Development of a Robust Algorithm to Detect Multi-Cloning (Copy-Move) Forgeries

Author

Muhammad Qasim Altaf

NUST201261239MCEME35512F

Supervisor

Brig. Dr. Javaid Iqbal

DEPARTMENT OF MECHATRONICS ENGINEERING

COLLEGE OF ELECTRICAL & MECHANICAL ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY

ISLAMABAD

AUGUST, 2016

Development of a Robust Algorithm to Detect Multi-Cloning (Copy-Move) Forgeries

Author

Muhammad Qasim Altaf

NUST201261239MCEME35512F

A thesis submitted in partial fulfillment of the requirements for the degree of

MS Mechatronics Engineering

Thesis Supervisor:

Brig. Dr. Javaid Iqbal

Thesis Supervisor's Signature: _____

DEPARTMENT OF MECHATRONICS ENGINEERING

COLLEGE OF ELECTRICAL & MECHANICAL ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY, ISLAMABAD

AUGUST, 2016

Declaration

I certify that this research work titled "*Development of a Robust Algorithm to Detect Multi-Cloning (Copy-Move) Forgeries"* is my own work. The work has not been presented elsewhere for assessment. The material that has been used from other sources it has been properly acknowledged / referred.

Signature of Student

Muhammad Qasim Altaf

2012-NUST-Ms-MECHT74

Language Correctness Certificate

This thesis has been read by an English expert and is free of typing, syntax, semantic, grammatical and spelling mistakes. Thesis is also according to the format given by the university.

Signature of Student

Muhammad Qasim Altaf

NUST201261239MCEME35512F

Signature of Supervisor

Brig. Dr. Javaid Iqbal

Copyright Statement

- Copyright in text of this thesis rests with the student author. Copies (by any process) either in full, or of extracts, may be made only in accordance with instructions given by the author and lodged in the Library of NUST College of E&ME. Details may be obtained by the Librarian. This page must form part of any such copies made. Further copies (by any process) may not be made without the permission (in writing) of the author.
- The ownership of any intellectual property rights which may be described in this thesis is vested in NUST College of E&ME, subject to any prior agreement to the contrary, and may not be made available for use by third parties without the written permission of the College of E&ME, which will prescribe the terms and conditions of any such agreement.
- Further information on the conditions under which disclosures and exploitation may take place is available from the Library of NUST College of E&ME, Rawalpindi.

Acknowledgements

All the praises and thanks be to Almighty ALLAH, the most gracious the most merciful, who gave me health, knowledge, courage and patience to accomplish this research.

I would specially thank my Supervisor Brig. Dr. Javaid Iqbal and Co- Supervisor Dr. Rab Nawaz for helping me through all the phases of the research. I also thank Dr. Kunwar Faraz khan and Dr. Umar Shabaz Khan for their guidance and support throughout the degree program.

I believe whatever I have achieved in my life is due to the unconditional support of my beloved Father (Late) and my loving Mother. All of my achievements in life are dedicated to them and their struggles they did while raising me to this level.

I thank Ms. Mariam Saleem, for helping me out during the research work, whenever I felt stuck. Completing this research could not have been possible, if your support was not there. I thank all other Faculty members, fellow students and University staff for helping me out during my time at worthy institute.

Dedicated to my beloved parents

Abstract

Digital image forensic is an emerging discipline that signifies a never ending struggle against image forgery. In this modern Era of advanced computation, tampering of images can be effortlessly accomplished by a number of available economical editing software's like Adobe Photoshop, Corel Draw etc. This poses a need for establishing techniques in order to verify the integrity and authenticity of digital images.

There are many types of image forgery; cloning/copy-move attack is one of them. Cloning forgery is specific type of tampering that uses a portion of original image as source to hide or duplicates certain features within the same image. This type of tampering is considered to be most advanced and has become researcher's point of interest in recent years. So the research in this dissertation is carried on this paradigm.

The research carried out during this dissertation is divided in two phases. The first phase focuses on development of a novel and robust model, which can identify single and multiple blind cloning forgeries in a given image while second phase deals with comparison of developed model with previously developed techniques in the field of digital multi-cloning detection.

The proposed methodology utilizes colored image unlike previously developed methods which worked on gray-scale. Local binary pattern label along with clustering is used to minimize false positive rate. The evaluation of the proposed method has been done on data set, MICC-F220, which is approved by IEEE transactions of information security and image forensics. All tests on algorithm and experiments have been carried out using MATLABR2012a. Afterwards, a comprehensive yet detailed comparison of previously developed copy-move methodologies with the proposed technique is presented. Parameters like time complexity, effects of post processing, rates, sensitivity, specificity & accuracy etc. are compared.

A concluding summary of this Master's Thesis together with an outlook on future suggestion completes this work.

Key Words: *Digital Forensic, Blind Cloning, Pixel-Based, Block Matching, Local Binary Pattern (LBP), Clusteri*

Table of Contents

List of Figures

List of Tables

CHAPTER 1: INTRODUCTION

The research carried out during this dissertation is divided in two phases. The first phase focuses on development of a novel and robust model, which can identify single and multiple blind cloning forgeries in a given image while second phase deals with comparison of developed model with previously developed techniques in the field of digital multi-cloning detection.

1.1 Background

A digital image can be best explained as a two dimensional array of finite digital values, known as pixels of an image. A digital image can be represented with several of its properties including Size (MxN), resolution, color, depth, hue, intensity levels, and formatting in which the data has been compressed. Size of any image is represented in pixels i.e. the two dimensional array (horizontal and vertical) of data storing blocks. The resolution of an image describes the number of total pixels in an image, and it can be calculated by multiplying number of vertical pixels to horizontal pixels. The images can be generally categorized in three types when it comes to coloring characteristics i.e. Black & White, Grey scale, and RGB images. Black and white images are composes of binary pixels i.e. 1 & 0, 1 representing white and 0 representing black. The advanced form of B&W images is grey scaled images which use shades of grey to represent different intensity levels of an image. There are in total 256 shades of grey ranging from pitch black to plain white. While RGB images utilize mixture of shades of prime colors i.e. Red, Green, and Blue for formation of an image. The resultant pixels can have millions of colors. Bit depth represents size of each pixel in different types of images. For example B&W images have bit depth of 1 bit while Grey scale images use 8 bits, and RGB images use 24 bits to represent each image pixel. *Hue* represents the scale of how "pure" or “saturated” a color is in a pixel. The intensity tells us how bright the colors are and how much they reflect the light back. Images are stored in different formats to maintain the data stored in each pixel. Each format carries its

encoding process. Some of famous formats are JPEG, BMP, and TIFF etc. Every format carries advantages and disadvantages. The selection is always based on the application where the image is to be used.

It has been found out in research that humans rely mostly on visual information for understanding and making decisions. The images are very powerful tool of communication and conveying information. The saying "A picture is worth a thousand words" is very true and we see its application every day in our life. As we can see in Figure 1.1 A Pakistani soldier is helping a kid in operation area. An essay can be written on the diverse range of messages that this image carries, but we can understand that all by just looking at the image.

Figure 1.1: A picture worth a thousand words

History of images is as old as humans are, the paintings in the caves from stone age are just one example. The images have always carried information for future generations to understand what their ancestors did to solve particular problems. In today's digital age the importance of digital images have grown to an extent where we can't even imagine our daily life to function without them. From newspapers to books from electronic media to social media, from formal events to informal ones, we can't ignore the importance of images. Digital images play their role in all walks of life in all kinds of applications including military, journalism, media, medical diagnosis, intelligence services, criminal investigation, scientific publications and surveillance systems.

Generally we believe in the information stated in any image but unfortunately in today's world the phrase that "A picture is worth of thousand words" has changed to "A picture may tell a thousand lies". The reason for that is, with the extensive development in modern technology and the quality of images, the field of photo-manipulation is also gaining mushroom like growth. These tools are not only low cost (Some even free) but also are very simple in application, and easily accessible. Most famous of such tools are Adobe Photoshop, GIMP, Corel Draw and Freehand etc. With these powerful tools even an immature can produce fake images that can change opinion of masses. These tools leave no marks of alteration behind that can be detected by naked eye and this has costed some experienced newspaper editors their jobs. The forgers employ the principal of "Seeing is believing" against the public, by changing the original details in their favor and leaving public/ viewer believing in a thing that does not exists in reality.

Keeping in mind all these facts, the need arises for development of powerful tools for authentication of digital images. Digital image forensics is the field that deals with analysis of authenticity of an image. The field is relatively new and a lot of research is needed to be carried out in this area. Digital image forensics by development of detection methodologies, analyzes the history of an image by checking different properties and characteristics and then gives the conclusion of weather the image has been altered or not i.e. image is forged or original. This allows users to take decisions and develop opinions in the light of true information [1].

1.2 Motivation

As mentioned earlier fake images are used to misrepresent information or to create negative influence and to manipulate public sentiment. Today's advanced tempering/ forging tools have changed the famous quote to "A picture unworthy a thousand true words".

The history of modern image forgery dates back to 1840 when the first fake photograph was created/manipulated by Hippolyta Bayrad. As we can see in Figure 1.2 he showed himself as committing suicide while in reality he was faking it.

Figure 1.2: First Fake Photograph

Photojournalists and forgers tamper images in such a way that they appear attractive and intense, same technique is used for propaganda images on print and electronic media where fake war heroes and fake crowds are created to change public's perception in the favor of one party. Figure 1.3 was published by Amilia Hill of "The Guardian" and is a doctored photograph, which was created to manipulate the public sentiment after death of Osama bin Ladin.

Figure 1.3: Tampered digital image (splicing of two images).

Magazines regularly use image forgery but not to cause harm, but to make their products look more attractive. This forgery is used to increase sales for all kinds of products, from cars to beauty products to food outlets. The Cars appear more sleek and shiny and stylish in magazines, the Big Mac appears much bigger and delicious than it actually is, and the model appear prettier, skinnier, and blemish free.

In Figure 1.4 the models face on right was enhanced using Photoshop by reducing her jaw width and neck size, by removing wrinkles and dark spots and by enhancing skin tone. Political parties use splicing to make their crowd looks much bigger for propaganda by copy-pasting the crowd multiple times in same image.

Figure 1.4: Tampered digital image (enhancing image smoothness).

The fields which are most dangerously affected by image forgery are medical and judiciary, since the decision if manipulated by false information in these fields can result devastatingly. Forgers manipulate medical reports to get jobs or to claim false insurance and alter critical evidence to change the decision by a judge in their favor. The alarming part is that the forgers are easily accessible and one can manipulate the documents with great ease at low cost. The Figure 1.5 shows a fake medical certificate of a student who forged the document by merging signature

and stamp from a real certificate to his created document. Such medical reports are also produced to hide age or mental conditions in court of a criminal to avoid serious sentence.


SWAZILAND GOVERNMENT

THE IMMIGRATION ACT, 1982

MEDICAL CERTIFICATE

DATE: 7 August, 2001

PLACE: Mbabane

I hereby certify that I have examined Jordan Mpofu

and find that he/she is not mentally or physically defective in any way except

that he/she is not an idiot, epileptic, insane, mentally deficient, deaf and dumb, deaf and blind or dumb and blind and that he/she is not suffering from leprosy, tuberculosis, or trachoma.

