ISSN No. 0976-5697
Volume 7, No. 3, May-June 2016
International Journal of Advanced Research in Computer Science
RESEARCH PAPER
Available Online at www.ijarcs.info

Passive Copy-Move Forgery Detection using SIFT, HOG and SURF Features

Shiv Prasad
Department of Computer Science and Engineering
Birla Institute of Technology, Mesra
Ranchi, India

S. Chaudhuri
Department of Computer Science and Engineering
Birla Institute of Technology, Mesra
Ranchi, India

B. Ramkumar
Department of Computer Science and Engineering
Birla Institute of Technology, Mesra
Ranchi, India

Abstract: Copy-move is a common type of digital image forgery. In an image, Copy-Move tampering might be done to hide an undesirable region, or to duplicate something in the image. These images might be used for necessary purpose like evidence in the court of law. So, authenticity verification plays a vital role for digital images. In this paper, we compare the CMFD (Copy-Move Forgery Detection) using Image features like SIFT (Scale Invariant Features Transform), HOG (Histogram Oriented Gradient) and SURF (Speed-Up Robust Features) and hybrid features (SURF-HOG and SIFT-HOG). The comparison results show that CMFD using SIFT features provide better results as compared with SURF and HOG features. Also, considering hybrid features, SIFT-HOG and SURF-HOG produce better results for CMFD using SIFT, SURF or HOG alone.

Keywords: Passive Forensics, Copy-Move Forgery Detection, Single Region CMFD, Multiple Region CMFD, Histogram Oriented Gradient, Scale Invariant Features Transform, Speed-Up Robust Features.

I. INTRODUCTION

Today, with the rapid growth of powerful computers, easy-to-use image editing software, and advancement in digital cameras, the authenticity of any digital image can no longer be trusted. The authenticity of these images has an important role as these are popularly used as supporting evidence and historical records for numerous applications related to law enforcements, defence, surveillance, insurance claims, medical imaging, journalistic photography and commercial applications. It's important to study and develop some creditable and robust methods to detect whether a digital image is authentic or tampered.

Copy-move is the simplest and well known method of image tampering, where for hiding or exposing some object or scene in a picture, a region of the image is copied and then pasted onto another region in the same image. For example, in Fig. 1, copy-move forgery is done (the cloned objects and original objects are encircled in red color and yellow color respectively).

There are many methods for detecting forgeries done on any digital image. These techniques directly or indirectly analyzes the pixel-level correlations resulting from a particular type of forgery. The different image tampering approaches are Splicing, Copy-Move, and Resampling. Lossy image compression methods such as JPEG may create a significant challenge for a forensic analyst.

Many CMFD techniques (copy-move forgery detection techniques) are based on the direct matching of an image pixels block or transform coefficients [9]. But they fail when the copy-move regions have undergone some transformation like rotation, scaling or illumination. Image features such as SURF (Speed up Robust Features) [2], SIFT (Scale Invariant Features Transform) [7] are invariant with respect to illumination and geometrical transformation. Hence, these features provide better results for CMFD with copy-move regions scaled or rotated. The overview of the CMFD method using image features can be easily understood (as in Fig. 2).

Reliable matching between different objects are performed using SIFT features within the scene. It extracts distinctive invariant features from images. The SIFT features are invariant concerning illumination, rotation and scaling [7]. It provides robust matching even with the addition of noise.

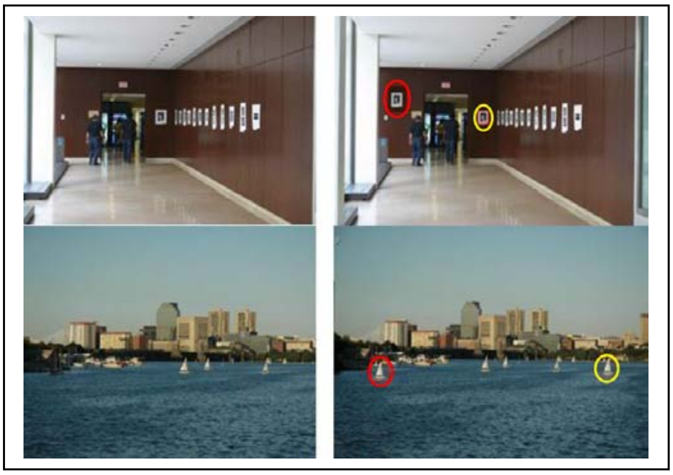

Figure 1. Examples of Copy-Move Image Forgery. Left Column: Original Images. Right Column: Tampered Images (cloned objects and original objects are encircled in red color and yellow color respectively).

