Multimedia Tools and Applications
https://doi.org/10.1007/s11042-019-08169-w

Copy-move forgery detection based on adaptive keypoints extraction and matching

Hong-Ying Yang[1] · Shu-Ren Qi[1] · Ying Niu[1] · Pan-Pan Niu[1] · Xiang-Yang Wang[1]

Received: 18 January 2019 / Revised: 15 July 2019 / Accepted: 2 September 2019
Published online: 14 September 2019

Abstract

Copy-move (region duplication) is one of the most common types of image forgeries, in which at least one part of an image is copied and pasted onto another area of the same image. The main aims of the copy-move forgery are to overemphasize a concept or conceal objects by duplicating some regions. Keypoint-based copy-move forgery detection (CMFD) schemes extract image keypoints and employ local image features to identify duplicated regions, which exhibits remarkable detection performance with respect to memory requirement, computational cost, and robustness. To enhance the performance of keypoint-based CMFD approaches, here are three issues that need to be solved: the non-uniform distribution of image keypoints, the low discriminatory power of local image descriptor, and the high computational cost and low matching efficiency of feature matching strategy. In order to overcome these issues, we propose a new copy-move forgery detection method based on adaptive keypoints extraction and matching in this paper. First, we extract the image keypoints using the adaptive uniform distribution threshold. Second, the binary robust invariant scalable keypoints (BRISK) descriptor is introduced to represent the local image feature of image keypoints. Afterwards, local BRISK descriptors are employed to match image keypoints by using embedded random ferns approach, which formulates the required matching as a discriminative classification problem. Finally, the falsely matched keypoints pairs are eliminated by utilizing the random sample consensus (RANSAC), and the fast mean-residual normalized intensity correlation (NNPROD) is employed to locate the tampering area. We evaluate the performance of the proposed CMFD method in detail by conducting several simulation experiments, and the experimental results have shown that the detection and localization accuracy of the proposed method is superior to that of the state-of-the-art approaches recently proposed in the literature, even in adverse conditions.

Keywords Copy-move forgery detection · Adaptive keypoints · BRISK descriptor · Embedded random ferns · Fast NNPROD

✉ Pan-Pan Niu
niupanpan3333@163.com

✉ Xiang-Yang Wang
wxy37@126.com

[1] School of Computer and Information Technology, Liaoning Normal University, Dalian 116029, People's Republic of China

1 Introduction

Today, digital image has been widely used in people's daily life and work. However, with the rapid development of computer network technology and portable digital equipment, the modification of digital image has become simpler and easier. If the forged image is applied in court, news report and scientific paper, this may bring a serious threat to political system and social stability. Therefore, it is very urgent to search the new technology, which can identify the authentic of digital image content. To solve this problem, many researchers have begun to pay close attention to the digital forged image and propose various solutions [18]. The existing methods of image forgery detection can be roughly divided into two categories: active and passive (blind). The active detection methods such as digital signature [45] and digital watermark [3] rely on the prior information of images. Because the prior information of test images are unavailable in some cases and the passive detection approaches may be applied to identify the authenticity of images, which made the passive approach become a research hotspot.

There are currently two main types of digital image forgeries: copy-move forgery [37] and splicing [43, 44]. Among them, copy-move forgery is one of the most common types of image forgeries, where a forger often copied one region of an image and then pasted it into another region of the original image. The purposes of copy-move forgery are to cover a particular object, or over-emphasize an object by copying certain areas. Here, in order to tamper the image without leaving obvious clues for human eyes, a simple modification may not be enough in some cases. The forger usually does a series of actions such as JPEG compression, rotation, and scaling for the parts to be copied. The rapid growth of image processing software has led to a large number of copy-move tampering images without obvious traces, making copy-move forgery detection (CMFD) the most important and hottest digital image authentication technology at present. In recent years, many passive forgery detection methods have been put forward, which can be divided into two major types: block-based methods and keypoint-based methods [8]. Both of them use feature matching approaches to look for similar areas in the copy-move forgery detection methods.

The block-based methods typically divide the images to be inspected into overlapping and fixed-size image blocks, and then every image block can be characterized by using different feature algorithms to enhance the robustness to JPEG compression, additive noise, rotation and scaling. However, there still remain some troubles with the block-based methods, such as low discriminative power of feature descriptors, poor affine transformation invariance and high computational cost [5, 11, 27, 35]. The keypoint-based methods were proposed to replace these block-based approaches to effectively overcome these troubles [28]. Firstly, the keypoint-based methods identify and select the keypoints from high contrast image areas and match the feature vectors of these keypoints. And then the false matched pairs are removed by post-processing operation. Finally, the forged areas are located utilizing the area correlation maps. The keypoint-based methods can get a perfect detection performance, especially in the non-smooth forged areas. Because the number of keypoints represents just a relatively small set of all the pixels in the test image, the computational complexity of these keypoint-based methods is comparatively lower. However, we shall never overlook the downsides of the keypoint-based methods, most of them are intrinsically less accurate than the block-based methods, especially when copy-move forgery includes smooth areas, and they are usually not robust to many post-processing operations.

To enhance the performance of keypoint-based methods when the copy-move forgery includes smooth areas and various attacks, here are three major issues that need to be solved.

(1) The non-uniform distribution density of keypoints will lead to issues of insufficient or even none keypoints in the smooth and small areas of test image, where the original and duplicated regions may be located. That is the reason why the detection performances of keypoint-based methods are failing when the copy-move forgery includes smooth areas. (2) The discriminative power of feature descriptors is low, which will be detrimental to the matching performance. (3) The high computational cost and low match performance of feature matching algorithm. When we concentrate on the detection of smooth forged areas, it is bound to cause an increase in the number of keypoints in the test image. An efficient matching algorithm not only can reduce the matching time, but also have an accurate match performance, which are useful for improving the efficiency of the error matched pairs removed.

In order to overcome the above issues, we propose a new copy-move forgery detection method based on adaptive keypoints extraction and matching in this paper. The novelty of the proposed method lies in the following aspects. (1) We design a simple yet effective way to ensure that the keypoints are uniformly distributed in the image, even in smooth and small regions, by using the adaptive threshold setting and block-based homogenization processing. (2) The binary robust invariant scalable keypoints (BRISK) descriptor is introduced for robust and discriminative keypoints representation; moreover, the obtained binary features can be quickly matched compared to the classical double-precision features. (3) As an improved method of KD-tree and random forests, embedded random ferns approach is employed to quickly and accurately match keypoints.

The remaining part of this paper is organized as follows. Section 2 briefly introduces related CMFD works. In Section 3, the proposed keypoint-based method is introduced in detail. And the simulation experimental results are presented in Section 4 to show the excellent performance of the proposed method. Finally, the conclusions of this paper are given in Section 5.

2 Related work

Over the past decades, many methods have been proposed for the passive copy-move forgery detection. From the statement above, we can know that the existing methods for copy-move forgery detection can be roughly categorized into two major classes: block-based techniques [7, 9, 10, 13, 26, 32, 39, 40, 47] and keypoint-based techniques [1, 2, 14–17, 19, 22, 24, 25, 30, 31, 33, 34, 38, 41, 42, 46].

2.1 Block-based techniques

The problem of passive copy-move forgery was firstly investigated by Fridrich et al. [13], who presented the block-based technique. First, the input image was divided into overlapping blocks with the same size. Then the discrete cosine transform (DCT) was used to indicate image blocks features in low dimensions. And they applied the lexicographic sorting algorithm to solve the problem of computational cost in the matching stage. Afterwards, a lot of robust block-based detection algorithms are proposed based on this work. The main difference between them is the way they represent blocks features, such as frequency domain-based method, the intensity-based method and moment-based method [8]. Various well-known techniques have been employed for this task, e.g., Discrete Wavelet Transform (DWT) [26], Fourier Transform (FT) [7], Singular Value Decomposition (SVD) [47] and Zernike moments (ZM) [32]. The results show that these feature representation methods are mainly robust to the

operations such as JPEG compression and additive white Gaussian noise. Such features, however, do not perform well in the presence of rescaling and rotation. Moreover, a common weakness of the above CMFD methods is the relatively high computational complexity because almost all pixels need to be matched. To overcome this problem, Cozzolino et al. [9] presented an efficient features matching method for CMFD, in which the the modified PatchMatch algorithm is used to compute efficiently a high-quality approximate nearest neighbor field. Lately, Cozzolino et al. [10] further improved their previous work. More rencently, Wang et al. [39, 40] proposed another strategy to reduce the complexity of matching. The features are grouped by a package clustering algorithm and then are matched separately in each group.

Although these proposed methods can significantly enhance the detection performance, there are still some problems like as high computational complexity and poor robustness to rescaling and rotation, which are the main reasons why these methods cannot get a perfect detection performance.

2.2 Keypoint-based techniques

Keypoint-based methods are more interesting alternative paths to solve the above problems. Interest points are firstly extracted in the host image; then, the keypoint features are matched for detecting the potential tamperings. Some recent works have shown that the CMFD on basic of Scale Invariant Features Transform (SIFT) has more applications. The SIFT detector was first employed in copy-move tampering detection by Huang et al. [15]. Following this path, a number of SIFT-based algorithms have appeared in the literature. Roughly speaking, these approaches differ mostly in the kind of matching algorithm taken into consideration and in the post-processing strategy used. Pan et al. [30] utilized the Best-Bin-First (BBF) approach to match the detected SIFT keypoints based on their features vectors. Amerini et al. [1] presented a new SIFT-based method, in which a generalized 2-nearest neighbor (g2NN) matching algorithm was proposed for dealing multiple duplicated areas. However, this method does not handle the forgery localization problem. Amerini et al. [2] also proposed a technique based on SIFT features, where a new clustering procedure, named J-Linkage clustering, was introduced in the post-processing. Li et al. [22] developed a new CMFD approach based on adaptive segmentation and SIFT feature point matching, which integrated both the traditional block-based CMFD methods and keypoint-based ones. Then, Pun et al. [31] proposed a somewhat similar scheme. Li et al. [24] clustered SIFT features based on the scale and gray level information of the keypoints, and then the hierarchical matching algorithm is performed efficiently.

Although such SIFT-based algorithms are popular, there are three main drawbacks: (1) there are not enough keypoints in small or smooth regions, causing detection failure; (2) descriptors are difficult to obtain the favorable robustness and discriminability performance, which may result in a significant number of mismatched pairs, especially on self-similar images; (3) they are time-consuming in the feature extraction and matching phases, and the matching accuracy still needs to be improved.

