STUDY OF IMAGE PROCESSING TECHNIQUES AND DETECTION OF COPY-PASTE REGIONS IN A DIGITAL IMAGE FORMED DUE TO IMAGE FORGERY

A Thesis submitted to the
SARDAR PATEL UNIVERSITY
for the degree of
Doctor of Philosophy
in
Computer Science

Submitted by
Brijesh Ramniklal Jajal

Research Guide:
Dr. Vipul Desai,
Charutar Vidya Mandal,
Vallabh Vidyanagar.

Research Center:
Sophisticated Instrumentation Center
for Applied Research & Technology
(SICART), Vallabh Vidyanagar.

JANUARY 2012

CERTIFICATE

Certified that the work incorporated in the thesis entitled “ **Study of Image Processing Techniques and Detection of Copy – Paste regions in a Digital Image formed due to Image forgery** “ submitted by **Mr. Brijesh Jajal** comprises the results of independents and original investigation carried out. The materials obtained (and used) from other sources have been duly acknowledged in the thesis.

BJajal

Date: 17.1.2012 **Signature of the Research Student**

Place: V. V. Nagar

Certified that the work mentioned above is carried out under my guidance.

Date : 17.1.2012 **Signature of the Research Guide**

Place : V.V.Nagar

Declaration

I declare that this thesis entitled "Study of Image Processing techniques and detection of copy-paste regions in a digital image formed due to image forgery" is the result of my own research except as cited in the references. This thesis has not been accepted for any degree and is not currently submitted in candidature of any other degree of any university.

Signature : BJajal

Name : Brijesh Ramniklal Jajal

Regn. No. : 4620

Date : 16.01.2012

ACKNOWLEDGMENTS

I offer my sincere gratitude to my research guide Dr. Vipul Desai for his invaluable advice, guidance and constructive criticism throughout the study. I attribute the research work to his encouragement and effort and without him, this thesis would not have been completed.

My acknowledgment is due to Dr. Ramji Makwana, who guided and motivated me for learning MATLAB tool and facilitated the implementation of algorithm ideas.

I am extremely grateful to Dr. C. L. Patel (Chairman, Charutar Vidya Mandal), who have always inspired us for achieving the excellence through knowledge.

My sincere thanks are due to Dr. D. B. Choksi (Director, MCA Department) for his administrative guidance, and most importantly for making my concepts clear during my study. I am really thankful to the members of MCA department, Dr. P. V. Virparia, Dr. B. B. Patel and Dr. P. S. Sajja, whose contribution created a very positive environment for research.

It is indeed during my service at Ashok & Rita Patel Institute of Integrated Study & Research in Biotechnology and Allied Sciences (ARIBAS), New Vidyanagar, that I am able to develop the zeal for Research work and Publication. I am thankful to Prof. P. M. Mehta, Dr. H. G. Vyas and Dr. P. S. Patel for their support at workplace.

I would like to admire the efforts of Dr. R. K. Patel for inspiring for research. I convey my sincere thanks to my colleagues Dr. Hiren, Dr. Kuldeep, Dr. Rajiv, Dr. Swagat, Dr. Saroja and Dr. Swati.

My parents deserve special mention for their inseparable support and prayers. I am beholden to my dear wife Sonal and angel daughter Moksha for their sacrifice and understanding.

I appreciate the concern of my brothers Ketan and Tarun during my research.

I dedicate this thesis to Shree Swaminarayan Bhagwan.

Brijesh Jajal

TABLE OF CONTENTS

ABBREVIATIONS

Form	*Meaning*
2D	Two dimensional
3D	Three dimensional
ASCII	American Standard Code for Information Interchange
CAD	Computer Aided Design
DCT	Discrete Cosine Transform
DWT	Discrete Wavelet Transform
IP	Internet Protocol
Mac	Macintosh
NTFS	Network Technology File System
OS	Operating System
PC	Personal Computer
TCP	Transmission Control Protocol
URL	Universal Resource Locator
USB	Universal Serial Bus

Chapter 1

INTRODUCTION

- ⊗ Background
- ⊗ Image Acquisition
- ⊗ Image Processing
- ⊗ File formats
- ⊗ Software and tools
- ⊗ Types of Image Forgery

CHAPTER - 1: INTRODUCTION

1.1 Background

The Digital era has made a layman to play with the image files in a way one desires. The image editing tools are good enough to provide the silver screen effects into multimedia and communications. However, the tampering or manipulation of images for the false purpose has to be detected by a forensic expert. There is a wide scope as well as challenges for development of new image forgery detection techniques, both in terms of a framework, as well as software solutions with specialized tools.

