Digital Forensic Technique for Multiple Compression based JPEG Forgery

A thesis Submitted by

Pankaj Malviya

710cs1028

In partial fulfilment of the requirements for the award of the degree

of

Master of Technology

In

Computer Science and Engineering

Department of Computer Science and Engineering
National Institute of Technology, Rourkela 769008

May 2015

Digital Forensic Technique for Multiple Compression based JPEG Forgery

A thesis Submitted by

Pankaj Malviya

710cs1028

In partial fulfilment of the requirements for the award of the degree

of

Master of Technology

In

Computer Science and Engineering

Under the Guidance of

Dr. Ruchira Naskar

Department of Computer Science and Engineering
National Institute of Technology, Rourkela 769008

May 2015

Dedicated to,

My Parents and teachers

Department of Computer Science and Engineering
National Institute of Technology, Rourkela 769008

CERTIFICATE

This is to certify that the thesis entitled **Digital Forensic Technique for multiple Compression based JPEG Forgery** submitted by **Pankaj Malviya** in partial fulfilment of the requirements for the award of the degree of **Master of Technology "Dual Degree"** in **Computer Science and Engineering** with specialization in **Computer Science** to the **National Institute of Technology, Rourkela** is an authentic record of research work carried out by him under my supervision. The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Thesis Advisor

Dr. Ruchira Naskar

Assistant Professor

Department of Computer Science and Engineering

ACKNOWLEDGEMENTS

As a matter of first importance, acclaims and on account of The God, The Almighty, for his showers of gifts all through my exploration work to finish the research effectively. I would like to express my most profound gratitude to my guide Dr. Ruchira Naskar for the direction and support. Her great vitality, innovativeness and incredible coding abilities have dependably been a steady wellspring of inspiration for me. The flawlessness that he conveys to every last bit of work that he does constantly roused me to do things comfortable time. She is an extraordinary individual and one of the best teacher, I generally be grateful to her. I might likewise want to say thanks to my friends for dedicating there time in helping me in my work. I also want to thank all PhD researchers in NIT Rourkela for making our lab such an incredible work environment. Most importantly I want to commit this proposal to my amazingly adoring and strong folks who have dependably been with me, regardless of where I am.

Besides my thesis advisor I extend my sincere thanks to **Dr. Sunil K. Sarangi, Director, National Institute of Technology, Rourkela**, **Prof. Santanu K. Rath, Head of the Computer Science and Engineering Department**, and all other faculties of the department for their timely co-operations during the project work. Such a large number of individuals have contributed to my research work, and it is with awesome joy to take the chance to express gratitude toward them. I apologize, in the event that I have overlooked anybody.

Pankaj Malviya

710cs1028

ABSTRACT

In today's digital world digital multimedia like, images, voice-notes and videos etc., are the major source of information/data exchange. The authenticity of these multimedia is greatly vital in the legitimate business, media world and broadcast industry. However, with enormous multiplication of ease, simple-to-utilize data manipulation tools and softwares lead to the faithfulness of digital images is in question. In our work, we propose a technique to identify digital forgery or tampering in JPEG (Joint Photographic Experts Group) images which are based on multiple compression factor. We are dealing with the JPEG images on the grounds because JPEG is the standard storage format used in almost all present day digital devices like digital camera, camcorder, mobile devices and other image acquisition devices. JPEG compresses a image to the best compression in-order to manage the storage requirement. JPEG is a lossy compression standard. At the point when an assailant or criminal modifies some region/part of a JPEG image by any image processing tools and save it, the modified region of the image is doubly-compressed. In our work, we exploit this multiple compression in JPEG images to distinguish digital forgery or falsification.

Keywords:- Digital Forensics, Image Forgery, Image Authentication, Cyber Crimes, JPEG Standard, JPEG Compression.

Contents

List of Figures

Chapter 1

Introduction

1.1 INTRODUCTION

In the present digital age cyber crime is the latest and perhaps the most complicated problem rising in the world. With the increasing proliferation of information technology, communication technology and the thriving opportunity for real-time borderless exchange of information, cyber crime is a sophisticated transnational issue in the growing economy. Any criminal activity/action that uses a computer either as an instrumentality, target or a means for perpetuating further crimes comes within the scope of cyber crime. Cyber crime may be said to be those species of which, forgery is the traditional crime, and where either the computer is an object or subject of the conduct constituting crime on digital information, such as manipulating the digital information. The specialized part of an investigation and analysis is partitioned into a few sub-branches, identifying with the kind of advanced gadgets included; PC legal sciences, system criminology, scientific information investigation and cell phone criminology. The average measurable procedure envelops the seizure, scientific imaging (securing) and analysis and investigation of digital media and the generation of a report into gathered proof. And in addition distinguishing direct confirmation of a wrongdoing, advanced legal sciences and technology can be utilized to credit proof to particular suspects, affirm justifications or explanations, focus plan, recognize sources (for instance, in copyright cases), or confirm archives. Investigations are much more extensive in degree than different zones of measurable examination (where the typical point is to give answers to a progression of easier inquiries) frequently including complex timetables or speculations.

The development in computer crimes amid the 1980's and 1990's brought on law enforcement organizations to start building up specific gatherings, as a rule at the national level, to handle the specialized parts of investigation and analysis. Since 2000, in light of the requirement for institutionalization, different bodies and organizations have distributed rules for advanced criminology. Amid the 1980s not very many specific digital measurable devices existed, and hence

specialists frequently performed live analysis on media, looking at computers from inside of the working framework utilizing existing framework head instruments to concentrate proof. This practice conveyed the danger of adjusting information on the circle, either unintentionally or something else, which prompted cases of proof altering. Various instruments were made amid the mid 1990's to address the issue.

An advanced measurable investigation normally comprises of three stages, procurement or imaging of shows, investigation, and reporting. In a perfect world procurement includes catching a picture of the PC's primary memory (RAM) and making a careful part level copy (or "measurable copy") of the media, regularly utilizing a compose blocking gadget to avoid adjustment of the first. Then again, the development in size of capacity media and advancements, for example, distributed computing have prompted more utilization of "live" acquisitions whereby a "coherent" duplicate of the information is obtained instead of a complete images of the physical stockpiling gadget. Both obtained images (or legitimate duplicate) and unique media/information are hashed the qualities contrasted with confirm the duplicate is precise.

1.1.1 JPEG Image

The term JPEG is an acronym for the "Joint Photographic Experts Group", the file format or standard is invented by the group. The standard JPEG(Joint Photographic Experts Group) [7, 8] is widely used lossy compression format for image acqusition all over the world in almost all the digital devices like camera. JPEG provides the best compression factor which optimize the space requirements. This standard helps in collecting information from the crime scenes very efficiently wiht very low space required of storage. It is very much possible to manually adjusted degree of compression, permitting a selectable trade-off between storage capacity and the quality of the image. With the slight perceptible loss of information and image quality JPEG achieves 10:1 compression. An example of JPEG image is shown in the fig. 1.1, which shows how the compressed image looks.

Figure 1.1: A image with the compression rate decreasing, and hence quality increasing, from left to right!