In the feature extraction stage, the existing improvements on speed and descriptive power are mainly reflected in three aspects. First, using different kinds of keypoints, like as Speed up Robust Feature (SURF) [25, 33, 34], Harris [17] and KAZE [41]. Second, different kind of keypoint detectors and feature descriptors have been combined in some methods, such as combining the SURF with BRISK descriptor [19] and combining a corner detector with Polar Cosine Transform (PCT) [46]. Third, the keypoints and their features vectors are selected in

different color spaces, for example, the technique of selecting SURF features in the reverse color space [14]. For smooth tampered regions, Wang et al. [38] presented a new keypoint-based forgery detection method. They first segmented original tampered images into non-overlapping and not fixed-size superpixels, which are furthly classified into smooth areas, texture areas and strong texture areas according to their local information entropy. Then, the robust keypoints were selected corresponding to different information entropy classification. Jin et al. [16] introduced the non-maximum suppression algorithm to ensure that the keypoints are uniformly distributed in the image. Yang et al. [42] developed a new keypoint-based CMFD method, in which a keypoints distribution strategy was established to insert the keypoints evenly throughout the test image. Also, a modified SIFT descriptor was applied to describe the keypoints. Li et al. [24] proposed a simple solution via reducing the threshold of SIFT and resizing the image. However, too many feature points in the texture area cause trouble for matching.

In the feature matching stage, most CMFD methods [1, 2, 15, 22, 24, 30, 31] utilize Euclidean distance to measure the correlation between the features. However, calculating the Euclidean distance of a large number of double-precision high-dimensional features requires expensive computational costs. Therefore, some methods based on Hamming distance and cosine distance have been proposed to reduce complexity. Kumar et al. [19] proposed a hybrid approach based on keypoints for CMFD. Since the BRISK features were applied to describe these keypoints detected by SURF algorithm. Therefore, the binary features were efficiently matched by using Hamming distance. Sachdev et al. [33] proposed a new method of copy-move forgery detection on the basic of SURF keypoints and SLIC segmentation, and they used the cosine similarity to calculate the distance between two feature vectors, which is a more efficient matching algorithm. In addition to distance metrics, for reducing the time complexity of K-Nearest Neighbors (KNN) search, the BBF KD-tree [15, 22, 30] is often adopted to replace the brute force and sorting methods. The KD-tree algorithm is very efficient in low-dimensional space, but its performance in high-dimensional space will drop rapidly. Although the BBF strategy makes the KD-tree algorithm possible for high-dimensional KNN search, the nearest neighbors are usually approximated in order to meet the speed requirements. This is not conducive to robust matching of keypoints when the image is attacked [29].

The research emphasis of keypoint-based CMFD has changed from the textural tampered areas into the smooth tampered areas. Therefore, it is urgent to obtain the keypoints with uniform distribution in the test image, which is beneficial for the detection performance and the time complexity of features matching stage. In the next section, the framework of the proposed keypoint-based CMFD method is introduced to solve these existing problems.

3 The proposed CMFD method

Keypoint-based detection methods have been reported to be very effective in revealing copy-move evidences, due to their robustness against various attacks, such as large-scale geometric transformations. However, these methods fail to handle the cases when copy-move forgeries only involve small or smooth regions, where the number of keypoints is very limited. In this paper, we propose a new copy-move forgery detection method based on adaptive keypoints extraction and matching, as shown in Fig. 1.

Firstly, the image keypoints are extracted based on the adaptive detection threshold selection and block-based homogenization processing. By this step, we can solve the problem

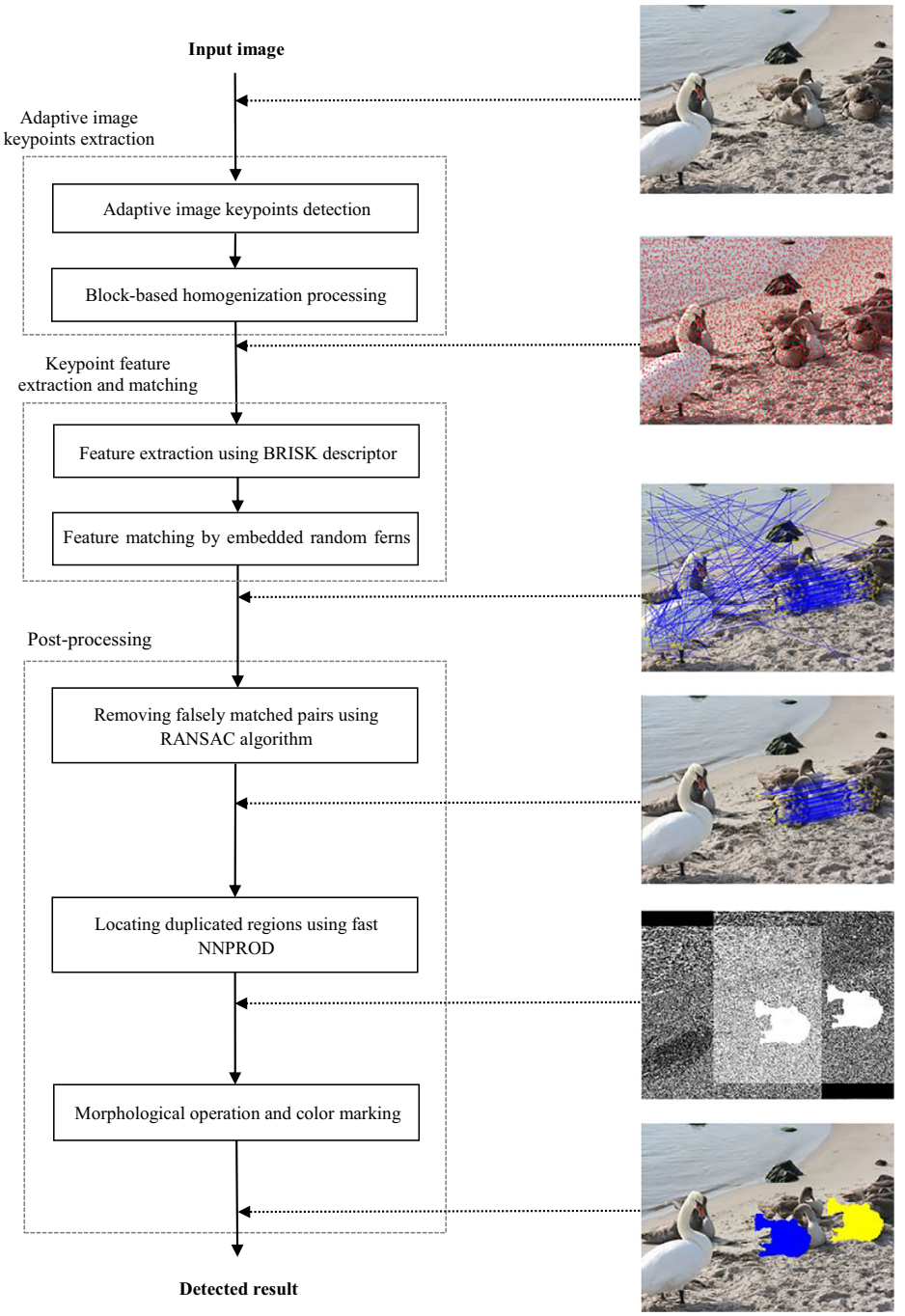


Fig. 1 Proposed Framework for keypoint-based copy-move forgery detection

that the image keypoints cannot be extracted sufficiently from small smooth regions. Then, the BRISK descriptor is introduced to represent the local image feature of image keypoints. By

this step, we can enhance the discriminative power and robustness of image keypoints features. Next, the embedded random ferns approach is employed to match image keypoints by using local BRISK descriptor. By this step, we can achieve higher matching accuracy than the approximate nearest neighbor searches, such as BBF KD-tree and PatchMatch. And finally, the falsely matched keypoints pairs are eliminated by utilizing the random sample consensus (RANSAC), and the fast mean-residual normalized intensity correlation (NNPROD) is employed to locate the tampering area. By this step, we can effectively improve the detection accuracy and reduce the detection time.

3.1 Adaptive image keypoints extraction

While maintaining the excellent properties of SIFT algorithm, the SURF algorithm [4] solves the shortcoming of high computational complexity and improves the aspects of keypoints extraction and the features vectors description. In this paper, we apply SURF detection to extract image keypoints. To ensure that the keypoints are uniformly distributed in the image, we propose an adaptive image keypoints detection approach based on adaptive detection threshold selection and homogenization processing. The details of our adaptive image keypoints detection are illustrated in Algorithm 1.

Algorithm 1 The adaptive image keypoints detection algorithm

Input: The input image

Output: The image keypoints

STEP-1: Initialize parameters.

STEP-2: Partition the input image into sub-blocks.

STEP-3: Detect SURF keypoints initially from the input image.

STEP-4: Detect SURF keypoints adaptively from the input image.

STEP-5: Homogenize processing on each sub-block.

3.1.1 Initialize parameters

We initialize the parameters of the adaptive image keypoints detection algorithm, such as input image size $M_I \times N_I$, image sub-block size $n \times n$, initial keypoints detection threshold δ_0, uniformity measurement threshold ζ of image keypoints, and denseness threshold ψ of image sub-block keypoints.

3.1.2 Partition the input image into sub-blocks

We partition the input image I into non-overlapping image sub-blocks $\mathbf{b_{ij}}$ with the size $n \times n$. Given an $M_I \times N_I$ image for example, we denote the non-overlapping sub-blocks with the size $n \times n$ as $\mathbf{b_{ij}}$ $(1 \le i \le \lceil M_I/n \rceil, 1 \le j \le \lceil N_I/n \rceil)$. As a result, the number of image sub-blocks is $\lceil M_I/n \rceil \times \lceil N_I/n \rceil$.

3.1.3 Detect SURF Keypoints initially from the input image

We detect initially the image keypoints from the input image I by employing SURF algorithm with initial keypoints detection threshold $\delta_0 = 10$.

3.1.4 Detect SURF keypoints adaptively from the input image

We detect adaptively the image keypoints from the input image I by employing SURF algorithm with the adaptive keypoints detection threshold δ.