SIGNATURE

NB: Any physical or mental defects should be stated with an indication of their nature and extent.

Figure 1.5: Tampered digital image (composite of two images)

Digital image forgery is a nightmare for celebrities and public figure who's fake images are created to defame them, many recent examples can be seen in local media and social media reports which were found to be fake later on. These fake images also boost ethnic and sectarian divide by proving that one group is oppressing the other.

The United States research integrity department has reported that there was less than 3% of fraud involving doctored images in 19th Century. The percentage went up to 44% by the end of 2007. This alarming condition demands a need of reliable image tampering detection system.

This thesis aims to target the above mentioned problems by developing tools and techniques that can readily counter the digital image forgery and give a final judgment on authenticity of an image so that one day these tools can be used in court of law which still does not accepts images as a proof.

1.3 Research Objectives

The main intentions of this research thesis are as follows:

- Develop a novel and robust Algorithm to detect multiple cloning forgeries.
- To merge this algorithm with Copy-Paste Algorithm for a powerful tool to be introduced.
- To draw a comparison between previously developed techniques and proposed technique.
- To publish Research publications using the outputs.

1.4 Organization of the Thesis

The taxonomy of the thesis is structured as follows: In Second chapter a detailed description of Digital Image Forensics, its types & algorithms developed so far will be discussed. Chapter 3 is related to the introduced methodology based on pros and cons of previous techniques described in chapter 2. Chapter 4 and 5 deals with the simulation result and comparison with previous techniques respectively whereas Chapter 6 concludes the dissertation.

CHAPTER 2: LITERATURE REVIEW

In this chapter we will try to understand what forgery is actually and what are its different types and what kind of techniques have been recently employed to detect blind cloning forgery. Researchers in the field of digital forensics consider it an important topic and have been paying more attention to it recently, which has resulted in development of some state of the art algorithms. Yet this area is comparatively new so the scope is limited and relatively fewer techniques exist in this domain.

2.1 Digital Image Forgery

What makes an image forged? If some content is added or removed from an image, changing its original details, whether the details of this stipulation are visible to naked eye or can be detected via computer software, the image will be called as forged or tempered image. Tempering in terms of can be defined as transformation of some pixel values in an image to new values in such a way that this transformation is unperceivable.

Digital image forgery detection has recently become a fast growing field amongst researchers, and it deals with development of authentication techniques of an image to intact our trust on digital images [2,3].

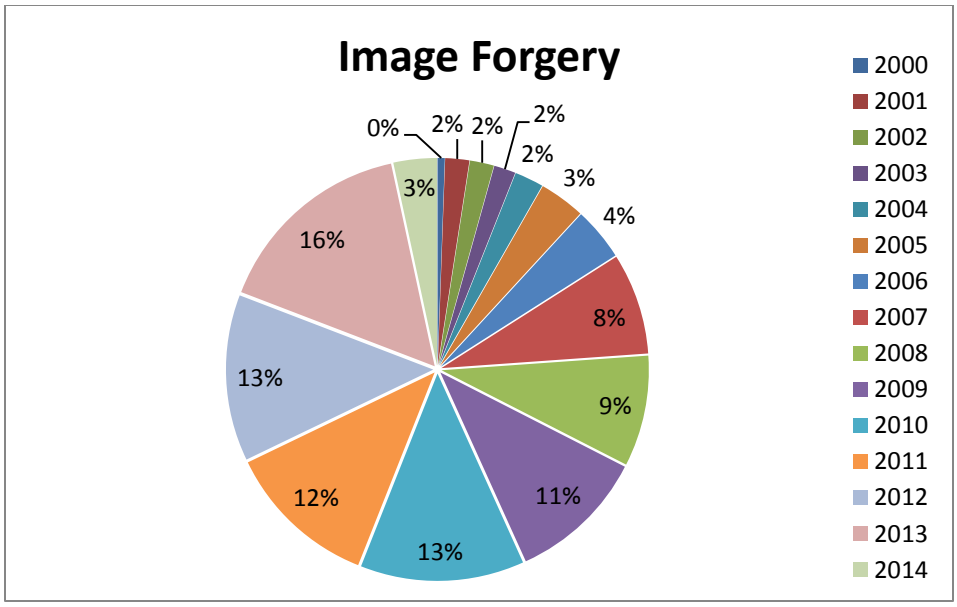


Fig2.1. Pie Chart of number of publications in Field of Digital Image Forgery

In recent history, tempered images have caused some serious problems on social, moral, legal and ethical grounds. These images have initiated anarchy amongst different groups, caused some serious legal problems and created medical complications. In our country where society is very polarized such images can create havoc religiously and politically. Therefore, the need for development of detection techniques is present now more than ever. The various types of detection techniques have been developed in recent years [4,5]. The chart in Figure 2.1 shows the number of publications published on this topic in different reputed international conferences and international journals.

2.2 Techniques of Detecting Digital Image Forgery

There are several algorithms to detect forged images but there general categorization is into two groups i.e. Active detection approach and passive detection approach. Further classification and complete hierarchy is shown in Figure 2.2.

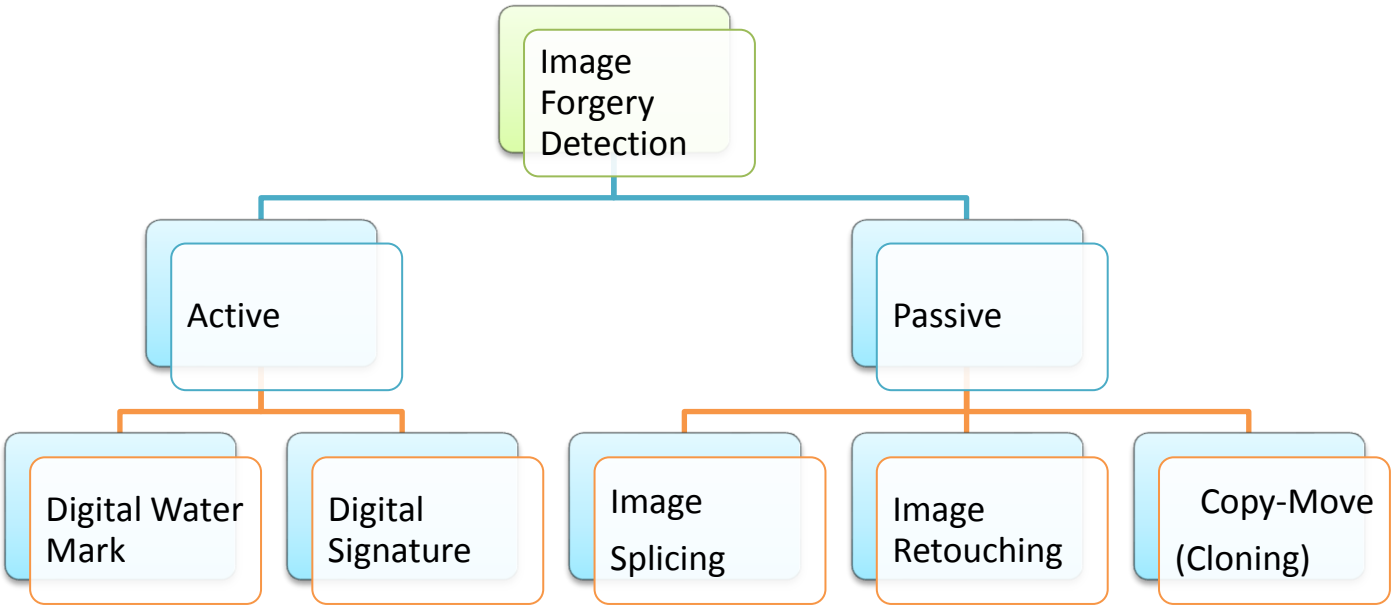


Fig2.2. Hierarchy of Image Forgery

2.2.1 Active Approach

In the active approach we use a known embedded code or we have an original image for the judgment of faked image that's why this approach is also known as informed approach. The originator of the image embeds a watermark (visible to eye) or digital signature (not visible through naked eye) in the image so that if some forging is done the watermark or signature changes its shape and when analyzed the verdict can be given on authenticity of the image. In the figure 2.3 the watermark has been installed by the photo studio while in figure 2.4 the part (b) depicts the secret signature while part (a) shows how the picture actually looks to the viewer.

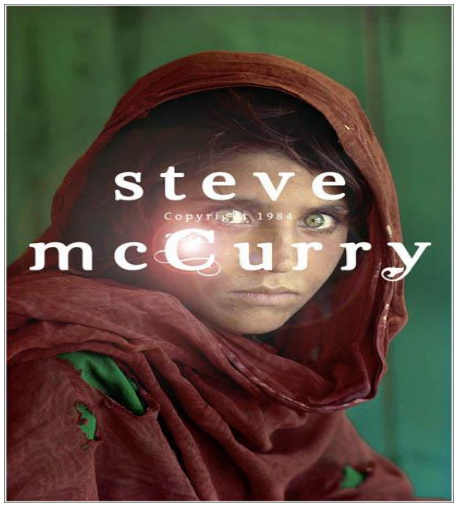


Fig2.3. An example of Digital Water Mark

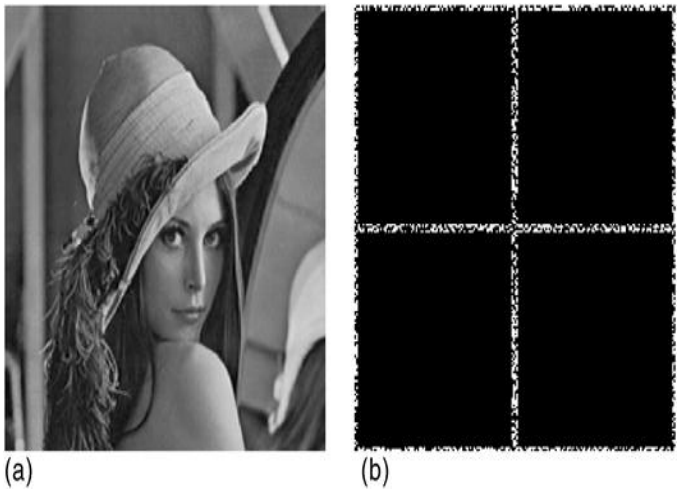


Fig2.4. An example of Digital Image Signature

The Active techniques have very limited scope (used for copy-right purposes) and can be applied to limited number of images where originator implants his mark. These techniques can also be applied on images taken from modern cameras which embed their noise signatures in images. But generally the images contain neither watermark nor digital signature. This arises need for special techniques that can cope up with those images.

2.2.2 Passive/Blind Approach

Unlike what is done in active approach, passive approach is used in the absence of any previous information about the image. That's why this approach is also called unknown image origin analysis. Passive approaches use algorithms that look for statistical changes and/or duplications in the image [6].

The research conducted in this field (passive techniques) is growing but it's still in the initial phase. Therefore we would like to further elaborate the work done in this regard.

2.3 Passive/Blind Techniques

As depicted in the figure 2.2 earlier the hierarchy of image forgery detection techniques is classified into active and passive approaches. The passive or blind techniques are further classified into three categories. Copy-Move (Image Cloning), Copy-Paste (Image Splicing), and Image Retouching/Enhancement.

