Copy-Move Forgery Detection Using the Segment Gradient Orientation Histogram

Ali Retha Hasoon Khayeat[1,2(✉)], Paul L. Rosin[1], and Xianfang Sun[1]

[1] School of Computer Science and Informatices,
Cardiff University, Cardiff, UK
{KhayeatAR,RosinPL,SunX2}@Cardiff.ac.uk
[2] Computer Science Department, College of Science,
Kerbala University, Karbala, Iraq
aliretha@gmail.com

Abstract. The ready availability of image-editing software makes ensuring the authenticity of images an important issue. The most common type of image tampering is cloning, or Copy-Move Forgery (CMF), in which part(s) of the image are copied and pasted back into the same image. One possible transformation is where an object is copied, rotated and pasted; this type of forgery is called Copy-Rotate-Move Forgery (CRMF). Applying post-processing can be used to produce more realistic doctored images and thus can increase the difficulty of forgery detection. This paper presents a novel segmentation-based Copy-Move forgery detection method. A new method has been developed to segment the Copy-Move objects in a consistent way that is more efficient than Simple Linear Iterative Clustering (SLIC) segmentation for CMF/CRMF. We propose a new method to describe irregular shaped blocks (segments). The Segment Gradient Orientation Histogram (SGOH), is used to describe the gradient distribution of each segment. The quality of initial matches is improved by applying hysteresis to grow the primary detection regions. We show that the proposed method can effectively detect forgery involving translation and rotation. Moreover, the proposed method can detect forgery in images with blurring, brightness change, colour reduction, JPEG compression, variations in contrast and added noise.

Keywords: Copy-Move Forgery · Segment Gradient Orientation Histogram (SGOH) · Simple Linear Iterative Clustering (SLIC) · Otsu thresholding

1 Introduction

Copy-move is the most common type of image forgery (copy and paste), where regions of the image are cloned to cover objects in the scene. If this is done with care, visual detection of cloning will be difficult. Moreover, because the cloned regions can be in any location or can have any shape, searching all possible image portions in different sizes and locations is computationally infeasible.

In Copy-Move Forgery (CMF), part(s) of the image are copied and pasted into the same image but in different places, possibly after a rotation. Moreover, because the copied-pasted region is from the same image, its characteristics (e.g. colour and noise) are compatible with that image. This type of forgery is more challenging to detect than other types, such as splicing and retouching. This is because the usual methods of detecting incompatibilities, using statistical measurements to compare different parts of the image, will be useless for CMF detection [6].

The most common approach to detect CMF consists of many steps, the most important step is the feature extraction. There are two different methods to extract features, either by tiling the image into blocks and then extracting the features from each block, or from interest points which are distributed over the image in different ways (e.g. SIFT, MSER, FAST, etc.). The block-based methods usually need a long time to extract features from the image. The keypoint-based methods are much faster than block-based methods, but can only detect part(s) of the duplicated region(s) because keypoints are distributed sparsely over the image.

We have considered a segmentation approach as a potential solution to overcome the problems of the block-based and keypoint-based methods. However, the main problem of the standard approach is: *there is no reliable method to segment identical objects consistently.* So, even with state-of-the art segmentation methods, there is no guarantee of segmenting Copy-Move objects in the same manner.

This paper presents a novel method to detect Copy-Move Forgery which performs a consistent image segmentation and extracts features from each segment. The image has been quantized into seven labels using the Rolling Guidance filter [18] followed by Otsu's thresholding [11]. Each quantized area (segment) has been described using a 3D colour histogram, the Segment Weighted Gradient Orientation Histogram (SWGOH), and its size. The 2nd Approximate Nearest Neighbour was used to detect the forgery at the segment level followed by a hysteresis technique to extend the primary detection. The proposed method is robust to rotation and the effects of post-processing method, and is relatively fast.

2   Related Works

Many papers have been proposed to detect CMF, and are mostly either block based or use keypoint based techniques. Recently two papers have used segmentation to detect CMF [8,13].

The authors of [13] tested four different image segmentation methods and used superpixels – the Simple Linear Iterative Clustering (SLIC) algorithm [1] – to over-segment the images. Then they extracted SIFT features from each segment, built a K-d tree for these features and used KNN to find the matches between patches. They computed the number of the matched feature points and identified suspicious pairs of patches which have many similar keypoints.

