DETECTION OF MANIPULATED IMAGES

A Thesis by

Santoshi Senapathi

Bachelor of Technology, JUNTU, India, 2006

Submitted to the Department of Electrical Engineering
and the faculty of the Graduate School of
Wichita State University
in partial fulfillment of
the requirements for the degree of
Master of Science

May 2009

DETECTION OF MANIPULATED IMAGES

The following faculty members have examined the final copy of this thesis form and content, and recommended that it be accepted in partial fulfillment of the requirement for the degree of Master of Science with a major in Electrical Engineering.

_____

Ravi Pendse, Committee Chair

_____

Edwin Sawan, Committee Member

_____

Krishna Krishnan, Committee Member

DEDICATION

My parents, family members and friends

ACKNOWLEDGEMENT

I would like to express my heartfelt and sincere thanks to my advisor Dr. Kamesh Namuduri for his guidance and remarkable patience all through my research and valuable suggestions that helped to complete my thesis successfully. I'm grateful to my committee chair Dr. Ravi Pendse for his generous help, guidance, and patience throughout my Master's program and also for his time to review the report.

I'm also grateful to the committee members who dedicated some time from their busy schedule to review my report and provide me with valuable suggestions

I shall always be indebted to my parents for making me capable to complete this thesis. I am also thankful to my friends, who helped me and made my stay at WSU memorable

ABSTRACT

When compared to the existing image formats, JPEG is the most widely used format that stores digital images using cameras and software based algorithms. Digital images have widespread use in applications such as surveillance, military, scientific, etc. With the availability of tools on the internet it is easy to manipulate the digital images. For instance, a true image when manipulated becomes useless in certain applications. An image that is supposed to be the source of evidence, if manipulated will be of no use in some applications. Hence there arises a need to verify the originality of digital images without the availability of the original image. In this thesis, we have developed a method to test the originality of a given image. The verification methods are based on computing DCT (Discrete Cosine Transform) for a given image. The pdf plots for the obtained DCT coefficients are plotted to observe the characteristics of the plot with the assumed Laplace distribution function. Later, the parameters of the distribution functions are estimated and computed using Maximum Likelihood Estimation, in order to test whether a given image is natural or manipulated.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

LIST OF FIGURES

Figure Page

LIST OF ACRONYMS

BAM	Blocking Artifact Measure
BSM	Binary Similarity Measure
CFS	Core Feature Set
DCT	Discrete Cosine Transform
DQ	Double Quantization
HOWS	Higher Order Wavelet Statistics
IQM	Image Quality Metrics
JFS	Joint Feature Set
JPEG	Joint Photographic Experts Group
LSB	Least Significant Bit
MLE	Maximum Likelihood Estimation
MSB	Most Significant Bit
PCA	Principal Component Analysis
PDF	Probability Distribution Function
PRCG	Photo Realistic Computer Generated
ROI	Region of Interest
SFFS	Sequential Forward Floating Search
SVM	Support Vector Machine

LIST OF SYMBOLS

μ	Mean
$\sum$	Summation
Log	Logarithm

CHAPTER 1
INTRODUCTION

In today's world, digital images are being widely used in numerous applications such as military, intelligence, surveillance, digital copy right applications, etc. Among the existing image formats, JPEG is the most widely used format that stores the digital images using digital cameras and software tools.

1.1 Digital Image Representation:

A digital image is represented in two-dimensional form and referred to as raster images. Raster images are represented with a finite set of elements called pixels and depending on the type of each pixel, digital images are classified into gray scale, color, index, binary, etc. A digital image is represented as:

$$F(x, y) = i(x, y)\, r(x, y) \qquad (1)$$

Where (x, y) represents the coordinates of the pixels, 'i' is the illumination and 'r' is the reflectance of an image. And a gray-scale image is depicted as below:

$$l = f(x_0, y_0) \qquad (2)$$

$$L_{min} \leq l \leq Lmax \qquad (3)$$

And the interval $[L_{min}, Lmax]$ is called the gray-scale.

1.2 Problem Description

Using the manipulation tools that are available on the internet it is easy to tamper the digital images without any trace. Therefore, verification of originality of images has become a challenging task. The early research in image forensics introduced digital watermarking and robust hashing in the original image for authentication.

From [1], “Digital watermarking is defined as a technique to embed the signal permanently into the digital data which is mainly used for copy right protection and can be extracted by means of computing operations to analyze the data”. It is also defined as, “a communication channel between a sender and a receiver to transmit the useful information”. Watermark was used successfully in digital rights management, where the usage of digital media and devices are limited by manufacturers and copyright holders.

Robust hashing is a method to insert time, date or model of the camera during the capture or creation of the digital images. Even if digital watermark or robust hashing is embedded into the digital image, there is a high probability that an attacker might attack the data. When the detector receives the data, the receiver has to detect whether the image embedded with watermark is original or manipulated.

1.3 Digital Watermarking Process:

In digital watermarking, the user embeds the data (which is to be hidden) into the original image ‘I’, with the help of a secret key ‘k’ to produce a watermarked data ‘Iw’. The watermarked data is transported through the communication channel where an attacker attacks the data intentionally or unintentionally and manipulates the data to produce a manipulated data $I_w`$. The receiver extracts the watermark present in the data with the help of a shared secret key to detect the watermark present in $I_w`$. The details of the watermark should be transparent to the receiver so that the true details of the watermark are hidden. This is obtained by executing the zero knowledge protocols. As the true details of the watermark are hidden, the receiver will not know whether the transmitted image is original or tampered. Zero-Knowledge protocols were proved to be applicable only if the embedded watermark is not modified. Most of the digital images created today do not contain digital watermark and robust hashing. Thus, there exists a need for an intensive approach to differentiate real and synthetic images, which in turn may or may not contain the digital watermark and robust hashing methods.

An image can be manipulated with a wide variety of manipulation techniques such as scaling, rotation, blurring, resampling, compression, cropping, filtering, image doctoring (i.e. copy/paste a part of an image into the original image) contrast and brightness adjustments, etc. The methods for differentiating real and tampered images should answer the following questions:

- Does the given image represent a true form of the original image or doctoral image?
- If doctored, what types of manipulations are done?

The above questions are just the instances, and there are many more things to be considered to prove the integrity of images. Methods to detect the problem of image tampering involves wavelet analysis of frequency sub-bands, image block analysis using DCT (Discrete Coefficient Transform), study of blocking artifacts and noise statistics variations. The methods addressed by several authors so far were successful in detecting original and manipulated images, but each had its own limitations.