2.3.1 Image Enhancement or Retouching

In this kind of forgery the forger enhances or reduces some features from the image to make it look more attractive. This kind of forgery is so common that every second image that we see on TV, magazines, papers etc. have gone through this forgery. This kind of forgery is also known as image retouching. The forger changes the features like brightness, sharpness or may apply smoothing filter on certain portions of the image. This kind of forgery is applied in advertisements to make the models look skinnier, and prettier than they actually are. This technique is becoming popular amongst teens especially after advent of social media. Where people like to show themselves as perfect as possible. So the apply changes to their facial features in easily available software like Piccasa, Photoshop, Coral draw and many more. These software use special filters to alter the image and make it look more attractive to human eye and brain [8].

An example of such forgery can be seen in figure 2.5 where wrinkles have been removed from lady's face and her fairness has also been improved. This makes her look less aged. The image retouching is considered naive form of forgery.

Fig2.5. An example of Digital Image Enhancement

2.3.2 Image Splicing or Copy-Paste

When forger gathers several parts form different images and forms a different individual image, the forgery is called as image splicing, photo montage or Copy-paste forgery. The parts gathered form different images are known as Regions of Interests (ROIs).

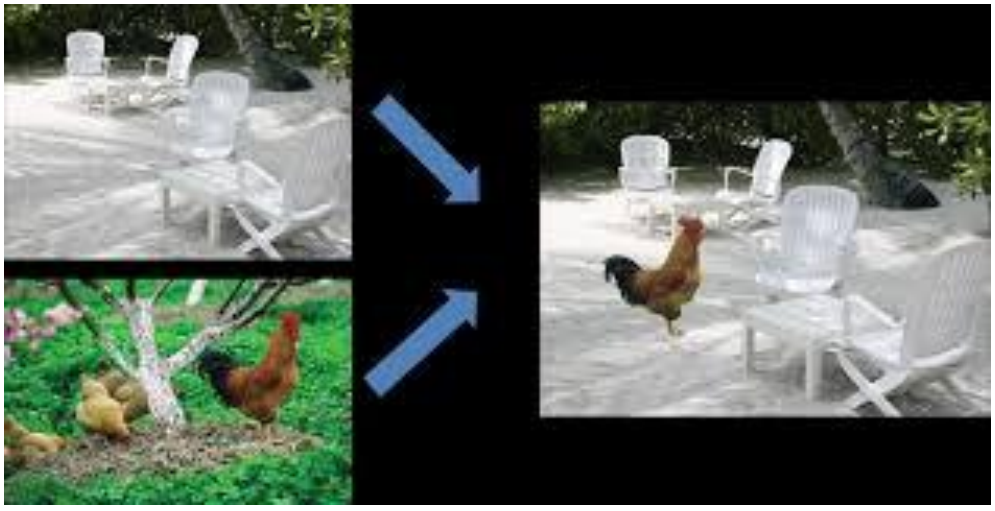


Fig2.6. An example of Digital Image Splicing

Fig 2.6 shows a how forger has taken hen from one image and then implanted in the other image to form a new image. This technique is really dangerous because it can be used to create serious propaganda images, as we have seen some in past made by agencies to defame a leader or community.

2.3.3 Cloning /Copy-Move Attack

The Copy-Move forgery is similar to Copy-paste in a way that image portions are combined to form a new image but the difference is that the portions are used from the same image instead of using multiple/external images. This type of forgery is done to hide some detail in the image by copy and pasting another region from the image. This type of forgery is really hard to detect because the forger leaves no trace to detect forgery visually and the computer software monitoring the difference in pixel density and intensity differences will be unable to detect any changes because ROI will be from same image.

An example can be seen in figure 2.7 where bushes from the image are duplicated and pasted to hide a military vehicle using forgery software.

Fig2.7. An example of Copy-Move Forgery

The cloning detection algorithms use certain key features or blocks and then match them to detect if two or more parts match, declaring if the image is forged or not.

In my research the cloning forgery is by far most advanced form of forgery and the algorithms are weak to detect it completely. Specially there have been only few publications regarding multiple-cloning detection. The digital image forensics and specially copy-move have becoming interest of many researchers, so I have decided to carry out my research in this paradigm of digital image forensics.

2.4 Types of Cloning Detection Techniques

The cloning forgery detection techniques are further classified into four sub-domains .Fig 2.8 shows the classification of detection algorithms.

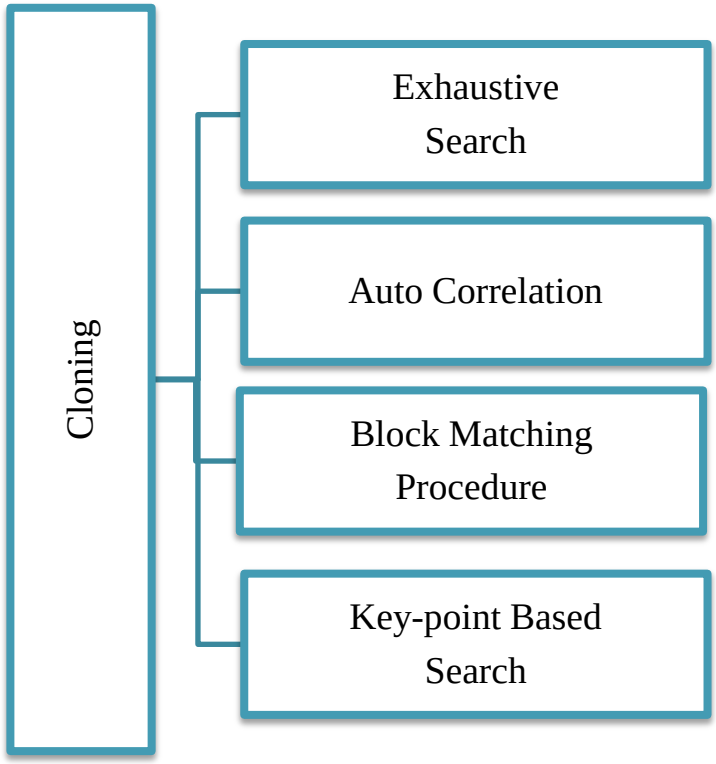


Fig2.8.Distribution of Copy-Move image Forgery

2.4.1 Exhaustive Search

Exhaustive search detection method uses the given image and then a circular shifted version of the same image is placed above the given image and subtraction is applied. If results are not close to zero in any portion the process is repeated again by shifting more pixels. Where ever the results are close to zero it means that the regions have matched. This methodology works fine only if the given images are of small dimension. Because for every MxN image, the number of computations will rise to $(MxN)^{2.}$ So the technique is highly un-recommended for medium and large sized images.

Fig2.9– An example of Exhaustive Search Technique

Figure 2.9 shows how this method works. On left side there is the given image whereas on right side the circular shifted version of the image is available.

If we consider a gray scale version of image of size M*N, and x is pixel under consideration i,j defines the position of the pixels then using the formula

$$\left|x_{ij} - x_{i+p} mod(M) j + q mod(N)\right| \qquad \text{p=0,1,…, M , q=0,1,..N for all i and j} \qquad (2.1)$$

We can calculate the areas copied and moved each time.

2.4.2 Autocorrelation

In this method the image is further divided into non overlapping blocks than each block is moved over the whole image as a filter. The block center is placed over each pixel and then subtraction is applied. The results are matched with a pre-selected threshold. If the resultant is below that threshold, the matching is declared.

An example of this technique can be seen in figure 2.10. Consider an image I of dimensions M*N, I(x, y) is the position of the pixel in an image where a and b are the pixels of pasted region then H (a, b) shows the auto-correlation function.

$$H(a, b) = \sum_{x}^{M} \sum_{y}^{N} i(x, y) * i(x - a, y - b) \qquad x=0,1,\ldots,M-1\ y=0,1,\ldots N-1 \qquad (2.2)$$

a) Original Image

b) Tampered Image

Fig2.10.

c) Auto Correlation computation

d) Detection result

This technique is not very accurate because the blocks are non-overlapping and it's hard to detect an accurate region around which forgery has been done. Also this technique fails badly in plainer regions like sky, lawn etc. Autocorrelation and Exhaustive search have become obsolete due to their computational expensiveness and inaccuracies.

2.4.3 Block-Based Techniques

This technique is advanced form of autocorrelation in a way that it also uses blocks for forgery detection but the blocks are first of all overlapping, secondly rather than application of simple subtraction operation, certain features from each block are extracted. These features are than matched with features from other blocks. A distance threshold is selected and every time a match is detected a distance check is applied on basis of the threshold, before announcing forgery. The block matching is further classified into several techniques depending upon which feature each technique uses. The general algorithm of detecting forgery using block matching is given in Figure2.11.

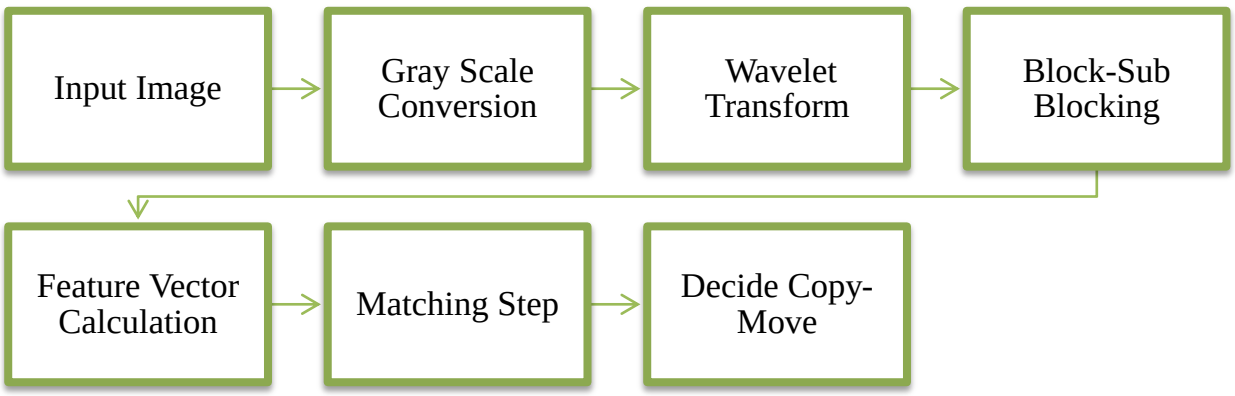


Fig2.11– Algorithm of Block Based Technique

2.4.4 Key-Point Based

The key-point based techniques are different in a way that rather than dividing image into blocks we pinpoint points on the image based on certain features, these points are known as key-points. Basically the points where there are several changes (in intensity, color, etc.) are taking place are considered key-points. The feature that makes the key-points special is that they are described in form a feature vector and this vector is robust against post processing operations. Several key-point based techniques have been developed so far but they all follow the algorithm described in Figure 2.12

