Fusion of Block and Keypoints Based Approaches for Effective Copy-move Image Forgery Detection

Jiangbin Zheng[1], Yanan Liu[1], Jinchang Ren[2], Tingge Zhu[1], Yijun Yan[2], and Heng Yang[3]

[1] Dept. of Computer Science and Engineering, School of Computers, Northwestern Polytechnical University, Xi'an, 710072, China
[2] Dept. of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, G1 1XW, U.K.

[3] Xi'an Communications Institute, 710106, Xi'an, China

Abstract. Keypoint-based and block-based methods are two main categories of techniques for detecting copy-move forged images, one of the most common digital image forgery schemes. In general, block-based methods suffer from high computational cost due to the large number of image blocks used and fail to handle geometric transformations. On the contrary, keypoint-based approaches can overcome these two drawbacks yet are found difficult to deal with smooth regions. As a result, fusion of these two approaches is proposed for effective copy-move forgery detection. First, our scheme adaptively determines an appropriate initial size of regions to segment the image into non-overlapped regions. Feature points are extracted as keypoints using the scale invariant feature transform (SIFT) from the image. The ratio between the number of keypoints and the total number of pixels in that region is used to classify the region into smooth or non-smooth (keypoints) regions. Accordingly, block based approach using Zernike moments and keypoint based approach using SIFT along with filtering and post-processing are respectively applied to these two kinds of regions for effective forgery detection. Experimental results show that the proposed fusion scheme outperforms the keypoint-based method in reliability of detection and the block-based method in efficiency.

Keywords: Image forensics, copy-move image forgery detection, adaptive fusion, SIFT, Zernike moments.

1 INTRODUCTION

With the increasing availability and functionalities of image processing software, image editing and refinement becomes an enjoyable hobby and popular trend for many people. However, this has also caused another issue in term of the reliability and data integrity of digital images. In some application areas, such as news reports, court certification and biometrics imaging, the authenticity of the image is particularly important (Christlein et al. 2012; Huynh et al. 2015; Sencar and Memon 2008).

In recent years, more and more researchers have begun to focusing on the problem of digital image forensics. According to (Sencar and Memon 2008), digital image forensics can be divided into three branches, i.e. image source identification to identify which device was used to capture an image (model or exemplar of scanner, of digital camera), discrimination of computer generated images (to detect if an image is natural or synthetic), and image forgery detection (to discern if an image has been intentionally modified by human intervention). Among many forgery techniques, copy-move is one of the most commonly used, by which the content of the image is copied from one region and pasted into another area in the same image. This forgery may be accompanied with geometric transforms and post-processing including rotation, scale, JPEG compression, noise addition, etc. Due to the local similarity in terms of color and texture, it is very difficult for human eyes to distinguish the forgery from the original.

Existing techniques for copy-move forgery detection can be classified into two categories, i.e. block-based and keypoint-based methods. The block-based methods usually extract features from overlapping blocks of the image, a number of features have been proposed for forgery detection.

In (Christlein et al. 2012), these features are grouped into four classes: frequency domain-based, dimensionality reduction-based, intensity-based, and moment-based features. Feature transformed approaches include block-based discrete cosine transform (DCT) (Fridrich et al. 2003), principal component analysis (PCA) (Popescu and Farid 2004), discrete wavelet transform (DWT) (Bashar et al. 2010), Dyadic Wavelet Transform (DyWT) (Muhammad et al. 2011), and Fourier-MellinTransform (FMT) (Bayram et al. 2009). In addition, some non-typical block-based features include singular value decomposition (SVD) (Kang and Wei 2008), and blur-invariant moments (BLUR) (Mahdian and Saic 2007).

For non-transform based approaches, usually features include rotation and flipping invariant uniform local binary patterns (LBP) (L. Li et al. 2013), average color and directional information of blocks (Luo et al. 2006). A 9-dimensional vector was proposed in Lin et al. (2009) which can detect the rotation with a fixed angle but not with arbitrary angles. In Wang et al. (2009) and Ryu et al. (2010), the first four Hu moments (HU) and rotation invariant Zernike moments are respectively used. After examining and assessing 15 most prominent feature sets, Christlein et al. (2012) recommended Zernike moments as block features due to its relatively small memory footprint and robustness to transformation.

With a large number of image blocks, the time consumption of block-based methods is high, especially with the increasing size of images. In addition, most of these methods show lack of robustness against generic geometric transformations, i.e. affine and projective transforms in particular. Keypoint-based methods, however, can relatively reliably compensate these shortcomings. In (Amerini et al. 2013; Amerini et al. 2011; Huang et al. 2008) Scale-invariant feature transform (SIFT) and (Bo et al. 2010; Jing and Shao 2012) Speed-up robust features (SURF) were used to detect copy-move forgery. Due to the number of keypoints is much less than the number of blocks, computational complexity is greatly reduced. In (Amerini et al. 2011) the authors also proposed to use g2NN to deal with multiple copies, the hierarchical clustering and Random Sample Consensus (RANSAC) were employed to filter out outliers and significantly improved the detection accuracy. In Kumar et al. (2015) a hybrid approach was proposed, where SURF was used to detect the keypoints in the image and Binary Robust Invariant Scalable Keypoints (BRISK) features were used to represent corresponding features at these keypoints.

Although keypoint-based methods are robust to geometric transformation and cost low, they may not work in dealing with smooth regions. At present, some researchers have proposed the combined methods to detect forgery. In Pun et al. (2015), SIFT features were extracted from each block and considered as block features, yet it still failed to detect forgery of copied smooth regions. In Zahra (2012) Zernike moments were combined with SIFT, yet how to identify smooth regions were missing.

In summary, keypoint-based methods may fail to work for copied smooth regions. On the contrary, block-based methods work well in such cases yet it will bring high computation cost. To this end, fusion of block-based method and keypoint-based method to respectively deal with smooth regions and keypoints regions becomes a natural choice in this context. After examining 15 most prominent feature sets, Christlein et al. (2012) recommended Zernike moments as the best block features due to its relatively

small memory footprint and high reliability. Moreover, their experimental results also showed that SIFT produced the best results among those keypoint-based methods. As a result, Zernike moment and SIFT are chosen in our fusion based approach for effective forgery detection. As shown in Fig. 1, our approach can successfully detect forgery image even in smooth regions.