The image in fig. 1.1 shows the compressed image at decreasing compression factor due to which the quality of the image is increasing from left to right. It says, lower the JPEG quality, lower the image quality, considering the quality factor of 0 implying the highest compression of the image, this means the image is completely degraded. With the increase in the value of quality factor, the image degradation decreases.

1.1.2 JPEG Forgery

These days modification of images through spontaneous programming is an operation of ingenuousness with minimal effort, in this manner each individual can orchestrate a fake image. It can be surprisingly more terrible in the equity when images are introduced as proof in the court of law. Along these lines, there is a solid interest for substantial and hearty confirmation system to observe whether a photo contains the true information or not. Digital Image Forgery may involves single image or multiple image, when an attacker try to modify the

image, the usual practice is to use any photo-editing tool in that case the image is re-saved. While re-saving the image goes through re-compression. Common image processing operations such as cropping, splicing, blurring etc., made widely available by such software tools, compel us to question the trustworthiness of the digital images and videos. In digital image frauds including two or more images, portion(s) of one image is malignantly transplanted into another, to give a thought to the viewers that the transplanted segment is a legitimate piece of the recent picture. In this process, when the region/part of the image which is manipulated is singly compressed while the remaining region of the image goes through multiple compression, the part which is replaced using some other part of the image is of different compression factor than the original image. This destroys the artifacts of the JPEG image as multiple compression in the same image. The fig. 1.2 and fig. 1.3 represents the digital image forgery involving multiple images where some particular region of the image modified using some part of the other image in-order to manipulate the facts and figures contained in the image.

In this work, we proposed a digital forensic technique for the detection of digital tampering/forgery in JPEG images. To clearly understand the concept behind JPEG forgery investigation technique, proposed in our work, let us consider a modification attack carried out on JPEG images. To deliver the attack, the attacker first opens the image in an photo editing software, manipulates some regions of the image, and finally re-saves the image back in JPEG format. In this entire process, some regions of the image get "doubly-compressed". Hence, degree of compression of the forged part is different from the rest of the image. This difference in degree of compression, although not perceptible to the Human Visual System (HVS), has been efficiently utilized in our work to provide evidence of image manipulation.

The measurable and in addition perceptual repetition in common pictures, are efficiently misused in JPEG compression. Besides, the JPEG configuration

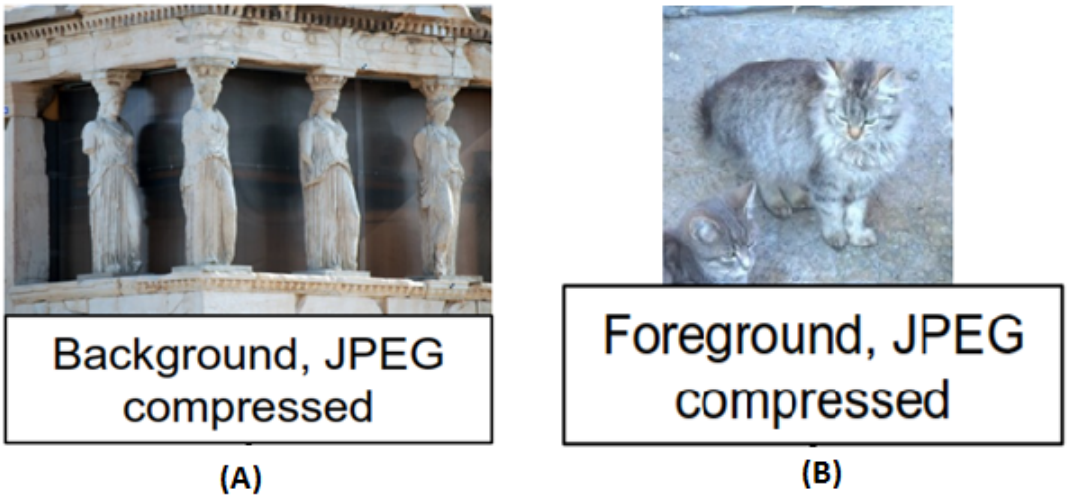

Figure 1.2: Image A This is original image (Background image which is modified), Image B This is the forground image which is adjusted to the original image.

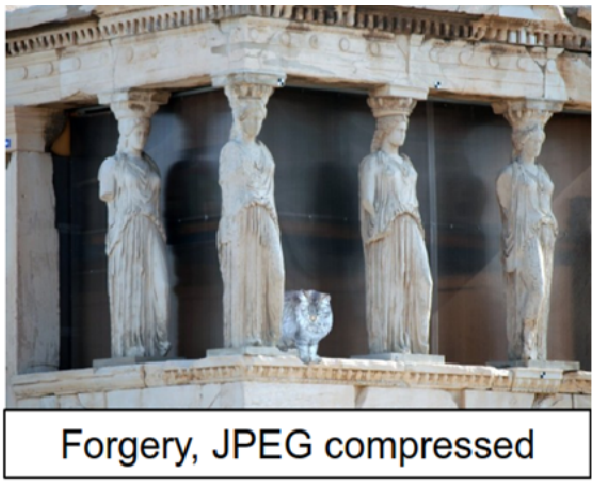

Figure 1.3: This is the resulting image which is obtained by adjusting forground image to the background(Original) image.

has a versatile compression conspire that permits sparing in differing levels of compression. Then again, every time we pack a image some space is spared yet in the meantime some data misfortune happens. Additionally when the image is reproduced from its JPEG compacted variant, it contains corruptions contrasted with the first image. In spite of the fact that the loss of data is disadvantageous while image forgery recognition, different JPEG highlights are

invaluable for identification of change or modification of JPEG pictures.

The image blocks which are transplanted may also be geometrically transformed by the attacker in many situations, e.g., an image block may be rotated by some angle by the attacker before it is transplanted. In such case, duplicate image regions may be detected by matching Scale-invariant feature transform (SIFT) [11] key points of the regions. Such forensic approaches to diagnose copy-move forgery having geometrically transformed image blocks, have been proposed in [13] and [12].

1.1.3 Motivation and Objective

In the present cyber world, digital images and videos being the substantial sources of evidence towards faithfulness of any event, maintenance of their reliability and trustworthiness is a major challenge in todays digital world. Digital images and videos are the source of information, which act as a evidence for law enforcement all around the world. These digital evidences like images are the most effective and efficient way of collecting information from the crime scene. Due to the availability of large number of electronic devices like Camera, Camcorder etc., it is very much efficient to collect information in the form of images and videos without alerting perpetrators during ongoing crime. Later, this multimedia files can be used as evidence in court of law by Law Enforcement. Though, the availability of voluminous software and tools which makes the information contained in the multimedia file vulnerable. In fact, many of these software and tools are in reach of a common man, which can be used to exploit the information/data. It is very much possible for an attacker to manipulate or hide the facts and figure in the multimedia file and delude the law enforcement in order to escape the court of law. It is extremely crucial to preserve such digital evidences against cyber crime for suitable presentation in court of law as well as for the media world and broadcast industry. The need for investigation and maintenance of the fidelity and reliability of digital multimedia, has given rise

to the field of Digital Forensics in the research and science community. Over the recent years, researchers and scientists have focused on the areas of image authentication, detection of tampering in the image, identification of image forgery as well as investigation of image sources. This is the motivation behind our idea of research, which can be helpful to identify the acts of forgery and help the law enforcement to solve the crime. Henceforth we are persuaded to work in the field of Digital Forensics [1] and [2] more specifically, on the JPEG forgery Detection Technique.