1.4 Significance of this thesis

The verification of originality of images is required in a variety of applications such as military, surveillance, copy right protection, forensic, scientific, etc. The reason for testing the natural and manipulated images is that an image that is supposed to be the source of authentication if manipulated will be useless. Thus, there exists a need for a method to prove whether a given image is natural or manipulated. The approach used in this thesis to differentiate original and manipulated images involves DCT computation for a given image. The pdf plots for the original and manipulated images are obtained and these plots are compared with Laplace distribution since, DCT coefficients follow Laplace distribution function. The parameters to test the originality of images are then computed using Maximum Likelihood Estimation.

1.5 Thesis Organization

The organization of this thesis is as follows.

Chapter 2 provides a brief survey of existing methods that distinguish original and synthetic images.
Chapter 3 presents my approach to detect the differences between original and synthetic images by computing DCT for a given image. Furthermore, this chapter explains the method used for testing the originality of an image.
Chapter 4 contains simulation results.
Chapter 5 draws conclusion and suggests future work

CHAPTER 2

LITERATURE REVIEW

This chapter provides a brief survey of relevant research efforts for the detection of original and tampered images.

The early research in the field of image processing to detect the original and manipulated images introduced the idea of embedding digital watermark and robust hashing in digital images (in [1]). With the availability of image manipulation tools and with the rapid increase in creation of digital images, it has become difficult to embed the digital watermarks. Hence, there arose a need for several new methods to differentiate the originality of images.

The methods discussed (in [2, 3]) followed different methods to detect the images tampered with doctoring but the underlying methods for testing the images remained the same. The method (in [2]) was based on the analysis that a doctored or tampered image might have undergone some manipulations such as rotation, scaling, re-sampling, cropping, etc. To detect the original and tampered images a classifier was designed to measure the distortion before and after the manipulation. The features used for the classifier were, the measurements obtained by measuring the distortion between real and tampered images. SFFS (Sequential Forward Floating Search) method was used to select the features that were obtained by measurements of two first-order moments of the angular correlation and Czenakowski measure. Features with independent content (in the original image) were manipulated, to enable the classifier to detect the differences. The features used for the classifier were employed by Mul1le image quality metrics. The reason for using Mul1le quality metrics is that an image might encounter different manipulations while doctoring, but some respond at pixel level, others at block level and few others to edge and spectral phase distortions. The obtained features were then trained to a linear regression classifier to detect the type of manipulations in an image. If the classifier detects a specific part of an image differently than the other part of the image, the image was classified as manipulated. The classifier was successful in detecting the

type of manipulations on images that were manipulated by blurring, rotating, scaling, brightness, doctoring and contrast adjustments. It was also successful in differentiating original and tampered images. In order to detect images generated by highly sophisticated tools and computing algorithms, and to detect the manipulation in color images, a highly sophisticated classifier is needed.

Popescu (In [3]) proposed a method to detect duplicate regions using PCA (Principal component Analysis) with a grayscale image of size N pixels, blocks of b pixels (which have $\sqrt{b}x\sqrt{b}$ dimensions but is less than the size of the duplicated regions). These blocks were represented in vector form. The blocks in vector were assumed to have zero mean and variance, and covariance of the matrix was computed. The Eigen vectors, e_j of the covariance matrix with corresponding Eigen values λ_j that define principal component values were calculated. These Eigen vectors form a basis for each image block i.e. represented as a vector. Using PCA, new N_t dimensional for each b pixel was calculated and $N_b$xb matrix was built and rows were quantized by a_i/Q where Q is the quantized coefficient. In the new matrix, rows were arranged in column order of the matrix $N_b$xb and s_i denoted the i^{th} row of this sorted matrix. The tuple (x_i , y_i) denoted the blocks of image coordinates. Finally, the row distance $|i-j|$ in the sorted matrix less than the threshold value was considered.

The duplicate regions in an image were detected using a duplication map. The duplication map was created by assuming a zero image which was same as the size of the original image and all the pixels that were suspected as duplicate were assigned a unique gray scale value. The method was successful in detecting the forgeries in a digital image but the complexity of this algorithm increased by N_tN logN.

Detection of manipulations on images requires a deep understanding of the pixel characteristics, as doctoring or any type of manipulations do not leave any trace of identification that an image is manipulated. Therefore, it becomes important to select the features that are susceptible to manipulations. To identify the differences between original and synthetic images, numerous tools have been developed based upon the

correlations between the bit planes and binary texture characteristics of the bit planes using BSM's (Binary Similarity Measures). The features obtained with BSM's were compared with other image forensic tools such as IQM's (Image Quality Measures) and HOW's (Higher Order Wavelet Statistics) and later a classifier was designed for classification.

The process of feature selection was done using SFFS method (as in [2]) which provides a way to select the best features leaving the plethoric ones for classification. The procedure used for selecting the features is as follows:

1. From the set of 'n' features, select two best features that result in best classification.
2. From the remaining (n-1) features available, add the most significant feature such that the selected feature supplements the classification.
3. Now, identify the least significant feature among the selected features and either add or remove that feature to see if it improves the selection. If yes, identify the other least significant feature until there is no improvement. If not repeat step 2.
4. If the number of features selected becomes equal to the number of features required, stop the search method.

The above method was run on each type of manipulation and also on a class of pooled manipulations and then trained to a classifier. When experiments were performed on different manipulations that were selected by SFFS algorithm, JFS (which is the combination of all the feature sets) outperformed the other forensic features. The works of Bayram et al. (in [6]) were also based on these forensic features which were trained to a classifier for classification. When experiments were conducted on images that were manipulated using various manipulation operations, JFS provided a detection accuracy of 90% among the entire features selected.

Farid et al approach (in [5]) explained the DQ (Double Quantization) effect in DCT coefficients to locate the doctored part in images. A given JPEG image was divided into

8x8 DCT blocks and DCT was computed for each block. The DCT coefficients were divided with different quantization steps and curved to nearest integer. Using DCT coefficients, histograms were computed and the quantization steps for different frequencies were stored in quantization matrices. With the histograms obtained by DCT coefficients, probabilities were fused to give the normality of that block. The normality map was set to a threshold value to decide whether an image is doctored or undoctored. With the segmentation, a feature vector was computed and the obtained features were trained using Support Vector Machine (SVM). This algorithm has several advantages when compared to other detection methods, i.e. it can detect images doctored with different kinds of manipulations such as alpha matting, texture synthesis and other tools. This method used less computational cost and memory.

