Detecting Forgery in Duplicated Region using Keypoint Matching

N. Suganthi* , N. Saranya, M. Agila*****

*Dept. of Computer Science
Vivekanandha College for Women
Namakkal, Tamilnadu, India
Suganthics1988@gmail.com

**Dept of CSE
SKCT
Coimbatore, Tamilnadu, India
saranvlbjcet@gmail.com

***Dept of CSE
SKCT
Coimbatore, Tamilnadu, India
agilavlb@gmail.com

Abstract**—**Region duplication is a common form of image manipulation where part of an image is pasted to another location to conceal undesirable contents. Mostly used region duplication detection methods are directly matching the block of image pixels. But it is not effective when the duplicated regions have geometrical and illumination distortions. The proposed method uses SR-PE (Scale rotation invariant pattern Entropy) for obtaining the matching pattern which is much faster compared to other algorithms. It generally aims to measure the spatial regularity of matching patterns formed by local keypoints. This method consist of various transformations from which the duplicated region can be identified, by estimating the transform between matched SIFT(Scale Invariant Feature Transform) keypoints. The SIFT algorithm along with the SR-PE and RANSAC algorithm helps to find the duplicated regions more effectively.

Index Terms - Digital image forensics; Region Duplication Detection, Image Feature Matching, Keypoints.

I. Introduction

The availability and sophistication of digital imaging technology (e.g., cameras, computers, software) and their wide use on the Internet have made digital images a main source of information. Rapid advancement in imaging technology has made it remarkably easy to manipulate digital image contents. With the Proliferation of digital cameras and computers, as well as software for image editing, the problem of digital image forgery is potentially very serious. Digital image counterfeiting has already appeared in many disturbing forms. However, concomitant with the ubiquity of digital images is the rampant problem of digital forgeries, which has seriously debased the credibility of photographic images as definite records of events. Accordingly, digital image forensics has emerged as a new research field that aims to reveal tampering operations in digital images

A common manipulation in tampering with digital images is known as region duplication, where a continuous portion of pixels is copied and pasted to a different location in the same image. To make convincing forgeries, the duplicated regions are often created with geometrical or illumination adjustments. In recent years, several methods have been proposed to detect region duplication for the purpose of image forensics. These methods are based on finding pixel blocks that are exact copies of each other in an image. Such methods are most effective for the detection of region copy-move, where a region of pixels is pasted without any change to another location in the image.

A common form of digital tampering is Copy-Move forgery, in which a part of the image itself is copied and pasted into another part of the same image to conceal an important object. Because the copied part come from the same image, its important properties, such as noise, color and texture, will be compatible with the rest of the image and thus will be more difficult to distinguish and detect. Several researchers have developed techniques for detecting this form of image forgery. Since the key characteristics of Copy-Move forgery is that the copied part and the pasted part are in the same image, a direct method to detect this forgery is exhaustive search, but it is computationally complex. Another approach for detecting copy-move forgeries is the block-matching procedure, which first divides the image into overlapping blocks. The aim of this approach is to detect connected image blocks that were duplicated, instead of detecting the whole duplicated region. Since the copied region would consist of many overlapping blocks and moving the region means moving all the blocks by the same amount, the distance between each duplicated block pair would be the same. Therefore, the decision of forgery can be made only if there are more than a certain number of duplicated image blocks within the same distance and these blocks are connected to each other.

Our method is based on image keypoints and feature vectors that are robust to typical image transforms We formulate region duplication detection as finding transformed identical regions in an image and use robust estimation to obtain correct keypoints matching and transforms between duplicated regions simultaneously. With the estimated transforms, our methods further obtain the precise location and extent of the detected duplicated regions. Pixel is the fundamental display element of an electronic screen or bitmap image. Screen pixel resolution is rated by the number of horizontal keypoints and vertical pixels.

