symmetry

MDPI

Article

An Image Copy-Move Forgery Detection Scheme Based on A-KAZE and SURF Features

Chengyou Wang, Zhi Zhang and Xiao Zhou *

School of Mechanical, Electrical and Information Engineering, Shandong University, Weihai 264209, China; wangchengyou@sdu.edu.cn (C.W.); zhi@mail.sdu.edu.cn (Z.Z.)

* Correspondence: zhouxiao@sdu.edu.cn; Tel.: +86-631-568-8338

Received: 3 October 2018; Accepted: 28 November 2018; Published: 3 December 2018

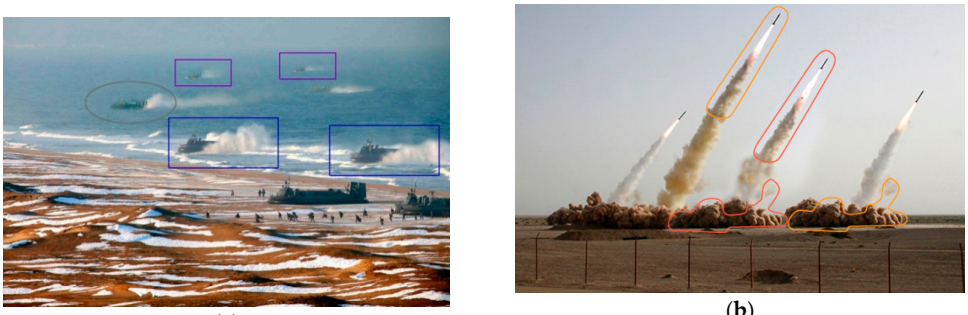
Abstract: The popularity of image editing software has made it increasingly easy to alter the content of images. These alterations threaten the authenticity and integrity of images, causing misjudgments and possibly even affecting social stability. The copy-move technique is one of the most commonly used approaches for manipulating images. As a defense, the image forensics technique has become popular for judging whether a picture has been tampered with via copy-move, splicing, or other forgery techniques. In this paper, a scheme based on accelerated-KAZE (A-KAZE) and speeded-up robust features (SURF) is proposed for image copy-move forgery detection (CMFD). It is difficult for most keypoint-based CMFD methods to obtain sufficient points in smooth regions. To remedy this defect, the response thresholds for the A-KAZE and SURF feature detection stages are set to small values in the proposed method. In addition, a new correlation coefficient map is presented, in which the duplicated regions are demarcated, combining filtering and mathematical morphology operations. Numerous experiments are conducted to demonstrate the effectiveness of the proposed method in searching for duplicated regions and its robustness against distortions and post-processing techniques, such as noise addition, rotation, scaling, image blurring, joint photographic expert group (JPEG) compression, and hybrid image manipulation. The experimental results demonstrate that the performance of the proposed scheme is superior to that of other tested CMFD methods.

Keywords: image forensics; copy-move forgery detection (CMFD); accelerated-KAZE (A-KAZE) feature; speeded-up robust features (SURF)

1. Introduction

Currently, even with the rapid development of technology, images and videos continue to be primary sources of information and have become important information carriers in research fields such as hyper-spectral image analysis [1], image registration [2], object tracking [3], and remote sensing and photogrammetry [4]. However, the increasing popularity of multimedia editing software such as GNU image manipulation program (GIMP) [5] and Adobe Photoshop [6] has made editing image content increasingly convenient. While multimedia that has been tampered with can make people's lives more interesting, such tampering also poses a threat in many fields [7], particularly those that involve legal and safety aspects such as insurance claims, court sentences, patent infringement, medical diagnoses, and so on. One recent event related to forged images involved the North Korea hovercraft landing photo shown in Figure 1a. There is speculation that some of the hovercrafts on the sea may have been copied and pasted onto other regions in the image; the similar regions are indicated by the purple and blue rectangles in Figure 1a. Another event involved the Iranian missile photo shown in Figure 1b, in which the third missile from the left was digitally appended to the original photo to obscure the fact that it did not fire. Forged images such as those shown in Figure 1 cause a great deal of commotion in

the world, demonstrating the urgency of developing image and video forensics techniques to avoid unnecessary problems.



(**a**) (**b**)

Figure 1. Actual events involving forged images: (**a**) North Korea hovercraft landing photo; (**b**) Iranian missile photo.

Image forensics techniques have been studied for many decades and can be classified into two classes: active and passive forensics techniques. Active forensics techniques verify the integrity of auxiliary information such as digital signature [8] and digital watermark [9] to help determine the authenticity of an image. However, this type of technology requires special software and hardware to insert the authentication information into the images before distribution and to extract authentication information from the images in the authentication phase. In contrast, passive forensics techniques verify the authenticity of an image by analyzing its content and structure, an approach that overcomes the disadvantages of active forensics techniques.

There are many ways to manipulate images, for example, resampling, retouching, splicing, copy-move, double joint photographic expert group (JPEG) compression [10], and filtering [11]. In Figure 1, the copy-move method was used to forge the image. Copy-move operations are usually performed with image editing software to obscure objects in smooth regions and to add new objects within an image. Copy-move is one of the most commonly used approaches for manipulating images and has already received considerable attention from researchers in various fields. In recent decades, many image forensics schemes for copy-move forgery detection (CMFD) have been proposed to judge whether an image has been tampered with via copy-move manipulation. These methods are mainly divided into two classes: block-based CMFD schemes and keypoint-based CMFD schemes.

Fridrich et al. [12] presented a CMFD method using discrete cosine transform (DCT), which is a landmark work in the field of block-based CMFD methods. They separated the image into overlapping image patches with a fixed size via a raster scan and then applied the DCT to each image block. A quantization feature vector was obtained by performing zigzag scanning on the quantized DCT coefficient matrix. The matrix, which consisted of feature vectors, was lexicographically ordered, and Euclidean distance was used to search for similar features. However, this method carries a high computational cost. To reduce the computational complexity of Fridrich et al.'s method [12], Huang et al. [13] truncated the feature vector by using a constant to reduce the feature dimensionality and presented a strategy for judging the similarity among adjacent feature vectors. To prevent distortion (e.g., rotation), Bi et al. [14] extracted the invariant moment descriptor and color texture descriptor calculated from the polar complex exponential transform (PCET) moments to solve the problem of finding duplicated regions. Zhong et al. [15] extracted discrete radial harmonic Fourier moments (DRHFMs) from each overlapping circular block to locate duplicated regions. Zhong and Gan [16] proposed a discrete analytical Fourier-Mellin transform (DAFMT)-based scheme to find duplicated regions. However, the DAFMT-based method is too complicated, particularly the process of invariant moment construction. To reduce the time required to perform the feature matching process, Cozzolino et al. [17] proposed a CMFD scheme that utilized invariant features and a modified PatchMatch algorithm. The block-based CMFD methods achieve good accuracy in detecting duplicated

regions without distortions. Deng et al. [18] combined DAFMT and locality sensitive hashing (LSH) to solve the problem of CMFD. Mahmood et al. [19] divided the image into overlapping circle blocks to extract local binary pattern variance (LBPV) to locate the duplicated region. Fadl and Semary [20] converted the image block into polar systems, and fast Fourier transform was used as the descriptor to find similar regions. However, most of the block-based schemes for searching for duplicated regions have high computational complexity.

Amerini et al. [21] proposed a CMFD scheme that used the scale-invariant features transform (SIFT) feature, which is insensitive to geometric transformation and illumination distortion. Their scheme extracted SIFT feature descriptors and then selected similar SIFT feature descriptors using a generalized 2 nearest-neighbor (g2NN) test. Agglomerative hierarchical clustering was used to find the regions with dense points. Next, the affine transformation matrix between the putative matched SIFT keypoints was estimated via the random sample consensus (RANSAC) algorithm. Amerini et al. [22] added the J-linkage algorithm, which implements a robust cluster in the space of a geometry transformation, to improve the performance of the scheme in Reference [21]. Jin and Wan [23] first utilized a non-maximum value suppression algorithm to choose keypoints, and color information was added to describe the feature descriptors. Clustering algorithms and superpixels were combined to locate the duplicated regions. In [24], Shivakumar and Santhosh Baboo proposed a CMFD scheme based on speeded-up robust features (SURF) [25] and used a k-dimensional tree (k-d tree) to search for similar SURF descriptors. However, the duplicated regions are indicated by lines, and the boundaries of the duplicated region are not explicit. Jaberi et al. [26] presented the mirror reflection invariant feature transform (MIFT), which is invariant to mirror reflection transformations. They presented a MIFT-based CMFD scheme that could detect reflection distortions. Yu et al. [27] supplemented redundant feature points and feature fusion to solve the problem that the keypoint-based CMFD methods failed to detect small and smooth tampered areas. In Reference [27], they presented a two-stage feature detection method to guarantee that enough points existed to cover an entire image. They proposed a fused feature obtained by using a multisupport region order-based gradient histogram and a hue histogram to improve feature matching. Uliyan et al. [28] proposed a Harris-based CMFD scheme combined with angular radial partitioning to find duplicated regions within forged images tampered by copy-move. Ulutas and Muzaffer [29] presented a CMFD scheme based on accelerated-KAZE (A-KAZE) [30], which can detect only the singly tampered region. In Reference [31], the authors combined the KAZE [32] and SIFT features to extract sufficient points and proposed a CMFD method that could detect multiple duplicated regions. After detecting interest points, Zandi et al. [33] used the polar cosine transform as feature descriptors and then used an iterative detection process for duplicated regions with regard to a priori information to enhance the output result. Yang et al. [34] utilized an adaptive SIFT detector to achieve CMFD; however, its detection results are marked with lines that cannot clearly mark duplicated regions. A maximally stable color region (MSCR) detector in Refence [35] was used to detect points, and the Zernike moment was used to describe the feature. An improved n best matching strategy was adopted to find multiple duplicated regions.

