Detection of Digital Image Forgeries for Forensics Applications

Lorin Enico Arthi. S[1], Student, E. A. Mohamed Ali[2], Associate Professor

Abstract —In this modern world, digital computing plays an important role in representing information as digital images. Due to the considerable improvement in imaging technologies the convenience, cleverness and transmission of digital images is easy today. However, the image editing technology is also being used for manipulating digital images and creating forgeries that are difficult to distinguish from authenticated photographs. Copy move forgery is a type of forgery in digital image forensics. A Gaussian filtering technique along with EM algorithm is proposed in this paper to detect the copy move forged area and its performance is compared with the block based method and keypoint based method. The parameters like precision and recall values are evaluated and plotted.

Keywords— Image forgery, image splicing, copy move image forgery, Keypoint, DCT.

I. Introduction

Today, the digital images are widely used by the people around the world. The digital image contains large amount of minute information that can be used as a witness in courts and evidence in news media. They are also used as a medium of communication since it can be easily understand by everyone. Many serious decisions are made in medical field based on the digital images. At the same time, each should pay a special attention to the field of digital image authenticity.

[1]S.LorinEnicoArthi, Author is currentlypersuing M.E (Applied Electronics) in National college of Engineering, Tirunelveli, Tamilnadu, India. (e-mail: geo.arthi@gmail.com).

[2]E. A. Mohamed Ali is currently working as an Associate Professor (Department of ECE) in National College of Engineering, Tirunelveli, Tamilnadu, India. (e-mail: ea_mdali@yahoo.com).

Nowadays the advancement in image editing and processing tools, cost and easy computer-human interface have become a great advantage to edit and manipulate the digital images even for ordinary users. It is feasible to change the content of the image and create forgeries, which are unable to differentiate by human naked eyes. Especially copy move forgeries are very difficult to identify.

Copy move image forgery is a type of digital image tampering, in which a part of the image is copied and pasted into another part of the same image to hide some important components of the image. This is also usually done in order to mask certain details or to duplicate certain aspect on an image. Recently, there are many different types of image forgery techniques are available. These techniques are based on pixelspresent in the image and some image components like brightness, hue, saturation, illumination and noises.

The goal of this paper is to compare the three types of image forgery techniques. The block based method, keypoint based method and Gaussian filtering methods are compared to predict the effective forgery detection algorithm. The block based method uses Discrete Cosine Transform (DCT) and the keypoint based method uses Scale Invariant Features Transform (SIFT) from which the feature are further extracted.

The proposed method uses Gaussian filter to estimate some of the components along with EM algorithm to detect the exact copy move forged area. The Gaussian filter estimates the Gaussian mixture model which contains the combination of several Gaussian components like mean and variance. Then these components are used for the extraction of features from the filtered image. Using EM (Estimation Maximization) algorithm the features are converted into a probability map that indicates the copy move forgery.

II. Review of Literatures

This section introduces the techniques and methods already exist in the field of digital image forgery detection. Several methods have been developed to detect copy-move forgeries. In [1], Fridrich first described about the in-depth search. The applicability of this method is limited mainly

because of its exponential complexity and it fails if there is any distortion. Resembling method is proposed by Popescu in [2]. In this method the features are extracted from the PCA (Principal Component Analysis). The usage of feature vectors was reduced by half in [3] and therefore the efficiency was improved. The drawback of this method is it has some difficulty in detection of rotated forged area. In [4] WeiqiLuo proposed a technique that has various post region duplication processing, blurring, noise and lossy compression steps. A DCT based image forgery detection technique is explained in the existing system of this paper.

In [5] a dyadic wavelet transform method was proposed and it has some drawbacks. It works on the images with simple background. XiaoBing Kang and ShengMin Wei [6] proposed a forgery detection method based on singular value decomposition. In [7] the block coefficients are evaluated using blur invariants. In [8], the authors propose a technique to detect cloning when the cloning is done using specific tools like Adobe Photoshophealing brush and Poisson cloning. The above methods do not provide the correct size and place of the detected region, but only displays the corresponding keypoints. The keypoint based method described in existing system detects the forged area by drawing lines in forge image.

To overcome the above mentioned drawbacks a new method is proposed in this paper. The proposed forgery detection algorithm is detection of forged area using Gaussian filter along with EM algorithm. The time required to run the proposed method is less than the block based method and key point based method.

