Detection of Region Duplication in Digital Images: A Digital Forensic Approach

Jatin Wadhwa (111CS0165)

Talib Ahemad (111cs0511)

Department of Computer Science and Engineering
National Institute of Technology Rourkela
Rourkela-769 008, Odisha, India.

Detection of Region Duplication in Digital Images: A Digital Forensic Approach

Thesis submitted for the degree of

Bachelor of Technology

In

Computer Science and Engineering

By

Jatin Wadhwa (Roll: 111CS0165)

Talib Ahemad (Roll: 111CS0511)

Under the guidance of

Prof. Ruchira Naskar

Department of Computer Science and Engineering

National Institute of Technology Rourkela

Rourkela-769 008, Odisha, India.

Department of Computer Science and Engineering
National Institute of Technology Rourkela
Rourkela-769 008, Odisha, India.

May 08, 2015

Certificate

This is to certify that the work in the thesis entitled **Detection of Region Duplication in Digital Images: A Digital Forensic Approach** by **Jatin Wadhwa and Talib Ahemad** is a record of an original research work carried out under my supervision and guidance in partial fulfillment of the requirements for the award of the degree of **Bachelor of Technology** in **Computer Science and Engineering**. Neither this thesis nor any part of it has been submitted for any degree or academic award elsewhere.

Prof. Ruchira Naskar
Department of Computer Science and Engineering
NIT Rourkela - 769008

Acknowledgment

This thesis has been possible due to the help and endeavor of many people.

Foremost, we would like to express our gratitude towards our project advisor, Prof. Ruchira Naskar, whose mentor-ship has been paramount, not only in carrying out the research for this thesis, but also in developing long-term goals for our career. Her guidance has been unique and delightful. She provided her able guidance whenever we needed it. Yet she always inspired us to be an independent thinker, and to choose and work with independence.

We would also like to extend special thanks to my project review panel for their time and attention to detail. The constructive feedback received has been keenly instrumental in improvising our work further.

We would like to thank other researchers in our lab and our friends for their encouragement and understanding.

Jatin Wadhwa
Talib Ahemad
Computer Science and Engineering
National Institute of Technology Rourkela

Abstract

Digital images are easy to manipulate and forge due to availability of powerful image processing and editing software. Region duplication is becoming more and more popular in image manipulation where part of an image is pasted to another location to conceal undesirable objects or sometimes to hide some useful information.

This thesis presents a detailed study and evaluation of one specific type of digital image forgery detection, known as the copy-move forgery detection. Over the past decade considerable number of technologies and solutions that have been proposed for detection of copy-move forgeries. We will be looking about different techniques - on the basis of time complexity, false positives and false negatives - for the detection of this type of forgery. Firstly, detecting the forgery using different techniques – Autocorrelation, Exact Block matching and exhaustive search – and then comparing these techniques on the basis of different parameters. Some results are very promising and can contribute to ongoing work in this field.

Contents

List of Figures

Chapter 1

Introduction

Digital Forensics is a branch of forensic science related to cyber-crime. Encompassing the recovery and investigation of material found in digital devices, often in relation to computer crime. Digital Forensics has expanded to cover investigation of all devices capable of storing digital data. Digital forensics has several applications which includes forensic investigation of digital media devices, intellectual property theft detection and investigation, fraud detection, e-discovery of potential digital evidences and testifying those in courtroom, confirm alibis or statements, determine intent identify sources (e.g., in copyright cases) and authenticate documents.

Digital images and videos are used as the principle form of digital evidences, in the court of law as well as media and broadcast industries. With the advancement in technologies and the availability of powerful image processing tools, the reliability of digital images is often under question. Hence maintenance of their fidelity becomes crucial. ‘

Many types of Digital forgery are being found now a days. Like Image Splicing, Image Retouching and Copy move etc. Image splicing is the process of making a composite picture by cutting and joining two or more photographs. The spliced image may introduce a number of sharp transitions such as lines, edges and corners. Image Retouching refers to manipulation for photo restoration or enhancement (adjusting colors / contrast / white balance (i.e. gradational retouching), sharpness, noise, removing elements or visible flaws on skin or materials).

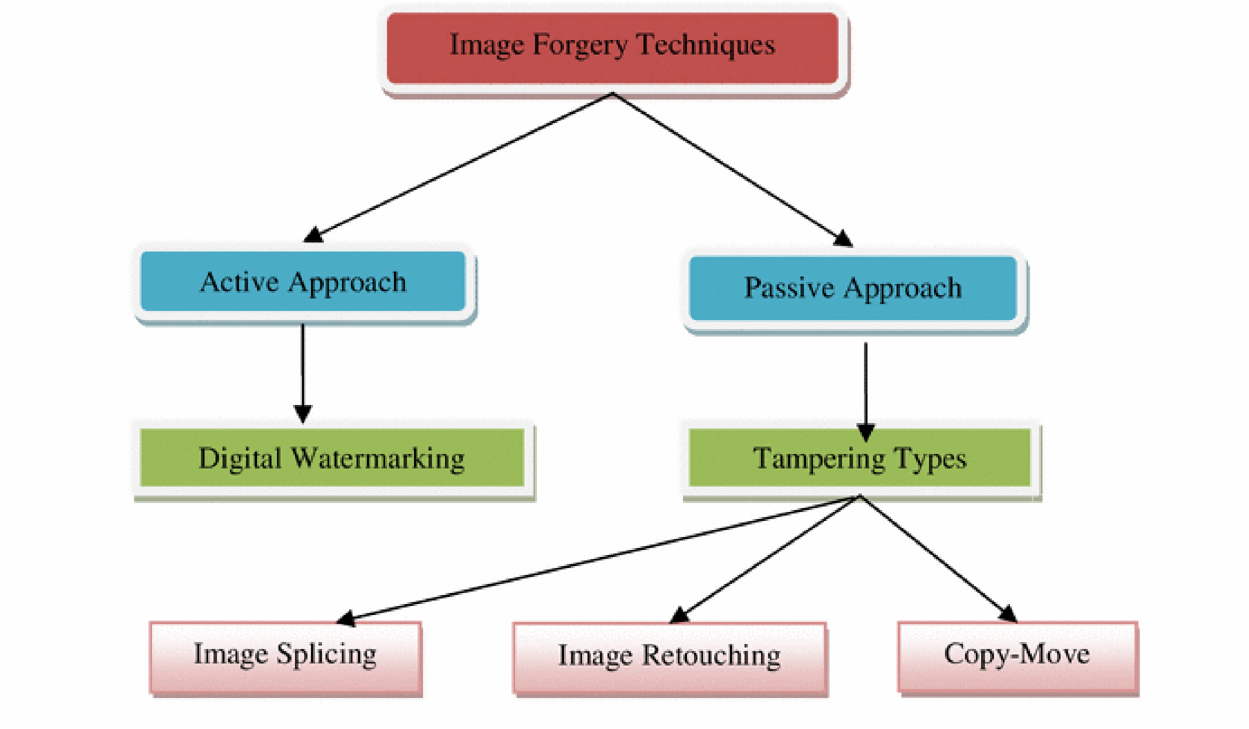
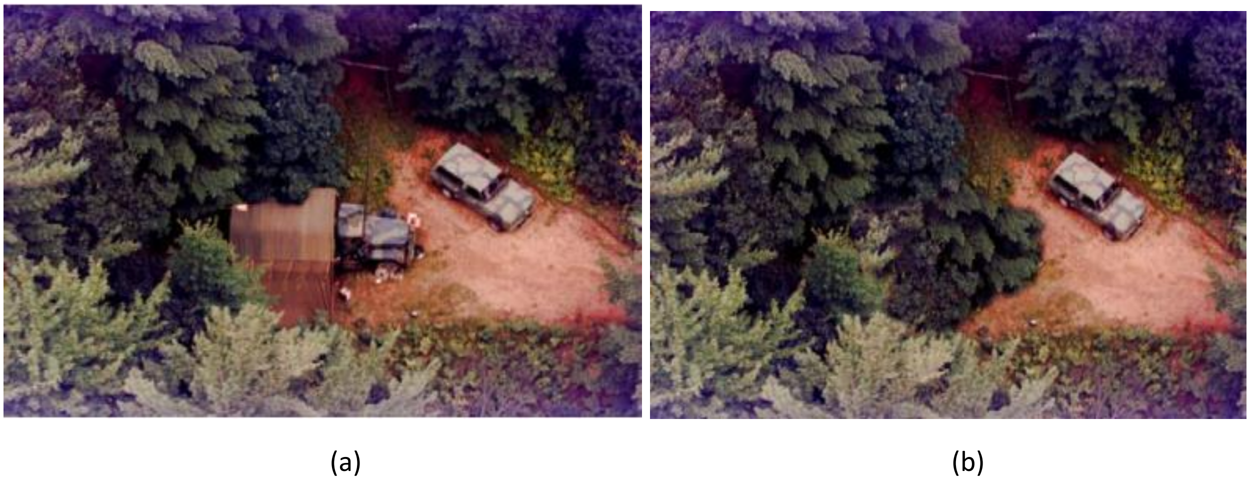


Figure 1.1: Classification of Image Forgery

Image reference: Kumar Sunil, Desai Jagan and Mukharjee Shaktidev, "DCT-PCA Based Method for Copy-Move Forgery Detection", Proceedings of the 48th Annual Convention of Computer Society of India – Vol II Advances in Intelligent System and Computing Volume249,2014,pp 577-583.

The most primitive form of cyber-attack on digital images is region Duplication. Aim of the attacker is to deceive the viewer with a manipulated image. This is also known as Copy-Move Forgery. In this, a part of the image is copied and pasted somewhere else in the image with the intent to obscure an important feature.



(a) (b)

Figure 1.2: Copy-Move Forgery (a) Original Image, (b) Forged image

Image reference: A.J. Fridrich, B.D. Soukal and A.J. Lukas, “Detection of copy-move forgery in digital images”, Proceedings of Digital Forensic Research Workshop, 2003.

Rest of the thesis has been organized in the following way: Chapter 2 presents a review of related research works. Chapter 3 presents different challenges in detecting a copy move forgery and techniques used to detect copy move forgery, along with a comparative performance analysis of all those techniques. Chapter 4 presents a parameterization procedure to evaluate the efficiencies of copy-move detection techniques, along with our simulation results. Finally we conclude our work in chapter 5.