II. RELATED WORKS

Most existing region duplication methods [2, 8, 7, and 5] assume a region is pasted to a new location without any change, i.e., $T\theta$ is an identity. This special case of region duplication is often known as region *copy-move*, for the detection of which it is sufficient to compare pixel blocks and find exact copies. As a brute-force match of all pixel blocks of a given size in an image will have a running time quadratic to the size of the image, most methods focus on using low dimensional representations of blocks, e.g., PCA [8, 7, 5] or DCT [2], an fast lexicographical sorting to improve efficiencySeveral general techniques in digital image forensics may be applied to detect duplicated regions. However, in practice, direct copy-move may not achieve desirable tampering, and the pixel regions are typically undergone further processing before or after being copied, such as scaling, rotation and boundary smoothing. The latter case has been recently discussed in [4]. For region duplications that involvescaling and rotating of the region before pasting, which can significantly disturb the pixel blocks, detection methods based on direct matching pixel blocks are unlikely to be effective.

Many other existing region duplication detection methods are based on matching blocks of image pixels or transform coefficients (e.g., [1], [4], [3], [6],[2], [3], [9], [5], [8]). While these methods can detect duplicated regions pasted to the target location without any change (a special case known as copy-move), they are largely ineffective to detect duplicated regions that are also distorted (such as examples in Fig.1). To alleviate this problem, a variant of the block matching region duplication method is proposed to handle duplicated regions rotated with 90, 180 and 270 angles. Another vein of works use blocks in the log-polar coordinate system [38], [8], [5], where rotation and scaling become translation and can be detected as copy-move. Another method has been proposed to detect duplicated regions with smoothing operation [7]. However, the flexibilities provided by these methods are limited and they cannot be extended for the detection of duplicated regions with general distortions.

As an alternative to the block matching based detection methods, several recent methods have explored the use of matched image keypoints to identify duplicated regions. In [2], keypoints and features based on the SIFT algorithm [3] are used to account for illumination changes in the detection of copy-move region duplication. However, the robustness of SIFT keypoints and features to image distortions are not fully exploited, which prevents this method from being extended to detect affine transformed duplicated regions. In our previous work [3], we describe a SIFT matching based detection method that can locate duplicated regions with rotation or scaling. Another recent work [3] uses SIFT keypoint matching to estimate the parameters of the affine transform and recover matched keypoints. But similar to, it does not provide the exact extent and location of the detected duplicated region, but only displays the matched keypoints. Furthermore, these detection methods are typically evaluated against simple forgeries where human viewers have no trouble to identify the duplicated regions, and their performance on challenging realistic forgery images is largely unknown.

As described in [6], the first step in collecting SIFT features is to identify *keypoints* that are locations with distinct image information and robust to scaling and rotation. This is achieved by searching for locations that are stable local extrema in the image scale space, followed by a computation of the dominant local orientation at the key points. Note that the number of keypoints is usually much less than the number of pixels, thus subsequent computation will not be wasted at locations with little image information. At each keypoint, a SIFT feature vector is generated from the normalized histograms of local gradients in a neighborhood of pixels of that keypoint. The size of the neighborhood is determined by the scale of the keypoint, and all gradients are aligned with the dominant orientation at the keypoint. These steps ensure that the obtained local descriptors are invariant to rotation and scaling. With the setting in [6], the final SIFT features are 128 dimensional vectors at each keypoint.

Fig1. Input image

III. PROPOSED SYSTEM

Proposed method detects the distorted duplicated region in the following modules:

- Finding Image Keypoints
- Keypoints Matching
- Eliminating Mismatched keypoints
- Estimation of Affine Transform
- Identifying Duplicated Regions

A. Finding Image Keypoints

In the preprocessing stage the RGB image is converted into grayscale image.Then we apply SIFT algorithm for finding the keypoints.SIFT algorithm consist of the following stages:

i. Scale-space extrema detection
ii. Keypoint localization
iii. Orientation assignment
iv. Generation of keypoint descriptors

Good keypoints and features should represent distinct locations in an image, be efficient to compute and robust to local geometrical distortion, illumination variations, noise and other degradations.

Here, we present a new region duplication detection method based on the image SIFT features.
Specifically, to detect the locations.

