Multimed Tools Appl (2017) 76:23353–23382
DOI 10.1007/s11042-016-4140-5

A new keypoint-based copy-move forgery detection for small smooth regions

Xiang-Yang Wang[1] · Shuo Li[1] · Yu-Nan Liu[1] · Ying Niu[1] · Hong-Ying Yang[1] · Zhi-li Zhou[2]

Received: 6 May 2016 / Revised: 21 September 2016 / Accepted: 7 November 2016 /
Published online: 18 November 2016

Abstract Copy-move forgery is one of the most common types of image forgeries, where a region from one part of an image is copied and pasted onto another part, thereby concealing the image content in the latter region. Keypoint based copy-move forgery detection approaches extract image feature points and use local visual features, rather than image blocks, to identify duplicated regions. Keypoint based approaches exhibit remarkable performance with respect to computational cost, memory requirement, and robustness. But unfortunately, they usually do not work well if smooth background areas are used to hide small objects, as image keypoints cannot be extracted effectively from those areas. It is a challenging work to design a keypoint-based method for detecting forgeries involving small smooth regions. In this paper, we propose a new keypoint-based copy-move forgery detection for small smooth regions. Firstly, the original tampered image is segmented into nonoverlapping and irregular superpixels, and the superpixels are classified into smooth, texture and strong texture based on local information entropy. Secondly, the stable image keypoints are extracted from each superpixel, including smooth, texture and strong texture ones, by utilizing the superpixel content based adaptive feature points detector. Thirdly, the local visual features, namely exponent moments magnitudes, are constructed for each image keypoint, and the best bin first and reversed generalized 2 nearest-neighbor algorithm are utilized to find rapidly the matching image keypoints. Finally, the falsely matched image keypoints are removed by customizing the random sample consensus, and the duplicated regions are localized by using zero mean normalized cross-correlation measure. Extensive experimental results show that the newly proposed scheme can achieve

✉ Xiang-Yang Wang
wxy37@126.com

✉ Hong-Ying Yang
yhy_65@126.com

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

[2] Jiangsu Engineering Center of Network Monitoring & School of Computer and Software, Nanjing University of Information Science & Technology (NUIST), Nanjing 210044, China

much better detection results for copy-move forgery images under various challenging conditions, such as geometric transforms, JPEG compression, and additive white Gaussian noise, compared with the existing state-of-the-art copy-move forgery detection methods.

Keywords Copy-move forgery detection · Superpixel · Adaptive feature points detector · Exponent moments · Reversed generalized 2 nearest-neighbor

1 Introduction

With the development of computer technology and image processing software tools, digital image forgery has been increasingly easy to perform. It has become a severe threat to security that anyone may access and modify the content of an image without leaving visually detectable traces. In recent years, more and more researchers have begun to focus on the problem of digital image forgery, and various methods have been developed to counter tampering and forgery in order to ensure the authenticity of images [9]. Current image forgery detection methods can be roughly categorized as active and passive (blind). Active methods such as watermarking or illegal image copy detection depend on prior information about the image. However, in many situations, prior information regarding an image is not available and passive, or blind methods should be used to authenticate the image. The practicality and wide applicability of passive methods have made them a popular topic of research [35].

Copy-move (or copy-paste) forgery is one of the most common types of image forgeries, where a region from one part of an image is copied and pasted onto another part, thereby concealing the image content in the latter region. Such concealment can be used to hide an undesired object or increase the number of objects apparently present in the image. Although a simple translation may be sufficient in many cases, additional operations are often performed in order to better hide the tampering. These include scaling, rotation, lossy compression, noise addition, among others. Because the wide availability of image processing software has made it easy to perform copy-move operations, passive copy-move forgery detection (CMFD) is becoming one of the most important and popular digital image forensic techniques currently [12, 39]. In this paper, we focus on passive image copy-move forgery detection. In previous years, many passive forgery detection methods have been proposed for copy-move forgery detection. According to the existing methods, the copy-move forgery detection methods can be categorized into two main categories: block-based algorithms and keypoint-based algorithms. They both try to detect the copy-move forgery through describing the local patches of one image.

The block-based forgery detection methods usually divide the input images into overlapping and regular image blocks, and then the tampered region can be obtained by matching blocks of image pixels or transform coefficients. Because the descriptor of the block is important for the algorithm, various description methods, like 1D-fourier transform (FT), discrete cosine transform (DCT), discrete wavelet transform (DWT), singular value decomposition (SVD), geometric moment, histogram, and principal component analysis (PCA) etc., have been introduced [32]. Although these block-based forgery detection methods are effective in forgery detection, they have three main drawbacks: 1) nearly all of these methods are based on a large number of blocks and the feature vectors extracted from the blocks are large, which results in high computational complexity due to the fact that multiple-index sorting is required to enable lexicographical sorting of all of the blocks. In addition, higher computational load for extracting feature vectors is also the weakness of some block-based forgery detection methods;

2) the host image is usually divided into overlapping rectangular blocks, which are fragile to geometrical transformations (e.g., scaling, rotation, etc.). So, the existing methods always cannot address significant geometrical transformations of the forgery regions; and 3) their recall rate is low under various noises and geometric transformations, the reason for this is that the extracted feature description usually cannot stably capture the image information.

As an alternative to the block-based methods, keypoint-based forgery detection methods were proposed. Unlike block-based algorithms, keypoint-based methods rely on the identification and selection of high-entropy image regions (i.e., the "keypoints"). A feature vector is then extracted from per keypoint. Consequently, fewer feature vectors are estimated, resulting in reduced computational complexity of feature matching and postprocessing [42]. The lower number of feature vectors dictates that postprocessing thresholds are also to be lower than that of block-based methods. Because the number of the keypoints is much smaller than that of the blocks divided in an overlapping way, the keypoint-based algorithms require less computational resource than the block-based ones. In this regard, the keypoint-based methods are faster and more favorable than the block-based ones. However, on the other hand, keypoint-based method also has the following three problems. Firstly, these methods are intrinsically less accurate than block-based ones. For keypoint-based algorithms, the copied regions are often only sparsely covered by matched keypoints. If the copied regions exhibit little structure, it may happen that the region is completely missed. Secondly, they usually cannot effectively detect the forgery regions in smooth image. The state-of-the-art intensity based image point detectors are always sensitive to texture region, and most detected keypoints gather in high contrast regions. So, they do not work if homogenous areas are used to hide an object, as keypoints cannot be extracted from those areas. Thirdly, these methods are often not robust to many postprocessing operations. The extraction of image keypoints is a key component of the CMFD schemes. However, the usually utilized image keypoints detectors, including Harris-Laplace, SIFT, and SURF, are not strongly resistant to some common image processing operations and geometric transformation, which prevents these method from being extended to detect the postprocessed duplicated regions.

In this paper, we propose a new keypoint-based copy-move forgery detection approach, which can identify effectively small hidden objects in smooth background area. The novelty of the proposed approach includes: 1) By utilizing the content based adaptive feature points detector, the stable image keypoints are extracted simultaneously from smooth and texture regions; 2) Robust local visual features, namely exponent moments magnitudes, are constructed for each image keypoint; 3) Best bin first (BBF) and reversed generalized 2 nearest-neighbor (RG2NN) are combined to find rapidly the matching image keypoints.

The rest of this paper is organized as follows. A review of previous related work is presented in Section 2. Section 3 recalls some preliminaries about speeded-up robust features (SURF) detector, and presents the improved SURF detector based on probability density gradient. Section 4 introduces the Exponent moments, and analyzes the invariant property of Exponent moments. Section 5 contains the description of our keypoint-based based copy-move forgery detection procedure. Simulation results in Section 6 will show the performance of our scheme. Finally, Section 7 concludes this presentation.

2 Related work

Over the past decades, passive copy-move forgery detection has been an active area of research in many applications, including: criminalization, surveillance systems, medical imaging, and

news media. Many approaches have been proposed to solve this problem. They can be roughly divided into two major categories: block-based algorithms and keypoint-based algorithms. Table 1 summarizes some inspiring and pioneering copy-move forgery detection algorithms.

Block-based algorithms These algorithms seek dependence between the image original area and the pasted one, by dividing the image into overlapping blocks and then applying a feature extraction process in order to represent the image blocks through a low-dimensional representation. Different block-based representations have been previously proposed in the literature such as FT, DCT, DWT, SVD, PCA, histogram, and geometric moment etc. [11]. Bravo-Solorio et al. [7] proposed a new method to detect duplicated regions, which used color-dependent feature vectors to reduce the number of comparisons in the search stage, and one-dimensional

Table 1 Survey of the state-of-the-art passive copy-move forgery detection (CMFD) algorithms

Category	Methods	Authors
Block-based CMFD algorithms	Using color-dependent feature vectors and 1-D descriptors	Bravo-Solorio et al. [7]
	Using Zernike moments of rectangular image blocks	Ryu et al. [33]
	Using multiresolution local binary patterns	Davarzani et al. [17]
	Using fast approximate nearest-neighbor search	Cozzolino et al. [14]
	Multi-level descriptor and hierarchical feature matching	Bi et al. [6]
	Using approximate 2D-DWT coefficients	Fattah et al. [18]
	Using Krawtchouk moments	Imamoglu et al. [19]
	Using histogram of orientated gradients	Lee et al. [25]
	Using dual tree complex wavelet transform	Wu et al. [38]
	Using DCT and SVD	Jie et al. [21]
	Using 1D-Fourier transform	Ketenci et al. [23]
	Using the histogram of orientated Gabor magnitude	Lee et al. [24]
	Using undecimated dyadic wavelet transform	Muhammad et al. [30]
	Using rotation-invariant features computed densely	Cozzolino et al. [15]
	Using Zernike-moments	Al-Qershi et al. [1]
Keypoint-based CMFD algorithms	Using SIFT features	Amerini et al. [3]
	Using multi-scale analysis and voting process	Silva et al. [34]
	Using multiscale Harris operator and MPEG-7 descriptor	Kakar et al. [22]
	Using the J-Linkage clustering algorithm	Amerini et al. [2]
	Two-stage feature detection Using MROGH and HH	Yu et al. [41]
	Using local warping algorithms	Caldelli et al. [8]
	By matching triangles of keypoints	Ardizzone et al. [4]
	Using distribution of SIFT keypoints	Costanzo et al. [13]
	Using segmentation and keypoints extraction	Li et al. [26]
	Using Harris corner detector	Chen et al. [10]
	Using color SIFT	Ustubioglu et al. [36]
	Using adaptive oversegmentation and feature point matching	Pun et al. [31]
	Using mirror reflection invariant feature transform (MIFT) features	Jaberi et al. [20]