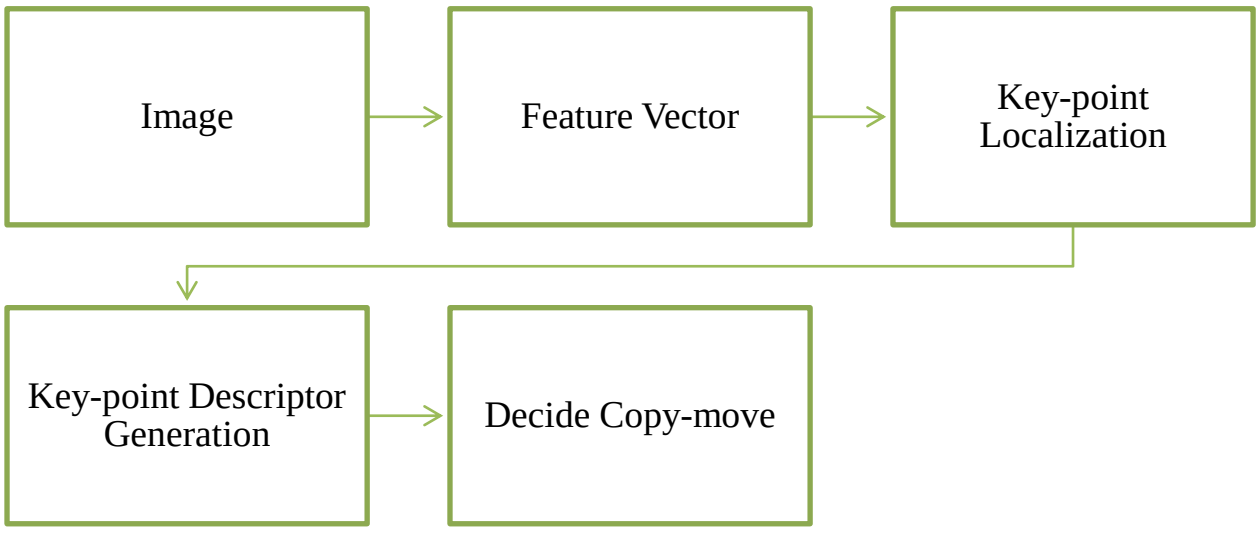


Fig2.12– Algorithm of Key-point based Technique

2.5 Related Studies

A number of significant researches have been done in the area of digital image forensics in the present century. A number of algorithms have been developed for different types of digital image forgery detection to recover people's confidence toward the reliability and trustworthiness of digital images.

It is clearly obvious from the graph in figure 2.13 which shows there had been noteworthy focus in the area of digital image forgery in major journals and conferences in past few years.

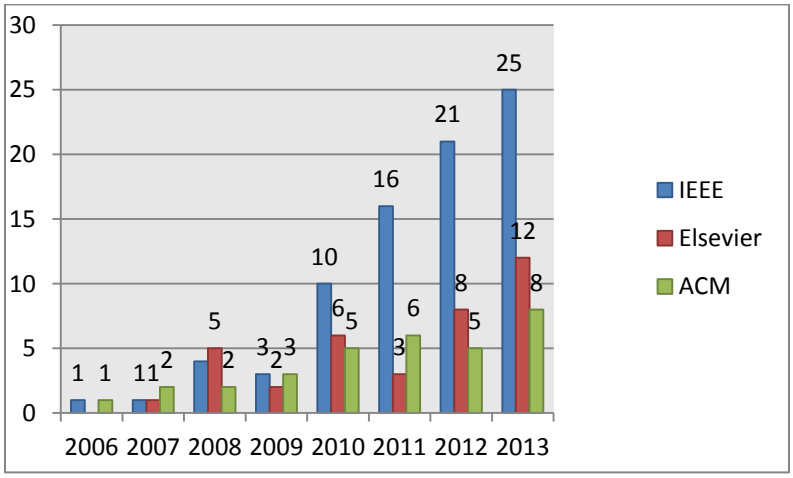


Fig2.13. Chart on Publications of Copy-Move Detection Algorithm

As discussed earlier that this research work attempts to focus on one detection technique i.e. Blind Cloning Forgery detection technique, so we would provide comprehensive literature review on this particular technique only.

2.5.1 Exhaustive Search Techniques

One method to identify cloning forgery is using an exhaustive search .A detection technique based on this particular approach was presented in Jan Lukas et al. [8]. In this technique the original image is compared with the circularly shifted form of tampered image. The differences for each shift of [p, q] is calculated by following the equation 2.1 described in section 2.4.1, the difference is then examined against a small threshold 't'. Settling threshold is a problem since there will be large amount of pixels pairs whose differences will be below 't'. This technique is effective for images of small size but is computationally expensive and impracticable for images having large size as it involves $(MN)^2$ times calculations for every single image of size MxN.

2.5.2 Auto-Correlation Techniques

An instinctive recommendation to avoid exhaustive search for the detection of copy-move forgery was Auto-Correlation. This technique implicates a correlation between the original and copied region of image. Following the equation 2.2 of section 2.4.2 several research methodologies have been introduced. Jessica Fridrich et al. [9] proposed that the algorithm can be proficiently implemented using Fourier transform. Although this technique is not computationally complex yet it so often fails to identify tampering unless the size of forged area is large .Another notable method introduced by Bo Xu et al. [10] employed the use of phase correlation. The tampered region was detected in a form of sharp peak of respective phase correlation coefficient. Although the results of this technique show that it is robust against noise and blurring and is efficient yet the false positive rate is large.

So a Block Based Cloning detection technique was introduced to detect tampering significantly

better and faster as compared to exhaustive and autocorrelation.

2.5.3 Block-Based Detection Techniques

[11] Suggested the first ever block based cloning detection technique. The technique involved extracting image features and matching with the help of Discrete Cosine Transform (DCT). It works by division of image of size M x N, into small overlapping blocks and then sliding a window k×k from the upper left corner to the lower right corner considering one pixel at each time. Afterwards, the quantized DCT coefficients are extracted from each block which represents the image features of these blocks. The quantized DCT coefficients are kept as one row in a matrix P of (M-k+1) × (N-k+1) rows and k× k columns, where k× k shows the block size. Two identical rows in the matrix "P, correspond to two identical blocks in the suspicious image. The rows of "P" are sorted lexicographically to make matching fast and effective. The shift vector is calculated for each sequential row pair and the occurrence counter of that shift vector (initially set to zero) is incremented by one. The threshold 't' is selected empirically. If the shift vector's occurrence value is greater than settled threshold then the respective block is said to be duplicated or tampered. The shift vector is suggested to solve the problem of many falsely matched blocks. Any block pair matching is taken in account if and only if there are numerous matching pairs in the similar mutual position. There is a clear drawback in this algorithm when dealing with uniform areas. Moreover, it's only tested with small number of images. The algorithm was robust in a sense that DCT coefficients comprises of high energy so any type of manipulations such as noise addition and compression will not affect energy coefficients. However above method fails if geometrical transformations are applied on image, such as rotation, scaling and affine transformation.

The second method was proposed by [12] and was quite similar to [11]. The only difference lies is the representation of image blocks. Instead of DCT, Principal Component Analysis (PCA) was employed. In this method the colored image is converted to the grayscale version firstly then divided into overlapping blocks. PCA is extracted and utilized to characterize those blocks as row vectors stored in a matrix. In order to reduce the dimensionality of the representation, truncating is used on the matrix rows. The authors stated that this method has been tested on tampered images with post processing operations such as joint photographic compression and the presence of additive noise, the accuracy of this method was found to be excellent except for small block sizes and low compression qualities. It's also stated that the method will be extended to work with RGB colored images. This method is efficient in a sense that it gives reduced dimensional representation of an image.

Numerous other block-based cloning detection algorithms have been developed previously. Some of them are based on Singular Value Decomposition, kernel-PCA or Discrete Wavelet Transform. Although these techniques can effectively detect copy moved attack region yet it needs more computational time and memory. If we focus on block-based techniques in the last five years following remarkable techniques have been developed.

An efficient block based forgery detection methodology was proposed by [13] in 2010. The technique involves Discrete Wavelet Transform (DWT) for reduction of image dimensions. The image is further divided into overlapping blocks. After application of DWT, phase correlation has been used to detect similar blocks. This process results in generation of duplication map that gives the number of forged pixels. The algorithm not only reduces time complexity but also improves the accuracy of detection. This approach also works well with relatively difficult post processing forgery operations like changes in JPEG quality levels and white noise. The method

is fragile against rotation and scaling operations on copied region.

Another effective methodology focusing on block based detection technique, which also resulted in most accurate results has been presented by [14] in the year 2010. This method was developed to detect copy-rotate-move forgery using Zernike moments which possesses qualities like invariance to rotation, robustness against noise etc. The average resultant accuracy is around 83.59%. Even against rotations of 30◦ on copied section the algorithm produced accurate results. Along with simple copy move operations some other post processing applications like JPEG compression, blurring effects and white Gaussian noise were also tested. This algorithm showed limitations when forger did scaling and affine transformation operations on copied region. The algorithm proposed usage of efficient data structures to make it computationally inexpensive.

An efficient and robust algorithm compared to the techniques presented above was presented by [15] in 2011. This method targets more efficient features while using DWT sub-blocking technique. The technique is efficient in terms of time complexity and can deal with slight rotation operations along with application of JPEG compression and Gaussian noise operations. The weak spots of this algorithm are arbitrary rotations and scaling operations.

[16] Targeted dyadic wavelets to develop their technique in 2012. Un-decimated dyadic possess the quality of shift invariance that's why they were chosen for this technique. The input image was processed with DyDWT and only LL1 and HH1 bands of transform are moved forward for further use. LL1 sub-band is used for approximating the original image while HH1 sub-band deals with the noise encoding of the image, which has been supposedly disturbed during forgery operation. Then both sub-bands are further subdivided into blocks of 16x16 pixels, than Euclidean distance was calculated amongst all blocks pair by pair in both LL1 and HH1. The distances of LL1 are arranged in ascending order while distances from HH1 are arranged in

descending order. The blocks which appeared on same location in both lists is considered forged (Copied-Moved). The technique avoids false positives because HH1 declines uniformed regions as similar. This technique was tested with and without rotation operation applied of copied region and also after application of different JPEG Q levels. This methodology has been found as more efficient than previous techniques.

The authors in publication [17] have chosen the detection technique that deals with luminance domain of the image. Since high frequency components are not robust against signal processing operations, the importance of low frequency components rises in feature matching. That's why a low-pass filter, like Gaussian low pass has been chosen to eliminate high components of image frequency. After filtration the image is divided in to circular blocks since it is helpful in keeping contents consistent even after rotation operation has been applied. Polar-sine Transform has been used to extract the features. Then comes the sorting of feature matrix. After this the Euclidean distance amongst the blocks is calculated and blocks a threshold is set. The blocks having less distance than this threshold are moved further. This technique has shown robustness against different post processing operations like scaling, rotation, scaling coupled with rotation, flipping of regions and some affine transformations like perspective projection and shearing along with signal processing operations including JPEG compression with different factors, white Gaussian noise, and Gaussian blur etc.