III. The Existing Method

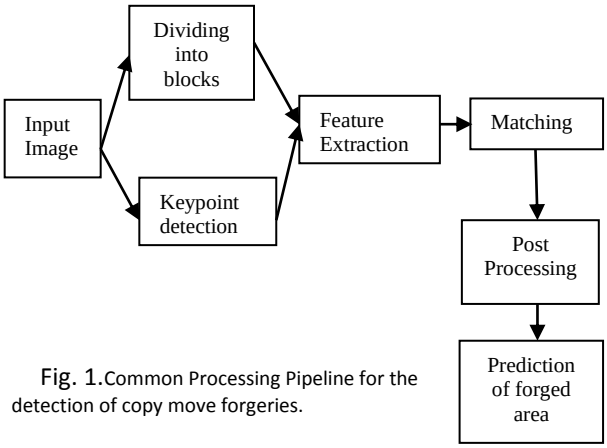

Fig. 1.Common Processing Pipeline for the detection of copy move forgeries.

The existing system consists of two approaches. They are

- ✓ Block based forgery detection and
- ✓ Keypoint based forgery detection.

The block diagram of joint copy move forgery detection method is shown in fig.1. The feature vectors are extracted from block based method and keypoint based method.

The joint copy move forgery detection algorithm steps are given below.

Given an M x N image, the detected regions are computed as follows:

Pre-processing:

1) The colour images are converted into gray scale images.

Extraction of feature vectors:

a) For block based method:
   - ✓ Divide the image in B_i overlapping blocks of size b x b, where $0 \le i < ((M-b+1).(N-b+1))$.
   - ✓ Compute a feature vector $\vec{f_t}$ for every block B_i.

b) For keypoint based method:
   - ✓ Scan the image for keypoints.
   - ✓ Compute the feature vector$\vec{f_t}$ using all keypoints .SIFT algorithm is used to find the keypoint and to extract the feature vector.

Matching

2) Match every feature vector by searching its approximate nearest neighbour. Let F_{ij} be a matched pair consisting of features $\vec{f_t}$ and $\vec{f_j}$, where *i, j* denote feature indices, and $i \neq j$. Let $c(\vec{f_t})$ denote the image coordinates of the block or keypoint from which $\vec{f_t}$ was extracted. Then $\overrightarrow{v_{tj}}$ denotes the translational difference ("shift vector") between positions c $(\vec{f_t})$ and c $(\vec{f_j})$.
3) Remove pairs $F_{ij}$where $|| \overrightarrow{v_{tj}} ||_2 < \tau_1$, where $|| . ||$ denotes the Euclidean norm.

Post-processing

4) Clustering of the remaining matches that remain to a joint pattern.
   - ✓ For block based method: let H(A) be the number of pairs satisfying the same affine transformation A. Remove all matched pairs where $H(A) < \tau_2$.
   - ✓ For keypoint based method: apply homography based clustering.

Prediction of forged area

5) If an image contains linked regions of more than τ_3 linked pixels, it is denoted as tampered.

It is quite common to set the thresholds τ_2 and τ_3 to the same value.

IV. Proposed Method

The block diagram of proposed system is shown below.

A. *CFA Interpolation*

Normally digital image forgeries will not leave any visual clues that alter the underlying statistics of an image. Digital colour images consist of three channels containing samples from different bands of colour spectrum. Almost digital cameras are equipped with a single colour sensor and it uses Colour Filter Array (CFA). The other two colour channels are estimated from the neighbours to obtain the three colour channels. The estimation of three colour channels is known as CFA Interpolation. The subset present in each channels are periodically correlated because all the three colour channels are arranged in periodic pattern. Here Bayes filter is used for CFA interpolation. The colour array sensor is shown in fig. 3.