Chapter 2

Review of Literature

In this chapter we first present the Detection of copy-move forgery in digital images proposed by Fridrich [1]. Having discussed the detection of copy-move forgery in digital images, we describe Exposing digital forgeries by detecting traces of re-sampling by Popescu [2]. Next we present Distinctive image features from scale-invariant key-points, by Lowe [3]. Finally the Detecting image duplication using SIFT features by Pan [4], has been illustrated.

One of the earliest digital forensic techniques for copy-move attack, proposed by Fridrich et. al. in [1], is based on the principle of matching of a region. The authors search for two image regions, having exactly identical pixel values. As we know that a standard image is too large and the complexity to compare pixel by pixel of the whole image will be far away from being practical. It works perfectly fine giving appropriate results for very small images.

However in [2], the authors further enhance the technique by dividing the image into fixed size blocks and sorting and then comparing block wise which will ultimately decrease the complexity but may produce some false positives.

In Popescu et. al. in [2], re-sampling (e.g., scaling or rotating) introduces specific statistical correlations, and describe how these correlations can be automatically detected in any portion of an image. This technique works in the absence of any digital watermark or signature. The attacker resizes the image by up-sampling or down-sampling the image signal. This kind of forgery is detected by searching for signal correlation among different image blocks. When a block is down-sampled by a factor of two, every alternate sample in the original block may be found in the re-sampled block. Alternatively, when a block is up-sampled by a factor of two, every sample in the original block may be found in the new block, along with alternate samples in the new block being linear combination of its neighbors. Re-sampling image blocks by any integer factor induces such periodic correlations among original and forged image blocks. Duplicate image regions may be detected by matching SIFT (Scale-invariant feature transform) [3] key points of the regions. Such forensic approaches to diagnose copy-move forgery having geometrically transformed image blocks, have been proposed in [4].

Chapter 3

Copy-Move Forgery Detection in Digital Images

A copy move forgery is created by copying and pasting the content within the same image, and potentially post processing it. In image region-duplication, the duplicated part(s) come from the same image. Hence its noise component, color palette, dynamic range, and most other important properties will be compatible with the rest of the image. Thus will not be detectable using methods that look for incompatibilities in statistical measures in different parts of the image. Feathered crop or retouch tools may be used to further mask any traces of the copied-and-moved segments. In recent years copy move forgery is emerged as hot research topic among the researchers of digital image forensics. A considerable number of different algorithms are proposed for detection of copy move forgery.

3.1 Detection based on Exhaustive Search:

In this method of copy-move forgery detection, the copy-moved segments are matched by circularly shifting and overlaying the forged image with the original forged image. Initially each pixel of both circularly shifted image and the original one is matched with one another and if there is a match, the corresponding segment of that pixel is checked in both images. If the corresponding pixels are greater and equal to predefined segment size B then the segment is said to be copy-moved.

Let us assume that $x_{i,j}$ is a pixel value of a grayscale image of size $M * N$ at position i, j. In exhaustive search following is examined.

$$\left|x_{i,j} - x_{i+k(modM),j+l(modN)}\right| = 0\,, \quad k = 0,1,2\,\ldots, M-1; l = 0,1,2,\ldots, N-1\ \forall\ i,j$$

While comparing the images, cyclic shift $[k, l]$ is found similar to cyclic shift $[k', l']$ where $k' = M - k$ and $l' = N - l$. Thus inspection of only those shifts with $1 \leq k \leq \frac{M}{2}, 1 \leq l \leq \frac{N}{2}$ will suffice, thus reducing the computational complexity by 4.

The time complexity of this algorithm is too high and hence it makes it impractical to detect the copy-moved segment of large images. It's implementation is simple and it is effective in finding major part of copy-moved segment in lossless image formats but for each shift every pixel pair must be compared. This comparison requires MN operations for each shift and there are $\frac{M}{2}, \frac{N}{2}$ total number of shifts and hence the complexity is proportional to $(MN)^2$. Thus, it is a viable option for small images only.

3.2 Detection based on Autocorrelation

Autocorrelation is degree of similarity between a series and a lagged version of itself over successive time intervals. The autocorrelation of image A of size MN can be defined as

$$r_{k,l} = \sum_{k=0}^{M-1} \sum_{0}^{N-1} x_{i,j}\, x_{i+k,j+l} \quad i = 0 \ldots M-1, j = 0 \ldots N-1$$

The logic behind the detection based on autocorrelation is that the original and the duplicated segments will introduce peaks in the autocorrelation for the shifts that corresponds to the copy-moved segments. If the autocorrelation is a spike, it says that the image has no elements that are correlated- that everything is unique. Lobes, away from (0,0), suggest repetition. The distance away from (0,0) is the periodicity of the repetition.

Let the minimal size of copy-moved segment is B, then to find the copy-moved segment following steps are undertaken.

1. Autocorrelation r of the forged image is computed.
2. Since the correlation is symmetric half of the correlation is removed.
3. $r = 0$ for exact overlapping of two images is set.
4. Maximum of r is ound from which shift vector is calculated and this shift is examined using exhaustive search.

5. If the detected area is greater than B, finish, else repeat step 5 with next maximum r.

This method computationally efficient as exhaustive search is not need to be performed for many different shift vectors.

Although this method is simple and is not of high computational complexity, it often fails to detect the forged area until forged area is large. It generates large number of false negatives.

3.3 Detection based on Exact Block matching

In exact matching algorithm forged image is divided into overlapping square blocks of size bxb. The square block is slid by one pixel from left corner to the right and from top to down. For each block position, the pixel values of the block are extracted in column-wise order into a row of a 2-D matrix A with b^2 columns and $(M - b + 1)(N - b + 1)$ rows, where $[M, N]$ is the size of the forged image.

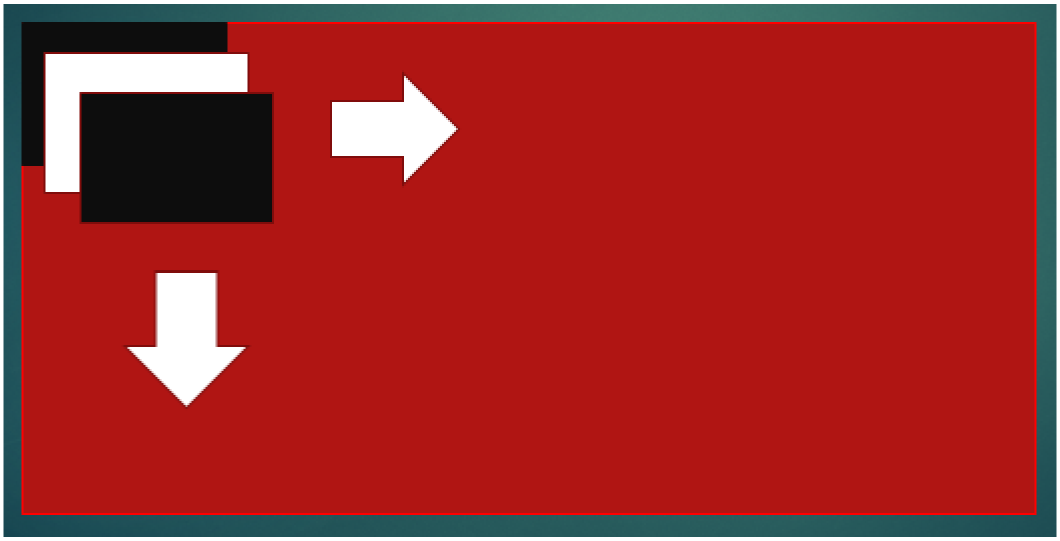


Fig 3.3.1: Exact Block Matching: Depicting how the image is divided into blocks

Each row in matrix A corresponds to one position of the sliding block B. Two identical rows in A corresponds to a copy moved segment of size equal to that of B. Identical rows can be easily extracted by lexicographically sorting the matrix A. Hence, identical rows comes next to one another in matrix A and can be easily identified by looking at two consecutive rows that are identical. Although, this algorithm is far more efficient than the previous two, but if the image is saved in lossy formats such as JPEG, a vast majority of identical blocks are lost.

3.4 Detection Using Principal Component Analysis

In PCA based approach, Images are divided into overlapping blocks, similar to block matching approach, each of which are considerably smaller than the size of duplicated region. The images represented in the form of a 2-D array of which each row contains a block arranged in row-wise manner. Let x_i ,i=1,2,3,…, N_b where N_b =(M+N-b)(M+N-b), be the the blocks in vectorized form. The covariance matrix of each block is calculated as follows:

$$C = \sum_{i=1}^{N_b} x_i x_i^T$$

The eigen vectors e_i of covariance matrix corresponding to eigen values λ_j defines the principal components where where j=1 … b and $\lambda_1 > \lambda_2 > \lambda_3 > \ldots > \lambda_b$. The eigenvectors, e_j, form a new linear basis for each image block,

$$x_i = \sum_{j=1}^{b} a_j x_j$$

where $a_j = x_i^T e_j$ is the new representation for each image block.

The dimensionality of each block is reduced by simply truncating the vector x_i to first N_t terms.

Below given is the algorithm for the same:

1) Let total number of pixels in a grayscale or colored image be N.
2) Let b be the number of pixels per block and N_b be the number of such blocks. N_n,N_f, N_d be the no. of neighboring rows to be searched in the lexicographically sorted matrix, minimum frequency threshold and minimum offset threshold respectively.
3) Build a $N_b * b$ matrix whose rows represents each block of N_b.
4) A new N_t-dimensional representation C is computed using PCA.
5) C is then sorted in lexicographic order to obtain a matrix S. Let S_i denotes a row of matrix S and (x_i, y_i) be co-ordinates that corresponds to S_i.

6) For every pair of rows S_i and S_j from S such that $|i-j| < N_n$. The coordinates corresponding to S_i and S_j are stored in a list.