(1-D) descriptors to perform an efficient search in terms of memory usage. Ryu et al. [33] proposed a forensic technique to localize duplicated image regions based on Zernike moments of small image blocks. The rotation invariance properties was exploited to reliably unveil duplicated regions after arbitrary rotations, and a novel block matching procedure was devised based on locality sensitive hashing. Davarzani et al. [17] presented an efficient method for copy-move forgery detection using multiresolution local binary patterns (MLBP). The proposed method is robust to geometric distortions and illumination variations of duplicated regions. Furthermore, the proposed block-based method recovers parameters of the geometric transformations. Cozzolino et al. [14] proposed a new algorithm for the accurate detection and localization of copy-move forgeries, based on rotation-invariant features computed densely on the image. Here, the PatchMatch algorithm was used to compute efficiently a high-quality approximate nearest neighbor field for the whole image. Bi et al. [6] proposed a multi-level dense descriptor (MLDD) extraction method and a hierarchical feature matching method to detect copy-move forgery in digital images. The MLDD extraction method extracts the dense feature descriptors using multiple levels, while the extracted dense descriptor consists of two parts: the color texture descriptor and the invariant moment descriptor. After calculating the MLDD for each pixel, the hierarchical feature matching method subsequently detects forgery regions in the input image. Fattah et al. [18] developed a copy-move image forgery detection scheme based on a block matching algorithm. Instead of considering spatial blocks, 2D-DWT is performed on the forged image and then DWT domain blocks are considered, where only approximate DWT coefficients are utilized. Imamoglu et al. [19] used Krawtchouk moments to extract features of non overlapping image blocks. For each block, the Krawtchouk moments of order (n + m) are calculated to form the feature vector. Then, blocks' similarities are tested by inspecting the lexicographically sorted array of features. Lee et al. [25] proposed a blind forensics approach to the detection of copy-move forgery. The input image is segmented into overlapping blocks, whereupon a histogram of orientated gradients is applied to each block. Statistical features are extracted and reduced to facilitate the measurement of similarity. Finally, feature vectors are lexicographically sorted, and duplicated image blocks are detected by identifying similar block pairs after post-processing. Wu et al. [38] presented a new forensic method to detect the replicated areas rotated by arbitrary angles, even by JPEG compression. To achieve this, overlapping blocks of pixels are decomposed using dual tree complex wavelet transform (DTCWT), and then channel energies are extracted from each subband at each decomposition level using the L1 norm. Finally, the anisotropic rotationally invariant features are extracted using magnitudes of discrete Fourier transform for these channel energies. Jie et al. [21] proposed a robust copy-move forgery detection method based on DCT and SVD. Firstly, the suspicious image is divided into fixed-size overlapping blocks and 2D-DCT is applied to each block, then the DCT coefficients are quantized by a quantization matrix to obtain a more robust representation of each block. Secondly, each quantized block is divided nonoverlapping sub-blocks and SVD is applied to each sub-block, then features are extracted to reduce the dimension of each block using its largest singular value. Finally, the feature vectors are lexicographically sorted, and duplicated image blocks will be matched by predefined shift frequency threshold. Ketenci et al. [23] presented a copy move forgery detection technique that uses 1D-Fourier Transform (FT) for feature extraction. Each block is transformed into frequency domain using 1D-FT over the rows. Average values of FT coefficients along the columns are calculated to extract the feature vectors from the blocks. Lee et al. [24] proposed a scheme for detecting instances of copy-move forgery and authenticating images based on the Gabor transform. The image is first divided into overlapping fixed-size blocks. The histogram of

orientated Gabor magnitude (HOGM) descriptor is then applied to each block for the extraction of local features. Finally, each feature vector is lexicographically sorted, and regions of image forgery are detected through the identification of similar block pairs. Muhammad et al. [30] proposed a blind method for copy move image forgery detection using undecimated dyadic wavelets, in which both the LL1 and HH1 subbands are utilized to find similarities and dissimilarities between the blocks of an image. Cozzolino et al. [15] proposed a new algorithm for the accurate detection and localization of copy-move forgeries, based on rotation-invariant features computed densely on the image. Here, a fast approximate nearest-neighbor search algorithm is utilized to compute the dense descriptor. Al-Qershi et al. [1] proposed an enhanced matching method that can be used to detect copy-move forgery based on Zernike-moments. By dividing the blocks into buckets and adopting relative error instead of Euclidean distance, the proposed method enhanced the detection accuracy significantly.

Keypoint-based algorithms As an alternative to the block-based methods, keypoint-based forgery detection methods were proposed, where image keypoints are extracted and matched over the whole image to resist some image transformations while identifying duplicated regions. Some works have recently appeared on copy-move forgery detection based on Scale Invariant Features Transform (SIFT) or Speed up Robust Feature (SURF) features. Amerini et al. [3] proposed a novel methodology to support image forensics investigation based on SIFT features, which are used to robustly detect and describe clusters of points belonging to cloned areas. After detection, these points are exploited to reconstruct the parameters of the geometric transformation. By using multi-scale analysis and voting processes of a digital image, Silva et al. [34] proposed a new approach to detect copy-move forgeries in digital images that is focused mainly on investigating and spotting out traces of copy-move forgeries aided by complex operation. Kakar et al. [22] proposed a copy-move forgery detection technique using multiscale Harris operator and MPEG-7 image signature tools. They also used a feature matching process that utilizes the inherent constraints in matched feature pairs to improve the detection of cloned regions. Amerini et al. [2] presented a novel approach for copy-move forgery detection and localization based on the J-Linkage algorithm, which performs a robust clustering after SIFT matching in the space of the geometric transformation. Yu et al. [41] proposed two-stage feature detection to obtain better feature coverage and enhance the matching performance by combining the multi-support region order-based gradient histogram (MROGH) and hue histogram (HH) descriptor. Caldelli et al. [8] investigated the effectiveness of some methodologies which introduce a local warping onto the copy-pasted patches in order to reduce the detection capability of SIFT-based approaches. They compared four different local warping algorithms in terms of removed matches after the attack and visual quality of the attacked patches. Ardizzone et al. [4] presented a very novel hybrid approach, which compares triangles rather than blocks, or single points. Interest points are extracted from the image and objects are modeled as a set of connected triangles built onto these points. Triangles are matched according to their shapes (inner angles), their content (color information), and the local feature vectors extracted onto the vertices of the triangles. Costanzo et al. [13] proposed three novel forensic detectors for the identification of images whose SIFT keypoints have been artificially removed and possibly reinserted. The proposed algorithms scan image regions with sufficiently high variance in search of suspect inconsistencies in the number and in the distribution of SIFT keypoints. By relying on such algorithms, the forensic analyst can decide on the authenticity of the image as a whole or localize tampered regions within the image by means of a sliding window approach. Li et al. [26] proposed a scheme to

detect the copy-move forgery in an image, mainly by extracting the keypoints for comparison. The main difference to the traditional methods is that the proposed scheme first segments the test image into semantically independent patches prior to keypoint extraction. As a result, the copy-move regions can be detected by matching between these patches. Chen et al. [10] proposed an effective method to detect region duplication based on the image interest points detected through the Harris corner detector. After the interest points are obtained, a rotation-robust image region description method based on step sector statistics is proposed to give a unique representation for each small circle region around the interest points. Then the matching of the representations of the interest points will reveal the duplicate regions in the forged digital images. Ustubioglu et al. [36] made the comparison between SIFT and color SIFT, and proposed a more effective forgery detection by obtaining more matching points using color SIFT than SIFT. Pun et al. [31] proposed a new copy-move forgery detection scheme using adaptive oversegmentation and feature point matching. Jaberi et al. [20] adopted keypoint-based features for copy-move image forgery detection, which employing a new set of keypoint-based features, called MIFT, for finding similar regions in an image. To estimate the affine transformation between similar regions more accurately, an iterative scheme was proposed which refines the affine transformation parameter by finding more keypoint matches incrementally.

3 Improved SURF detector based on probability density gradient

In recent years, some feature points detectors, including Harris-Laplace, SIFT, and SURF operators etc., have been proposed and applied in object recognition and image processing [5, 27]. However, these state-of-the-art feature points detectors are based on intensity gradient, which cannot describe efficiently image texture information. So, they can usually detect the feature points at image textured portions or on the image edges. In other words, these detectors cannot always extract image feature points from smooth area. Besides, they are not strongly resistant to some common image processing operations and geometric transformation, which prevents these methods from being extended to detect the postprocessed duplicated regions.

In this paper, we will extract the stable and uniform image keypoints by utilizing improved SURF detector, in which the probability density gradient is considered.

3.1 Introduction to SURF detector

Recently, the SURF image feature points detector has appeared as an alternative to SIFT [5]. Its main advantage is its fastest computation, while keeping a high descriptive power (including repeatability, robustness, and distinctiveness). It is partially inspired by SIFT, but instead of using the gradient image, SURF algorithm [5] use the determinant of the approximate Hessian matrix as the base of the feature points detector. To locate the feature point, the blob-like structures are detected at locations where the determinant is at maximum. Integral images are used in Hessian matrix approximation, which reduce computation time drastically.

Given a point $\mathbf{x} = (x, y)^T$ in an image I, the Hessian matrix $H(\mathbf{x}; \sigma)$ in $\mathbf{x}$ at scale σ is defined as follows

$$H(\mathbf{x}, \sigma) = \begin{bmatrix} L_{xx}(\mathbf{x}, \sigma) & L_{xy}(\mathbf{x}, \sigma) \\ L_{xy}(\mathbf{x}, \sigma) & L_{yy}(\mathbf{x}, \sigma) \end{bmatrix} \quad (1)$$

where $L_{xx}(\mathbf{x}, \sigma)$ is the convolution of the Gaussian second order derivative $\partial^2 / \partial x^2 g(\sigma)$ with the image I in point $\mathbf{x}$, $g(\sigma) = \frac{1}{2\pi\sigma^2} e^{-\left(x^2+y^2\right)/2\sigma^2}$, and similarly for $L_{xy}(\mathbf{x}, \sigma)$ and $L_{yy}(\mathbf{x}, \sigma)$.

To reduce computation time, the convolution of 9×9 box filters, $D_{xx}(\mathbf{x}, \sigma)$, $D_{yy}(\mathbf{x}, \sigma)$, and $D_{xy}(\mathbf{x}, \sigma)$, are introduced to compute the approximations of Hessian's determinant

$$\det\left(H_{approx}\right) = D_{xx}D_{yy} - \left(0.9D_{xy}\right)^2 \tag{2}$$

Then, a pyramid scale-space is constructed by utilizing box filters and integral images, and a non-maximum suppression in a $3 \times 3 \times 3$ neighborhood is applied to localize candidate feature points in the image and over scales. Finally, image feature points can be obtained by removing the candidate feature points when their approximate determinants of Hessian matrix are smaller than the Hessian response threshold T.

3.2 Probability density gradient computation

The SURF detector serves as a powerful feature points extraction tool. For SURF detector, an image is regarded as a two-dimensional intensity function, and the local image regions are classified into three categories: 1) uniform region where the intensity gradient of each pixel is very small or nearly zero, 2) edge region where the intensity gradient magnitude of each pixel is large and the gradient direction is perpendicular to edge, and 3) textured region where the intensity gradient magnitude of each pixel is large and the gradient direction is different from each other. And then, the feature points are chosen inside the textured image regions with intensity variations in scale space. However, broad categories of real-world images have non-textured regions. It is very likely that regular SURF detector will miss stable feature points there. Furthermore, SURF detector is built by using local image intensity gradients and thus loses its distinctiveness when built on non-textured areas. In this paper, we will utilize the improved SURF detector to extract the uniform and robust feature points, in which the probability density gradient is used instead of intensity gradient.