The algorithm developed by Farid et al (in [5]) had a limitation in which 1. If the original image that is doctored or undoctored is not a JPEG image 2. If the compression quality of the original JPEG image is Q1 and after image forgery, the image is saved with a different compression quality Q2 and the ratio of Q2/Q1 is smaller, then the DQ effect becomes invisible and difficult to detect the doctored part in image. 3. If the original image itself is compressed twice, then the presence of DQ effect may not be useful to detect the doctored part. An alternative method was proposed by observing the DCT coefficients in detail. Also, it was observed that if a JPEG image does not contain both the doctored and undoctored part, the histogram of undoctored part will still exhibit DQ effect because of the compression effect of the original image. The algorithm also tested that the histogram of the doctored part sometimes will not have DQ effect under following conditions:

1. If the doctored part of the image is copied from a bitmap image and pasted onto a JPEG image, then the doctored image will not undergo the first compression of the JPEG image.

2. Suppose in the original JPEG image itself, if the blocks of the image were swapped with other blocks, neither of the blocks would exhibit DQ effect.

3. There is less chance that the doctored part will have 8x8 blocks. Along the boundaries of the doctored part, there will be a few pixels which will be in the doctored part and the undoctored part. So, any processing operations along the boundaries of the doctored part will not follow the DQ effect.

On the whole, if a doctored image is manipulated with image manipulation operations, it will cause the loss of DQ effect. Therefore, histograms of the doctored image can be considered as superimposed of high peaks and valleys. Thus, DQ effect was proved to detect doctored images, but it had a converse effect that the designed algorithm was easily attacked.

Works of Lyu et al (in [6]) were also based on statistics of the features extracted from HOW's forensic feature set for three-level discrete wavelet coefficients. In this, the features were extracted by computing mean, variance, skewness and kurtosis of sub-band coefficients at each orientation and scale. The errors in coefficient magnitudes of all spatial, scale and orientation were computed to get higher order statistics for each scale and orientation, thus resulting in 72 features. When experiments were performed on real and PRCG (Photo Realistic Computer Generated) images produced an accuracy of 67% with 1% false-alarm rate.

Neg et al works (in [6]) showed the differences between real and synthetic images based on geometrical characteristics. His work focused on differences among objects and surface models and the acquisition differences between PRCG and natural images. He extracted features from the PRCG images based on polygonal surface models and simple light transport models, that do not have acquisition characteristics i.e. device name, scanners etc. The selected features were extracted from local patch statistics, local fractal dimensions and normalized geometric characteristics such as surface gradient and Beltrami flow. Experiments when conducted on PRCG and real images, they produced an accuracy of 80%, compared to the approach proposed by Lyu et al.

Dehnie at el (in [6]) focused on noise patterns of the imaging sensors. Dehnie observed that each digital camera will have unique noise patterns, but the basic technology of the image sensors remains the same since they exhibit unique statistical characteristics. It was observed that PRCG images which require some tools and methods to create or manipulate, add additive noise to the images. Therefore, they differ in their statistical characteristics. When experiments were conducted on PRCG and denoised real images, the resultant noise residues were calculated for first order statistics like skewness and kurtosis. This showed noticeable differences and hence it was easy to distinguish two images with an accuracy of 75%. Methods proposed by different authors concentrated on variations in image features that also aided in differentiating real and synthetic images. The following section deals how to detect the variations in image features that would help one to authenticate the original images.

2.1 Variations In Image Features:

The above approaches corresponded to the features that were susceptible to manipulations and provided an insight to differentiate original and manipulated images. The values obtained were preserved and later used on PRCG images to make a decision, whether the features obtained from these images deviated from the actual preserved features. These approaches were based on designing a classifier to make a decision. To discriminate real and PRCG images, Avcibas proposed a method to obtain image quality metrics as discussed (in [2]). The features which were selected using image quality metrics were susceptible to image manipulations. These features were used in conjunction with the classifiers on the basis of pixel and block level differences, edge distortions, and spectral phase distortions to ensure that the selected features were susceptible to image manipulations only. Experiments on altered images with various manipulation operations provided an accuracy of 80% and later training the PRCG images to classifiers yielded an accuracy of 74%.

Based on the inspection that non-linear processing operations exhibit high order correlations, Neg et al (in [6]) identified the effects of image splicing on magnitude and

phase of bicoherence spectrum, i.e. normalized bispectrum of Fourier Transform of a third order moment signal. The authors observed that at splicing point bipolar signals deviate from a smooth curve and proved that bipolar signals correspond to the changes in bicoherence spectrum of the signal. The experiments on magnitude and phase features of images provided an accuracy of 62%. Subsequent works of authors on the new bicoherence features that included edge pixel density and variations in bicoherence features in original and manipulated images along with the existing bicoherence features provided an accuracy of 72%. Authors also showed that along with variations in image features, inconsistencies in image variations can also be used to detect the original and manipulated images.

2.2 Image Feature Inconsistencies:

This section deals with an approach that includes features based on the variations obtained by the abrupt deviations from the norm of the image or unexpected similarities over the image. These methods elevate the presence of double JPEG compressed images. Compressing an image that is already compressed reduces the smoothness of DCT coefficient histograms and creates eye-catching patterns in DCT coefficient histograms. If the step size of second quantization is made smaller, few bins in the resultant histogram becomes zero. If the quantization step size is made higher than the second quantization step size, the DCT coefficients of histogram show periodic peak patterns. This method has been used to determine the initial compression parameters and detect the double compressed images. When images were spliced at different quality factors the spliced parts in the re compressed images had different double compression characteristics when compared to compressed images. He et al (in [6]) designed an algorithm to detect or identify the locations of manipulations in images by analyzing the histogram coefficients of each DCT channel for double compressed image and assigned the probabilities to each 8x8 DCT blocks. Later, the probabilities for each block were fused to obtain the normality map of the blocks to detect the presence of manipulations on the map. Algorithm used on manipulated images was successful. Popescu et al (in [6]) developed a method to detect the images that were resized. The basic idea of this method was that up-

sampling introduces periodic interpolation coefficient correlations and re-sampling at different rates in turn requires up-sampling and down-sampling operations to attain the required rate. The presence of correlations between pixels was used to determine the images that were resized. To study the exact form of correlations, authors proposed probabilistic models to predict the errors of interpolated and un interpolated coefficients. Experiments on images that were subjected to different types of manipulations provided accuracy close to 100%.