ISSN 2250-3153

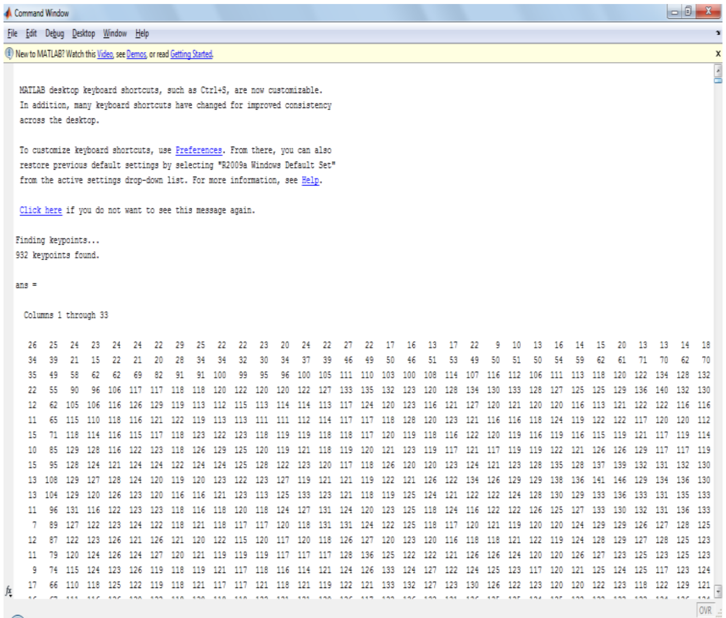

Fig2. Detected Keypoints

of potential duplicated regions, we first detect SIFT keypoints in an image like as shown in the figure (b). And compute the SIFT features for such keypoints. At each keypoint, a 128 dimensional feature vector is generated from the histograms of local gradients in its neighborhood. To ensure the obtained feature vector invariant to rotation and scaling, the size of the neighborhood is determined by the dominant scale of the keypoint, and all gradients within are aligned with the keypoint's dominant orientation dominant orientation. Furthermore, the obtained histograms are normalized to unit length, which renders the feature vector invariant to local illumination changes.

B. Keypoints Matching

The detected keypoints are matched using Scale-Rotation invariant Pattern Entropy (SR-PE) algorithm. It generally aims to measure the spatial regularity of matching patterns formed by local key points. Its effectiveness is illustrated with a large-scale image set. It mainly consists three components: bag-of-words representation, local key point matching and matching pattern evaluation. In the first step of the SR-PE algorithm, an exhaustive estimation of transformation parameters for all pairs of matching lines is carried out. The parameters are clustered with mean-shift algorithm [1], respectively, in the scale s and rotation μ channels. In the procedure, each cluster in a channel corresponds to one ND region pair. Because each matching pair formed by (pi, pj) involves multiple estimations of s and μ in the first step, each (pi, pj) basically can be clustered into more than one group. The cluster membership of (pi, pj) is further defused by assigning it to the most likely cluster where (pi, pj) resides most of the times. Consequently, we obtain two sets of region pairs corresponding to scale and rotation channels respectively. The following steps are:

i. Read the sift keypoints from the given input images.
ii. Compare the keypoints of one image with the other image and if the keypoints matches draw a line indicating the matched keypoints.

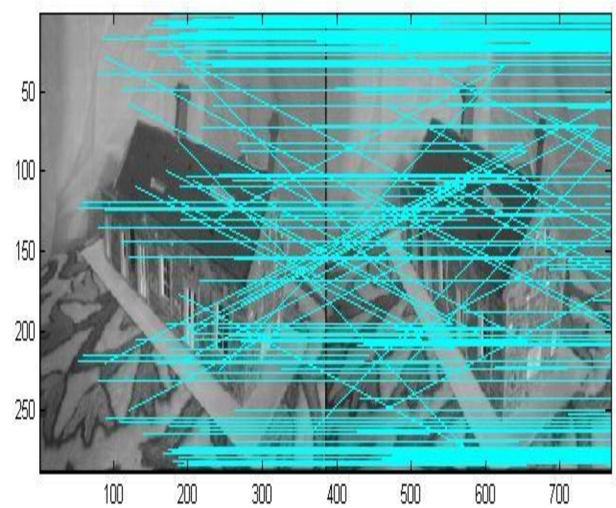

Fig3. Matched Keypoints

iii. The distance ratio is adjusted such that it gives out the best matches between them.
iv. Append the two images and then lines are drawn indicating the matches.