In this thesis work, our goal is to come up with the effective and efficient method to detect the acts of forgery in the JPEG images. The technique which helps not only in matters of detection but also in the localization of the tampered region/part of the image.

1.1.4 Our Contribution

1. We have proposed a new Digital Forensics Technique to detect the forgery in the JPEG images.

2. A technique which can help the forensic analyst to identify and localize the tampered region(s) in the image.

1.1.5 Thesis Organization

The remaining section of the thesis is organized as follows chapter-2 summarizes the related work in which we works on the detection of JPEG Ghost, chapter-3 comprise of the proposed work, we have come with the new forgery detection technique,chapter-4 provides the results and discussion, chapter-5 provides the Conclusion and Future work.

1.2 Summary

Here, in this chapter we have discussed briefly about the basic fundamentals of the digital forensics. More specifically we discussed about the JPEG image Forgery in the digital images.

Chapter 2

Related Work

2.1 Related Work

Digital forensics is the branch of computer science and technology which deals with detection of cyber crime by investigation and analysis of digital information which act as the evidences involved in the cyber crime. In this field of information technology, we deals with the study of digital multimedia like images, videos, voice notes etc.., In this work we manage recognition of digital image forgery, uncommonly JPEG images. A digital image forgery may include a solitary or different images. In a solitary image imitation, some segment of the image is supplanted by some other part. Consequently a few items may be erased from, modified or rehashed in the picture. Moreover such picture modifications are generally joined by smoothening of the item edges by smearing or obscuring chose districts of the picture. As specified already, in today's situation, simple accessibility of easy to understand picture preparing softwares and tools has made such picture controls greatly minor, notwithstanding for novice users.

Numerous computerized scientific techniques misuse the factual attributes, intrinsically display in normal images (for e.g., high pixel-value correlation) to distinguish produced or unnatural images. In common images, visual descriptors are highlights used to evaluate the visual boost that the picture creates in the HVS (Human Visual System). One sample of visual descriptors utilized as a part of to recognize regular pictures, is shading properties of the picture. The authors [10] have utilized the quantity of unmistakable hues involving the images also spatial variety of hues in the picture. Normality in shading organization is another intrinsic regular picture measurable highlight, misused by the authors in [4], by considering that the frontal area (questions) and foundation are profoundly color-compatible in characteristic pictures. Samples of other characteristic picture insights utilized as a part of digital image crime scene investigation are shadow surface, surface unpleasantness or smoothness, power range of picture and so on., Wavelet space coefficients and different snippets of

the wavelet circulation, for example, mean and change, are additionally utilized by a few specialists as common picture measurements in advanced criminology. In advanced image forgery including numerous pictures, portion or portions of one image is vindictively transplanted into another, to give a thought to the viewers that the transplanted bit is a legitimate piece of the last picture. This class of falsification is alluded as copy-move forgery. One of the earliest advanced scientific systems for copy-move forgery, proposed by Fridrich et. al. in [8], is in view of the rule of cloning identification. In [8], the authors hunt down two picture locales, having precisely indistinguishable pixel values. However standard images comprising of thousands of pixels, it is computationally very infeasible to complete a savage power pursuit to and such indistinguishable (image region) sets. To make the seeking efficient, the authors separate the whole picture into fixed-sized pieces and after that sort the squares lexicographically. Post-sorting, extraction of indistinguishable picture obstructs from this sorted rundown is not relevant.

Whenever the attacker is transplanting a image region onto some other region of the image, ordinarily the foe needs to resize the square, to match the quantity of the area to darken. One such forgery technique has been introduced in [9]. Re-inspecting (by a number variable) actuates intermittent relationships among unique and manufactured picture pieces. Such connection among sign specimens, a marvel not pervasive in normal pictures, is abused in to recognize transplantation of re-inspected picture squares.

2.1.1 JPEG Ghost

Digital forensic techniques for JPEG forgery detection, based on analysis of JPEG ghost classification are proposed in [14] and [15]. JPEG ghost classification is a forensic analysis technique that enables detection of multiple JPEG compressions in an image. In any present-day digital camera, widely used and the standard configuration for picture format is JPEG (Joint Photographic

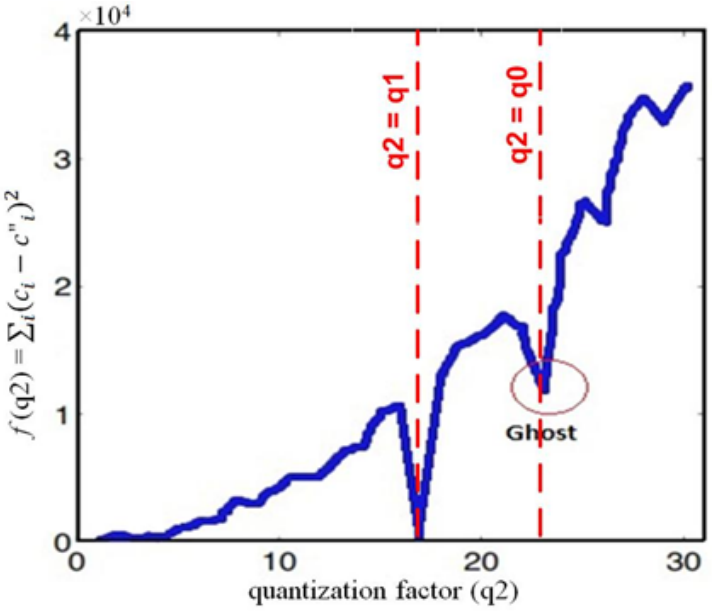

Figure 2.1: The plot represents the formation of JPEG Ghost in multiple compressed images.

Experts Group). This standard is utilized because of the way that JPEG arrangement delivers the best compression, thus ideal space prerequisite for picture format. The measurable and also perceptual repetition in common pictures, are very much abused in JPEG pressure. Additionally, the JPEG configuration has a versatile pressure plot that permits sparing in fluctuating levels of pressure. Be that as it may, each time we pack a picture some space is spared, however in the meantime some data adversity happens. Likewise, when the image is reproduced from its JPEG packed version, it contains degradations contrasted with the first picture. In spite of the fact, that the loss of data is disadvantageous while picture imitation location, different JPEG highlights are beneficial for recognizable proof of modification or change of JPEG pictures.