The rest of this paper is organized as follows. In Section 2 the proposed approach is presented in detail. In Section 3, the experimental results are discussed. Finally, some concluding remarks are drawn in Section 4.

2 The proposed approach

In the proposed approach, an image is first adaptively divided into non-overlapped regions, using simple linear iterative clustering (SLIC) algorithm. Then, SIFT is used to detect keypoints in the whole image, and based on whether the ratio of keypoints' number to the pixels' number is less than a threshold, a region is classified as smooth region or keypoints region. Afterwards, a multiple keypoints matching procedure is performed in keypoints regions to decide candidate forgery regions, and RANSAC is used to prune outlier. Finally, if there are more than two smooth regions in the image, Zernike moments are used as block features to detect forgery in smooth regions. The technical implementation of the proposed approach is presented in detail as follows.

2.1 Adaptive Image Segmentation

As the copy-moved regions always are semantically meaningful, we firstly segment the image into semantically independent non-overlapping regions by using the simple linear iterative clustering algorithm (SLIC) (Achanta et al. 2012). We choose SLIC for its low computational complexity, which provides a simple and efficient k-means clustering method to segment the image into visually homogeneous regions.

In our implementation, we employ VLFeat toolbox- vl_slic (Vedaldi and Fulkerson 2010) to segment images. The vl_slic approach has two parameters, i.e. *regionsize* and *regularizer*. *Regionsize* is the initial size of the superpixels (regions). *Regularizer* is the trades-off appearance for spatial regularity during clustering, where a larger value results in more spatial regularization. *Regularizer* in our experiments is set to 10000. In practice, images have different content and size, the initial size of the superpixels has

considerable influence on the segmentation result. In general, when the texture of the image is simple, the initial size of the superpixels can be set to be relatively large, which can ensure the superpixels get close to the edges. Furthermore, larger initial size implies a smaller number of blocks, which can reduce the computational cost when processing smooth region. In contrast, when the texture of the image is complicated, the initial size of the superpixels can be set to be relatively small to ensure good forgery detection results.

We adopt the approach proposed in Pun et al. (2015) to obtain the initial size of superpixels, adaptively. Four level Discrete Wavelet Transform (DWT) (Shensa 1992) is employed to analyze the frequency distribution of the image.

$$E_{LF} = \sum |\mathbf{CA}_4| \tag{1}$$

$$E_{HF} = \sum_{i} (\sum |\mathbf{CD}_i| + \sum |\mathbf{CH}_i| + \sum |\mathbf{CV}_i|) \qquad i = 1,2,3,4 \tag{2}$$

Where $\mathbf{CA}_4$ indicates the approximation coefficients at the 4-th level of DWT; $\mathbf{CD}_i$, $\mathbf{CH}_i$, $\mathbf{CV}_i$ indicate the detailed coefficients at the i^{th} level of DWT.

$$P_{LF} = \frac{E_{LF}}{E_{LF} + E_{HF}} \tag{3}$$

$$S\_init = \begin{cases} \sqrt{0.02 \times row \times col} & P_{LF} > 0.5 \\ \sqrt{0.01 \times row \times \mathrm{col}} & P_{LF} \le 0.5 \end{cases} \tag{4}$$

Where S_init means the initial size of the superpixels; row and col denote respectively the number of rows and columns of the image. Segmentation result is an array containing the superpixel identifier for each image pixel and each superpixel corresponds to a value, this array is named **Seg** which will be used in the formula (6). The visualization of the segmentation results is shown in Fig. 2.

2.2 Keypoints extraction and regions classification

In David G Lowe (1999), SIFT was used to detect keypoints in the whole image. In our work we employ siftDemoV4Toolbox (David G. Lowe 2004) to extract keypoints from the image. Fig. 3 (a) illustrates the results of the keypoints extraction, where we can find few such points in smooth regions.

We classify these non-overlapping regions using the following principals. If the ratio between the number of keypoints and the number of pixels in the region is less than a threshold T , it is classified as a smooth region. Otherwise, it is classified as a keypoints region.

$$R = \begin{cases} 1 & \text{if} \quad \dfrac{N_f}{N_p} \le T \\ 0 & \text{if} \quad \dfrac{N_f}{N_p} > T \end{cases} \tag{5}$$

Where N_f , N_p indicate the number of keypoints and pixels in the inspected region, respectively. $R = 1$ denotes that the region is smooth, and $R = 0$ signifies that it is keypoints region.

2.3 Detection in keypoints region

The flowchart of detection in keypoints region is given in Fig. 3, which contains multiple keypoints matching and outlier filtering as detailed below.

1) Multiple keypoints matching in keypoints region

In a real world situation, the same image region is cloned over and over (for example, Fig. 7 (H1)). In order to detect a region pasted in multiple regions, we use the g2NN strategy (Amerini et al. 2011) to find matching keypoints in keypoints regions. For a given keypoint, the similarity values (inverse cosine distance in the SIFT space) between it and other keypoints are calculated and form a similarity vector. The vector is sorted (lexicographic sort) and named as $\mathbf{S} = \{s_1, s_2, ...s_{n-1}\}$, where n is the number of keypoints. The detailed steps of g2NN strategy are shown in Table 1.

Here T_r is an empirical value, and according to Amerini et al. (2011) it is set to 0.5. If k ($i = k$) is the value where the procedure stops, the corresponding keypoints of $s_1, s_2 ....s_k$ are considered as candidate matching of the given keypoint. This scheme considers not only the single best-matching keypoint but also the k best-matching keypoints, thus it can detect multiple copies in keypoints regions.

2) Filtering outliers

As the similarity of adjacent keypoints is high, in the matching process adjacent keypoints should be excluded. To determine the shortest distance between two keypoints, existing approach is to set a threshold t_d , where only if the Euclidean distance between the two keypoints is greater than t_d the keypoints

will be used in comparison (Cheng et al. 2015; Cheng et al. 2014). However, the choice of threshold t_d is subjective, which neglects its relationship with the image size and content.