RANSAC was applied to estimate the transformation matrix between each pair of patches. Each pixel was represented by a dense SIFT descriptor. Then, the patch level matches were refined by applying matching to all the pixels in the matched patches and applying RANSAC to remove outliers.

The authors found that the segmentation method did not significantly influence the Copy-Move Forgery detection. Their approach is similar to finding matches between different images using SIFT features, considering each segment as a different image. Segmentation was used to divide the original image into various patches (small images). Their approach depends on extracting SIFT features from segments instead of from the whole image.

Li et al. [8] also used the SLIC method to segment the images. They used different scales of segmentation depending on the image contents itself. They set a large initial superpixel size for smooth images, and a small initial size for detailed images. The Discrete Wavelet Transformation (DWT) was used to analyse the frequency distribution of the image. According to their approach, the image is smooth when the majority of the energy of the host image is low-frequency; otherwise, the image is considered to be detailed. Then they extracted SIFT features from each segment and computed the Euclidean distance between features. If the number of matched points is more than a threshold, a correlation coefficient map generated to find the matched patches. They used SLIC to segment each matched patch to a smaller size and measured the local colour feature for each sub-patch. Neighbouring sub-regions (patches) are merged when the colour features are similar and the morphological close operation applied to generate detected forgery regions.

The above two methods [8,13] rely on using the keypoint-based approach (SIFT) in addition to image segmentation in order to detect CMF. In homogeneous regions there will be few SIFT features detected, causing matches to be missed. Moreover, the computational complexity of the two proposed methods is high.

3   Background

As the keypoint based techniques are likely to fail to detect interest points in relatively homogeneous regions, we instead use a dense descriptor to represent each segment. However, this will require repeatable segmentation, i.e. instances of the copied objects should be segmented in the same way, since otherwise, their descriptors will not match. As demonstrated in Sect. 5.2, SLIC is not repeatable: variations in an object's surroundings prevents the consistent segmentation of duplicated objects. We will present a new segmentation method that can segment the image consistently.

3.1   Threshold Selection Using Otsu's Method

The most common and simplest method to segment an image is using thresholding. Otsu [11] suggested a method to find the best threshold to binarize the grayscale image.

The binary image which is generated from applying a single Otsu threshold on the grayscale image contains a lot of small segments, and these would be unstable for matching in CMF detection, see the binary image in Fig. 1. Also, using the multi-level thresholding version of Otsu does not produce a better result, see the coloured images in Fig. 1.
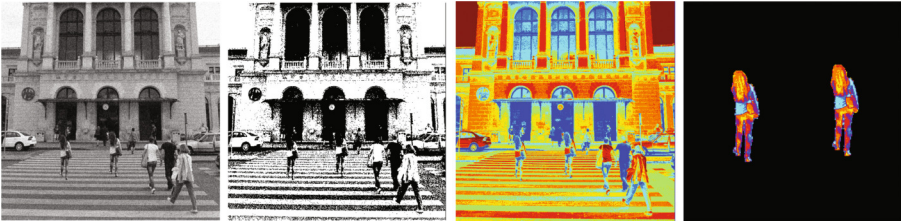

Fig. 1. (from left to right) The grayscale input image, the binary image generated by Otsu's method, the segmented image using 7 Otsu thresholds, zoomed CMF areas.

3.2   Rolling Guidance Filter

In order to produce more meaningful segments, we will filter the image to remove noise and unnecessary details. Images contain significant structures and edges over a range of scale [18]. Many filtering techniques have been proposed to smooth the image while maintaining those structures. Edge-aware filters have been used to remove the low-contrast edges (gradual changes) and preserve the high-contrast edges, e.g. bilateral filter [15], guided filter [4] and weighted median filters [18]. In this work we use the Rolling Guidance Filter because it has been shown to be effective at removing small-scale structures while preserving large-scale structures by the use of scale-aware local operations. It can cope with irregular shapes and furthermore has low computation cost.
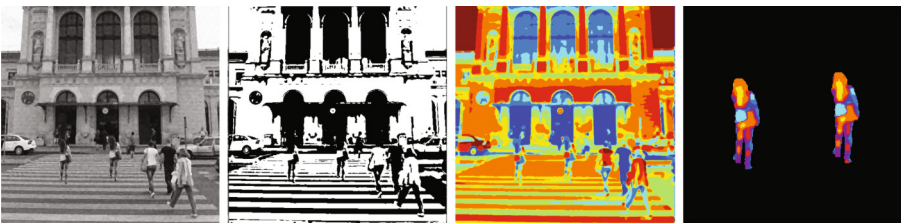

Fig. 2. (from left to right) The rolling guidance grayscale image, the rolling guidance binary image generated by Otsu's method, the rolling guidance segmented image using 7 Otsu thresholds, zoomed CMF areas in the rolling guidance image after threshing.