An automated and effective detection technique was proposed in [18]. This approach uses DWT along with Fast Walsh-Hadamard Transform (FWHT) for feature extraction. Walsh-Hadamard Transform is composed of Walsh coded Nth order matrix. This helps in reduction of required computational power. The use of FWHT effectively recues false positive detection. This method also works well against post processing operations like added Gaussian noise, distorting and

blurring of images. The technique worked well against 75% of duplicated blocks with rotation operation was detected. Also this method has accurately dealt with JPEG compression with Q factors between 50 and 100.

In [19] feature extraction from blocks has been done using two dimensional Fourier Transform. A predetermined set of Fourier coefficients collects data from blocks. In the end, the matching between adjacent feature vectors is performed to determine the forgery. Results highlight that this method can identify duplicated regions with higher accuracy even if the post processing operations like blurring or JPEG compression (for even qualities lower than 50) have been applied. The size of feature vector has also been reduced in this technique. So this method ensures higher accuracy while keeping smaller vector size. It is also good against multiple cloning forgeries.

A relatively recent methodology which concentrates on block expansion was proposed by [20] in 2013. Direct block comparison is done instead of block feature comparison. The input image is subdivided into overlapping blocks along with a dominant feature (it averages the grey level values in that block) is calculated for each block. All blocks are sorted than buckets are formed to group them, each bucket containing blocks with similar features. The blocks in a bucket are compared with other blocks in same bucket. A mean value based statistical hypothesis is performed on pixels. Elimination is performed on block if it doesn't matches any other in the bucket. Comparison begins with smaller regions and continues to expand while searching for possible matches. With the expansion in region the number of blocks reduces in the bucket. The remaining blocks are considered as copied region. This technique was further tested in comparison to other techniques based on DCT, DWT and PCA on images containing blurring and JPEG compressions on different levels along with irregularities in shaped region. This

technique was also able to detect illumination operations performed on copied region. This type of forgery was never dealt before.

2.5.4 Key-point based Detection Techniques

Quite lot researchers have utilized the method of key-points matching in order to recognize copy-moved attack section.

Xu Bo [21] proposed a fast robust copy-moved detection method based on the Speed up Robust Features (SURF). The technique extracts the image key-points by employing fast hessian detector and then the possible cloned regions were found by matching the key-point descriptor vectors. The performance analysis of this technique showed that it can detect cloned region effectively as well as efficiently.

An automatic and accurate cloned detection technique was proposed by Li Jing [22]. The proposed technique employed the use of speeded up robust feature (SURF). The given image was converted to an integral image by recursive computation and hessian matrix was employed. Besides SURF the algorithm also used binary k-d tree for effective key-point matching. The k-d-binary tree is ordinarily used binary tree structure for storing key-points and examining the nearest neighbors. Although this technique productively detects tampering yet it was unable to detect smaller copied regions.

[23] Suggested an innovative algorithm based on Scale invariant Feature Transform (SIFT) features in order to detect and locate cloned regions. The SIFT features descriptors are invariant to changes in any geometrical distortion or illumination. The method was found to be effective and efficient, compared to previous key-point detection techniques based on SURF. The suggested method also illustrates helpfulness in images subjected to splicing, montage and

multiple cloning. However, the technique was found to be frail against regions having clusters or which had high uniform textures and flat surfaces.

Another method based on key-point detection was introduced by [24]. Though the method was similar to previously used key-point techniques which were based on SIFT yet it employed a Scale Rotation invariant Pattern Entropy (SR-PE) in order to successfully achieve the matching pattern efficiently. After localizing key-points and creating feature descriptors the detected feature points were matched using SR-PE. It measures the spatial regularity of matching patterns of key-points. In the first phase an exhaustive approximation of parameters for all pairs of matching point is carried out. Afterward these parameters are assembled using the mean-shift technique. The technique was found to be efficient and robust even in the presence of degradation and compression.

The next cloned detection technique was presented by [25]. The methods employed SIFT to detect copy moved region. The image features were extracted using SIFT instead of matching through feature descriptors vectors, generated by SIFT the matching of image features was based on KD- tree and Best-Bin-First method. The k-d tree is a binary tree that stores key-points in k-dimensional space, than BBF is employed to search closest neighbor. This algorithm identifies neighbors using a limited amount of computations. Though the computational complexity of the suggested technique was similar to the block-matching detection methods yet it had improved detection accuracy. The experimental analysis showed that this method can detect copy-moved regions effectively, even when these areas are affected with geometrical distortion.

A methodology was proposed by [26] to overcome the weakness of previously developed techniques. The technique was based on extraction of image features by utilizing scale invariant feature transform algorithm and then application of Zernike moments. SIFT effectively extracts

the image feature points even in regions that were geometrically modified like scaled or rotated. However it fails to detect feature points in areas which are flat so this problem was solved by utilization of Zernike moment. The given image was binarized as Zernike moments are defined only over on unit disk having radius r. After binarization the shapes were resampled in order to normalize the size of image to 2r*2r. The Zernike moments are extracted from the developed normalized image. Different moments of order of Zernike are extracted. The magnitude of Zernike moments is used as a descriptor to extract features from flat surfaces. Zernike moments have remarkable properties i.e. they are invariant to scaling, rotations and robust against noise whether it is Gaussian or additive white noise. Moreover besides detection of key-points in flat region the utilization of SIFT and Zernike combinatory also increases robustness, precision and helpful in time reduction.

Another effective cloned detection method which employed the use of speeded up robust features (SURF) and hierarchical agglomerative clustering (HAC) was proposed by [27]. In this algorithm the image features and their corresponding feature descriptors was perceived using SURF. Although this key-point extraction technique is efficient as compared to scale invariant feature transform yet it has computational complexity. HAC or hierarchy of clusters is then employed to form the groups or sets of key-points. In this algorithm at the beginning every extracted feature point behaves as single cluster. Afterwards the distance between every key-point is calculated and checked against a certain threshold. If the key-point clusters are found to be different they are merged by using averaging or ward linkage. This process continues until all the key-points are grouped into a single cluster. The performance of this algorithm is found to effective as compared to previous techniques which were based on either SIFT or SURF. This technique is helpful in detecting multiple cloned region in an image and robust against

degradation like noise or geometric modification.

Recently [28] employed the use of an effective key-point based technique named as Mirror Reflection Invariant Feature (MIFT). Although this technique is much similar to scale invariant feature transform yet it has some additional properties i.e. it is also invariant to mirror reflection modification which is helpful to estimate if any geometrical modification has been applied to the cloned area. This technique is also effective in improving the detection and localization of copy-moved region as compared to SIFT. Reason being the use of hysteresis threshold as compared to empirically settled threshold which results in accurate detection. Hysteresis thresholding is based on selection of two threshold; low and high. Low threshold is opted for weak regions and high threshold is opted for strong pixels i.e. edges etc. Afterwards erosion and dilation is also employed to further reduce the false positives and negatives also the utilization of an increment iterative scheme to find key-point match helps in correct localization of cloned areas. For this purpose cross correlation algorithm is generally employed which is helpful in correct localization. This technique is supportive for small sized images or images in which the cloned region is very small.

Though the proposed technique is effective and accurate in many ways as compared to SIFT yet it cannot extract or specify key-points in areas having flat regions.

CHAPTER 3: METHODOLOGY

This Chapter briefly describes the proposed technique for detection of Blind cloning forgery. A brief detailed review of previously developed methodologies in Chapter 2 leads into further exploration and development of an improved technique. A comprehensive description and explanation of each step of the recommended methodology is given.

3.1 Introduction

After the extensive analysis & over view of existing state-of-art cloning detection methodologies in Chapter 2 we find that all the algorithms aim for the same goal, i.e. to detect copy-move forgery however there are some aspects where they differ considerably. So based on the pros and cons of previous techniques the proposed methodology should have four elementary characteristics: less false positive rate, robust, efficient and small computational time.

3.2 Algorithm

The schematization of proposed methodology is shown in Figure 3.1. A detailed description of each block is given in the following subsections.

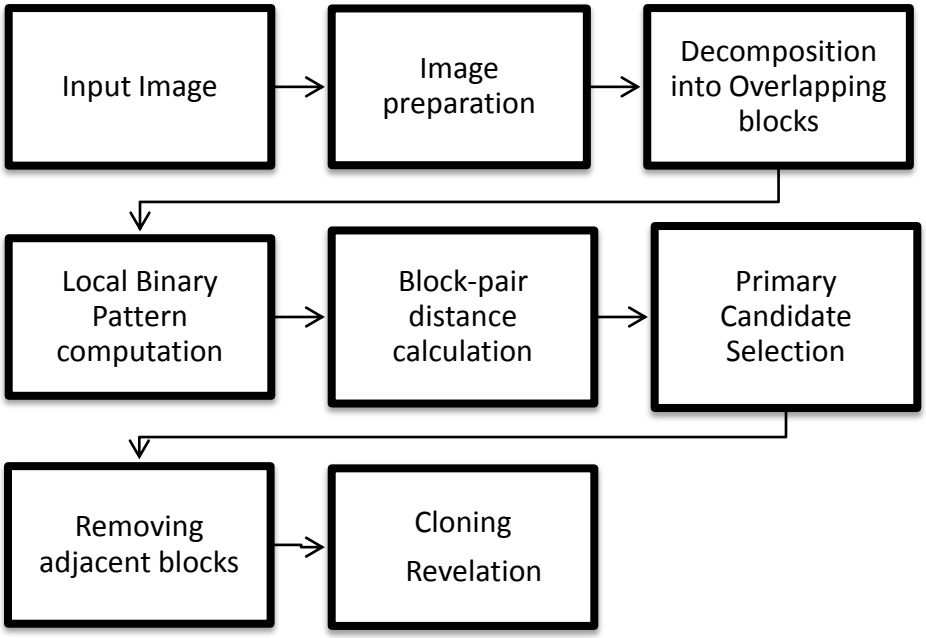


Fig3.1. A layout of Proposed Methodology

3.2.1 Input Image

An RGB image of unknown origin is input to detect the presence of copy-paste forgery.

3.2.2 Image Preparation

This stage is also known as the pre-processing stage, in this stage the color image is first decomposed into three color (red, green, and blue) components. The purpose behind decomposing image into components is to exploit information present in different color

components. As we have surveyed previously developed methodologies, nearly in every algorithm first step was to convert color image into gray-scale by using approximated weights of R, G & B. During this conversion, some weak but important bits of tampering can be lost. So in this proposed algorithm the color image is converted to components only.

3.2.3 Decomposition into Overlapping block

Next stage is to decompose the tampered image into overlapping block. In this step each color component is divided into square overlapping blocks of specified pixel size. The decomposition is done by following rule: State the specification of tampered image, such as, the size of block n, the no. of pixels we have to shift for overlapping let say p and number of blocks in neighborhood m. So each component is divided into square overlapping blocks of size n × n pixels. The overlapping is of size p × p pixels; if p = n, then there is no overlapping. The block size in this algorithm must be maximum one-third of the forged area. In this theory, if the forged area size is less than the size of three blocks, then the algorithm cannot detect the forgery. For example, if the algorithm uses 20 × 20 block sizes with overlapping of order 10, it can detect forgery larger or equal to 60 × 60.