For the traditional CMFD algorithm, a fixed keypoint threshold is usually selected to remove the keypoints with low contrast. As the forgery area may locate in smooth areas, which may have a negative impact on the final result of image forgery detection, so the extraction of low contrast keypoints is particularly important for keypoints-based CMFD scheme. Here, we propose an adaptive iterative threshold selection method for the SURF keypoints detection. Its flowchart is shown in Fig. 2. Where, keypoints uniformity measurement (KUM) value Φ is developed for evaluates the distribution level. If the KUM value Φ of the image keypoints is less than ζ, it means that the distribution of image keypoints is good. The detection threshold δ is adjusted by KUM values to adaptively determine the detection threshold of the SURF detector. In this paper, the uniformity measurement threshold ζ is set to 0.5. The value of KUM is calculated by the following formula

$$\Phi = \sqrt{\sum_{i=1}^{\lceil M_I/n \rceil} \sum_{j=1}^{\lceil N_I/n \rceil} \left(NoK_{ij} - \phi\right)^2 / \sum_{(i,j)} NoK_{ij}} \tag{1}$$

Where, NoK_{ij} represent the number of keypoints in image sub-block $\mathbf{b_{ij}}$, and standard keypoints number ϕ defined as

$$\phi = \frac{\sum_{(i,j)} NoK_{ij}}{\lceil M_I/n \rceil \times \lceil N_I/n \rceil} \tag{2}$$

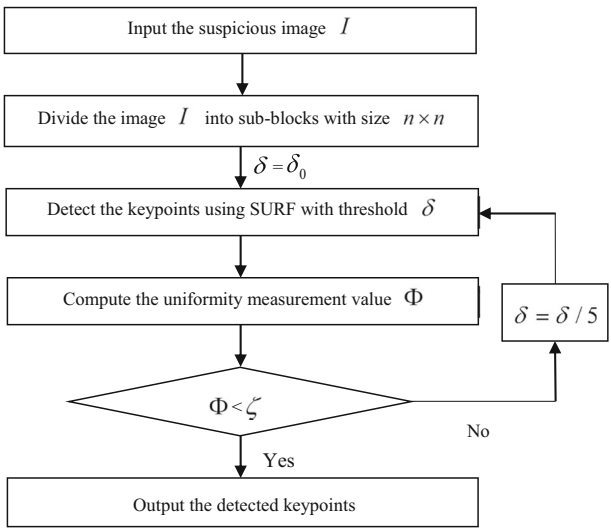


Fig. 2 The flowchart of adaptive iterative detection threshold selection for SURF algorithm

3.1.5 Homogenize processing on each sub-block

We perform homogenize processing on each non-overlapping image sub-blocks $\mathbf{b_{ij}}$. For each image sub-block $\mathbf{b_{ij}}$, if its keypoints number is larger than the denseness $\psi = 1\%$, the redundant weak keypoints are removed. Finally, we get the SURF keypoints set $\mathbf{K} = \{\mathbf{k_i} = (x, y, s)_i\}$, where (x, y) represents the coordinates and s represents the scale.

Figure 3 shows the adaptive image keypoints detection process.

3.2 Keypoint feature extraction and matching

3.2.1 Keypoint feature extraction

Accurate feature description is important for capturing the subtle image details for identifying the copy-move regions. To decrease the time consumption of feature representation and improve the efficiency of feature matching phase, we extract the binary robust invariant scalable keypoints (BRISK) [21] features in a neighbourhood of the detected adaptive image keypoints. BRISK is an effective texture descriptor. Besides having low computational complexity, BRISK achieves orientation-invariance and scale-invariance. BRISK descriptor is composed as a binary string by concatenating the results of simple brightness comparison tests. BRISK feature descriptors can be generated as follows [21].

Gaussian filtering The core idea of the BRISK descriptors is to utilize a pattern used for sampling the neighborhood of the image keypoint. The concentric circles with different radii are constructed with a keypoint **k** as the center, and a certain number of equal interval sampling points are obtained on each circle. Set *NoS* is the number of all sampling points. In order to avoid the aliasing effects caused by this neighborhood sampling pattern, we need to do Gaussian

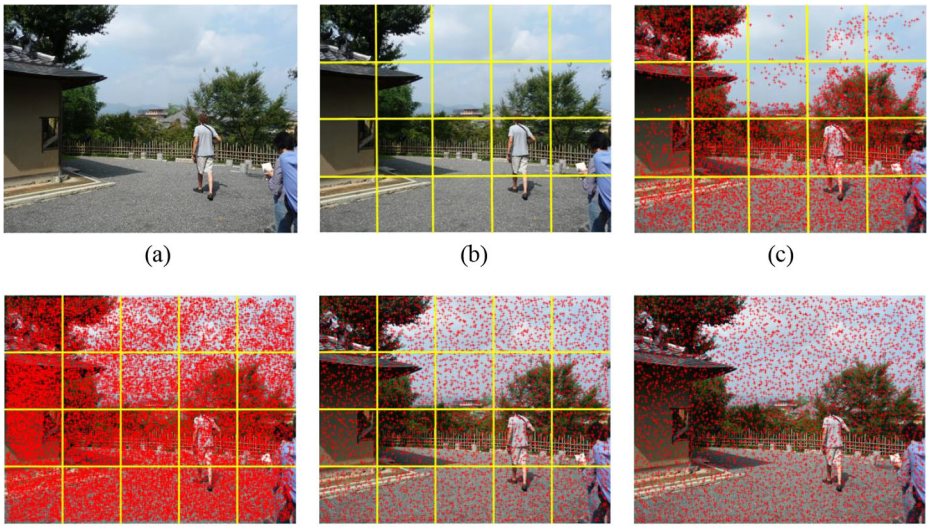


Fig. 3 The adaptive image keypoints detection process: **a** The input image, **b** Partition the input image, **c** The initial image keypoints detection, **d** The adaptive image keypoints detection, **e** The image sub-blocks homogenize processing, **f** The adaptive image keypoints detection results

filtering on the sampling points. The pattern is illustrated in Fig. 4, the blue circles denote the sampling locations. The red dashed circles are drawn at a radius σ corresponding to the standard deviation of the Gaussian kernel used to smooth the intensity values at the sampling points.

3.2.2 Local gradient calculation

For a total of *NoS* sampling points, the $NoS(NoS-1)/2$ sampling-points pairs $(\mathbf{sp_i}, \mathbf{sp_j})$ should be considered. The $I(\mathbf{sp_i}, \sigma_i)$ and $I(\mathbf{sp_j}, \sigma_j)$ respectively present the gray values of sampling points after Gaussian filtering, which are applied to calculate the local gradient $\mathbf{g}(\mathbf{sp_i}, \mathbf{sp_j})$

$$\mathbf{g}\left(\mathbf{sp_i}, \mathbf{sp_j}\right) = \left(\mathbf{sp_j} - \mathbf{sp_i}\right) \frac{I\left(\mathbf{sp_j}, \sigma_j\right) - I(\mathbf{sp_i}, \sigma_i)}{\left\|\mathbf{sp_j} - \mathbf{sp_i}\right\|^2} \tag{3}$$

Definiting the set $\mathbf{A}$ is all the sampling points pairs

$$\mathbf{A} = \left\{ \left(\mathbf{sp_i}, \mathbf{sp_j}\right) \in \mathbb{R}^2 \times \mathbb{R}^2 \mid i < NoS, j < i, \ \ i, j \in \mathbb{N} \right\} \tag{4}$$

We can divide the set $\mathbf{A}$ into a subclass with short distance pairs $\mathbf{S}$ and the other subclass with long distance pairs $\mathbf{L}$ based on the distance between two sampling points.

$$\mathbf{S} = \left\{ \left(\mathbf{sp_i}, \mathbf{sp_j}\right) \in \mathbf{A} \mid \left\|\mathbf{sp_i} - \mathbf{sp_j}\right\| < \theta_{\max} \right\} \subseteq \mathbf{A} \tag{5}$$

$$\mathbf{L} = \left\{ \left(\mathbf{sp_i}, \mathbf{sp_j}\right) \in \mathbf{A} \mid \left\|\mathbf{sp_i} - \mathbf{sp_j}\right\| > \theta_{\min} \right\} \subseteq \mathbf{A} \tag{6}$$

The threshold $\theta_{\max}$ and $\theta_{\min}$ are respectively set to $9.75s$ and $13.67s$, and here s presents the scale of keypoint $\mathbf{k}$. We use the above information to calculate the principal feature pattern directions of the keypoints (Note: this is only a long distance subclass $\mathbf{L}$), as follows

$$\begin{pmatrix} g_x \\ g_y \end{pmatrix} = \frac{1}{|\mathbf{L}|} \times \sum_{\left(\mathbf{sp_i}, \mathbf{sp_j}\right) \in \mathbf{L}} \mathbf{g}\left(\mathbf{sp_i}, \mathbf{sp_j}\right) \tag{7}$$

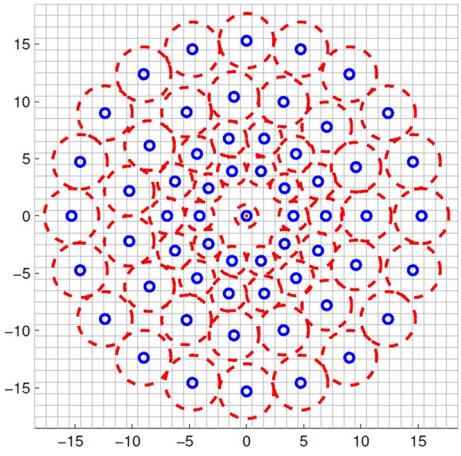


Fig. 4 The BRISK descriptors extraction of sampling points

Building the descriptors To solve the problem of rotation and scale invariance and form the rotation and scale normalized descriptors, BRISK uses the sampling pattern rotated by $\alpha = \arctan(g_y/g_x)$ around the keypoint $\mathbf{k}$. The bit-vector descriptor is assembled by performing all the short distance intensity comparisons of point pairs $\left(\mathbf{sp}_\mathbf{i}^\alpha, \mathbf{sp}_\mathbf{j}^\alpha\right)$ (i.e. in the rotated pattern), such that each bit b corresponds

$$b = \begin{cases} 1 & I\left(\mathbf{sp}_\mathbf{j}^\alpha, \sigma_j\right) > I\left(\mathbf{sp}_\mathbf{i}^\alpha, \sigma_i\right) \\ 0 & otherwise \end{cases} \qquad \forall\left(\mathbf{sp}_\mathbf{i}^\alpha, \mathbf{sp}_\mathbf{j}^\alpha\right) \in \mathbf{S} \tag{8}$$

From this we can obtain a binary string of length 512 bits. If one byte (i.e., 8 bits) is used to represent a feature element f, we can define the feature descriptor $\mathbf{f}$ of keypoint $\mathbf{k}$ as

$$\mathbf{f}(\mathbf{k}) = (f_1, f_2, \ldots, f_{63}, f_{64}) \in [0,\ \ 255]^{64} \tag{9}$$

Generally, BRISK descriptor not only has low computational complexity, but also achieves orientation-invariance and scale-invariance, so we introduce BRISK descriptor to depict image keypoint feature in this paper. Figure 5 shows the process of keypoints features extraction. To demonstrate the process, we select five keypoints as samples and illustrate the features descriptors of one of the keypoints, which is expressed in the form of histogram.

