ISSN:0975-9646

Correlation Based Image Tampering Detection

Priya Singh
M. Tech. Scholar
CSE Dept. MIET
Meerut, India

Ms. Shalini Sharma Goel
Assistant Professor
CSE Dept. MIET
Meerut, India

Abstract-The current era of digitization has made it easy to manipulate the contents of an image. Easy availability of image processing tools on the internet allows modification to any image with no difficulty. Image format can be changed easily from one format to another and even the altering in image can be performed pixel by pixel transforming it to greater extends. This scenario has left the digital images prone to great threats and the validity of image is beyond the trust. To regain the trust in the reality of digital images has become a greater challenge in this digital world. Prior to this digital era, detection of the altered photographs was easy as there were no specific tools to change the images to such greater extends. But now with the arrival of latest software in the field of photo editing like Corel PaintShop Pro X7, Picasa, Adobe Photoshop Lightroom 5, Adobe Photoshop CC, etc. tampering of photographs, image forgery can be carried out without any noticeable sign of changes in the image. Even the authentic parts of the image cannot be found easily and it becomes difficult to expose the forgery. As the dependency on the digital images has increased now and various information exchanges occurs over internet, it has become necessary to keep the digital images safe and keep a check on their authenticity. Considering a tampered image a real image can cause various issues. An image can be tampered by hiding some information into its contents, by summing it with some templates or by other means, there can be any possibility. However, the consistency of the image is lost during the process of tampering. This paper identifies an active approach of forgery detection in the copy move image forgeries. The image is subdivided into smaller fixed size patches overlapping each other and then tampering areas are identified. This paper discusses the detection of tampering through correlation method to find out the tampered parts in the image.

General Terms-Image Forgery, Image Tampering, Copy Move Forgery, Active Approach, Correlation Coefficient, Mask/Block, False Accept, False Reject.

1. INTRODUCTION

As the world today has moved to a new digital era in which manipulating the image and adding or removing any element from it may result to greater number of forgeries, there is a great need to develop methods that identify such forgeries. The use of manipulation tools available over internet made it easy to tamper any image. This makes the verification of the image more challenging. Techniques such as cropping, filtering, blurring, scaling, resampling, rotation, etc. are some examples of image manipulation techniques [1].

Image tampering detection is required to prevent image forgery and protect the copyrights in various fields like media, glamour, forensics, military, etc. For detecting the tampered images it is necessary that possible correlations are identified which have changed due to the process of tampering. It is a rising research to detect forgeries in digital images.

Majorly there are three types of image forgeries; copy move forgery, retouching and image compositing. In copy move forgery a part of the same image is copied and pasted on the image to some other location [2]. It is one of the difficult types of forgeries to be detected because copying the same part of the image does not bring significant change in the attributes of the image as coping from any other image can bring. The next is image retouching which is used widely these days for various commercial purposes. In image retouching the image features are enhanced or reduce to bring attention towards the certain aspect of the image [2]. The third type of tampering is the image composting also called image splicing. Image splicing is the result of cutting and joining two or more different images to form a single composite image which looks like a single real image [2]. The process of composting images is carried out with seamless transition without leaving any traces or clues about the joining of the images.

Fig 1(a): Example of a Digital Image Tampering

The above figure shows the example of a copy move forgery. The above image is the tampered image and the bottom image is the original image. In the tampered image the truck is covered with the foliage on the left side of the truck. The tampering is done so flawlessly that there is no

suspicion of the presence of the truck in the image. The foliage on the left side of the truck was clipped from this image and further it was pasted over the truck to hide its presence in the original image.

2. Digital Image Tampering Detection Techniques

The tampering detection techniques for the digital images are broadly classified into two categories, active approaches and passive approaches [2]. In active approach we prepare the image at the time of capturing by some preprocessing like signature or watermarking so that it can be kept safe from being tampered [2]. Passive approach applies when there is no watermark or signature embedded into the original image. It involves the processes like statistical anomalies, measurement of attributes, compressions, correlations, etc. [2] to detect the parts tampered in the image.