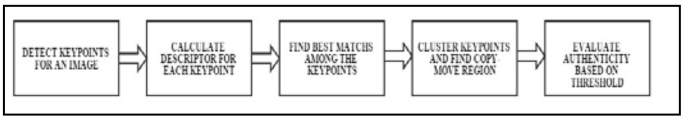

Figure 2. Overview of the CMFD Method using Image Features

SURF, concerning repeatability, distinctiveness, and robustness, outperforms previously proposed methods. Further, SURF features can be compared and computed much faster than SIFT. SURF leads to a combination of feature description, detection, and matching.

HOG determines the number of occurrences of gradient orientation of a digital image in localized portions. HOG method uses overlapping of local contrast normalization for improved accuracy. HOG works well in combination with SURF or SIFT for CMFD process [4].

In this paper, we investigate the Copy-Move attack of an image forgery. The proposed method can detect single and multiple Copy-Move (cloned region) within a digital image.

The whole paper is divided into seven sections. The second section describes the related work of CMFD. The third section describes the use of SIFT features in CMFD; the fourth section describes the use of SURF features for CMFD, and the fifth section describes the use of HOG features in CMFD. The experimental results are shown in the sixth section. Finally, the seventh section provides the conclusion and future direction of our work.

II. RELATED WORK

Multimedia Forensics has developed some methods that test the authenticity in the absence of watermarks [8, 5, 10] and are defined as Passive Methods of Image Forensics.

In CMFD, a correlation between the pasted region and the original image is analyzed by separating the image into overlapping blocks with low dimensional representation by applying a feature extraction process. Copy-move forgery detection should be highly robust to scaling, illumination, rotation. For e.g. [9] was not able to detect rotation or scaling transformation, unlike in [3, 6], where some of the rotation and scaling were detected. [11] using Zernike moments, made an attempt to overcome the problem. However, the forgery detection takes place only when the rotation of the copied area took place. [12] proposed a technique using block description invariant to rotation and reflection to detect forged regions. [1] applied CMFD using SIFT features and also estimate the transformation by RANSAC algorithm.

III. CMFD USING SIFT FEATURES

Considering an input image I, using scale–space representation method, we can detect SIFT features implemented in the form of an image pyramid. We select the interest points as the local extrema and we can obtain the image pyramid levels by sub-sampling of the image resolution and Gaussian smoothing. Using Difference of Gaussians (DoG) which is a computable approximation of the Laplacian of Gaussian, we can extract. We can represent a DoG Image D by the following equation:

$$\begin{aligned} D(x,y,\sigma) &= (GB(x,y,k\sigma) - GB(x,y,\sigma)) \\ *I(x,y) &= LG(x,y,k\sigma) - LG(x,y,\sigma) \end{aligned} \tag{1}$$

where LG(x, y, kσ) is the convolutions of the original image I(x,y) and GB(x, y, kσ) is the Gaussian blur at scale kσ. For feature extraction to be made invariant concerning rotations, canonical orientation o is applied by the SIFT algorithm to each keypoint a. To show this orientation, in the neighborhood of the keypoint, a HOG is computed. In a case of given an image sample LG(x, y, σ) at scale σ, the orientation Θ(x, y) and gradient magnitude m(x, y) are pre-computed using pixel differences by the following equations:

$$m(x,y) = \begin{pmatrix} \left(LG(x+1,y) - LG(x-1,y)\right)^2 + \\ \left(LG(x,y+1) - LG(x,y-1)\right)^2 \end{pmatrix}^{1/2} \tag{2}$$

$$\theta(x,y) = \tan^{-1}\left(\frac{LG(x,y+1) - LG(x,y-1)}{LG(x+1,y) - LG(x-1,y)}\right) \tag{3}$$

