Digital Image Forgery Detection using Correlation Coeficients

Chhaya Saini
M. Tech. Scholar
CSE Dept. KEC
Ghaziabad, India

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

Pramod Kr. Sethy
Assistant Professor
CSE Dept. KEC
Ghaziabad, India

Raj Kumar Saini
P.hd. Research Scholar
CSE Dept. IIT Roorkee
Roorkee, India

ABSTRACT
In digital era, it has become easy to modify any image. Due to this the trust and validation of it is going to lose. Now it has become major problem of digital world to regain the lost trust. The background behind the modification and any changes in an image is easy availability of software tools on internet. Images can be transformed from one image format to another and any part of image can be altered pixel by pixel. Before the digital age, it was literally easy to detect the altered photographs. But now with the advent in the commercial software like various image photo editing software like Adobe Photoshop, XnView; ProShow Gold etc. make image forgery simple, the tampering of the photographs have become very easy, can be carried out without any noticeable signs of tampering and it is becoming harder to expose and mark the authentic ones. With the increased dependency over the digital images for exchanging the information, the need to keep their authenticity increases and digital images also use as authenticated facts for an offence. If it will not contain the authenticity then a problem will arise. An image forgery is made either by summing some templates, or hiding some kind of information in an image, in which the consistency is lost. This paper identifies the key methods for detecting forgery in the digital images. To identify and detect the forged areas, the image is divided into overlapped patches of some fixed size. In our paper we will discuss the correlation method, that how it find outs the forged part in an image. Firstly, the digital image tampering process is discussed. After that, it shows that different algorithms have different approaches to detect the forgery.

Keywords
Image Forgery; Mean Vector Method, Correlation Coefficient, Templates.

1. INTRODUCTION
In the present world which has become digital computing world, the exchanging and representing the information in visual manner has become more essential. Due to great evolution in digital computation and networking technologies, the earlier period have showed a significant hike in the accessibility, and broadcasting of digital images using digital image processing software's. However, manipulation and forgeries are also created by these technologies in digital images and due to this it become difficult to tell between original image and forged image. Forgery of images contains pasting one part of an image onto another image, expertly manipulated to avoid any notion. Each image changes may be a forgery based upon the perspective in which it is used. The advanced and inexpensive software of digital era enable the manipulation of digital images with undetectable hints. On an image, manipulation includes these processes like rotation, scaling, brightness changes, contrast enhancement, blurring, etc. or any combinations of them. Now it has become more complex to establish image authenticity and this problem is harder to sort out due to the availability of digital images and free image editing tools


INTERNET

Spies turn to high-tech info ops

PCs, Internet used for manipulating images, public opinion

BY DANIEL VERTON

ALTERED IMAGE

Federal intelligence agencies are studying ways to use computers and the Internet, rather than just leaflets and radio broadcasts, to shape and disseminate information designed to sway public opinion in the world's hot spots.

As part of its so-called "perception management" program, the intelligence community has for decades created misinformation to trigger political change without direct political or military involvement in countries where the United States has vested interests, such as Iraq and North Korea.

Acting on congressional recommendations to bolster research and development in information technology, intelligence agencies are turning to PCs to develop more sophisticated means of manipulating and delivering digital photos, video clips and recorded sound to portray fictitious events in hopes of provoking desirable outcomes.

Altered Egos
The top photo, of a fictitious meeting between President Clinton and Saddam Hussein, was generated by digitally altering and combining elements from three photos, shown below the altered image. Intelligence agencies plan to use similar techniques to create images and then disseminate them via the Internet in their efforts to influence events in foreign countries, such as Iraq.

Fig 1: Example of a Digital Image Forgery

The Figure above shows a famous example of digital image forgery. In which a newspaper cutout shows, three different pictures were collected from different sources and merged together to create a forged image: Pictures of Saddam Hussein, The White House, and Bill Clinton. Here White House has been blurred to show a real effect a farther focus background. Then, the images of Mr. Bill Clinton and Mr. Saddam Hussein were clipped from two distinct pictures and imposed on the White House. The image is forged with realistic effects, it care the correct shadows of speaker stands with microphones.