In spite of the varieties of professional experts and software tools available worldwide, the image forgery detection techniques are often a failure, since every method is specific to a given type of problem only. Considering the aspect of India, the jurisdictions require lot of modifications in order to consider such cases, being an example of cyber crime or digital crime.

An extraordinary effort of Hany Farid, Dartmouth College, USA in the field of image forensic has aided the researchers across the globe to implement his ideas and work for enhancement of the same.

The dual fold approach for the detection of image forgery from a digital image, reported in the current research, can aid to the forensics to analyse the suspected digital documents.

With an amount of increase in the image forgery and their consequences, it becomes important to the investigators to make such systems handy to them, which may be in form of a software tool, scientific equipment or customized device. Out of these available solutions to determine the forgery, software tools may be considered to preferable, due to its flexibility for improvement.

1.2 Image Acquisition

"A picture is worth thousand words" is universally accepted truth identifying the importance of a picture or an image. An image can be defined as a two-dimensional function having a scalar value for each of its position. It is described in the form of co-ordinate point (x, y) of a plane. A digital image refers to the information stored by means of a digital camera, wherein the co-ordinate points or locations are referred in the form of an image element or picture element, popularly termed as pixel.

The scene can be acquired or captured permanently by using digital cameras, and stored in form of a file. All the digital cameras available today, be it a basic one, or the DSLR (Digital Single Lens Reflex) used by the professionals, make use of the CCD or CMOS technology to sense the image components.

CCD (Charge Coupled Device) and CMOS (Complementary Metal Oxide Semiconductor) image sensors are two different technologies for capturing images digitally. CCDs and CMOS imagers were both invented in the late 1960s and 1970s. CCD became dominant, primarily because they gave far superior images with the fabrication technology available.

Each has unique strengths and weaknesses giving advantages in different applications. Neither is categorically superior to the other, although vendors selling only one technology have usually claimed otherwise. In the last five years much has changed with both technologies, and many projections regarding the demise or ascendance of either have been proved false. The current situation and outlook for both technologies is vibrant, but a new framework exists for considering the relative strengths and opportunities of CCD and CMOS imagers.

Charge-coupled device has three major components:

- A silicon diode photo-sensor (a Pixel) that receives photons of various intensity. If the incident photons have sufficient energy to agitate an electron motion away from the silicon layer, it generates a charge.
- The charge moves to a down-stream charge storage region, generating an analogous signal.
- The quantity of the accumulated charge, the sum of 0 or 1 (depending on the incident light intensity), is then amplified and transmitted through a clock signal.

The CMOS works on the following principle:

- Each column of photo sensors has an amplifier associated with it, and the image can be transmitted above the noise.
- A row of pixels can be readout in parallel with the row selected by an addressing register (Y-addressing) or an individual pixel can be selected by addressing column multiplexer (X-addressing).
- A CMOS device is essentially a parallel readout device and therefore can achieve higher readout speeds. However, compensating for the variations in the current state of the art CMOS devices is difficult.

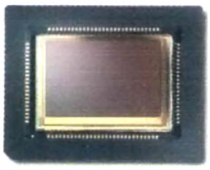


Figure - 1.1: (a) CCD Sensor (b) CMOS Sensor

The CFA (Color Filter Array) pattern arrangement depends on the manufacturer, although Bayer's filter mosaic is often preferred. As a result, the sensor output is a mosaic of e.g. red, green and blue pixels arranged on a single layer. To obtain the canonical 3-channels representation, the signal needs to be interpolated. Demosaicing algorithms are applied to this purpose; the missing pixel values in each layer are estimated based on the values of existing neighbors. Before the eventual storage, additional processing is performed, such as white balance, gamma correction, and image enhancement ***(Redi et al, 2011)***. Finally, the image is recorded in the memory device, as in Figure - 1.2. Here, the file formats of image can vary.