Deep learning techniques have been extensively adopted due to the advance of modern technology and production techniques, and have been applied in many fields, including machine health monitoring [36], medical diagnosis [37], and target detection [38,39]. Rao and Ni [40] utilized a convolutional neural network (CNN) to learn hierarchical representations from RGB images. After obtaining the discriminative features, they applied a support vector machine (SVM) to differentiate authentic and tampered images. However, the CNN-based method requires very large amounts of training data samples, and emphasizes classification accuracy. Deep learning techniques will require further exploration in the field of image forensics in the future.

A CMFD method based on A-KAZE and SURF is proposed in this paper. A-KAZE [30] is an accelerated version of KAZE [32] that reduces the time KAZE requires for feature extraction and description. Similar to SIFT extraction, KAZE builds a nonlinear scale space instead of performing

Gaussian blurring. This approach makes it possible to avoid the issue of Gaussian blurring, which ignores natural boundaries of objects in images. A-KAZE also improves localization accuracy and uniqueness when smoothed to the same level of detail and noise. To obtain sufficient points, point detection response thresholds are set to small values, which ensures that sufficient points remain after SURF and A-KAZE feature extraction. In the feature matching stage, the g2NN test [21] is adopted to improve the feature matching precision. To enhance the robustness of the method presented, the affine transformation matrix is estimated via RANSAC. To demarcate the duplicated regions using closed regions rather than lines and points, the estimated affine transformation matrix is used to locate the forged regions. A new correlation coefficient map is calculated, and filtering and mathematical morphology operations are combined to refine the detected duplicated regions and eliminate isolated points. The experimental results demonstrate the effectiveness of the proposed scheme for detecting single and multiple tampered regions and show its robustness against scaling, rotation, image blurring, noise addition, JPEG compression, and hybrid image manipulation.

The rest of this paper is organized as follows. The A-KAZE and SURF feature detection and description are fully described in Section 2. The proposed CMFD scheme based on these two features is described in Section 3. Section 4 presents the performance indexes of various CMFD schemes and experimental results of image CMFD and robustness tests. Finally, conclusions and suggestions for future work are provided in Section 5.

2. Review of A-KAZE and SURF

In this section, the processes of A-KAZE and SURF detection and description are described in Sections 2.1 and 2.2, respectively. The proposed CMFD scheme using hybrid features is based on these two features; it uses their invariance to geometric transformation to ensure that the proposed scheme can detect distortions such as scaling and rotation.

2.1. A-KAZE

2.1.1. Nonlinear Scale Space Construction

The divergence of a certain flow function can express the nonlinear diffusion filter to represent the luminance change of images at different scales. The classic nonlinear diffusion equation is shown in Equation (1):

$$\frac{\partial L}{\partial t} = \mathrm{div}(c(x,y,t) \cdot \nabla L), \tag{1}$$

where div is the divergence operator, ∇ is the gradient operator, L is the brightness of the image, and $c(x,y,t)$ is a conductivity function defined in Equation (2):

$$c(x,y,t) = g|\nabla L_\sigma(x,y,t)|, \tag{2}$$

where ∇L_σ is the gradient of a Gauss smooth version of L, and time t is a scale parameter. The smaller the value of t is, the more complex the obtained image representations are. Alcantarilla et al. [30] offered four kinds of conductivity functions in their project. This study adopted the g_2 function, which is defined in Equation (3):

$$g_2 = \frac{1}{1 + \frac{|\nabla L_\sigma|^2}{\lambda^2}}, \tag{3}$$

where λ is a contrast factor that regulates the diffusion level and is relevant to the marginal information. The smaller the value of λ, the larger the amount of retained edge information.

The nonlinear scale space construction resembles that of SIFT. An image pyramid is constructed using fast explicit diffusion (FED) [41], which is discretized in a battery of S octaves and O sublevels. The octaves and sublevels are mapped into the corresponding scale σ using Equation (4):

$$\sigma_i(o,s) = 2^{o+\frac{s}{S}},\ o \in [0,1,\cdots,O-1],\ s \in [0,1,\cdots,S-1],\ i \in [0,1,\cdots,M], \tag{4}$$

where M is the total number of filtered images. The transformation between the scale parameter (in pixels) and the nonlinear scale space (in time units) is complete when $\sigma_i \to t_i$:

$$t_i = \frac{1}{2}\sigma_i^2,\ i = \{0,1,\cdots,M\}, \tag{5}$$

where t_i is the evolution time. All images in the nonlinear space can be obtained using different evolution times.

Equation (1) is discretized into Equation (6) through an explicit scheme:

$$\frac{L^{i+1} - L^i}{\tau} = \boldsymbol{A}(L^i)L^i, \tag{6}$$

where $\boldsymbol{A}(L^i)$ is a matrix encoding the conductivity of a picture, and τ is a time constant. The solution L^{i+1} is obtained as follows:

$$L^{i+1} = [\boldsymbol{I}_{\text{identity}} + \tau\boldsymbol{A}(L^i)]L^i, \tag{7}$$

where $\boldsymbol{I}_{\text{identity}}$ is the identity matrix. In consideration of the prior estimation $L^{i+1,0} = L^i$, an FED cycle with n alterable step sizes τ_j can be acquired, as defined in Equation (8). The value of τ_j is calculated as shown in Equation (9):

$$L^{i+1,j+1} = [\boldsymbol{I} + \tau_j\boldsymbol{A}(L^i)]L^{i+1,j}, j = 0,1,\cdots,n-1, \tag{8}$$

$$\tau_j = \frac{\tau_{\max}}{2\cos\left(\pi\frac{2j+1}{4n+2}\right)}. \tag{9}$$

Figure 2 shows scale space images in both Gaussian scale space and nonlinear scale space. An experiment is conducted as follows. By adjusting the parameters of the octaves and levels in the SIFT and A-KAZE algorithms in Microsoft Visual Studio 2012 using the OpenCV library, we executed SIFT and A-KAZE on Lena with a size of 512 × 512 to generate 6 individual Lena images at a resolution of 256 × 256. The first, third, fifth, and sixth space images are depicted in Figure 2, where the first and second rows show the 256 × 256 Lena images in Gaussian scale space and nonlinear scale space, respectively. These images demonstrate that nonlinear filtering retains more details than linear filtering. A similar conclusion was obtained in Reference [42].

(**a**) (**b**) (**c**) (**d**)

Figure 2. The 256 × 256 Lena images in Gaussian scale space (top) and nonlinear scale space (bottom): (**a**) the first space; (**b**) the third space; (**c**) the fifth space; (**d**) the sixth space.

2.1.2. A-KAZE Feature Detection

Similar to SIFT extraction, A-KAZE features were extracted using a Hessian matrix, which was calculated using Equation (10):

$$L^i_{\text{Hessian}} = \sigma^2_{i,\text{norm}}(L^i_{xx}L^i_{yy} - L^i_{xy}L^i_{xy}), \tag{10}$$

where $\sigma_{i,\text{norm}}$ is a normalized scale factor, defined as $\sigma_{i,\text{norm}} = \sigma_i/2^{o^i}$. Here, L^i_{xy}, L^i_{yy}, and L^i_{xx} are the second-order cross, vertical, and horizontal derivatives, respectively. Next, Hessian extreme points are detected among the 26 points around the detected point in a $3 \times 3 \times 3$ neighborhood between the 3×3 rectangle windows of the neighbor scales and the current scale. A detected point is considered to be a keypoint if it is the extreme value and its Hessian extreme value is higher than the pre-threshold $T_{\text{A}-\text{KAZE}}$.

Taking the extreme point as the centrality of the neighborhood, the principal direction can be found by searching over a radius of $6\sigma_i$ with a sampling step of σ_i to guarantee that A-KAZE features are rotation invariant. First-order differential values of all the adjacent points in a circular region centered at the interest point are weighted with a Gaussian weighting. These weighted values are regarded as the response values of pixels of the image. In the sliding window with a sector region of $\pi/3$, all the internal response values are summed. After traversing the entire circle, the direction of the sector region with the highest value provides the main orientation of the feature point.