2. IMAGE TAMPERING DETECTION TECHNIQUES
The forgery detection techniques are divided into two types, Active approaches and Passive approaches. These approaches identify the forged digital images. In the active methods, we add some data or signature into the original image to keep it safe from forgery, but in passive methods we don't have original image, so we can't insert any data and we should perform operations on some features of the image e.g. on correlations, compressions, statistical anomalies and measurements of attributes in the given image to detect forgery. The passive approaches are subdivided into five categories. These are camera-based, format-based, pixel-based, geometric-based and physical-based. Active approaches can be divided into two types, the embedding position spatial domain or frequency domain data.

2.1 Active Approaches

The Active approach considered to insert the data and signature into the original image at the resource side and make sure the image reliability at the detection side. Basically, Active approach have original image to embedding the data and signature into the image.

2.2 Passive Approaches

In Passive approaches, there is no availability of original images. It works in the absence of the data and signature. I use the attributes of an image to detect forgery in it. Passive approach is divided into five sub methods: pixel-based method, Formats-based method, Physically-based method, camera-based method, and Geometry- based method.

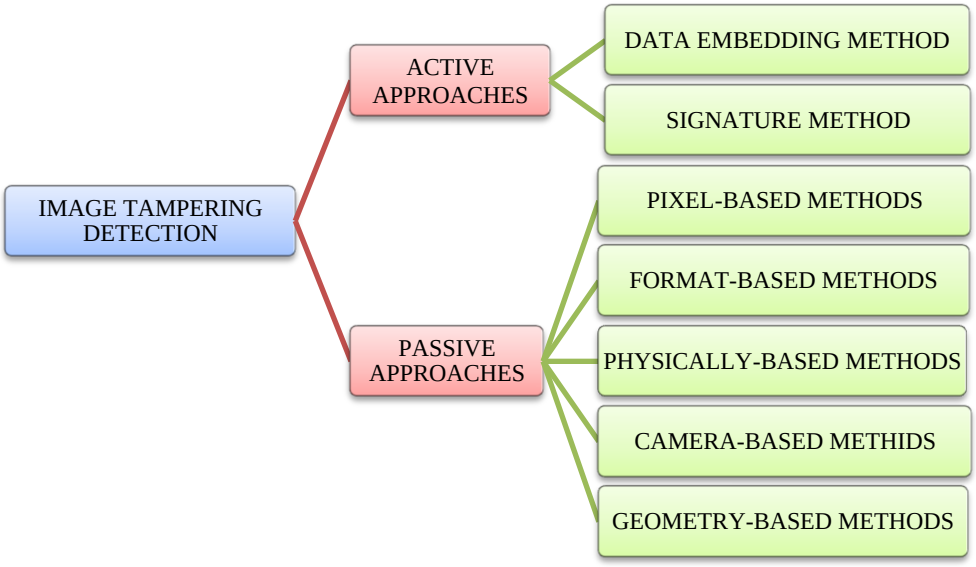

Figure 2(a): Digital image forgery detection approaches

3. RELATED WORK

In the field of digital image processing, a lot of work is done to detect the tampered images. There are so many techniques to create a forgery but copy move is one of easy and famous technique. 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. [1], 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. Ashima Gupta1 et al [3] 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. One more approaches Pradyuman Deshpandey in which there are two methods; here first algorithm works effectively for copy-move to detect the copied part (copied without any changes) at different region in same image. Second algorithm is failed to detect very tiny copied part and it can't handle rotated images.

Auto regressive coefficient as element vector and artificial neural network (ANN) classifier method is developed by the Gopi et al [5] 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 Hit percentage in this experiment to detect the forgery is 77.67% in which forged images were used to train ANN and 94.83% in experiment in which a database of forged images was used. The forgery detection approach using 3D lighting system is given by Fan et al. [4], 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.

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 [2], [6], moments [7], [8] color properties [9], frequency domain transform [1].

