j) = \{J_h(i, j) + J_v(i, j) + J_d(i, j) + J_{md}(i, j)\}/4 \tag{10}$$

Step4. Extract an AC coefficient from each $b \times b$ block with the same frequency position to form a JPEG 2-D array. Since there are $(b^2 - 1)$ AC coefficients in each block except one DC coefficient, we can construct the mode 2-D arrays (denoted as $F_1 - F_{b^2-1}$ in the scanning order, as shown in Fig. 3). For each mode 2-D array $F_i (i = 1, ..., b^2 - 1)$, calculate the difference mode 2-D arrays F_{ih}, F_{iv}, F_{id} and F_{imd} along the horizontal, vertical, diagonal and minor diagonal directions similar with (1)–(4), respectively. Then, compute four JPDMs J_{ih}, J_{iv}, J_{id} and J_{imd} similar with (5)–(8), respectively.

Step5. Compute the averaged JPDM J_{iave} of J_{ih}, J_{iv}, J_{id} and J_{imd}, which is similar with (10). The ranges of i and j are also limited within $[-4, +4]$. Thus, there are also $(2 \times 4)^2 + 1 = 81$ inter-block correlation features for blind forensics.

$$J_{\text{iave}} = \frac{\sum_{i=1}^{b^2-1} (J_{ih} + J_{iv} + J_{id} + J_{ibd})/4}{b^2 - 1} \tag{11}$$

After extracting both intra-block and inter-block correlation features, both feature sets are combined into a unified correlation feature set in DCT domain. In summary, there are also 81 features for blind forensics.

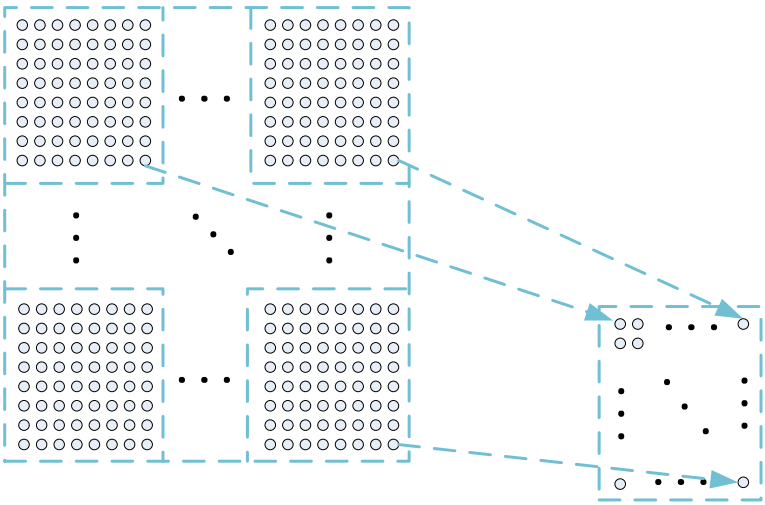


Fig. 3 Construction of mode 2-D array

3.1.2 The effectiveness analysis of forensics features for inpainted images without post-processing

Most existing exemplar-based image inpainting techniques share the same mechanism of copying the best-match patch to the target block to be filled [6, 20], their differences are the priority calculation and the searching strategy for best-matching block. However, hole-filling usually introduces two kinds of footprints into the original image: one is the similar abnormalities between image block pairs, and the other is splicing artifacts. That is, the tampered regions are replicated by multiple reference regions. Figure 4 shows an example of splicing artifacts, in which A is the inpainted region, and B, C, D, E and F are five reference regions. They are located at different image regions, but found by the searching for best-matching patches.

Apparently, the forged region includes several copies of reference regions with different color distributions. This brings at least two side artifacts on the correlations among adjacent pixels. First, there are great changes along the boundaries of removed object, which has influence on the intra-block and inter-block correlations. Second, there are also great changes for the boundary pixels between different patches within the forged object, which may destroy the intra-block and inter-block correlations in DCT domain as well. Figure 5 shows the changes of inter-block correlation caused by object removal, where four green small circles refer to the AC coefficients with the same frequency position in adjacent DCT blocks, and the yellow and purple areas are tampered regions. It is apparent that both AC coefficients in two tampered regions will be modified, which inevitably brings some changes to the correlation among these four AC coefficients.