2.1.3. A-KAZE Feature Description

In the feature description phase, feature descriptors are described using a modified-local difference binary (M-LDB) descriptor. The M-LDB descriptor is obtained by modifying the local difference binary (LDB) descriptor [43]. To ensure that the descriptor is rotation invariant, Alcantarilla et al. [30] subsampled the grids in the steps as a function of the σ of the feature rather than the mean value of all the pixels in each sub-division of the grid. An image patch centered at the feature point is selected. Then, each image patch is divided into $q \times q$ grids with a fixed size, and representative information is extracted from each grid unit. A binary test operation is performed on a pair of grid units, as indicated in Equation (11):

$$\overline{\omega}(F_{\text{function}}(i), F_{\text{function}}(j)) = \begin{cases} 1, & \text{if}(F_{\text{function}}(i) - F_{\text{function}}(j)) > 0,\ i \neq j, \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

where $F_{\text{function}}(\cdot)$ denotes the function used to extract information from the grid unit. The function is defined as follows:

$$F_{\text{function}}(\cdot) = \left\{F_{\text{intensity}}(\cdot), F_{\text{d}x}(\cdot), F_{\text{d}y}(\cdot)\right\}, \tag{12}$$

$$F_{\text{intensity}}(i) = \frac{1}{m}\sum_{k=1}^{m} Intensity(k),\ F_{\text{d}x}(i) = \text{Gradient}_x(i),\ F_{\text{d}y}(i) = \text{Gradient}_y(i), \tag{13}$$

where m is the total number of pixels in grid unit i, $\text{Intensity}(k)$ is pixel value, and $\text{Gradient}_x(i)$ and $\text{Gradient}_y(i)$ are the regional gradients of the grid units in x and y, respectively.

Upon the completion of the A-KAZE feature description phase, 61 dimensional descriptors are obtained.

2.2. SURF

SURF [25] is not only invariant to scaling and rotation but also robust to illumination variation and affine transformation. SURF uses the determinant of the Hessian matrix to select the scale and location. The Hessian matrix $\boldsymbol{H}(\boldsymbol{x};\sigma)$ at scale σ is shown in Equation (14):

$$\boldsymbol{H}(\boldsymbol{x};\sigma) = \begin{bmatrix} C_{xx}(\boldsymbol{x},\sigma) & C_{xy}(\boldsymbol{x},\sigma) \\ C_{xy}(\boldsymbol{x},\sigma) & C_{yy}(\boldsymbol{x},\sigma) \end{bmatrix}, \tag{14}$$

where $C_{xx}(\boldsymbol{x},\sigma)$ denotes the convolution of the Gaussian second-order derivative $\partial^2 g(\sigma)/\partial x^2$ of the image in $\boldsymbol{x}$, and $C_{xy}(\boldsymbol{x},\sigma)$ and $C_{yy}(\boldsymbol{x},\sigma)$ are treated similarly.

To save computation time, the box filter is used to simulate the Gaussian second-order derivative, $D_{xx}(\boldsymbol{x},\sigma)$, $D_{xy}(\boldsymbol{x},\sigma)$, and $D_{yy}(\boldsymbol{x},\sigma)$, which can improve the calculation speed of convolution and reduce the time complexity of the entire process. The approximation of Hessian's determinant is computed as shown in Equation (15):

$$\det(\boldsymbol{H}_{\text{approx}}) = D_{xx}D_{yy} - (0.9D_{xy})^2. \tag{15}$$

The box filters and integral images are used to construct the image pyramid scale space. Interest points are determined using non-maximum suppression with a $3 \times 3 \times 3$ neighborhood. Finally, the image points are determined by removing the candidate points whose approximate determinate of the Hessian matrix is smaller than the Hessian response threshold T_{SURF}.

To obtain the rotation invariant descriptors, the main direction needs to be determined. The direction of the point depends on the circular region response centered at the point with a radius of 6σ. A Haar wavelet filter is used to process the circular region to obtain the Haar wavelet response. The entire circular region of Haar wavelet response is scanned by a fan-shaped area centered at the point with an angle of $\pi/3$. The vector sum of the Haar wavelet response in each fan-shaped region is calculated. The longest vector among all the vectors is chosen as the main direction. Finally, after determining the direction of the point, the feature descriptor at the point is generated, as described in the following steps:

Step 1. Build a square area centered at the point. The length of the square region is 20σ. Rotate the square region to the main direction of the point;

Step 2. Divide the square region into a 4×4 sub-region, and calculate a 4-dimensional vector from the 5×5 regular grid space in each sub-region that includes the sum of the Haar wavelet responses to the horizontal and vertical direction, and the absolute value of the sum of the Haar wavelet responses to the horizontal and vertical direction;

Step 3. Calculate the 4-dimensional feature vector of each sub-region; then, a 64-dimensional feature descriptor is obtained by combining the feature vectors calculated from the 16 sub-regions.

3. The Proposed Image Copy-Move Forgery Detection Scheme

The entire process of the proposed CMFD method can be summarized in the following stages: A-KAZE and SURF feature extraction, g2NN feature matching, eliminating false matching points with RANSAC, calculating the correlation coefficient map, and filtering and mathematical morphology operations, as shown in Figure 3.

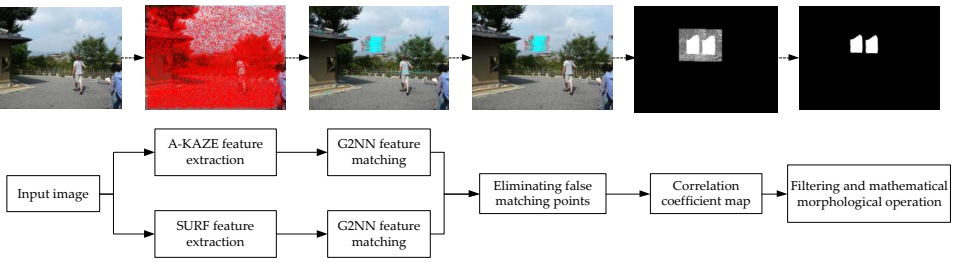


Figure 3. Flow diagram of the proposed image copy-move forgery detection method.

3.1. Feature Extraction

A-KAZE and SURF, as described in Section 2, are performed on the input image. Next, feature descriptors are extracted. It should be noted that most keypoint-based CMFD methods fail to detect duplicated regions. In addition, most keypoint extraction methods executed with the default parameters cannot obtain sufficient keypoints in smooth tampered regions. Inspired by Reference [44], in this paper, the detector response thresholds, $T_{\mathrm{A-KAZE}}$ and T_{SURF}, are set to small values to obtain sufficient interest points. As is depicted in Figure 4b, the image of the Japan tower is hidden by the sky region within the tampered same image. Using their default parameters, the SIFT, SURF, KAZE, and A-KAZE feature extraction algorithms are unable to extract points in the tampered region. However, those points can be obtained with the hybrid features (SURF and A-KAZE with small response thresholds). Obtaining a sufficient number of points is the basis of keypoint-based CMFD methods.

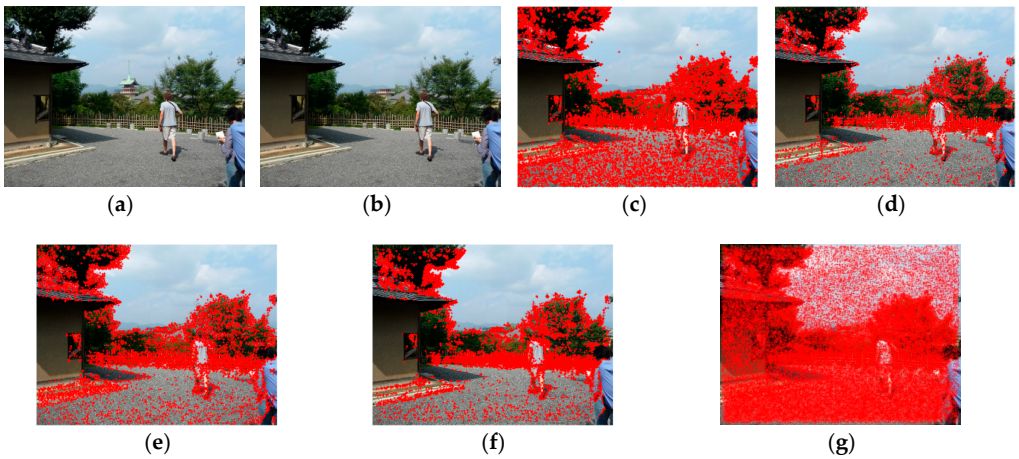


Figure 4. Feature extraction with different keypoint detectors: (**a**) original Japan tower; (**b**) tampered Japan tower; (**c**) SIFT detector; (**d**) SURF detector; (**e**) KAZE detector; (**f**) A-KAZE detector; (**g**) A-KAZE and SURF detectors with small response thresholds.

3.2. Feature Matching

After the A-KAZE and SURF feature extractions, the g2NN test [21] is performed on the feature descriptors obtained to identify similar descriptors in an image. A group of feature keypoints $\{x_1, x_2, \cdots, x_n\}$ and their corresponding feature descriptors $\{f_1, f_2, \cdots, f_n\}$ are generated. To select similar feature descriptors, Euclidean distance is used to evaluate the similarity between two keypoint

descriptors. The similarity vector $\boldsymbol{D} = \{d_1, d_2, \cdots, d_{n-1}\}$ represents the sorted Euclidean distances regarding other feature descriptors. In the 2NN test, the ratio is calculated by dividing the closest distance by the second-closest distance; then, that value is compared with a threshold value V_t (set to 0.6 in this paper, an example is shown in Table 1). Based on this idea, the feature keypoints are considered to match it if the constraint in Equation (16) is satisfied. The g2NN test consists of iterating the 2NN test between d_i and d_{i+1} until the ratio is larger than V_t. If the procedure stops at iteration k, those k points can be viewed as a match of the inspected point.

$$d_1/d_2 < V_t,\ V_t \in (0,1). \tag{16}$$

Table 1. Number of pairs of matched points with different V_t.