4. PROPOSED WORK

There are two methods that can spy forgery in BMP images and then place the forged elements. Our methods detects forged region by partitioning the image into overlapping patches and then test for the forged area. The efficiency and robustness of this technique at realistic forgeries level has been computed in this paper. Here, it can be verified that the algorithm is efficient for noisy images too.

4.1 Mean Vector Method

From the given image figure 4(a) first we calculate two mean vectors namely row mean vector and column mean vector that may be calculated as follows. Let image has M rows and N columns, and then the row mean vector is calculated as.

$$\text{Row Mean Vector} = \frac{1}{M}\sum_{i=1}^{M} ri$$

$$\text{Column Mean Vector} = \frac{1}{N}\sum_{j=1}^{N} Cj$$

Where r_i and c_j represent ith row and jth column in the image.

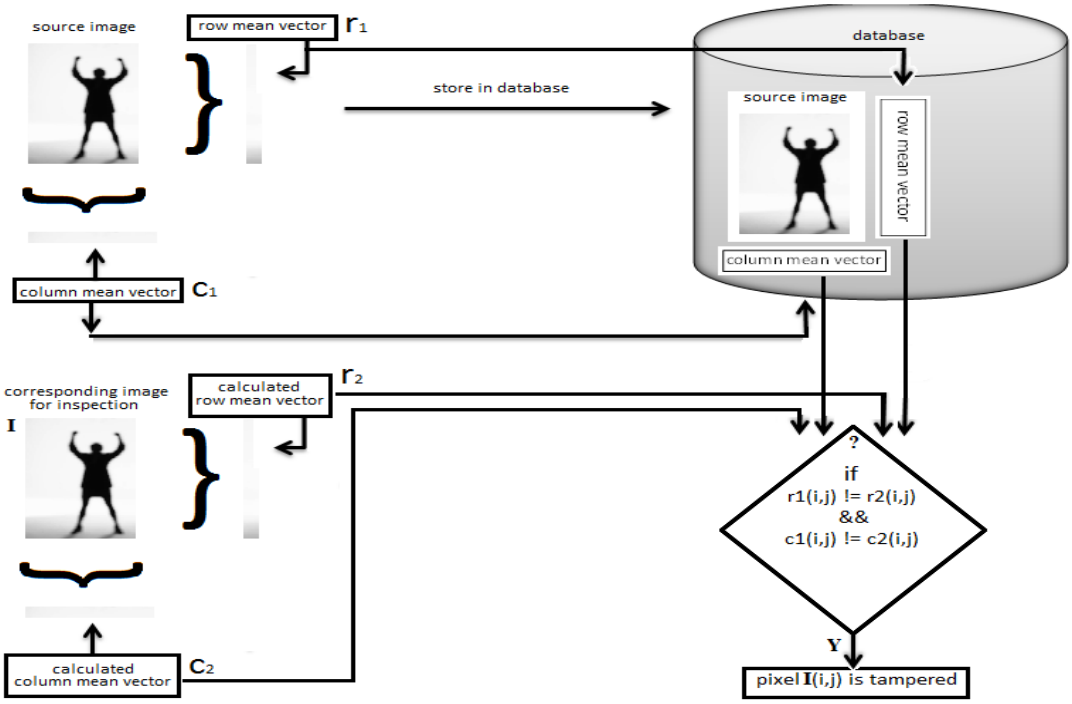

Fig 4(a): Mean Vector Testing

To identify the forgery, the Row and Column Mean Vector are calculated for both images, i.e., the Original image (r1, c1) and the Forged image (r2, c2). In case the image is forged, then these values for two images do not match up and then the pixel I (i, j) is tampered in the image.

Fig 4(b): Original image

Fig 4(c): Forged image

Fig 4(d): Result Image

In given example, Fig 4(b) is original image and Fig 4(c) is forged image in which some division of image is forged.

After comparing the row mean vector and column mean vector of both images it find out the tempered part in the image which showed in fig 4(d).

4.2 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:

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

1. Where, r varies within the range -1 and 1 i.e. $-1 \leq r \leq 1$
2. Where A and B are 2-D data set and,
3. $\overline{A}$, $\overline{B}$ are means of sets A and B respectively.
4. m=1,2,3,…,M and n=1,2,3,…, N.