The nth SIFT keypoint x_n (x, y, σ, o, f) have the following information: *x, y* represent coordinates of the image plane, σ represent scale, and *f* represent the final descriptor [1]. We summarize the method of CMFD using SIFT features as follows:

1) SIFT features for an image is calculated, and the Euclidian Distance is determined between each pair of SIFT keypoints.
2) By selecting the appropriate threshold based on the minimum distance obtained from above step, best matches are determined.
3) Cluster Centres are defined by the best match Keypoints, and by using a threshold for the Euclidian Distance and Cluster size, match the clusters.
4) Display the clusters if both clusters have a minimum of two points to decide whether the image is forged or authentic.

IV. CMFD USING SURF FEATURES

Concerning processing speed, SURF is a Hessian matrix based unique scale and rotation-invariant interest point descriptor and detector. The Haar wavelet distribution responses within the interest point neighborhood is described by the descriptor. For speed, integral images are utilized. The time for feature matching and computation is reduced as it uses only 64 dimensions, and there by simultaneously increasing the robustness. The sign of the Laplacian is the basis of the indexing step. In an image, for a given point x = (x_i, y_i), the Hessian matrix H(x, σ) at scale σ can defined as:

$$H(x,\sigma) = \begin{vmatrix} L_{xx}(x,\sigma) & L_{xy}(x,\sigma) \\ L_{xy}(x,\sigma) & L_{yy}(x,\sigma) \end{vmatrix} \tag{4}$$

where $L_{xx}(x,\sigma)$ is the convolution of the image I at point x with Gaussian second order derivative $\frac{\partial^2}{\partial x^2} g(\sigma)$ and similarly for $L_{xy}(x,\sigma)$ and $L_{yy}(x,\sigma)$.

The 9 × 9 box filters Gaussian second order derivatives approximations with σ = 1.2 with lowest scale representation. The estimates are denoted by D_{xx}, D_{yy}, and D_{xy}. For computational efficiency, we balance further the expression relative weights for the Hessian's determinant as shown below:

$$\frac{\left|L_{xy}(1.2)\right|_F \left|D_{xx}(1.2)\right|_F}{\left|L_{xx}(1.2)\right|_F \left|D_{xy}(1.2)\right|_F} \cong 0.9 \tag{5}$$

where, $|x|_F$ is the Frobenius norm which yields:

$$Det(H_{approx}) = D_{xx}D_{yy} - (0.9D_{xy})^2 \quad (6)$$

To calculate the descriptor, select an orientation of a square region which is centered at the interesting point and then divided into sub-regions of 4 × 4 square block. The Haar-wavelet response in a vertical direction (dy) and horizontal direction (dx) for each sub region is recorded with a filter size 2s (s-scale) at sample points of regularly spaced 5 × 5 block. Summing up the wavelet responses over each sub-region will form a feature vector. For the underlying intensity structure of each sub-region, we define a four-dimensional descriptor vector v as shown below:

$$v = (\sum dx, \sum dy, \sum |dx|, \sum |dy|) \quad (7)$$

This results in a descriptor vector of length 64 for all 4 × 4 sub-regions. The wavelet responses are consistent with illumination (offset) of bias. We summarize the method of CMFD using SURF features as:

1) SURF features for an image is calculated and Euclidian Distance determined between each pair of SURF keypoints.
2) By selecting the appropriate threshold based on the minimum distance obtained from above step, best matches are determined.
3) Cluster centres are defined by the best match keypoints and by using a threshold for the Euclidian Distance and Cluster size, match the clusters.
4) Display the clusters if both clusters have a minimum of two points to decide whether the image is forged or authentic.