3.1.3 The analysis of robustness against post-processing

After removing undesirable objects from digital image by exemplar-based image inpainting, sophisticated counterfeiters often conceal the forgery traces by further post-processing operations including JPEG compression, Gaussian noise and/or Gaussian blurring. Please note that any post-processing can also be regarded as an image forgery. That is, there are at least two forgery operations for those doctored images after post-processing. The tampering traces left by multiple operations might be mixed together, which leads to ambiguous

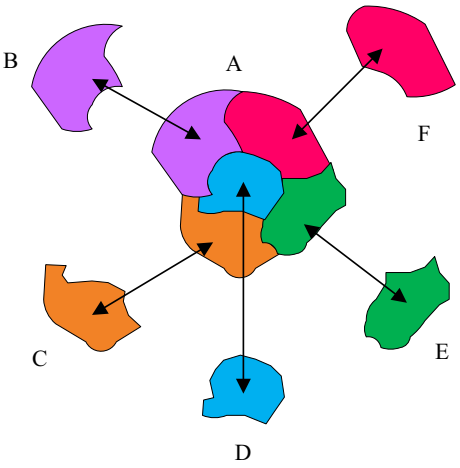


Fig. 4 Inpainted region and its reference regions

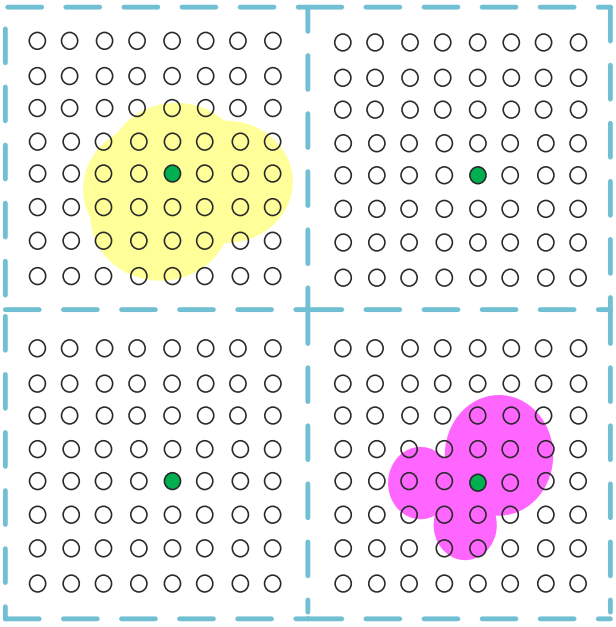


Fig. 5 The changes of correlation caused by object removal

processing artifacts. Luckily, we found by experiments that post-processing further destroys the inherent intra-block and inter-block correlations, which makes the devised forensics features much easier to distinguish the forged image from the original ones. Figure 6 shows the influences of post-processing towards the forgery detection of object removal by exemplar-based inpainting, where Fig. 6a is the original image and Fig. 6b is the tampered image after object removal by exemplar-based inpainting. Figure 6c and d are JDPMs before and after tampering, respectively. Figure 6e is the differences of forensic features between the original image and tampered one without post-processing, whereas Fig. 6f–g are the feature difference arrays when the tampered image are further processed by JPEG compression ($QF = 75$), Gaussian noise addition ($SNR = 35dB$) and Gaussian blurring (standard deviation is 0.5), respectively. By comparing Fig. 6e with Fig. 6f–g, we observe that JPEG compression, Gaussian noise and Gaussian blurring increase the feature differences between original image and the tampered ones. As a result, the devised forensic features are effective for the detection of object removal and robust to post-processing.

3.2 Ensemble classifier for passive forensics

After extracting the JPDMs features, a pattern recognition classifier is required to determine whether a candidate image is exemplar-based inpainted image or not. In this paper, the ensemble classifier is adopted as the classifier, simply because of its good compromise between computational complexity and detection accuracy [13]. In the training phase, two classes of images (inpainted or not-inpainted) are represented by the JPDM-based features. Then, several base learner are employed, which is a linear discriminant analysis (LDA) classifier, to train with a randomly selected subset of all the training samples on a