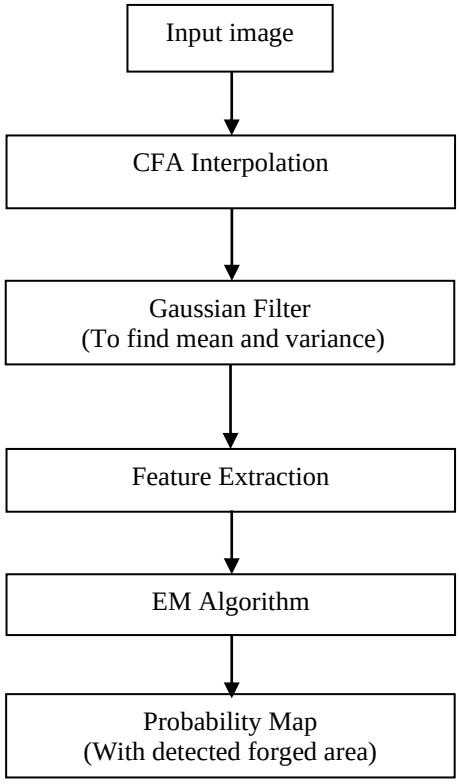

Fig. 2. Block diagram of Proposed System

B. *CFA Interpolation*

Normally digital image forgeries will not leave any visual clues that alter the underlying statistics of an image. Digital colour images consist of three channels containing samples from different bands of colour spectrum. Almost digital cameras are equipped with a single colour sensor and it uses Colour Filter Array (CFA). The other two colour channels are estimated from the neighbours to obtain the three colour channels. The estimation of three colour channels is known as CFA Interpolation. The subset present in each channels are periodically correlated because all the three colour channels are arranged in periodic pattern. Here Bayes filter is used for CFA interpolation. The colour array sensor is shown in fig. 3.

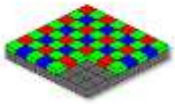

Fig. 3. Colour Filter Array sensor.

C. *Gaussian Filters*

Gaussian filters are used to estimate the Gaussian Mixture Models. The parameters include the estimation of mean and variance. The parameter chosen in this paper is estimation of variance map. Gaussian mixture model contains the combination of several Gaussian components. The Gaussian filter is also used to blur images and to remove noise. The expression for Gaussian filter and variance is given below.

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-x^2}{2\sigma^2}} \quad \text{..... (1)}$$

$$var(x) = \frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N-1} \quad \text{..... (2)}$$

D. *Feature Extraction*

The variance map of the acquired pixels and interpolated pixels are chosen as features. The variance map separates the acquired and interpolated pixels to estimate the local variance of these pixels.

E. *EM Algorithm*

EM algorithm is two step iterative algorithm that performs 'Expectation' and 'Maximization' operations to calculate the likelihood among features. The log-likelihood filter that works based on an EM algorithm is used to find the relation between features extracted in previous step. The likelihood function is given as

$$p(Model|Data) = \frac{p(Data|Model)p(Model)}{p(Data)} \text{..... (3)}$$

The expression for log-likelihood filter is given as.

$$l(\theta) = \log p(D|\theta) = \log \sum_H p(D, H|\theta) \text{..... (4)}$$

D – Set of n observed features.
H – Set of n values of hidden variable z.
Z(i) corresponds to x(i).

F. *Probability Map*

The probability map is the plot that shows the detected output of our forged image. The output image will be shown in experimental results.

V. Experimental Results

The proposed approach has been evaluated using datasets containing different types of tampered images. The original image and the output of the block based, keypoint based and

the proposed Gaussian filter with EM algorithm is shown below.

Fig. 4.Original image for keypoint method.

Fig. 5.Forged image for keypoint method.

Fig. 6.Output image for keypoint method.

Fig. 7. Original image for block based method.

Fig. 8. Forged image for block based method.

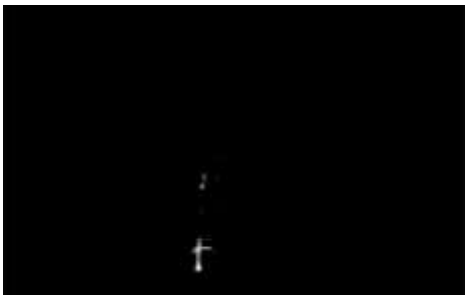

Fig. 9. Output image for block based method.

Fig. 10.Original image for proposed method.

Fig. 11.Forged image for proposed system.

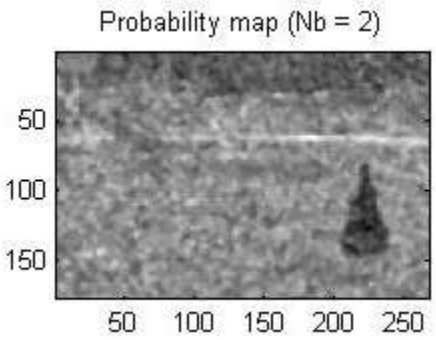

Fig. 12.Output image for proposed method.