2.1 Active Approach

Active approach is based on hiding data into the image at the source side. This means that secondary data like signature or watermark is embedded into the image while digitizing it at the source side like scanner. Further this secondary data is retrieved at the destination point for verifying the authenticity of the image. If the image is tampered then the secondary data cannot be retrieved at the destination and hence the forgery in the image can be identified. Active approach is based on two types of data retrievals, frequency domain data and spatial domain data.

2.2 Passive Approaches

Passive approach applies when there is no watermark or signature embedded into the original image. Although no visual clues of tampering is seen in a tampered image, but the attributes or the underlying statistics of that image changes which are the main area of focus in passive approach. Passive approach is a great challenge in identifying the image tampering and there is no particular method for all cases. There can be several methods to detect the tampering of special kinds. Passive approach is divided into five sub methods: pixel-based method, Formats-based method, Physically-based method, camera-based method, and Geometry- based method.

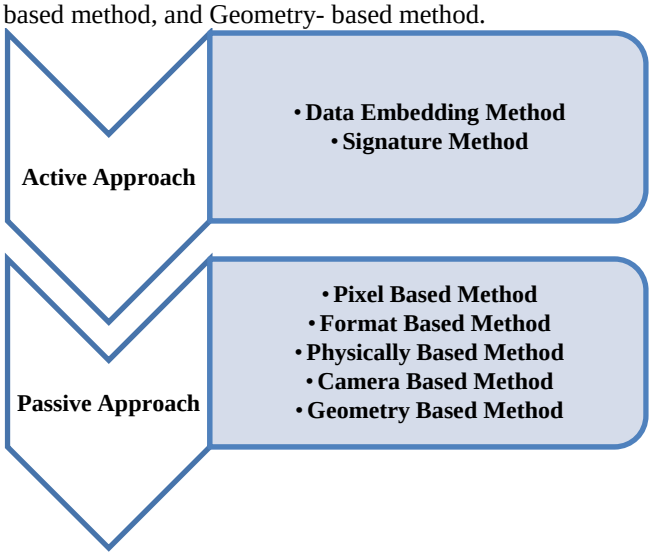

Fig 2(a): Active and Passive Approaches of identifying Digital image tampering

3. Related Work

In the field of digital image processing, a lot of work is done to detect the tampered images. There are three main techniques to create a forged image but copy move is one of the easy and famous techniques. In copy move forgery, one of the portions of the image is copied and moved to another part in the same image. Lots of methods are there to detect these types of forgeries. Fridrich et al. [3], recommended a technique to identify copy-move tampering, it works on analyzing the image to each and every cyclic shifted version. Due to the high complexity, it needs $(mn)^2$ steps to execute an image of size M×N. Due to the high complexity, it become typical to implement. In an approach proposed by Popescu and Farid [4], two methods are given; first algorithm works effectively for copy-move tampering to detect the copied part (copied without any changes) at different region in same image. Second algorithm fails to detect very tiny copied part and it can't handle rotated images. Ashima Gupta et al. [5] proposed the technique to detect the region duplication with the help of Discrete Cosine Transform (DCT). In the technique of DCT, the forgery is detected by dividing the image in the overlapping blocks and the duplicated blocks are identified. But it fails in small copied area to detect forged blocks. Fan et al. [6] has developed a tampering detection techniques based on 2D lightening coefficients. Further 3D lightening coefficients were involved to advocate the intermediately result and identify the forgeries. The forgery detection approach using 3D lighting system is given by Fan et al. [6], based on the shape by shading. It's a hopeful technique in detecting the forgery through 3D lighting system but problem with it is assessment of 2D figures of object leftover. Auto regressive coefficient as element vector and artificial neural network (ANN) classifier method is developed by the Gopi et al. [7] to detect image tampering. In it, 300 attributes vectors were used (form different images) to train an ANN. Another 300 attributes vector used to test an ANN.

The process of detecting a copy move forgery is similar to the process of feature extraction. Other methods are also used and are currently worked on reducing dimensionality [8], moments [9], color properties [10], region duplication [11], discrete wavelet transformation [12] and frequency domain transform [13].