3.2.3 Keypoint matching

After the BRISK feature extraction, the CMFD method identifies some of the potential tampered pairs by finding the keypoints with similar features, and these keypoints pairs will be further processed. Nowadays, with the rapid development of electronic data information, fast similarity clustering and search have been considered as one of the most basic problems in multimedia platforms. The embedded random ferns approach [12] has been proposed to solve the problems due to their multilevel processing capabilities, and high efficiency. In this paper, we introduce the embedded random ferns approach to match image keypoints, which formulates the required matching as a discriminative classification problem.

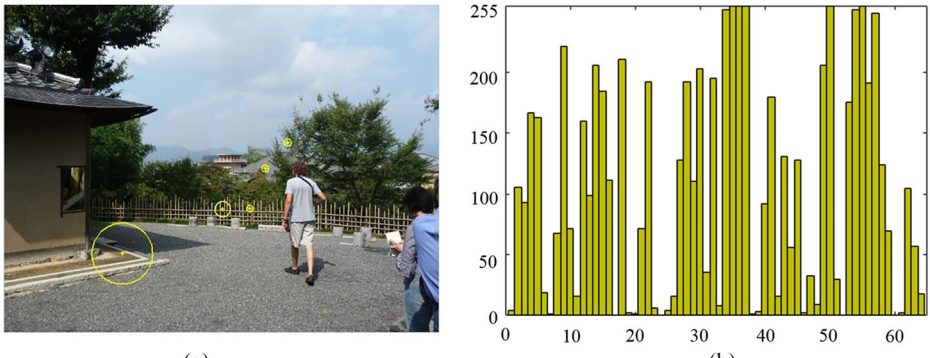


Fig. 5 Image keypoints features extraction: **a** Keypoints samples, **b** Feature descriptors for an image keypoint

Embedded random ferns In 2013, Lepetit et al. [12] developed the embedded random ferns method based on random ferns [6, 20, 36], which is a kind of machine learning classification approach. Generally, the embedded random ferns transforms the traditional matching problem into the classification problem, and its core idea is that employing a supervised dimensionality reduction to map randomly selected feature dimensions into a subspace. Embedded random ferns includes training and testing, and Fig. 6 shows the embedded random fern concept for a single fern [12].

The details about the embedded random ferns are expressed in Algorithm 2.

Algorithm 2 Embedded random ferns [43]

Sample Training:

STEP-1 Feature dimensions reduction. The $N \times D$ input data matrix $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N\}$ ($\mathbf{x}_i \in \mathbb{R}^D$) is reduced to a fern-specific $N \times B$ $(B \le D)$ matrix $\mathbf{X}^m = \{x_{id} \mid i \in \{1,2,\cdots,N\}, d \in \mathbf{S}^m\}$, where $\mathbf{S}^m = \{d_{1m}, d_{2m}, \cdots, d_{Sm}\}$ is a set of B randomly chosen feature IDs.

STEP-2 Associating dataset label. A dataset label matrix $\mathbf{Y} = \{\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_N\} \in \{0,1\}^{N \times C}$ is built, where the label $\mathbf{y}_i \in \{1,2,\cdots,C\}$ is associated to the corresponding input data $\mathbf{x}_i$, and the corresponding entry y_{il} is 1 if input data $\mathbf{x}_i$ belongs to class c and 0 otherwise.

STEP-3 Subspace projection. Based on the provided features $\mathbf{X}^m$ and the corresponding label $\mathbf{Y}$, the Canonical Correlation Analysis (CCA) is employed to produce a new embedding space $\mathbf{W}^m$, and the projection enables the assignment of each training sample to a bin $\mathbf{o}^m$. So, a set of random ferns are obtained by their projection matrices and the offset vectors ($\mathbf{W}^m, \mathbf{o}^m$).

Data Testing:

STEP-4 Mapping data points and computing the class-conditional probabilities. By employing the same feature dimension $\mathbf{X}^m$, the data points $\mathbf{X}$ are mapped to the different projection spaces per fern using ($\mathbf{W}^m, \mathbf{o}^m$), and the learned projection is applied to assign the sample to a bin (binary vector $\mathbf{b}^m$). Then, the binary vectors $\mathbf{b}^m$ are utilized to compute

Keypoint matching using embedded random ferns In this paper, we match the image keypoints by employing the BRISK feature descriptors and embedded random ferns classifier. First, for the keypoints feature set $\mathbf{F} = \{\mathbf{f}(\mathbf{k})\}$, a set of the embedded random fern classifiers with the number $T_{NoF} = 30$ and depth $T_{DoF} = 12$ is established, where each image keypoint corresponds to a fern class. Then, all keypoints feature descriptor $\mathbf{f}(\mathbf{k})$ are mapped to the

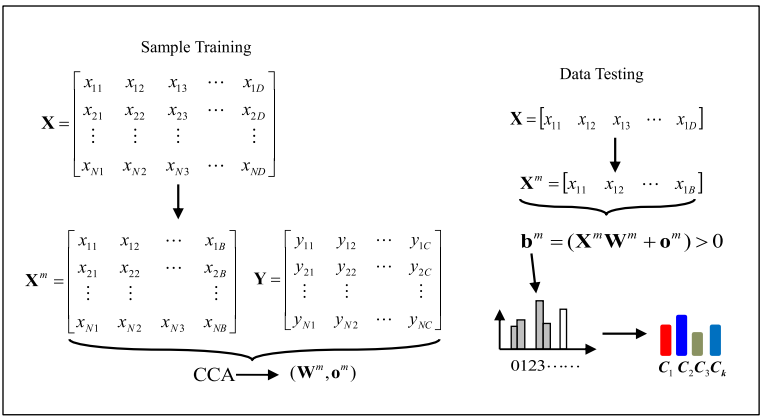


Fig. 6 Illustration of the embedded random ferns for a single base classifier

different projection spaces per embedded random fern classifier, and the conditional fern probabilities p_{l,c_i} and the final classification results can be obtained by the semi-naive Bayesian method.

Generally, the image keypoints with the maximum conditional fern probability is the classification result. However, the input feature descriptors $\mathbf{F} = \{\mathbf{f}(\mathbf{k})\}$ are the same in the training and test steps, and each keypoint $\mathbf{k}$ corresponds to a fern class, which results that the maximum probability calculated by the embedded random fern classifier comes from the keypoints itself. To solve this issue, we select the image keypoint with the *second* largest probability as the classification result in this paper. Therefore, the embedded random fern is skillfully applied to the stage of feature matching in forgery detection, which makes the matching strategy more efficient and feasible.

Figure 7 shows the image keypoint matching procedure for image "Japan_tower".

3.3 Post-processing

3.3.1 Removing falsely matched pairs using RANSAC

Although the local BRISK descriptor has higher robustness and discriminative power, and the embedded random fern based matching approach can efficiently detect the image keypoints pairs. However, the falsely matched pairs are inevitable because of the continuous nature of the image. To remove the wrong pairs, the affine transformation parameters can be robustly estimated by using the RANSAC method, which can estimate high-precision parameters from a data set that contains a large number of outliers. The RANSAC method [30] was first presented by Fischler and Bolles in 1981, which is often used to solve the problems of the image matched keypoints pairs and the calculation of the basic matrix in the field of computer vision. The basic assumption of the RANSAC algorithm is that the sample contains the correct data (inliers) and also contains abnormal data (outliers). These abnormal data may be generated due to incorrect measurements, incorrect assumptions, incorrect calculations, etc. At the same time, RANSAC also assumes that, for a given set of correct data, there is a way to calculate model parameters that fit these data. The RANSAC algorithm is a good choice for

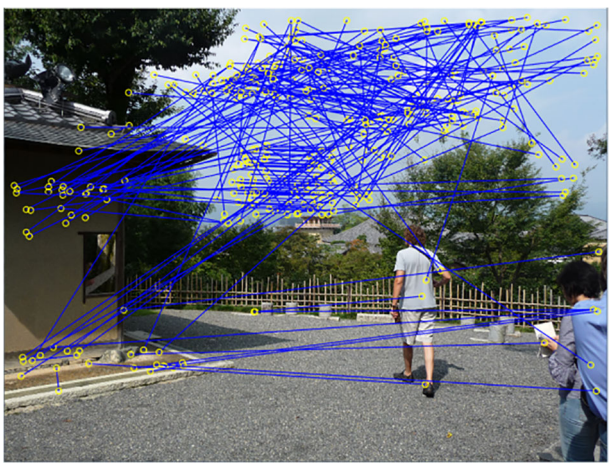


Fig. 7 The image keypoint matching procedure for image "Japan_tower"

robust estimation of model parameters because of the ability to handle a large number of outliers.

Let $\mathbf{M} = \{(\mathbf{k}, \mathbf{k}')\}$ denote the set of the matched keypoints pairs obtained by the embedded random fern. Where, $\mathbf{k} = (x, y, s)$ and $\mathbf{k}' = (x', y', s')$. We select three spatially adjacent collinear pairs from $\mathbf{M}$ to infer their affine matrix $\mathbf{H}$.