V_t	Non-Distortion			Blurring		
	N_{dup}	$N_{non-dup}$	N_{sum}	N_{dup}	$N_{non-dup}$	N_{sum}
0.1	172	0	172	125	0	125
0.2	186	0	186	134	0	134
0.3	196	0	196	139	1	140
0.4	202	1	203	145	3	148
0.5	208	6	214	153	4	157
0.6	214	17	231	157	23	180
0.7	224	87	311	161	97	258
0.8	237	630	867	175	546	721
0.9	367	5944	6311	256	4920	5176

The full set of matched keypoints is obtained by combining the results of A-KAZE and SURF feature matching.

Table 1 lists the data obtained with different V_t after the g2NN test, in which the Japan tower is taken as an example to indicate the influence of threshold value V_t. The data in Table 1 are the number of pairs of matched points N_{sum}, the number of pairs of matched points in duplicated region N_{dup}, and the number of pairs of matched points in non-duplicated region $N_{non-dup}$. From Table 1, V_t is suitable for being set to 0.6 for feature matching considering the distorted region with rotation and scaling, and blurring is added to Table 1 to demonstrate that 0.6 is reasonable. The false matching points can be eliminated using RANSAC, as mentioned below. In this paper, the k-d tree is used to implement the g2NN test to search for similar features.

3.3. Affine Transformation Estimation

The pasted region is often processed by distortions such as rotation and scaling before being moved to another region within the same image. The distortion is modeled as an affine transformation of image coordinates. Two coordinates from the copied region and pasted region are $\boldsymbol{x} = (x, y)^{\mathrm{T}}$ and $\widetilde{\boldsymbol{x}} = (\widetilde{x}, \widetilde{y})^{\mathrm{T}}$, respectively. Their relation is shown in Equation (17):

$$\begin{pmatrix} \widetilde{x} \\ \widetilde{y} \end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \rightarrow \begin{pmatrix} \widetilde{x} \\ \widetilde{y} \\ 1 \end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} & x_0 \\ t_{21} & t_{22} & y_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \rightarrow \widetilde{\boldsymbol{X}} = \boldsymbol{T}\boldsymbol{X}, \tag{17}$$

where (x_0, y_0) is a shift vector, (x, y) is the coordinate of the copied region, $(\widetilde{x}, \widetilde{y})$ is the coordinate of the pasted region, and t_{11}, t_{12}, t_{21}, and t_{22} are the affine transformation parameters.

To obtain $\boldsymbol{T}$, at least three pairs of corresponding non-collinear coordinates are needed. RANSAC is widely used to estimate the affine transformation matrix $\boldsymbol{T}$, which it can achieve with a high degree of accuracy, even though many mismatched pairs are included in the input data. Using the matched features obtained in Section 3.2, the following steps are executed N times:

(1) Randomly choose three non-collinear matched points (as described above). Based on the selected point pairs, estimate $\boldsymbol{T}$ by minimizing the objective function in Equation (18):

$$L(\boldsymbol{T}) = \sum_{i=1}^{N} \left\| \widetilde{\boldsymbol{X}}_i - \boldsymbol{T}\boldsymbol{X}_i \right\|_2^2. \tag{18}$$

(2) Divide all the pairs of matched points obtained above into outliers and inliers by using the estimated matrix $\boldsymbol{T}$. A pair of matched points $\{\boldsymbol{x}, \widetilde{\boldsymbol{x}}\}$ belongs to the inlier group if $\|\widetilde{\boldsymbol{X}} - \boldsymbol{T}\boldsymbol{X}\|_2 \leq \varepsilon$; otherwise, it belongs to the outlier group.

In this scheme, the maximum number of iterations is set to $N = 1000$, and the evaluation error is set to $\varepsilon = 10^{-6}$.

3.4. Filtering and Post-Processing

Most keypoint-based CMFD methods terminate at the keypoint feature matching stage using the RANSAC algorithm. When regions marked with dense points exist in the detection result, the images can be considered as tampered images. In this paper, we adopt a strategy that can demarcate the duplicated regions with closed regions. In the scope of the entire image, all the points are forward transformed using the estimated matrix $\boldsymbol{T}$, $\widetilde{\boldsymbol{X}}_{\mathrm{f}} = \boldsymbol{T}\boldsymbol{X}$. The similarity between the region centered at the original point and that centered at the estimated point is evaluated using correlation coefficients, as shown in Equation (19). These pixel values at locations l and $\widetilde{l}$ are $I_{\mathrm{o}}(l)$ and $I_{\mathrm{t}}(\widetilde{l})$, respectively. The correlation coefficient is calculated as follows:

$$c(l) = \frac{\sum_{\mu \in \Omega(l), \widetilde{\mu} \in \Omega(\widetilde{l})} [I_{\mathrm{o}}(\mu) - \overline{I}_{\mathrm{o}}][I_{\mathrm{t}}(\widetilde{\mu}) - \overline{I}_{\mathrm{t}}]}{\sum_{\mu \in \Omega(l), \widetilde{\mu} \in \Omega(\widetilde{l})} \sqrt{[I_{\mathrm{o}}(\mu) - \overline{I}_{\mathrm{o}}]^2 [I_{\mathrm{t}}(\widetilde{\mu}) - \overline{I}_{\mathrm{t}}]^2}}, \tag{19}$$

where $\Omega(l)$ and $\Omega(\widetilde{l})$ are 5×5 pixel neighbor regions centered at l and $\widetilde{l}$, respectively, $\overline{I}_{\mathrm{o}}$ and $\overline{I}_{\mathrm{t}}$ are the mean values of $\Omega(l)$ and $\Omega(\widetilde{l})$, respectively, and $c(l)$ is the correlation coefficient, which ranges from 0 to 1. A smaller value of $c(l)$ indicates less similarity between the original and the transformed regions. The other correlation coefficient map is obtained using a similar approach, i.e., $\widetilde{\boldsymbol{X}}_{\mathrm{b}} = \boldsymbol{T}^{-1}\boldsymbol{X}$. In this paper, the correlation coefficient map is only calculated in the square region determined by $x_{\min} - N_{\mathrm{pixel}}$, $x_{\max} + N_{\mathrm{pixel}}$, $y_{\min} - N_{\mathrm{pixel}}$, and $y_{\max} + N_{\mathrm{pixel}}$ instead of the whole image scope, where $x_{\min}$ and $y_{\min}$ are the smallest coordinates of the region of dense points, $x_{\max}$ and $y_{\max}$ are the biggest coordinates of the region of dense points, and N_{pixel} is the number of expanded pixels in the x and y direction. Combining these two maps, the entire correlation map is obtained. An example is shown in the fifth image of the first row of Figure 3.

After obtaining the correlation map, a binary image is obtained by transforming each correlation map with the threshold T_{bin}. Then, a filtering scheme is adopted to roughly locate the tampered regions. A 5×5 square area centered at a point in the binary image is obtained. If the sum of this area is larger than 80% of this area, the point belongs to a duplicated region. Finally, isolated points are removed, and mathematical morphology operations are used to fill in holes in the binary image.

To make the proposed method more understandable, a pseudocode for the entire scheme process is presented in Algorithm 1. Several of the symbols in Algorithm 1 need to be defined: $\boldsymbol{P}_{\mathrm{A}_1}$ and $\boldsymbol{P}_{\mathrm{A}_2}$ are the positions of the matched points from the g2NN test of the A-KAZE descriptors. $\boldsymbol{P}_{\mathrm{S}_1}$ and $\boldsymbol{P}_{\mathrm{S}_2}$ are the positions of the matched points from the g2NN test of the SURF descriptors. $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ are the positions of the matched points after eliminating the points that are too close between $[\boldsymbol{P}_{\mathrm{A}_1}, \boldsymbol{P}_{\mathrm{A}_2}]$ and $[\boldsymbol{P}_{\mathrm{S}_1}, \boldsymbol{P}_{\mathrm{S}_2}]$.

Algorithm 1. Proposed CMFD Scheme.