4. Proposed Work

The method proposed here is an active approach to identify the copy move tampering in the images. This method was used to detect the tampering in the BMP images by partitioning the image into overlapped patches and then testing the correlation coefficients of the forged area by comparing them with the correlation coefficients of the original image. The efficiency of this algorithm at realistic forgeries has been computed for different mask sizes and the time consumed by each mask in identifying the tampering in an image was also calculated and is discussed in this paper.

4.1 Correlation Method

Correlation method is used as a statistical tool to establish the association between two variables. The 2-D correlation is defined as follows:

$$r = \frac{\sum_m \sum_n \left(A_{mn} - \overline{A}\right)\left(B_{mn} - \overline{B}\right)}{\sqrt{\left(\sum_m \sum_n \left(A_{mn} - \overline{A}\right)^2\right)\left(\sum_m \sum_n \left(B_{mn} - \overline{B}\right)^2\right)}}$$

Here, the value of 'r' ranges from -1 to 1 as $-1 \leq r \leq 1$. A & B represents the 2D data sets while $\overline{A}$ & $\overline{B}$ are means of sets A and B respectively. Further the size of A and B is M×N. Also m= 1,2,3,4,……., M and n= 1,2,3,4,……..,N.

The dimensions of the original image are M×N and it is further partitioned into the smaller overlapping blocks of dimension m×n. This makes the total number of blocks to be (M – m + 1) × (N – n + 1). After partition the image into blocks, the correlation coefficients are calculated through above given formula between the adjacent overlapping blocks. This experiment is done at source side (on original image) and then same formula is applied on the destination side (on forged image). There is a threshold value 0.025 to establish the forgery level between two images. All adjacent blocks are traced to calculate the values of correlation coefficient for both images (original image and forged image) and the difference of value of corresponding correlation coefficients from original and forged images are taken. If calculated correlation coefficient is greater than the threshold value 0.025, then there is forgery in an image.

In case of working with two 1D data sets, the 1D correlation may be defined as follows:

$$r = \frac{\sum_i \left(Ai - \overline{A}\right)\left(Bi - \overline{B}\right)}{\sqrt{\sum_i \left(Ai - \overline{A}\right)^2 \sum_i \left(Bi - \overline{B}\right)^2}}$$

The correlation calculation for two 1D data sets can be found by putting the values of these data set in the above given formula. The value of r should vary within the range -1 and 1. If value of r is greater or less than the given interval then there is no correlation between them.

5. EXPERIMENTAL RESULTS

For this research an odd mask is taken of block size of an odd number for this method. The odd masks used are 3×3, 5×5, 7×7, 11×11, 13×13 and 15×15 and the respective output images are generated showing the tampered parts of the digital images. The reason behind taking an odd mask is to easily achieve the value of the central pixel which cannot be obtained through an even mask.

Fig 5(a): Original image 242 x 362

Fig 5(b): Forged Image 242 x 362

Fig 5(c): Result with 3 x 3 mask

Fig 5(d): Result with 5 x 5 mask

Fig 5(e): Result with 7 x 7 mask

Fig 5(f): Result with 9 x 9 mask

Fig 5(g): Result with 11 x 11 mask

Fig 5(h): Result with 13 x 13 mask

Fig 5(i): Result with 15 x 15 mask

It is observed that as the mask size is increasing the fault accepts are also increasing significantly and fault reject decreases to minor extends.

5.1 Database Preparation

To accomplish the research work, a database of digital images is required. The database should be of high quality and scalable images. Thus, a collection of stable scene images was gathered to work upon. BMP image format was used and preferred because it is the simplest image format that directly stores the intensity at each pixel in the image. It does not require compression technique. The database was collected and it covered mostly greenery and landscape images.

The database consists of:

- A set of fifty original color images and fifty corresponding forged images with identical dimensions of *1600* x *1600*. All images belong to the class uint8.
- Another set of *50* x *3* original images and *50* x *3* corresponding forged images with identical dimensions of *242* x *362*.

Major work was done on MATLAB; some work was done on MS-paint and trail version of Adobe Photoshop cs2. More than 2000 digital images with different zooming using Nikon 16MP camera were clicked and it took around 6 months in collecting all data. Removal of noisy and poor images in terms of visibility of objects was carried out. For training samples the images were forged with MATLAB 8.1.0. For the sake of simplicity the forgery shape is taken to be square and only one region is forged to create forged images.

