Region Duplication Forgery Detection using Histogram of Oriented Gradients

Revathy G S
Department of Electronics and Communication
Marian Engineering College
Trivandrum, India

Dhanya Mathew
Department of Electronics and Communication
Marian Engineering College
Trivandrum, India

Abstract—Digital images are powerful means of communication and it is widely used in many fields. With the availability of powerful image editing software packages altering of digital images is done. The use of digital images in law enforcement and forensics necessitates the development of efficient forgery detection technique to prove the reliability and authenticity of image. The most common type of image forgery is copy move forgery or region duplication. In this type of forgery, one region of image is copied and pasted somewhere else in the image. Main aim of this type of forgery is to hide some important information. Tampered images may affect the statistical properties of the image which is used to detect forgery. In this paper, we propose a block based approach in which discrete wavelet transform is applied to the input image. Then the image is divided into overlapping blocks. Discriminative features are extracted by applying histogram of gradients to each block. These features are lexicographically sorted and duplicated blocks are detected after applying block matching step. Experimental results show that the proposed method can identify the forged region in the images very efficiently.

Keywords—Image forgery, region duplication, discrete wavelet transform, histogram of oriented gradients, lexicographically sorting, duplicated region detection.

I.INTRODUCTION

Digital images have often served as evidence in courts. With the development of powerful computer graphics editing software creation of fake images has become simple. This necessitates the need of efficient forgery detection methods to verify the integrity and reliability of the images. Vast range of image tampering methods and tools has made it difficult to have a single tool which can detect of all types of image forgery.

Existing digital forgery detection methods are divided into

- Active
- passive (blind)

Active approaches can be divided into digital watermark and signatures. The scope of this method is limited because every digital device will not have the ability to embed digital signatures. In contrast to active approaches, passive approaches do not need any embedded and priori information about the image. In passive approach it is assumed that forgery

may affect the statistical properties of the image which is used to detect forgery. Passive approaches are commonly divided into

- Image retouching – It enhances or reduces certain feature of an image which does not significantly change an image.
- Image splicing – This technique combines two or more images to create a fake image.
- Copy move attack or region duplication- In this method some part of information from an image is copied and pasted in the same image.

The methods used for copy move forgery detection can be mainly classified as either block-based methods or key point based methods. The main idea behind block based approach is that image is initially divided into overlapping or non-overlapping blocks and the blocks are compared against each other in order to find which blocks are matching.

Fridrich et al. [2] proposed a block matching detection scheme based on discrete cosine transform (DCT). But this method cannot differentiate between large identical textures of a natural image. Popescu and Farid [3] proposed a copy move forgery detection method, which is similar to DCT approach. It uses principal component analysis (PCA) instead of DCT. This method is better in capturing discriminating features. However, the efficiency decreases as the block size decreases. The method of copy-move forgery detection based on intensities was proposed by Luo, Huang and Qiu [4] which provide robust to attacks, such as JPEG compression, Gaussian blurring, and additive noise but fails when the images are highly distorted. Mahdian et al. [5] used blur invariant moments to detect forged regions, which provided robustness against post processing. The computation time needed for this algorithm is comparatively high. Zhao et al. [6] exploited DCT and SVD for the detection of duplicated regions which results in high computational complexity. Zhang, Feng and Su [7] applied DWT to find the duplicated regions. But this method depends on the location of the forged region and cannot be applied to an image which has a copied region at the center of the image.

Key point based methods selects key points from an image. The feature descriptors are calculated from these key points. Duplication matching is done on these feature descriptors. A study by Huang, Guo and Zhang [8] introduced SIFT (Scale Invariant Feature Transform) algorithm for detection but they

faced difficulties in implementing high scale image. A SIFT based method for the detection of copy move attacks and transformation recovery was proposed by Amerini et al. [9].SIFT-based schemes are still limited in their performance because it is only possible to extract key points from specific locations in an image and they are susceptible to a number of post-processing operations.

The histogram of oriented gradients (HOG) is a robust tool for detection [10]. This paper proposes a block based approach, which uses HOG descriptors to extract features for forgery detection. Experimental results demonstrate that the proposed method achieves better performance than the DWT based method [1].