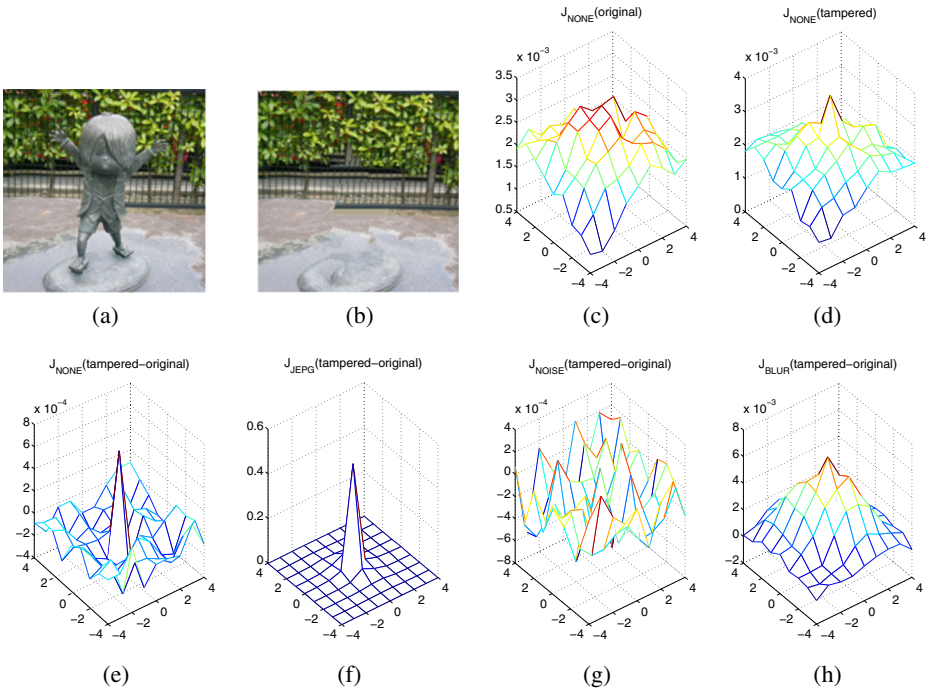


Fig. 6 Influences of post-processing. **a** Original image, (**b**) tampered image, (**c**) JPDMs before tampering (**d**) JPDMs after tampering, (**e**) feature differences without post-processing, (**f**) feature differences after JPEG compression, (**g**) feature differences after adding Gaussian noise, (**h**) feature differences after Gaussian blurring

randomly selected subspace of the JPDM-based feature space. To further obtain satisfactory classification accuracy, these sufficient number of base learners are combined to construct the ensemble classifier. In the final classification stage, the JPDM-based feature vectors extracted from the candidate image are input into the classifier after training. Therefore, the candidate image will be classified into either class of image: inpainted or non-inpainted.

4 Experimental results and analysis

4.1 Experimental setup

To evaluate the performance of the proposed approach, a series of experiments are conducted. The hardware configuration is a desktop computer (Intel(R) Core(TM) i3-3120M CPU @ 2.5GHz, 4.0GB RAM). The forensics detector is implemented in Matlab 2010b, and the Ensemble classifier is directly downloaded[1] An image dataset, which is composed of 1438 images, is created from two popular image databases including UCID[2] with 1338

[1]Available at http://dde.binghamton.edu/download/ensemble/.

[2]Available at http://homepages.lboro.ac.uk/cogs/datasets/ucid/ucid.html.