7) Offset frequency of each element in the list is calculated as follows:

   $(x_i - x_j, y_i - y_j)$ if $x_i - x_j > 0$

   $(x_j - x_i, y_i - y_j)$ if $x_i - x_j < 0$

   $(0, y_i - y_j)$ if $x_i = x_j$

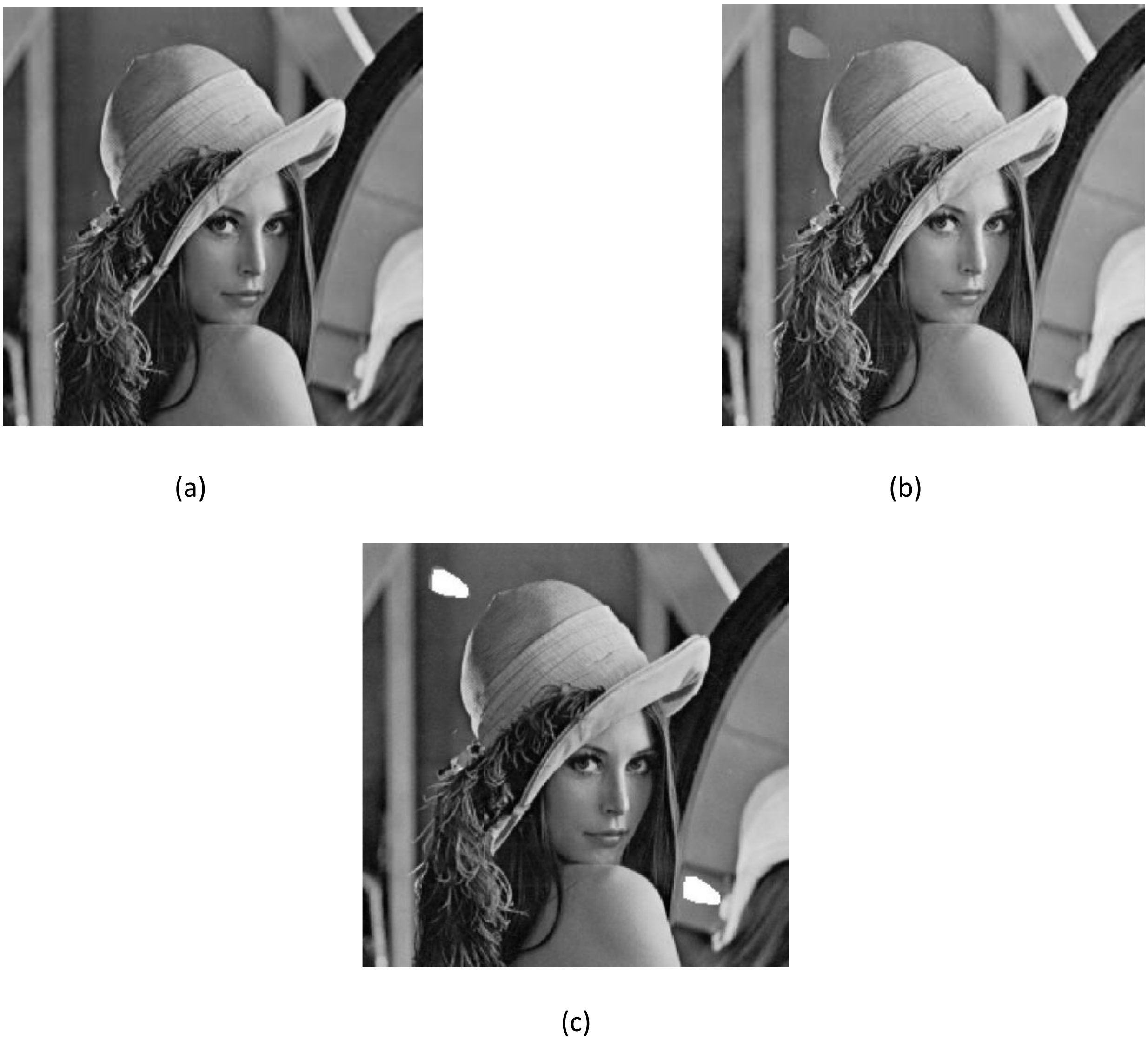
8) The pair with offset frequency less than N_f and the pair whose offset magnitude$\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$, is less than N_d are discarded.

9) The remaining pairs of block are used to color the image with different grayscale values. The colored part represent the copy-moved portion of the image.

3.5 Performance Analysis

All simulation was done in matlab 2013a. In our experiment, we tampered images by copying and pasting one image block over another, in the same image.

In our first method, we took a 128x128 pixel image and 8x8 to be the smallest block size. We circularly shifted the image by one pixel each time and matched for the common region of minimal block size. The computational complexity of this method was 128x128/4 . The original image is shown in figure 3.4.1 (a), the tampered image is shown in figure 3.4.1 (b) and the final output image of this method is shown in figure 3.4.1 (c) .



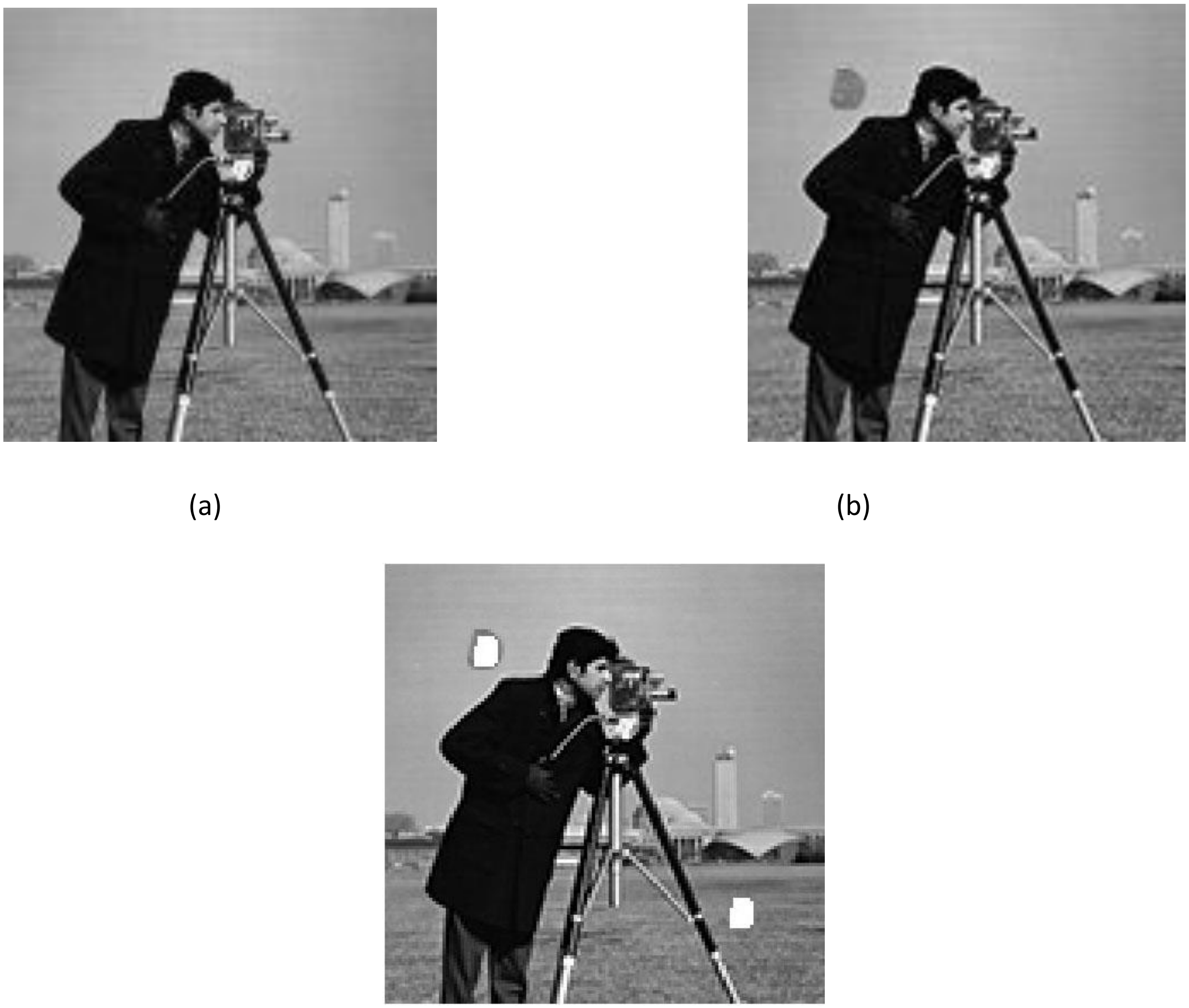
(a)

(b)

(c)

Fig. 3.4.1: Exhaustive Search: (a) Original Image, (b) Tampered Image, (c) Output Image.

In the second method a 128x128 image is taken and its autocorrelation is calculated. We took the highest peak of autocorrelation and shift is calculated. Then exhaustive search is applied to this shift. Figure 3.4.2(a) is the original image for this method. Tampered image is shown in figure 3.4.2(b) and output for this method is shown in 3.4.2(c). The computational complexity is better than previous method as we don’t have to match exhaustively for every shift.