Based on the assumption that copy-move forgery does not occur in the same superpixel, a new solution is put forward in this paper. As long as (x, y) and (x_{match}, y_{match}) are not in the same superpixel, it determines they are comparable, i.e.,

$$\mathbf{Seg}(x, y) \sim= \mathbf{Seg}(x_{match}, y_{match}) \tag{6}$$

Seg is an array which is the segmentation result of SLIC algorithm. Obviously, this solution avoids the selection of the threshold t_d.

After detecting of matching keypoints in keypoints region, we employ hierarchical clustering (Amerini et al. 2011) and Random Sample Consensus algorithm (RANSAC) (Fischler and Bolles 1981) to filter outliers and remove false alarms. Fig. 3 (b) and Fig. 3 (c) illustrate the comparison of the matching keypoints before and after filtering, respectively.

2.4 Detection in smooth region

Considering that keypoints are scarce in smooth regions, special process is needed to deal with these regions to allow accurate detection. Actually, block-based approach is utilized for forgery detection in these regions as detailed in several steps below.

Step 1. To count the number of smooth regions. If there are more than two smooth regions we will perform the following steps, otherwise we will go directly to the post-processing stage.

Step 2. These regions are divided into overlapping blocks of $b \times b$, (b was set to 16 in our experiment). As shown in Fig. 4, the region in the red box is classified as smooth, and we calculate the bounding box of this region using x_{min}, y_{min}, x_{max}, y_{max}. This external rectangle is divided into overlapping blocks, where x_{min}, x_{max} indicate the minimum and maximum value of the region in horizontal coordinate, respectively. y_{min}, y_{max} indicate the minimum and maximum value of the region in vertical coordinate, respectively.

Step 3. As recommended in Christlein et al. (2012), Zernike moments are chosen as block features due to its relatively small memory footprint and high reliability. Zernike moments (Khotanzad and Hong

1990; Ryu et al. 2013; Ryu et al. 2010; Teh and Chin 1988) are rotation invariant and robust to noise, JPEG compression and even blurred.

The Zernike moments of order n with repetition m for a continuous image function $f(\mathrm{x},\mathrm{y})$ that vanishes outside the unit disk is:

$$\mathbf{Z}_{nm} = \frac{n+1}{\pi}\int\int_{unitdisk} f(\mathrm{x},\mathrm{y})\,\mathrm{V}^{*}_{nm}(\rho,\theta)dxdy \tag{7}$$

Where n is a non-negative integer and m is an integer subject to $n-|m|$ is non-negative and even. The complex-valued functions $V_{nm}(\rho,\theta)$ is defined as:

$$V_{nm}(\rho,\theta) = \mathrm{R}_{nm}(\rho)\exp(\mathrm{j}m\theta) \tag{8}$$

Where $V_{nm}(\rho,\theta)$ is a complex conjugate of $\mathrm{V}^{*}_{nm}(\rho,\theta)$. ρ and θ indicate polar coordinates over the unit disk and R_{nm} is Zernike polynomials of ρ given by:

$$\mathrm{R}_{nm}(\rho) = \sum_{s=0}^{(n-|m|)/2} \frac{(-1)^{s}[(\mathrm{n}-s)!]\rho^{n-2s}}{s!(\frac{n+|m|}{2}-\mathrm{s})!(\frac{n-|m|}{2}-\mathrm{s})!} \tag{9}$$

$|\mathbf{Z}_{nm}|$ denotes the magnitude of the Zernike moments, and it can be used as feature of image (or block). A N_m dimensional vector is obtained from block.

$$N_m = \sum_{i=0}^{n}(\left\lfloor \frac{i}{2} \right\rfloor + 1) \tag{10}$$

Where n is the degree of Zernike moments, and n is set to 5 according to the suggest in Ryu et al. (2010). Therefore a $N_m = 12$ dimension vector is obtained from each block.

Step 4. Block matching.

The feature vectors of blocks compose matrix $\mathbf{M}$, then we sort $\mathbf{M}$ by lexicographic sorting, the sorted matrix is denoted as $\mathbf{S}$. In $\mathbf{S}$ the similar blocks are in the adjacent rows, when the Euclidean distance between two adjacent feature vectors is smaller than the predefined threshold D they are considered to be a pair of candidates for the forgery, i.e.,

$$\left\|\mathbf{S}^{p} - \mathbf{S}^{p+1}\right\|_{2} \le D \tag{11}$$

Where $\mathbf{S}^{p}$, $\mathbf{S}^{p+1}$ indicate the p and $p+1$ row of $\mathbf{S}$, respectively.

Because matching is done in smooth regions where the similarity of adjacent blocks is high, threshold D should be set smaller to reduce false alarms (Khotanzad and Hong 1990; Y. Li 2013).

Step 5. Filtering

Due to the fact that the similarity of adjacent blocks is high, we need to remove candidate matching blocks when

$$\sqrt{(i-k)^2+(j-l)^2} \le D1 \tag{12}$$

Where (i,j) and (k,l) indicate the coordinates of candidate matching blocks.

The shortest distance $D1$ between two comparable blocks is related to image size, we set the empiric value in our experiment as follows.

$$D1 = \begin{cases} 60 & U \ge 3000 \times 2000 \\ 30 & \text{otherwise} \end{cases} \tag{13}$$

Where U indicates the image size.

2.5 Postprocessing

Due to the fact that keypoints are by nature very sparse, it needs further post-processing to locate forgery regions. We apply the SLIC algorithm once again to segment test image into smaller regions. Here the initial size of the small superpixels is set to $S\_init/6$ by experiments, where S_init is the initial size we used in previous segmentation. The new regions that exist matching points are marked as forged.

Then we combine the regions which are detected in the keypoints regions and smooth regions. Finally, morphological operations are used to fill the small cracks and remove the connected region whose area is too small.

3 Experimental results and discussion

3.1 Databases

In this section, we evaluate the reliability and efficiency of the proposed approach using two image databases.

Benchmark database for CMFD evaluation: This database was constructed by Achanta et al. (2012), which consists of 48 high-resolution base images without compression and 87 copied snippets

that were pasted in the same image. These snippets are carefully selected such that forgery trace is inconspicuous, and the average size of an image is about 3000×2300 pixels.