Image manipulations often include sharpness or blurriness adjustments. The characteristics of blurriness on altered images were expected to vary from an unaltered image. Works of Sutcu et al (in [6]) made use of regularity properties of wavelet transform coefficients to detect the variations in edges due to sharpness or blurriness. This method used edge detection algorithm to analyze the edge locations, followed by a multi-scale wavelet decomposition of the image. The edge locations were identified by analyzing the edge image and their corresponding maximum amplitude values of wavelet sub band signals. Using the linear curve fitting the log of the maximum amplitude values were plotted to test the goodness of the fit. This method was tested on images that were globally blurred and altered with local adjustments.

The most commonly used method in image/video processing applications, such as bit allocations and pre-processing techniques is Blocking Artifact Measurement (BAM). Various methods have been proposed to estimate the quantization table on the basis of total blocking artifact at a particular DCT frequency and at a given quantization step using MLE (Maximum Likelihood Estimation). Using Benford's law, the probability distribution for the first digits of the JPEG coefficients was used to estimate the JPEG quantization tables which were time consuming and require an extensive search algorithm. Hence the proposed method (in [7]) to estimate the quantization table was faster than the normal MLE or Benford's law.

2.3 Estimation of Quantization Table:

It was assumed that if the DCT coefficients are quantized with a certain step size, they exhibit periodic patterns in the histogram. If Fourier transform is computed for the histogram of DCT coefficients, then sharp peaks at mid and high frequencies are observed. Hence the quantization table (in [7]) was estimated by finding the derivative of the power spectrum of the histogram of DCT coefficients. Following are the steps performed to estimate the quantization table for the whole image:

1. The image was first divided into 8x8 blocks and DCT was computed for each block.
2. Histogram was calculated for each of the DCT coefficients.
3. Power Spectrum of the histogram of the DCT coefficients was calculated.
4. Then, the second derivative for the power spectrum computed in step 3 was calculated and low pass filter was applied.
5. Local minimum number (NUM) of the second derivative power spectrum was calculated.
6. The Quantization step for the DCT frequency was estimated as NUM + 1.

2.4 Image Forensics Analysis using Blocked Artifact Measure:

Image authentication was done with the help of blocking artifacts that were introduced during JPEG compression. When an image was manipulated using different manipulation operations, inconsistencies of blocking artifact was used to differentiate whether an image is original or tampered. Synthetic images can be identified with the help of inconsistencies of blocking artifact. For authenticating the images, the image was first divided into blocks and then blocking artifact consistencies for each block was verified. The region or part in an image suspected as manipulated was evaluated and other blocks

were estimated using the quantization table. Then, BAM for the entire image was calculated with the estimated quantization table. If there were any blocking artifact inconsistencies, then the image was classified as manipulated else, original. Therefore, with the inconsistencies in blocks of the image, one could test whether an image is original or manipulated.

Although the methods discussed (in [3, 4] and [7]) for detecting specific type of manipulations and the classifier based approach (in [2, 4]) which was used to differentiate real and tampered images were successful, still there was a need for an extensive search method to prove the originality of the digital images. These methods provided a foundation for digital image analysis irrespective of the type of manipulations involved. Features such as Joint Feature Set (JFS) and Higher-Order wavelet statistics (in [4]) were used to build the classifiers. The method (in [8]) used statistical noise features to discriminate real and tampered images. The idea behind this approach was based on the observation that specific type of manipulations alter noise statistics in images which were (obtained) constituted from image denoising, wavelet analysis and neighborhood prediction.

The given image was denoised using an averaging filter, linear filter with Gaussian filtering, wiener adaptive filtering and median filtering. The mean and standard deviation at a particular pixel location for all denoising operations was calculated. Using wavelet analysis a given image was decomposed to four sub-bands as low-low, low-high, high-low and high-high, since it was believed that wavelet coefficients at higher frequency sub-bands do not follow Gaussian distribution and if an image is manipulated these coefficients are affected. So, the given image was first normalized with unit energy, after which the mean and standard deviation was computed for low-high, high-low and high-high sub-band coefficients. A Gaussian distribution function was built to test the goodness of the plot with the parameters from sub-band coefficients. The horizontal and vertical gradient values were computed with an estimated threshold value set on the pixel intensity value and the smooth region was further partitioned into dark and bright regions. The mean and standard deviation were computed for neighborhood pixel estimation. The

features collected were mean and variance of denoised image, wavelet sub-band coefficients and neighborhood pixels of both natural and tampered images. These features were in turn subjected to SVM as (in [6]) for classification. When experiments were performed on 500 natural and tampered images, the classifier detection results were successful with a detection accuracy of 90%.

With the increase in digital images, detection of manipulations in JPEG images has become a problem. The artifacts introduced into the digital images when manipulated can be used as a source of authentication to prove the integrity of digital images. For instance, the methods proposed (in [5]) using double quantization effect and ([7]) using Blocking Artifacts Measure (BAM) were based on DQ effect in histograms of DCT coefficients. Later for the image which was suspected to be manipulated, the quantization table was estimated and the BAM for entire image to identify a region or a part of the image was calculated. The DQ effects and the blocking Artifacts Measures cannot be used because in few scenarios these methods failed to verify the integrity of images. Experiments showed that once a JPEG image was manipulated by double compression and cropping, inconsistencies in blocking artifacts were introduced and the symmetry of the whole image was disturbed. JPEG images vary in their content and symmetric strength in terms of inconsistencies in blocking artifacts from an original to a tampered image and it is very difficult to differentiate whether an image is original or manipulated. Thus, the experiments conducted (in [9]) showed that these inconsistencies do not always provide reliable results. Therefore, to differentiate original and cropped or recompressed JPEG images, blocking periodicity in JPEG images in spatial domain was used, in which the image pixels were classified as across-block and within-block pixels and for each pixel the local pixel difference was computed as its blockiness. It was assumed that the local pixel difference for across-block pixels was larger and equal to zero for within-block pixels. Then, the dependency of local pixel difference was modeled and derived for different dependency models such as across-block and within-block pixels and the dependency of neighborhood pixels within-block were assumed to be correlated with a Gaussian model. Later the posterior probability of each pixel belonging to the within-block pixel was calculated.