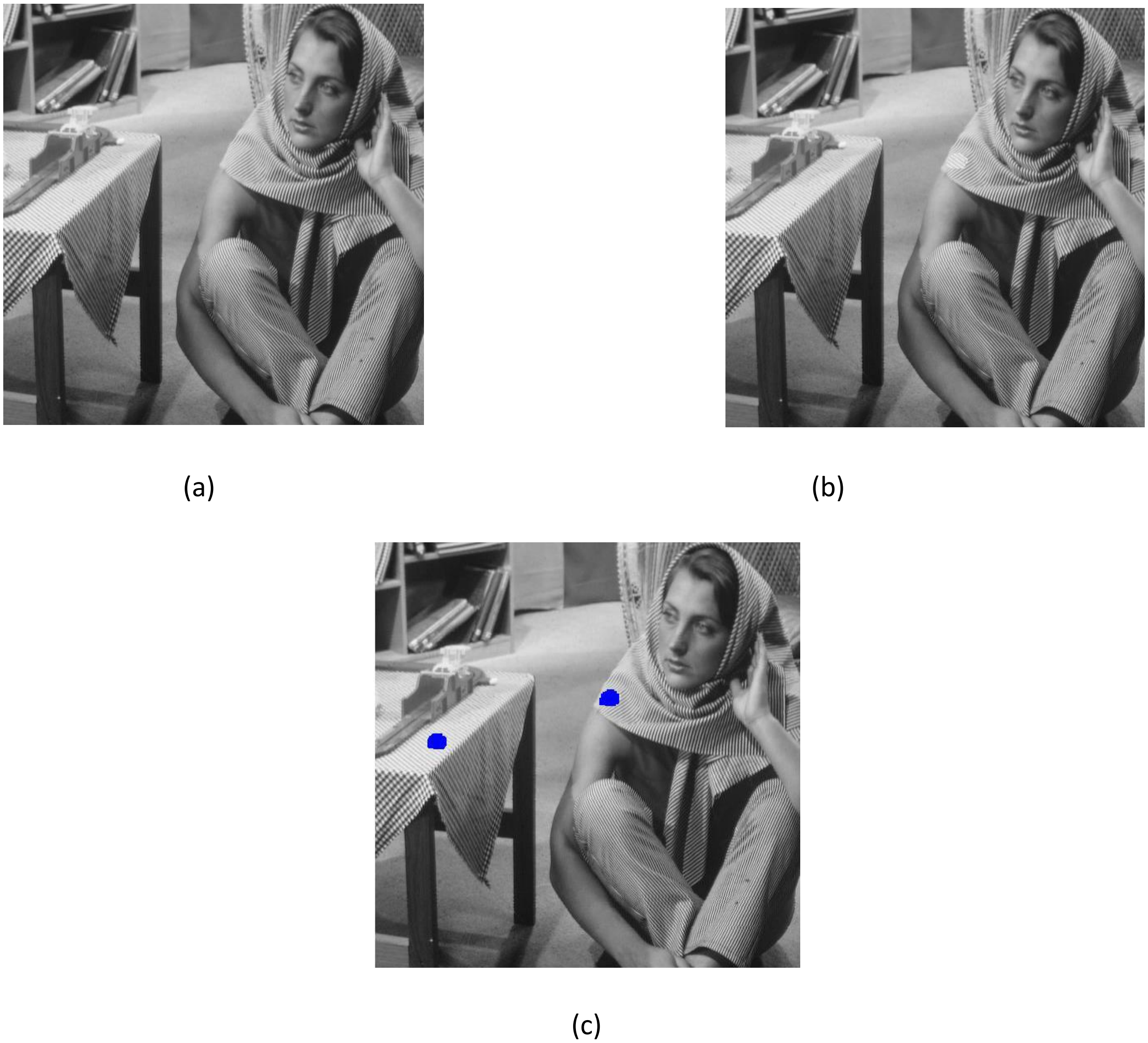
(a)

(b)

(c)

Fig. 3.4.2: Method of Autocorrelation: (a) Original Image, (b) Tampered Image, (c) Output Image.

In the third method of copy move forgery detection 128x128 pixel image was taken and our minimum block size was set to 8x8 pixels. The image was then subdivided into (128-8+1)(128-8+1) sub blocks off size 8x8 and stored in column order in an array. The array was then lexicographically sorted to find the adjacent matching blocks and the matching block was then mapped to the forged image to find the copy-moved segment. Figure 3.4.3(a) shows the original 128x128 image. Figure 3.4.3(b) shows the corrupted image and figure 3.4.3(c) is the output for this method.



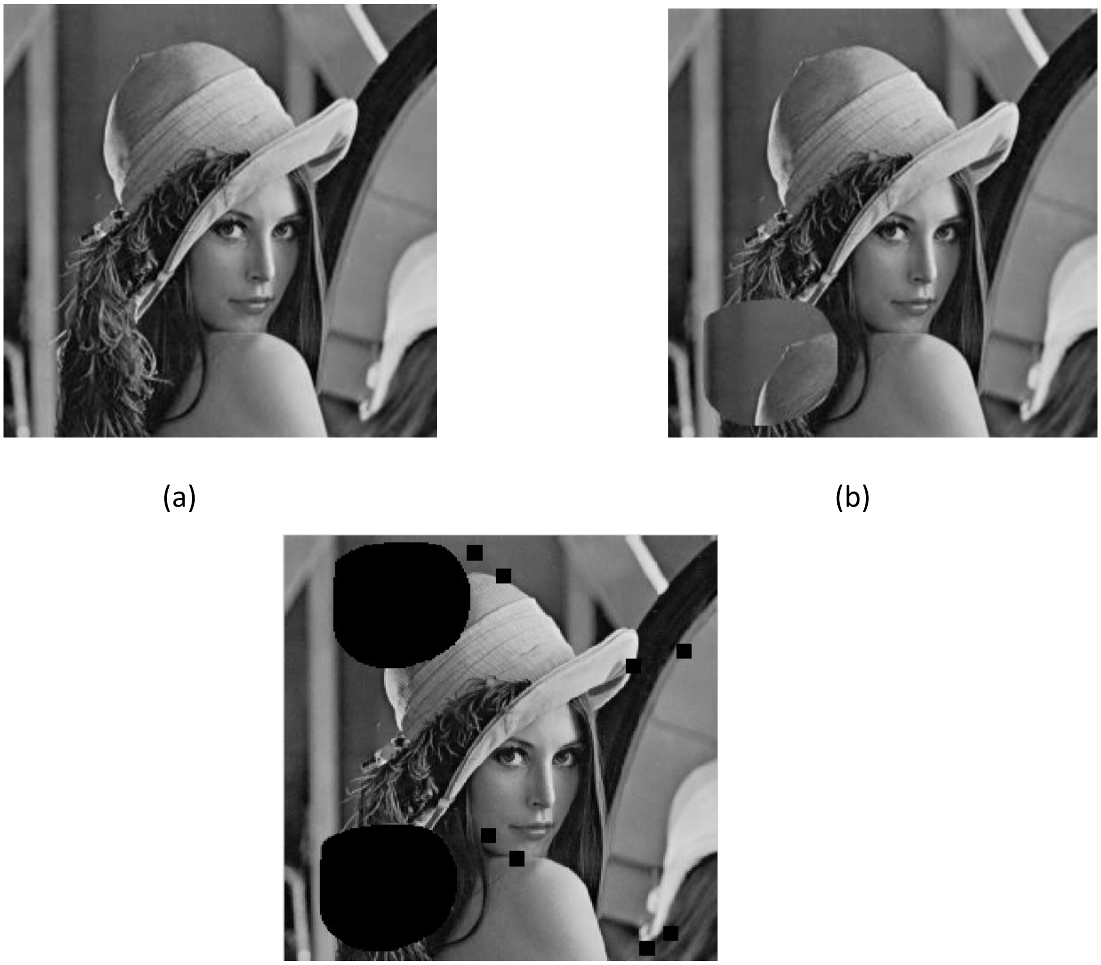
(a)

(b)

(c)

Fig. 3.4.3: Exact Block Matching: (a) Original Image, (b) Tampered Image, (c) Output Image.

In the fourth method of copy move forgery detection i.e. Principal of component analysis 128x128 pixel image was taken and our minimum block size was set to 9x9 pixels. The image was then subdivided into (128-9+1)(128-9+1) sub blocks off size 9x9 and stored in column order in an array. Using the method of principal of component analysis dimensions of the data are reduced that will help to sort the whole list and then further comparison easy. The array was then lexicographically sorted and find the adjacent matching blocks and the matching block was then mapped to the forged image to find the copy-moved segment. Figure 3.4.4(a) shows the original 128x128 image. Figure 3.4.4(b) shows the corrupted image and figure 3.4.4(c) is the output for this method.



(a)

(b)

(c)

Fig. 3.4.4: Method of Principal of component analysis: (a) Original Image, (b) Tampered Image, (c) Output Image.

Chapter 4

Parameterization and Simulation Results

4.1 False Positive/Negative and Detection Accuracy

There are few factors on the basis of which we can compare the images and our results.

False Positive detection or indication of a condition in data reporting when in reality it is not present, while a **False Negative** fails to detect or indicate a test result when in real there is some error present in the data.

We will be concentrating only on the improving the false positives. In copy move forgery, sometimes we will detect a forgery even if there is no forgery at all. This generally happens in case there is some smooth pattern like sky or grass etc.

Our aim is to make the method more robust and less prone to false positive. In earlier methods, we were taking exact matching of the image pixels and the image was not subjected to additive noise or lossy JPEG compression. So we were not facing any kind of false positives. While making the method more robust, there is some tradeoff between robustness and false positive rate. We will be focusing on how to reduce the false positive while increasing the robustness of the method. We would be calculating **False Positives Rate** v/s **Copied-Moved block Size**.

In our context, the False Positive rate is proportional to number of falsely detected pixels i.e., the pixels which are not copy-moved but are still detected by our detection algorithms.

$$FPR = \frac{Number\ of\ falsely\ detected\ pixels}{Total\ number\ of\ copy\ moved\ pixels} * 100$$

False negative rate is proportional to number of undetected pixels which were from copy-moved part of the image but are not detected by the detection algorithms

$$FNR = \frac{Number\ of\ undetected\ copy\ moved\ pixels}{Total\ number\ of\ copy\ moved\ pixels} * 100$$

A method can be made robust by taking the features of the image in consideration instead of the image pixel value itself. Features of an image such as Principal component analysis (PCA), Discrete Cosine Transform (DCT) etc. are not much prone to changes caused due to any alteration during forging of an image. They can be slightly deviated from the original but are useful in finding the copy-moved block in the image.

Due to the deviation caused by the enhancement in the copy-moved image, similar feature vector are considered as from copy moved segment even if they are not exact. This raises the possibility of detection of false segment.

Other factor is the **Detection accuracy**, which is the percentage of total number of pixels correctly detected from the forged part by the forgery detection technique.

$$\boldsymbol{DA = \frac{Number\ of\ Correctly\ detected\ pixels\ (in\ copy - moved\ area)}{Total\ number\ of\ copy\ moved\ pixels} * 100}$$

In the simulation process, we took five different forgery sizes for our analysis. We took a 256x256 image ‘lena.tiff’ and forged it for different forgery sizes as 5%, 10%, 20%, 30% and 40%. Then for each forgery size different detection techniques were applied for 8 different unit block sizes. Block sizes taken in consideration are 12x12, 16x16, 20x20, 24x24, 28x28, 32x32 and 36x36. Following are the graphs obtained after the simulation process.