Our small database: we select 4 images from MICC-F220 (Ng et al. 2004) and 4 images from CoMoFoD_small_v2 database (Tralic et al. 2013). All of these images are tampered, and their sizes vary from 512×512 to 800×600. Most images in this database have been tampered by copying the smooth regions. Each tamper image corresponds to a mask image (ground truth), and they are shown in the first and second columns of Fig. 7.

3.2 Assessment criteria

In practical applications, there are two main requirements in image forensics: one is to distinguish forgery and original image (image level), and the other is to correctly locate tampered region (pixel level). Therefore we evaluate the proposed scheme at two levels: image level and pixel level. It's worth noting that when evaluating at image level, if forgery has been found in keypoints regions, the smooth regions will not be detected, otherwise we further detect forgery in smooth regions.

Metrics: In this paper, we adopt precision, recall and F1 measurements as metrics which are often used in the field of forgery detection and information retrieval. They are defined as:

$$precision = \frac{T_P}{T_P + F_P} \tag{14}$$

$$recall = \frac{T_P}{T_P + F_N} \tag{15}$$

$$F1 = 2\times\frac{precision\times recall}{precision + recall} \tag{16}$$

T_P: The number of correctly detected forged images (pixels).

F_P: The number of original images (pixels) that have been erroneously detected as the forged.

F_N: The number of forged images (pixels) that have been falsely missed.

As seen, the precision is a measure for the probability that a detected forgery is truly a forgery, and the recall denotes the probability that a forgery is detected. $F1$ is a trade-off between $precision$ and $recall$. These measures are used at image level and pixel level, in general a higher $F1$ indicates superior performance.

3.3 Threshold determination

The ratio of keypoints' number to pixels' number in the same superpixel T : The selection of threshold T has a great influence on the reliability and efficiency of the algorithm. If T is set too large, although the detection reliability is guaranteed, the calculation costs much more time. If T is small, the algorithm will degenerate to only use keypoint-based method in test image, and it is likely to miss the forgery in smooth regions. Therefore, it is crucial to choose an appropriate T . To determine the appropriate value of T , we test on Fig. 7 (b) and analyze the precision, recall, F1 and running time (at pixel level) by changing values of T. Note that the experiments are tested with the machine of 3.20GHz processor, 32GB RAM, and simulated by Matlab. Fig. 5 illustrates that different values of T corresponds to different smooth regions (blue marked). As shown in Table 2, with the increase of T , the precision, recall and F1 tend to be stable, but time for feature extraction is increased. Therefore we set T =0.0025 in the following experiments.

Similarity threshold between matching pairs D **:** Larger D may introduce more false alarms, yet smaller D may result in missed detection. As matching is done in smooth regions where the similarity between blocks is high, threshold D should be set smaller. To decide the value of D , we test on the second database (our small database) and draw the curves of precision, recall and F1 (on average) with different value of D . As seen in Fig. 6, the precision has the downward trend when D increases to 40. Because some original blocks are classified as forged when D is larger. We also notice that a small D cannot detect forgery in smooth regions. Therefore, we set D to 30 in the following experiments.

Block size b : reduced block size will result in increased computation expense, however, it is essential to avoid missed detection of small copy-move forged areas. The majority of block-based methods propose a block size of 16×16 , and that is why we also used $b=16$ in our experiments.

3.4 Evaluation at image level

In this test we examine the ability of the proposed scheme on the first database (Benchmark database for CMFD evaluation). We test on 96 images, 48 images of which are original and others of which are the plain copy-move forgery. Since the original sizes of these images are rather large, we have to resize all the images to reduce the computation cost, even though resizing may make the detection more difficult. We compare our scheme with two kinds of keypoint-based methods: SIFT and SURF. Due to high com-

putational complexity of Zernike moments, it is impractical to use Zernike moments method in this database. So another block-based method, HU moments method is added into the comparison. Additionally, a state-of-the-art method (J. Li et al. 2015) is added for benchmarking, in which the test images were segmented with a fixed initial size. Keypoints were matched between patches and Expectation Maximization Algorithm (EM) (Bilmes 1998) was designed to filter false alarm patches.

As seen Table 3, our proposed approach produces the highest F1 among all methods, which indicates its robustness in dealing with large scale changes. In terms of recall rate, ours is 0.9375 at the image level, which performs similarly to the block-based method (HU moments) owing to applying block-based method in smooth regions. The proposed method outperforms J. Li et al. (2015) in terms of both precision and recall rates. The average computation times (in seconds) per image are also compared and shown in Table 3. In comparison with block based methods, our approach has significantly reduced the computational cost, as block-based processing is only applied to selected smooth regions for improved efficiency and efficacy, though it is still much higher than keypoints based methods. For the precision rate, ours is slightly lower than those from SIFT and SURF. The reason behind is that the similarity between some blocks is less than the threshold D in the original images. As a result, they are falsely classified as forgery.

3.5 Evaluation at pixel level

As the process of large scale resizing makes the detection difficult, in the following test we evaluate the scheme at pixel level using the second database without resizing. The second database is smaller, so we can compare our scheme with Zernike moments method in this database. The third, fourth and fifth columns of Fig. 7 show the detection results of SIFT scheme, our proposed scheme and Zernike moments method, respectively.

From the results shown in Fig. 7 we can see that (B1), (C1), (E1) cannot be detected using SIFT, though it can detect a portion of the forgery regions in (A1), (F1) and (G1). The reason for poor performance of SIFT method in these images is that smooth regions are copied and few keypoints can be extracted from them. Our approach, however, can detect forgery even when smooth regions are copied. At the same time, it performs similarly to Zernike moments method and improves the detection speed greatly, owing to selectively applying block-based method to the smooth regions rather than all regions of the image. Note that our approach and the Zernike moments method bring false alarms when detecting (E1)

and (G1), due to the self-similarity of image regions. A region pasted in multiple regions (H1) can also be detected by our method due to the g2NN strategy is applied in the keypoints matching stage.

The average precision, recall, F1 and running time on the second database are shown in Table 4, it is obvious that our scheme performs much better than SIFT method. As most of the images in the second database have been tampered by copying the smooth regions, resulting in a poor performance of the SIFT method. In terms of reliability, the F1 measurement of our fusion method is 0.8717, which is as good as the Zernike moments method, yet the average running time has been reduced from 5018 to 179 seconds. When smooth regions are copied block-based method has excellent performance, but the extremely high running time makes it impractical. Thanks to the fusion based approach, our proposed methodology has successfully produced satisfactory results from different test images.