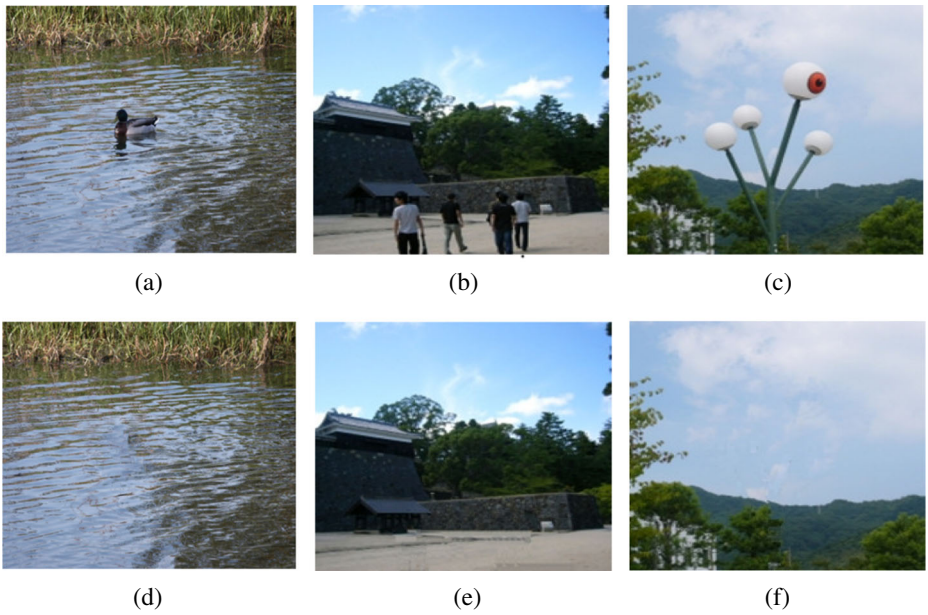


Fig. 7 Sample images

images and yokoya[3] with 100 images. These images have diverse types of contents including landscape, buildings, animals, human, and so on. Then, Wang's inpainting approach [20] is exploited for object removal simply because it achieves relative better visual quality than the existing approaches when a large object is removed. Figure 7 shows some sample images, where Fig. 7a–c are the original images and Fig. 7d–f are the corresponding forged images after object removal. To highlight the robustness of the proposed approach, several groups of experiments are conducted under the following scenarios: (1) object removal without post-processing; (2) object removal integrated with post-processing; (3) original images and inpainted images without post-processing; (4) original images and inpainted images with post-processing. To make comparisons with the existing blind detection approaches, two state-of-the-art works (the Chang's method [4] and our recent work [15]) are selected as baselines for performance evaluation because they are the most representative works for image inpainting forgery detection. That is, there are three detectors for experimental comparisons, which are tested with the same hardware/software environment and test image datasets. Moreover, a 3-fold cross validation strategy is exploited for training and testing.

4.2 Choice of block size *b*

For the proposed approach, there is only one parameter required to be determined, which is the block size b. It has direct influence on the accuracy of the feature J_{iave}. Obviously, if b is too big or too small, the correlations among the coefficients with the same position in different blocks, which are modeled with JPDM, might be insufficient to distinguish inpainted images from original ones. To assess the influence of block size b, we have conducted some

[3] Available at http://yokoya.naist.jp/research2/inpainting/.

experiments on those inpainted images without post-processing, and the tampering ratios are about 30%. The experiment results are summarized in Fig. 8. From it, it is apparent that with the increase of b, the detection accuracy first increases and then decreases as we expect. In this paper, the block size b is set with 8, which achieves the highest detection accuracy.

4.3 Detection for object removal without post-processing

For inpainted images with tampering ratios from 10% to 50% with a step size of 10%, we test the effectiveness of the proposed method. All the uncompressed original images and the tampered ones are converted in JPEG images with the best quality factor, where all the elements in the quantization table are 1s. Table 2 summarizes the detection accuracies of three detectors. Please note that since the proposed approach combines both our earlier method [15] and the JPDM-based new features, their contributions to the overall detection accuracy are also provided separately. With the increase of tampering ratios, the three detectors achieve higher detection accuracies because the similar abnormalities between image block pairs and the splicing artifacts between the inpainted region and its neighboring region increase as well. Moreover, the Chang's method [4] and Yang's method [15] achieve higher detection accuracies than solely JPDM-based features. The reasons are two-folds: Firstly, both methods are specially designed for the forgery detection of exemplar-based image inpainting without post-processing, and they are based on similar abnormalities between image block pairs. The patch matching based approaches achieve especially desirable detection results when the forged images are not further processed by post-processing after object removal. Secondly, For inpainted images with relatively low tamper ratios, the disturbance of inherent correlations among adjacent pixels, which locates between the inpainted region and its neighboring region, is not particularly significant. From Table 2, it can be observed that the proposed approach achieves a relatively higher detection accuracy than individual method [15] or JPDM-based features. That is, the JPDM-based features can further identify those undetected tampered images by the earlier method [15], and they both improve the detection accuracy.

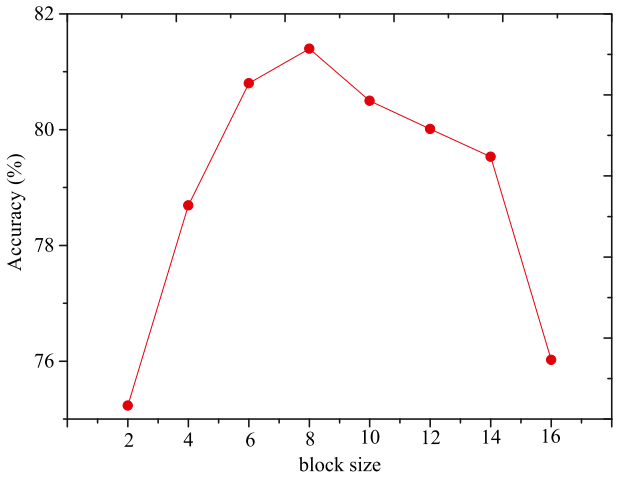


Fig. 8 Experiment results with different block size

Table 2 Detection performance among three detectors for object removal without post-processing (%)