3.2.4 Local Binary Pattern (LBP) computation

After decomposing multi-channel image into overlapping block the next step is to compute LBP histogram of each channel. LBP is basically a powerful visual descriptor operator used for in computer vision for texture classification. The algorithm operates by thresholding the defined image pixel i.e. m × m pixels in a cell here m=3; compare the pixel to each of its 2^3=8 neighbors

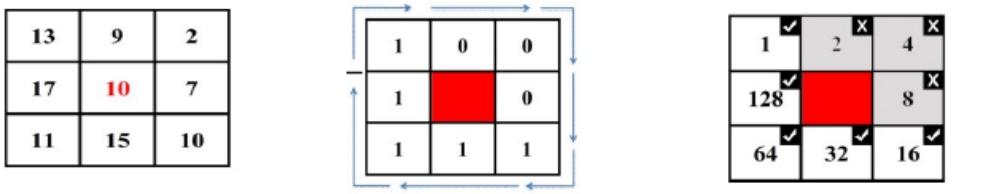
(left-top, left-middle, left-bottom, right-top, etc.). Follow the pixels along a circle, i.e. either clockwise or counter-clockwise. If the value of center pixel is greater than the neighbor's value, indicate it with a zero otherwise one. Compute the histogram, over the cell, of the frequency of each binary digit which is occurring, this histogram can be demonstrated as a 2^8=256-dimensional feature vector. Normalize the histograms of all cells. This gives a feature vector for the entire window which can be used as texture spectrum.

Given a pixel (x,y), LBP can be expressed mathematically as:

$$\text{LBP} = \sum_{n=0}^{n-1} 2^n (i_n - c)$$

Here, n=8 neighborhood pixel of cell, c is the center pixel value of the cell and i_n designates surrounding 8 pixel-values.

The center value is deducted from each of its eight neighbors. If the deduction outcome is positive or equal to zero the resultant block is designated in binary format as 1 otherwise as 0. The eight binary values are normalized either clockwise or anti-clockwise to form an 8-bit binary number as shown in figure 3.2. The corresponding decimal value of the generated binary number is then used as a label for the given pixel.



a. Example of pixels in cell **b. Thresholding results in binary 10001111** **c. weight computation of cell**

Fig3.2. Calculating Local Binary Pattern

3.2.5 Block-pair distance calculation

A wide variety of methods have been developed in the past to compute the dissimilarity (distance) between the blocks of the image as discussed in literature review. Of all these techniques our focus would be on Euclidean distance and city-block since the outcome is simple and is compatible to be used in the cluster process. The procedure uses best bin first criteria to find neighborhood or nearest Euclidean distance is rationed to the next Euclidean distance, if the percentage is less than a set threshold, then this pair is accepted as matched. An appropriate threshold is selected to reduce the number of incorrect matches.

Given a test image I(x, y) having set of key-points K= {k_1, k_2,…, k_n} with their corresponding key-point descriptors {j_1,j_2,…,j_n} using Euclidean distance formula

$$D(i, j) = \sqrt{\sum_{k=1}^{N}(i_k - j_k)^2}$$, N=number of elements

So D= {d_1, $d_{2,\ldots,}d_n$} gives a sorted Euclidean distance of key-point w.r.t to other key-points. Let

T be a settled Threshold, The key-point will be matched only if $d_1/d_2 < T$.

However, many false key-matches would be formed if the threshold is not selected accurately So there is a need to avoid these false matches by some approach to increase the robustness.

3.2.6 Primary Candidate Selection

The above technique is repeated to all the three color components i.e. R,G & B. The retained block pairs are checked across all the color channels. If a block pair is present in all these components, we put it in a separate list, which is called primary candidate list. Only one-fifth (assuming trial and error) block pairs are retained in this step. The entries in this list are the candidates of copy-move, because the existing block pairs have similar texture in all the three color components. However, after this step, we may find some false positives (non-duplicated blocks are detected as duplicated blocks) due to homogeneity in the image content. In the next step, the primary candidates are refined using neighborhood clustering.

In the first attempt we will set a threshold t and will start following the algorithm steps described before. Due to geometric modification there is a chance that Bd of each of block pairs may not be equivalent , so we will look up for the block difference which exceed the settled threshold. If the differences exceed the threshold the algorithm is preceded else the blocks are classified as replicated or forged blocks.

3.2.7 Cloning Revelation

After the application of proposed algorithm the last step of the method is to fill the copy-move blocks with black or white color for sake of visualization. If any black /white blocks are spotted, it means that the input digital image has cloning forgery; moreover the position of block will clearly indicate the part of tampered image and rest as unaltered.

CHAPTER 4: Experimental Outcomes and Discussion

This particular section will cover the test results of the techniques discussed in chapter 3. The quantitative performance of proposed copy-move forgery technique will be evaluated through simulation experiments on dataset acquired from internet.

4.1 Data Set

To examine the performance of different blind cloning detection techniques a bench mark Data Set MICC-F220 is used [29]. It was released in 2011 by I.Amerini. This Data set is permitted by IEEE Transactions on Information Forensics and Security. It comprises a total of 220 images out of which 110 are authentic or original images whereas 110 are forged or tampered images. These images constitute photos of animals, scenery, plants, characters, texture and architecture. Figure 4.1 demonstrates some of the images include in this Data set.

Figure 4.1: Examples of images form database

Here the top row exhibits the authentic original images whereas bottom row is of tampered images.

The description of images present in data set in term of types of forgeries is shown in table 4.1.

Table 4-1 Properties of images of data set

Number of images	Original	Just Cloned	Translation	Scaling s[0.8,1.5]	Rotation θ=[0°,360°]
110	√	x	x	x	x
20	x	√	x	x	x
10	x	√	x	√	x
10	x	√	√	x	x
10	x	√	x	x	√
10	x	√	√	√	x
10	x	√	√	x	√
20	x	√	x	√	√
10	x	√	√	x	√
10	x	√	√	√	√

4.2 Experimental Configuration

In order to implement the proposed techniques discussed in previous section MATLAB R 2012a is employed with a machine having an Intel Core i5, 3.2 GHz processor with 4 GB memory. The reason of using MATLAB as compared to other high level programming language is that it is highly optimized in handling matrices and we already know that digital images are composed of two dimensional matrixes.

4.3 Experimental Result for Cloning forgery

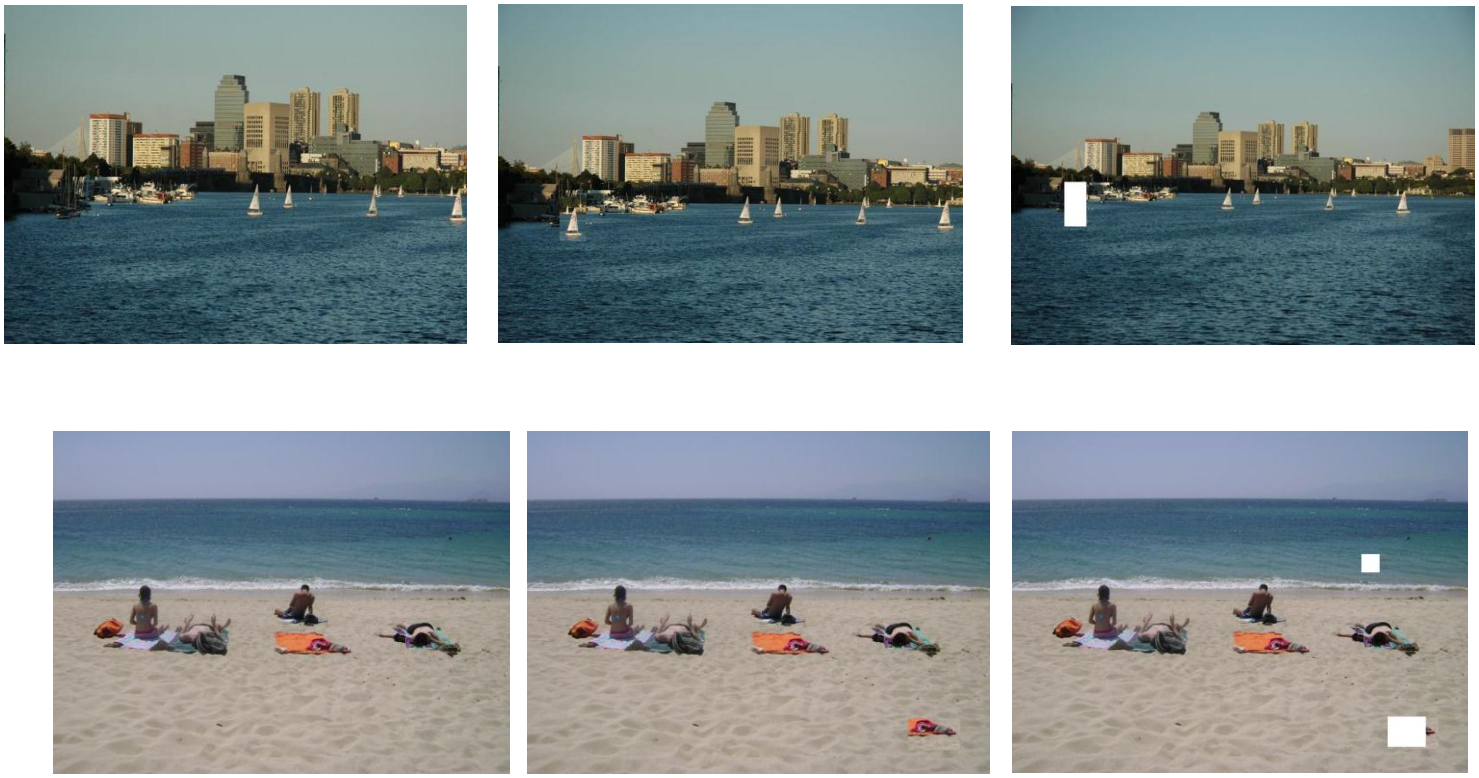


Figure 4.2: Examples of cloning detection

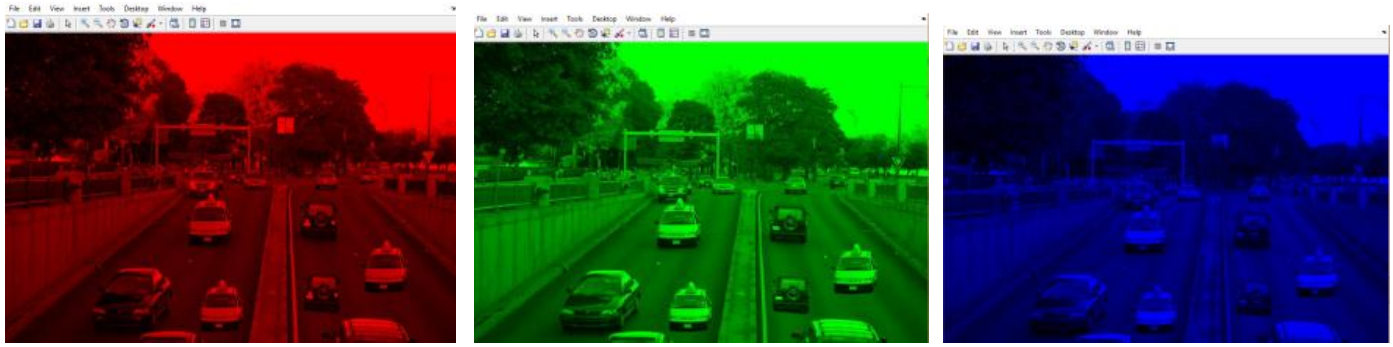
In this section we will also observe the quantitative performance of the technique proposed in section 3 through simulation experiments on the tampered images of data set. Following the proposed methodology we randomly select authentic images from data set. Afterwards a small portion of pixels from same image is copied and then pasted to new location. It should be kept in mind that no further geometric modification like scaling or rotation should be done to the moved region. Some example of cloned detection using this algorithm is intimated in following figure 4.2.