Although the linear dependency model proved to be reliable than the local pixel difference, some times within-block pixels along texture or edge boundary region do not follow a linear dependency. To overcome this dependency model, dynamic block selection model was adopted and only those blocks were selected that provided a reliable blockiness measurement. The blocks were selected such that they weren't overly-smooth or overly-complex because the variations introduced in overly-complex blocks become greater than the blocking artifacts and are difficult to analyze. Probability map was obtained by computing Fourier domain for periodic analysis. Since JPEG image was divided into 8x8 blocks to compute DCT for each block, it was expected that for every eight pixels along the horizontal and vertical directions, periodic blocks will appear repeatedly i.e. 8x8 peaks in the power spectrum. For accuracy, the probability map was up-sampled by a factor of two using a high-pass filter and after this pre-processing operation, 16x16 peaks were obtained from a power spectrum. Hence the variations in these 16x16 peaks were observed when an image was tampered and analysis was made based on lower frequency area. The features from single compressed, cropped and re-compressed images were obtained by calculating mean, variance and entropy for each pixel within the blocks selected and trained to SVM for classification. When experiments were conducted on original and tampered images (which were tampered by cropping and re-compressing) it produced accurate results than the ones obtained (in [5] and [7]) which used DQ and Blocking Artifact Measures. This method also had a limitation, i.e. if the compression quality of a single compressed image is larger than the cropped and recompressed image the detection results will not be accurate.

The methods proposed (in [2]) by Popescu showed that image resampling introduces statistical correlations and also defined a way to detect the correlations to authenticate the image. A. Swaminathan (in [8]) showed a method to detect tampering in digital images using three sets of statistical noise features. The proposed methods do not apply to images which were manipulated with small manipulation strength. Hence the approach discussed (in [10]) proposed a simple way of investigating the images by dividing the image into partitions and identifying the noise levels as a source of image tampering

detection. Works were based on the periodic properties observed in the covariance structure of interpolated images and their n^{th} derivatives using geometrical manipulations such as resampling, resizing and rotation.

The working model (in [10]) was based on a simple, linear and stochastic model that reflects the dependency of noise levels independent from the part of the image. The mean of the random part of the image was assumed to zero; the covariance for the assumed model was computed to calculate the finite differences between adjacent samples. The periodic properties of interpolation were estimated in two steps, first the physical arrangement of the pixels was done in spatial transformation and then the interpolation step was performed. The pixel intensity value for the image that was transformed in the first step was assigned using a low-pass interpolation filter and the corresponding weights were multiplied for signals at various levels. The Interpolation method was detected using Region of Interest (ROI) selection, Signal derivative computation, Radon transformation and Signal search for periodicity. In ROI method, the regions were investigated to make sure if they were resampled and were selected by a block of RxR pixels. In Signal derivative, the presence of periodic properties in interpolated image was analyzed and n^{th} derivative of a pixel belonging to the block selected in ROI method was calculated and derivative operator was applied to the rows of the block in ROI. In order to highlight the traces of scaling and rotation Radon transformation was used, in which the magnitudes of the n^{th} derivative of a pixel multiplied by the weights of low-pass interpolation filter with an angle 'θ' were estimated. The main goal of this method was to identify the strong periodic peaks in the covariance matrix for the blocks selected in ROI selection method. To detect the periodicity, first order derivative filter was applied and the magnitude of the Fast Fourier transform was computed for the obtained sequences.

Now the image to be investigated was divided into four sub-bands i.e. LL, LH, HL and HH. The HH sub-band was selected since it gave the diagonal details of the image which were mapped to the non-overlapping blocks selected in ROI method. Then, the noise levels for each block were estimated using a wavelet-based mean absolute estimator. Once the estimated noise standard deviation was computed, the blocks were further sub-

divided into sub blocks and the standard deviation between the neighboring blocks was calculated with a simple merging algorithm. The limitation of this method was that the interpolation/resampling detector was highly sensitivity to noise. Also, it became difficult to detect the periodic correlations that were introduced into the signal by interpolation process.

CHAPTER 3
EXPERIMENTS AND DISCUSSIONS

In order to differentiate original and manipulated images, a database containing six images (each 256X256) is created. A set of five manipulations, are applied on the images chosen randomly from this database. The different manipulation processes are; modifying the 4th MSB of every pixel, adding salt & pepper noise, white Gaussian noise, contrast level adjustment and embedding an image of size 128x128. As manipulations do not leave any trace of identification, methods are required to test the originality of images.

The testing method includes computing DCT of original and manipulated images. The pdf plots for the DCT coefficients of both original and manipulated images are plotted and the pdf plot of original image is tested with the assumed Laplace distribution function because DCT coefficients for JPEG images exhibit Laplace distribution. The Laplace distribution function is:

$$f(I_i/\mu, v) = (1/2*v) \exp[-| I_i - \mu | / v] \qquad (4)$$

where μ is the mean of the pixels and v is the variance. Using maximum likelihood estimation (MLE) the parameters of the distribution function in equation (4) are estimated. As the pixel values in I_i are random and are identically distributed samples of the form $i_1, i_2 ... i_N$., the mean and the variance of the pixel values become estimators of the distribution function. Therefore, the parameters of the function are estimated by differentiating the equation (4) and equating the maximum value of the derivative function to zero as follows:

1) Parameter estimation for variance v':

Differentiating w.r.t variance parameter v'

$$(1/ 2*v')[(| I_i - \mu| / v'^2) * \exp(-| I_i - \mu| / v)] + \exp (-| I_i - \mu| / v) (-1/2* v'^2) = 0 \quad (5)$$

$$(1/ 2*v'* v'^2)[| I_i - \mu| * \exp (-| I_i - \mu| / v)] = \exp (| I_i - \mu| / v) (1/2* v'^2) = 0 \quad (6)$$

$$| I_i - \mu| / v' = 1 \quad (7)$$

$$v' = (1/N)* \Sigma | I_i - \mu| \quad (8)$$

2) Parameter estimation for mean μ':

Differentiating w.r.t variance parameter μ',

$$(1/ 2*v') * \exp (-| I_i - \mu'| / v)] = f (I_i/\mu,v) \quad (9)$$

$$(1/ 2*v') *(| I_i - \mu'| / v)] * \exp (-| I_i - \mu'| / v)] = 0 \quad (10)$$

$$\exp (-| I_i - \mu'| / v)] = 0 \quad (11)$$

$$(-| I_i - \mu'| / v)] = 1 \quad (12)$$

$$I_i = \mu'; \mu' = (1/N)* \Sigma I_i \text{ , for } i = 1 \ldots N \quad (13)$$

Thus, the parameters μ' and v' estimated using MLE, are average of the image pixel values and average of the expected mean value respectively. The corresponding μ' and v' are calculated for images with different types of manipulations and the corresponding standard deviation is calculated. The measure of the standard deviation is used to estimate by what value the pixels of the manipulated image deviates when compared to the original image.