Tampering ratio	Chang's [4]	Yang's [15]	JPDMs	Proposed method
10	87.63	88.92	72.41	89.94
20	90.86	92.33	76.33	93.83
30	94.31	94.50	81.57	95.98
40	96.93	97.02	85.85	98.36
50	98.54	98.54	88.46	99.18
Average	93.65	94.26	80.92	95.45

4.4 Detection of inpainted images with post-processing

4.4.1 Detection for object removal integrated with JPEG compression

Digital images are usually kept in JPEG format for storage. The tampered images after object removal may also be compressed with JPEG. It is well-known that quantization and inverse quantization in JPEG compression are lossy operations. Because of the rounding errors, JPEG compression not only changes the original values of pixels, but also leads to information loss. Thus, the correlations among DCT coefficients are changed as well. To evaluate the robustness performance of the proposed approach, the original images in the test image database are compressed into JPEG images, in which the quality factors vary from 75 to 95 with a step size of 5. For the tampered images after object removal with different tampering ratios, they are saved into JPEG images with the same quality factors. Please note that those inpainted images with different tampering ratios are mixed together for performance evaluation in the experiments. Table 3 summarizes the detection results. From it, the proposed approach, which combines our earlier method [15] and JPDM-based features, achieves the highest detection accuracy among the three detectors. Moreover, with the decrease of image quality, there are significant decreases of detection accuracies for two existing approaches including Chang et al. [4]. and Yang et al. [15]. Especially, both can not achieve desirable detection accuracy when the quality factor of JPEG compression is 75. The reason behind this is that JPEG compression alleviates the similar abnormalities between block pairs, which destroys the clues for patch matching-based approaches. However, the correlations among DCT coefficients, which are located between the inpainted region and its neighboring region, are further distorted. This increases the detection accuracies of the proposed approach by introducing the JPDM-based features. That is, the mixture

Table 3 Performance comparisons among three detectors for object removal together with JPEG compression(%)

QF	Chang's [4]	Yang's [15]	JPDMs	Proposed method
95	91.06	88.32	85.06	93.03
90	82.43	80.33	83.79	91.81
85	79.18	78.41	89.50	90.73
80	65.02	65.30	87.31	89.59
75	50.48	50.01	86.07	87.25
Average	73.63	72.47	86.34	90.48

Table 4 Comparisons among three detectors for object removal with Gaussian noise addition(%)

SNR	Chang's [4]	Yang's [15]	JPDMs	Proposed method
40 dB	92.16	89.73	86.77	93.31
35 dB	84.63	82.13	88.46	91.91
30 dB	75.18	74.66	90.35	91.55
Average	83.99	82.17	88.52	92.25

of JPEG compression artifacts and the artifacts left by object removal make the proposed method achieve the best detection accuracies.

4.4.2 Detection for object removal integrated with Gaussian noise addition

Sophisticated counterfeiters, who have background knowledge about both image forgery and image forensics, may attempt to conceal the forgery traces by adding Gaussian noise. Thus, it is essential to validate the robustness of the proposed forgery detection algorithms against adding Gaussian noise. In the experiments, the forged images after object removal are added with Gaussian noise. The SNRs of the forged images after adding Gaussian noise are kept with 30dB, 35dB and 40dB, respectively. Table 4 reports the detection results of three detectors. It is apparent that with the decrease of SNRs, there are significant performance degradations for two existing approaches including Chang et al. [4]. and Yang et al. [15], while the proposed approach gradually improves its detection accuracies. When the SNRs of the forged image are less than 35dB, the JPDM-based features achieve better accuracies than those of two existing approaches [4, 15]. The reason behind this is that adding Gaussian noise to the forged images is actually another image manipulation, which further modifies the pixels in the forged images. This inevitably destroys the similar block pairs between inpainted area and the rest areas, yet also simultaneously disturbs the correlation among neighboring pixels. That is, the changes of correlation among neighboring pixels caused by Gaussian noise are mixed with the traces left by object removal using exemplar-based inpainting. This leads to more significant changes to the correlation among neighboring pixels. Therefore, the proposed method obtains desirable detection results.