The results of algorithm following the steps of methodology described in lay out 3.1 are shown

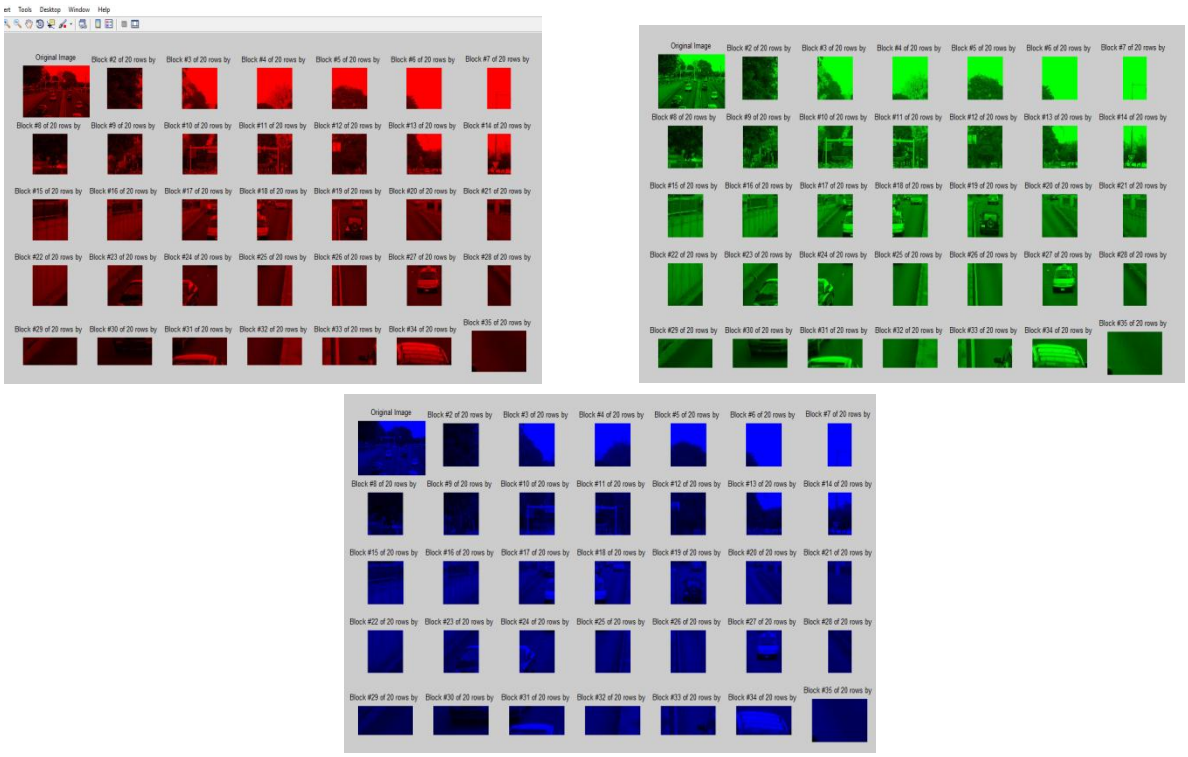
in following figures.

a) Input image

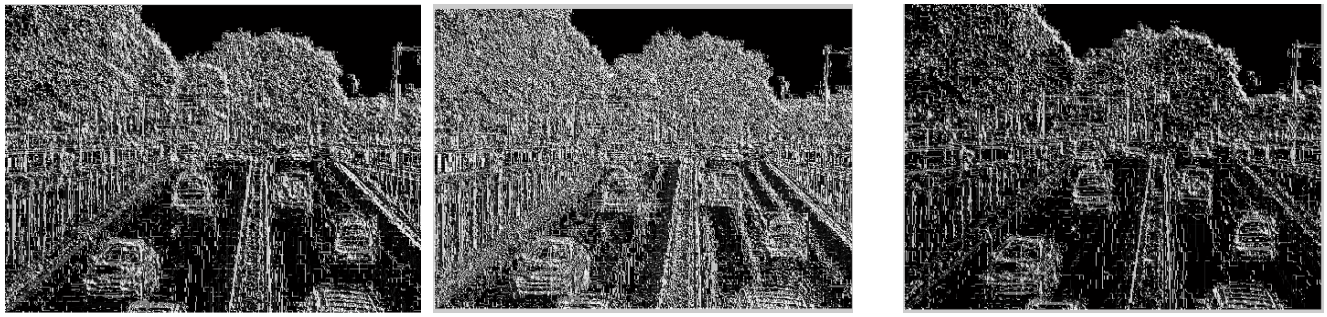
b) Cropping input image by 200*200 pixels to create tampered image



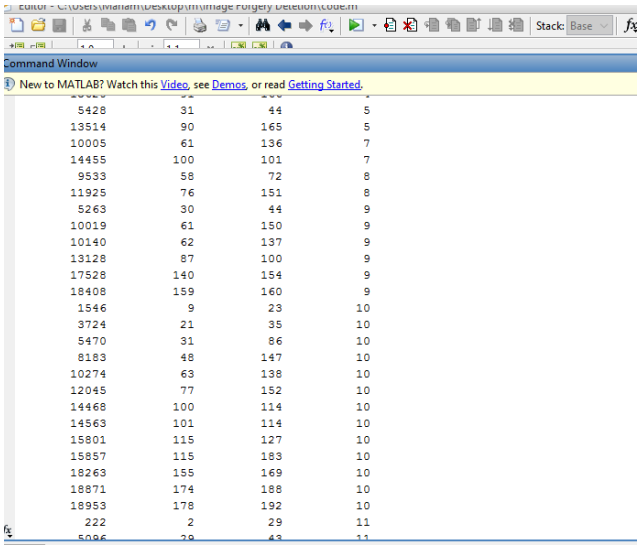
c) Separating color channels (R,G & B) of tampered image



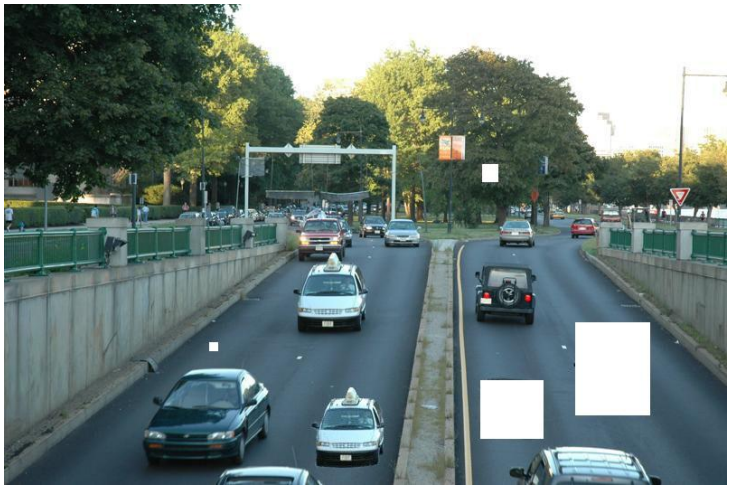
d) Decomposing R, G & B into overlapping block



e) LBP of channel R, G & B



f) Block pair distance calculation and sorting it in array, removing adjacent block keeping 1/5 only



g) Filling cloned blocks with white colors

Figure 4.3: A detailed example of cloning detection using proposed methodology

Next step is to compare our proposed technique with some of the previously developed methodologies on the same data sets to compute rates, accuracy, sensitivity and processing time. Chapter 5 will cover the above said statement.

CHAPTER 5: PERFORMANCE COMPARISON

Now moving towards the second phase, in this chapter a comprehensive yet detailed comparison of previously developed copy-move methodologies with the proposed technique is presented. We used five formerly developed research publications & proposed in order to develop the general comparison.

Properties like time complexity, effects of post processing**,** rates, sensitivity, specificity & accuracy etc. are compared. On the basis of analysis of these parameters we have derived some key results too. Here we would like to present our key findings about all these methodologies including their pros & cons in Table 5-1 & Table 5-2.

TABLE 5-I PARAMETER COMPARISON OF PROPOSED TECHNIQUE WITH PREVIOUSLY DEVELOPED METHODOLGIES

Methodology	**Pre-Processing**	**Time Complexity (sec)**	**Effect of Geometric Transformation**	**Advantages**	**Drawbacks**
Haar-like Features [30]	Grey-scale only	8.86	Weak against geometric transformation	Time efficient due to integration of image.	Weak against geometric transformation since every time different filter should be used, FPR very large as identical features computed using hit & trial basis
Sub-Blocking [31]	Grey-scale/un-compressed	18.15	Handles scaling & known rotation only	Less FPR, Effective result due to erosion followed by dilation	Shift vector computed on hit & trial basis which can increase the chance of tampering data loss & FNR

Multi-Resolution LBP [32]	Colored	25.43	Tested for all geometric irregularities	Addressed affine transforms that were Rarely considered before.	Time Consuming, Complex algorithm due to lexographic sorting for each block, Failed for high resolution
2D-Fourier Transform [33]	Grey Scale/Smoothing	7.45	Handles geometric irregularities	Effective for multiple cloning	Complex & less accurate due to large number of Fourier transform coefficients
Expanding Blocks [34]	Colored	22.54	Handles all geometric modification	Can handle any irregular shape even darkened or lightened duplicated region	Slow in execution as blocks are sorted on basis of mean variance,
Proposed Methodology	Colored	9.32	Handles any geometric transformation	Efficient due to descriptor Low FPR due to neighbourhood clustering	Failed for affine transformation

The following quantities are then computed using data set for further comparison: True Positive which determines the number of tampered regions categorized as tampered, False Positive which shows the number of un-tampered regions categorized as tampered, True Negative to determine the number of un-tampered regions categorized as un-tampered & False Negative which shows the number of tampered regions categorized as un-tampered.