The proposed algorithm to test whether a given image is tampered or original is as follows:

1. Read a JPEG image from a file
2. Convert the image into gray scale.

3. Compute the DCT for the obtained image.
4. Compute the pdf for the obtained DCT coefficients.
5. Assume the distribution function as Laplace distribution.
6. Using Laplace distribution function compute the probability value of each pixel.
7. Plot the probability distribution function for the result obtained in step 6
8. Compare the DCT plots of the original and manipulated images with the pdf plot in step 7.
9. Compute the variance and standard deviation for the pdf plots of the DCT coefficients.
10. Compute the kurtosis for the original and manipulated images

3.1 Results for Image Manipulations:

It is observed from the figure 4.1, that the pdf plot for the DCT coefficients of the original image follows the Laplace distribution. The figures 4.7 through 4.12 shows the images which are manipulated by (1) adding salt & pepper noise with a noise density d = 0.05, (2) adding Gaussian noise with mean = 0 and variance v = 0.01, (3) modifying the 4^{th} MSB of every pixel (4) embedding an image of size 128x128 and (5) increasing the contrast levels of the original image. Also, from figures 4.2 through 4.6 it is observed that the pdf plots for the DCT coefficients of the manipulated images become flatter. The computed standard deviation value is greater than 8.0 for the images manipulated by (1) modifying the 4^{th} MSB of every pixel (2) adding salt & pepper noise and (3) white Gaussian noise. The image modified by adjusting the contrast levels, the calculated standard deviation is greater than 6.0 and the image modified by embedding an image is greater than 7.5. Thus, standard deviation can be used as a parameter to test the originality of images

3.2 Observations:

The experimental results show that tampered images do not follow Laplace distribution function, i.e. as the strength of the manipulation increases the pdf becomes flatter. The

variance for all the manipulated images is found to be greater than the original images. Also, it is observed that kurtosis for image modified by the addition of salt & pepper noise is greater than the original image since it sharpens the noise, whereas for all types of manipulations kurtosis is less than the original image.

CHAPTER 4
SIMULATION RESULTS

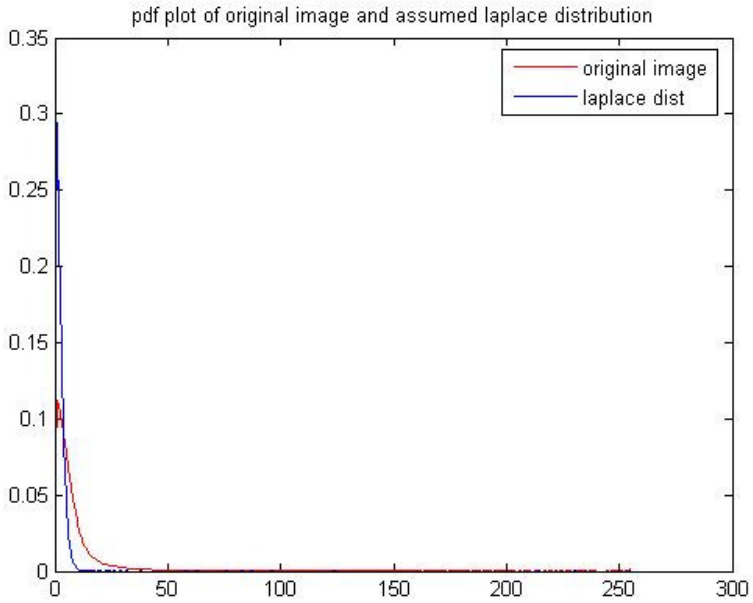


Figure 4.1: The pdf plot for the DCT coefficients of the original image shows that it follows the assumed Laplace distribution function.

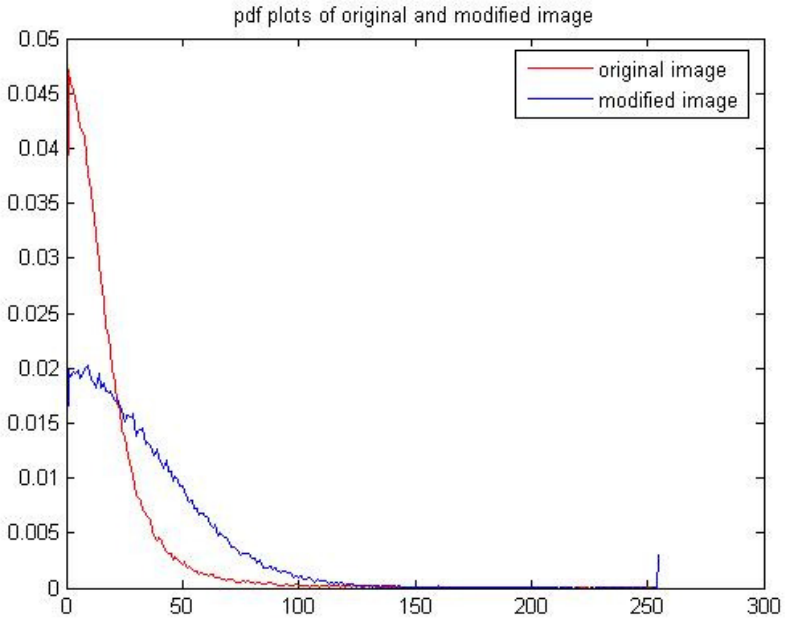


Figure 4.2: The pdf plots for the DCT coefficients of original and the image manipulated by the addition of salt and pepper noise with a noise density, d = 0.05