4.4.3 Detection for object removal integrated with Gaussian blurring

It is claimed that the patch-based hole filling mechanism in exemplar-based image inpainting introduces blocking artifacts around the holes to be filled. To obtain a visually plausible

Table 5 Performance comparisons among three detectors for object removal together with Gaussian blurring(%)

Standard deviation	Chang's [4]	Yang's [15]	JPDMs	Proposed method
0.3	93.56	94.48	85.04	95.18
0.4	86.33	86.92	87.45	94.35
0.5	81.77	81.83	89.36	94.06
0.6	74.26	75.00	90.86	93.80
0.7	68.02	69.38	92.75	93.71
Average	80.79	81.52	89.09	94.22

forged image, counterfeiters might conduct Gaussian blurring as post-processing to conceal this kind of inconsistency in textural regions. However, Gaussian blurring also modifies the original pixels in the forged image after object removal, which might destroy the similarity between tampering region and its referenced patches. Thus, Gaussian blurring should be considered when evaluating the robustness of the proposed forgery detection approach. In the experiments, the forged images after object removal are further processed with Gaussian blurring, in which the standard deviation varies from 0.3 to 0.7. Table 5 compares the detection accuracies among the three detectors. It can be observed that with the increase of standard deviation of Gaussian blurring, the two existing approaches including Chang et al. [4] and Yang et al. [15] degrades their accuracies. However, the proposed approach increases its detection accuracy when the standard deviation of Gaussian blurring increases. This implies that Gaussian blurring strengthens the changes of correlations among neighboring pixels between the inpainted region and its adjacent region for object removal. Meanwhile, our proposed approach outperforms individual the JPDM-based features as well.

From Tables 2–5, the proposed hybrid approach achieves better detection accuracies than Chang's [4] and Yang's [15] approaches and individual JPDM-based features. The reasons behind this are summarized as follows: First, when the tampered images are less affected by the post-processing operations (e.g., the tampered ratio is small, or the QF for JPEG re-compression is large, etc), there are less modified pixels involved in image tampering. Thus, the consistencies among adjacent pixels will be less destroyed. It will be more difficult to differentiate the forged images from the original ones because the JPDM-based features are extracted from the difference 2D array by exploiting the joint probability density. Meanwhile, the similarities between the forged patches and their referenced ones still keep relatively high. That is, since the existing approaches are specifically designed for the detection of exemplar-based image inpainting, they are still effective for blind detection and higher detection accuracies can be achieved. Actually, the main motivation of the proposed approach, which introduces the JPDM-based features into our earlier method [15], is to further judge those undetected images to improve the detection accuracy. Second, when the tampered images are highly affected by post-processing (e.g., large standard deviation of Gaussian blurring, small QFs of JPEG compression, or low SNR values of Gaussian noise, etc), the situation will be contrary to the afore-mentioned phenomena. Meanwhile, the JPDM-based features are motivated by the literatures [5, 10, 12, 19, 25], so as to to improve the detection accuracy of our recent method [15] and especially its robustness to post-processing.

4.5 Identification between original images with post-processing and inpainted images

Since both post-processing operation and exemplar-based image inpainting have influences on the correlations among adjacent pixels, we also conduct an additional experiment to identify between original images with post-processing and inpainted images. The 1438 original

Table 6 Performance comparisons among four detectors for original images with post-processing and inpainted iamges(%)

Post-processing operation	Chang's [4]	Yang's [15]	JPDMs	Proposed method
JPEG compression	82.26	81.48	90.18	92.26
Gaussian noise	89.36	88.26	94.62	96.36
Gaussian blur	83.07	84.34	91.24	95.18

Table 7 Performance comparisons among four detectors for original images with post-processing and inpainted images with post-processing(%)

Post-processing operation	Chang's [4]	Yang's [15]	JPDMs	Proposed method
JPEG compression	81.56	83.84	84.04	85.32
Gaussian noise	86.53	87.29	88.84	90.72
Gaussian blur	83.87	85.38	87.76	89.61

images are post-processed with JPEG compression, adding Gaussian noise and Gaussian Blurring, respectively. The quality factor for JPEG compression is 85, the SNRs of the forged images after adding Gaussian noise are 35dB, and the standard deviations of Gaussian blurring are 0.5. Then, these images are mixed with the corresponding inpainted images with different tampering ratios as experimental dataset. Table 6 compares the detection accuracies among four detectors. From it,we can find that the proposed approach achieve the highest detection accuracy among the four detectors. That is,the proposed approach can also correctly identify the original images with post-processing and inpainted images. The reasons behind this are two-folds. Firstly, the average modification ratios are different for these two types of operations, as shown in Table 1. This leads to different JPDMs. Secondly, the proposed approach includes our earlier work [15], which exploits the similar block pairs between inpainted region and the rest regions, and thus can efficiently judge inpainted images without post-processing.