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \mathbf{H} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} \mathbf{SR} & \mathbf{t} \\ 0^T & 1 \end{pmatrix}, \quad \mathbf{SR} = \begin{pmatrix} s_x & 0 \\ 0 & s_y \end{pmatrix} \times \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}, \quad \mathbf{t} = \begin{pmatrix} t_x \\ t_y \end{pmatrix} \tag{10}$$

Where $\begin{pmatrix} s_x & 0 \\ 0 & s_y \end{pmatrix}$, $\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$ and $\begin{pmatrix} t_x \\ t_y \end{pmatrix}$ are scaling matrix, rotation matrix, and shift vector, respectively. All pairs in $\mathbf{M}$ are classified into inliers or outliers by checking the condition:

$$\left\| \mathbf{H} \cdot (x, \ y, \ 1)^T - \left(x', \ y', \ 1\right)^T \right\|_2^2 < T_\varepsilon \tag{11}$$

for classification threshold $T_\varepsilon = 2$. This procedure is repeated $T_{IT} = 500$ times, each time initialized with a triple of keypoints pairs randomly drawn from set $\mathbf{M}$. The RANSAC algorithm outputs the set of pairings with the largest number of inliers.

Figure 8 shows the false matches removal result for image "Japan_tower".

3.3.2 Localizing duplicated regions

In this paper, we localize the duplicated regions by employing a new template matching measure, which is based on the fast version of the mean-residual normalized production correlation (NNPROD) [23]. Template matching based accuracy duplicated regions localization is a crucial and challenging step in copy-move forgery detection, and the NNPROD approach has been widely used for template matching. However, direct calculation by using the prevalent NNPROD approach is usually computationally expensive because of the large number of redundancies that directly affect execution time. In order to lower the time complexity of the NNPROD algorithm, Li et al. [23] proposed a fast NNPROD method by

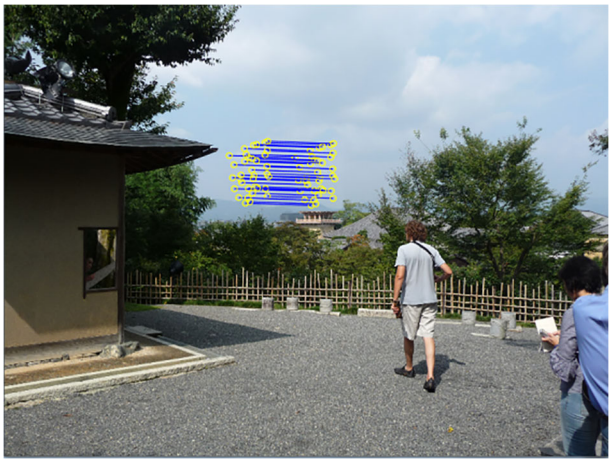


Fig. 8 The false matches removal results for image "Japan_tower"

expanding and optimizing the NNPROD through reducing the number of division calculations.

After obtaining the affine transform matrix $\mathbf{H}$, the warped images W_f and W_b can be obtained:

$$W_f = \mathbf{H} \cdot I, W_b = \mathbf{H}^{-1} \cdot I \tag{12}$$

For the sake of notation simplicity, we use W to refer to W_f or W_b. Next, block-wise correlation checking is performed between the original image I and warped one W for localizing duplicated areas. The NNPROD coefficient is used as measure of correlation:

$$\mathrm{NNPROD}(\mathbf{x}) = \frac{\sum_{\mathbf{z}\in\Omega(\mathbf{x})}\left(I(\mathbf{z})-\overline{I}\right)\left(W(\mathbf{z})-\overline{W}\right)}{\sqrt{\sum_{\mathbf{z}\in\Omega(\mathbf{x})}\left(I(\mathbf{z})-\overline{I}\right)^2\sum_{\mathbf{z}\in\Omega(\mathbf{x})}\left(W(\mathbf{z})-\overline{W}\right)^2}} \tag{13}$$

Here, $\Omega(\mathbf{x})$ represents a 7 pixels neighboring region centered at each pixel $\mathbf{x}$; $I(\mathbf{z})$ and $W(\mathbf{z})$ are the pixel values at the coordinate $\mathbf{z}$; $\overline{I}$ and $\overline{W}$ denote the mean pixel values of I and Win the region $\Omega(\mathbf{x})$. Next, with a threshold χ, the correlation map is converted to binary image G. We can rewrite the above formula as

$$\mathrm{NNPROD}(\mathbf{x}) = \frac{\sum_{\mathbf{z}\in\Omega(\mathbf{x})}\left(I(\mathbf{z})W(\mathbf{z})-\overline{W}I(\mathbf{z})-\overline{I}W(\mathbf{z})+\overline{I}\,\overline{W}\right)}{\sqrt{\sum_{\mathbf{z}\in\Omega(\mathbf{x})}\left(I^2(\mathbf{z})-2\overline{I}I(\mathbf{z})+\overline{I}^2\right)\sum_{\mathbf{z}\in\Omega(\mathbf{x})}\left(W^2(\mathbf{z})-2\overline{W}W(\mathbf{z})+\overline{W}^2\right)}} \tag{14}$$

Obviously, the calculation of NNPROD consists of 10 independent items, and their computational complexity is roughly equal. To reduce the complexity, we give the following approximate relationship based on the continuity of the natural image.

$$I(\mathbf{z}) \approx \overline{I}, W(\mathbf{z}) \approx \overline{W} \tag{15}$$

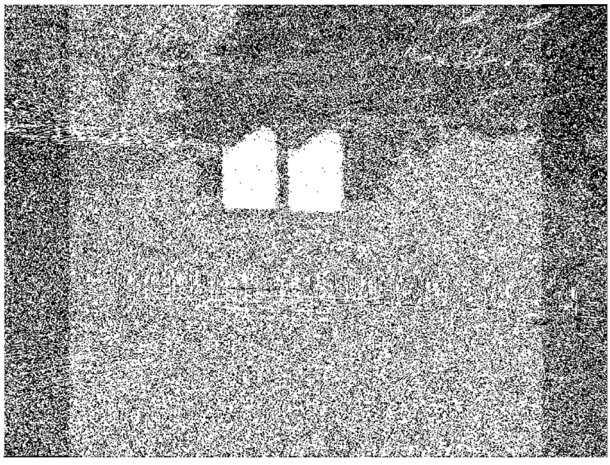


Fig. 9 The duplicated regions localizing result for image "Japan_tower"

And we can get the simplified formula, namely fast NNPROD, as

$$\mathrm{NNPROD}(\mathbf{x}) \approx \frac{\sum_{\mathbf{z}\in\Omega(\mathbf{x})}\left(I(\mathbf{z})W(\mathbf{z}) - \overline{I}\,\overline{W}\right)}{\sqrt{\sum_{\mathbf{z}\in\Omega(\mathbf{x})}\left(I^2(\mathbf{z}) - \overline{I}^2\right)\sum_{\mathbf{z}\in\Omega(\mathbf{x})}\left(W^2(\mathbf{z}) - \overline{W}^2\right)}} \tag{16}$$

Here, the calculation of the fast NNPROD contains only 6 independent items, which is significantly less than the classical algorithm. After obtaining the correlation maps, we can also reduce the noise by using a Gaussian filter and set the binarization threshold χ to 0.4. Figure 9 shows the duplicated regions localizing result for image "Japan_tower".

3.3.3 Morphological filtering operation and marking

In order to get the final copy-move areas positioning results, we first eliminate the small isolated areas in the binary correlation map G, by using a tampered area threshold $\gamma = 0.05\%$ which represent the minimum proportion of the tampered area. Then, we employ the morphological filtering operation to fill and merge the "holes" presented in the region for the purpose of smoothing and connecting the boundary of detected tampered areas. According to the identified location area, a color mark is made on the tampered areas of image. Figure 10 shows the final detection results for image "Japan_tower".

4 Experimental results

In this Section, a series of the simulation experimental results for the proposed method are presented. Firstly, we will present the image datasets needed for the experiment, as well as the evaluation metrics for the experimental results. Next, the setting of various parameter values of the proposed method will be described in detail. Finally, we compare the experimental results of the proposed method with some state-of-the-art approaches [1, 2, 8–10, 31]. The method is implemented in MATLAB R2011a on a computer with Inter Core i7 3.6 GHz Pentium CPU and 16 GB RAM.

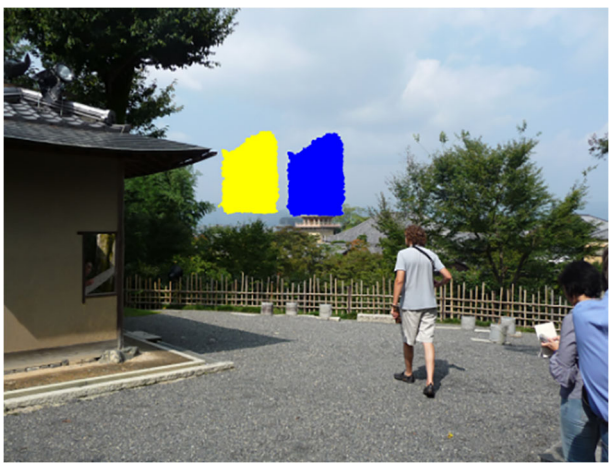


Fig. 10 The final detection results for image "Japan_tower"

4.1 Datasets and evaluation metrics

In order to assess the efficiency of the proposed CFMD method, the evaluation should be implemented through various copy-move image datasets. Christlein et al. [8] constructed the Image Manipulation Dataset based on 48 original color images, called FAU. These images in this dataset are quite large, and the typical size of an image is about 3000 × 2400 pixels and the tampered regions occupying about 6% of pixels of an entire image. The copied regions are created by a series of attacking operations such as rotation, scale, JPEG compression and additive white Gaussian noise, and its categories can be divided from smooth to textured area, which contains man-made, living, nature and mixed. A further dataset composed by 80 images is used, which is called GRIP [10]. All these images have the size of 768 × 1024, with arbitrary shapes in forgeries areas, which aimed at obtaining visually satisfactory results, while duplicated areas covering less than 1% of each image.

To evaluate the CMFD performance in this paper, the two metrics Recall and Precision are used. Recall is the probability that the relevant regions are detected, and it denotes the ratio of the number of correctly detected tampered pixels to the number of tampered pixels in the ground-truth image. Precision is the probability that the detected regions are relevant, and it denotes the ratio of the number of correctly detected tampered pixels to the number of totally tampered pixels. We computed the error metrics Precision and Recall as the following formulas

$$\text{Recall} = \frac{|G \cap GT|}{|GT|} \quad , \qquad \text{Precision} = \frac{|G \cap GT|}{|G|} \tag{17}$$

where G represents the detected region, and GT is the ground-truth forgery region. In additional, another metric F that combines both Recall and Precision is also computed, which can be calculated using (18)

$$\text{F} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{18}$$

In this paper, we measure the efficiency of proposed CMFD method both at image level and pixel level. Recall and Precision are computed by checking the number of pixels in the corresponding area at the pixel level. At the image level, Recall is the probability that a forged image is detected, and Precision is the probability that a detected forgery is truly a forgery. In general, the goal of CMFD method is to achieve the higher Recall and higher Precision simultaneously, which can indicate the superior performance of the method.