2.1.2 JPEG Ghost Detection

Now, lets consider an arrangement of coefficients c_0 quantized by a amount q_0, trailed by quantization by a amount $q_1 < q_0$ to yield c_1. Further quantizing c_1 by

q_2 yields the coefficients c_2. As some time recently, the contrast in the middle of c_1 and c_2 will be negligible when $q_2 = q_1$. Yet, since the coefficients were at first quantized by q_0, where $q_0 > q_1$, we hope to locate a second least when $q_2 = q_0$. Demonstrated in Fig. 2.1 is the aggregate of squared contrasts in the middle of c_1 and c_2, as an element of q_2, where $q_0 = 23$ and $q_1 = 17$. As some time recently, this distinction increments as an element of expanding q_2, achieves a base at $q_2 = q_1 = 17$, and most interestingly has a second nearby least at $q_2 = q_0 = 23$. We allude to this second least as a JPEG GHOST, as it uncovers that the coefficients were already quantized (compacted) with a bigger quantization (lower quality). Review that the JPEG compression conspire independently quantizes every spatial recurrence inside of a 8 ×8 pixel square. One way to deal with recognizing JPEG Ghost would be to independently consider every spatial recurrence in each of the three luminance/shading channels. On the other hand, review that different minima are conceivable when contrasting number different quantization values. On the off chance that, then again, we consider the total impact of quantization on the hidden pixel values, then this issue is far less inclined to emerge. Hence, as opposed to registering the distinction between the quantized DCT coefficients, we consider the distinction figured straightforwardly from the pixel value. Note that it is expected here that the same JPEG qualities were utilized as a part of the creation and testing of a picture, and that there is no shift in the altered area from its unique JPEG piece grid. The effect of these presumptions will be investigated underneath, where it is demonstrated that they are not basic to the adequacy of the location of JPEG apparitions. So as to make a consistent match with whatever remains of the picture, it is likely that the controlled area will be changed after it has been embedded. Any such post-preparing may disturb the discovery of JPEG phantoms. To test the effectively to such post-preparing, the altered locale was either obscured, honed, or histogram leveled subsequent to being embedded into the picture. For altered districts of size 100 ×100, the recognition enhanced marginally (with the same false positive rate of 1 percent).

Here, JPEG Ghost is considered as the forgery in the image. This technique helps us understand that the image is being forged or tampered with, since we get multiple minima during the experiment. If the image is doubly compressed, we will get second minima, on the other hand if the image is multiply compressed then multiple minima is obtained which represents the image is multiply compressed. While performing the same operation with the original image which is not being forged or tampered with, there is only fist minima is observed. This observation represents that the image is authentic, which is not doubly compressed. This technique is only applicable for the JPEG images.

2.2 Summary

Here, in this chapter we have discussed briefly about the related work and the literature survey. We discussed about the JPEG Ghost technique to detect the forgery in the image.

Chapter 3

Proposed Work

3.1 Proposed Work

In any JPEG standard image, at whatever point any altering is done on the image and it is composed back to memory, the image experiences re–compression. This highlight is abused in the proposed work [16] to recognize any illicit adjustment or altering in JPEG pictures. To make the idea more understandable, let's consider the 512×512 *Lena* images shown in Fig. 3.1 (a) and (c), both of which are JPEG images. Now let's take into consideration a modification attack which modifies/tampers a region/part at the center of the original/untampered image shown in Fig. 3.1(a). This central part/region of size 200×200 pixels, shown in Fig. 3.1(b), has been extracted, re–saved at a different JPEG quality/factor and transplanted into the original image of Fig. 3.1(a) to produce the tampered image of Fig. 3.1(c). It is clearly evident from Fig. 3.1 that the regions saved at different JPEG quality factors, are not perceptibly distinguishable.

Figure 3.1: *Lena* JPEG images: (a) Original Lena 512×512 image; (b) Central 200×200 portion/region which is re–saved at a different degree of compression; (c) Forged lena image in which the central portion/region is modified.

3.1.1 Proposed JPEG Forgery Scheme

Next, we have come up with a forensic technique to detect such JPEG forgeries with double/multiple degrees of compression (the degree/factor of

compression which is varying from one region to other) within the same image. To detect such acts of forgery in a JPEG image, the following procedure is followed:

1. First, we compress the entire JPEG image iteratively, at varying degrees of compression. We refer to this degree of compression as the JPEG quality factor. The JPEG quality factor may range from 0 to 100. Higher the JPEG quality, higher is the image quality, with quality factor 100 implying no compression of the image at all. With decrease in the value of quality factor, the image degradation increases. In our work, we vary the quality factor between [40,90].

2. Let the quality factor of the original JPEG image I be q, which is a constant, and the quality factor used for re–compression be q', which is varied in [40,90] in steps of 1. For each q', we find the squared–error matrix of the image. The squared–error matrix D_1 is defined as:

$$D_1(i,j) = [I(i,j) - I_{q'}(i,j)]^2, \ \forall 512 \leq i,j \leq 512 \tag{3.1}$$

   where image $I_{q'}(i,j)$ is produced by compressing I at quality q'.

3. From each squared–error matrix $I_{q'}$, we again compute the consecutive horizontal pixel–pair differences, row–wise. The absolute pixel–pair differences are computed as:

$$D_2 = \{|I_{q'}(i,j) - I_{q'}(i,j+1)| : 1 \leq i \leq 512, 1 \leq j \leq 511\} \tag{3.2}$$

   where D_2 contains 512×511 elements.

4. We consider D_2 as the vector $[D_2(1), D_2(2), \ldots D_2(512 \times 511)]$. And we consider another vector P of pixel–pair positions in a 512×512 matrix; $P = [1, 2, \ldots 512 \times 511]$.

5. Finally, we plot the vector of absolute differences, D_2 against P. We investigate the variation of the elements of D_2 over the entire 512×512

image matrix, from the D_2 vs. P plot. Our key observation in this paper is that, for forged JPEG images (containing multiple degrees of compression within the same image), for certain values of $q' \in [40, 90]$, the D_2 vs. P plot demonstrates a sudden rise, which remains persistent over a range of P (pixel–pair positions), corresponding to the area or region of image tampering. That is, if 200×200 pixels are tampered, the D_2 values remain persistent for $(200 \times 199 =)39,800$ positions, after the sudden rise.

3.1.2 Observation

The D_2 vs. P plot for the modified *Lena* image in Fig. 3.2 is shown in Fig. 3.2. Fig. 3.2(a) which shows a sudden rise in the pixels values in the central part of image, which is persistent for the certain range of pixels which is undergone through double-compression due to modification in the image. This characteristic of JPEG images, provides an evidence of image forgery, involving double or multiple compression. Authentic JPEG images in which have no sub–part or region manipulated (hence doubly-compressed), demonstrate neither such a sudden rise of D_2 values nor its persistence. This is evident from Fig. 3.2(b), which shows the D_2 vs. P characteristics of the authentic or original JPEG *Lena* image, without any modification or tampering.