4.6 Identification between original images with post-processing and inpainted images with post-processing

In this experiment, we further discriminate between original images with post-processing and the inpainted images with post-processing as well. Both the original images and the inpainted images are post-processed with JPEG compression, adding Gaussian noise and Gaussian Blurring, respectively. The quality factor for JPEG compression is 85, the SNRs of the forged images after adding Gaussian noise are 35dB, and the standard deviations of Gaussian blurring are 0.5. The original images and the inpainted images after post-processing are mixed together as experimental dataset. Table 7 summarizes the experimental results among the four detectors. From it, we observe that the proposed approach achieves the best detection accuracy among the four detectors, though the proposed detection accuracies in this situation are slightly less than results reported in Table 6. The reason behind this is that the artifacts left by object removal is more or less alleviated by post-processing.

5 Conclusions

In the paper, a robust forgery detection algorithm is proposed for object removal by exemplar-based image inpainting, especially when it is further suffered from some common post-processing. Since post-processing disturbs the correlations among adjacent pixels to some extent, we are inspired to exploit JPDM to characterize such correlation. Specifically, a hybrid detection strategy is presented by introducing the JPDM-based features into our earlier method [15]. JPDM is computed for each difference JPEG 2-D array to model the intra-block correlation, and the averaged JPDMs for those difference mode 2-D arrays is computed to model the inter-block correlation. All the elements of these matrices are used

as feature vectors, which are input into the ensemble classifier for forgery detection. Experimental results show that with the help of the JPDM-based features, the proposed approach can effectively detect object removal by exemplar-based inpainting either with or without post-processing. However, there are still some limitations in this work. First, the proposed approach provides only a binary judgement towards whether a candidate image is forged or not. It will be better if it can locate the forged areas of object removal. A possible solution is to further exploit the structural inconsistency and geometric discrepancy of the forged image after image inpainting. Second, though the proposed approach is robust to some post-processing, we do not differentiate the ambiguous processing artifacts left by examplar-based inpainting and post-processing. In future research, we will attempt to estimate the types of image forgeries by separating the intrinsic fingerprints left by multiple forgeries and introducing multi-class classifier [7, 8, 22–24]. This is still an open issue of processing history estimation/recovery in the field of digital image forensics.

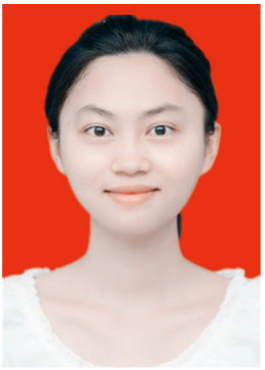
Acknowledgements We would like to thank the anonymous reviewers for their professional comments and valuable suggestions. This work is partially or fully sponsored by National Natural Science Foundation of China (61572183, 61379143), the Specialized Research Fund for the Doctoral Program of Higher Education (20120161110014), the Scientific Research Fund of Hunan Provincial Education Department of China (14C0029), Natural Science Foundation of Hunan Province (2016JJ2005). The authors appreciate the nice help from Mr Moses Odero for his improving the English usages.

References

1. Amerini I, Ballan L, Caldelli R et al (2013) Copy-move forgery detection and localization by means of robust clustering with J-linkage. Signal Process Image Commun 28(6):659–669
2. Bacchuwar KS, Ramakrishnan KR (2013) A jump patch-block match algorithm for multiple forgery detection. In: Proc. of IEEE international multi-conference on automation, computing, communication, control and compressed sensing (iMac4s), pp 723–728
3. Birajdar GK, Mankar VH (2013) Digital image forgery detection using passive techniques: A survey. Digit Investig 10(3):226–245
4. Chang I, Yu J, Chang CC (2013) A forgery detection algorithm for exemplar-based inpainting images using multi-region relation. Image Vis Comput 31(1):57–71
5. Cozzolino D, Gragnaniello D, Verdoliva L (2014) Image forgery detection through residual-based local descriptors and block-matching. In: IEEE International conference on image processing (ICIP), pp 5297–5301
6. Criminisi A, Prez P, Toyama K (2004) Region filling and object removal by exemplar-based image inpainting. IEEE Trans Image Process 13(9):1200–1212
7. Gu B, Sheng VS, Wang Z et al (2015) Incremental learning for v-support vector regression. Neural Netw 67:140–150
8. Gu B, Sheng VS, Tay KY et al (2015) Incremental support vector learning for ordinal regression. IEEE Trans Neural Netw Learn Syst 26(7):1403–1416
9. Guillemot C, Meur OL (2014) Image inpainting: overview and recent advances. IEEE Signal Process Mag 31(1):127–144
10. He Z, Lu W, Sun W et al (2012) Digital image splicing detection based on Markov features in DCT and DWT domain. Pattern Recog 45(12):4292–4299
11. Kakar P, Sudha N (2012) Exposing postprocessed copy-paste forgeries through transform-invariant features. IEEE Trans Inf Forens Secur 7(3):1018–1028
12. Kirchner M, Fridrich J (2010) On detection of median filtering in digital images. SPIE Electron Imag Int Soc Opt Photon 754110–754110
13. Kodovsky J, Fridric J, Holub V (2012) Ensemble classifiers for steganalysis of digital media. IEEE Trans Inf Forens Secur 7(2):432C444
14. Li J, Li XL, Yang B, Sun XM (2015) Segmentation-based image copy-move forgery detection scheme. IEEE Trans Inf Forens Secur 10(3):507–518
15. Liang Z, Yang G, Ding X et al (2015) An efficient forgery detection algorithm for object removal by exemplar-based image inpainting. J Vis Commun Image Represent 30:75–85