5.2 Database Pre-Processing

In this step all the color images were firstly converted to color bmp images and then all the images were converted from color to grayscale images using the following formula (MATLAB uses this formula to convert color image into corresponding gray image):

$$Grayscale=0.2989 * R + 0.5870 * G + 0.1140 * B$$

Here R, G, B implies Red, Green, Blue component of corresponding color image. After gathering the data of gray scale bmp images, the images the proposed methodology was applied to the images for detecting the copy move forgeries.

5.3 Experiment Configuration

The experiment was performed on a Core to Duo (32-bit) machine with 2.1 GHz processor speed using 2 GB of DDR2 RAM. MATLAB 8.1.0 was used to run the research and perform coding of the algorithm. The image extension was taken as 'bmp'. All images are in color (RGB) and also converted into grayscale images. Image resolution (Dimension) is 1600×1600, 242 ×242. Tampered Shape Square used has the dimensions 100×100, 50×50 pixels. Here one can take a 50(1600×1600) + 50(242×242) original images and 50(1600×1600) + 50(242×242) Tampered images. Number of forged region in image one. Camera used to take the pictures is Nikon 16MP camera.

5.4 Result Analysis

The average false reject and false accept for zero zoom, 2x zoom and 4x zoom were calculated for each of the mask/block size and the efficiency of the algorithm for each mask/block size was calculated. The Average false reject and false accept were calculated as follows:

TABLE 1. Comparison of Average False reject and False Accept for different mask/block sizes (using Correlation coefficient)

Mask/ Block Size	Projected Average false reject (0X+2X+4X)/3	Projected Average false accept (0X+2X+4X)/3
3 x 3	162.286163	163.897800
5 x 5	90.097484	236.407250
7 x 7	65.732704	336.515750
9 x 9	49.449685	446.624200
11 x 11	38.496855	563.548700
13 x 13	30.355345	685.287700
15 x 15	23.377358	813.040850

TABLE 2. Time analysis for each mask/block size (each block =50 sets)

Mask/Block Size	Average Time Taken by Each Mask to find out the Tampered Region in the Images
3 x 3	25.770668 seconds
5 x 5	25.749802 seconds
7 x 7	25.631235 seconds
9 x 9	25.381463 seconds
11 x 11	24.873647 seconds
13 x 13	25.117938 seconds
15 x 15	24.751047 seconds

5.5 Algorithm Efficiency

Observing the table of Time analysis, it can be concluded that the time difference among the block sizes is minor. The forgery detection in each block size is approximately same. It can be noticed that the mask of 3 x 3 takes the maximum time. This may be because the shorter the size of the mask is taken, more are the number of blocks to be checked, however when the mask size is larger than the number of blocks to be checked are less. Considering the other aspect of time it can be seen that it could be easier to find out the correlation coefficient of 3 x 3 mask as compared to the 15 x 15 mask because the increase in size will increase the time for calculation of the correlation coefficient. Hence, even though the mask of 3 x 3 is taking the maximum time, but other masks also take similar time with minor differences. So mask of 3 x 3 can be used to find out the forgery in images.

In Figure below it can be seen that as block size increases the false reject is decreases to some aspect, however there is a great increase in the false accept. Further, both the lines of average false accept and average false reject coincides for the block size 3 x 3 which shows that mask of 3 x 3 can be used more efficiently for this algorithm.