TABLE 5- II COMPARISON OF FPR, TPR, TNR & FNR

Methods	TPR	FPR	TNR	FNR
[30]	88	56.31	65.67	12.45
[31]	94	27.34	76.68	39.38
[32]	91.49	30.58	65.4	23.12
[33]	77.84	45.12	55.67	10.54
[34]	71.23	37.65	45.29	20.21
Proposed	96	20.18	87.31	9.14

Now using the general formulae for accuracy and sensitivity by plugging in the value of rates computed above the comparison of accuracy & sensitivity is given in table II. Sensitivity designates the effectiveness of technique in extracting tampered area; accuracy determines the extraction of tampered area with least rate of false positive whereas specificity narrates the capability of methodology to detect trustworthy image appropriately as trustworthy. Therefore a high value of specificity & sensitivity yields better performance of the algorithm.

$$\text{Sensitivity} = \frac{\text{True Positive}}{\text{True Positive+False Negative}} * 100$$

$$\text{Accuracy} = \frac{\text{True Positive+True Negative}}{\text{True Positive+False Negative+True Negative+False Positive}} * 100$$

$$\text{Specificity} = \frac{\text{True Negative}}{\text{True Negative } + \text{False Negative}} * 100$$

TABLE 5-III COMPARISON OF ACCURACY, SENSITIVITY & SPECIFICITY

Methods	Sensitivity	Accuracy	Specificity
[30]	87.6	69.08	84.06
[31]	70.47	71.89	66.06
[32]	79.82	74.50	73.88
[33]	88.07	61.59	84.08
[34]	77.89	66.81	69.14
Proposed	91.30	86.21	90.52

The performance analysis of these techniques is shown in graphical form in figure **5.1**.

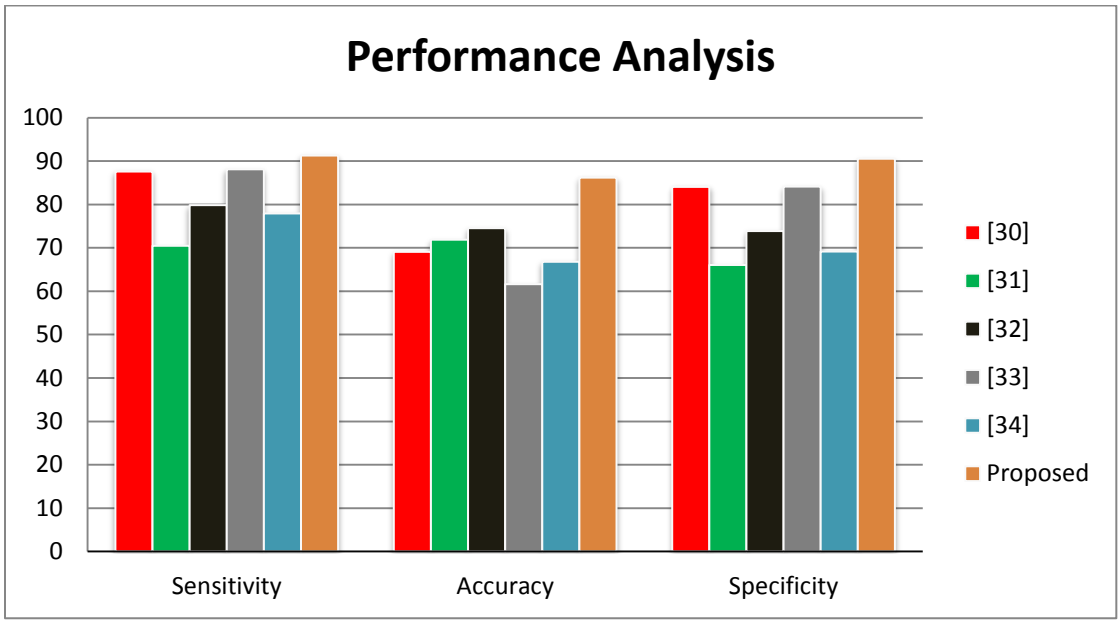


Figure 5.1 Performance Analysis of Methodologies

The results indicate that the proposed method works more effectively and efficiently as compared to other methods.

LIST OF PUBLICATIONS

1. Mariam Saleem , Qasim Altaf, Dr.Qaiser Chaudry, “A Comparative Analysis on Pixel-Based Blind Cloning Techniques ”,in IEEE 4th International Conference on Control System, Computing & Engineering , Penang (Malaysia), November 2014
2. Qasim Altaf, Javaid Iqbal, Rabnawaz, “A Block-Based Robust Algorithm to detect Multi-Cloning Forgeries”, 19th IEEE International Multi-Topic Conference, (under review)

U7

CHAPTER 6: CONCLUSION & FUTURE SUGGESTION

In this dissertation, we thoroughly discussed Digital Image Forensics and proposed an efficient & effective tampering detection technique. The primary focus of our research was to develop a copy-move technique which can detect multi-cloned tampering too. After analyzing the previous methodologies in literature review we were able to pin point the pros and cons of each technique. Completing this investigation enabled us to recognize the domains which were missed by previous researchers. Keeping in mind all these outcomes we developed an algorithm based on the local binary pattern & clustering techniques. More over the technique follow these initial steps. Firstly, the image is decomposed into 3 color components. As we have surveyed previously developed methodologies, nearly in every algorithm first step was to convert color image into gray-scale by using approximated weights of R, G & B. During this conversion, some weak but important bits of tampering can be lost. So in this proposed algorithm the color image is converted to components only. Afterwards each component is divided into overlapping blocks, followed by extracting characteristics of each block using powerful visual descriptor i.e. Local Binary Pattern. Feature vector also known as texture spectrum is computed automatically due to which time complexity is reduced as compared to previously developed methodology which worked on cumbersome methods. After computing the dissimilarity (distance) between the blocks of the image neighborhood clustering is employed to reduce false positive rate.

The simulations results, comparative/performance analysis between proposed & previously developed techniques in chapter 4 & chapter 5 respectively showed that the method can accurately and efficiently detect the cloned region with the processing time greatly reduced. The

accuracy, sensitivity & specificity of algorithm using images of data set were found to be 86.21, 91.30 and 90.52% respectively. However the proposed technique failed to accurately detect cloned region if some post processing operations like affine transformation had been done to diminish the visual artifacts of tampering.

Based on the performance of proposed techniques it is highly recommended to extend this research in future.

a. Ability to handle severe Geometric Transformation i.e. Affine Modification.

b. Applying wide variety of texture descriptors to compute vector beside LBP.

c. Current research in cloning detection techniques is limited to a two dimensional image and can be extended to audio and video.

d. Ability to detect tampering in case of compressed image data.

REFERENCES

[1] “PhotoTampering”,at FourandSix.comhttp://www.fourandsix.com/photo-tampering-history/category.

[2] H.T. Sencar, and N.Memon, “Overview of State of- the Art in Digital image Forensics”, World Scientific Press, 2008.

[3] Judith A. Redi & Wiem Taktak & Jean-Luc Dugelay, "Digital image forensics: a booklet for beginners”, Multimedia Tools and Applications, Vol. 51, Oct. 2010, pp.133-162.

[4] H. Farid: "A Survey of Image Forgery Detection", IEEE Signal Processing Magazine, 26(2):16-25, 2009

[5] Luo W, Qu Z, Pan F, Huang J. A survey of passive technology for digital image forensics. Front Comput Sci China 2007a; 1(2):166–79.

[6] Bayram S, Sencar HT, Memon NA. Survey of copy-move forgery detection techniques. In: Proc. of IEEE Western New York image processing workshop 2008.

[7] http://www.adobephotoshop.com

[8] Jessica F., David S., Jan Lukas., “Detection of Copy-Move Forgery in Digital Images”, in Proceedings of Digital Forensic Research Workshop, Cleveland, OH, August 2003.

[9] J. Fridrich, “Methods for "Methods for Tamper Detection in Digital Images", in Proc. ACM Workshop on Multimedia and Security, Orlando, FL, October 30−31, 1999, pp. 19−23.

[10] Bo Xu, Guangjie Liu and Yuewei Dai, “ A Fast Image Copy-move Forgery Detection method using Phase Correlation” , in Fourth International Conference on Multimedia Information Networking and Security,2012.

[11] Fridrich J. “Digital image forensics”. IEEE Signal Process Mag 2009; 2 (26):26–37.

[12] Popescu A, Farid H. “Exposing digital forgeries by detecting duplicated image regions”. TR2004-515. DC,2004

[13] Saiqa Khan, Arun Kulkarni, "Robust method of detection of copy-move forgery in digital images," in IS&I, IEEE , 2010

[14] S. Ryu, M. Lee and H. Lee, "Detection of copy-rotate-move forgery using Zernike moments," in. Int. Workshop Information Hiding, Springer, pp. 51–65, 2010.

[15] Vivek Kumar Singh, R.C.Tripathi, "Fast and efficient region duplication detection in digital images using sub-blocking method", IJAST (Vol.35), Oct, 2011.

[16] Ghulam Muhammad, Muhammad Hussain, George Bebis., "Passive copy move image forgery detection using undecimated dyadic wavelet transform", Digital Investigation (9), 2012.p.49-5790

[17] Li L, Li S, Zhu H, Wu X. "Detecting copy-move forgery under affine transforms for image forensics", Compt.Elect.,2013.

[18] Bin YANG, Xingming SUN, Xianyi Chen, Jianjun Zhang "An efficient forensic method for copy–move forgery detection based on DWT-FWHT", VOL. 22, NO. 4, DECEMBER 2013

[19] Seniha Ketenci, G.Ulutas, "Copy-move forgery detection in images via 2D-Fourier transform", TSP, IEEE, 2013

[20] L Gavin, S Frank, "An efficient expanding block algorithm for image copy-move forgery detection", Information Sciences 239, 2013

[21] XuBo, Wang Junwen, Liu Guangjie & Dai Yuewei "Image copy-move forgery detection based on SURF", ICMIN&S, 2010

[22] Li Jing , Chao Shao, "Image copy-move forgery detection using local invariant feature" in JM,Vol.7 ,#1,2012.

[23] I. Amerini, Lumberton Ballan, "A SIFT-based forensic method for copy-move attack detection and transformation recovery", IEEE Transaction on IF&S, Vol.4, No. 3, September 2011

[24] N. Suganthi, "Detecting forgery in duplicated region using key-point matching". International Journal of Scientific and Research Publication, 2010, vol. 2, no. 11, Nov, 2012

[25] B.L.Shivakumar, Dr.S.S.Baboo, "Detection of region duplication forgery in digital images using SURF", IJCSI 8(4) (2011)199–205

[26] Mohamadian Zahra, "Image duplication detection using two robust features", Res. J. Recent Science, Vol.1 (12), 1-6, Dec2012.91

[27] H. Huang, Tseng, Y. Zhang, Parul Mishra "Detection techniques based on SURF and HAC," in HPC, 2013.

[28] Maryam Jaberi, George Bebis, M.Hussain, Ghulam Muhammad "Accurate and robust localization of duplicated region in copy-move image forgery", Springer, June 2013

[29] http://lci.micc.unifi.it/labd/2015/01/copy-move-forgery-detection-and-localization/

[30] Ghorbani, irouzmand, "Detection of Copy-Move Forgery using Haar-like features", IIJCS, ISSN 2321-5992, Volume 2, Issue 1, January 2014.

[31] R.C.Tripathi, "Fast and efficient region duplication detection in digital images using sub-blocking method", IJAST (Vol.35), Oct, 2011.

[32] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns," IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 7, pp. 971-987, 2002.

[33] Seniha Ketenci and Guzin Ulutas, "Copy-Move Forgery Detection in Images via 2D-Fourier Transform" 978-1-4799-0404-4/13/$31.00 ©2013 IEEE.

[34] L Gavin, S Frank, "An efficient expanding block algorithm for image copy-move forgery detection", Information Sciences 239, 2013