16. Liu Q, Chen Z (2014) Improved approaches with calibrated neighboring joint density to steganalysis and seam-carved forgery detection in JPEG images. ACM Trans Intell Syst Technol 5(4):63
17. Muhammad G, Hussain M, Bebis G (2012) Passive copy move image forgery detection using undecimated dyadic wavelet transform. Digit Investig 9(1):49–57
18. Stamm MC, Wu M, Liu KJR (2013) Information forensics: an overview of the first decade. IEEE Access 1:167–200
19. Shi YQ, Chen C, Xuan G et al (2007) Steganalysis versus splicing detection. Int Workshop Digit Watermark 158–172
20. Wang J, Lu K, Pan D et al (2014) Robust object removal with an exemplar-based image inpainting approach. Neurocomputing 123:150–155
21. Wu Q, Sun SJ, Zhu W et al (2008) Detection of digital doctoring in exemplar-based inpainted images. In: Proc. of IEEE international conference on machine learning and cybernetics, vol 3, pp 1222–1226
22. Xia ZH, Wang XH, Sun XM, Wang BW (2014) Steganalysis of least significant bit matching using multi-order differences. Secur Commun Netw 7(8):1283–1291
23. Xia Z, Wang X, Sun X et al (2016) Steganalysis of LSB matching using differences between nonadjacent pixels. Multimed Tools Appl 75(4):1947–1962
24. Xia Z, Wang X, Zhang L et al (2016) A privacy-preserving and copy-deterrence content-based image retrieval scheme in cloud computing. IEEE Trans Inf Forens Secur 11(11):2594–2608
25. Zhao X, Wang S, Li S et al (2013) Image splicing detection based on noncausal markov model. In: IEEE International conference on image processing, pp 4462–4466

Dengyong Zhang is a PhD student in Hunan University, China. He is also a lecture in the Changsha University of Science and Technology, China. His research interests is digital media forensics and video passive forensics.



Zaoshan Liang is a master student in Hunan University, China. His research interests is digital media forensics.

Gaobo Yang is a professor in Hunan University, China. He is also a key member of Hunan Provincial Key Laboratory of Networks and Information Security. He received the Ph.D. degree in Communication and Information System from Shanghai University in 2004. He is the PI of several projects such as Natural Science Foundation of China (NSFC), Special Pro-phase Project on National Basic Research Program of China (973) and program for New Century Excellent Talents (NCET) in university. Currently, his research interests are in the area of image and video signal processing, digital media forensics.

Qingguo Li is a professor in Hunan University, China. He received the Ph.D. degree in Mathematics from Hunan University in 1997. Currently, his research interests are in the area of Topology and signal processing.

Leida Li is an associate professor in the School of Information and Electrical Engineering, China University of Mining and Technology. He obtained his PhD degree from Xidian University, China in 2010. During Jan 2014 to Jan 2015, he made his academic visit to Nanyang University of Technology, Singapore. His research interests include image quality assessment and image forensics.

Xingming Sun is currently a Professor with the School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing, China. He received the B.S. degree in mathematics from Hunan Normal University, Hunan, China, in 1984, the M.S. degree in computing science from the Dalian University of Science and Technology, Dalian, China, in 1988, and the Ph.D. degree in computer science from Fudan University, Shanghai, China, in 2001. His research interests include network and information security, digital watermarking, cloud computing security, and wireless network security.