Variable Declaration:

$\boldsymbol{I}$: test image
$\boldsymbol{D}_i$: extracted descriptors, $i = \mathrm{A-KAZE, SURF}$
$\boldsymbol{P}_j$: positions of detected points, $j = \mathrm{A-KAZE, SURF}$
$\boldsymbol{P}_k$: positions of matched points with g2NN test, $k = \mathrm{A_1, A_2, S_1, S_2}, 1, 2$
$\boldsymbol{T}$: estimated affine transformation matrix
$\boldsymbol{P}_{\mathrm{inliners}}$: matched points with RANSAC
$\boldsymbol{M}_{\mathrm{map}}$: correlation coefficient map
$\boldsymbol{M}_{\mathrm{mask}}$: final binary image with demarcated duplicated regions

Proposed CMFD Scheme:

1. Read Image

$\boldsymbol{I} \leftarrow$tested image
$\boldsymbol{M}_{\mathrm{mask}} \leftarrow$image whose pixel values are 0

2. Feature Extraction

$[\boldsymbol{D}_{\mathrm{A-KAZE}}, \boldsymbol{P}_{\mathrm{A-KAZE}}] \leftarrow \mathrm{A-KAZE}(\boldsymbol{I})$
$[\boldsymbol{D}_{\mathrm{SURF}}, \boldsymbol{P}_{\mathrm{SURF}}] \leftarrow \mathrm{SURF}(\boldsymbol{I})$

3. Feature Matching

$[\boldsymbol{P}_{\mathrm{A_1}}, \boldsymbol{P}_{\mathrm{A_2}}] \leftarrow \mathrm{g2NN}(\boldsymbol{D}_{\mathrm{A-KAZE}}, \boldsymbol{P}_{\mathrm{A-KAZE}})$
$[\boldsymbol{P}_{\mathrm{S_1}}, \boldsymbol{P}_{\mathrm{S_2}}] \leftarrow \mathrm{g2NN}(\boldsymbol{D}_{\mathrm{SURF}}, \boldsymbol{P}_{\mathrm{SURF}})$
$[\boldsymbol{P}_1, \boldsymbol{P}_2] \leftarrow \mathrm{pos\_combination}(\boldsymbol{P}_{\mathrm{A_1}}, \boldsymbol{P}_{\mathrm{A_2}}, \boldsymbol{P}_{\mathrm{S_1}}, \boldsymbol{P}_{\mathrm{S_2}})$

4. Eliminating False Matches with RANSAC

$[\boldsymbol{T}, \boldsymbol{P}_{\mathrm{inliers}}] \leftarrow \mathrm{RANSAC}(\boldsymbol{P}_1, \boldsymbol{P}_2)$

5. Calculate Correlation Coefficient Map

$\boldsymbol{M}_{\mathrm{map}} \leftarrow \mathrm{corr\_map}(\boldsymbol{I}, \boldsymbol{T})$

6. Filtering and Mathematical Morphology Operation

$\boldsymbol{M}_{\mathrm{mask}} \leftarrow \mathrm{post\_processing}(\boldsymbol{M}_{\mathrm{map}})$

7. Judgment

if $\boldsymbol{M}_{\mathrm{mask}}$ is black
    $\boldsymbol{I}$ is an authentic image
else
    $\boldsymbol{I}$ is a tampered image
end if

4. Performance Analysis

In this paper, all the experiments were conducted using MATLAB R2015b and a computer with an Intel Core i5-4690 processor at 3.50 GHz and 8 GB of memory. The GRIP dataset created by Cozzolino et al. [17] and the FAU dataset created by Christlein et al. [45] were used to evaluate the performance of the proposed method. The GRIP dataset included 80 plain tampered images and 80 corresponding authentic images with a size of 1024×768 in PNG format. The tampered regions were also separately saved as images in PNG format, and the location of the tampered regions in tampered images was saved in text form. The FAU dataset included 48 groups of tampered images with various distortions and 48 corresponding authentic images of different sizes in both PNG and JPEG formats. Binary images that demarcated the tampered regions were provided that correspond to the tampered

images in these two datasets. In the following experiments, the parameter settings were as follows: $T_{\mathrm{A-KAZE}} = 0.0001$, $T_{\mathrm{SURF}} = 0.1$, $V_{\mathrm{t}} = 0.6$, $T_{\mathrm{bin}} = 0.8$, and $N_{\mathrm{pixel}} = 50$.

4.1. Performance Indexes

The precision p, recall r, and F score [45] metrics were used to evaluate the performance of the CMFD methods. The calculations are shown in Equation (20):

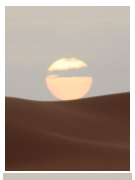
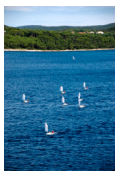
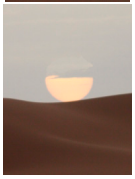
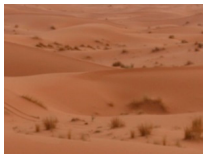
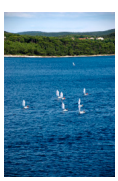
$$r = \frac{N_{\mathrm{DD}}}{N_{\mathrm{DD}} + N_{\mathrm{DA}}}, \; p = \frac{N_{\mathrm{DD}}}{N_{\mathrm{DD}} + N_{\mathrm{AD}}}, \; F = 2 \cdot \frac{p \cdot r}{p + r}, \tag{20}$$

where N_{DD} is the number of correctly detected doctored pixels, N_{DA} is the number of falsely detected authentic pixels, and N_{AD} denotes the number of falsely detected doctored pixels. Similar indexes were also used to evaluate the performance of the CMFD schemes at the image level as follows: N_{DD} denotes the number of correctly detected doctored images, N_{DA} denotes the number of falsely detected doctored images, and N_{AD} denotes the number of falsely detected authentic images. A larger recall r, precision p, and F score indicate higher accuracy of the CMFD scheme.

4.2. Copy-Move Forgery Detection

To compare the proposed method with other keypoint-based CMFD methods, an experiment was conducted in which the proposed method is performed as described above except for the feature extraction step. For the feature extraction, the SIFT, SURF, A-KAZE, BRIEF (binary robust independent elementary features) [46], BRISK (binary robust invariant scalable keypoints) [47], and hybrid features used in this paper were evaluated to compare the performance of the method. The number of pairs of matched points within duplicated regions (as shown in the fourth image of the first row of Figure 3) and related data are listed in Table 2. The four images shown in Table 2 are from the datasets provided by References [17,45]. The data demonstrate that the hybrid features consisting of A-KAZE and SURF obtain more matched points within the duplicated regions than SIFT, SURF, A-KAZE, BRIEF, and BRISK do with their default parameters. Thus, the hybrid features can be used to estimate the geometric transformation matrix more accurately. It should be noted that KAZE is not included in Table 2 because KAZE and A-KAZE perform similarly in this scheme. The first group of images was correctly identified as tampered images using the hybrid features; however, the correct judgment cannot be obtained using the other tested keypoints.

Table 2. Number of pairs of matched points within duplicated regions.

Features \ **Images**				
				
SIFT	0	17	515	16
SURF	0	2	145	4
A-KAZE	0	24	390	5
BRIEF	2	21	188	5
BRISK	0	2	11	0
Hybrid features	99	158	691	16

In addition, the duplicated regions were located using the post-processing method mentioned above. We selected several tampered images and corresponding ground-truth maps from the tested datasets [17,45] to demonstrate the effectiveness of the proposed scheme; these are shown in Figure 5, and their names are given in the subtitles of Figure 5. The proposed method can detect duplicated regions in doctored images with both single and multiple image copy-move forgeries, illustrating the effectiveness of the presented scheme. In addition, the detection results of other CMFD algorithms are also shown in Figure 5, demonstrating that regions demarcated by the proposed scheme are more accurate than the other methods [17,33] and demarcating the locations of tampered regions more clearly than the block-based CMFD method [17]. The precision p, recall r, and F score were also calculated for the image examples in Figure 5, and the results are listed in Table 3. The data in Table 3 also show that the proposed scheme achieves state-of-the-art CMFD performance at the pixel level. In addition, the precision p of the proposed method is higher than those of the other tested CMFD methods. To evaluate the overall use of the proposed scheme at the image level, we conducted an experiment in which the presented scheme was performed on every plain tampered image and the corresponding authentic image in the GRIP [17] and FAU [45] datasets. Similar to Table 3, the precision p, recall r, and F score were calculated and are listed in Table 4. The proposed method achieves a similar performance to that of the other recent tested CMFD schemes at the image level. The data in Table 4 illustrate the advantages of the block-based CMFD scheme [17], in which the most relevant evaluation criteria are higher than those of other schemes. In particular, the recall r of the presented scheme on FAU dataset is higher than that of other methods. However, several images exist in datasets [17,45] that cannot be judged correctly.

Table 3. Precision p, recall r, and F score of the images shown in Figure 5.

Methods		Figure 5a	Figure 5b	Figure 5c	Figure 5d	Figure 5e
Zandi et al. [33]	p	0.9456	0.9662	0.8589	0.7428	0.9468
	r	0.9672	0.9909	0.9845	0.9646	0.9695
	F	0.9563	0.9784	0.9174	0.8393	0.9580
Cozzolino et al. [17]	p	0.9721	0.9688	0.9577	0.9491	0.9791
	r	0.9656	0.9850	0.9658	0.9761	0.9571
	F	0.9688	0.9769	0.9617	0.9624	0.9680
Amerini et al. [22]	p	0.5598	0.5747	0.9081	0.4844	0.9528
	r	0.9767	0.9890	0.9814	0.9659	0.9925
	F	0.7117	0.7269	0.9433	0.6452	0.9722
Proposed	p	0.9799	0.9948	0.9954	0.9972	0.9761
	r	0.9396	0.9431	0.9472	0.9394	0.9797
	F	0.9594	0.9682	0.9707	0.9674	0.9779

Table 4. Precision p, recall r, and F score of the images in the GRIP and FAU datasets.