4.2 A Comparative Study of techniques

The techniques that were taken under consideration were block based and they were compared in terms of detection accuracy and false positive/negative. Comparison were done by varying unit block size and varying the image forgery sizes.

4.2.1 On the basis of Detection Accuracy

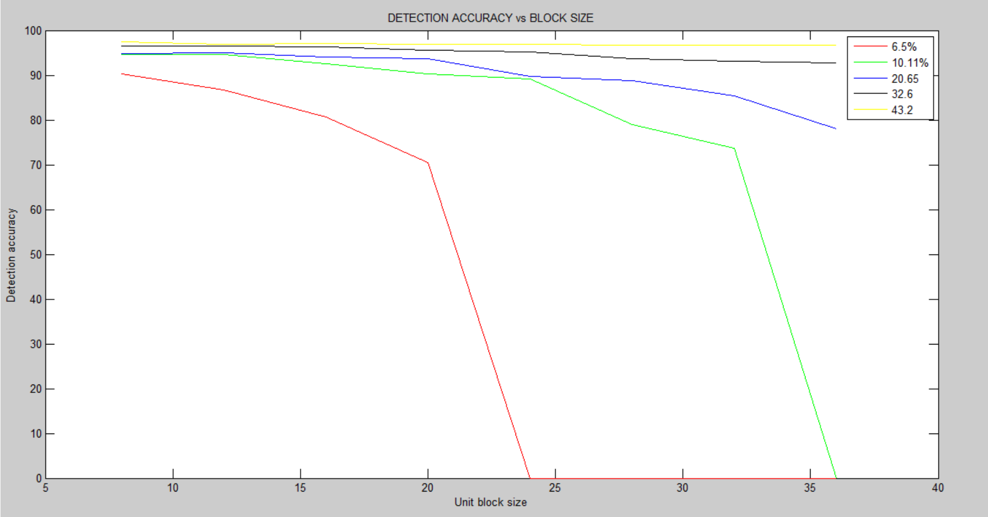


Figure 4.2.1.1: Simulation result and plot for Exact Block Matching Technique - Detection Accuracy with unit block size for 5 different set of Forgery size.

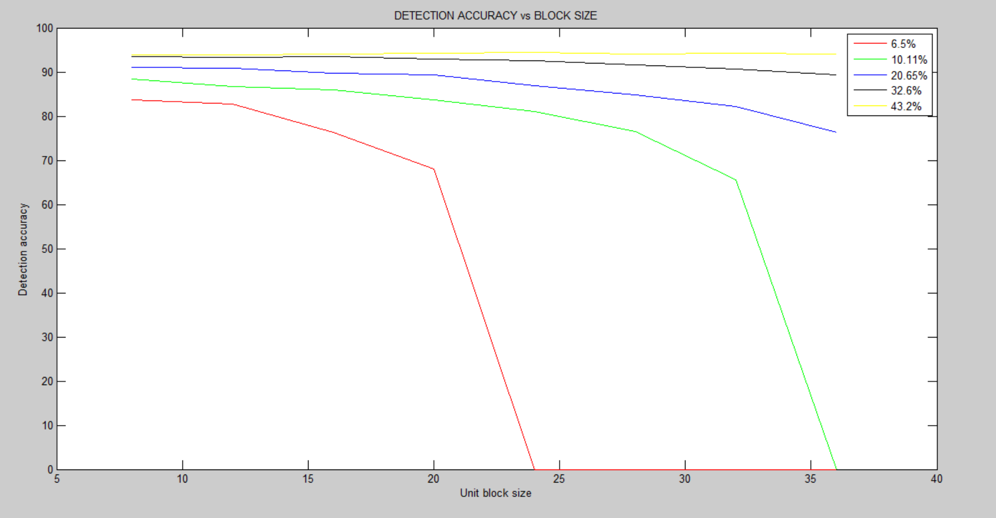


Figure 4.2.1.2: Simulation result and plot for Exhaustive Search - Detection Accuracy with unit block size for 5 different set of Forgery size.

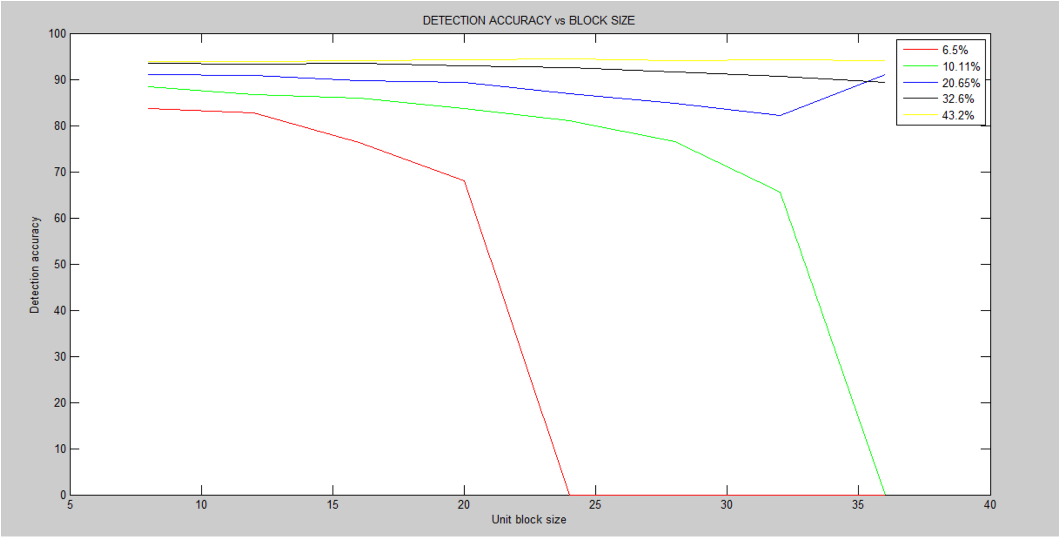


Figure 4.2.1.3: Simulation result and plot for Autocorrelation Based Detection - Detection Accuracy with unit block size for 5 different set of Forgery size.

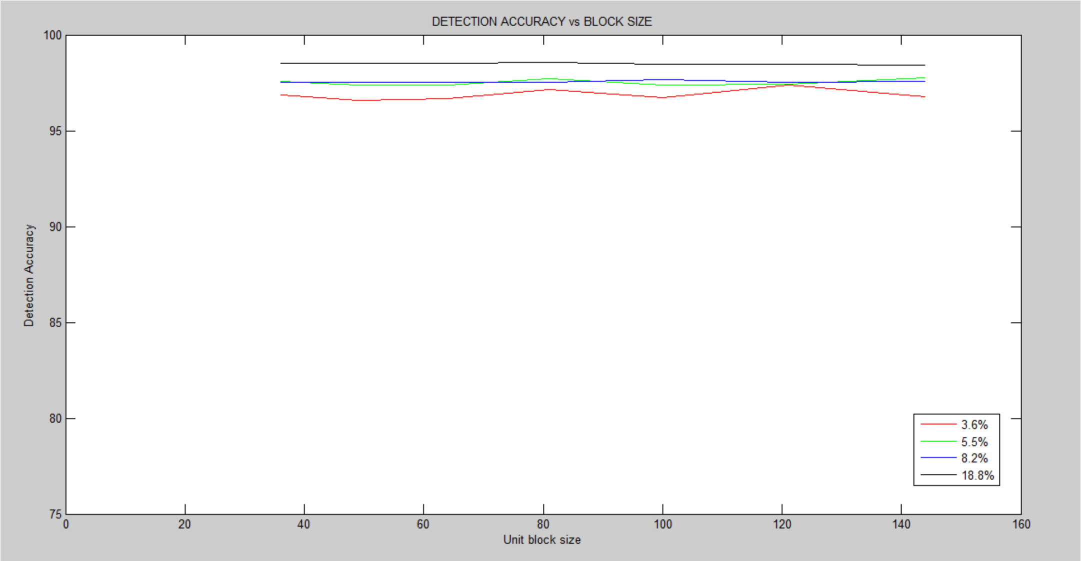
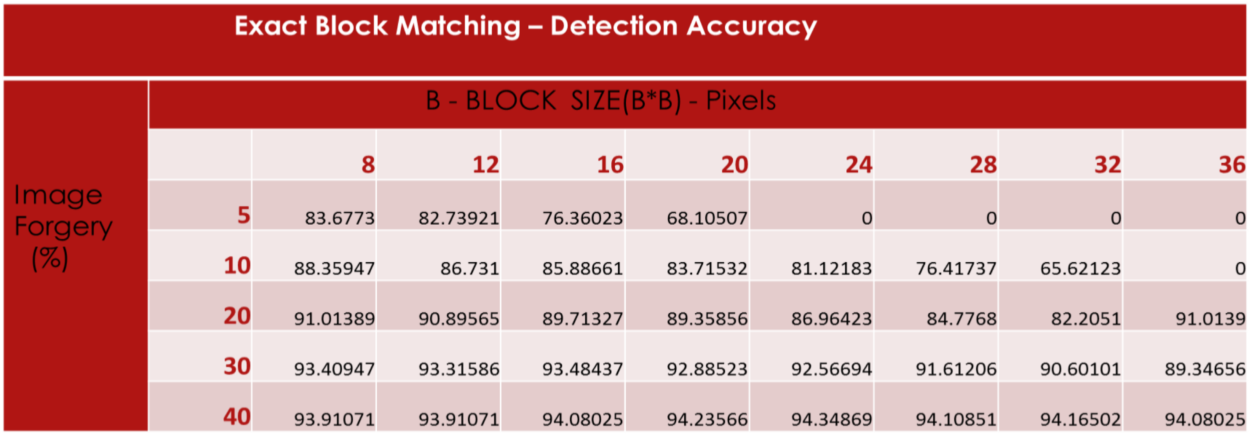


Figure 4.2.1.4: Simulation result and plot for PCA based detection - Detection Accuracy with unit block size for 4 different set of Forgery size.