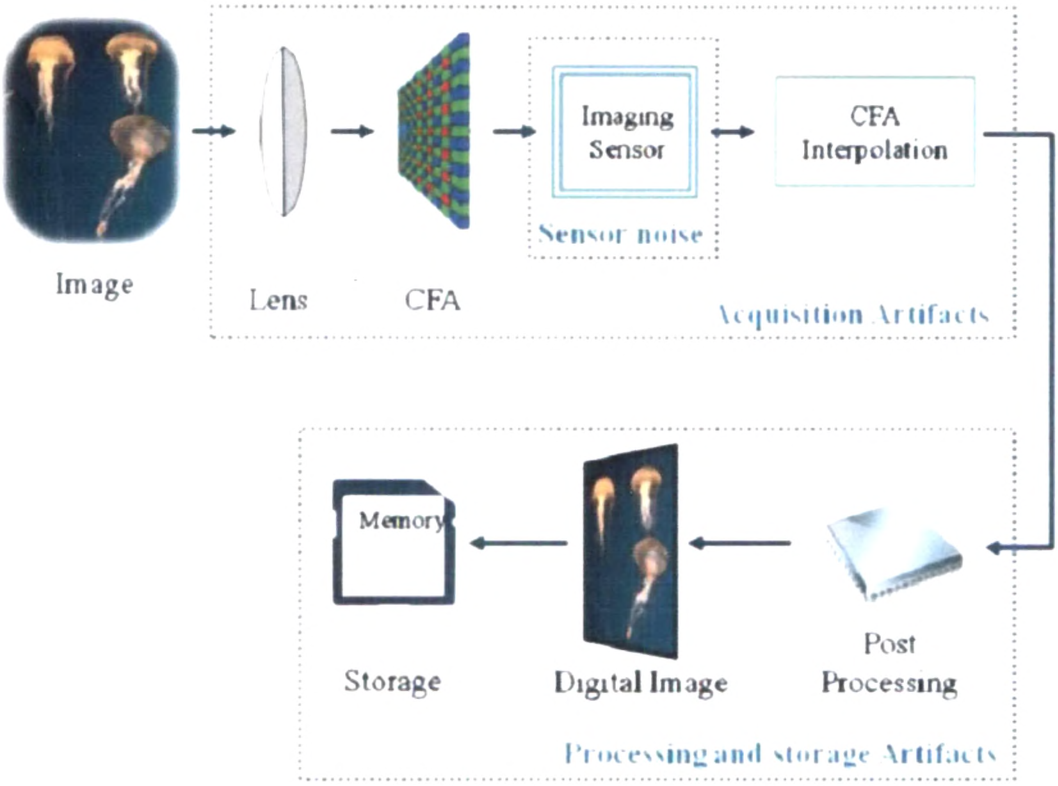
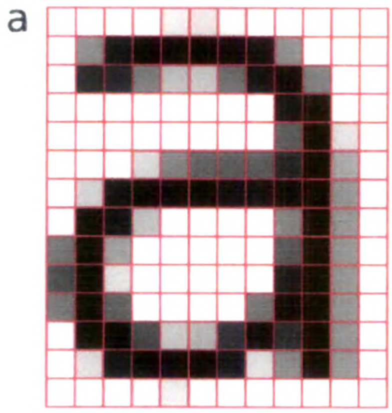


Figure - 1.2: Image Acquisition processing from camera

After the sensing of scene elements by the sensors of a camera, each element of an image, called pixel, is stored in form of its color equivalent value, in case of Color representation, while it is a gray scale value in the range of 0 (black) to 1 (white), for the gray image, as described in Figure - 1.3.



1.0	1.0	1.0	0.9	0.6	0.6	0.6	1.0	1.0	1.0	1.0	1.0
1.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.5	1.0	1.0	1.0
1.0	0.2	0.2	0.5	0.6	0.6	0.5	0.0	0.0	0.5	1.0	1.0
1.0	0.9	1.0	1.0	1.0	1.0	1.0	0.9	0.0	0.0	0.9	1.0
1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.5	0.0	0.5	1.0
1.0	1.0	1.0	0.5	0.5	0.5	0.5	0.5	0.4	0.0	0.5	1.0
1.0	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	1.0
0.9	0.0	0.0	0.6	1.0	1.0	1.0	1.0	0.5	0.0	0.5	1.0
0.5	0.0	0.6	1.0	1.0	1.0	1.0	1.0	0.5	0.0	0.5	1.0
0.5	0.0	0.7	1.0	1.0	1.0	1.0	1.0	0.0	0.0	0.5	1.0
0.6	0.0	0.6	1.0	1.0	1.0	1.0	0.5	0.0	0.0	0.5	1.0
0.9	0.1	0.0	0.6	0.7	0.7	0.5	0.0	0.5	0.0	0.5	1.0
1.0	0.7	0.1	0.0	0.0	0.0	0.1	0.9	0.8	0.0	0.5	1.0
1.0	1.0	1.0	0.8	0.8	0.9	1.0	1.0	1.0	1.0	1.0	1.0