To measure the robustness of our fusion method we also test it against two kinds of attacks, i.e., noise addition and JPEG compression, as shown in Table 5. In total 72 (8×9) images are used for robustness test with the results shown in Fig. 8. The performance of Zernike moments drop down obviously with increased intensity of attacks. The proposed method and SIFT method are more robust to these attacks. Owing to fusion of the two best supplementary features, our proposed approach outperforms the other two methods in terms of F1 measurement even when the added noise or JPEG attack is large. Although our experiments only involve noise and JPEG attacks, it can also work well against geometric attacks, theoretically. Because the keypoints are extracted by SIFT which is robust to scale and rotation, block features are extracted by Zernike which is invariant against rotation.

4 Conclusion

This paper presents a novel fusion based approach for image forgery detection by adaptive combination of keypoint-based method and block-based method. For each image, our scheme can adaptively determine an appropriate initial size of regions, and divide the image into smooth region and keypoints region. By applying different methods to these two types of regions, our approach can effectively detect forgery from both smooth regions and non-smooth ones whilst reducing the computation cost. When detecting forgery in smooth regions, the selection of threshold D has a great influence on the results. In our future work, we will investigate how to optimally determine the value of D based on the image content and further improve the detection speed and accuracy. In addition to image-splicing detection-by

trace, integration of additional most state-of-the-art feature extraction approaches such as sub-pixel image matching (Jiang et al. 2011; Ren et al. 2010), sparse representation (Zhao et al. 2013), saliency detection and deep learning (Han et al. 2015a; Han et al. 2015b) as well as singular spectrum analysis et al (Ren et al. 2014; Zabalza et al. 2014) will also be focused for future study.

5 Acknowledgments

This study was supported by the Science and Technology Innovation Project of Shaanxi Province (Nos. 2015KTTSGY04-05 and 2015KTZDGY04-01), Pre-research Project: target detection project. The authors also wish to greatly thank the editors and anonymous reviewers for their constructive comments to further improve the clarity and quality of this paper. We also thank J. Li et al. (2015) for providing their source codes to enable us for more accurate performance assessment in term of not only precision, recall and F1 but also the running time.

6 REFERENCES

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., & Susstrunk, S. (2012). SLIC superpixels compared to state-of-the-art superpixel methods. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 34*(11), 2274-2282.

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 Processing: Image Communication, 28*(6), 659-669.

Amerini, I., Ballan, L., Caldelli, R., Del Bimbo, A., & Serra, G. (2011). A sift-based forensic method for copy–move attack detection and transformation recovery. *Information Forensics and Security, IEEE Transactions on, 6*(3), 1099-1110.

Bashar, M., Noda, K., Ohnishi, N., & Mori, K. (2010). Exploring Duplicated Regions in Natural Images. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society, iccia*(99), 1-1.

Bayram, S., Sencar, H. T., & Memon, N. An efficient and robust method for detecting copy-move forgery. In *IEEE International Conference on Acoustics, Speech & Signal Processing, 2009* (pp. 1053-1056)

Bilmes, J. A. (1998). A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models. *International Computer Science Institute, 4*(510), 126.

Bo, X., Junwen, W., Guangjie, L., & Yuewei, D. Image copy-move forgery detection based on SURF. In *Multimedia Information Networking and Security (MINES), 2010 International Conference on, 2010* (pp. 889-892): IEEE

Cheng, G., Han, J., Guo, L., Liu, Z., Bu, S., & Ren, J. (2015). Effective and Efficient Midlevel Visual Elements-Oriented Land-Use Classification Using VHR Remote Sensing Images. *IEEE Transactions on Geoscience & Remote Sensing, 53*(8), 4238-4249.

Cheng, G., Han, J., Zhou, P., & Guo, L. (2014). Multi-class geospatial object detection and geographic image classification based on collection of part detectors. *Isprs Journal of Photogrammetry & Remote Sensing, 98*(1), 119–132.

Christlein, V., Riess, C., Jordan, J., Riess, C., & Angelopoulou, E. (2012). An Evaluation of Popular Copy-Move Forgery Detection Approaches. *IEEE Transactions on Information Forensics & Security, 7*(6), 1841-1854.

Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM, 24*(6), 381-395.

Fridrich, J., Soukal, D., & Lukas, J. (2003). Detection of copy-move forgery in digital images. *Proceedings of Digital Forensic Research Workshop*.

Han, J., Zhang, D., Cheng, G., Guo, L., & Ren, J. (2015a). Object Detection in Optical Remote Sensing Images Based on Weakly Supervised Learning and High-Level Feature Learning. *IEEE Transactions on Geoscience & Remote Sensing, 53*(6), 3325-3337.

Han, J., Zhang, D., Hu, X., Guo, L., Ren, J., & Wu, F. (2015b). Background Prior-Based Salient Object Detection via Deep Reconstruction Residual. *IEEE Transactions on Circuits & Systems for Video Technology, 25*(8), 1309-1321.

Huang, H., Guo, W., & Zhang, Y. Detection of copy-move forgery in digital images using SIFT algorithm. In *Computational Intelligence and Industrial Application, 2008. PACIIA'08. Pacific-Asia Workshop on, 2008* (Vol. 2, pp. 272-276): IEEE

Huynh, T. K., Huynh, K. V., Le-Tien, T., & Nguyen, S. C. A survey on Image Forgery Detection techniques. In *Computing & Communication Technologies - Research, Innovation, and Vision for the Future (RIVF), 2015 IEEE RIVF International Conference on, 2015*

Jiang, J., Kohler, J., Macwilliams, C., Zaletelj, J., Guntner, G., Horstmann, H., et al. (2011). LIVE: An Integrated Production and Feedback System for Intelligent and Interactive TV Broadcasting. *IEEE Transactions on Broadcasting, 57*(3), 646-661.

Ren, J., Jiang, J., & Vlachos, T. (2010). High-accuracy sub-pixel motion estimation from noisy images in Fourier domain. *IEEE Transactions on Image Processing, 19*(5), 1379-1384.