5. M×N is size of A and B.

Here the whole original image having dimension $M * N$ is partitioned into the smaller overlapping blocks of dimension$m * n$. Thus the total no. of blocks will be$(M - m + 1) * (N - n + 1)$. After partition an image into blocks, the correlation coefficients were 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 take the difference of value of corresponding correlation coefficients from original and forged image.

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 1-D data set, the 1-D correlation may be defined as follows:

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

The correlation calculation for two 1-D data set can be find out by putting the values of these data set in the given above formula. The value of r should be varies 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

Here a mask or blocks of size 2×2, 4×4, 6×6, 8×8, 10×10, 12×12, 14×14 and 16×16 was taken and relative output images have been generated that are showing the sign of forgery in the digital images.

Fig 5(a): Original image

Fig 5(b): Forged Image

Fig 5(c): Result image (2×2)

Fig 5(d): Result image (4×4)

Fig 5(e): Result image (6×6)

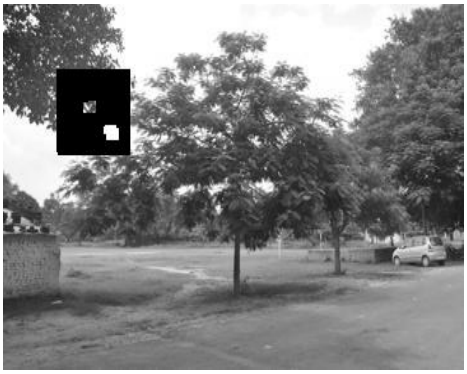

Fig 5(f): Result image (8×8)

Fig 5(g): Result image (10×10)

Fig 5(h): Result image (12×12)

Fig 5(i): Result image (14×14)

Fig 5(j): Result image (16×16)

In the observed results if the mask size is increased the fault accepts increases and fault reject decreases.

5.1 Data Base Preparation

To accomplish the research work, a database of digital images is required. For this a word processing software for documentation of work is required .The database should be highly quality and scalable. Thus, an arrangement of collected and stable scene images was gathered to work upon. BMP image format can be preferred because it is simplest image format that directly store 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. Major work was done on MATLAB, some work was done on MS-paint and trail version of Adobe Photoshop cs2.We clicked more than 2000 digital images with different zooming using Nikon 16MP camera. It took around 6 months in collecting all data. Removal of noisy and those images those were very poor in terms of visibility of objects in the images.

Brightness is raised just by adding a constant "K" to the image.

$$I' = I + K$$

Intensity normalization for making the dynamic range identical

$$I' = (I - a)\frac{(d - c)}{(b - a)} + c$$

Where (a,b) current dynamic is range and (c,d) is new dynamic range.

I and I' are current and new image.

5.2 Experiment Configuration

To execute the program, one needs a dual core (32-bit) machine with 3.2GHz processor speed using 1GB of DDR2 Ram. MATLAB 7.0.1 is used to run the research as MATLAB software is used for all Algorithm coding. Extension is Bmp. All images are in color (RGB) and also converted into grayscale, Binary. Image resolution (Dimension) is 1600×1600, 242 ×242.Tampered shape Square 100×100, 50×50 pixels. Here one can take a 50(1600×1600) +53(242×242) original images and 50(1600×1600) +53(242×242) Tampered images. Number of forged region in image one. Camera used to take the pictures is Nikon 16MP camera.

5.3 Result Analysis

TABLE 1. Comparison of Avg. False reject and False Accept with different block sizes (using Correlation coefficient)

Mask Size	2×2	4×4	6×6	8×8	10×10	12×12	14×14	16×16	Average
Average Reject per mask	223.801 9	118.3396	88.66038	68.12264	53.79245	43.45283	34.51887	26.53774	82.1533
Average accept per mask	164.3028	245.884	368.7748	506.166	655.0569	810.1698	971.5783	1142.328	608.0326

TABLE 2. Time analysis for each block / mask size (each block =53 sets)

Mask Size	2×2	4×4	6×6	8×8	10×10	12×12	14×14	16×16
Average Time per mask	20.5263 sec	20.8106 sec	19.3596 sec	15.6688 sec	20.3119 sec	15.6546 sec	15.6728 sec	14.8580 sec

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.