4.2 Settings for forgery detection

The parameter setting is an important part of the CMFD methods. Since the performance at different steps of forgery detection is high dependency, the setting of one parameter is often closely related to other parameters. And the finally results of experiment will be affected by the compatibility of parameter settings. Therefore, some important parameter settings are essential for obtaining valid test results. There are three important parameters to be adapted in this paper: the size of image sub-block n, the uniformity measurement threshold ζ, and NNPROD binarization threshold χ. We empirically determine the appropriate parameter values and test the impact of each parameter on the proposed method.

The size of image sub-block The size of the image sub-blocks n can directly affect the distribution of SURF keypoints, further affecting the number of matched pairs. When sub-

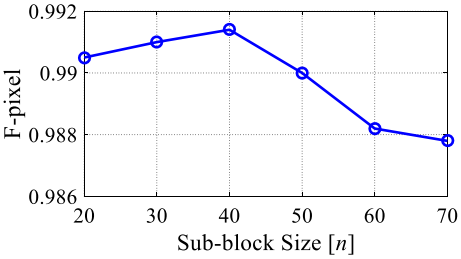


Fig. 11 The F-measure for different size of sub-block

block size is larger, the distribution of the extracted keypoints is not uniform. On the other hand, the keypoints cannot be effectively extracted from the smooth regions with a smaller sub-block, which will lead to the increase of time complexity. Therefore, the appropriate size of sub-blocks n is determined by experiment. We employ different sizes of sub-block ranging from 20 to 70 in increments of 10 and calculate the corresponding F-measure. As can be seen from Fig. 11, the F is well performance with the size of block $n = 40$.

The uniformity measurement threshold If the uniformity measurement value Φ is smaller than the threshold ζ, the keypoints are well distributed. Otherwise, if the threshold ζ is too small, the time complexity of the keypoints extraction will be increased. So we use different size thresholds ζ ranging from 0.3 to 0.8 in increments of 0.1, and compute the corresponding average F-measure and average CPU time of uniformly keypoints extraction by utilizing twenty forged image from dataset FAU. As can be seen from Fig. 12, in order to achieve the higher F-measure and the relatively appropriate time, we can set the uniformity measurement threshold $\zeta = 0.5$.

Binarization threshold The fast NNPROD approach has lower time complexity than the original NNPROD algorithm, while maintaining the same high accuracy. The Gaussian filter is used to remove the Gaussian white noise, and the binarization threshold χ is used to evaluate binarization of the correlation bitmap results. The F-measure is well performance with the threshold $\chi = 0.4$, as shows in Fig. 13, so the binarization threshold χ is set to 0.4.

4.3 Detection results at pixel level

To evaluate the detection results at pixel level in this experiment, we use the metric F to measure the mark accuracy of the tampered areas. In practical experiment, firstly, the proposed method is evaluated under ideal conditions (the test images without any post-processing

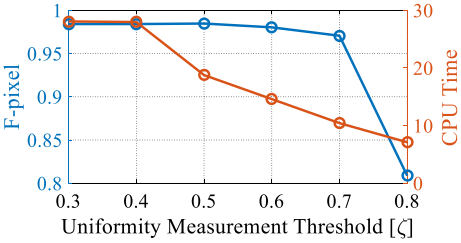


Fig. 12 The F-measure and CPU time for different uniformity measurement threshold

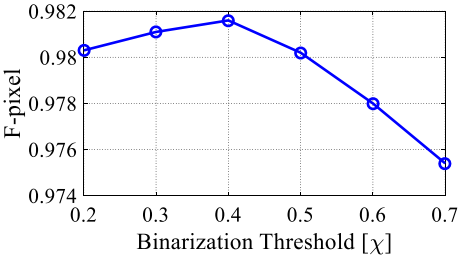


Fig. 13 The F-measure for different binarization thresholds

operations), that is, only a simple copy-move transformation. Then, those images are measured after a series of attacks on the copied areas such as: JPEG compression, additive white Gaussian noise, down-sampling, rotation and scaling.

Plain copy-move We first evaluate the performance of the proposed CMFD method without any post-processing operations of the test images. Here, we have the 128 original images and 128 forgery images from FAU and GRIP datasets, in which only a simple copy-move transformation is performed. In this case, we must distinguish the original images and the forgery images.

JPEG compression We utilize the forgery images that are JPEG compressed with the intensity factors from 20 to 100 in steps of 10. We have to test 128 forged images of FAU dataset and

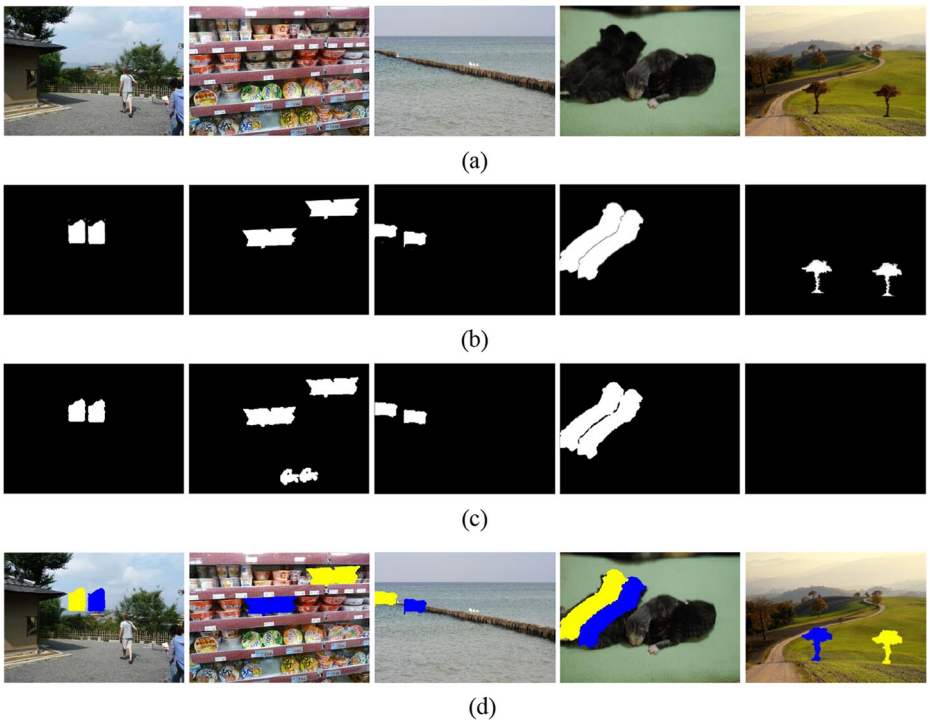


Fig. 14 Detection results: **a** Several forged images from FAU dataset, **b** Ground-truth, **c** The detection result using Scheme [10], **d** The detection result using our method

GRIP dataset for each of the assessed intensity levels, so we need to test $128 \times 9 = 1152$ forgery images.

Rotation In rotation tampering, the forged regions are rotated with the angle from 2° to 10° in steps of 2°. Here, we need to test a total of $128 \times 5 = 1024$ forgery images from FAU dataset and GRIP dataset.

Scaling In the FAU dataset, we scaled the forged regions with the scale factor from 91% to 109% in steps of 2%. In the GRIP dataset, we scaled the forged regions with the scale factor from 50% to 200% with varying steps. Here, a total of $48 \times 10 + 80 \times 9 = 1200$ forgery images from FAU dataset and GRIP dataset must be tested.

Additive white Gaussian noise (AWGN) The forged regions are blurred by noise of the 80 forgery images from GRIP dataset. The intensity of distortion is measured by two parameters: noise's standard deviation and filter's radius. To this end, we consider zero-mean AWGN with standard deviation 2, 4, 6, 8 and 10, and a fixed filter's radius.

Down-sampling The forged images in the FAU dataset are scaled down with the scale factors from 90% to 10% in steps of 20%. Here, we need to test a total of $48 \times 5 = 240$ images.

Figure 14 reports the experimental results of several images from FAU dataset [8], and Fig. 15 gives the experimental results of several images from GRIP dataset [10]. From these results, we can clearly see that the proposed method has more excellent detection performance

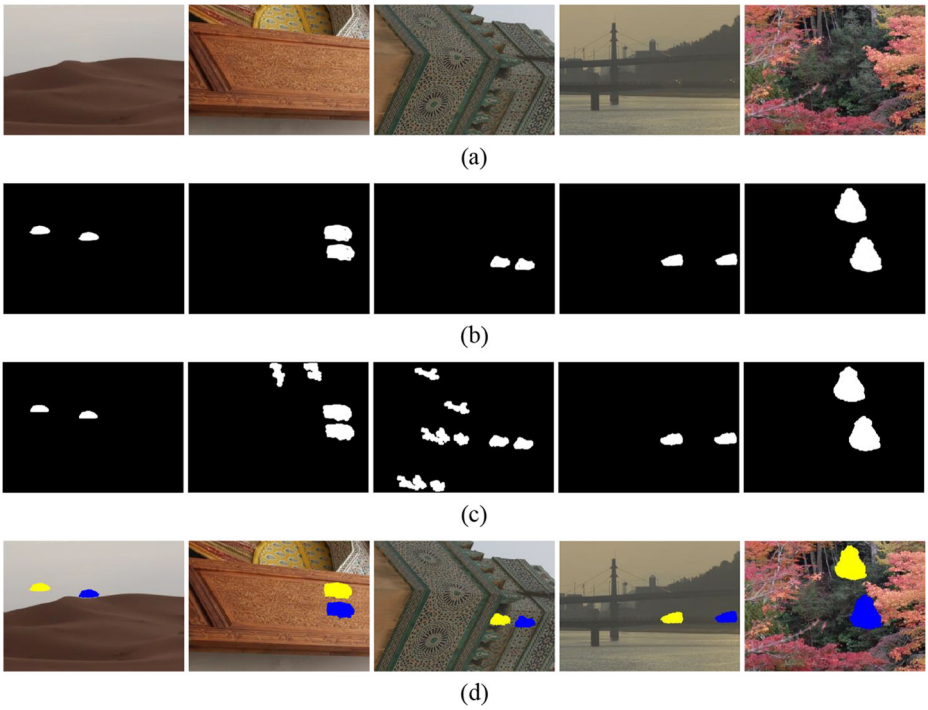


Fig. 15 Detection results: **a** Several forged images from GRIP dataset, **b** Ground-truth, **c** The detection result using Scheme [10], **d** The detection result using our method

than the method proposed by Cozzolino et al. [10]. Please note the undetected forged area and wrongly detected genuine region in the algorithm [10], which is mainly caused by the weak description ability of the feature and the approximation of the matching algorithm. Figure 16 demonstrates the pixel-level detection performance of the images in the FAU dataset under various attacks. Figure 17 demonstrates the pixel-level detection performance of the images in the GRIP dataset under various attacks. The comparison results shows that our technique outperforms the existing state-of-the-art schemes when under the various attacks. This phenomenon is mainly due to the following two factors. (1) Limited by the high computational cost, the block-based algorithms [9, 10] typically use approximate nearest neighbor searches and simple post-processing strategies, which have a negative impact on the robustness of the algorithm. (2) In the FAU and GRIP datasets, some tampered regions are very smooth, which are challenging for the keypoint-based methods [1, 2, 8, 31] because the number of features extracted is often very limited in such regions.