The rolling guidance method includes two main steps:

1. Remove small structures: In this step, the Gaussian filter is used. However, as well as removing the edges of structures smaller than the smoothing scale, it also blurs large-scale structures instead of preserving them.
2. Edge recovery: A joint bilateral filtering of the given input image and the image from the previous iteration is used to recover the edges. This can be understood as a filter that smooths the input image guided by the structure of the previous iteration image.

The binary image which is generated from applying the single Otsu threshold to the Rolling Guidance filtered image (smoothed image) produces a reasonably segmented image. As shown in the binary image of Fig. 2, the detrimental or unwanted content has been removed and the pixels have been clustered appropriately.

However, this does not adequately segment the Copy-Move objects in the image. Under-segmentation has caused the objects to become merged with the background. Therefore, we use the multi-level thresholding version of Otsu applied after the Rolling Guidance filtering, see the coloured images of Fig. 2.

4   Methodology

4.1   Segment Gradient Orientation Histogram (SGOH)

SIFT/DSIFT is restricted to regular blocks. Here, we have developed the Segment Gradient Orientation Histogram (SGOH) to describe the gradient for each segment (irregular block). The SIFT descriptor [9] has a 128 element feature vector for each keypoint. It considers a 16×16 neighbourhood around each keypoint and divides it into 4×4 sub-regions. For each sub-region, an eight-bin histogram of gradient magnitude weighted orientations is computed. DSIFT [2] follows the same approach however it considers all the pixels in the image as keypoints.

The steps to build the Segment Gradient Orientation Histogram (SGOH) are as follows:

1. The moment method (intensity centroid measure) [14] is used to find the canonical orientation for each segment.
2. Rotate each segment according to its canonical orientation to make the descriptor rotation invariant.
3. Construct a gradient magnitude weighted orientation histogram containing $18 = 360°/20°$ bins.
4. Normalize the generated feature vector (SGOH) between zero and one.

4.2   Proposed Algorithm

The Rolling Guidance filter [18] is used to smooth the image and preserve the strong edges, then the Otsu method [11] is used to find 7 different thresholds

on the filtered image. We have tried a different number of thresholds (5, 7, 9, 11 and 13), and experimentally we found that using 7 thresholds segments the Copy-Move objects in the most consistent way. The 7 thresholds have been used to quantize the Rolling Guidance filtered image into 8 different labels. Next, connected component labelling is applied in each different intensity threshold range and properties (e.g. area, pixel list, etc.) are computed for each object. All segments of size less than 70 pixels ($T_1 = 70$) are removed; this threshold has been chosen experimentally, and is fixed for all experiments. A 3D colour (3DRGB) histogram has been used to describe the colour distribution of each segment, and a Segment Gradient Orientation Histogram (SGOH) has been built to represent the gradient of each segment, see Sect. 4.1. The SGOH is concatenated with 3DRGB and the segment area to form a feature vector. A K-d tree is built from the segment feature vectors, and for each segment its 2ANN is found. Matched segments are kept if the Euclidean distance between their concatenated feature vectors is less than a threshold ($T_2 = 0.002$) and the ratio between their sizes is less than a threshold ($T_3 = 1.5$). Make the size of the two matched segments is equal, save its coordinates in two separated lists and call RANSAC to remove the outliers. Finally, Hysteresis technique is applied to grow the detected Copy-Move regions.

4.3   Hysteresis Technique

To produce the best possible result, we use a hysteresis technique in the CMF/CRMF detection. Hysteresis thresholding is based on using two thresholds, one low and one high, and it considers the spatial information to improve the result. This technique is commonly employed in edge detection [3]. Recently, hysteresis has been used in forgery detection [5]; the "strong" matches detected using the high threshold, and the low threshold were rejected very "weak" matches. The main drawbacks of Jaberi et al. [5] are that they used a window to search for the new matched features which may contain parts outside of the CMF areas. Also, they recompute all the MIFT feature for all pixels in the detection window with each iteration.