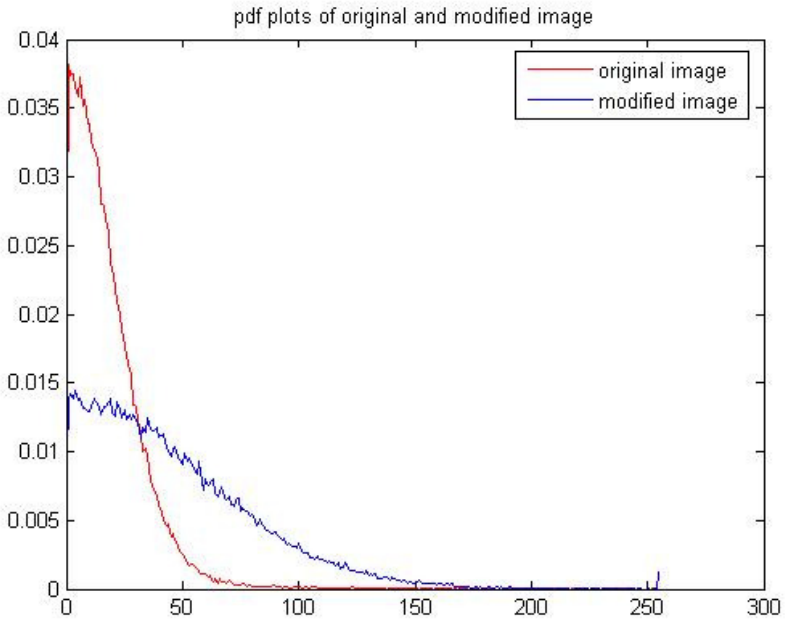


Figure 4.3: The pdf plots for the DCT coefficients of original and image modified by the addition of Gaussian noise with mean = 0 and variance = 0.01

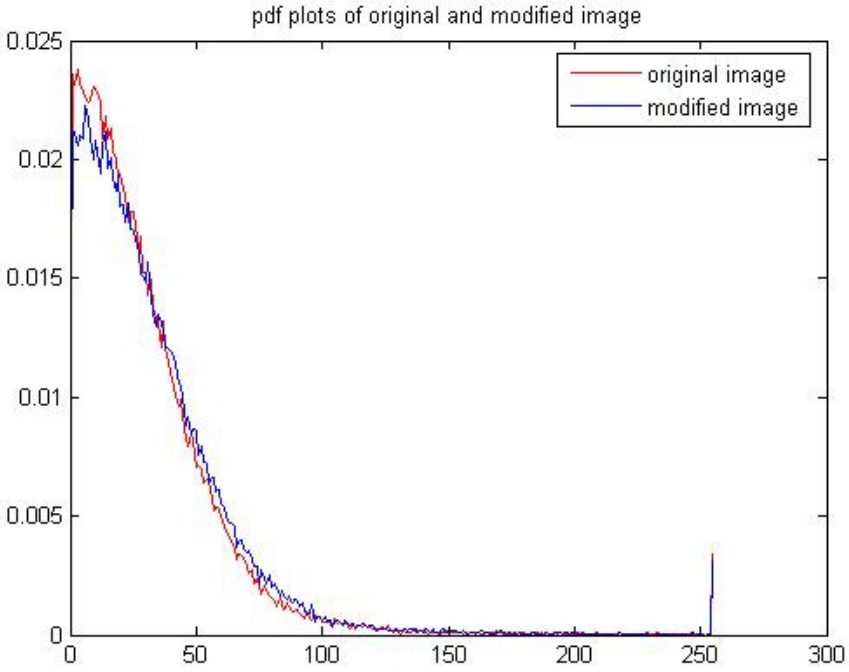


Figure 4.4: The pdf plots for the DCT coefficients of original and image manipulated by modifying the 4th MSB of every pixel.

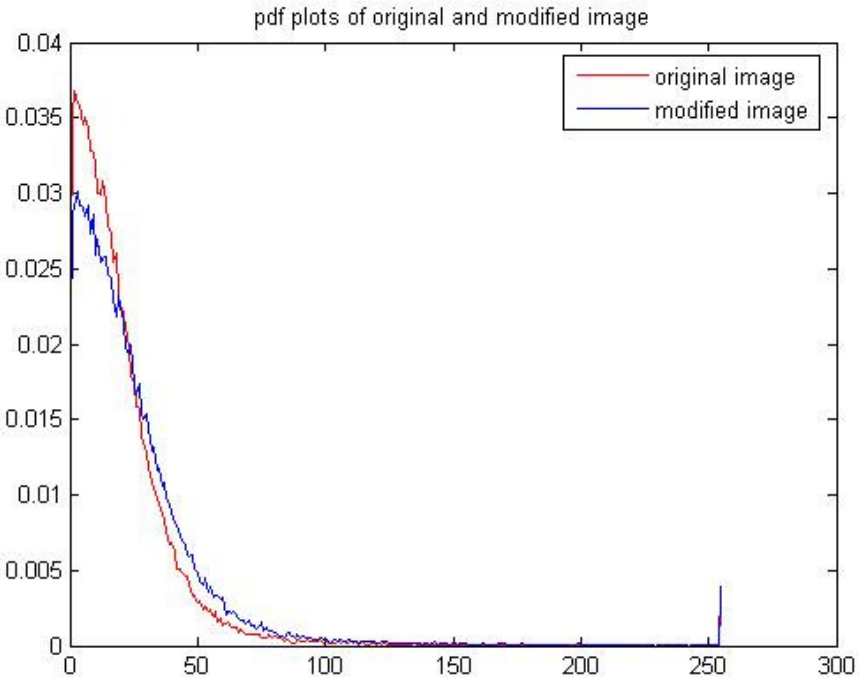


Figure 4.5: The pdf plots for the DCT coefficients of original and image tampered by embedding with another image of size 128x128.

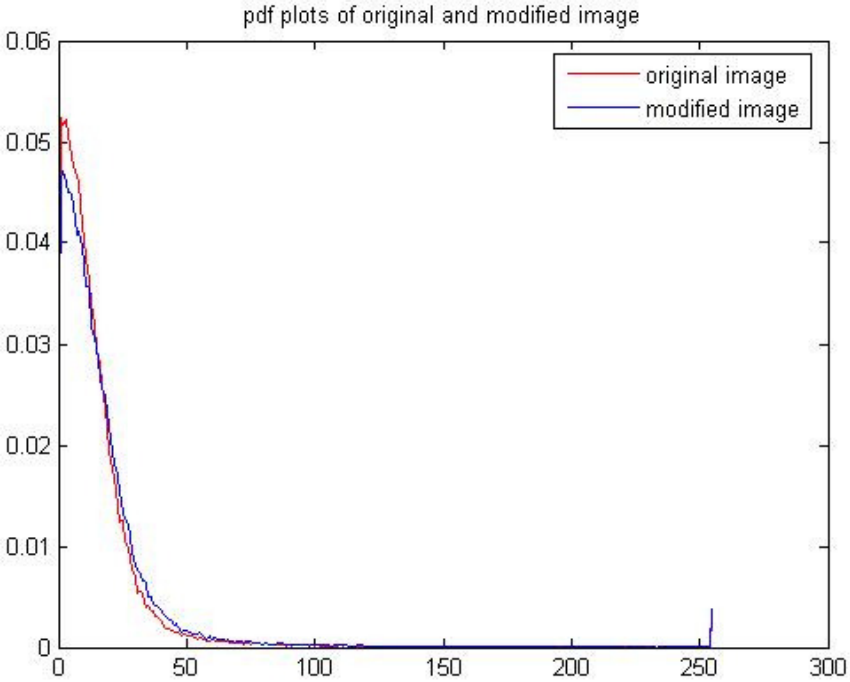


Figure 4.6: The pdf plots for the DCT coefficients of original and image manipulated by increasing the contrast levels of the pixels.