In our work, the technique proposed is very much effective in the detection the tampered JPEG image which are double or multiply compressed during the editing. The resulting plot of the test image clearly represents that the part of image is being tampered with, because the part shows the sudden hike in the pixel values due to the JPEG characteristics. While the image is undergone forgery the tampered part of the image which is part of some other image originally with the different compression quality/factor is transplanted over the image which is of different compression quality. Thus, while saving the image to the memory the whole image is gone through the compression due to the JPEG characteristics. But, the central part of the image which is the tampered part is singly compressed

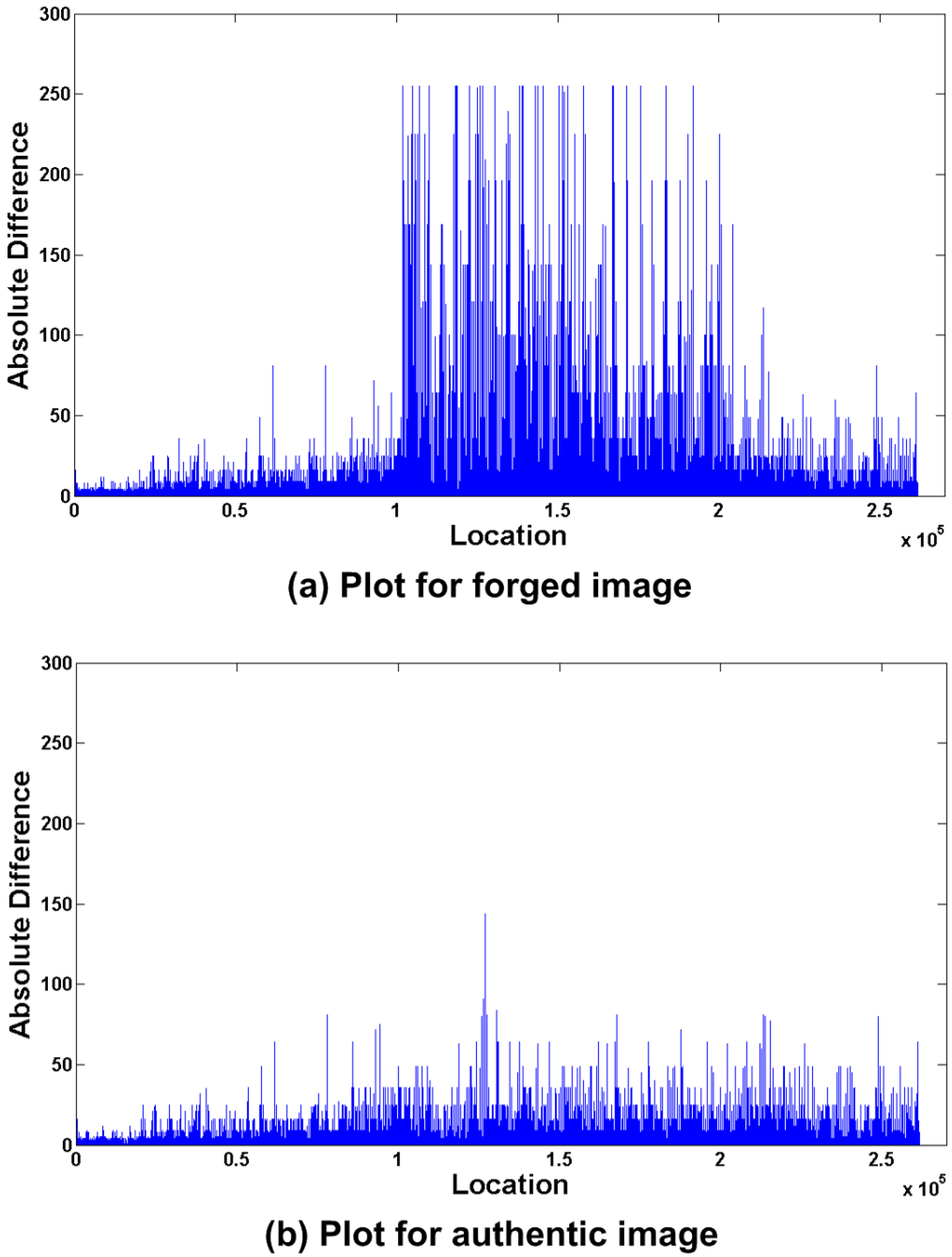

Figure 3.2: *Lena* JPEG plots: (a) Original 512×512 image plot; (b) Forged Lena image plot in which central part of the image is tampered which represents sudden hike.

while the rest of the part of the image is doubly compressed because it is already gone through compression when it is first saved to the memory.

Note that, the technique which we propose in our work [16] is a *blind* forgery detection technique, where the forensics analyst need neither the original image nor any pre–computed information from the original image, for the detection of forgery. The only information a forensic analyzer need in this *blind* forgery detection technique is the (possible) forged image.

3.2 Summary

In this chapter, we discussed about the proposed scheme "Digital Forensic Technique for Multiple compression based JPEG Forgery in the image" and the application of the technique.

Chapter 4

Results and Disscussion

4.1 Results

The technique which we proposed is implemented in the Matlab. This is excellent tool for implementing the image related problems as the image acts as a matrix(2-Dimensional Matrix). We have used the Matlab Image Processing Toolbox in our experiment in order to re-write the image in the memory, with different quality factor by using imwrite command. Also, for other operations which in necessary for the working of the algorithm and producing appropriate results. We use standard images for the experimental purpose which are the standard images in the image processing environment. We have taken six different standard images on which we applied our technique, which is shown in the following Fig. 4.1.