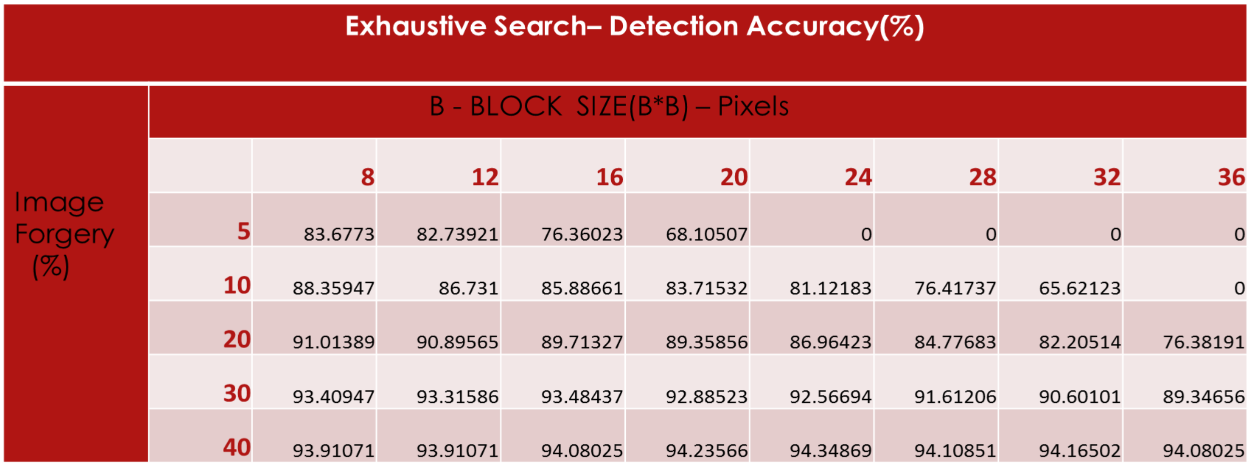
From the plots, it can be easily observed that the detection accuracy of a forged image increases with the increase in forged area and can decrease with the increase in the size of unit block use in the detection algorithm. Also of all the four techniques it was found that detection using the PCA approach gave the best result in terms of detection accuracy and also were efficient in terms of time complexity. Variations in detection accuracy can be observed in the following tables.



Exact Block Matching – Detection Accuracy

		B - BLOCK SIZE(B*B) - Pixels							
Image Forgery (%)		8	12	16	20	24	28	32	36
	5	83.6773	82.73921	76.36023	68.10507	0	0	0	0
	10	88.35947	86.731	85.88661	83.71532	81.12183	76.41737	65.62123	0
	20	91.01389	90.89565	89.71327	89.35856	86.96423	84.7768	82.2051	91.0139
	30	93.40947	93.31586	93.48437	92.88523	92.56694	91.61206	90.60101	89.34656
	40	93.91071	93.91071	94.08025	94.23566	94.34869	94.10851	94.16502	94.08025

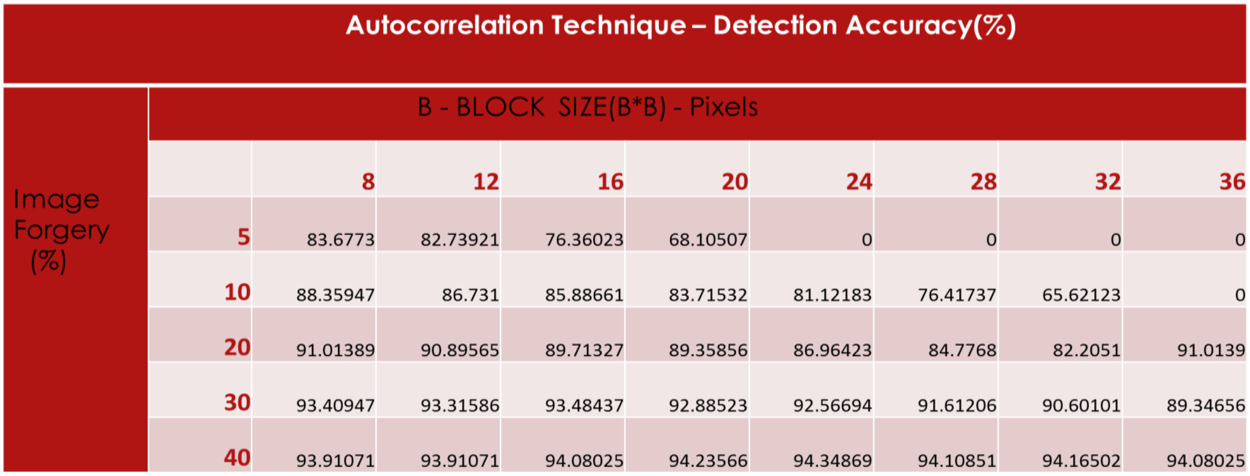
Table 1: Detection accuracy of exact block matching



Exhaustive Search– Detection Accuracy(%)

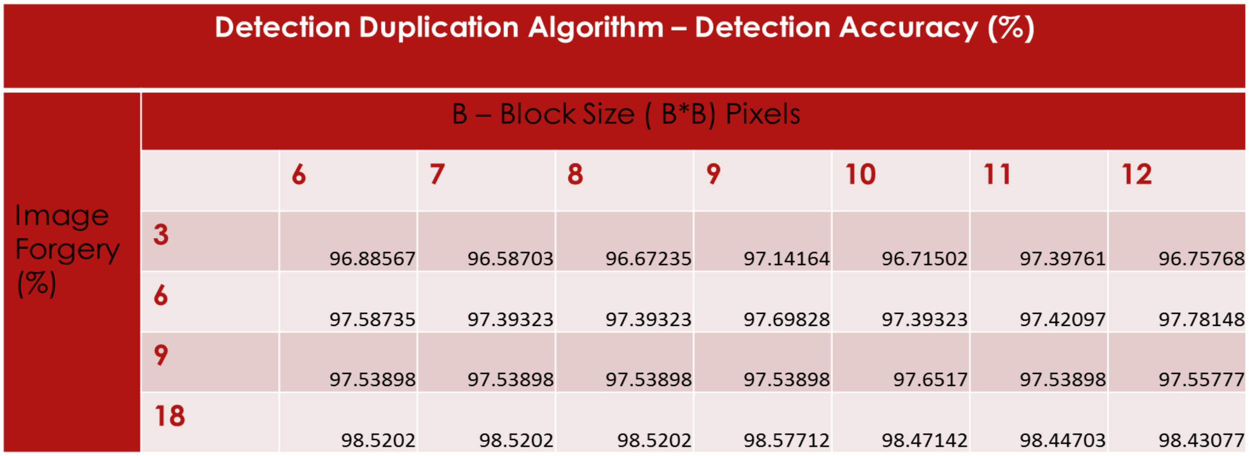
		B - BLOCK SIZE(B*B) – Pixels							
Image Forgery (%)		8	12	16	20	24	28	32	36
	5	83.6773	82.73921	76.36023	68.10507	0	0	0	0
	10	88.35947	86.731	85.88661	83.71532	81.12183	76.41737	65.62123	0
	20	91.01389	90.89565	89.71327	89.35856	86.96423	84.77683	82.20514	76.38191
	30	93.40947	93.31586	93.48437	92.88523	92.56694	91.61206	90.60101	89.34656
	40	93.91071	93.91071	94.08025	94.23566	94.34869	94.10851	94.16502	94.08025

Table 2: Detection accuracy of exhaustive matching



Autocorrelation Technique – Detection Accuracy(%)									
Image Forgery (%)	B - BLOCK SIZE(B*B) - Pixels								
		8	12	16	20	24	28	32	36
	5	83.6773	82.73921	76.36023	68.10507	0	0	0	0
	10	88.35947	86.731	85.88661	83.71532	81.12183	76.41737	65.62123	0
	20	91.01389	90.89565	89.71327	89.35856	86.96423	84.7768	82.2051	91.0139
	30	93.40947	93.31586	93.48437	92.88523	92.56694	91.61206	90.60101	89.34656
	40	93.91071	93.91071	94.08025	94.23566	94.34869	94.10851	94.16502	94.08025

Table 3: Detection accuracy of Autocorrelation Based matching



Detection Duplication Algorithm – Detection Accuracy (%)								
Image Forgery (%)	B – Block Size (B*B) Pixels							
		6	7	8	9	10	11	12
	3	96.88567	96.58703	96.67235	97.14164	96.71502	97.39761	96.75768
	6	97.58735	97.39323	97.39323	97.69828	97.39323	97.42097	97.78148
	9	97.53898	97.53898	97.53898	97.53898	97.6517	97.53898	97.55777
	18	98.5202	98.5202	98.5202	98.57712	98.47142	98.44703	98.43077

Table 4: Detection accuracy PCA based matching

4.2.2 On the basis of False Positives/Negatives

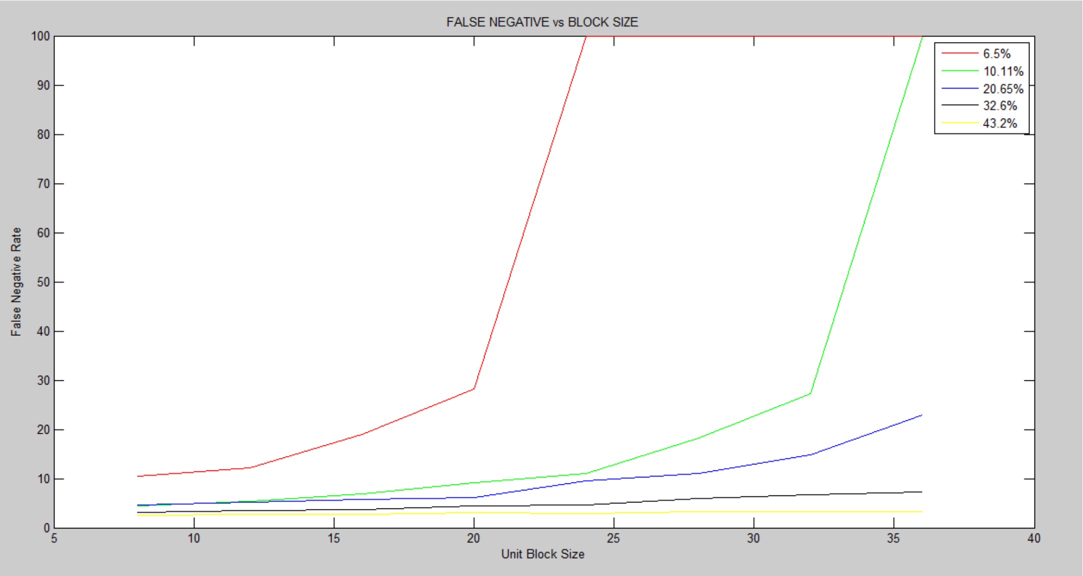


Figure 4.2.2.1: Simulation result and plot for Exact Block Matching – False Negative with unit block size for 5 different set of Forgery size.