Figure - 1.3: The internal representation of a letter 'a'

1.3 Image Processing

A field of Digital Image Processing (DIP) has shown revolutionary applications in all areas of human life. It constitutes the use of a digital image as input as well as output. A digital image acquisition process, being the first operation on digital images, refers to pre-processing part of image storage into a digital storage media. Consequently, DIP refers to the operations performed on an existing digital image. The image processing can be categorized into three major aspects. The low level processing is useful for the basic functions like removal of noise, segmentation and identifying objects from an image. The mid level processing deals with attributes of an image, whereas the advanced level constitutes inference and vision. The image analysis is assumed to be an application between DIP and Image vision, with the later one performing complex vision related functions. However, there is no clear cut boundary to all these digital image aspects.

The study of DIP is of utmost significance, with one of its basic reason of being applied in the types of images which are not considered to be visible by the human eye. The DIP is able to cover the whole range of electromagnetic (EM) spectrum.

Assuming an example of a person taking a digital image of an event by using a digital camera, one may adjust the colour or resolution settings or lens mode. These functions are considered as a part of image acquisition. Once a digital image is created and accessible, it is possible to identify the region of interest (ROI), which is called low level processing. The image size settings for a print, indicates mid level processing. Finally, determining the size of actual object from the image is considered as high level processing.

Classification of Digital Image Processing:

In today's world, the digital image processing is applicable in each and every area of technology. Hence, it can be classified based on different criteria.

According to Signal Frequency: The digital images can be captured for all the parts of EM spectrum. The EM waves are sinusoidal waves with different wavelengths and moving at the speed of light. These are wave like particles which contain an energy measured in photons. Figure - 1.4 indicates the types of EM waves based on its energy of photon.

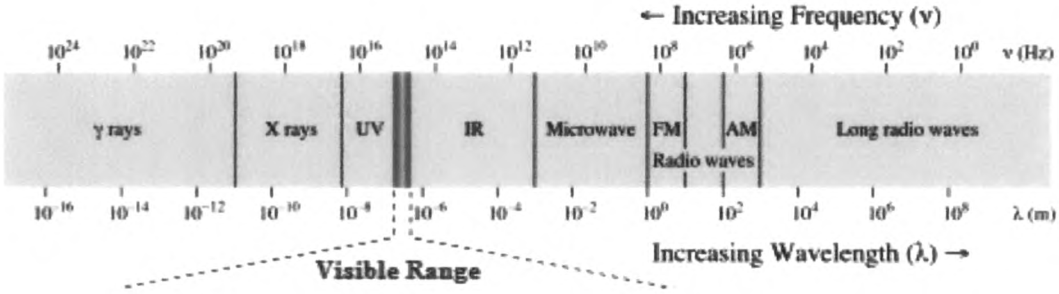


Figure - 1.4: Types of Electromagnetic waves

Gamma ray imaging

The gamma ray band is useful in two major areas viz. nuclear medicine and astronomical observations. The nuclear medicine contains the radioactive isotope injected into the patient, which emits gamma rays as it decays. The image can be acquired from the body by using gamma ray detectors. The PET (Positron

Emission Tomography) is a similar method of producing an image by detection of the flow of positrons emitted from the body.