To use the hysteresis thresholding in CMF/CRMF detection, we developed the following approach. Find the 2nd Approximate Nearest Neighbour (2ANN) for each feature vector (SGOH) within the strict low threshold (T_2) [5,19] and this will decrease the false matches. The low threshold is used to detect "strong" similar features, which represent the pixels from the original and the duplicated regions (segments). Apply RANSAC to remove the outliers and find the coordinate's transformation of the matched features. For each coordinate in the transformation list, recolour the block that takes this coordinate as a center. In the next step, dilate each region using a disk with a one-pixel radius size. For each of the newly added pixels, compute the improved DSIFT [6]. Build a K-d tree and find the 2ANN for each new feature vector. If the Euclidean distance between the matched features is less than the high threshold ($T_4 = 2T_2$), we store the coordinates of these features. Apply RANSAC to remove any new outliers and keep the new coordinates within the previously found transformation. Add the new pixels
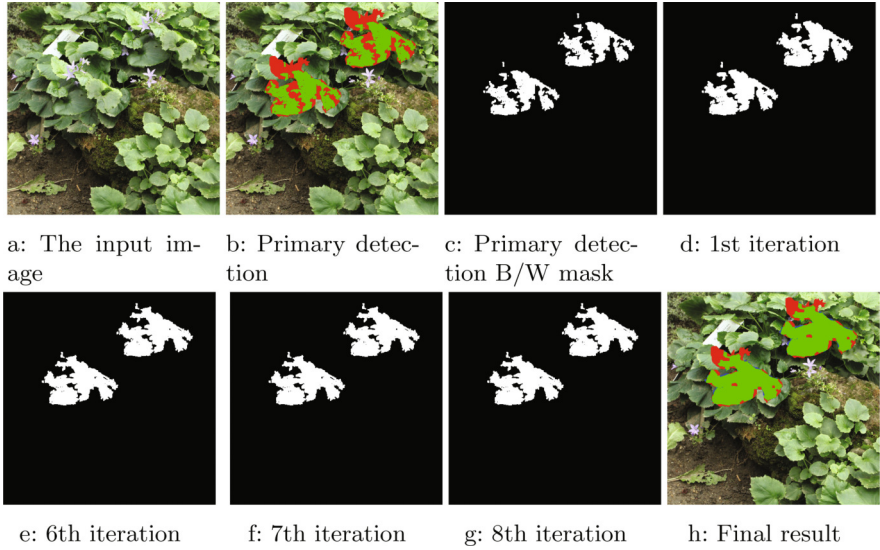
a: The input im- b: Primary detec- c: Primary detec- d: 1st iteration
age   tion   tion B/W mask

e: 6th iteration   f: 7th iteration   g: 8th iteration   h: Final result

Fig. 3. An example of growing the detection regions with hysteresis thresholding.

to the matching list and update the transformation matrix. Grow the detection regions by adding a new block located at the center of the new matched pixel. Repeat this process until no more pixels can be added to the primary detection. This region growing technique depends on the primary detection of the strong features matching and the spatial information, to iteratively add one block to the edges of the primary detection. As illustrated on Fig. 3 the detection regions have been grown, which increased the F-measure from 0.72 to 0.88.