Jing, L., & Shao, C. (2012). Image copy-move forgery detecting based on local invariant feature. *Journal of Multimedia, 7*(1), 90-97.

Kang, X. B., & Wei, S. M. Identifying Tampered Regions Using Singular Value Decomposition in Digital Image Forensics. In *2008 International Conference on Computer Science and Software Engineering, 2008* (pp. 926-930)

Khotanzad, A., & Hong, Y. H. (1990). Invariant image recognition by Zernike moments. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 12*(5), 489-497.

Kumar, S., Desai, J., & Mukherjee, S. (2015). A Fast Keypoint Based Hybrid Method for Copy Move Forgery Detection. *Int. J. Com. Dig. Sys, 4*(2).

Li, J., Li, X., Yang, B., & Sun, X. (2015). Segmentation-based image copy-move forgery detection scheme. *Information Forensics and Security, IEEE Transactions on, 10*(3), 507-518.

Li, L., Li, S., Zhu, H., Chu, S. C., Roddick, J. F., & Pan, J. S. (2013). An efficient scheme for detecting copy-move forged images by local binary patterns. *Journal of Information Hiding & Multimedia Signal Processing, 4*, 46-56.

Li, Y. (2013). Image copy-move forgery detection based on polar cosine transform and approximate nearest neighbor searching. *Forensic science international, 224*(1), 59-67.

Lin, H.-J., Wang, C.-W., & Kao, Y.-T. (2009). Fast copy-move forgery detection. *WSEAS Transactions on Signal Processing, 5*(5), 188-197.

Lowe, D. G. Object recognition from local scale-invariant features. In *Computer vision, 1999. The proceedings of the seventh IEEE international conference on, 1999* (Vol. 2, pp. 1150-1157): Ieee

Lowe, D. G. Distinctive Image Features from Scale-Invariant Keypoints. In *International Journal of Computer Vision, 2004* (pp. 91-110)

Luo, W., Huang, J., & Qiu, G. Robust detection of region-duplication forgery in digital image. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on, 2006* (Vol. 4, pp. 746-749): IEEE

Mahdian, B., & Saic, S. (2007). Detection of copy–move forgery using a method based on blur moment invariants. *Forensic science international, 171*(2), 180-189.

Muhammad, N., Hussain, M., Muhammad, G., & Bebis, G. Copy-Move Forgery Detection Using Dyadic Wavelet Transform. In *Proceedings of the 2011 Eighth International Conference Computer Graphics, Imaging and Visualization, 2011* (pp. 103-108)

Ng, T., Chang, S., Hsu, J., & Pepeljugoski, M. (2004). Columbia photographic images and photorealistic computer graphics dataset, ADVENT. *Columbia University*.

Popescu, A. C., & Farid, H. (2004). Exposing Digital Forgeries by Detecting Duplicated Image Regions. *Comput.sci.dartmouth College Private Ivy League Res.univ, 646*.

Pun, C. M., Yuan, X. C., & Bi, X. L. (2015). Image Forgery Detection Using Adaptive Over-Segmentation and Feature Points Matching. *IEEE Transactions on Information Forensics & Security, 10*, 1-1.

Ren, J., Zabalza, J., Marshall, S., & Zheng, J. (2014). Effective Feature Extraction and Data Reduction in Remote Sensing Using Hyperspectral Imaging [Applications Corner]. *IEEE Signal Processing Magazine, 31*(31), 149-154.

Ryu, S.-J., Kirchner, M., Lee, M.-J., & Lee, H.-K. (2013). Rotation invariant localization of duplicated image regions based on Zernike moments. *Information Forensics and Security, IEEE Transactions on, 8*(8), 1355-1370.

Ryu, S.-J., Lee, M.-J., & Lee, H.-K. Detection of copy-rotate-move forgery using Zernike moments. In *Information Hiding, 2010* (pp. 51-65): Springer

Sencar, H. T., & Memon, N. (2008). Overview of state-of-the-art in digital image forensics. *Algorithms*.

Shensa, M. (1992). Discrete Wavelet Transform: Wedding the A Trous and Mallat Algorithms. *IEEE Transactions on Signal Processing, 40*(10), 2464-2482.

Teh, C.-H., & Chin, R. T. (1988). On image analysis by the methods of moments. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 10*(4), 496-513.

Tralic, D., Zupancic, I., Grgic, S., & Grgic, M. CoMoFoD—New database for copy-move forgery detection. In *ELMAR, 2013 55th International Symposium, 2013* (pp. 49-54): IEEE

Vedaldi, A., & Fulkerson, B. VLFeat: An open and portable library of computer vision algorithms. In *Proceedings of the 18th ACM international conference on Multimedia, 2010* (pp. 1469-1472): ACM

Wang, J., Liu, G., Zhang, Z., Dai, Y., & Wang, Z. (2009). Fast and robust forensics for image region-duplication forgery. *Acta Automatica Sinica, 35*(12), 1488-1495.

Zabalza, J., Ren, J., Wang, Z., Marshall, S., & Wang, J. (2014). Singular Spectrum Analysis for Effective Feature Extraction in Hyperspectral Imaging. *IEEE Geoscience & Remote Sensing Letters, 11*(11), 1886-1890.

Zahra, M. (2012). Image duplication forgery detection using two robust features. *Res. J. Recent Sci, 1*(12), 1-6.

Zhao, C., Li, X., Ren, J., & Marshall, S. (2013). Improved sparse representation using adaptive spatial support for effective target detection in hyperspectral imagery. *International Journal of Remote Sensing, 34*(24), 8669-8684.

List of Figure Captions:

Fig. 1. Results of forgery detection using our approach.

Fig. 2. Adaptive segmentation results

Fig. 3. Forgery detection in keypoints regions

Fig. 4. Illustrate how to use block-based method in smooth region

Fig. 5. The blue marked regions are smooth region

Fig. 6. Average precision, recall and F1 curves with different values of D.

Fig. 7. Detection results of the SIFT scheme and proposed scheme. From left to right, the five columns show the test images, the ground truth of the forged regions, and detected results using SIFT scheme, our proposed scheme and Zernike moments method, respectively.