G. *Performance Analysis*

The performance of the three forgery detection algorithm can be evaluated by calculating the precision 'p' and recall 'r' values. Let Tp be the number of correctly detected forged images, Fp be the number of images that has been incorrectly detected as forged and Fn be the falsely missed forged images. From these parameters the precision and recall values are computed.

$$precisionp = \frac{Tp}{Tp+Fp} \quad \text{..... (5)}$$

$$recallr = \frac{Tp}{Tp+Fn} \quad \text{..... (6)}$$

Precision is said to be the probability that a detected forgery is truly a forgery and recall denotes the probability that a forged image is detected. Recall is also known as true positive. Measure that combines the precision and recall is called as harmonic mean 'F_1' of precision and recall. It is also known as the traditional F-measure or balanced F-score.

$$F_1 = \frac{2.p.r}{p+r} \quad \text{..... (7)}$$

The tabulation and bar chart for the evaluation of three forgery techniques is given below.

Method	Precision (in percentage)	Recall (in percentage)	F1 (in percentage)
Gaussian Filter with CFA	88.89	89.24	89.06
Block Based Method	77.77	87.5	82.34
Keypoint Based Method	75	33.33	46.15

Table. 1. Comparison of three detection algorithm

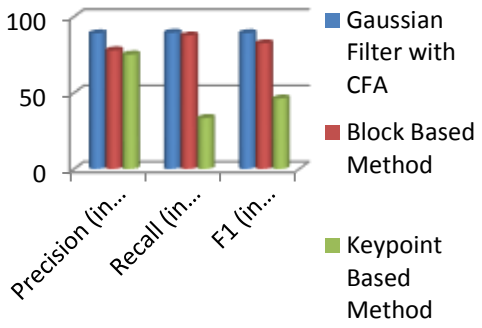

Fig. 13. Comparison chart for three detection algorithm

VI. Conclusion

The performance of different widely-used features for copy move image forgery is evaluated. The performance of Gaussian filter with EM algorithm produce better result when compared with the block based method and keypoint based method. Also the time taken by the proposed method is very less than other two methods. The block based method uses DCT features and the keypoint method uses SIFT features whereas, the proposed method extracts features from CFA interpolated Gaussian filtered image. And the precision and recall values for proposed method are more than the other methods. Thus the harmonic mean of proposed algorithm is better than the block based and keypoint based method.

REFERENCES

[1]. J. Fridrich, D. Soukal, J. Lukas, *"Detection of copy-move forgery in digital images",Proc. of the digital forensic research workshop*, 2003.

[2]. A.C. Popescu, H. Farid, "Exposing *digital forgeries by detecting duplicated image regions*", *Proc. Of Technical report TR2004-515*, Dartmouth College, 2004.

[3]. W. Q. Luo, J. W. Huang, G. P. Qiu, *"Robust detection of region-duplication forgery in digital image", Chinese Journal of Computers*, vol. 30, no. 11, pp. 1998-2007, 2007.

[4]. Y. P. Huang, W. Lu, W. Sun, D. Y. Long, *"Improved DCT-based detection of copy-move forgery in images",Journal of Forensic Science International*, vol. 206, no.1-3, pp. 178-184, 2011

[5]. M. G. M. Najah, H. Muhammad and B. George, "*Copy-move forgery detection using dyadic wavelet transform," Eighth International Conference Computer Graphics, Imaging and Visualization*, pp. 103–108, 2011.

[6]. S. W. X. Kang, "*Identifying tampered regions using singular value decomposition in digital image forensics*," International Conference on Computer Science and Software Engineering, 2008.

[7]. B. Mahdian and S. Saic, "*Detection of copy-move forgery using a method based on blur moment invariants," Forensic Sci. Int.*, vol. 171, no. 2–3, pp. 180–189, 2007.

[8]. B.Dybala, B. Jennings, andD. Letscher, "Detecting *filtered cloning in digital images," in Proc. ACM Int.WorkshoponMultimedia& Security (MM&Sec),* New York, 2007.