C. Eliminating Mismatched keypoints

We can use the matched SIFT keypoints to estimate the affine transform parameters, but the obtained results are inaccurate due to the large number of mismatched keypoints.To find out the unreliable keypoints we use Random Sample Consensus(RANSAC)algorithm.We run the ransac algorithm N times repeatedly to detect the duplicated region.It executes the following steps N times:

i. Randomly select three or more pairs of matche keypoints that are not collinear. Using the chosen pairs of keypoints, estimate T and shift vector x0 by minimizing the objective function given in Eq.(2).

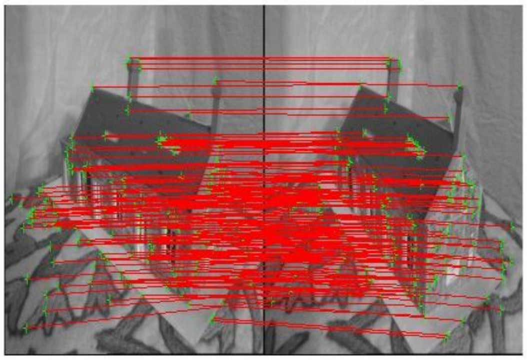

Fig 4.RANSAC Matched

ii) Using the estimated T and x0, classify all pairs of matched SIFT keypoints into inliers or outliers. Specifically, a pair of matched keypoints (x, ~x) is an inlier if ||~x-Tx-x0||2<=β, otherwise, it is an outlier.

The RANSAC algorithm returns with the estimated transform parameters that lead to the largest number of inliers. In our experiment, we choose default values for N = 100 and β = 3 as they lead to better empirical performance.

D. Estimation of Affine Transform

Based on the putative keypoint matching, we estimate the possible geometric distortions of the duplicated regions. To generalize transforms such as rotation, scaling and shearing that are supported in most photo-editing software, we model the distortion as affine transform of pixel coordinates. Given two corresponding pixel locations from a region and its duplicate as

$$x = (x, y)^T \quad \ldots(1)$$

$$\tilde{x} = (\tilde{x}, \tilde{y})^T , \quad \ldots(2)$$

respectively, they are related by a 2D affine transform specified by a 2x2 matrix T and a shift vector x0 as:

$$\tilde{x} = Tx + x0, \quad \ldots(3)$$

or more explicitly. We need at least three pairs of corresponding keypoints that are not collinear. In practice, due to imprecise matching, it may not be satisfied exactly, and we form the least squares objective function using matched keypoints and searching for T and x0 that minimize it.

E. Identifying Duplicated Regions

With the estimated region transform, we can establish the correspondence between all pixels in the original region and their counterparts in the duplicated region. A map of region correlations is then created to identify the original and the duplicated regions. In doing so, we first segment the image into overlapping *contour blocks* of 4 × 4 pixels. Using the parameters estimated in section 3.3, we transform the tampered image and compute the correlation coefficient between each pair of corresponding contour blocks which generates a correlation map. We process the correlation map by first applying a Gaussian filter of 7 × 7 to remove the artifacts at the edge and then obtaining all possible original and duplicated regions where the correlation coefficient is larger than a pregiven correlation threshold. Next, we binarize the correlation map by resetting the value to one for all locations where the correlation coefficient is larger than a threshold value and zero otherwise. This is followed by removal of regions with area smaller than an area threshold so to reduce the effect of noise.Finally, the contours of the potential original and duplicated regions are connected with mathematical morphological operation to the duplication regions that (1) dilated then eroded to eliminate holes in the detected regions, and (2) eroded then dilated to smooth the region contours [9]..