Methods	GRIP			FAU		
	p	r	F	p	r	F
Zandi et al. [33]	0.7692	1	0.8695	0.7188	0.9583	0.8214
Cozzolino et al. [17]	0.9286	0.9750	0.9512	0.9167	0.9167	0.9167
Amerini et al. [22]	0.8837	0.9500	0.9157	0.8936	0.8750	0.8842
Proposed	0.9176	0.9750	0.9454	0.8824	0.9375	0.9091

Figure 5. Detection results of different CMFD methods and the proposed scheme. First row: the authentic images; Second row: the tampered images; Third row: ground-truth maps of the tampered images; Fourth row: the detection results of Zandi et al.'s method [33]; Fifth row: the detection results of Cozzolino et al.'s method [17]; Sixth row: the detection results of Amerini et al.'s method [22]; Seventh row: the detection results of the proposed scheme. (**a**) IMG_C01_010; (**b**) IMG_C01_018; (**c**) barrier; (**d**) IMG_C02_041; (**e**) kore.

The experiments mentioned above tested the performance of the proposed method and other CMFD schemes [17,22,33] only with respect to plain copy-move forgery images. However, the intruder may process the copied region (rotation, scaling, etc.) before pasting it to another region. Thus, it is also necessary to evaluate the ability of CMFD schemes against rotation and scaling. The data listed in Figures 6 and 8 are the average values of precision p, recall r, and F score obtained by calculating the precision p, recall r, and F score of the detection results of the image with corresponding distortions, barrier, clean walls, extension, fountain, and supermarket. In the "rotation experiment", the copied region is rotated by 2° to 10° with a step size of 2° and pasted into another region within the same image. In the "scaling test", the copied region is scaled by different factors and pasted into another region within the same image. The results of the rotation and scaling are presented in Figure 6a,b, respectively. From Figure 6, it can be seen that the precision p and F scores of the proposed scheme are noticeably higher than those of the other CMFD schemes [17,22,33]. The region located using the proposed method was more accurate than those of the other tested CMFD methods under rotation and scaling. However, the recall values r of the proposed method were lower than those of the other tested

CMFD methods [17,22,33] after the copied region distorted by rotation and scaling. Two distorted images with rotation and scaling and their corresponding detection results of proposed method are given in Figure 7.

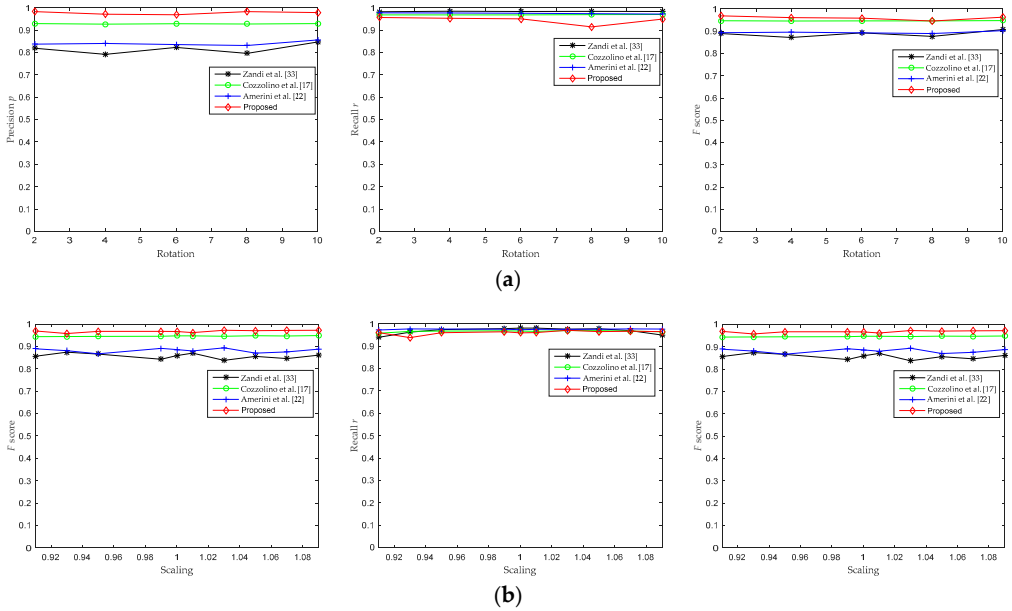


Figure 6. Detection results (precision p, recall r, and F score) of the proposed scheme and different CMFD methods under different distortions: (**a**) rotation; (**b**) scaling.

Figure 7. Detection results of the proposed scheme for distorted images with rotation and scaling: (**a**) barrier with rotation (10°); (**b**) detection result of (**a**); (**c**) clean walls with scaling (1.09); (**d**) detection result of (**c**).

4.3. Robustness Test

In addition, image post-processing manipulations such as image blurring, noise addition, JPEG compression, and hybrid image manipulation are usually used to conceal the evidence of image copy-move tampering. Therefore, robustness experiments were conducted to test the ability of CMFD methods to detect tampering in images with post-processing manipulations. In the following experiments, image blurring, noise addition, and JPEG compression were used to process the copy-move forgery images with individual different parameters. The images in the dataset were resaved in JPEG format with different quality factors (QFs); QF ranged from 30 to 100 with a step size of 10. Next, the various CMFD methods were performed on these images, and the detection results were displayed in Figure 8a. In the "noise addition test", Gaussian noise was added to images with zero-mean and different variances (ranging from 0.005 to 0.02 with a step size of 0.005). The detection results of noise addition of CMFD methods are shown in Figure 8b. In the "blurring experiment", a filter with a radius ranging from 0.5 to 2.5 with a step size of 0.5 was used to process the forged images. The detection results of image blurring using the proposed method and the other tested CMFD schemes are shown in Figure 8c. Three distorted images with different distortions and their corresponding detection results of proposed method are given in Figure 9.

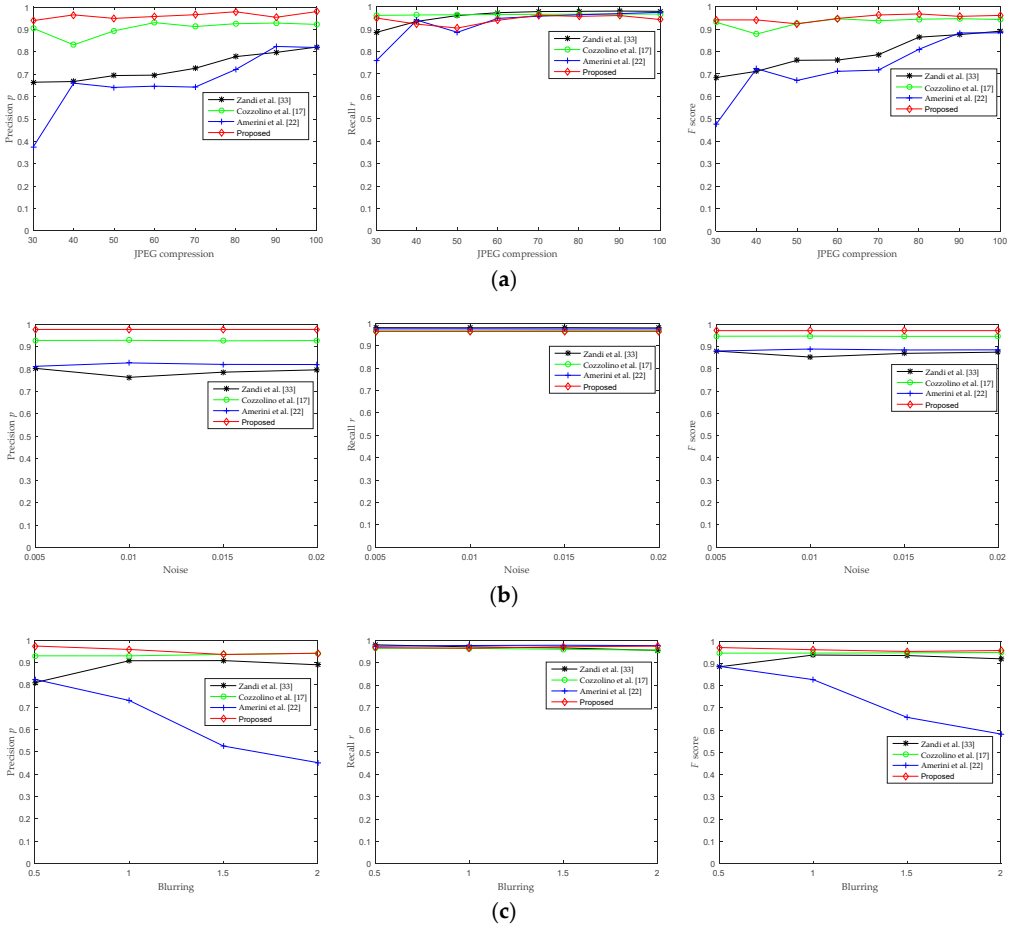


Figure 8. Detection results (precision p, recall r, and F score) of the proposed method and tested schemes under distorted copy-move images with different image post-processing manipulations: (**a**) JPEG compression; (**b**) noise addition; (**c**) image blurring.