In fact, not only can an image be regarded as a two-dimensional intensity function, but also can be described as a random distribution of pixel intensities in an even grid [16, 37]. So, we can redefine second order moment matrix by using intensity probability density gradient instead of intensity gradient, and compute the stability measure of SURF detector to extract the uniform and robust feature points.

Kernel density estimation is the most popular density estimation method [16]. The intensity probability density of each image pixel can be obtained by Kernel density estimation, and the intensity probability density $P(\mathbf{X})$ in the point $\mathbf{X}(x, y)$ is given by

$$P(\mathbf{X}) = \sum_{W} \frac{C}{h_s^2 h_r} k_1\left(\left\|\frac{\mathbf{X}-\mathbf{X}_j}{h_s}\right\|\right) k_2\left(\left\|\frac{c-c_j}{h_r}\right\|\right) \tag{3}$$

where W denotes the viewing window centered at $\mathbf{X}$, $\mathbf{X}_j$ is an image pixel in window W, c and c_j are the intensities of points $\mathbf{X}$ and $\mathbf{X}_j$ respectively, h_s denotes the employed kernel bandwidths in image space, h_r denotes the employed kernel bandwidths in intensity space, C is the corresponding normalization constant, k_1 and k_2 are the kernel functions (generally, Gaussian function), and

$$k_1 = g(\sigma) = \frac{1}{2\pi\sigma^2} e^{-\left(x^2+y^2\right)/2\sigma^2} \quad k_2 = g_t(\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-t^2/2\sigma^2}$$

From the Eq. (3), we can derive the intensity probability density gradient estimator $G(x, y)$ of image pixel $\mathbf{X}$

$$G(x,y) = \nabla_s P(x,y) = \left(P_{xx}, P_{yy}\right) \tag{4}$$

$$P_{xx} = \frac{d^2P(\mathbf{X})}{d\mathbf{x}^2} = \sum_W \frac{C}{h_s^2 h_r} \frac{d^2}{d\mathbf{x}^2} \left(k_1 \left(\left\| \frac{\mathbf{X}-\mathbf{X}_j}{h_s} \right\| \right) k_2 \left(\left\| \frac{c-c_j}{h_r} \right\| \right) \right) \tag{5}$$

$$P_{yy} = \frac{d^2P(\mathbf{X})}{d\mathrm{y}^2} = \sum_W \frac{C}{h_s^2 h_r} \frac{d^2}{d\mathrm{y}^2} \left(k_1 \left(\left\| \frac{\mathbf{X}-\mathbf{X}_j}{h_s} \right\| \right) k_2 \left(\left\| \frac{c-c_j}{h_r} \right\| \right) \right) \tag{6}$$

3.3 Improved SURF detector based on probability density gradient

The SURF detector is a powerful feature points extraction tool. It has been proven that feature points extracted by SURF detector are invariant to image rotation, scaling, viewpoint changes, and JPEG compression. In this paper, we will extract the stable and uniform image feature points by utilizing improved SURF detector, in which the probability density gradient is considered.

The new image feature point detector contains three key steps [37]: (1) Compute the probability density gradient of the image, (2) Compute the new second order moment matrix (Hessian matrix H) by using the probability density gradient of the image, (3) Detect the feature points by using SURF detector with the new second order moment matrix.

Figure 1 shows the image feature points detection results by using SURF detector and our probability density-based SURF detector. Obviously, for our probability density-based SURF

Fig. 1 The feature point detection results using SURF detector and our probability density-based SURF detector (Lena): **a** SURF detector and no attack, **b** SURF detector and JPEG compression (QF = 50), **c** SURF detector and image blurring (1.0), **d** SURF detector and viewpoint change (10°), **e** Probability density-based SURF detector and no attack, **f** Probability density-based SURF detector and JPEG compression (QF = 50), **g** Probability density-based SURF detector and image blurring (1.0), **h** Probability density-based SURF detector and viewpoint change (10°). Note: *yellow circle* denotes the repeated feature points, *red plus* denotes the detected useful feature points

detector, the feature points are chosen inside both the textured regions and non-textured areas, so the feature points are distributed evenly. Additionally, our probability density-based SURF detector is more robust than SURF detection methods under various attacks.

4 Introduction to exponent moments

In 2011, Meng & Ping [29] extended radial harmonic Fourier moments and introduced a new moment named Exponent moments (EMs). Compared with other orthogonal moment, EMs has many desirable properties such as better image reconstruction and lower computational complexity. The EMs modulus has good robustness against various noises, and geometric transformations. Besides, the EMs is free of numerical instability issues so that high order moments can be obtained accurately. So, the EMs modulus is suitable for descripting image content, especially for small image regions.

A function set $P_{n,m}(r, \theta)$ defined in a polar coordinate system (r, θ) contains the radial function $A_n(r)$ and Fourier factor in angle direction $\exp(jm\theta)$

$$P_{n,m}(r,\theta) = A_n(r)\exp(jm\theta) \tag{7}$$

where $A_n(r) = \sqrt{2/r}\exp(j2n\pi r)$, $n, m = -\infty, \cdots, 0, \cdots, +\infty$, $0 \le r \le 1$, $0 \le \theta \le 2\pi$. According to the characteristic of radial function and Fourier factor in angle direction, the set of $P_{n,m}(r, \theta)$ is orthogonal and sound over the interior of the unit circle

$$\int_0^{2\pi}\int_0^1 P_{n,m}(r,\theta)P^*_{k,l}(r,\theta)rdrd\theta = 4\pi\delta_{n,k}\delta_{m,l} \tag{8}$$

where 4π is the normalization factor, $\delta_{n,k}$ and $\delta_{m,l}$ are the Kronecker symbols, and $P^*_{k,l}(r, \theta)$ is the conjugate of $P_{k,l}(r, \theta)$.

The image $f(r, \theta)$ can be decomposed with the set of $P_{n,m}(r, \theta)$ as

$$f(r,\theta) = \sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} E_{n,m}A_n(r)\exp(jm\theta) \tag{9}$$

where $E_{n,m}$ is the EMs of order n with repetition m, whose definition is

$$E_{n,m} = \frac{1}{4\pi}\int_0^{2\pi}\int_0^1 f(r,\theta)A^*_n(r)\exp(-jm\theta)rdrd\theta \tag{10}$$

here, $A^*_n(r)$ is the conjugate of $A_n(r)$.

Following the principle of orthogonal function, the image function $f(r, \theta)$ can be reconstructed approximately by limited orders of EMs ($n \le n_{\max}$, $m \le m_{\max}$). The more orders used, the more accurate the image description

$$f'(r,\theta) = \sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} E_{n,m}A_n(r)\exp(jm\theta) \approx \sum_{n=-n_{\max}}^{n_{\max}}\sum_{m=-m_{\max}}^{m_{\max}} E_{n,m}A_n(r)\exp(jm\theta) \tag{11}$$

where $f'(r, \theta)$ is the reconstructed image. The basis functions $A_n(r)\exp(jm\theta)$ of the EMs are orthogonal over the interior of the unit circle, and each order of the EMs makes an independent contribution to the reconstruction of the image.

Below, we will derive and analyze the geometric invariant property of EMs. Let $f^r(r, \theta) = f(r, \theta + \alpha)$ denote the rotation change of an image $f(r, \theta)$ by the angle α, then EMs of $f(r,\theta + \alpha)$ and $f(r, \theta)$ have the following relations

$$E_{n,m}(f^r) = E_{n,m}(f)\exp(\mu m\alpha)$$

where $E_{n,m}(f^r)$ and $E_{n,m}(f)$ are the EMs of $f^r(r, \theta)$ and $f(r, \theta)$, respectively. According to above equation, we know that a rotation of the image by an angle α induces a phase shift $e^{\mu m\alpha}$ of the $E_{n,m}(f)$. Taking the norm on both sides, we have

$$\left|E_{n,m}(f^r)\right| = \left|E_{n,m}(f)\exp(\mu m\alpha)\right| = \left|E_{n,m}(f)\right|\left|\exp(\mu m\alpha)\right| = \left|E_{n,m}(f)\right|$$

So, the rotation invariance can be achieved by taking the norm of the images' EMs. In other words, the EMs modulus $|E_{n,m}(f)|$ are invariant with respect to rotation transform. Besides, the EMs modulus is invariant to scaling if the computation area can be made to cover the same content. In practice, this condition is met because the EMs are defined on the unit disk.

Figure 2 gives the reconstructed images and reconstruction errors using EMs for standard image Lena (moment orders $n = 5$, 10, 15, 20, 25, 30), and Fig. 3 shows the EMs modulus distribution for image Lena under various attacks. It can be seen that, EMs can effectively capture the image information, and the reconstructed images get closer to the original images when more moments are added to the reconstruction process. Also, the EMs modulus has good robustness against various noises, geometric transformations, and image variations.

5 The proposed robust copy-move forgery detection

Figure 4 shows original image and tampered image Japan tower from image dataset FAU [9]. The proposed framework for robust copy-move forgery detection shown in Fig. 5 is carried out in four key steps here. Firstly, superpixels segmentation (See Fig. 5a). The given image is segmented into nonoverlapping and irregular superpixels by using the entropy rate superpixel (ERS) algorithm, wherein the number of superpixels is assigned adaptively according to image size and content. Secondly, superpixels classification and adaptive keypoints extraction (See Fig. 5b and c). The superpixels are classified into smooth, texture and strong texture ones based on local information entropy and the stable image keypoints are extracted from each superpixel by utilizing probability

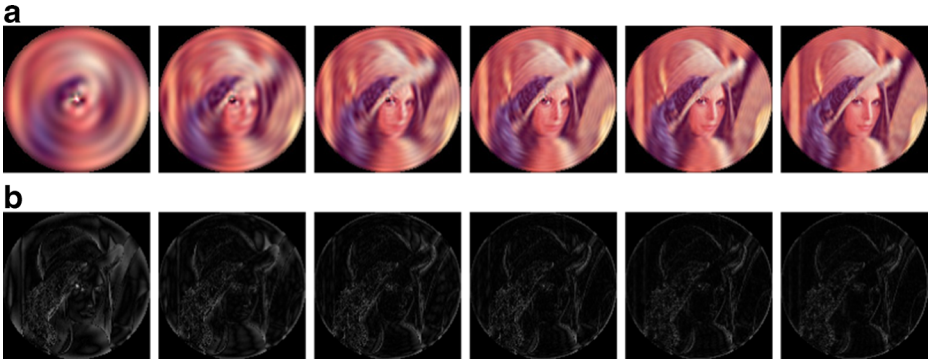

Fig. 2 The reconstructed images and reconstruction errors for image Lena of size 128×128 (moment orders $n = 5$, 10, 15, 20, 25, 30): **a** Reconstructed images, **b** Reconstruction errors