5.4 Algorithm Efficiency

In Fig 5(k), it can be seen that as block size increases the false reject is decreases and false accept is increases. Thus, large block size should be used to detect the forgery correctly.

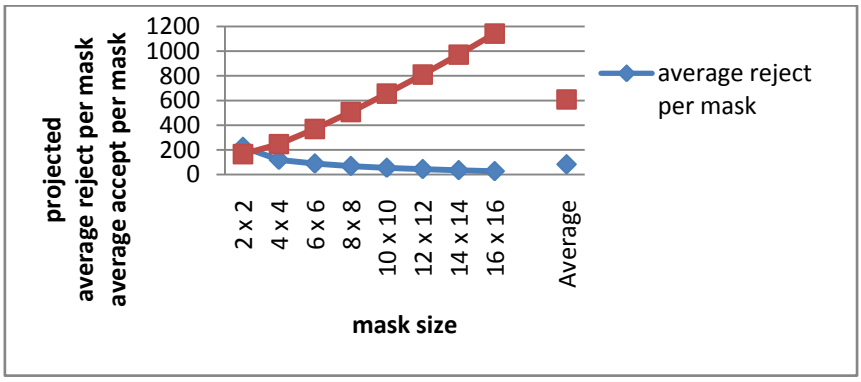

Fig 5(k): Graph shows False reject per mask size and false accept per mask size

This graph Fig 5(l) shows that the false reject is almost minor, near to zero for all sets (There were 53 sets in the experiment) and false accept is higher and same for all sets as per the blocks size.

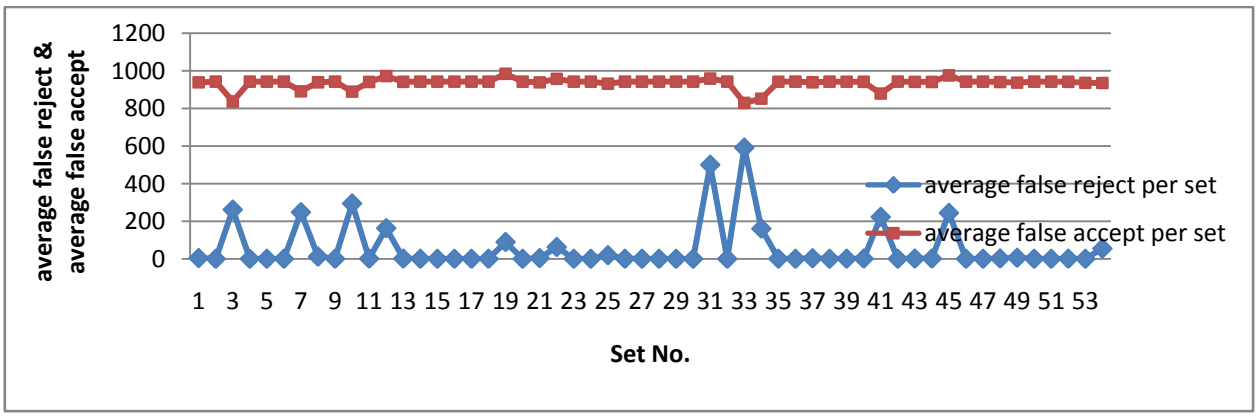

Fig 5(l): Graphs shows false reject and false accept per set number

6. CONCLUSION

Digital image forgery has become a common technique and is amongst the top most forgeries carried out in the current era. This report 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. Here, in this bmp images have been used for dissertation. Mean Vector method works well with grey as well as with binary images and detects the forged region. Some tests were performed on the algorithm over the images to check our results.Correlation method detects forgery with some false acceptances and some false rejections. The experiment results in improved detection rate in forgery and also improved the detection time of the Digital image forgery hit uncovering algorithm that is used. Future work is to mature the second 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., improving the running time of technique.

7. REFERENCES

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

[2] 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.

[3] 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

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.