II. PROPOSED METHOD

In our proposed method, Hog descriptors are used to extract discriminative features which are further compared to find the region duplication. The method can be divided into 7 steps as shown in fig.1.
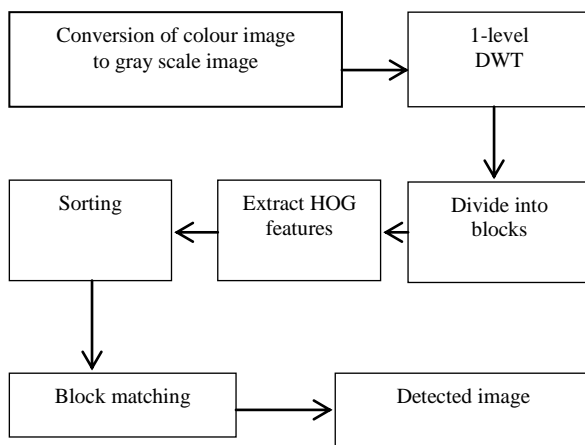
Fig.1. The work flow of HOG based region duplication Detection

A. Conversion of Colour Image into Gray Scale Image

If the input image is coloured then it is converted into gray scale image (I) by calculating weighted average of the Red (R), Green (G) and Blue (B) components according to the formula:

$$I = 0.2989\times R + 0.5870\times G + 0.1140\times B \qquad (1)$$

B. Applying DWT

The DWT has been found to be an effective signal and image processing tool which is advantageous to inherit multi resolution characteristics of images. Discrete wavelet transform helps reduction of size of the image at each level. At

each level, image is divided into one low frequency sub image and several sub images bearing the high frequency details along the horizontal, vertical and diagonal directions. Further steps will be performed on the resulting image of this step.
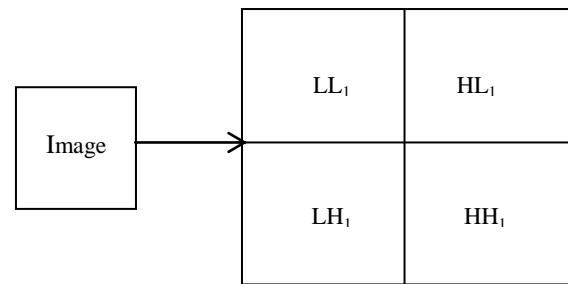
Fig.2. Pyramidal decomposition of 1-level 2-D DWT

C. Dividing Image into Overlapped Blocks

The input image is divided into overlapped blocks. The gray scale image (I) of size M×N will generate (M-B+1) × (N-B+1) overlapping blocks when a square block of size B×B is slided over the image.

D. Extracting HOG Features

HOG is applied to each block. Here four bins are considered for the local histogram. Each histogram bin corresponds to an orientation interval of 45^0 because the channels of the histogram are evenly spread between 0^0 and 180^0. The resulting cell histograms are then combined into a descriptor vector for each block, where four features can be used to represent each block. For an image of size M×N, matrix H will include (M-B+1) × (N-B+1) rows and four columns.

E. Lexicographical sorting

The feature vectors obtained from this are lexicographically sorted so that similar feature vectors are stored in neighboring rows which will reduce the time required for block matching.
Let A be the sorted matrix.

F. Block matching

Let Ai denote the ith row of the sorted matrix A and (xi, yi) denote the block's image coordinates. Consider adjacent rows Aj, whose row distance, |i- j|, in the sorted matrix A is less than a threshold (Nn). The offset of all such pairs is given by:

$$(xi - xj, yi -yj) \text{ if } xi - xj > 0$$
$$(xj - xi, yi - yj) \text{ if } xi - xj < 0$$
$$(0, |yi - yj|) \text{ if } xi = xj \qquad (2)$$

Note the offsets with high occurrence which denotes the duplicated regions i.e., a large duplicated region will consist of many smaller blocks, and these blocks will appear in neighbouring places after lexicographical sorting. They have same offset. In order to avoid false detection, offset magnitudes below a specified threshold (Nd) are ignored.

G. Output

As a part of post processing step erosion followed by dilation is performed in order to remove false detection. Remaining block pairs belongs to the duplicated region and an output image is created to mark all the duplicated regions.