V. CMFD USING HOG FEATURES

In HOG descriptor, first, horizontal and vertical gradients are calculated to compute gradient orientations and magnitude [4]. For a 64 x128 detection window, divide the image into 16 x 16 sized blocks with 50% overlap. Divide each block into cells of size 8 x 8 each. Quantize the gradient magnitudes in 9 bins. Then all the histograms for all the blocks in a window are concatenated to compute the descriptor that is 3780-dimension vector.

Computation of the HOG descriptor requires the following basic configuration parameters: computational masks for derivatives and gradients, normalization parameter, geometry of the image splitting into cells and grouping cells into a block, block overlapping. According to [4], the recommended values for the HOG parameters are: 1D centered derivative mask [-1, 0, +1]. detection window size of 64 x 128, cell size of 8 x 8, block size of 16 x 16 (2 x 2 cells). We summarize the method of CMFD using HOG features as:

1) Apply Level 1-DWT for getting the Low-Pass Approximation after converting RGB image to Gray-Scale Image.
   After applying 1-DWT, the image is divided into sub images of one low-frequency and several high-frequency details along the diagonal, vertical, horizontal directions.
2) Divide the image resulted from Step 1 into Overlapped Blocks.
   A square block of size B × B will generate (M-B+1) × (N-B+1) when slid over the image of size M × N.
3) HOG features for each block are extracted.
   The resulting cell histograms are then combined into a descriptor vector for each block. For an image of size M × N, matrix H will include (M-B+1) × (N-B+1) descriptors, where B is the size of each block.
4) Lexicographically Sort the feature vectors obtained from step 3.
   The lexicographically sorting is done so as to reduce the time required for block matching such that similar feature vectors are stored in neighboring rows. Let FV be the sorted Matrix.
5) Block Matching.
   Let FVi denote the ith row of the sorted matrix FV and (xi, yi) denote the block's image coordinates. Consider adjacent rows FVj, whose row distance, |i- j| in the sorted matrix FV is less than a threshold (Nn). The offset of all such pairs is given by:
   (xi – xj, yi – yj) if xi > xj,
   (xj – xi, yi – yj) if xi < xj,
   (0, |yi-yj|) if xi = xj
6) Cluster the Matched Pairs according to the offset for each pair and filter out small clusters.
   To denote the duplicated regions, offsets with high occurrence are determined, i.e., a largely duplicated region will consist of many smaller blocks, and these blocks will appear in neighboring places after lexicographical sorting (Step 4). They have the same offset. To avoid false detection, offset magnitudes below a specified threshold (Nd) are ignored.
7) Display the clusters on the image and decide whether the image is forged or authentic.

VI. RESULTS AND DISCUSSION

We evaluate results for each of the proposed methods and determine the accuracy for each of them. The algorithms for comparison were implemented using MATLAB and executed in a system with the following specification:

- Processor: Intel® Core™ i7-4710HQ CPU @ 2.50 GHz
- RAM: 8.00 GB
- Operating System: Windows 10 (64-bit)

For evaluating the results, we have considered MICC–F220 [1] image dataset. The dataset is composed of 220 images, out of which 110 are tampered images and 110 are originals images. For the images, a Cluster Size of 20 was selected. For larger sized images, Cluster Size can vary between 50 and 100. We define accuracy, precision and recall for proposed methods as given below:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

$$Precision = \frac{TP}{TP + FP} \quad (9)$$

$$Recall = \frac{TP}{TP + FN} \quad (10)$$

where, TP is True Positive Matches; TN is True Negative Matches; FP is False Positive Matches; FN is False Negative Matches. Also, we can calculate the False Positive Rate (FPR) as:

$$FPR = \frac{FP}{FP + TN} \quad (11)$$

In Fig. 3. results of various CMFD techniques using SIFT, SURF, HOG, SIFT-HOG and SURF-HOG features are shown We can visualize that original objects are encircled in yellow color and the objects which got tampered by applying copy-move forgery of the original object are encircled in red color. Original objects and tampered objects are mapped using keypoint descriptors which can be visualized using blue colored lines.