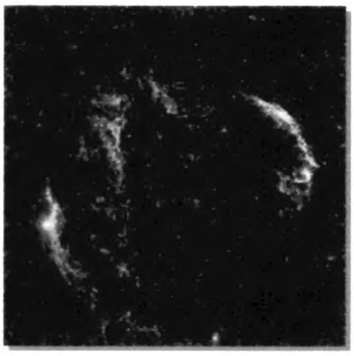
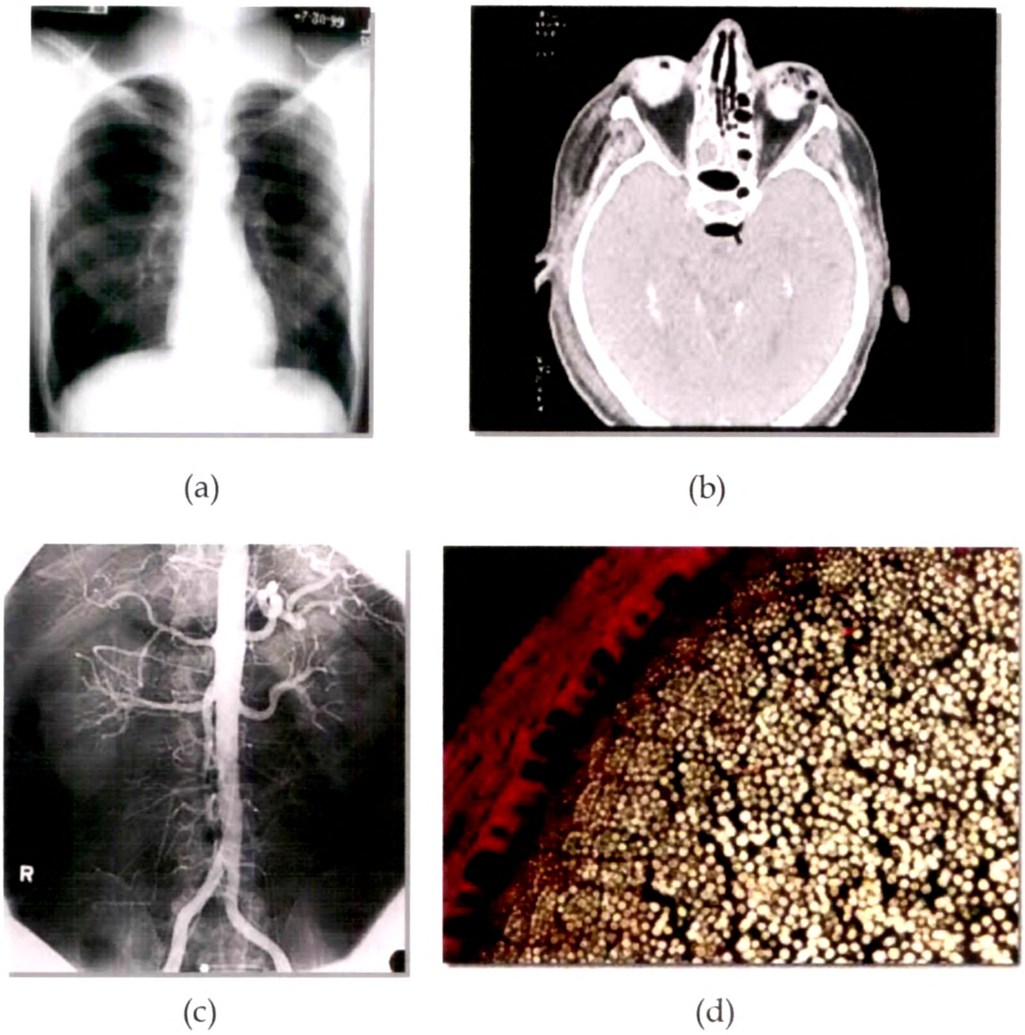


Figure - 1.5: Photograph of Cygnus Loop

Figure - 1.5 shows astronomical observation of a famous Cygnus Loop, which glows due to the explosion, occurred before about 15,000 years.

X-ray imaging

It is widely used concept for medical diagnostics, especially for bone images to detect the fractured areas. X-rays are also having applications in industry and astronomy. The X-rays are generated by using X-ray tube with anode and cathode, in which electrons get released through the cathode heating. The penetration power of X-rays can be varied by applying the required voltage across anode. The image is obtained in gray-scale format on the X-ray sensitive film with the patient between the X-ray source and film. Figure - 1.6 (a) represents the chest X-ray of patient. In order to acquire blood vessel image, the Angiography, also called contrast enhancement radiography, is used. The catheter is placed into the blood vessel into the required area and X-ray contrast medium is inserted. This results into visible irregularities or blockages, represented in Figure - 1.6 (b). Another example using in higher energy X-rays is Computerized Axial Tomography (CAT) scan. The machine is capable of taking two dimensional slices of the required area, which consequently adds the third dimension (3D) of depth depending on the duration between two scans. Figure - 1.6 (c) shows the CAT image of the head.