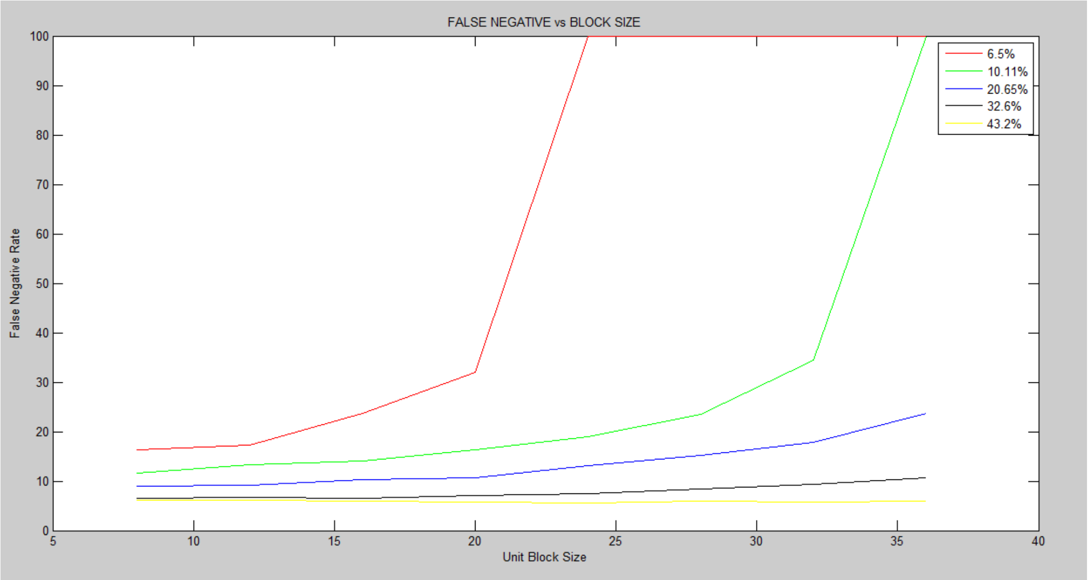


Figure 4.2.2.2: Simulation result and plot for Exhaustive Search – False Negative with unit block size for 5 different set of Forgery size.

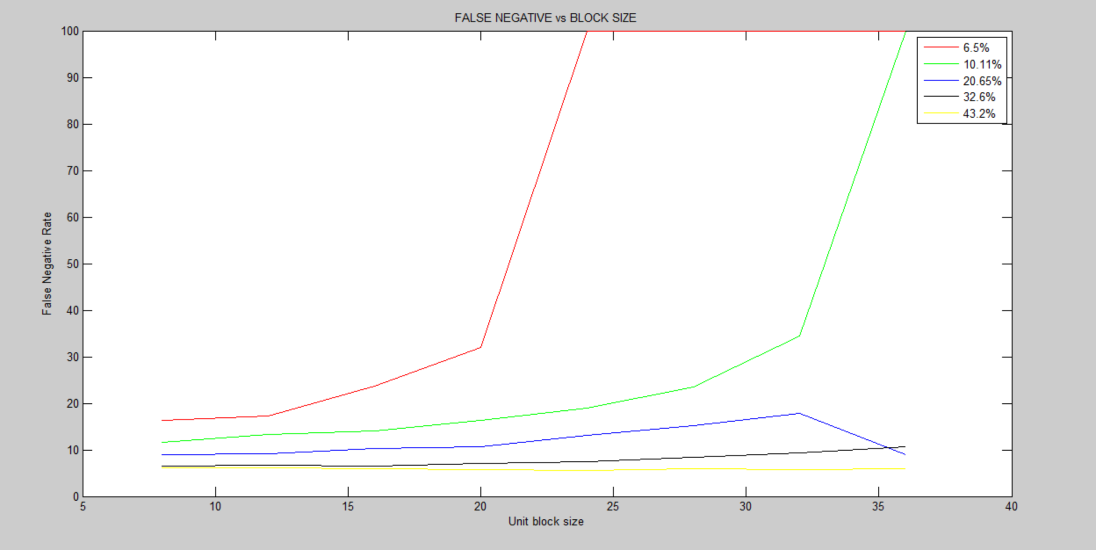


Figure 4.2.2.3: Simulation result and plot for Autocorrelation based detection– False Negative with unit block size for 5 different set of Forgery size.

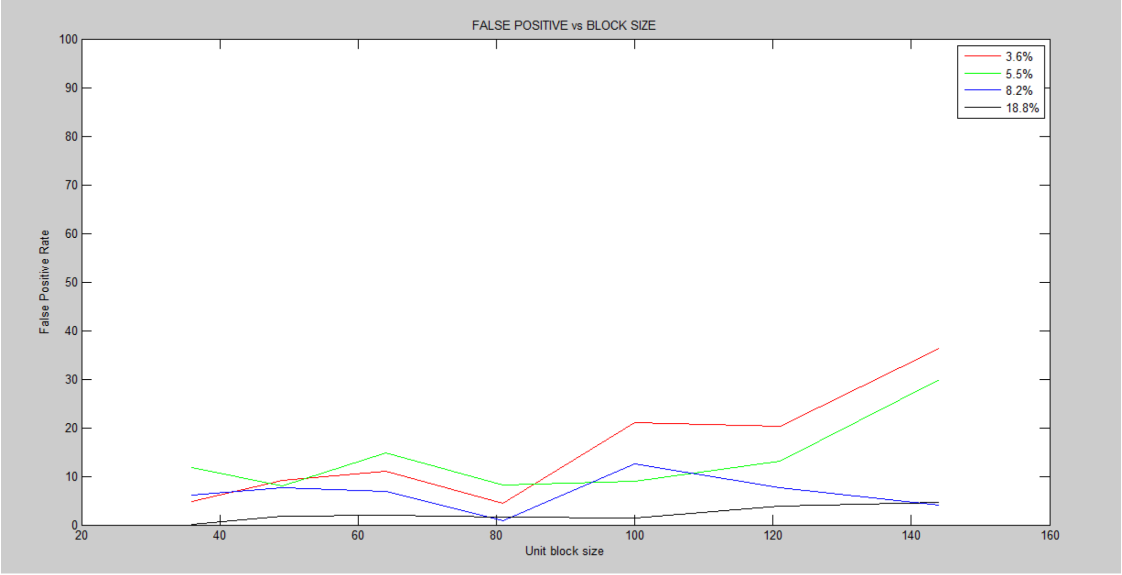
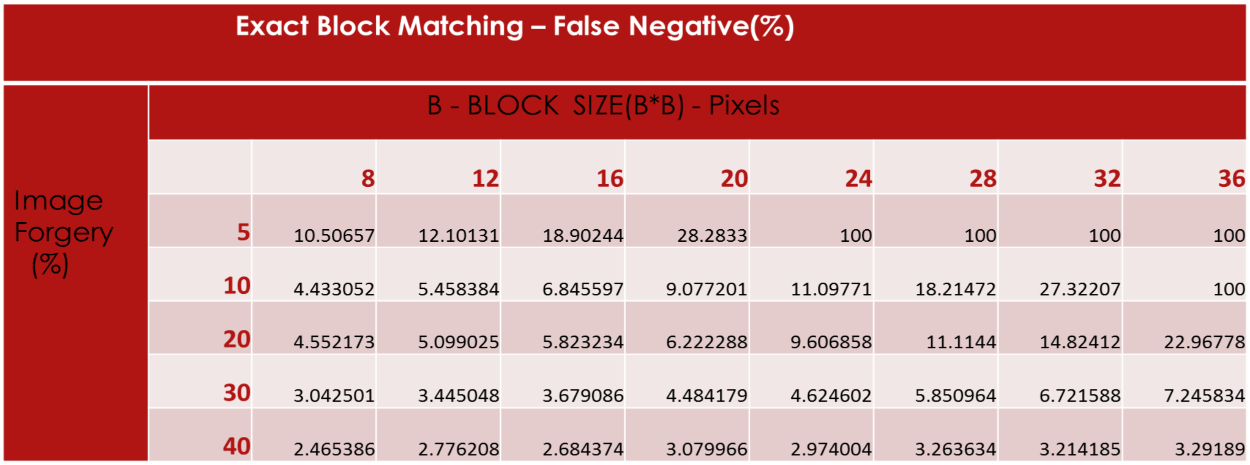


Figure 4.2.2.4: Simulation result and plot for PCA based detection – False Negative with unit block size for 5 different set of Forgery size.