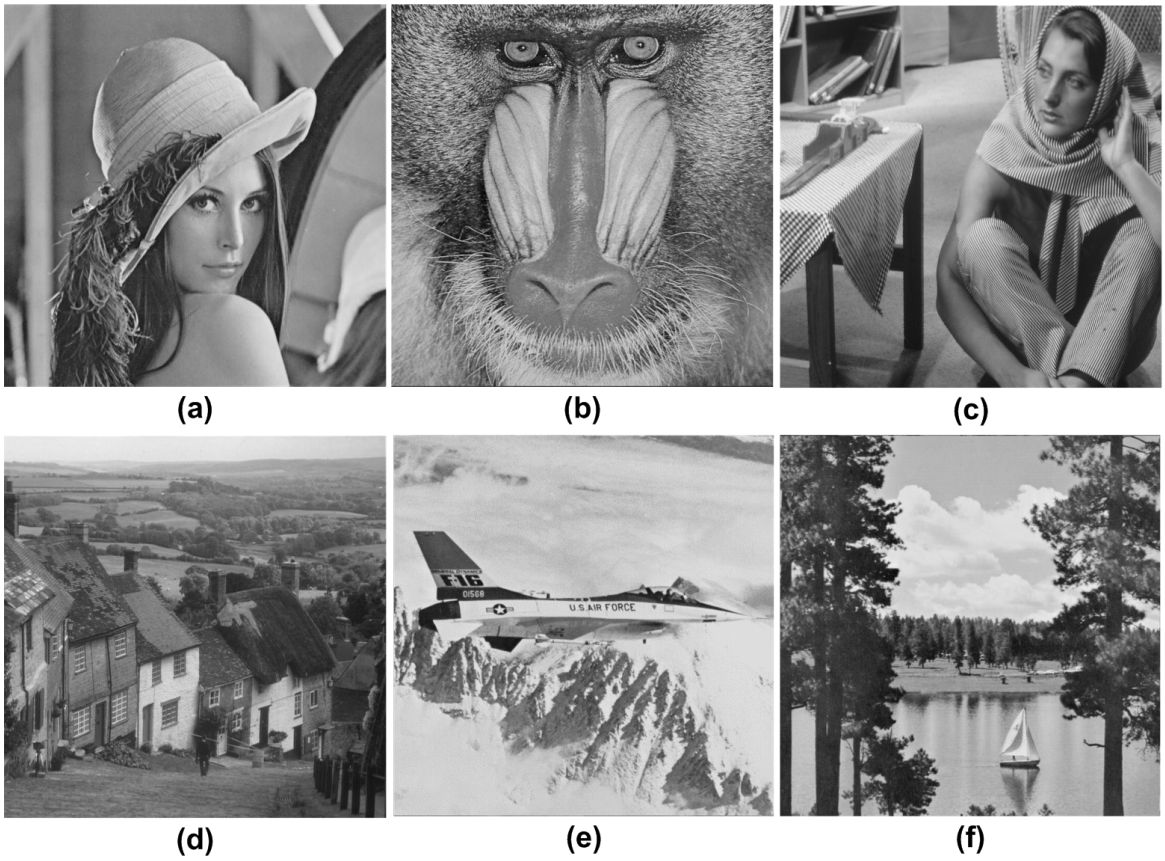

Figure 4.1: 512×512 *Test Images* JPEG images: (a) Lena ; (b) Mandrill ; (c) Barbara ; (d) Goldhill ; (e) Plane ; (f) Sailboat.

4.1.1 Manual Tampering in Image

For the experimental purpose in our technique, we manually tampered the image in order to test the algorithm. We tampered the central region of the image with some part of the other image with the different compression quality for the detection of the image forgery in the image. For this manual tampering in the image and to produced the tampered image we have gone through these steps :

1. Firstly, we extract the central region of the image which is 200×200 from the testing images.

2. Then we save the extracted part of the image separately in the memory with the different compression quality than the original image using the **imwrite** function of the Matlab.

3. The part which we have re-compressed and saved separately is now transplanted in the original image.

Now, with these manually forged or modified images we have analyzed our detection technique to identify the tampering in the image. We have analyzed the technique with the six different standard image processing images.

4.1.2 Selection of Compression Quality

As we discussed earlier in the section 3, different values which we get from the forged image with the re-compression factor of q', different squared error matrix is provided that is $I_{q'}$. Here, we determined that the optimal squared error matrix is one from which the tampered or modified regions are clearly identifiable or visible. For all our test image which we analyzed the algorithm with, the re compression quality factor which generates the optimal squared error matrix lies somewhere between [60,85]. For the test image Lena, the squared error matrix is shown in the fig. 4.2 at the varying degree of re compression quality q'. Here in this image, the optimal squared error matrix is found in which the quality factor is 70, where the tampered region is clearly visible. For each of the test image, we have investigated

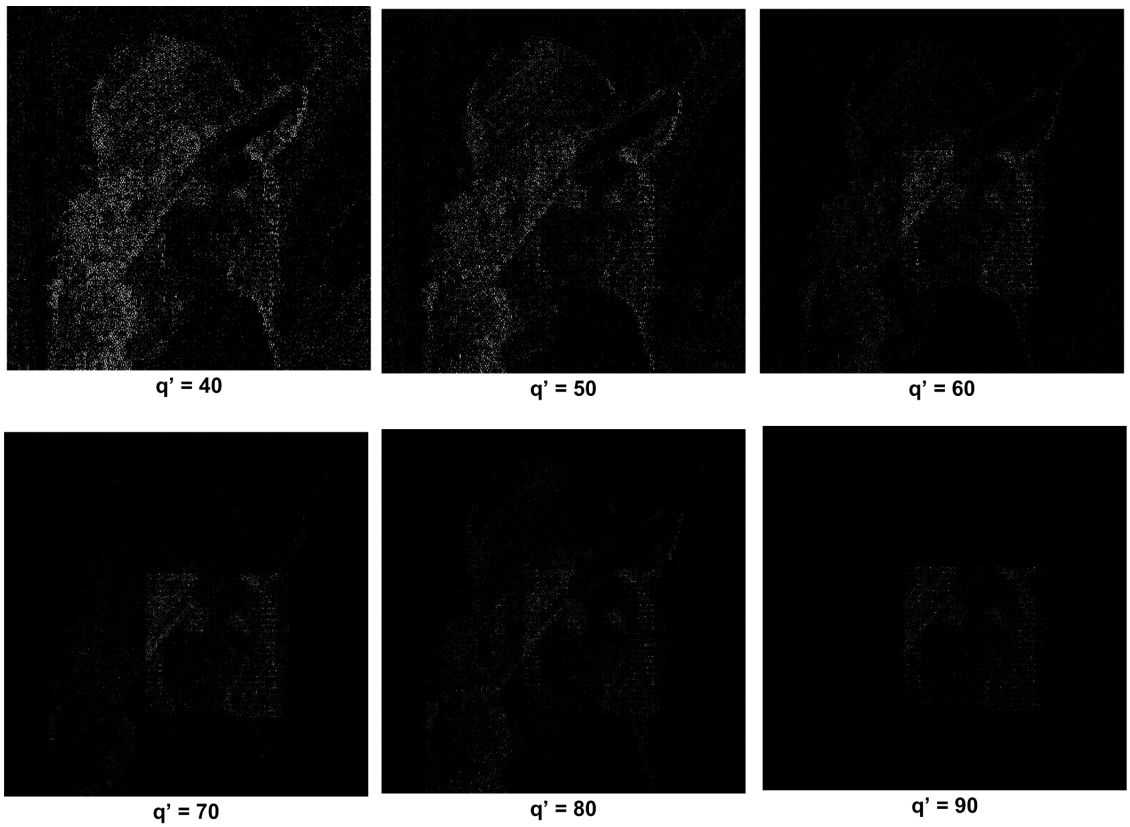

Figure 4.2: Shows the Squared error matrix of the test image : lena, at varying degrees of compression quality q'

the JPEG forgery based on the double or multiple compression by analyzing and studying the D_2 vs P plot characteristics, which is corresponding to the optimal squared error matrix which is $q' \in [65, 85]$. The plots for each image which we generated is shown in fig. 4.1 for our test images which are Lena, Mandrill, Barbara, Goldhill, Plane, Sailboat respectively.

4.1.3 Detection of Forgery

We have analyzed our technique for all the standard image processing images and generated results are shown in the fig. 4.3. These Plots proves that the images which is being tampered(the central part/region of the image is modified) has a sudden hike in the pixel values(absolute squared error pixel pair difference) and this hike is persistent for the range of pixels which has been modified during the forgery. However this JPEG characteristics is inborn in JPEG images containing

regions with double or multiple compression. Such characteristics is absent in the JPEG images which are not modified. This is evident from the D_2 vs. P plots corresponding to our original test images and authentic test images, as shown in Fig. 4.4.