(a) (b)

(c) (d)

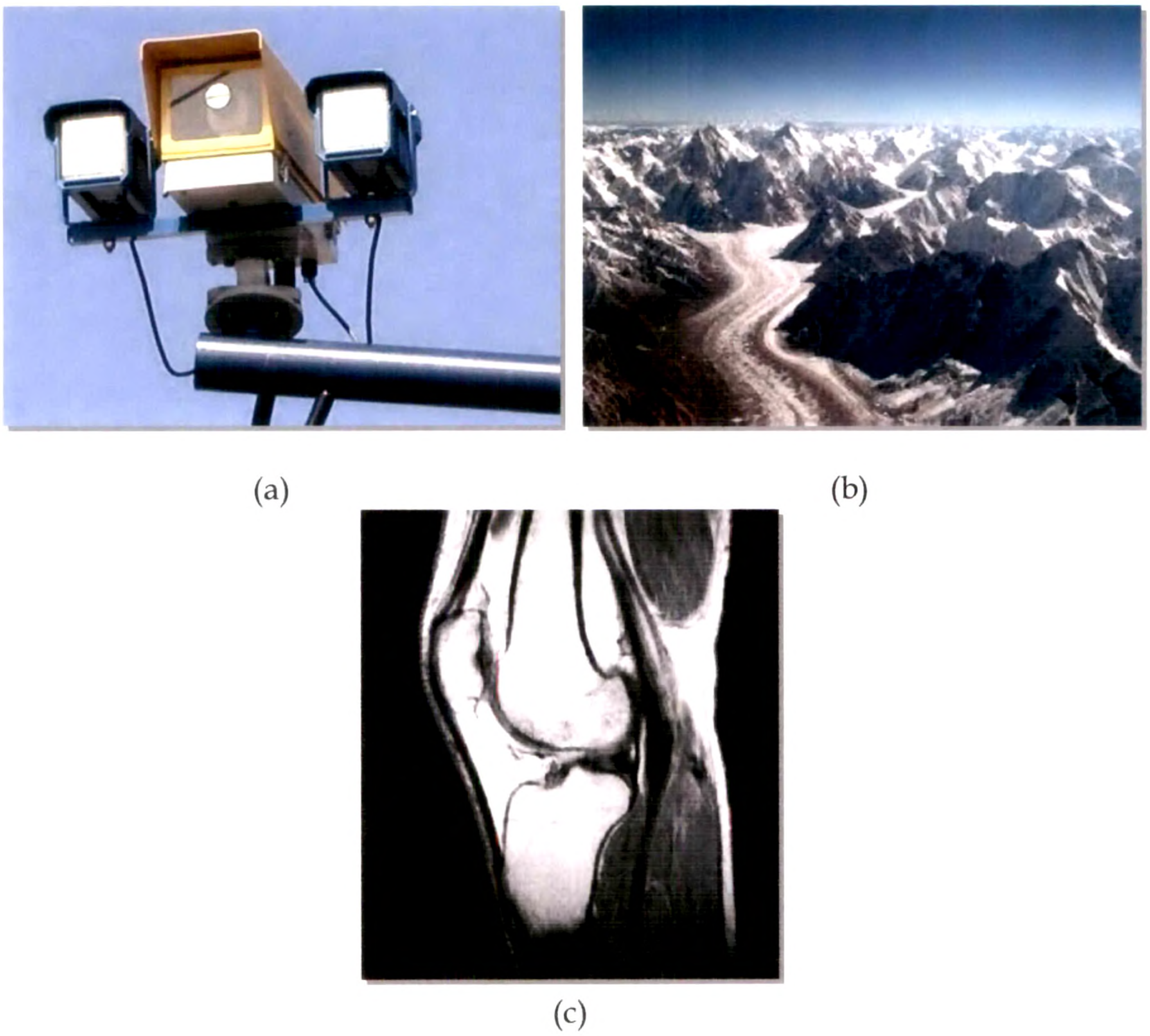
Figure - 1.6: (a) X-Ray of chest (b) CAT image of head
(c) Aortic Angiogram (d) Fluorescent microscopy of corn

Ultraviolet imaging

The ultraviolet rays are useful in microscopic images. The fluorescence microscopy can be used to visualize the specimen, which can fluoresce either naturally or by the use of chemicals. Figure - 1.6 (d) represents the smart corn image achieved by fluorescence microscopy.