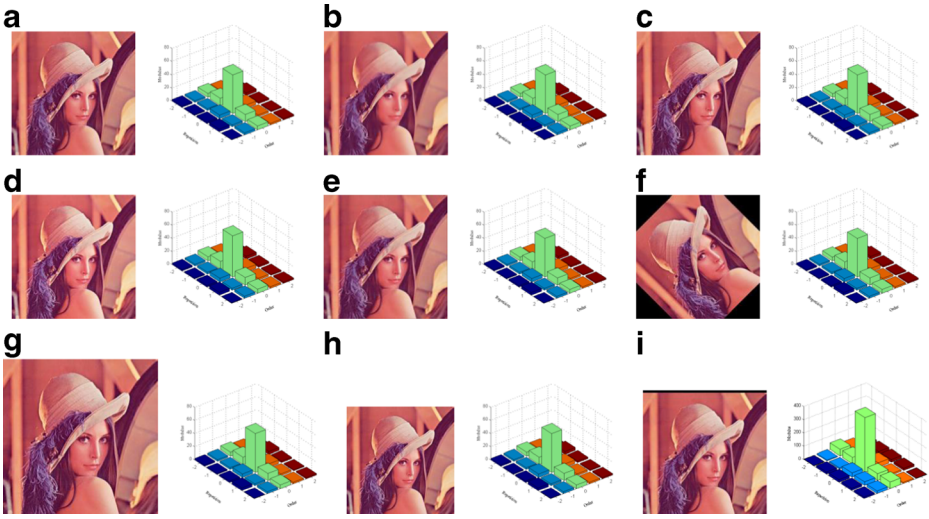

Fig. 3 The EMs modulus for image Lena under various noises, geometric transformations, and image variations: **a** Original image, **b** Blurring, **c** Sharpening, **d** Contrast changing, **e** Average filtering, **f** Rotation (45°), **g** Scaling (1.2), **h** Scaling (0.8), **i** Translation

density-based SURF detector. Thirdly, the keypoint feature construction and keypoints matching (See Fig. 5d and e). The feature vectors, namely exponent moments magnitudes, are constructed for each keypoints, and the Rg2NN algorithm is utilized to find rapidly the matching keypoints. Finally, detected copy-move forgery regions postprocessing (See Fig. 5f and h). The RANSAC algorithm is utilized to estimate the homographies corresponding to the transformations that the copied regions have undergone. Also, morphological operation is applied to generate the detected forgery regions. The following sections provide detailed steps in this proposed framework.

5.1 Superpixels segmentation

In this paper, we employed the entropy rate superpixel (ERS) algorithm [28] to segment the host image into meaningful irregular superpixels, which are further used for content based adaptive feature points detection. The ERS algorithm adapts a new objective function for superpixel segmentation. This objective function consists of two components: entropy rate of a random walk on a graph and a balancing term. The entropy rate favors compact and

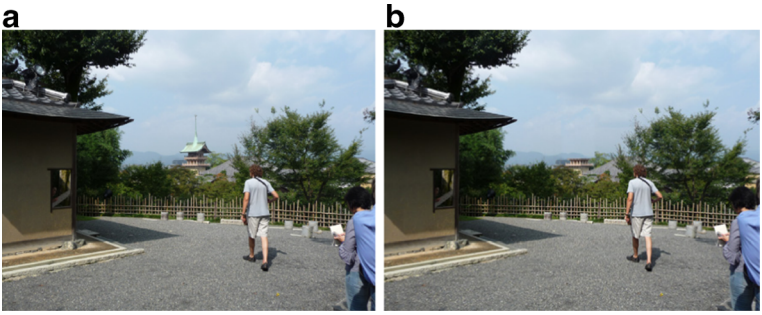

Fig. 4 Original image and tampered image (Japan tower): **a** Original image, **b** Tampered image

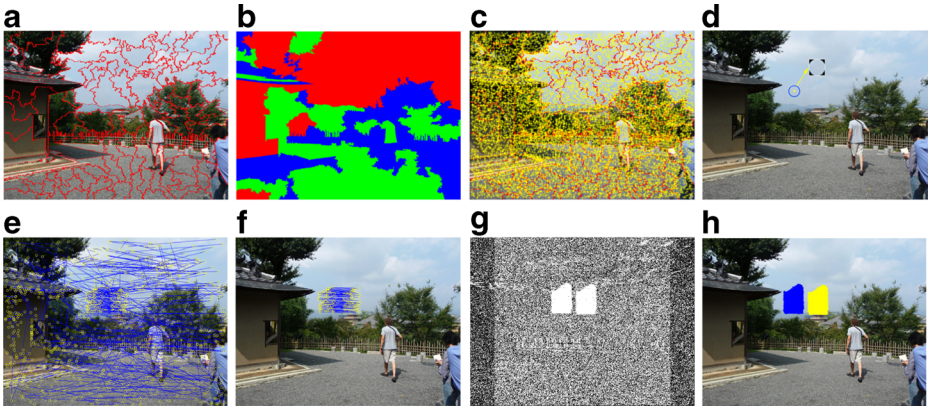

Fig. 5 The key steps of the proposed copy-move forgery detection approach: **a** Superpixels segmentation, **b** Superpixels classification (*red* denotes smooth, *green* is texture, and *blue* denotes strong texture), **c** Adaptive keypoints extraction, **d** Keypoint feature construction, **e** Keypoints matching, **f** Falsely matched keypoints removing, **g** Localizing the duplicated regions, **h** Morphology processing

homogeneous clusters-encouraging division of images on perceptual boundaries and favoring superpixels overlapping with only a single object; whereas the balancing term encourages clusters with similar sizes-reducing the number of unbalanced superpixels.

For our superpixel content based adaptive feature points detection, the selection of optimal number of the superpixels in ERS segmentation is a key component. This is because that the original tampered image and forgery regions are always of different sizes and have different content, and different number of the superpixels can usually produce different feature points detection results. For example, the smaller number of the superpixels will lead to a significantly higher computational load, but the larger one will make the feature points detection results not be sufficiently accurate. So, we should select the number of the superpixels by considering both the computational expense and the detection accuracy. Here, we assign adaptively the optimal number K of superpixels according to the tampered image size and content as follow

$$K = \frac{n_1}{n_2} \times \left\lfloor \frac{Max(M,N)}{5} \right\rfloor$$

where n_1 and n_2 denote the number of pixels with maximum gray value, and the number of all pixels in the gray version of the tampered color image, respectively. $M \times N$ denotes the host image size. $\lfloor x \rfloor$ returns the lower integer part of x.

Using the above selection method, we can obtain adaptively the optimal number of the superpixels, $K = 118$, 193, for tampered image Red tower (1632×1224) and Ship (1632×1224). Figure 6 shows the non-overlapping and irregular superpixels, which are further used for content based adaptive feature points detection.

5.2 Superpixels classification and adaptive keypoints extraction

5.2.1 Superpixels classification

In order to extract the stable and uniform image feature points from each superpixel, we must analyze the superpixel content to define the adaptive Hessian response threshold. In this paper, we firstly classify the superpixels into smooth, texture and strong texture ones according to

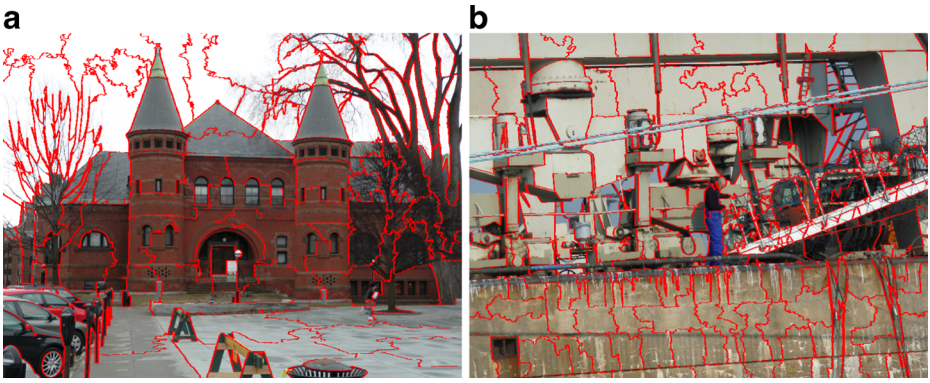

Fig. 6 Non-overlapping and irregular superpixels using ERS algorithm: **a** Red tower, **b** Ship

their local information entropy, and then define adaptively the Hessian response threshold based on superpixel content.

Suppose that the given image I have been segmented into K nonoverlapping and irregular superpixels $S_k(k=1,\cdots,K)$, and we firstly define the local information entropy E_k of superpixels S_k as

$$E_k = -\sum_i P_{S_k}(i)\log_2 P_{S_k}(i)$$

where $P_{S_k}(i)$ is the probability of the pixel value i in superpixels S_k.

Then, we formulate a superpixel type code, denoted as $B_k(k=1,\cdots,K)$, to classify the superpixels S_k into smooth, texture and strong texture ones according to their local information entropy, where

$$B_k = \begin{cases} 0 \quad (smooth) & if\ E_k < EA \\ 1 \quad (texture) & if\ EA < E_k < EB \\ 2\ (strongtexture) & if\ E_k > EB \end{cases}$$

And

$$EA = \min_k(E_k) + \frac{E}{2},\ EB = \min_k(E_k) + \frac{3E}{4},\ E = \max_k(E_k) - \min_k(E_k)$$

Finally, the adaptive Hessian response threshold for superpixel S_k feature points detection can be defined according to their local information entropy, where

$$T_k = \begin{cases} T \times M^2 \times \min\limits_{B_k=0}(S_k) & if B_k = 0 \\ T \times M^2 \times \min\limits_{B_k=1}(S_k) & if B_k = 1 \\ T \times \min\limits_{B_k=2}(S_k) & if B_k = 2 \end{cases}$$

where $T=0.0002$ is the original Hessian response threshold for SURF detector, and $M=0.01$ is a constant factor for adjusting Hessian response threshold.

5.2.2 Adaptive keypoints extraction

Robust image feature points extraction is a key component of the keypoint-based CMFD schemes. In recent years, some image feature points detectors have been proposed, and have

been widely applied in image forgery detection. For example, scale-invariant feature transform (SIFT) by Lowe et al. and speeded-up robust features (SURF) by Bay et al. However, these state-of-the-art feature points detectors are based on intensity gradient, which cannot describe efficiently image texture information. So, they can usually detect the feature points at image textured portions or on the image edges. In other words, these detectors cannot always extract image feature points from smooth area. Besides, they are not strongly resistant to some common image processing operations and geometric transformation, which prevents these methods from being extended to detect the postprocessed duplicated regions.

In this paper, we extract evenly image feature points from each superpixel using our probability density-based SURF detector and adaptive Hessian response threshold, which have been proven to be robust against common image processing operations and geometric distortion, such as rotation, scaling, and lossy compression etc. Figure 7 shows the superpixels classification and adaptive keypoints extraction procedure for image Red tower and Ship.

5.3 Keypoint feature construction and keypoints matching

5.3.1 Keypoint feature construction

Keypoint feature construction is a prerequisite step for CMFD and crucial to detection accuracy. It is desired that the keypoints in a copy-move pair can be mapped to similar features even in the presence of post-processing. At the same time, the features should correctly distinguish distinct keypoints in the image, even under some severe conditions. From the foregoing, we know that EMs can effectively capture the image contents, especially for small image regions. Besides, EMs modulus has good robustness against various noises and geometric transformations. So, we employ the EMs modulus to extract keypoint features from the local feature region in this paper. Our process for constructing the EMs based keypoint feature can be summarized as follows.