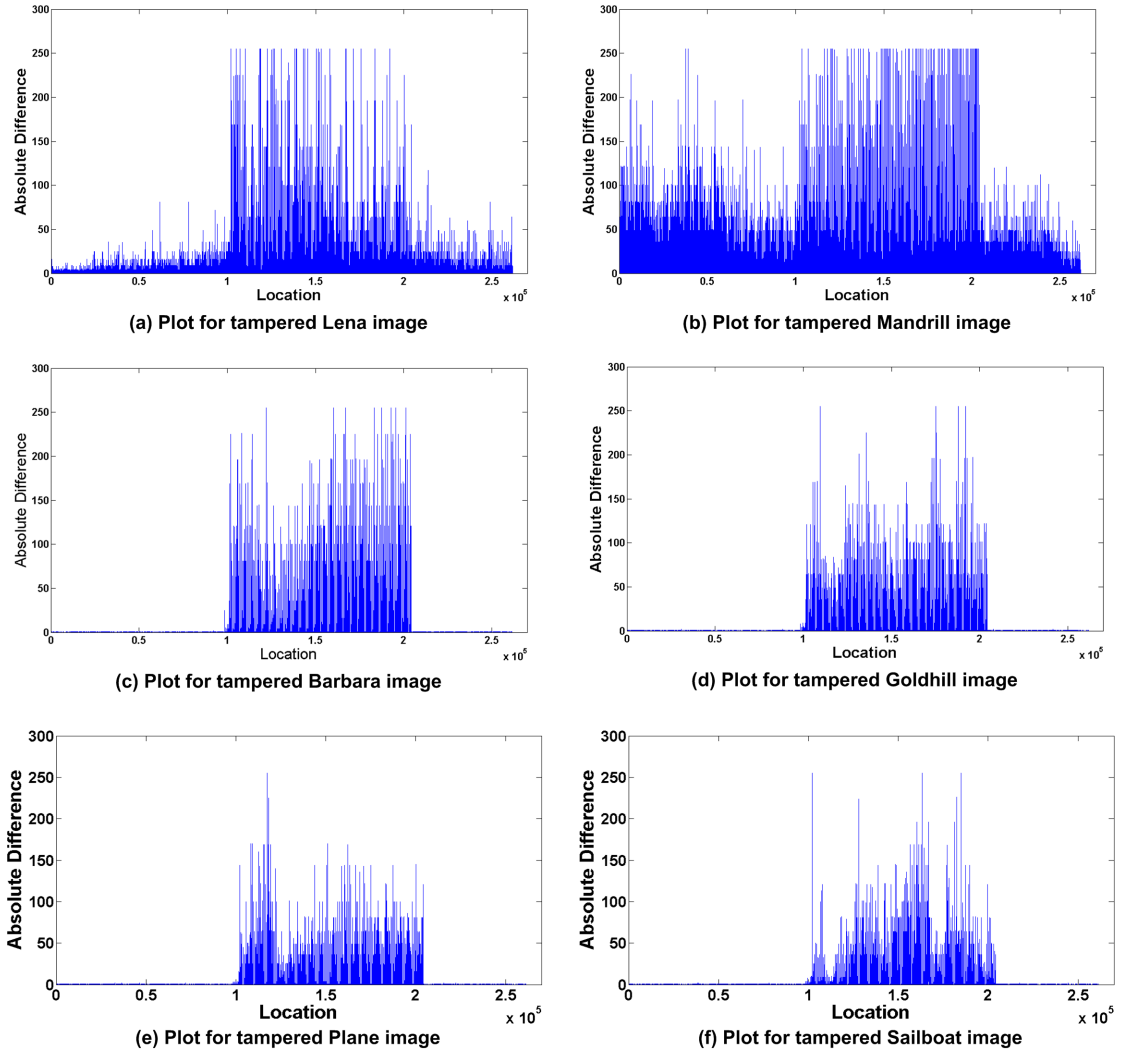

Figure 4.3: Absolute squared–error pixel–pair differences vs. pixel–pair locations: (a) Plot for forged *Lena* image; (b) Plot for forged *Mandrill* image; (c) Plot for forged *Barbara* image; (d) Plot for forged *Goldhill* image; (e) Plot for forged *Plane* image; (f) Plot for forged *Sailboat* image.

Note that the plots corresponding to authentic *Barbara* and *Goldhill* images (Fig. 4.4 (c) and (d) respectively) demonstrate considerable uniformity in