Fig. 8. Detection results of different methods against 2 kinds of attacks.

List of Table Captions:

Table 1. The detailed steps of g2NN strategy

Table 2. Precision, recall, F1 and running time with different values of T

Table 3. Detection results at image level on the first database

Table 4. Average precision, recall and F1 on the second database

Table 5. Setting of attacks

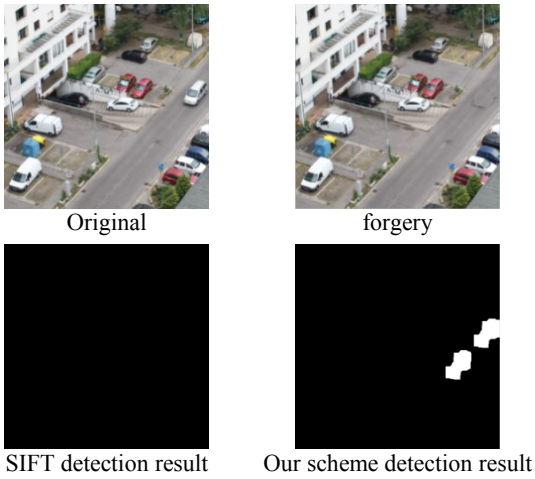


Fig.1. Results of forgery detection using our approach

Fig.2. Adaptive segmentation results

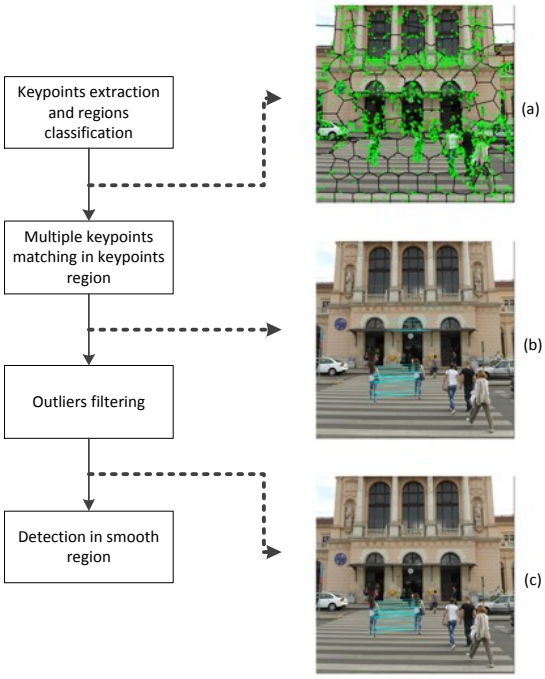


Fig.3. Forgery detection in keypoints regions

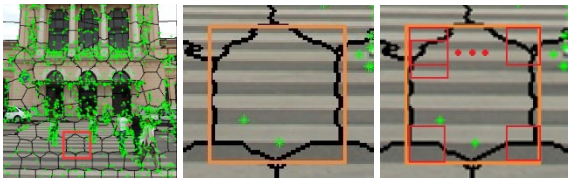
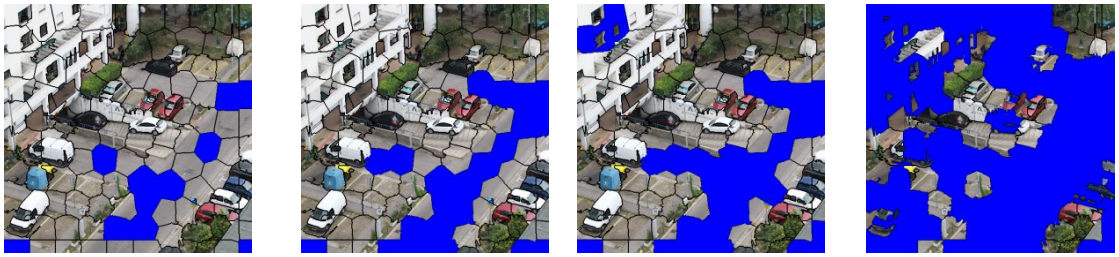


Fig.4. Illustrate how to use block-based method in smooth region



T =0.001 T=0.0025 T =0.004 T =0.01

Fig.5. The blue marked regions are smooth region

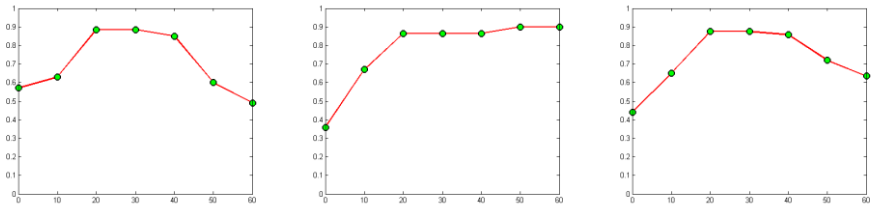


Fig.6. Average precision, recall and F1 curves with different values of D .