5 Experiments

5.1 Dataset and Evaluation Method

We tested our method using the image database for Copy-Move Forgery Detection (CoMoFoD) [16]. CoMoFoD consists of 260 forged images categorized into two categories (small 512×512, and large 3000×2000). The small category consists of 200 original images with different types of forgery. We considered only the small images in our work. In the small category, images are divided into 5 different groups according to the applied manipulations, as follows: translation, rotation, scaling, distortion and a combination of all previous manipulations. Moreover, various types of post-processing methods (e.g. blurring, brightness change, color reduction, JPEG compression, contrast adjustments and added noise), are applied to all forged images in each group. The total number of images in the small group is 10400 with different types of manipulations. We used the F-measure [20] at the pixel level to evaluate the accuracy of our results.
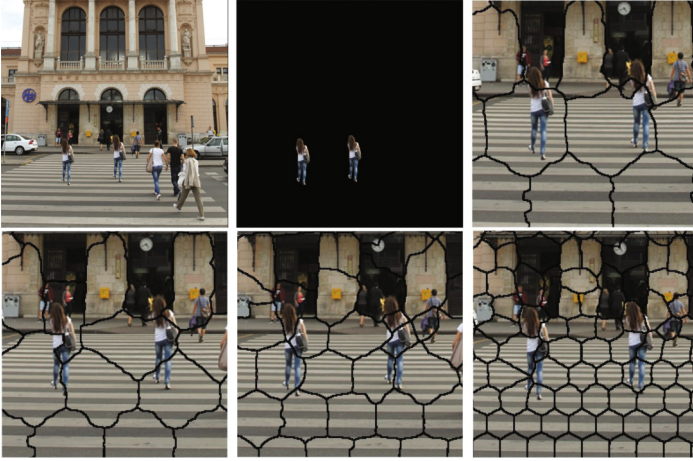
Fig. 4. (top to bottom, left to right) The input forged image, the Copy-Moved objects, zoomed CMF areas segmented using SLIC with $K = 30, 55, 100$ and 300.

5.2 Experiment to Detect CMF Using SLIC

In our initial work we tried to use superpixel segmentation as the basis for detecting CMF; SLIC was applied, and a set of features was densely extracted from each segment.

The size of duplicated objects can vary from one image to another, as they can form a small or large part of an image. SLIC divides the image into irregular blocks which exhibit state-of-the-art boundary adherence [1]. The required number of approximately similar-sized superpixels (K) is the parameter to control the SLIC algorithm.

Problems with this approach are evident in Fig. 4. When the SLIC method is applied to this image it does not segment the Copy-Move objects (i.e. the two ladies) consistently, due to differences between their backgrounds, which reveals the unreliability of this approach.

So, instead of using SLIC to segment the image, we have used our proposed method to segment the Copy-Move objects, see Sect. 4.2, that is more consistent than SLIC and produces better CMF detection results, see Figs. 4 and 5.

5.3 Experiment to Detect CMF with Translation and Post-processing

We used our suggested method to test 40 different images with plain CMF (without post-processing), and obtained an F-measure = 0.79, see Table 1 and Fig. 5. Moreover, the proposed method is less complicated than the other suggested methods [8, 13] which use segmentation; it takes about 45 s to process one image.

We tested our proposed method on 280 images with different types of post-processing (image blurring, brightness change, colour reduction, JPEG compression, contrast adjustments and added noise), see Table 1.
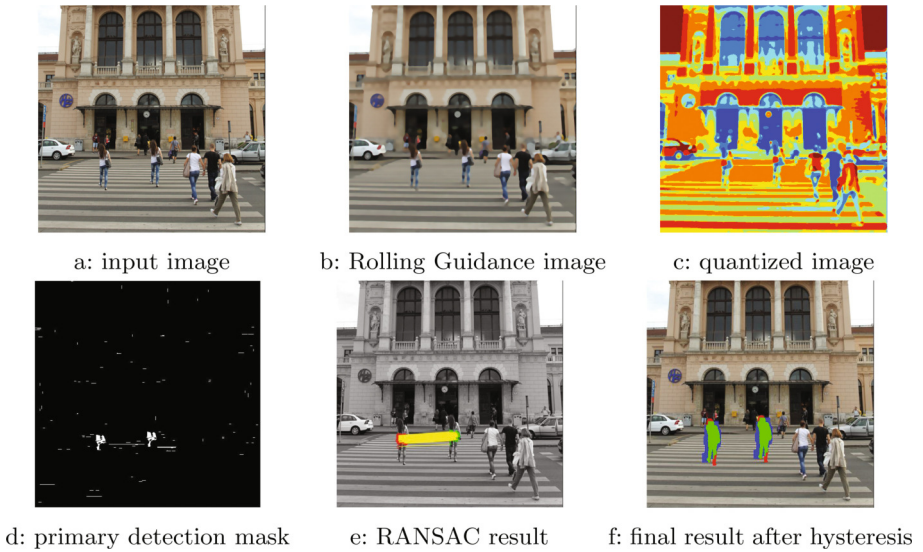

a: input image   b: Rolling Guidance image   c: quantized image

d: primary detection mask   e: RANSAC result   f: final result after hysteresis

Fig. 5. An example of plain CMF detection (True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN)).