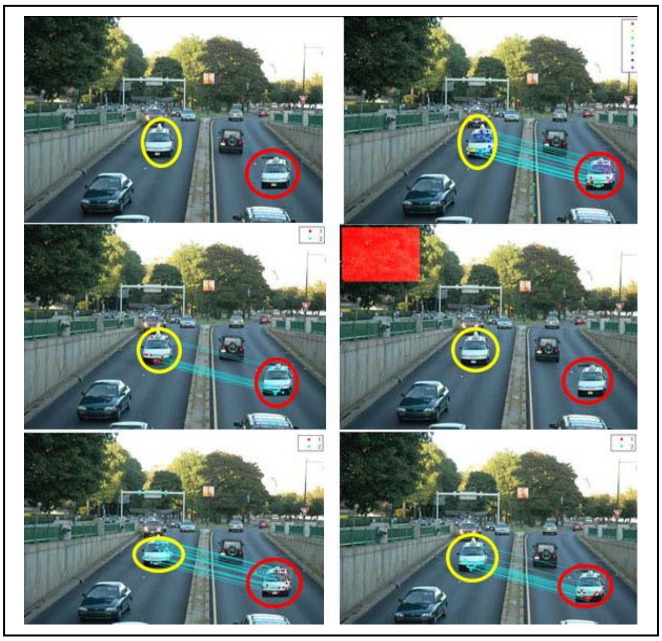

Figure 3. CMFD Results. Top Row: Input Image (Left), CMFD Result using SIFT Features (Right). Middle Row: CMFD Result using SURF Features (Left), CMFD Result using HOG Features (Right). Bottom Row: CMFD Result using SIFT-HOG Hybrid Features (Left), CMFD Result using SURF-HOG Hybrid Features (Right).

Table I. Performance Comparison of CMFD using SIFT, SURF and HOG Image Features

Image Features	*Performance (%)*			
	Accuracy	*Precision*	*Recall*	*False Positive Rate*
SIFT	98.64	98.20	99.09	3.36
SURF	97.27	97.27	97.27	0
HOG	49.09	45.45	9.09	10.91

Table II. Performance Comparison of CMFD using SIFT-HOG and SURF-HOG Hybrid Features

Image Features	*Performance (%)*			
	Accuracy	*Precision*	*Recall*	*False Positive Rate*
SIFT-HOG	99.09	98.21	100	1.81
SURF-HOG	97.72	96.46	99.09	3.36

Table I shows the experimental results for each of the proposed methods. We can observe that:

- CMFD using SIFT features (Fig. 3, Top Row, Right Column) provides the best result with an accuracy, precision and recall of about 98.64%, 98.20% and 99.09% respectively when compared with SURF (Fig. 3, Middle Row, Left Column) which resulted with an accuracy, precision and recall of about 97.27%, 97.27% and 97.27% respectively, and HOG (Fig. 3 Middle Row, Right Column) features which yielded comparatively inferior results in terms of accuracy, precision and recall of about 49.09%, 45.45% and 9.09% respectively.
- In terms of false positive rate SURF features produced ideal result of about 0% when compared to SIFT with false positive rate of 3.36%, and HOG features with false positive rate of 10.91%.

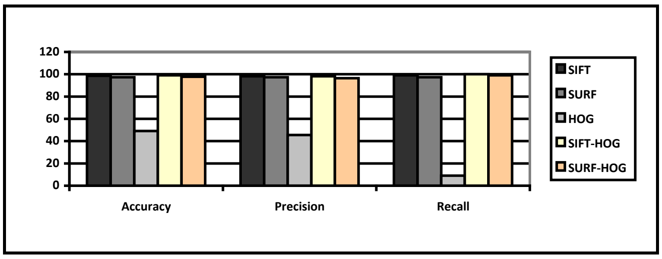

Figure 4. A Comparison among SIFT, HOG, SURF and Hybrid (SIFT-HOG and SURF-HOG) Image Features based CMFD.