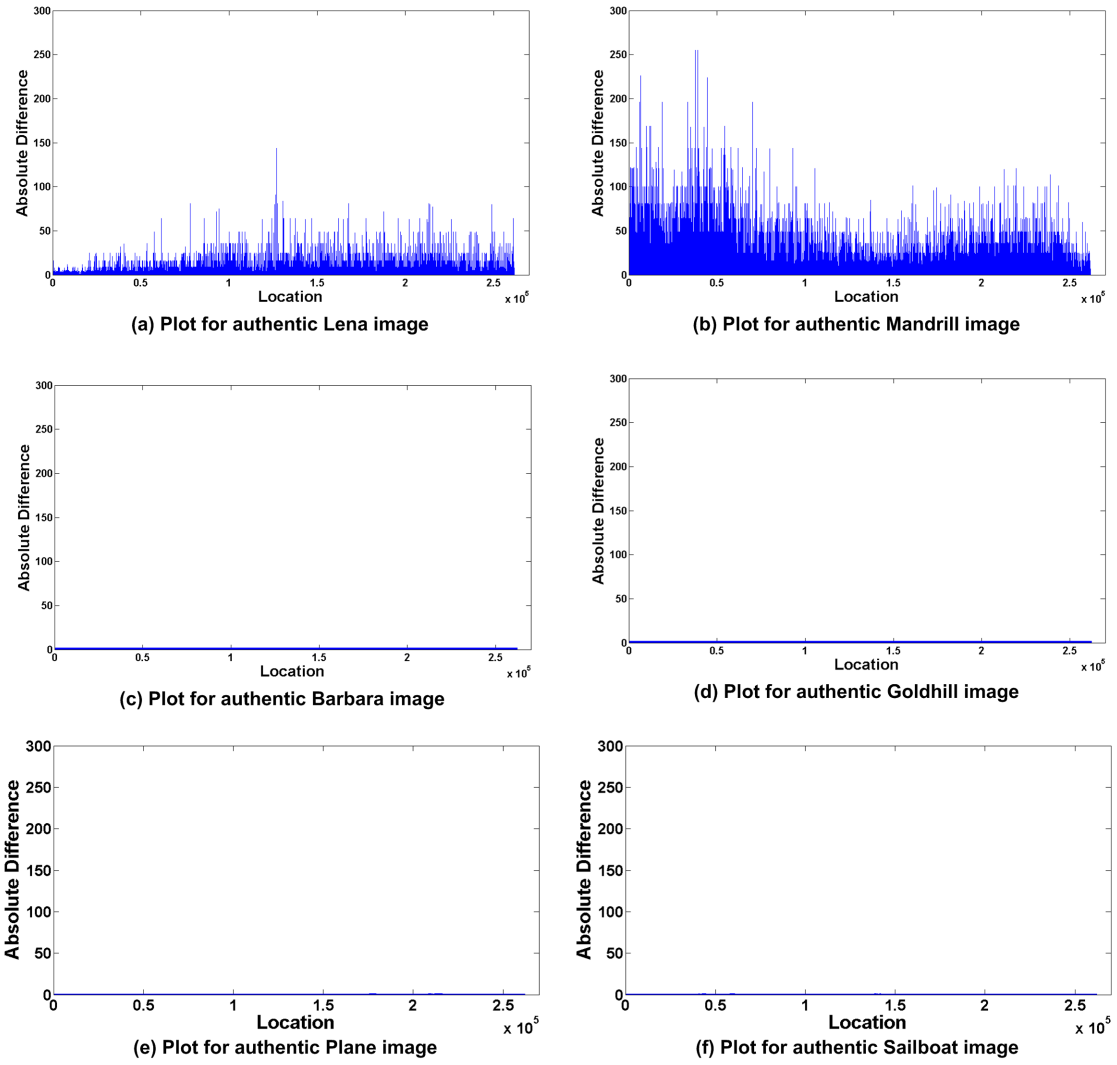

Figure 4.4: Absolute squared–error pixel–pair differences vs. pixel–pair locations: (a) Plot for authentic *Lena* image; (b) Plot for authentic *Mandrill* image; (c) Plot for authentic *Barbara* image; (d) Plot for authentic *Goldhill* image; (e) Plot for authentic *Plane* image; (f) Plot for authentic *Sailboat* image.

difference values (≈ 0) over the entire range of pixel–pair locations, as compared to *Lena* and *Mandrill* (Fig. 4.4 (a) and (b) respectively). This is due to high correlation among neighboring pixels, present in *Barbara* and *Goldhill* images. Due to minimum inter–pixel correlation in *Mandrill*, the plot of Fig. 4.4(b) shows the maximum randomness in variation of absolute difference. However, irrespective of inter–pixel correlation among neighboring pixels, a forged JPEG

image always demonstrates consistent D_2 vs. P characteristics, as is evident from Fig. 4.3 that the image is being tampered with because the forged portion of the image shows the high values between the pixel pairs and this values are persistent.

4.2 Summary

In this chapter, we discussed about the results which we get as output on test image and application of the proposed scheme into some standard image processing images and their results.

Chapter 5

Conclusion and Future Work

5.1 Conclusion

In this thesis, a digital forensics technique for the detection of the JPEG image forgery is presented, which is a resulting image from modifying the image and re-writing it to memory by an adversary. The proposed technique exploits the feature of multiple or double compression, inherent in tampered JPEG images. Our experimental data is a proof of the forgery detection efficiency of the proposed technique. The proposed scheme is completely blind that is there is no need of any pre-computed information for the scheme to work. the only requirement is forged image.

5.2 Future Work

Automation of quality factor determination is the future direction for this research work. Utilizing the forensic results and data related to JPEG forgery, recovery of forged JPEG regions may also be investigated in the future. Reconstruction of the forged portion close to the original may also be an option because the we know exactly which part of the image is being tampered.

In our work, we are able to identify the tampered region of the image. Here, we have been able to collect some information about exactly which part, how much part and what part of the image is tampered. Now, it may be possible to to reconstruct the tampered part of the image on the basis of the rest of the part/region which is authentic. We are aware of the fact that the JPEG is the lossy compression standard, due to the compression in the image the information is already lost and cannot be retrieved back. In future, it may be possible to reconstruct the tampered part of the image back, it is not entirely not possible to obtain 100 percent of the data back because of the compression in the image. It may be possible to reconstruct the tampered region of the image with the help of the neighbouring pixels of the image. On the basis of the neighbouring pixel values, it is possible to create the tampered region pixel values

DISSEMINATION OF WORK

1. **P. Malviya** and R Naskar, "Digital Forensic Technique for Double Compression based JPEG Image Forgery Detection", International Conference on Information Systems Security (ICISS), Hyderabad, India, 2014.Published in Lecture Notes in Computer Science, Springer, vol. 8880, pp. 437-447, 2014.

2. R. Naskar, **P. Malviya** and R. S. Chakraborty, "Digital Forensics: State-of-the-Art and Open Problems", in Dr. R. Pal (ed.), "Innovative Research in Attention modeling for Computer Vision Applications", IGI Global. (Forthcoming)

Bibliography

[1] H.T. Sencar and N. Memon, (eds.), "Digital Image Forensics: There is More to a Picture than Meets the Eye", New York, NY, USA: Springer, 2013.

[2] J. Redi, W. Taktak, and J.L. Dugelay, "Digital Image Forensics: A Booklet for Beginners", *Multimedia Tools and Applications*, vol. 51, no. 1, pp. 133-162, Jan. 2011.

[3] H.T. Sencar and N. Memon, "Overview of state-of-the-art in digital image forensics", *Indian Statistical Institute Platinum Jubilee Monograph series titled Statistical Science and Interdisciplinary Research*, Singapore: World Scientific, 2008.

[4] J.F. Lalonde and A.A. Efros, "Using color compatibility for assessing image realism", *Proceedings of the International Conference on Computer Vision*, 2007.

[5] N. Wang and W. Doube "How real is really a perceptually motivated system for quantifying visual realism in digital images", *Proceedings of the IEEE International Conference on Multimedia and Signal Processing*, pp. 141-149, 2011.

[6] H. Farid and S. Lyu, "Higher–order wavelet statistics and their application to digital forensics", *IEEE Workshop on Statistical Analysis in Computer Vision*, 2003.

[7] G. Wallace, "The JPEG still picture compression standard", *IEEE Transactions on Consumer Electronics*, vol. 34, no. 4, pp. 30-44, 1991.

[8] A.J. Fridrich , B.D. Soukal , A.J. Luk, "Detection of copy-move forgery in digital images", *Proceedings of Digital Forensic Research Workshop*, 2003.

[9] 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.

[10] J. Wu, M.V. Kamath, S. Poehlman, "Detecting differences between photographs and computer generated images", *Proceedings of the 24th IASTED International conference on Signal Processing, Pattern Recognition, and Applications*, pp 268-273, 2006.

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

[12] X. Pan and S. Lyu, "Detecting image duplication using SIFT features", *Proceedings of IEEE ICASSP*, 2010.

[13] H. Huang, W. Guo and Y. Zhang, "Detection of copy–move forgery in digital images using SIFT algorithm", *IEEE Pacific–Asia Workshop on Computational Intelligence and Industrial Application*, 2008.

[14] H. Farid, "Exposing digital forgeries from JPEG ghosts", *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 1, pp. 154-160, Mar. 2009.

[15] F. Zach, C. Riess and E. Angelopoulou, "Automated Image Forgery Detection through Classification of JPEG Ghosts", *Proceedings of the German Association for Pattern Recognition (DAGM 2012)*, pp. 185-194, Aug. 2012.

[16] P. Malviya and R. Naskar, "Digital Forensic Technique for Double Compression based JPEG Image Forgery Detection", *International Conference on Information Systems Security (ICISS) , LNCS,* vol. 8880, pp. 437-447, 2014.