III. EXPERIMENTAL RESULTS

Detection examples of the proposed method are shown in Fig.3, 4 and 5. The parameters used for our method are set as: B=16, Nn=5, Nd=15.
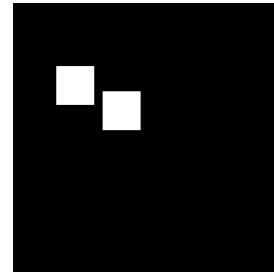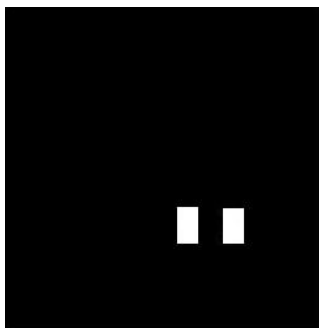

(c)

Fig.4.Detection example: (a) Original image, (b) forged image, (c) detection result

(a)   (b)

(c)

Fig.3.Detection example: (a) Original image, (b) forged image, (c) detection result

(a)   (b)

(a)

(b)

(c)

Fig.5.Detection example: (a) Original image, (b) forged image,
(c) detection result

IV. PERFORMANCE MEASUREMENT

Two evaluation criteria are used to illustrate the performance of the proposed algorithm namely the correct detection ratio, F_c and the false detection ratio, F_f.

a) Correct detection ratio (F_c)

F_c denotes the ratio in which the algorithm can correctly locate the pixels of copy move regions in the region duplicated image.

$$F_c = \frac{|\mu \cap \mu^c| + |w \cap w^c|}{|\mu| + |w|} \qquad (3)$$

b) False detection ratio (F_f)

F_f denotes the ratio of false positive i.e., pixels incorrectly identified as duplicated regions.

$$F_f = \frac{|\mu^c - \mu| + |w^c - w|}{|\mu| + |w|} \qquad (4)$$

where μ and w respectively denote pixels of original region and tampered region in forged image, μ^c and w^c respectively denote pixels of original region and tampered region in detected results. $\|$ denote area of the region, $\cap$ denote the intersection of two regions and $-$ denote the difference between two regions. The method will be more precise if F_c is closer to 1 and F_f is closer to 0. Table I shows the performance of proposed method and DWT based method for a 256 x 256 image.

TABLE 1: Performance results

PARAMETER	PROPOSED METHOD	DWT BASED METHOD
NO. OF 16 X 16 BLOCKS	13456	58081
FEATURE VECTOR LENGTH	16	30
CORRECT DETECTION RATIO	0.9963	0.9971
FALSE DETECTION RATIO	0.0032	0.0273
ESTIMATED TIME	92.52 Sec	1645.87 Sec

V. CONCLUSION

In this paper major focus was given for region-duplication or copy move forgery. The proposed block based approach is capable of finding the duplicated regions without any prior information about the image and it works in the complete absence of digital watermarks or signatures. Experimental results demonstrate that the proposed method achieves better performance than the DWT based method. Even though this method is efficient in terms of performance and time complexity is reduced still there is a scope of future work. Neural network, combinatorial optimization solution methodology can be associated with this method.

REFERENCES

[1] Y. Wang, K. Gurule, J. Wise, J. Zheng, "Wavelet Based Region Duplication Forgery Detection", *Proceedings of IEEE International Conference on information technology*, Las Vegas, NV, 2012
[2] J. Fridrich, D. Soukal, J. Lukas, "Detection of copy-move forgery in digital images," in *Proceedings of Digital Forensic Research Workshop*, Cleveland, OH, August, 2003.
[3] A.C. Popescu and H. Farid, "Exposing digital forgeries by detecting duplicated image regions", Department of Computer Science, Dartmouth College, Tech. Rep. TR2004-515, 2004.
[4] W. Luo, J. Huang, G. Qiu, "Robust detection of region duplication forgery in digital images," in *International Conference on Pattern Recognition*, vol. 4, Aug. 2006, pp. 746–749
[5] B. Mahdian, S. Saic, "Detection of copy-move forgery using a method based on blur moment invariants," *Forensic Science International*, vol. 171, no. 2, pp. 180–189, Dec. 2007.
[6] J. Zhao, J. Guo, Passive forensics for copy–move image forgery using a method based on DCT and SVD, *Forensic Science International,* vol. 233, pp.158–166, 2013
[7] J. Zhang, Z. Feng, and Y. Su, "A new approach for detecting copy move forgery in digital images," *in International Conference on Communication Systems*, Nov. 2008, pp. 362–366.
[8] H. Huang, W. Guo, and Y. Zhang, "Detection of Copy-Move Forgery in Digital Images Using SIFT Algorithm," in *Pacific-Asia Workshop on Computational Intelligence and Industrial Application*, vol. 2, Dec. 2008, pp.272–276.
[9] I. Amerini, L. Ballan, R. Caldelli, A. Del Bimbo, G. Serra, "A SIFT-based forensic method for copy–move attack detection and transformation recovery," *IEEE Transactions on Information Forensics and Security,* vol. 6, no. 3, pp. 1099–1110, Sep. 2011.
[10] N. Dalal, B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2005, pp. 886–893.