> Step 1 The local feature regions (LFRs) are constructed. Here, LFR refers to the image region around keypoints in the given image, which reflects the important image keypoint semantics. Theoretically, the LFRs can be any shapes such as triangle, rectangle, and circle etc., but it is important to assure that LFRs should be invariant to various noises and geometric transformations. In this paper, we construct the circular LFRs to extract keypoint feature, which can achieve better robustness property. Based on the fact that the feature scales of keypoints vary with the local image feature, we chose the radius of the LFRs in accordance with

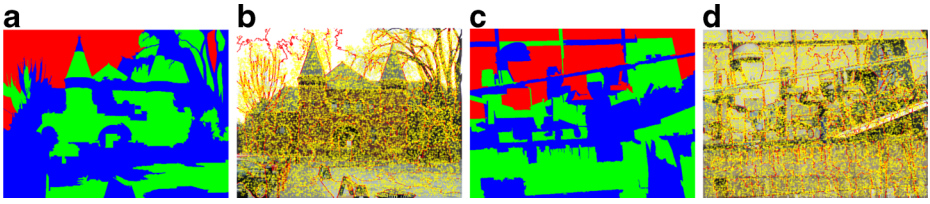

Fig. 7 The superpixels classification and adaptive keypoints extraction for image Red tower and Ship: **a** Superpixels classification (*Red tower*), **b** Adaptive keypoints extraction (*Red tower*), **c** Superpixels classification (*Ship*), **d** Adaptive keypoints extraction (*Ship*)

the feature scale of the keypoint. As a result, a set of adaptive circular regions (LFRs) are generated. The radius of the circular LFRs is defined as:

$$\Re = \tau \cdot \mathrm{round}(\delta)$$

where $\Re$ is the radius of the LFR, δ is the feature scale of the keypoint, and τ is a positive integer, which is used to adjust the size of the local image regions.
Step 2 The keypoint features, EMs modulus coefficients, are computed. In order to compute the EMs of circular LFR, we need to apply the zero-padding operation to the circular LFR so as to obtain the square LFR. Then, we can compute the EMs of the square LFR, and select EMs modulus as the keypoint feature.

According to the better reconstruction property of EMs, we know that image features can normally be captured by just a few low-frequency EMs. Further, the EMs modulus coefficients $|E_{n,-m}| = |E_{n,m}|$, and the values of $|E_{0,0}|$ and $|E_{1,0}|$ are nearly constant for all images, so we only chose some low-frequency EMs modulus coefficients as the keypoint feature $\mathbf{E}^{(i,j)}$ in this paper.

$$\mathbf{E}^{(i,j)} = \left(\left|E_{1,1}^{(i,j)}\right|, \left|E_{2,2}^{(i,j)}\right|, \left|E_{2,1}^{(i,j)}\right|, \left|E_{2,0}^{(i,j)}\right|, \left|E_{3,3}^{(i,j)}\right|, \left|E_{3,2}^{(i,j)}\right|, \left|E_{3,1}^{(i,j)}\right|, \left|E_{3,0}^{(i,j)}\right|\right)$$

5.3.2 Keypoints matching

After we have obtained the image keypoints and keypoint features, we must locate the matched keypoints through the keypoint features. Generally, we consider that two image keypoints are matched each other when their Euclidean distance of the EMs features is smaller than the predefined threshold, which means that the image keypoint (i, j) is matched to the image keypoint (k, l) only if they can meet the following condition

$$\left\|\mathbf{E}^{(i,j)} - \mathbf{E}^{(k,l)}\right\| \le D$$

where $\mathbf{E}^{(i,j)}$ and $\mathbf{E}^{(k,l)}$ denote the EMs feature vector of image keypoints (i, j) and (k, l), respectively. D is feature Euclidean distance threshold. But, due to the high-dimensionality of the feature space, the above simple matching approach usually obtains a low accuracy because some keypoint features are much more discriminative than others. For this reason, more effective nearest-neighbor matching techniques, including 2NN and g2NN, are proposed by using the ratio between the distance of the closest neighbor to that of the second-closest one, and comparing it with a threshold [40]. However, they have one main drawback: they are unable to manage multiple keypoint matching. This is a key aspect in case of copy-move forgeries since it may happen that the same image area is cloned over and over. In other words, they only find matches between keypoints whose features are very different from those of the rest of the set. Therefore, the case of cloned patches is very critical since the keypoints detected in those regions are very similar to each other. In this paper, we improve the structure of image keypoint pairs matching algorithm and propose an enhanced model based on the reversed-generalized 2 nearest-neighbor (Rg2NN) [40], in which reverse order is used in keypoints matching, so that all keypoints that match with the detected point can be calculated accurately.

Rg2NN is the improved version of g2NN algorithm, which considers the ratio between the nearest neighbor and the second nearest one. For a given keypoint, $D = \{d_n, d_{n-1}, \cdots, d_2\}$ is defined as the sorted Euclidean distances in reverse order with respect to other keypoints. The keypoint corresponding to d_i is considered a match of the given keypoint only if the ratio $T_i = {}^{d_{i-1}}/_{d_i}$ $(n \geq i \geq 2)$ is larger than a fixed threshold T_{R2gNN} and T_{i-1} is smaller than T_{R2gNN} (in our experiments we set T_{R2gNN} to 0.6). A very high ratio (e.g. greater than T_{R2gNN}) means two random features.

The Rg2NN algorithm iterates the g2NN test in reverse order between ${}^{d_{i-1}}/_{d_i}$ until this ratio is greater than T_{R2gNN}, which is able to match multiple copies. That is to say, if k is the value in which the Rg2NN procedure stops, each keypoint in correspondence to a distance in $\{d_n, d_{n-1}, \cdots, d_k\}$ (where $n \geq k \geq 2$) is regarded as a match for the inspected keypoint. After using the Rg2NN test on all the keypoints, we can obtain a set of matched keypoint pairs.

Figure 8 shows the image keypoints matching procedure for image Red tower and Ship.

5.4 Reducing false matching

After the set of matched image keypoints x_m has been obtained, it is necessary to cluster these data in such a way so as to be able to distinguish the different matched regions. For clustering on the spatial locations (i.e., x, y coordinates) of the matched keypoints, an agglomerative hierarchical clustering is used in this paper. Hierarchical clustering creates a hierarchy of clusters which may be represented by a tree structure. The algorithm starts by assigning each keypoint to a cluster; then it computes all the reciprocal spatial distances among clusters, finds the closest pair of clusters, and finally merges them into a single cluster. Such computation is iteratively repeated until a final merging situation is achieved. Considering that we have at least two matched areas, the result of agglomerative hierarchical clustering provides $P \geq 2$ different sets of matched keypoints M_p, so $x_m = M_1 \cup M_2 \cup \cdots \cup M_p$, and this allows the definition of the different duplicated regions.

When an image has been classified as nonauthentic, we utilize the RANSAC algorithm to determine which geometrical transformation was used between the original area and its copy-moved version. Besides, another benefit is that we can further filter out mismatches through this procedure. RANSAC algorithm was first introduced by Fischler and Bolles [2], which is a simple, yet powerful parameter estimation approach designed to cope with a large proportion of outliers in the input data. In essence, RANSAC is a resampling technique that generates candidate solutions using the minimum number data points required to estimate the underlying model parameters. This algorithm estimates a global relation that fits the data, while

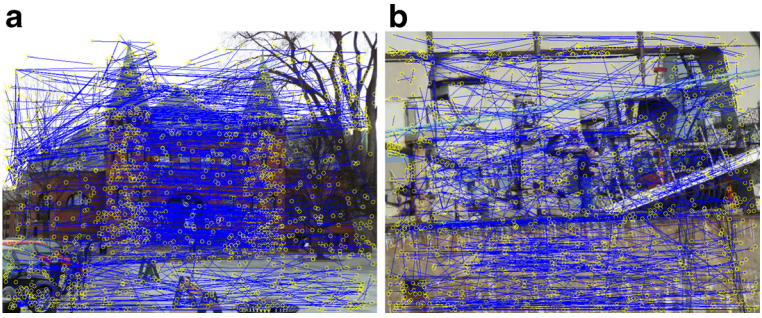

Fig. 8 The image keypoints matching results: **a** Red tower, **b** Ship

simultaneously classifying the data into inliers and outliers. Due to its ability to tolerate a large fraction of outliers, RANSAC is a popular choice for a variety of robust estimation problems.

Denoting the set of all Rg2NN-matched image keypoints pairs as P, and a reduced set P^* is constructed from P by keeping only those pairings for which at least one spatially adjacent image keypoint pair is also included in P. Assuming that a pair of matched image keypoints are (i, j) and (k, l), we select three spatially adjacent collinear pairs from P^* to infer their 2×2 affine transformation $\mathbf{R}$ in the spatial domain,

$$(i,j)^T = \mathbf{R}\cdot(k,l)^T + \mathbf{t}$$
$$\mathbf{R} = \begin{pmatrix} s_x & 0 \\ 0 & s_y \end{pmatrix} \times \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}, \mathbf{t} = \begin{pmatrix} t_x \\ t_y \end{pmatrix}$$

where $\begin{pmatrix} t_x \\ t_y \end{pmatrix}$, $\begin{pmatrix} s_x & 0 \\ 0 & s_y \end{pmatrix}$, and $\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$ are shift vector, scaling matrix and rotation matrix, respectively.

All image keypoints pairs in P^* are classified into inliers or outliers by checking the condition

$$\left\| (i,j)^T = \mathbf{R}\cdot(k,l)^T + \mathbf{t} \right\| < T$$

for classification threshold T. This procedure is repeated N_{iter} times, each time initialized with a triple of keypoints pairs randomly drawn from set P^*. The algorithm outputs the set of pairings with the largest number of inliers as duplicated region. In this paper, the RANSAC parameters are set to classification threshold $T = 2$ and iteration times $N_{iter} = 100$, which resemble literature settings [5].

Figure 9 shows the false matching removal result for image Red tower and Ship.

5.5 Localizing duplicated regions

In order to localize the duplicated regions, we use a block-wise correlation measure based on zero mean normalized cross-correlation (ZNCC) between the gray-scale of the original image and the warped image [2]. It is computed as

$$ZNCC(\mathbf{x}) = \frac{\sum_{\mathbf{v}\in\Omega(\mathbf{x})} \left(I(\mathbf{v})-\bar{I}\right)\left(W(\mathbf{v})-\overline{W}\right)}{\sqrt{\sum_{\mathbf{v}\in\Omega(\mathbf{x})} \left(I(\mathbf{v})-\bar{I}\right)^2 \left(W(\mathbf{v})-\overline{W}\right)^2}}$$

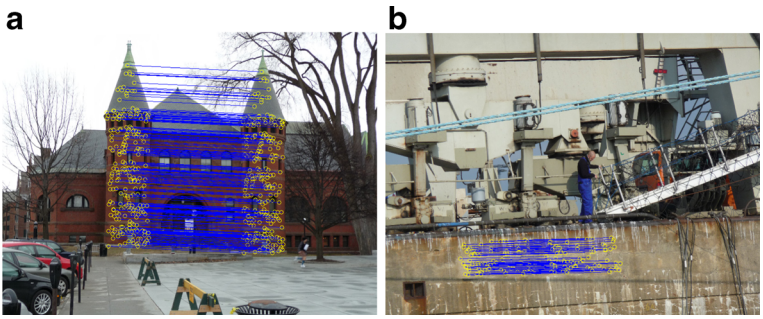

Fig. 9 The false matching removal results: **a** Red tower, **b** Ship

where $\Omega(\mathbf{x})$ is a 7 pixels neighboring area centered at every pixel $\mathbf{x}$ of I; $I(\mathbf{v})$ and $W(\mathbf{v})$ denote the pixel intensities at the location $\mathbf{v}$; $\bar{I}$ and $\overline{W}$ are the average pixel intensities of I and W, respectively, computed on $\Omega(\mathbf{x})$. Once the correlation map is obtained, we apply a Gaussian filter of size 7 pixels with standard deviation 0.5 in order to reduce the noise.