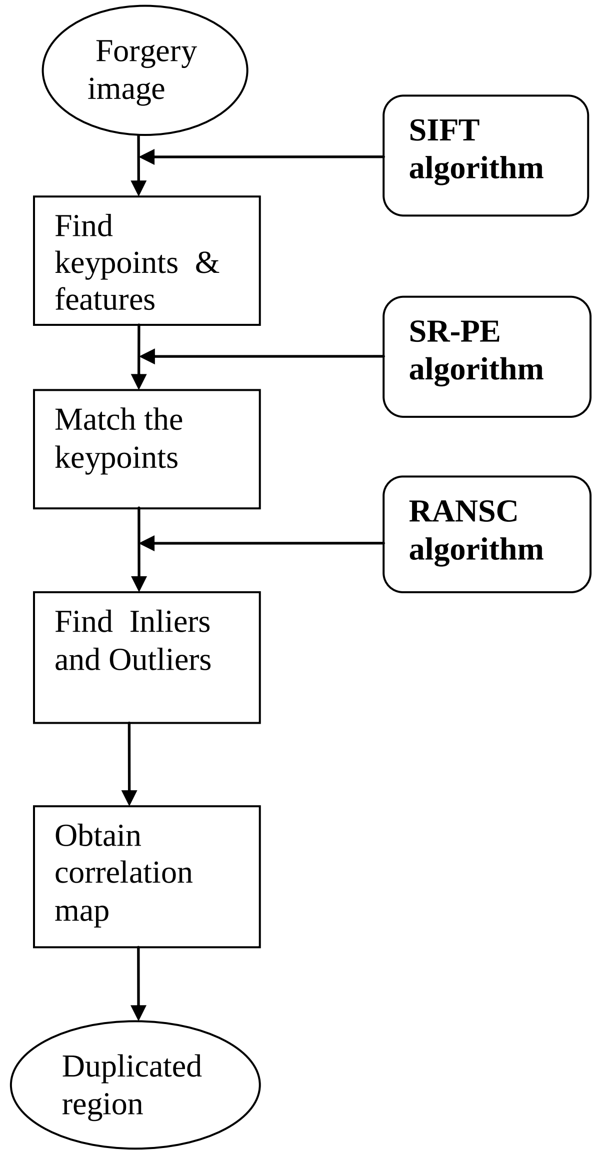

Fig 5: Block diagram

IV. CONCLUSION

Region duplication is an important problem in the field of digital image forensics. In this paper, we describe an effective method to detect image region duplication. Our method is based on local image SIFT features, which makes it applicable to the detection of general region duplications with region scaling and rotation. Experimental results demonstrate that this method is effective and robust in the presence of additive noise and different JPEG qualities. Compared to other method where only matched key points are shown as detection results, we further estimate the transform between duplicated regions based on SIFT features and recover the complete region contours using correlation map.As an important future work, we will consider several approaches to improving the detection performance for such cases, including incorporating other features such as PCA-SIFT or histograms of oriented gradients, and combining with other detection schemes based on intrinsic signal statistics/patterns to provide strong cues when image keypoints and features are not sufficient.

REFERENCES

[1] Xunyu Pan and Siwei Lyu, "Region duplication detection using image feature matching",IEEE Transaction on Information forensics and security, VOL. X, NO. X, XX 2010.
[2] Xunyu Pan and Siwei Lyu ,"Detecting Image Region Duplication using sift features ", IEEE Transaction on Information forensics and security,2010
[3] Weiqi Luo, Jiwu Huang and Guoping Qiu "Robust Detection of Region-Duplication Forgery in Digital Image", IEEE Transaction on Information forensics and security, 2006.
[4] Sevinc Bayram, Husrev Taha Sencar , Nasir Memon "An efficient and robust method for detecting copy move forgery" IEEE Transaction on Information forensics and security,2009.
[5] B.L.Shivakumar1 Lt,Dr. S.Santhosh Baboo," Detecting Copy move forgery in digital images:" A Survey and Analysis of current Method, Global Journal of computer science and technology,Vol. 10 Issue 7 Ver.1.0 September 2010.
[6] A.C. Popescu and H.Farid. "Exposing digital forgeries by detecting duplicated image regions". Technical Report TR2004-515, Department of Computer science, Dartmouth College, 2004.
[7] A.Langille and M.Gong."A efficient match-based duplication detection algorithm". In The Canadian conference on computer and robust vision, 2006.
[8] D.Letscher,"Detecting filtered cloning in digital images". In ACM workshop on MM & Sec,2007.
[9] H-J. Lin, C-W. Wang and Y-T. Kao," Fast copy-move forgery detection". SEAS Trans,on sig.proc.. 5(5):188–197, 2009.
[10] D.Lowe,"Distinctive image features from scale-invariant keypoints".IJCV,60(2):91-110,2004.

AUTHORS

First Author – N.Suganthi, M.Sc.,M.phil, Vivekanandha college for women , suganthics1988@gmail.com and 9487203508.

Second Author – N.Saranya, BE, SKCT , saranvlbjcet@gmail.com and 7708828385.

Third Author – M.Agila, BE, SKCT , agilavlb@gmail.com and 9790347694