4.4 Detection results at image level

When the tested image is known as a forgery, we can use the metrics such as Recall, Precision and F to measure the performance of detection at pixel level. However, the tested images in practice are generally not known in advance, so we will conduct image-level performance

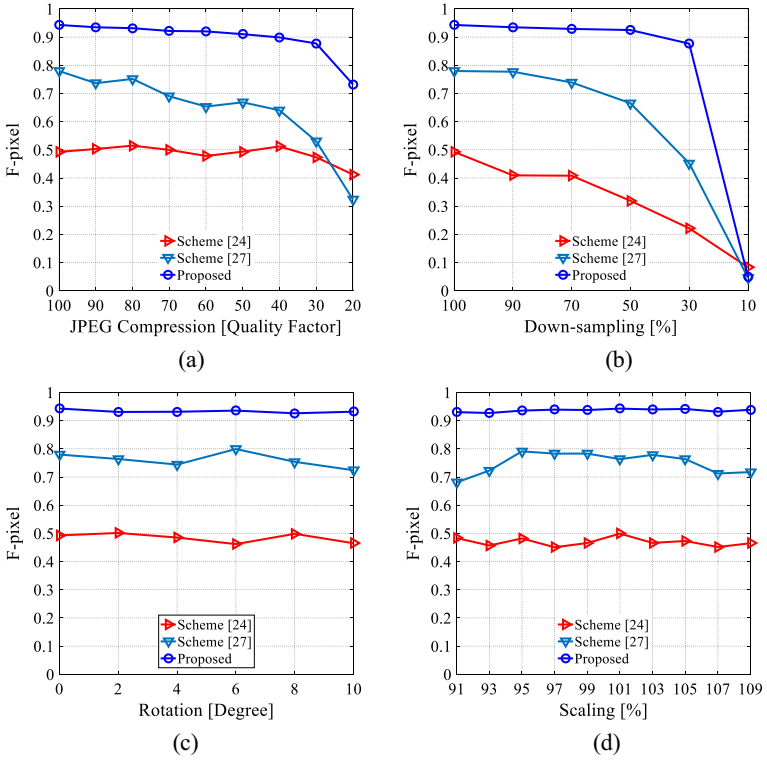


Fig. 16 Average F-measure at pixel level on 48 forged images from FAU dataset: **a** JPEG compression, **b** Down-sampling, **c** Rotation, **d** Scaling

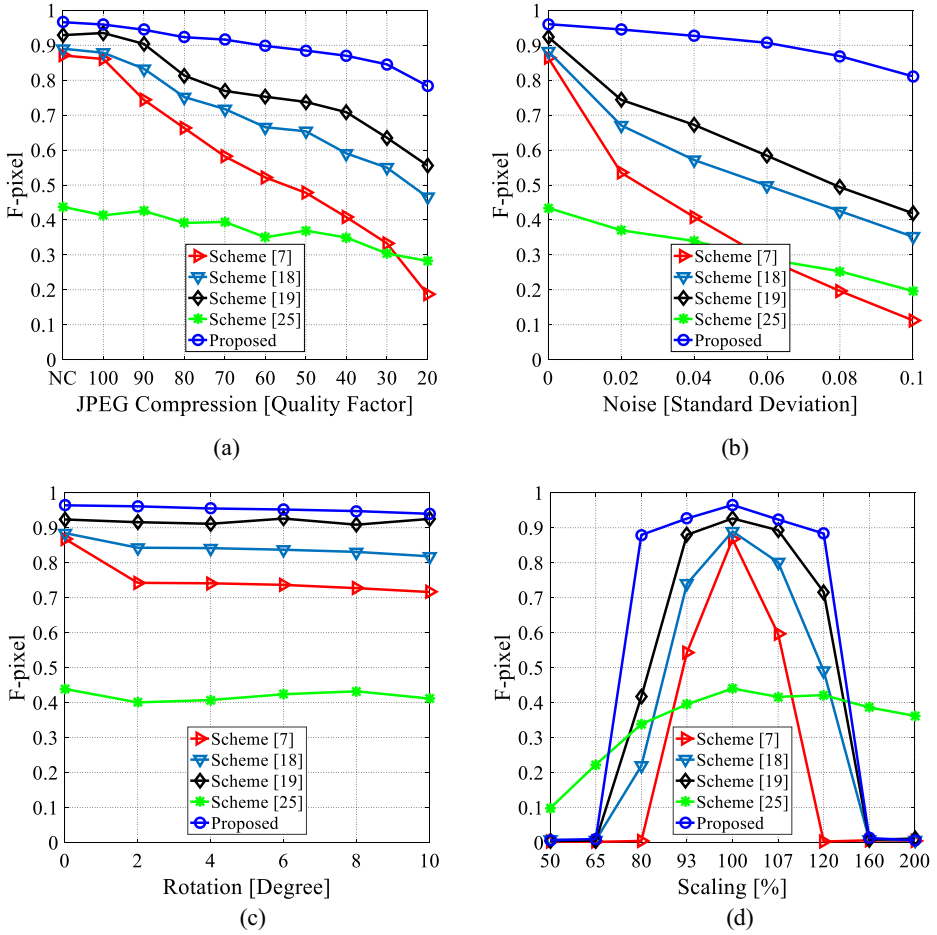


Fig. 17 Average F-measure at pixel level on 80 forged images from GRIP dataset: **a** JPEG compression, **b** Additive white Gaussian noise, **c** Rotation, **d** Scaling

testing in the next set of experiments. Specifically, a successful detection for a forgery image is that the scheme tests a forged area greater than the region threshold. For an un-tampered image, the real downside occurs when the scheme cannot detect any forged areas.

To measure the detection performance at image level, the proposed method is assessed under ideal conditions (the test images without any post-processing operations), which is shown in Tables 1 and 2. These tables illustrate the average image-level and pixel-level F-measure for FAU and GRIP dataset, respectively. We compare the proposed method with

Table 1 Average pixel-level and image-level F-measure for plain copy-move on the FAU dataset

Methods	F-image	F-pixel
Scheme [8]	0.9302	0.9352
Scheme [9]	0.9485	0.8977
Scheme [10]	0.9399	0.9262
Scheme [2]	0.7407	0.5011
The proposed method	0.9621	0.9396

Table 2 Average pixel-level and image-level F-measure for plain copy-move on the GRIP dataset

Methods	F-image	F-pixel
Scheme [8]	0.9816	0.8744
Scheme [9]	0.9467	0.8867
Scheme [10]	0.9340	0.9267
Scheme [2]	0.6772	0.4441
The proposed method	0.9753	0.9640

several promising technologies in recent years. From these tables, we can see that all forgery detection results are very well aligned on the two dataset. However, the proposed method has more excellent detection performance than the recent promising technologies.

From the above test results, it is can be observed that our proposed method has more excellent detection performance under both ideal conditions and various attack conditions compared with some state-of-the-art approaches recently proposed in the literature. This is because that: (1) By using the adaptive threshold setting, we extracted evenly and robustly the SURF keypoints from the forged image, including the smooth and small region; (2) We introduced the binary robust invariant scalable keypoints (BRISK) descriptor to represent the local image feature of image keypoints, which enhanced the discriminative power and robustness of keypoints features; (3) We employed the embedded random ferns approach to match BRISK features, which can provide higher matching accuracy compared to the approximate nearest neighbor algorithm such as BBF KD-tree and PatchMatch; (4) We use the RANSAC and fast NNPROD to locate the tampering area, which can effectively improve the detection accuracy and reduce the detection time.

Although the proposed image forgery detection method has better detection performance, there is still a lot of room for improvement. For example, the proposed detection method requires more computing power for the FAU dataset on a desktop computer with Dual Core 3.6-GHz Pentium CPU and 16 GB RAM, which causes the proposed detection method ineffective for larger images in real-time programs.

5 Conclusion

In this paper, we presented an improved method for keypoints based copy-move forgery detection. The main novelties of this work consist of the following four aspects: First, the adaptive uniform distribution threshold was introduced to solve the problem that the image keypoints cannot be extracted sufficiently from small smooth regions. Second, the BRISK descriptor was employed to represent the binary feature descriptors of SURF keypoints, which enhanced the discriminative power and robustness of image keypoints features. Third, we matched the BRISK features of image keypoints by using embedded random ferns approach, which provided the higher matching accuracy than other classical algorithms. Forth, the falsely matched keypoints pairs are eliminated by utilizing RANSAC method and the fast NNPROD approach are used to locate the tampering area. To test the performance of the proposed CMFD approach, we conducted experiments on the FAU and GRIP datasets. The results show that the proposed CMFD approach is the most satisfying one in comparison to state-of-the-art methods.

The limitations of our current method are mainly reflected in the higher time consumption, which make it impossible to be effectively applied in real-time detection. In future work, we plan to combine the hierarchical matching strategy with the embedded random ferns KNN searches to significantly reduce computational complexity.