We also considered hybrid features by taking keypoints of SIFT and SURF and calculating the HOG descriptors for both. After considering hybrid features (SIFT-HOG and SURF-HOG) (Fig. 3, Bottom Row, Left and Right Column respectively), from Table I and Table II, we can observe:

- There are improvements in the results using SIFT-HOG with an accuracy, precision, recall and false positive rate of about 99.9%, 98.21%, 100% and 1.81% respectively when compared with SIFT features.
- There are improvements in the results using SURF-HOG with an accuracy and recall of about 97.72% and 99.09% respectively but decline in terms of precision and false positive rate of about 96.46% and 3.36% respectively when compared with SURF features.
- The performance of SIFT-HOG is better than SURF-HOG features.

A graphical representation of Table I and Table II are shown in Fig. 4.

VII. CONCLUSION

Copy-move is the simplest and well known method of image tampering, where for hiding or exposing some object or scene in a picture, a region of the image is copied and then pasted onto another region in the same image. We compared copy-move forgery detection using SIFT, SURF and HOG image features and hybrid features (SIFT-HOG and SURF-HOG). The results were recorded using SURF, HOG and SIFT image features. By our experimental results, we observed that:

- SIFT provided the best outcome in the form of accuracy and precision when compared with SURF and HOG image features.
- After considering hybrid features (SURF-HOG or SIFT-HOG), we are getting better results for CMFD in comparison to SIFT, SURF or HOG features when used alone. Hence, if we want commendable precision and accuracy in CMFD, then we should select SIFT or Hybrid features.
- In terms of time, the CMFD using SURF is very fast compared to SIFT or hybrid features. Hence, if we want fast CMFD in term of time, then we should choose SURF.
- If we want better localization of Copy-Move region, then we should choose Hybrid Features.

VIII. REFERENCES

[1] I. Amerini, L. Ballan, R. Caldelli and A. D. Bimbo, "A SIFT-Based Forensic Method for Copy-Move Attack Detection and Transformation Recovery", IEEE Transactions on Information Forensics Security, Vol. 6, No. 3, pp. 1099–1110, 2011.

[2] H. Bay, T. Tuytelaars and L. V. Gool, "SURF: Speeded Up Robust Features [CVIU]", Journal of Computer Vision and Image Understanding, Vol. 110, No. 3, pp. 346–359, 2006.

[3] S. Bayram, H. T. Sencar and N. Memon, "An Efficient and Robust Method for Detecting Copy-Move Forgery", Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 1053-1056, 2009.

[4] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection", Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 886-893, 2005.

[5] H. Farid, "A Survey of Image Forgery Detection", IEEE Journal of Signal Processing Magazine, Vol. 2, No. 26, pp. 16–25, 2009.

[6] J. Fridrich, D. Soukal and J. Lukas, "Detection of Copy-Move Forgery in Digital Images", Proceedings of the Digital Forensic Research Workshop , 2003

[7] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", International Journal of Computer Vision, Vol. 60, No. 2, pp. 91-110, 2004.

[8] S. Lyu and H. Farid, "How Realistic is Photorealistic?", IEEE Transactions on Signal Processing, Vol. 53, No. 2, pp. 845–850, 2005.

[9] A. Popescu and H. Farid, "Exposing Digital Forgeries by Detecting Duplicated Image Regions", Computer Science, Technical Report, TR2004-515, Dartmouth College, 2004.

[10] J. A. Redi, W. Taktak and J. L. Dugelay, "Digital Image Forensics: A Booklet for Beginners", Multimedia Tools and Applications, Vol. 51, No. 1, pp. 133–162, 2011.

[11] S. J. Ryu, M. J. Lee and H. K. Lee, "Detection of Copy-Rotate-Move Forgery using Zernike Moments", Proceedings of the 12th International Conference on Information Hiding, pp. 51-65, 2010.

[12] B. Solorio and A. K. Nandi, "Passive Method for Detecting Duplicated Regions Affected by Reflection, Rotation and Scaling," Proceedings of the European Signal Processing Conference (EUSIPCO), pp. 824-828, 2009.