After the above postprocessing to reduce missing detections, we can construct a duplication map to visualize the forgery detection result. We firstly create an all-zeros matrix M that has the same size as the tampered image, and set the entry $M_{i,j}$ to one if the coordinate (i, j) in the image is covered by a copy-move pair. Then, we use a forged area threshold A, a value denoting the minimum area of the duplicated region, to remove small isolated regions. Finally, we use mathematical morphological operations to smooth and connect the boundaries of the detected duplicated regions, and the duplication map can be obtained by multiplying the binary matrix with the tampered image. Figure 10 shows the duplicated regions localizing, morphological operations, and final forgery detection results for image Red tower and Ship.

6 Simulation results

In this section, we present the results of our proposed copy-move forgery detection approach. We present quantitative results and examples for the detection of copy-move forgeries subjected to many image processing operations. Also, experimental results are compared with schemes in [2, 14, 15]. All measurements are performed on a desktop computer with Dual Core 3.4-GHz Pentium CPU and 8 GB RAM memory, running Matlab R2011a.

6.1 Test image dataset and error measures

In this paper, the public available image dataset FAU is used to evaluate the performance of different CMFD schemes, which was constructed by Christlein et al. [9]. The image dataset is formed based on 48 high-resolution uncompressed PNG true color images. In the dataset, the

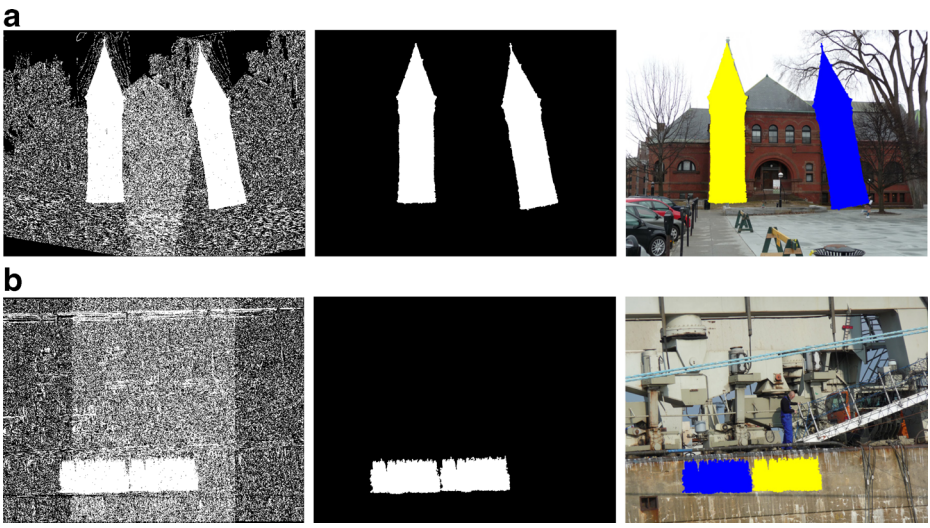

Fig. 10 The duplicated regions localizing, morphological operations, and final forgery detection results: **a** Red tower, **b** Ship

copied regions are from the categories of living, nature, man-made and mixed, and they range from overly smooth to highly textured; the copy-move forgeries are created by copying, scaling and rotating semantically meaningful image regions. Table 2 shows the assignment of images to categories. In summary, the dataset has 1826 images in total, which are realistic copy-move forgeries. These images are quite large, with typical size 3000×2400 pixels, with tampered areas covering about 6 % of each image on average. We also perform experiments over 80 images from GRIP image database [15]. All these images have size 768×1024 pixels, while the forgeries have arbitrary shapes aimed at obtaining visually satisfactory results, with size going from about 4000 pixel (less than 1 % of the image) to about 50000 pixels.

For practical use, the most important aspect is the ability to distinguish tampered and original images. However, the power of an algorithm to correctly annotate the tampered region is also significant, especially when a human expert is visually inspecting a possible forgery. In this work, we evaluate the CMFD performance at two levels: at image level, where we focus on whether the fact that an image has been tampered or not can be detected; at pixel level, where we evaluate how accurately tampered regions can be identified. At image level, the important measures are the number of correctly detected forged images, T_p, the number of images that have been erroneously detected as forged, F_p, and the falsely missed forged images F_N. From these we computed the error measures *P*recision and *R*ecall. They are defined as

$$\textit{P}\text{recision} = \frac{T_P}{T_P + F_P}, \text{ and } \textit{R}\text{ecall} = \frac{T_P}{T_P + F_N}$$

*P*recision denotes the probability that a detected forgery is truly a forgery, while *R*ecall shows the probability that a forged image is detected. *R*ecall is often called true positive rate. We also compute another criterion F as a measure which combines *P*recision and *R*ecall in a single value.

$$F = 2 \cdot \frac{\textit{P}\text{recision} \cdot \textit{R}\text{ecall}}{\textit{P}\text{recision} + \textit{R}\text{ecall}} \tag{12}$$

We used these measures at pixel level, too. In that case, *Tp* is the number of correctly detected forged pixels. *Fp* denotes the number of falsely detected forged pixels and F_N are the falsely missed pixels. The previous definition of *P*recision, *R*ecall and F measures also hold on the pixel level.

In this paper, we evaluate the proposed CMFD method at the pixel level and image level simultaneously. This is because that, the pixel-level metrics are useful for assessing the general

Table 2 Categorization of Christlein's FAU dataset by texture properties

Category	Assigned images	
	Small copied area	Large copied area
Smooth	ship, motorcycle, sailing, disconnected s2hift, noise pattern, berries, sails, mask, cattle, swan, Japan tower, wading	four babies, Scotland, hedge, tapestry, Malawi
Rough	supermarket, no beach, fisherman, barrier, three hundred, writing history, central park	lone cat, kore, white, clean walls, tree, christmash edge, stone ghost, beach wood, red tower
Structured	bricks, statue, giraffe, dark and bright, sweets, Mykene, jellyfish chaos, Egyptian, window, knight moves	fountain, horses, port, wood carvings, extension

localization performance of the algorithm when the ground-truth data are available, and the image-level decisions are especially interesting with respect to the automated detection of manipulated images. In general, a higher precision, a higher recall, and a higher F-measure indicate superior performance.

6.2 Detection results at pixel level

The goal of this experiment is to evaluate how precisely a copy-moved region can be marked. So, we focused mainly on the number of detected (or missed, respectively) copied-moved matches. For each detected match, we check the centers of two matched blocks against the corresponding (pixelwise) ground truth image. All boundary pixels are excluded from the evaluation. Please note that all the measures, e. g. false positives and false negatives, are computed using all the pixels in the tampered images only. In the practical test, we first evaluate the proposed scheme under ideal conditions (no postprocessing) of the pixels. Subsequent experiments examine the cases of: JPEG compression on the copied region, noise on the copied region, rotation and scaling of the copied region.

1) Plain Copy-Move: As a baseline, we evaluated how the CMFD methods perform under ideal conditions (no postprocessing). Here, we used the 128 forgery images from FAU and GRIP dataset without any additional modification to evaluate the detection performance. Note that we calibrated the thresholds for all methods in a way that yields very competitive (comparable) detection performances.
2) JPEG Compression: We introduced a common local disturbance on the copied region, lossy JPEG compression, for which the quality factors varied between 30 and 100 in steps of 10. Per evaluated quality level, we applied the same JPEG compression to the copied regions of the 128 forgeries images, so a total of $128 \times 8 = 1024$ forgeries images are tested. For very low quality factors, the visual quality of the image is strongly affected. However, we consider at least quality levels down to 70 as reasonable assumptions for real-world forgeries.
3) Additive White Gaussian Noise (AWGN): We applied additive white Gaussian noise to the copied regions of the 128 forgeries images from FAU and GRIP dataset. The strength of distortion is parametrized via the noise's standard deviation and the filter's radius. We considered zero-mean AWGN with standard deviation 2, 4, 6, and 8, as well as average filters of radius 0.5, 1, 1.5, 2, and 2.5, respectively.
4) Rotation: One question that recently gained attention was the resilience of CMFD algorithms to affine transformations, like scaling and rotation. We rotated the copied regions with the rotation angle varying from 2 to 10°, in increments of 2°, and with the rotation angles of 20°, 60° and 180° as well. In this case, we tested a total of $128 \times 8 = 1024$ forgeries images.
5) Scaling: We conducted an experiment where the copied region was slightly rescaled, as is often the case in real-world image manipulations. Specifically, we rescaled the copied region between 91 and 109 % of its original size, in steps of 2 %. We also evaluated rescaling by 50, 80, 120 and 200 % to test the degradation of algorithms under larger amounts of copied region resizing.

Figures 11 and 12 shows the detection results on some test images from FAU dataset [9] with copy-move regions fused in the background, and Fig. 13 shows the detection results on some test images from GRIP database [15]. It can be observed that the proposed scheme detects most copy-move regions (But the scheme [15] cannot detect correctly some forged images, such as tree, giraffe

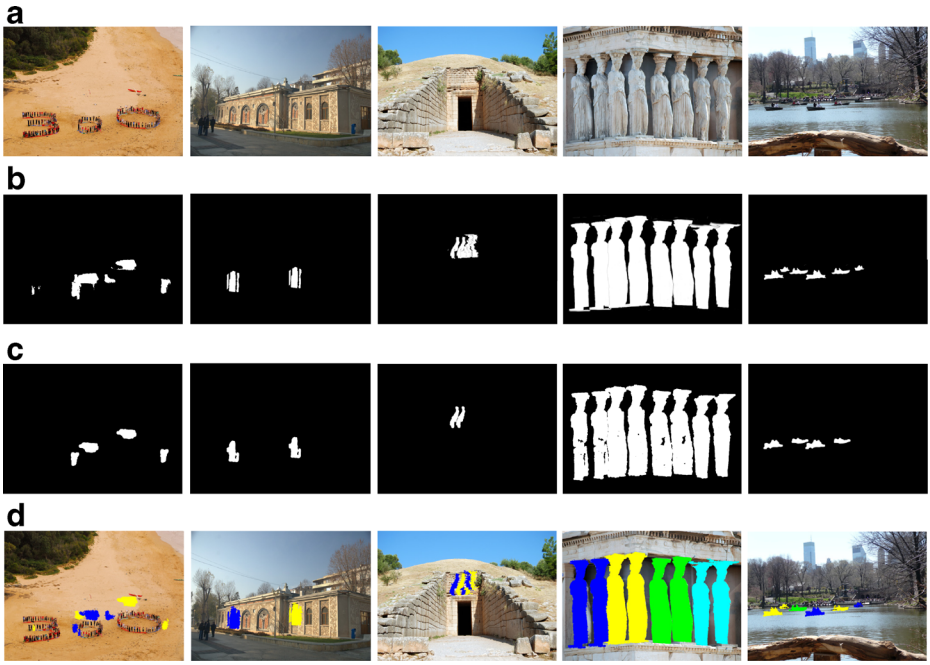

Fig. 11 Detection results on some test images from FAU dataset: **a** Forged images, **b** Ground-truth forgery regions, **c** Detected forgery regions using Cozzolino's method (2015), **d** Detected forgery regions using our approach

from FAU dataset). Figure 14 shows the average CMFD performance at pixel level under various attacks on the copied region, including common signal processing and geometric distortions.