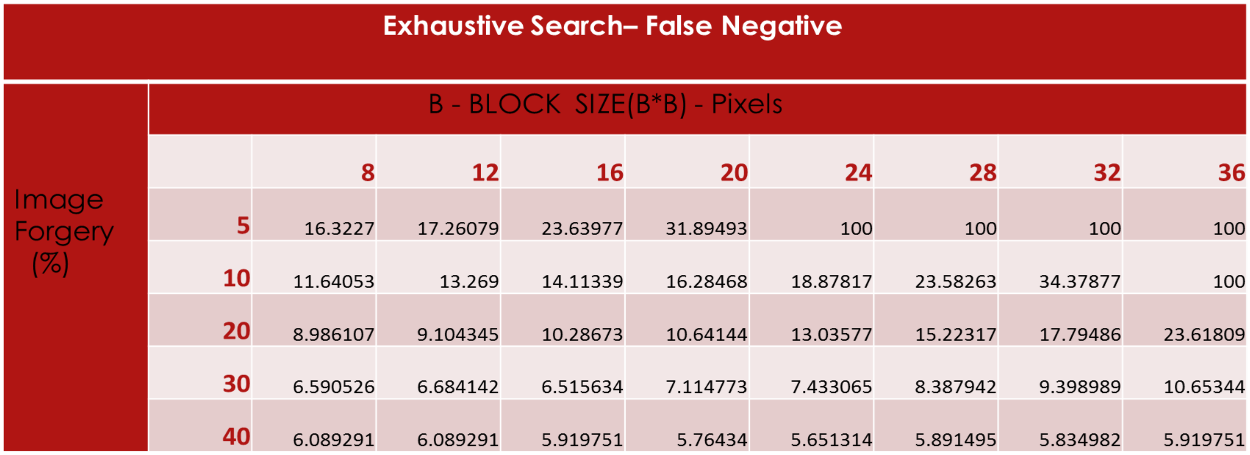
From the plots, it can be easily observed that the False Negative/Positive of a forged image decreases with the increase in forged area and can increase with the increase in the size of unit block use in the detection algorithm. Also of all the three techniques it was found that detection using the Exact matching approach gave the best result in terms of False negative and there were no False positives generated during these techniques. But in PCA based detection technique there were some instances of false positive which had similar behavior as of false negatives. Variations in False Negatives/Positives can be observed in the following tables.



Exact Block Matching – False Negative(%)

Image Forgery (%)		B - BLOCK SIZE(B*B) - Pixels							
		8	12	16	20	24	28	32	36
	5	10.50657	12.10131	18.90244	28.2833	100	100	100	100
	10	4.433052	5.458384	6.845597	9.077201	11.09771	18.21472	27.32207	100
	20	4.552173	5.099025	5.823234	6.222288	9.606858	11.1144	14.82412	22.96778
	30	3.042501	3.445048	3.679086	4.484179	4.624602	5.850964	6.721588	7.245834
	40	2.465386	2.776208	2.684374	3.079966	2.974004	3.263634	3.214185	3.29189

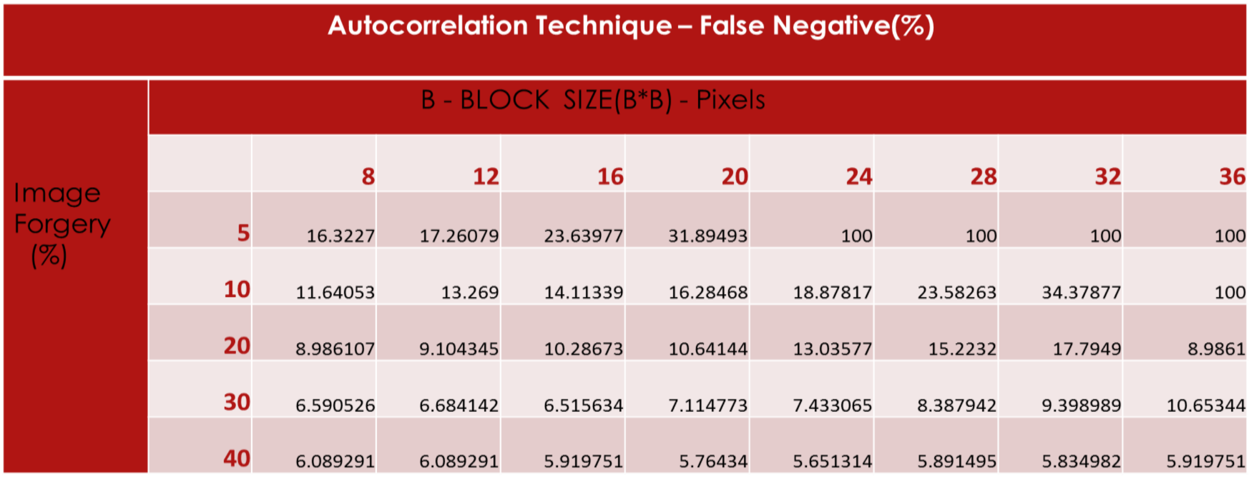
Table 5: False Negative of Exact Block Matching



Exhaustive Search– False Negative

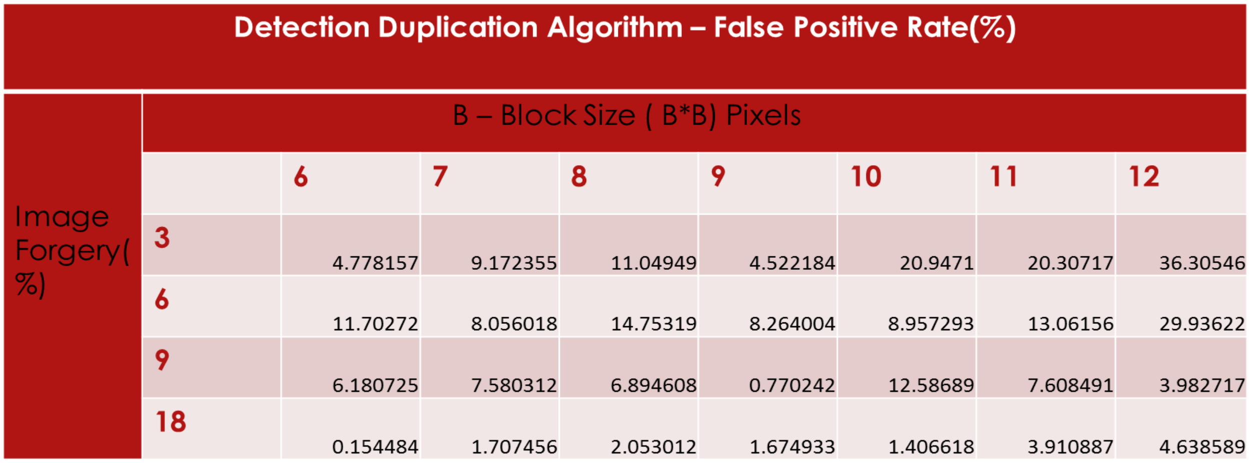
Image Forgery (%)		B - BLOCK SIZE(B*B) - Pixels							
		8	12	16	20	24	28	32	36
	5	16.3227	17.26079	23.63977	31.89493	100	100	100	100
	10	11.64053	13.269	14.11339	16.28468	18.87817	23.58263	34.37877	100
	20	8.986107	9.104345	10.28673	10.64144	13.03577	15.22317	17.79486	23.61809
	30	6.590526	6.684142	6.515634	7.114773	7.433065	8.387942	9.398989	10.65344
	40	6.089291	6.089291	5.919751	5.76434	5.651314	5.891495	5.834982	5.919751

Table 6: False Negative of Exhaustive Matching Technique



Autocorrelation Technique – False Negative(%)									
Image Forgery (%)	B - BLOCK SIZE(B*B) - Pixels								
		8	12	16	20	24	28	32	36
	5	16.3227	17.26079	23.63977	31.89493	100	100	100	100
	10	11.64053	13.269	14.11339	16.28468	18.87817	23.58263	34.37877	100
	20	8.986107	9.104345	10.28673	10.64144	13.03577	15.2232	17.7949	8.9861
	30	6.590526	6.684142	6.515634	7.114773	7.433065	8.387942	9.398989	10.65344
	40	6.089291	6.089291	5.919751	5.76434	5.651314	5.891495	5.834982	5.919751

Table 7: False Negative of Autocorrelation Technique



Detection Duplication Algorithm – False Positive Rate(%)								
Image Forgery(%)	B – Block Size (B*B) Pixels							
		6	7	8	9	10	11	12
	3	4.778157	9.172355	11.04949	4.522184	20.9471	20.30717	36.30546
	6	11.70272	8.056018	14.75319	8.264004	8.957293	13.06156	29.93622
	9	6.180725	7.580312	6.894608	0.770242	12.58689	7.608491	3.982717
	18	0.154484	1.707456	2.053012	1.674933	1.406618	3.910887	4.638589

Table 8: False Positive of PCA based Technique

Chapter 5

Conclusion

Copy-Move forgery is a dynamic field in which a lot of researchers are working on and have successfully implemented different techniques. In our work we have considered four techniques – Exhaustive Search, Autocorrelation, Exact Block Matching and Detection Duplication Algorithm using PCA. All these techniques are block based techniques and have various positives and negatives like Exhaustive Search suffers from high time complexity and PCA approach has very good accuracy results but it suffers from False Positives. So, when it comes to - which technique to choose for a particular type of forged image - then this parameterization will help the user. On the basis of requirement, one can use a particular unit block size and a particular technique. From our simulation results, it is clear that Detection Accuracy is inversely proportional to unit block size and for larger forgery the detection accuracy will be high. For False negative it is just opposite to that of detection accuracy i.e. directly proportional to block size. For the first three techniques that we discussed False Positive for them was zero but in the PCA approach we have high accuracy but we face a small False Positive rate so there is a tradeoff between Detection Accuracy and False positive. Hence, parameterization of the copy-move forgery detection techniques, proposed here, help by giving us options to prioritize the selection of a particular detection scheme over others.

Bibliography

[1] A.J. Fridrich , B.D. Soukal , A.J. Lukáš, “Detection of copy-move forgery in digital images”, Proceedings of Digital Forensic Research Workshop, 2003.

[2] A.C. Popescu and H. Farid, “Exposing digital forgeries by detecting traces of re-sampling”, IEEE Transactions on Signal Processing, vol. 53, no. 2, pp. 758–767, 2005.

[3] D. Lowe, “Distinctive image features from scale-invariant key-points”, Internaional Journal of Computer Vision, vol. 60, no. 2, pp. 91–110, 2004.

[4] X. Pan, S. Lyu, “Detecting image duplication using SIFT features”, Proceedings of IEEE ICASSP, 2010.

[5] A. Popescu and H. Farid, Exposing digital forgeries by detecting duplicated image regions. Computer Science, Dartmouth College, Tech. Rep. TR2004-515, 2004.

[6] Weiqi Luo; Jiwu Huang; Guoping Qiu, "Robust Detection of Region-Duplication Forgery in Digital Image," Pattern Recognition, 2006. ICPR 2006. 18th International Conference on , vol.4, pp.746,749