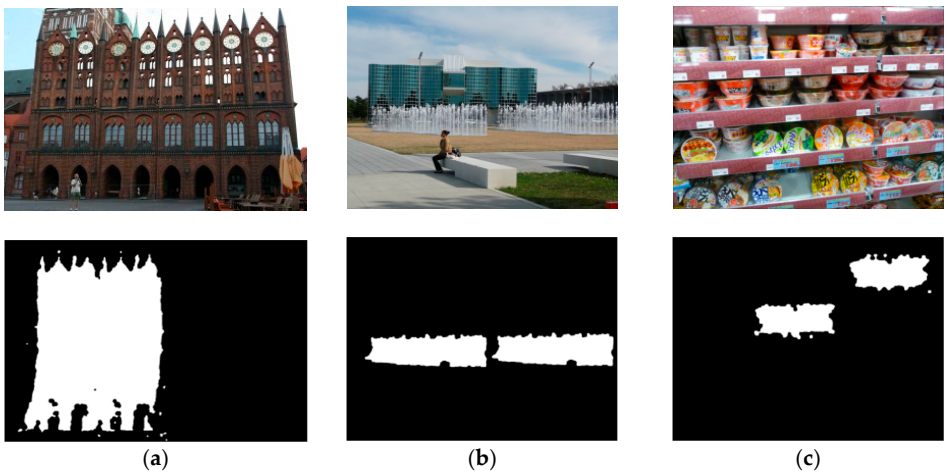


Figure 9. Detection results of the proposed scheme for distorted images with different distortions: (**a**) extension with JPEG compression (40) and its detection result; (**b**) fountain with noise (0, 0.015) and its detection result; (**c**) supermarket with blurring (2) and its detection result.

First, the proposed method can detect duplicated regions containing forged images with distortions such as image blurring, noise addition, and JPEG compression. These results demonstrate that the proposed method is robust to image post-processing manipulations. Figure 8 shows that the

proposed scheme is superior to the method [22] with respect to JPEG compression, noise addition, and image blurring. However, the proposed scheme has both advantages and disadvantages compared to other tested CMFD methods [22,33]. The precision p and F scores of the proposed method are higher than those of other CMFD methods, demonstrating that the proposed scheme is superior to other tested CMFD schemes [17,22,33] against JPEG compression, noise, and image blurring.

The proposed scheme also detects tampered images where the duplicated regions are distorted with hybrid image manipulations. However, the post-processing method is not as effective with these manipulations as with the images in previous experiments; therefore, the duplicated regions are indicated with lines and points in Figure 10. The detection result of the proposed method for the tampered image distorted with scaling (0.7) and rotation (50°) is shown in Figure 10a. The detection result of the proposed scheme for the tampered image distorted with rotation (90°) and JPEG compression (80) is shown in Figure 10b. The detection result of the proposed method for the tampered image distorted with scaling (0.6) and blurring (1.5) is shown in Figure 10c. The detection result of the proposed scheme for the tampered image distorted with scaling (0.8), rotation (120°), and JPEG compression (70) is shown in Figure 10d.

(**a**) (**b**) (**c**) (**d**)

Figure 10. Detection results of the proposed scheme with distorted copy-move images with hybrid image post-processing manipulations: (**a**) scaling (0.7) + rotation (50°); (**b**) rotation (90°) + JPEG compression (80); (**c**) scaling (0.6) + blurring (1.5); (**d**) scaling (0.8) + rotation (120°) + JPEG compression (70).

In summary, the proposed method can detect duplicated regions in forged images with single and multiple copy-move forgeries. In addition, the proposed scheme can detect pasted regions altered by rotation and scaling, and it is robust to image post-processing manipulations such as image blurring, noise addition, JPEG compression, and hybrid image manipulation. Compared to the other tested CMFD methods [17,22,33], the proposed method exhibits both advantages and disadvantages. The proposed method overcomes the defect of most keypoint-based CMFD methods, which are unable to detect sufficient points in smooth tampered regions. In particular, the detected forged regions of the proposed method are more accurate than those of the other tested methods with respect to the precision p and F score in most situations. In addition, the proposed scheme is more robust against image blurring than the other tested methods in some situations. Although the proposed method can correctly judge the forged images, the performance indexes of the proposed scheme are lower than other tested CMFD schemes in some situations. These discrepancies form one of the directions for investigating ways to improve the proposed method.

The comparison with other CMFD methods shows that current CMFD methods have several issues. Too many similar regions in images make detecting duplicated regions difficult, and they can easily falsely detect regions. In particular, forged images with small duplicated regions are difficult to detect. In addition, clearly locating the duplicated regions found by keypoint-based CMFD methods is an area that requires further exploration because mathematical morphology operations with fixed

parameters are difficult to generalize in various situations. Most importantly, a faster feature matching algorithm is also necessary to save time when finding similar features. With the development of deep learning techniques such as CNN and SVM, their applications in the image forensics field are also worth exploring.

5. Conclusions

In this paper, a CMFD scheme based on A-KAZE and SURF was proposed. A-KAZE and SURF features have the advantages of fast feature extraction and robustness, which were used in this scheme to find regions within the tampered image using copy-move forgery. To obtain sufficient points in the smooth regions, the response thresholds for A-KAZE and SURF were set to small values instead of their default parameters. This approach allows for the detection of the duplicated regions within a tampered image even in smooth regions. In particular, a new correlation map was presented in this paper that can demarcate the duplicated regions with closed regions in tampered images. However, when the tampered region is distorted by image manipulation, this may not be as effective as that in plain image copy-move forgery detection. The experimental results demonstrate that the proposed scheme can detect forged regions in tampered images even when the tampered region is distorted by image manipulations such as image blurring, rotation, noise addition, scaling, JPEG compression, and hybrid image manipulations. Compared to other tested CMFD methods, the proposed method exhibits both advantages and disadvantages. For example, it is more accurate than the other tested CMFD methods in some aspects but inferior to other CMFD schemes in others. However, there are several directions by which the proposed method could be improved in future work. The process of obtaining sufficient points is time-consuming. This issue could be solved by finding a faster feature matching algorithm to identify similar features. In addition, similar features may be found by using image-matching techniques from other fields [48,49]. In keypoint-based CMFD methods, the duplicated regions should be demarcated with closed regions with a clear boundary; however, this remains a problem to be solved. As deep learning techniques become increasingly popular, their applications in the field of multimedia forensics will continue to be explored.

Author Contributions: C.W. and Z.Z. conceived the algorithm and designed the experiments; Z.Z. performed the experiments; C.W. and X.Z. analyzed the results; Z.Z. drafted the manuscript; C.W., Z.Z., and X.Z. revised the manuscript. All authors read and approved the final manuscript.

Funding: This work was funded by the National Natural Science Foundation of China (Nos. 61702303 and 61201371) and the Shandong Provincial Natural Science Foundation, China (Nos. ZR2017MF020 and ZR2015PF004).

Acknowledgments: The authors thank P. F. Alcantarilla et al. for providing the open-source code of A-KAZE. The authors thank V. Christlein et al. and D. Cozzolino et al. for providing copy-move forgery datasets.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Samadzadegan, F.; Hasani, H.; Schenk, T. Determination of optimum classifier and feature subset in hyperspectral images based on ant colony system. *Photogramm. Eng. Remote Sens.* **2012**, *78*, 1261–1273. [CrossRef]
2. Nagarajan, S.; Schenk, T. Feature-based registration of historical aerial images by area minimization. *ISPRS J. Photogramm. Remote Sens.* **2016**, *116*, 15–23. [CrossRef]
3. Klinger, T.; Rottensteiner, F.; Heipke, C. Probabilistic multi-person localisation and tracking in image sequences. *ISPRS J. Photogramm. Remote Sens.* **2017**, *127*, 73–88. [CrossRef]
4. Janowski, A.; Nagrodzka-Godycka, K.; Szulwic, J.; Ziółkowski, P. Remote sensing and photogrammetry techniques in diagnostics of concrete structures. *Comput. Concr.* **2016**, *18*, 405–420. [CrossRef]
5. GIMP—GNU Image Manipulation Program. Available online: https://www.gimp.org/ (accessed on 30 November 2018).
6. Photoshop. Available online: https://www.photoshop.com/ (accessed on 30 November 2018).