Table 3 shows the average F-measure of 128 test images under ideal conditions (no postprocessing), which combine both the precision and recall into a single value, for the proposed scheme compared with the existing methods. This table indicates that the forgery detection result of the proposed scheme is better than that of the existing state-of-the-art methods when under plain copy-move. Table 4 reports the run times of the proposed copy-move forgery detection, including the CPU-time of each key steps and the whole approach, for color images with different size.

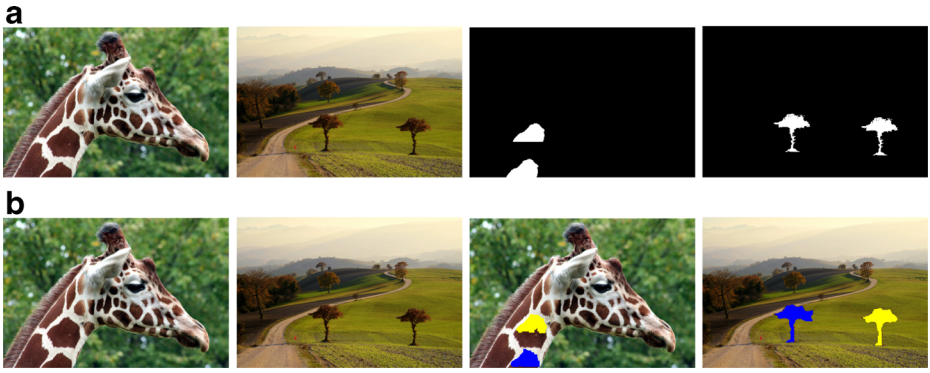

Fig. 12 Detection results on some test images from FAU dataset: **a** Forged images, **b** Ground-truth forgery regions, **c** Detected forgery regions using Cozzolino's method (2015), **d** Detected forgery regions using our approach

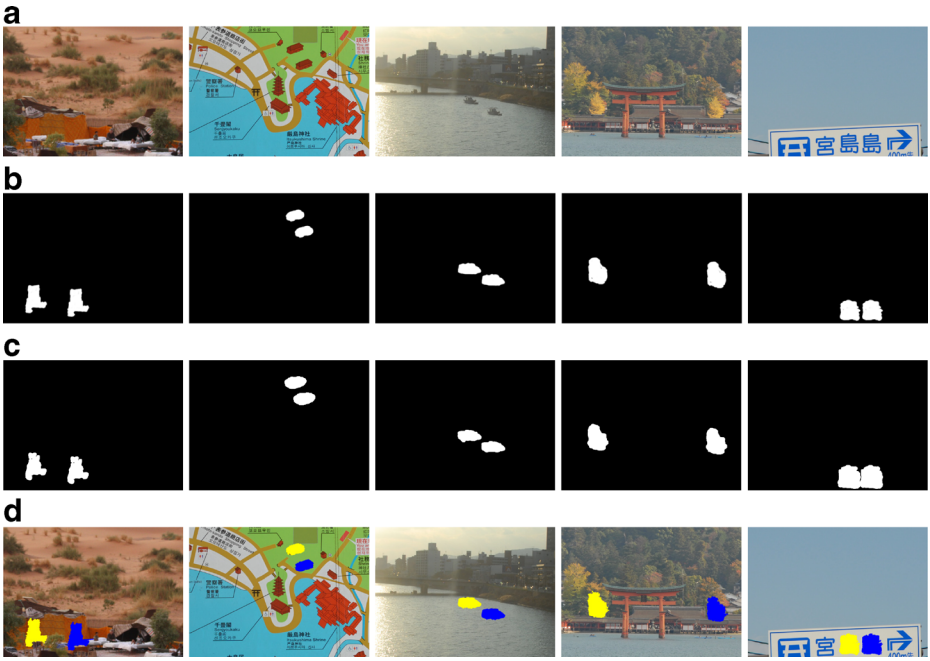

Fig. 13 Detection results on some test images from GRIP database: **a** Forged tampered images, **b** Ground-truth forgery regions, **c** Detected forgery regions using Cozzolino's method (2015), **d** Detected forgery regions using our approach

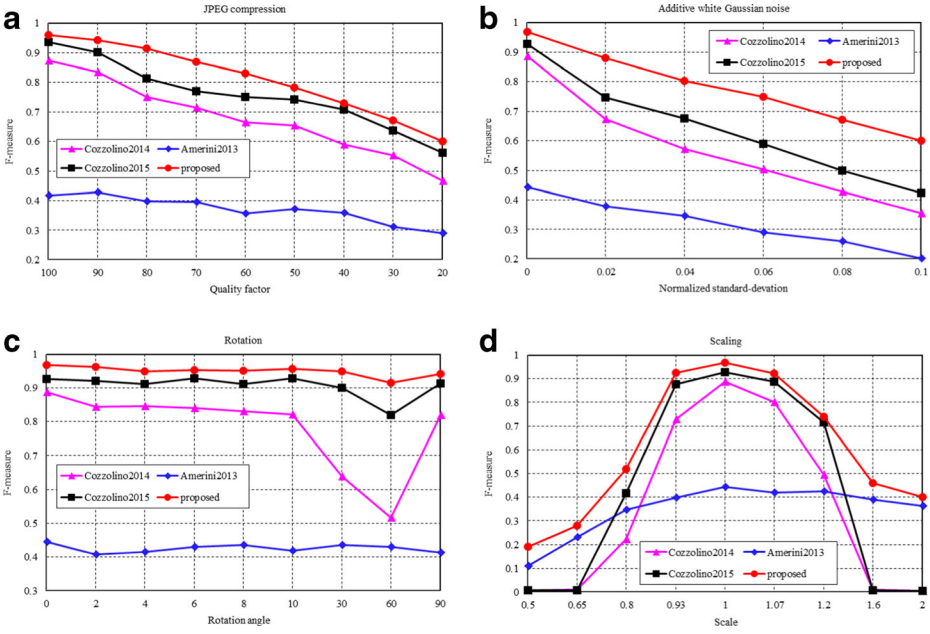

Fig. 14 Average CMFD performance (F-measure) at pixel level on 80 test images from GRIP dataset: **a** JPEG compression, **b** Additive white Gaussian noise, **c** Rotation, **d** Scaling

Table 3 Average detection results under plain copy-move at pixel level and image level

Methods	F-measure image level	F-measure pixel level
Cozzolino [14]	0.9470	0.8740
Amerini [2]	0.9304	0.7973
Li [26]	0.9489	0.8909
Cozzolino [15]	0.9426	0.8785
Our scheme	0.9680	0.9226

6.3 Detection results at image level

Area based pixel level performance measures such as *P*recision, *R*ecall, and *F* rates are useful when we know that the tested image is a forgery. Yet, in practice, this is usually not known a priori. In the next set of experiments, we test the overall image-level detection performance of our method. Specifically, for a forgery image, a successful detection is deemed when our method detects a duplicated region larger than the area threshold. For an untampered image, a true negative occurs when our method does not detect any duplicated region.

We split these image level experiments in a series of separate evaluations. We first evaluate the proposed approach under ideal conditions (plain copy-move); in other words, we have 128 original images and 128 forgery images, in which a one-to-one copy-move is implemented. We must distinguish the original and forgery images in this case. Then, the proposed copy-move forgery approach is evaluated under different types of attacks: the geometric transforms, such as scaling and rotation; common signal processing, such as JPEG compression and additive white Gaussian noise.

Figure 15 shows the average CMFD performance at image level on 128 test images from FAU and GRIP dataset under various attacks.

1) Plain Copy-Move: We evaluated the proposed approach under ideal conditions. We used the 128 original images and 128 forgery images, and chose per-method optimal thresholds for classifying these 256 images. Similarly to the experiment at pixel level, all regions have been copied and pasted without additional disturbances.

Table 4 CPU-time of each key steps for color images with different size (seconds)

Image sizes	Key steps					Total
	Superpixels segmentation	Adaptive keypoints extraction	Keypoint feature construction	Feature points matching	Postprocessing	
Giraffe (400 × 266)	5.29	6.42	6.95	24.99	82.42	126.07
Tree (512 × 342)	5.81	12.36	6.40	25.66	99.05	149.28
Cattle (640 × 427)	7.60	15.96	7.11	44.63	217.21	292.51
Window (1520 × 1007)	13.06	30.67	13.56	76.84	113.05	267.18
Red tower (1632 × 1224)	13.77	34.85	15.28	82.96	94.83	241.69
Kore (1936 × 1296)	17.55	41.06	18.96	185.51	185.98	449.06
Sails (1944 × 1296)	19.93	41.76	19.51	197.15	90.02	368.37

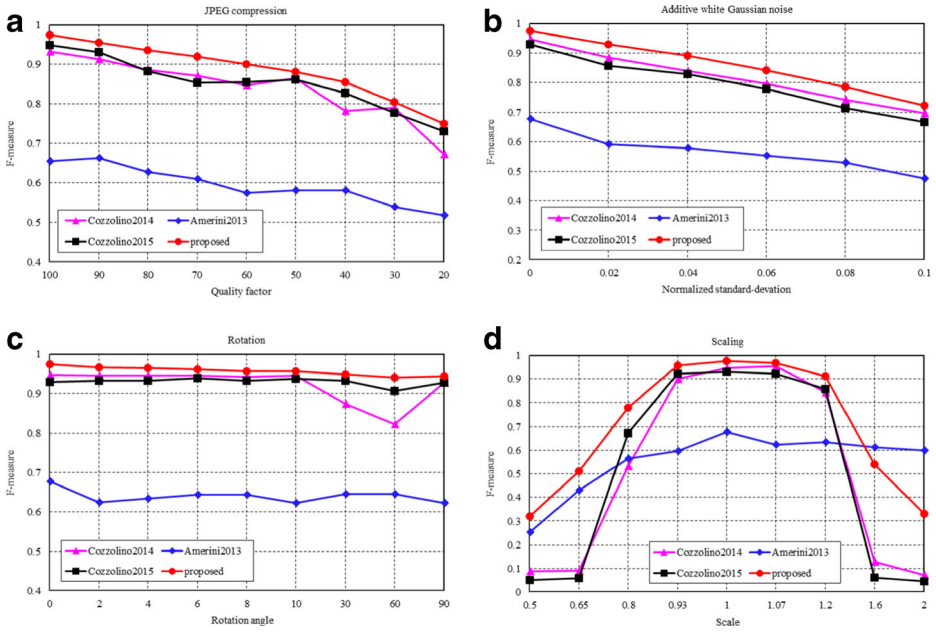

Fig. 15 Average CMFD performance (F-measure) at image level on 80 test images from GRIP dataset: **a** JPEG compression, **b** Additive white Gaussian noise, **c** Rotation, **d** Scaling

2) JPEG Compression: We used the same experimental setup as in the pixel level evaluation, i. e. added JPEG compression between quality levels 30 and 100 in steps of 10.
3) Additive White Gaussian Noise (AWGN): We again used the same experimental setup as in the pixel level evaluation, i. e. zero-mean AWGN with standard deviation 2, 4, 6, and 8, as well as average filters of radius 0.5, 1, 1.5, 2, and 2.5, has been inserted snippets before splicing.
4) Rotation: We evaluated cases where the copied region has been rotated between 2 and 10° in steps of 2°, and also tested three larger rotation angles of 20°, 60° and 180°. We assumed this to be a reasonable range.
5) Scaling: The experimental setup is the same as on the pixel level analysis. The copied region is scaled between 91 and 109 % of its original size in steps of 2 %. Additionally, more extreme scaling parameters were evaluated, including 50, 80, 120 and 200 %.