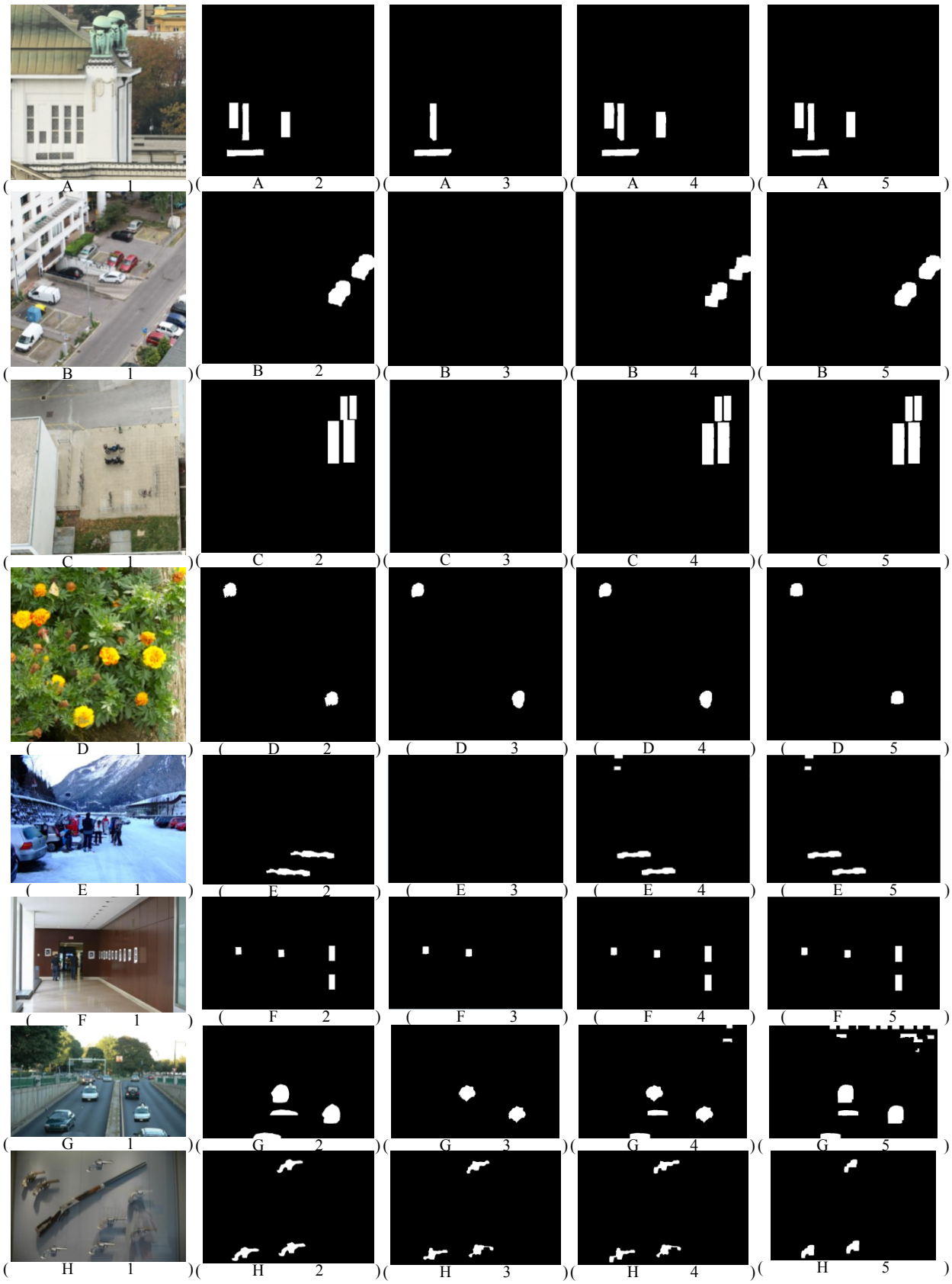


Fig. 7. Forgery detection results comparison. From left to right, the five columns show the test images, the ground truth of the forged regions, and detected results using SIFT, our proposed scheme and Zernike moments, respectively.

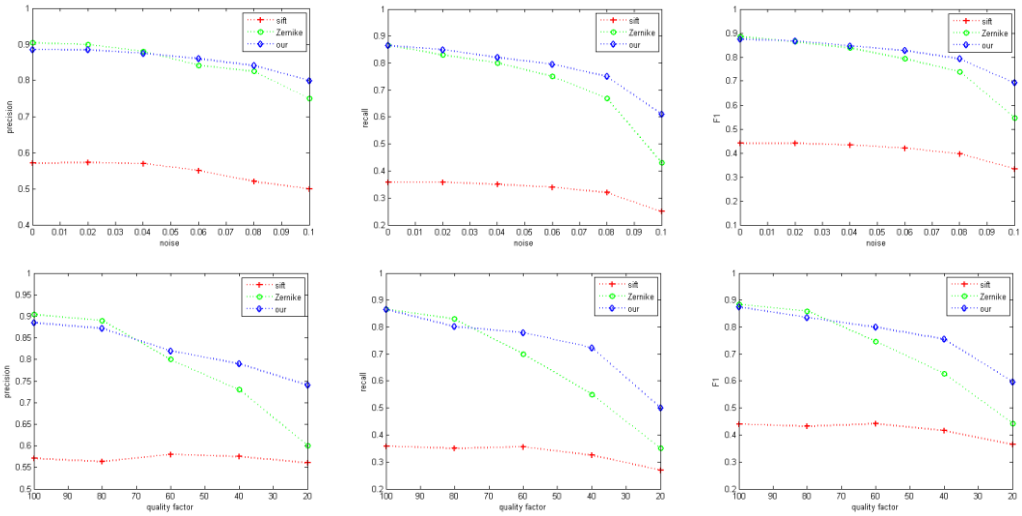


Fig. 8. Detection results at pixel level of different methods against 2 kinds of attacks.

Table 1. The detailed steps of g2NN strategy

Input: $\mathbf{S} = \{s_1, s_2, ...s_{n-1}\}$
Output: i
Initial: $i = 1$
While $\frac{s_i}{s_{i+1}} <= T_r$
$i = i + 1;$
end

Table 2. Precision, recall, F1 and running time with different values of T

	precision	**recall**	**F1**	**Running time (s)**
T =0.001	0.9518	0.3512	0.5131	77
T =0.0025	0.9545	0.8545	0.9017	144
T =0.004	0.9545	0.8545	0.9017	227
T =0.01	0.9545	0.8545	0.9017	1040

Table 3. Detection results at image level on the first database

	Precision	**Recall**	**F1**	**Running time (s)**
SIFT	0.7872	0.7708	0.7789	13.7
SURF	0.8387	0.5417	0.6582	7.62
HU moments	0.6571	0.9583	0.7796	1300
J. Li et al. (2015)	0.7017	0.8333	0.7618	289.26
Ours	0.7759	0.9375	0.8491	862.5

Table 4. Average precision, recall and F1 at the pixel level on the second database

	Precision	Recall	F1	Running time
SIFT	0.5706	0.3585	0.4221	8.4
Zernike moments	0.9039	0.8657	0.8805	5018
Ours	0.8851	0.8648	0.8717	179

Table 5. Setting of attacks

Attacks	Parameters
Adding Noise	Standard deviation (0.02:0.02:0.1)
JPEG Compression	Quality factor (100: -20:20)