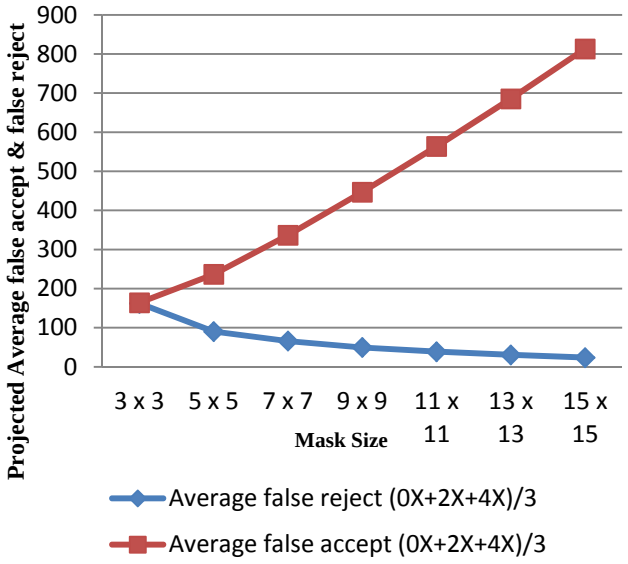

Fig 5(j): Graph shows average False reject and false accept per mask size for zero zoom, 2x zoom and 4x zoom

6. Conclusion

Digital image forgery has become a common technique and is amongst the top most forgeries carried out in the current era. This research work establishes what exactly the digital image forgery is. Some of the major approaches for digital image authentication and forgery detection are defined. The method described in this image is a robust approach to find out the forged part of an image. In this research work bmp images were used. Correlation method detects forgery with some false acceptances and some false rejections. The experiment results in improved detection rate in forgery and also improves the detection time of the Digital image forgery hit uncovering algorithm that is used. Future work is to mature the correlation method and to produce better result with more than one forged region in the image. With more than one and complex, irregular shapes of forged region like circle, ellipse, convex hull etc. and also improving the running time of proposed algorithm.

REFERENCES

[1] D. Sharma and P. Abrol, "Digital Image Tampering – A Threat to Security Management", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, Issue 10, pp. 4120-4123, October 2013 ISSN (Online): 2278-1021.

[2] S. K. Mankar and A. A. Gurjar, "Image Forgery Types and Their Detection: A Review", International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 5, Issue 4, pp. 174-178, April 2015 ISSN: 2277 128X.

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

[4] C. Popescu and H. Farid, "Exposing Digital Forgeries by Detecting Duplicated Image Regions," Technical Report, TR2004-515, Department of Computer Science, Dartmouth College, pp. 758-767, 2006.

[5] Ashima Gupta , Nisheeth Saxena , S.K Vasistha, "Detecting Copy move Forgery using DCT", International Journal of Scientific and Research Publications, Volume 3, Issue 5, May 2013 1 ISSN 2250-3153.

[6] W. Fan, K. Wang, F. Cayre and Z. Xiong, "3D Lighting-Based Image Forgery Detection Using Shape-From-Shading", 20th European Signal Processing Conference EUSIPCO, (2012), pp. 1777-1781.

[7] E. Gopi, N. Lakshmanan, T. Gokul, S. Ganesh and P. Shah, "Digital image forgery detection using artificial neural network and auto regressive coefficients", Proc. Canadian conference on electrical and computer engineering, (2006), pp. 194–7.

[8] X. Kang and S. Wei, "Identifying Tampered Regions Using Singular Value Decomposition in Digital Image Forensics," International Conference on Computer Science and Software Engineering, pp. 926-930, 2008.

[9] B. Mahdian and S. Saic, "Detection of copy-move forgery using a method based on blur moment invariants.," Elsevier Forensic Science International, vol. 171, no. 2-3, pp. 180-189 Sep. 2007.

[10] S.-jin Ryu, M.-jeong Lee, and H.-kyu Lee, "Detection of Copy-Rotate- Move Forgery Using Zernike Moments," IH, LNCS 6387, vol. 1, pp. 51-65, 2010.

[11] W. Luo, J. Huang, and G. Qiu, "Robust Detection of Region-Duplication Forgery in Digital Image," 18th International Conference on Pattern Recognition (ICPR'06), pp. 746-749, 2006.

[12] Kwang-Fu Li, Tung-Shou Chen and Seng-Cheng Wu, "Image tamper detection and recovery system based on discrete wavelet transformation," *Communications, Computers and signal Processing, 2001. PACRIM. 2001 IEEE Pacific Rim Conference on*, Victoria, BC, vol.1, pp. 164-167, 2001.

[13] A. Sharma and P. Singh, "A comparative study of frequency domain based approaches for image tamper detection." TENCON 2015 - 2015 IEEE Region 10 Conference, Macao, pp. 1-4., 2015. ISSN: 2159-3442.