5.4 Experiment to Detect CRMF and Post-processing

Using rotation invariant features is the primary requirement of the Copy-Rotate-Move Forgery (CRMF) detection. The Segment Gradient Orientation Histogram (SGOH) is rotation invariant as each segment is rotated to its canonical orientation before computing the weighted histogram.

The experimental work illustrated that the suggested algorithm can detect rotated duplicated objects with acceptable performance. The algorithm detected forgery on 35 images out of 40 with F-measure = 0.71, see Table 1.

280 images with rotated duplicated objects and different types of post-processing (image blurring, brightness change, colour reduction, JPEG compression, contrast adjustments and added noise) have been tested. The suggested algorithm successfully detect forgery on 243 images. We have shown experimentally that our proposed method is not affected by the post-processing methods and it can detect forgery even on compressed or noisy images.

We found that the under-segmentation is the main reason that the proposed algorithm cannot detect forgery on some images (Fig. 6).

We compared our proposed method with Zernike moments (ZM) [6,21] to demonstrate the robustness of our method. Table 1 shows that Zernike moments
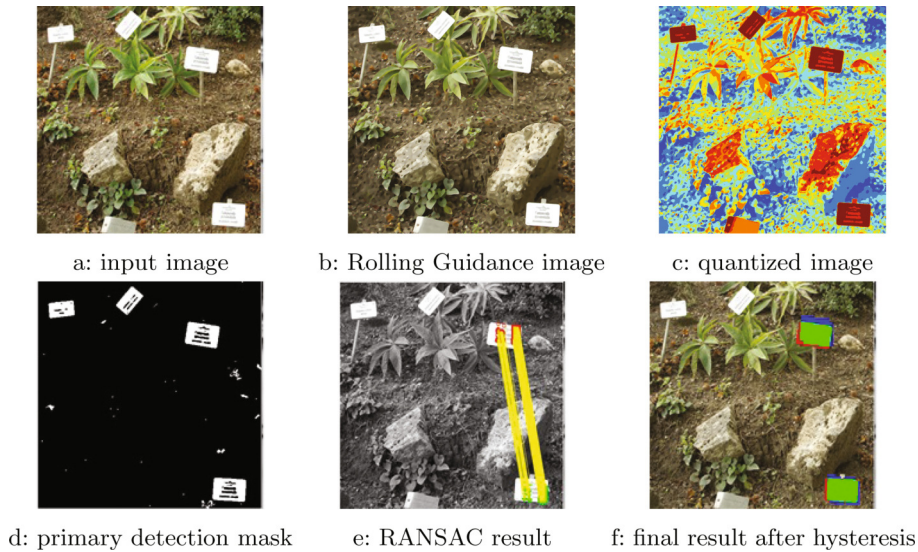
a: input image	b: Rolling Guidance image	c: quantized image
d: primary detection mask	e: RANSAC result	f: final result after hysteresis

Fig. 6. An example of Copy-Rotate-Move Forgery detection (True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN)).

Table 1. CMF/CRMF detection with post-processing.

Post-processing	Translation		Rotation	
	F-measure (SGOH)	F-measure (ZM) [6]	F-measure (SGOH)	F-measure (ZM) [6]
Without post-processing	0.79	0.88	0.71	0.68
Brightness change range (0.01, 0.8)	0.78	0.45	0.70	0.40
Contrast adjustment range (0.01, 0.8)	0.78	0.48	0.70	0.38
Colour reduction (32 intensity levels)	0.78	0.49	0.70	0.40
Image blurring (5×5 average filter)	0.79	0.56	0.70	0.49
Adding noise ($\mu = 0$, $\sigma^2 = 0.005$)	0.69	0.63	0.64	0.61
JPG compression (quality factor $= 40$)	0.78	0.60	0.69	0.62

are not robust to post-processing, which confirms the results of Fig. 9 in [21]. In contrast, our proposed method produces consistent results with post-processing.

6   Conclusions

In this paper, we considered Copy-Move forgery incorporating translation and rotation. A new segmentation method was suggested to segment the Copy-Move objects in a more consistent way than SLIC. We obtained good results on translation and reasonable results with rotation.