From the above experimental results, we can see that the newly proposed approach can achieve much better detection results for copy-move forgery images under various challenging conditions, such as geometric transforms, JPEG compression, and additive white Gaussian noise, compared with the existing state-of-the-art copy-move forgery detection schemes. This is because that 1) By utilizing the content based adaptive feature points detector, the stable image keypoints are extracted simultaneously from smooth and texture regions; 2) Robust local visual features, namely exponent moments magnitudes, are constructed for each image keypoint; 3) Best bin first (BBF) and reversed generalized 2 nearest-neighbor (RG2NN) are combined to find rapidly the matching image keypoints; 4) The RANSAC algorithm is utilized to estimate the homographies corresponding to the transformations that the copied regions have undergone. Also, morphological operation is applied to generate the detected forgery regions.

Despite the present advances, there is still much room for improvements. As an example, the proposed copy-move forgery approach is computationally more demanding, and the

average detection time of 128 forgery images from FAU and GRIP database, on a desktop computer with Dual Core 3.4-GHz Pentium CPU and 8 GB RAM, is 221 s, which makes the proposed approach cannot be used effectively in real-time applications.

7 Conclusion

In this paper, we presented a new keypoint-based copy-move forgery detection for small smooth regions. The main novelty of the work consists in introducing the superpixel content based adaptive feature points detector, robust EMs-based keypoint features, and fast Rg2NN based keypoint matching. We demonstrate the effectiveness of the proposed approach with a large number of experiments. Experimental results show that the proposed approach can achieve better detection results for copy-move forgery images if the forged image is rotated, scaled or highly compressed. We compared the robustness of our method with the previously proposed block-based and keypoint-based forgery detection schemes, and we showed that our method is more robust to various types of processing.

The limitation of the proposed copy-move forgery detection is that it has higher computational complexity, which makes the proposed approach cannot be used effectively in real-time applications. In our future work, we will focus on eliminating these drawbacks. Also, we will investigate the use of our approach in detecting regions which have undergone nonaffine transformations.

Acknowledgments This work was supported by the National Natural Science Foundation of China under Grant No. 61472171 & 61272416, the Natural Science Foundation of Liaoning Province of China under Grant No. 201602463, A Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions, and Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology.

References

1. Al-Qershi OM, BE Khoo (2015) Enhanced matching method for copy-move forgery detection by means of Zernike moments. 13th International Workshop on Digital-Forensics and Watermarking (IWDW 2014), LNCS 9023, pp 485–497
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. Amerini I, Ballan L, Caldelli R (2011) A SIFT-based forensic method for copy-move attack detection and transformation recovery. IEEE Trans Inf Forensics Secur 6(3):1099–1110
4. Ardizzone E, Bruno A, Mazzola G (2015) Copy-move forgery detection by matching triangles of keypoints. IEEE Trans Inf Forensics Secur 10(10):2084–2094
5. Bay H, Ess A, Tuytelaars T, Gool LV (2008) Speeded up robust features (SURF). Comput Vis Image Underst 110(3):346–359
6. Bi X, Pun CM, Yuan XC (2016) Multi-level dense descriptor and hierarchical feature matching for copy-move forgery detection. Inf Sci 345:226–242
7. Bravo-Solorio S, Nandi AK (2011) Exposing duplicated regions affected by reflection, rotation and scaling. 2011 I.E. International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, pp 1880–1883
8. Caldelli R, Amerini I, Ballan L (2012) On the effectiveness of local warping against SIFT-based copy-move detection. Proceedings of the 5th International Symposium on Communications, Control and Signal Processing, Rome, Italy, pp 1–5
9. Chambers J, Yan W, Garhwal A (2015) Currency security and forensics: a survey. Multimedia Tools Appl 74(11):4013–4043

10. Chen L, Lu W, Ni J, Sun W (2013) Region duplication detection based on Harris corner points and step sector statistics. J Vis Commun Image Represent 24(3):244–254
11. Chen B, Shu H, Coatrieux G (2015) Color image analysis by quaternion-type moments. J Math Imaging Vision 51(1):124–144
12. Christlein V, Riess C, Jordan J (2012) An evaluation of popular copy-move forgery detection approaches. IEEE Trans Inf Forensics Secur 7(6):1841–1854
13. Costanzo A, Amerini I, Caldelli R (2014) Forensic analysis of SIFT keypoint removal and injection. IEEE Trans Inf Forensics Secur 9(9):1450–1464
14. Cozzolino D, Poggi G, Verdoliva L (2014) Copy-move forgery detection based on patchmatch. 2014 I.E. International Conference on Image Processing (ICIP), Paris, France, pp 5312–5316
15. Cozzolino D, Poggi G, Verdoliva L (2015) Efficient dense-field copy-move forgery detection. IEEE Trans Inf Forensics Secur 10(11):2284–2297
16. Da S (2009) Research on density based image processing algorithms and application. Harbin Institute of Technology, Harbin
17. Davarzani R, Yaghmaie K, Mozaffari S (2013) Copy-move forgery detection using multiresolution local binary patterns. Forensic Sci Int 231(1–3):61–72
18. Fattah SA, Ullah MMI, Ahmed M (2014) A scheme for copy-move forgery detection in digital images based on 2D-DWT. IEEE 57th International Midwest Symposium on Circuits and Systems (MWSCAS), College Station, TX, pp 801–804
19. Imamoglu M, Ulutas G, Ulutas M (2013) Detection of copy-move forgery using Krawtchouk moment. 2013 8th International Conference on Electrical and Electronics Engineering, Bursa, Turkey, pp 311–314
20. Jaberi M, Bebis G, Hussain M (2014) Accurate and robust localization of duplicated region in copy-move image forgery. Mach Vis Appl 25(2):451–475
21. Jie Z, 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
22. Kakar P, Sudha N (2012) Exposing postprocessed copy-paste forgeries through transform-invariant features. IEEE Trans Inf Forensics Secur 7(3):1018–1028
23. Ketenci S, Ulutas G, Ulutas M (2014) Detection of duplicated regions in images using 1D-Fourier transform. International Conference on Systems, Signals and Image Processing, Dubrovnik, Croatia, pp 171–174
24. Lee JC (2015) Copy-move image forgery detection based on Gabor magnitude. J Vis Commun Image Represent 31:320–334
25. Lee J, Chang C, Chen W (2015) Detection of copy–move image forgery using histogram of orientated gradients. Inf Sci 321:250–262
26. Li J, Li X, Yang B (2015) Segmentation-based image copy-move forgery detection scheme. IEEE Trans Inf Forensics Secur 10(3):507–518
27. Li Y, Wang S, Tian Q (2015) A survey of recent advances in visual feature detection. Neurocomputing 149: 736–751
28. Liu MY, Tuzel O, Ramalingam S, Chellappa R (2011) Entropy rate superpixel segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Colorado Springs, Colorado, USA, pp 2097–2104
29. Meng M, Ping ZL (2011) Decompose and reconstruct images based on exponential Fourier moments. J Inner Mongolia Norm Univ (Nat Sci Ed) 40(3):258–260
30. Muhammad G, Hussain M, Bebis G (2012) Passive copy move image forgery detection using undecimated dyadic wavelet transform. Digit Investig 9(1):49–57
31. Pun CM, Yuan XC, Bi XL (2015) Image forgery detection using adaptive oversegmentation and feature point matching. IEEE Trans Inf Forensics Secur 10(8):1705–1716
32. Qureshi MA, Deriche M (2015) A bibliography of pixel-based blind image forgery detection techniques. Signal Process Image Commun 39(Part A):46–74
33. Ryu SJ, Kirchner M, Lee MJ (2013) Rotation invariant localization of duplicated image regions based on Zernike moments. IEEE Trans Inf Forensics Secur 8(8):1355–1370
34. Silva E, Carvalho T, Ferreira 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. Sitara K, Mehtre BM (2016) Digital video tampering detection: an overview of passive techniques. Digit Investig 18:8–22
36. Ustubioglu B, Ayas S, Doganl H (2015) Image forgery detection based on color SIFT. The IEEE Signal Processing and Communications Applications Conference, Malatya, Turkey, pp 1741–1744
37. Wang X, Liu Y, Li S, Yang H, Niu P, Zhang Y (2015) A new robust digital image watermarking using local polar harmonic transform. Comput Electr Eng 46:403–418
38. Wu YJ, Yu D, Duan HB (2014) Dual tree complex wavelet transform approach to copy-rotate-move forgery detection. Sci China Inf Sci 57(1):1–12

39. Xia Z, Wang X, Sun X (2016) Steganalysis of LSB matching using differences between nonadjacent pixels. Multimedia Tools and Applications 75(4):1947–1962
40. Yan L, Nian L, Bin Z, Kai-guo Y, Yang Y (2015) Image multiple copy-move forgery detection algorithm based on reversed-generalized 2 nearest-neighbor. J Electron Inf Technol 7:1767–1773
41. Yu L, Han Q, Niu X (2016) Feature point-based copy-move forgery detection: covering the non-textured areas. Multimedia Tools Appl 75(2):1159–1176
42. Zhou Z, Wang Y, Jonathan Wu QM, Yang C-N, Sun X (2006) Effective and efficient global context verification for image copy detection. IEEE Trans Inf Forensics Secur. doi:10.1109/TIFS.2016.2601065

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.

Shuo Li received the B. E. degree from the School of Computer and Information Technology, Liaoning Normal University, China, in 2014, where she is currently pursuing the M. S. E. degree. Her research interests include digital watermarking and forgery detection.

Yu-Nan Liu received the B. E. degree from the School of Computer and Information Technology, Liaoning Normal University, China, in 2014, where he is currently pursuing the M. S. E. degree. His research interests include digital watermarking 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.

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.

Zhi-li Zhou received his BS degree at Communication Engineering from Hubei University in 2007, and his MS and PhD degrees in Computer Application at the School of Information Science and Engineering from Hunan University, in 2010 and 2014, respectively. He is an assistant professor at Nanjing University of Information Science and Technology. His current research interests include near-duplicate image/video detection, image/video copy detection, coverless information hiding, digital forensics, and image processing.