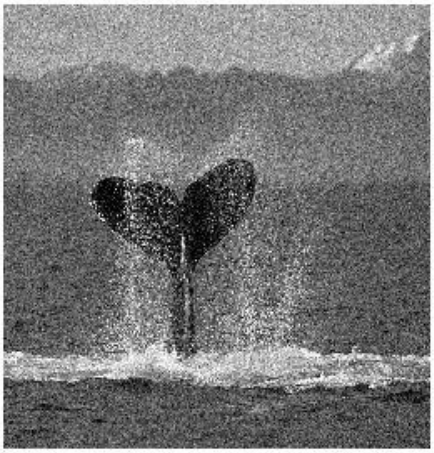
Table 4.1 shows the variance, standard deviation and kurtosis for different types of original and manipulated images for various manipulation tools.

Type of Manipulation	Original image			Manipulated image		
	Variance	Stand dev	Kurtosis	Variance	Stand dev	Kurtosis
Adding salt & pepper noise	42.7	6.53	3.5	70.5	8.4	4.0
Adding Gaussian noise	49.73	7.05	3.54	82.64	9.1	2.34
Modifying the 4th MSB of every pixel in the original image	63.7	7.98	2.2	67.0	8.2	2.04
Embedding of 128x128 image into original image	50.3	7.1	2.8	55.3	7.5	2.4
Adjusting the contrast levels of the original image	35.7	6.0	5.23	39.7	6.3	2.8

Figure 4.7: original image used for comparison with the assumed Laplace distribution function

(a) (b)

Figure 4.8(a): original image and Figure 4.8(b): image manipulated by the addition of salt & pepper noise with a noise density, d = 0.05



(a) (b)

Figure 4.9(a): original image and Figure 4.9(b): image modified by the addition of Gaussian noise with mean = 0 and variance = 0.01

(a) (b)

Figure 4.10(a): original image and Figure 4.10(b): manipulated image, in which the 4^{th} MSB of every pixel is modified.

(a)

(b)

(c)

Figure 4.11(a): original image before embedding and Figure 4.11(b): manipulated image embedded with an image of size 128x128 and Figure 4.11(c): image used for embedding.

(a)

(b)

Figure 4.12(a): original image and Figure 4.12(b) image manipulated by increasing the contrast levels of the pixels.

CHAPTER 5
CONCLUSIONS AND FUTURE WORK

5.1 Conclusion

In this paper we have introduced a method to detect whether a given image is original or manipulated. To differentiate the images, a database containing images of size 256x256 are created and a set of manipulations are performed on each image selected from this database. The computed pdf plot for the DCT coefficients of the original image is compared with the assumed Laplace distribution function as it is assumed that DCT coefficients exhibit a Laplace distribution. For all the original and manipulated images we have computed the DCT and the corresponding pdf plots are plotted. The pdf plots for manipulated images are found to be flatter as the strength of the manipulation increases on an image. Further, Using MLE method the parameters of the distribution function are calculated, from which the standard deviation is computed to test how much the pixel values of the manipulated image deviates when compared to the original image. Therefore, from the simulation results we can say that the pdf plots for the DCT coefficients of the manipulated images become flatter as the strength of the manipulation increases and also the standard deviation value is greater than that of original images. Finally the kurtosis for image tampered by addition of salt & pepper noise is greater than the images manipulated by other types of manipulations.

The limitation of this method lies in detecting the images with little manipulation strength and the detection methods apply to JPEG images only. Also, the images used for testing are gray scale.

5.2 Future Work:

Research can be carried out on the above method, to test images with little manipulation strength, different image formats (not just JPEG) and color images.

In order to test the images manipulated with little manipulation strength, highly sophisticated techniques are required. So far, all the proposed methods to detect the synthetic and real images were successful but had limitations. Lastly, if all the proposed methods are put together then the detection of natural and synthetic images can be achieved accurately with little computational power and in a short span of time.

REFERENCES

REFERENCES

1. On the Integration between Digital Watermarking and cryptography.

2. Avcibas, S. Bayram, N. Memon, B. Sankur and M. Ramkumar, A Classifier Design for Detecting Image Manipulations, Proc. of IEEE ICIP 2004.

3. Popescu and H. Farid, Exposing Digital forgeries by detecting duplicated image regions. Computer Science, Dartmouth College, Tech. Rep. TR2004-515, 2004.

4. S. Bayram, I. Avcibas, B. Sankur and N. Memon, Image Manipulation Detection, Journal of Electronic Imaging, vol. 15, no. 4 2006.

5. J. He, Z. Lin, L. Wang and X. Tang, Detecting Doctored JPEG Images via DCT Coefficient Analysis, Proc. of ECCV 2006.

6. H. T. Sencar and N. Memon, overview of state-of-the-art Digital image forensics, Proc. of IEEE ICIP 2007.

7. Shuiming Ye, Qibin Sun, Ee-Chien Chang, Detecting Digital Image Forgeries by Measuring Inconsistencies of Blocking Artifact. Proc. ICME, pp. 12-15, July 2007.

8. H. Gou, A. Swaminathan, and M. Wu, "Noise features for image tampering detection and steganalysis," IEEE Trans. on Image Processing, vol. 6, no. 10, pp.97-100, 2007**.**

9. Yi-Lei Chen, Chiou-Ting Hsu, "Image tampering detection by blocking periodicity analysis in JPEG compressed images," IEEE proc, pp. 803-808, Oct 2008.

10. Mahdian, B.; Saic, S, "Detection of Resampling Supplemented with Noise Inconsistencies Analysis for Image Forensics**,**" ICCSA apos; 08. International Conference, pp. 546 – 556, July 2008.

11. R.C. Gonzalez and R.E. Woods (2002). Digital Image Processing, 2nd edition, Prentice Hall.

12. Matlab Simulator version 7.0, June, 2004.

13. http://en.wikipedia.org/wiki/Kurtosis, April 2009.