7. Lin, X.; Li, J.H.; Wang, S.L.; Liew, A.W.C.; Cheng, F.; Huang, X.S. Recent advances in passive digital image security forensics: A brief review. *Engineering* **2018**, *4*, 29–39. [CrossRef]
8. Ziaullah, M.; Shetty, P.; Kamal, S. Image feature based authentication and digital signature for wireless data transmission. In Proceedings of the 6th International Conference on Computer Communication and Informatics, Coimbatore, India, 7–9 January 2016.
9. Liu, Z.H.; Huang, J.W.; Sun, X.M.; Qi, C.D. A security watermark scheme used for digital speech forensics. *Multimedia Tools Appl* **2017**, *76*, 9297–9317. [CrossRef]
10. Wang, Q.; Zhang, R. Double JPEG compression forensics based on a convolutional neural network. *EURASIP J. Inf. Secur.* **2016**, *2016*, 23. [CrossRef]
11. Chen, J.S.; Kang, X.G.; Liu, Y.; Wang, Z.J. Median filtering forensics based on convolutional neural networks. *IEEE Signal Process Lett.* **2015**, *22*, 1849–1853. [CrossRef]
12. Fridrich, J.; Soukal, D.; Lukáš, J. Detection of copy-move forgery in digital images. In Proceedings of the Digital Forensic Research Workshop, Cleveland, OH, USA, 6–8 August 2003; pp. 55–61.
13. Huang, Y.P.; Lu, W.; Sun, W.; Long, D.Y. Improved DCT-based detection of copy-move forgery in images. *Forensic Sci. Int.* **2011**, *206*, 178–184. [CrossRef]
14. Bi, X.L.; Pun, C.M.; Yuan, X.C. Multi-level dense descriptor and hierarchical feature matching for copy-move forgery detection. *Inf. Sci.* **2016**, *345*, 226–242. [CrossRef]
15. Zhong, J.L.; Gan, Y.F.; Young, J.; Huang, L.; Lin, P.Y. A new block-based method for copy move forgery detection under image geometric transforms. *Multimedia Tools Appl.* **2017**, *76*, 14887–14903. [CrossRef]
16. Zhong, J.L.; Gan, Y.F. Detection of copy-move forgery using discrete analytical Fourier–Mellin transform. *Nonlinear Dyn.* **2016**, *84*, 189–202. [CrossRef]
17. Cozzolino, D.; Poggi, G.; Verdoliva, L. Efficient dense-field copy-move forgery detection. *IEEE Trans. Inf. Forensics Secur.* **2015**, *10*, 2284–2297. [CrossRef]
18. Deng, J.H.; Yang, J.X.; Weng, S.W.; Gu, G.S.; Li, Z. Copy-move forgery detection robust to various transformation and degradation attacks. *KSII Trans. Internet Inf. Syst.* **2018**, *12*, 4467–4486.
19. Mahmood, T.; Irtaza, A.; Mehmood, Z.; Tariq Mahmood, M. Copy-move forgery detection through stationary wavelets and local binary pattern variance for forensic analysis in digital images. *Forensic Sci. Int.* **2017**, *279*, 8–21. [CrossRef]
20. Fadl, S.M.; Semary, N.A. Robust copy-move forgery revealing in digital images using polar coordinate system. *Neurocomputing* **2017**, *265*, 57–65. [CrossRef]
21. Amerini, I.; Ballan, L.; Caldelli, R.; Del Bimbo, A.; Serra, G. A SIFT-based forensic method for copy-move attack detection and transformation recovery. *IEEE Trans. Inf. Forensics Secur.* **2011**, *6*, 1099–1110. [CrossRef]
22. Amerini, I.; Ballan, L.; Caldelli, R.; Del Bimbo, A.; Del Tongo, L.; Serra, G. Copy-move forgery detection and localization by means of robust clustering with J-Linkage. *Signal Process Image Commun.* **2013**, *28*, 659–669. [CrossRef]
23. Jin, G.N.; Wan, X.X. An improved method for SIFT-based copy-move forgery detection using non-maximum value suppression and optimized J-Linkage. *Signal Process Image Commun.* **2017**, *57*, 113–125. [CrossRef]
24. Shivakumar, B.L.; Santhosh Baboo, S. Detection of region duplication forgery in digital images using SURF. *Int. J. Comput. Sci. Issues* **2011**, *8*, 199–205.
25. Bay, H.; Ess, A.; Tuytelaars, T.; Van Gool, L. Speeded-up robust features (SURF). *Comput. Vis. Image Underst.* **2008**, *110*, 346–359. [CrossRef]
26. Jaberi, M.; Bebis, G.; Hussain, M.; Muhammad, G. Accurate and robust localization of duplicated region in copy-move image forgery. *Mach. Vis. Appl.* **2014**, *25*, 451–475. [CrossRef]
27. Yu, L.Y.; Han, Q.; Niu, X.M. Feature point-based copy-move forgery detection: Covering the non-textured areas. *Multimedia Tools Appl.* **2016**, *75*, 1159–1176. [CrossRef]
28. Uliyan, D.M.; Jalab, H.A.; Wahab, A.W.A.; Sadeghi, S. Image region duplication forgery detection based on angular radial partitioning and Harris key-points. *Symmetry* **2016**, *8*, 62. [CrossRef]
29. Ulutas, G.; Muzaffer, G. A new copy move forgery detection method resistant to object removal with uniform background forgery. *Math. Probl. Eng.* **2016**, *2016*, 3215162. [CrossRef]
30. Alcantarilla, P.F.; Nuevo, J.; Bartoli, A. Fast explicit diffusion for accelerated features in nonlinear scale spaces. In Proceedings of the 24th British Machine Vision Conference, Bristol, UK, 9–13 September 2013; pp. 1–11.
31. Yang, F.; Li, J.W.; Lu, W.; Weng, J. Copy-move forgery detection based on hybrid features. *Eng. Appl. Artif. Intell.* **2017**, *59*, 73–83. [CrossRef]

32. Alcantarilla, P.F.; Bartoli, A.; Davison, A.J. KAZE features. In Proceedings of the 12th European Conference on Computer Vision, Florence, Italy, 7–13 October 2012; pp. 214–227.
33. Zandi, M.; Mahmoudi-Aznaveh, A.; Talebpour, A. Iterative copy-move forgery detection based on a new interest point detector. *IEEE Trans. Inf. Forensics Secur.* **2016**, *11*, 2499–2512. [CrossRef]
34. Yang, B.; Sun, X.; Guo, H.; Xia, Z.; Chen, X. A copy-move forgery detection method based on CMFD-SIFT. *Multimedia Tools Appl.* **2018**, *77*, 837–855. [CrossRef]
35. Li, J.; Yang, F.; Lu, W.; Sun, W. Keypoint-based copy-move detection scheme by adopting MSCRs and improved feature matching. *Multimedia Tools Appl.* **2017**, *76*, 20483–20497. [CrossRef]
36. Zhao, R.; Yan, R.Q.; Chen, Z.H.; Mao, K.Z.; Wang, P.; Gao, R.X. Deep learning and its applications to machine health monitoring. *Mech. Syst. Signal Process.* **2019**, *115*, 213–237. [CrossRef]
37. Bakator, M.; Radosav, D. Deep learning and medical diagnosis: A review of literature. *Multimodal Technol. Interact.* **2018**, *2*, 47. [CrossRef]
38. Ammour, N.; Alhichri, H.; Bazi, Y.; Benjdira, B.; Alajlan, N.; Zuair, M. Deep learning approach for car detection in UAV imagery. *Remote Sens.* **2017**, *9*, 312. [CrossRef]
39. Li, S.X.; Zhang, Z.L.; Li, B.; Li, C.W. Multiscale rotated bounding Box-based deep learning method for detecting ship targets in remote sensing images. *Sensors* **2018**, *18*, 2702. [CrossRef] [PubMed]
40. Rao, Y.; Ni, J.Q. A deep learning approach to detection of splicing and copy-move forgeries in images. In Proceedings of the 8th IEEE International Workshop on Information Forensics and Security, Abu Dhabi, UAE, 4–7 December 2016. no. 7823911.
41. Grewenig, S.; Weickert, J.; Bruhn, A. From box filtering to fast explicit diffusion. *Lect. Notes Comput. Sci.* **2010**, *6376*, 533–542.
42. Qu, Z.; Bu, W.; Liu, L. The algorithm of seamless image mosaic based on A-KAZE features extraction and reducing the inclination of image. *IEEJ Trans. Electr. Electron. Eng.* **2018**, *13*, 134–146. [CrossRef]
43. Yang, X.; Cheng, K.T. LDB: An ultra-fast feature for scalable augmented reality on mobile devices. In Proceedings of the 11th IEEE and ACM International Symposium on Mixed and Augmented Reality, Atlanta, GA, USA, 5–8 November 2012; pp. 49–57.
44. Wang, X.Y.; Li, S.; Liu, Y.N.; Niu, Y.; Yang, H.Y.; Zhou, Z.L. A new keypoint-based copy-move forgery detection for small smooth regions. *Multimedia Tools Appl.* **2017**, *76*, 23353–23382. [CrossRef]
45. Christlein, V.; Riess, C.; Jordan, J.; Riess, C.; Angelopoulou, E. An evaluation of popular copy-move forgery detection approaches. *IEEE Trans. Inf. Forensics Secur.* **2012**, *7*, 1841–1854. [CrossRef]
46. Calonder, M.; Lepetit, V.; Strecha, C.; Fua, P. BRIEF: Binary robust independent elementary features. In Proceedings of the 11th European Conference on Computer Vision, Heraklion, Crete, Greece, 5–11 September 2011; pp. 778–792.
47. Leutenegger, S.; Chli, M.; Siegwart, R.Y. BRISK: Binary robust invariant scalable keypoints. In Proceedings of the IEEE International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 2548–2555.
48. Dickscheid, T.; Förstner, W. A trainable markov random field for low-level image feature matching with spatial relationships. *Photogramm. Fernerkund. Geoinf.* **2013**, *4*, 269–283. [CrossRef]
49. Abduljabbar, Z.A.; Jin, H.; Ibrahim, A.; Hussien, Z.A.; Hussain, M.A.; Abbdal, S.H.; Zou, D.Q. SEPIM: Secure and efficient private image matching. *Appl. Sci.* **2016**, *6*, 213. [CrossRef]