The Segment Gradient Orientation Histogram (SGOH), which was inspired by SIFT [9], was used to describe the gradient for each segment (irregular block).

The hysteresis technique was used to grow the detection region(s) and improve the primary detection result. Also, our method can detect CMF in images with blurring, brightness change, color reduction, JPEG compression, variations in contrast and added noise.

References

1. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Süsstrunk, S.: SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Trans. Pattern Anal. Mach. Intell. **34**(11), 2274–2282 (2012)
2. Bosch, A., Zisserman, A., Munoz, X.: Image classification using random forests and ferns. In: 11th International Conference on Computer Vision, pp. 1–8. IEEE (2007)
3. Canny, J.: A computational approach to edge detection. IEEE Trans. Pattern Anal. Mach. Intell. (PAMI) **8**(6), 679–698 (1986)
4. He, K., Sun, J., Tang, X.: Guided image filtering. IEEE Trans. Pattern Anal. Mach. Intell. **35**(6), 1397–1409 (2013)
5. Jaberi, M., Bebis, G., Hussain, M., Muhammad, G.: Accurate and robust localization of duplicated region in copy-move image forgery. Mach. Vis. Appl. **25**(2), 451–475 (2014)
6. Khayeat, A.R.H., Sun, X., Rosin, P.L.: Improved DSIFT Descriptor Based Copy-Rotate-Move Forgery Detection. In: Image and Video Technology. LNCS, vol. 8334, pp. 642–655. Springer (2016)
7. Klette, R.: Concise Computer Vision. Springer, London (2014)
8. Li, J., Li, X., Yang, B., Sun, X.: Segmentation-based image copy-move forgery detection scheme. IEEE Trans. Inf. Forensics Secur. **10**(3), 507–518 (2015)
9. Lowe, D.G.: Distinctive Image Features from Scale-Invariant Keypoints. Int. J. Comput. Vis. **60**(2), 91–110 (2004)
10. Muja, M., Lowe, D.G.: Fast approximate nearest neighbors with automatic algorithm configuration. In: International Conference on Computer Vision Theory and Applications, pp. 331–340. INSTICC Press (2009)
11. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. **9**(1), 62–66 (1979)
12. Pietikäinen, M., Hadid, A., Zhao, G., Ahonen, T.: Computer Vision Using Local Binary Patterns. Springer, London (2011)

13. Pun, C.-M., Yuan, X.-C., Bi, X.-L.: Image forgery detection using adaptive over-segmentation and feature point matching. IEEE Trans. Inf. Forensics Secur. **10**(8), 1705–1716 (2015)
14. Rosin, P.L.: Measuring corner properties. Comput. Vis. Image Underst. **73**(2), 291–307 (1999)
15. Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. In: Sixth International Conference on Computer Vision, pp. 839–846 (1998)
16. Tralic, D., Zupancic, I., Grgic, S., Grgic, M.: CoMoFoD - new database for copy-move forgery detection. In: Proceedings of 55th International Symposium ELMAR, pp. 25–27. IEEE (2013)
17. Yang, J., Jiang, Y.-G., Hauptmann, A.G., Ngo, C.-W.: Evaluating bag-of-visual-words representations in scene classification. In: Proceedings of the International Workshop on Multimedia Information Retrieval, vol. 63, p. 197. ACM Press (2007)
18. Zhang, Q., Shen, X., Xu, L., Jia, J.: Rolling guidance filter. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014. LNCS, vol. 8691, pp. 815–830. Springer, Cham (2014). doi:10.1007/978-3-319-10578-9_53
19. Zhang, T., Boult, T.E., Johnson, R.: Two thresholds are better than one. In: Conference on Computer Vision and Pattern Recognition, pp. 1–8. IEEE (2007)
20. DM Powers: Evaluation: from precision, recall and f-measure to ROC, informedness, markedness & correlation. J. Mach. Learn. Technol. **2**(1), 37–63 (2011)
21. Ryu, S.-J., Lee, M.-J., Lee, H.-K.: Detection of copy-rotate-move forgery using Zernike moments. In: Böhme, R., Fong, P.W.L., Safavi-Naini, R. (eds.) IH 2010. LNCS, vol. 6387, pp. 51–65. Springer, Heidelberg (2010). doi:10.1007/978-3-642-16435-4_5