Acknowledgments This work was supported partially by the National Natural Science Foundation of China (Nos. 61701212 & 61472171), China Postdoctoral Science Foundation (No. 2017 M621135, 2018 T110220), and High-level Innovation Talents Foundation of Dalian (No.2017RQ055).

References

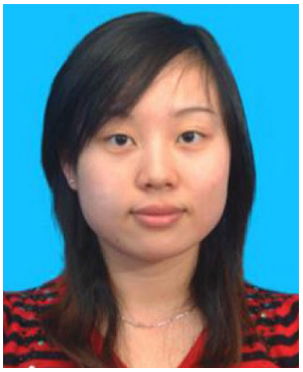
1. Amerini I, Ballan L, Caldelli R (2011) A SIFT-based forensic method for copy-move attack detection and transformation recovery. IEEE Trans on Information Forensics & Security 6(3):1099–1110
2. Amerini I, Ballan L, Caldelli R, Del Bimbo A, Del Tongo L, Serra G (2013) Copy-move forgery detection and localization by means of robust clustering with J-Linkage. Signal Process Image Commun 28(6):659–669
3. Asikuzzaman M, Pickering MR (2017) An overview of digital video watermarking. IEEE Trans on Circuits and Systems for Video Technology 28(9):2131–2153
4. Bay H, Ess A, Tuytelaars T et al (2008) Speeded-up robust features (SURF). Comput Vis Image Underst 110(3):346–359
5. Birajdar GK, Mankar VH (2013) Digital image forgery detection using passive techniques: A survey. Digit Investig 10(3):226–245
6. Bosch A, Zisserman A, Munoz X (2007) Image Classification using random forests and ferns. IEEE 11th International Conference on Computer Vision, Rio de Janeiro, pp 1–8
7. Bravo-Solorio S, Nandi AK (2011) Exposing duplicated regions affected by reflection, rotation and scaling. 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, pp 1880–1883
8. Christlein V, Riess C, Jordan J (2012) An evaluation of popular copy-move forgery detection approaches. IEEE Trans on Information Forensics & Security 7(6):1841–1854
9. Cozzolino D, Poggi G, Verdoliva L (2014) Copy-move forgery detection based on patchmatch. 2014 IEEE International Conference on Image Processing (ICIP), Paris, pp 5312–5316
10. Cozzolino D, Poggi G, Verdoliva L (2015) Efficient dense-field copy-move forgery detection. IEEE Trans on Information Forensics and Security 10(11):2284–2297
11. Dixit R, Naskar R (2017) Review, analysis and parameterization of techniques for copy–move forgery detection in digital image. IET Image Process 11(9):746–759
12. Donoser M, Schmalstieg D (2014) Discriminative feature-to-point matching in image-based localization. 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, pp 516–523
13. Fridrich J, Soukal D, Lukas J (2003) Detection of copy-move forgery in digital images. In: Proceedings of Digital Forensic Research Workshop, Cleveland
14. Gong J, Guo J (2016) Image copy-move forgery detection using SURF in opponent color space. Transactions of Tianjin University 22(2):151–157
15. Huang H, Guo W, Zhang Y (2008) Detection of copy-move forgery in digital images using SIFT algorithm, vol 2. Pacific-Asia Workshop on Computational Intelligence and Industrial Application (PACIIA), Wuhan, pp 272–276
16. Jin G, Wan X (2017) An improved method for SIFT-based copy–move forgery detection using non-maximum value suppression and optimized J-Linkage. Signal Process Image Commun 57:113–125
17. Kakar P, Sudha N (2012) Exposing postprocessed copy-paste forgeries through transform-invariant features. IEEE Trans on Information Forensics & Security 7(3):1018–1028
18. Korus P (2017) Digital image integrity–a survey of protection and verification techniques. Digital Signal Processing 71:1–26
19. Kumar S, Desai JV, Mukherjee S (2016) A fast keypoint based hybrid method for copy move forgery detection. Computer Vision and Pattern Recognition 4(2):91–99
20. Lepetit V, Fua P (2013) Keypoint recognition using random forests and random ferns. Decision Forests for Computer Vision and Medical Image Analysis:111–124
21. Leutenegger S, Chli M, Siegwart RY (2011) BRISK: Binary Robust invariant scalable keypoints. 2011 International Conference on Computer Vision, Barcelona, pp 2548–2555
22. Li J, Li X, Yang B, Sun X (2015) Segmentation-based image copy-move forgery detection scheme. IEEE Trans on Information Forensics & security 10(3):507–518
23. Li C, Yu F, Lin Z, Kang X (2016) A novel fast target tracking based on video image. Proceedings of the 35th Chinese Control Conference, Chengdu, pp 10264–10268
24. Li Y, Zhou J (2018) Fast and effective image copy-move forgery detection via hierarchical feature point matching. IEEE Trans on Information Forensics and Security 14(5):1307–1322

25. Manu VT, Mehtre BM (2016) Detection of copy-move forgery in images using segmentation and SURF. Advances in Intelligent Systems and Computing, Cham, pp 645–654
26. Muhammad G, Hussain M, Bebis G (2012) Passive copy move image forgery detection using undecimated dyadic wavelet transform. Digit Investig 9(1):49–57
27. Muhammad AQ, Mohamed D (2015) A bibliography of pixel-based blind image forgery detection techniques. Signal Process Image Commun 39(Part A):46–74
28. Muzaffer G, Karaagacli ES (2017) Recent keypoint based copy move forgery detection techniques. 2017 International Artificial Intelligence and Data Processing Symposium (IDAP), Malatya, pp 1–7
29. Ozuysal M, Calonder M, Lepetit V, Fua P (2009) Fast keypoint recognition using random ferns. IEEE Trans on Pattern Analysis and Machine Intelligence 32(3):448–461
30. Pan X, Lyu S (2010) Region duplication detection using image feature matching. IEEE Trans on Information Forensics and Security 5(4):857–867
31. Pun C, Yuan X, Bi X (2015) Image forgery detection using adaptive over-segmentation and feature point matching. IEEE Trans on Information Forensics Security 10(8):1705–1716
32. Ryu SJ, Lee MJ, Lee HK (2010) Detection of copy-rotate-move forgery using Zernike moments. International Workshop on Information Hiding, Berlin, Heidelberg, pp 51–65
33. Sachdev K, Kaur M, Gupta S (2017) A robust and fast technique to detect copy move forgery in digital images using SLIC segmentation and SURF keypoints. Proceeding of International Conference on Intelligent Communication, Control and Devices. Springer Singapore, 479: 787-793
34. Silva E, Carvalho T, Ferreira A, Rocha A (2015) Going deeper into copy-move forgery detection: Exploring image telltales via multi-scale analysis and voting processes. J Vis Commun Image Represent 29:16–32
35. Soni B, Das PK, Thounaojam DM (2018) CMFD: a detailed review of block based and key feature based techniques in image copy-move forgery detection. IET Image Process 12(2):167–178
36. Sun L, Ji S, Ye J (2011) Canonical correlation analysis for multilabel classification: A least-squares formulation, extensions, and analysis. IEEE Trans on Pattern Analysis & Machine Intelligence 33(1): 194–200
37. Teerakanok S, Uehara T (2019) Copy-Move Forgery Detection: A State-of-the-Art Technical Review and Analysis. IEEE Access 7:40550–40568
38. Wang XY, Li S, Liu YN, Niu Y et al (2017) A new keypoint-based copy-move forgery detection for small smooth regions. Multimed Tools Appl 76(22):23353–23382
39. Wang H, Wang HX (2018) Perceptual Hashing-Based Image Copy-Move Forgery Detection. Security and Communication Networks
40. Wang H, Wang HX, Sun XM, Qian Q (2017) A passive authentication scheme for copy-move forgery based on package clustering algorithm. Multimed Tools Appl 76(10):12627–12644
41. Yang F, Li J, Lu W, Weng J (2017) Copy-move forgery detection based on hybrid features. Eng Appl Artif Intell 59:73–83
42. Yang B, Sun X, Guo H, Xia Z, Chen X (2018) A copy-move forgery detection method based on CMFD-SIFT. Multimed Tools Appl 77(1):837–855
43. Yao H, Cao F, Tang Z, Wang J, Qiao T (2018) Expose noise level inconsistency incorporating the inhomogeneity scoring strategy. Multimed Tools Appl 77(14):18139–18161
44. Yao H, Wang S, Zhang X, Qin C, Wang J (2017) Detecting image splicing based on noise level inconsistency. Multimed Tools Appl 76(10):12457–12479
45. Yavuz AA, Mudgerikar A, Singla A, Papapanagiotou I, Bertino E (2017) Real-time digital signatures for time-critical networks. IEEE Trans on Information Forensics Security 12(11):2627–2639
46. Zandi M, Mahmoudi-Aznaveh A, Talebpour A (2016) Iterative copy-move forgery detection based on a new interest point detector. IEEE Trans on Information Forensics and Security 11(11):2499–2512
47. Zhao J, Guo J (2013) Passive forensics for copy-move image forgery using a method based on DCT and SVD. Forensic Sci Int 233(1-3):158–166

Hong-Ying Yang is currently a professor with the School of Computer and Information Technology at the Liaoning Normal University, China. Her research interests include signal processing and communications, digital multimedia data hiding.

Shu-Ren Qi received the B. E. degree from the School of Computer and Information Technology, Liaoning Normal University, China, in 2017, where he is currently pursuing the M. S. E. degree. His research interests include image processing and forgery detection.

Ying Niu received the B. E. degree from the School of Computer and Information Technology, Liaoning Normal University, China, in 2015, where she is currently pursuing the M. S. E. degree. Her research interests include digital watermarking and forgery detection.



Pan-Pan Niu received the Ph. D. degree from the School of Computer and Information Technology, Dalian Maritime University, China, in 2013. She is currently an associate professor with the School of Computer and Information Technology at the Liaoning Normal University, China. Her research interests include image processing and pattern recognition.

Xiang-Yang Wang is currently a professor with the Multimedia and Information Security Laboratory, School of Computer and Information Technology, Liaoning Normal University, China. His research interests lie in the areas of information security, image processing, pattern recognition, and computer vision. He is the author of two books. He has published over 80 papers in international journals (including IEEE/ACM Transactions) and 25 papers in international conferences and workshops. Mr. Wang is a Reviewer for many leading international and national journals and conferences, including IEEE/ACM Transactions.