Visible and IR Imaging

The images in visible spectrum actually ranges from the microscopic images zoomed from the factor of 2X to 1000X, to the very far images like satellite remote sensing. Figure - 1.7 (a) shows the example of traffic monitoring or surveillance.



(a)

(b)

(c)

Figure - 1.7: (a) Traffic Surveillance System (b) Radar image of Himalayas
(c) MRI Image of human knee

Microwave imaging

The most dominant part of microwave range is RADAR (Radio Detection and Ranging), which is able to penetrate all the objects including clouds. The radar image of Himalayan Mountains of India is depicted in Figure - 1.7 (b).

Radio imaging

Similar to the gamma ray imaging, the radio waves are also used in medical imaging and astronomy. The MRI (Magnetic Resonance Imaging) uses the short pulses of a powerful magnet applied on the patient's body, where the corresponding pulse is detected by the computer in form of a 2-D picture. Figure-1.7 (c) represents the MRI image of human knee.

According to Functionality

Image acquisition and sensing: There are three major sensor arrangements used to transfer the illumination energy of any object into a digital image format. The acquisition process refers to this conversion.

Single sensor acquisition

The photodiode is the most commonly used sensor, which converts the light energy into its equivalent output voltage, due to the presence of silicon material used in it. In order to achieve the 2D image, the sensor can be moved in linear direction with the aid of mechanical arrangements. Alternatively, a flat-bed or a LASER source can also be used for imaging.

Sensor strip acquisition

A sensor strip comprises of an in-line arrangement of sensors, which enables the detection of image along one dimension, covering about 4000 elements of an image. A sensor strip can be linear as in Figure - 1.8(a), or circular as in Figure - 1.8(b) depending on the type of application, being 2D or 3D respectively. The scanning techniques like CAT (computerized Axial Tomography), MRI (Magnetic Resonance Imaging) and PET (Positron Electron Tomography) are also using a similar imaging modality.

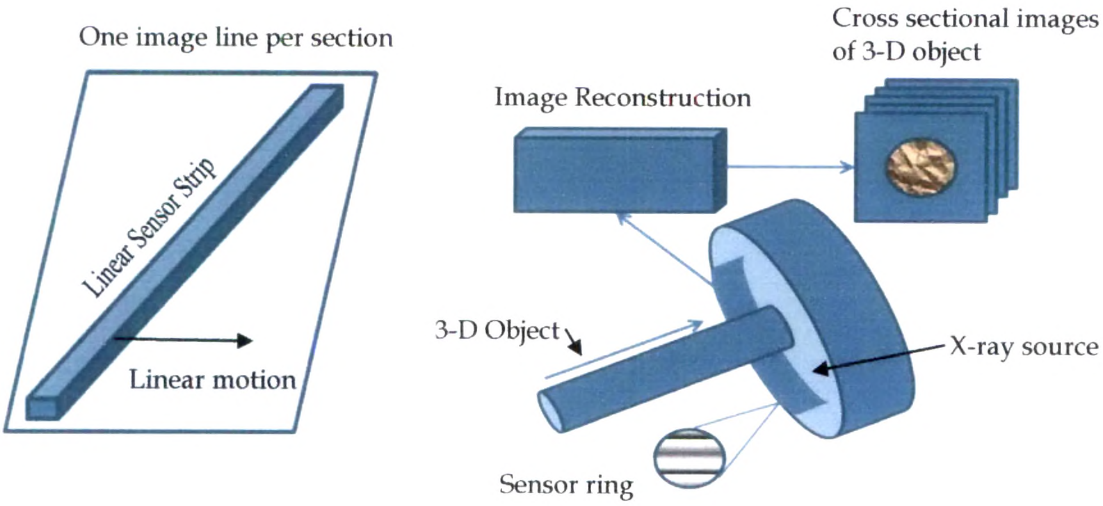


Figure - 1.8: (a) Linear Sensor strip (b) Circular Sensor strip

Sensor array acquisition

The set of sensors are arranged in form of a 2D array of typical size of 4000 x 4000 elements. The advantage of array sensors is that since it covers horizontal as well as vertical elements of an image at the same time, there is no displacement required. The digital cameras are using CCD array sensors used with broad range of sensing properties.

a. Image sampling and quantization: In order to convert an image into a digital form, the co-ordinate values are digitized, and the process is termed as sampling. The digitizing of amplitude is called quantization.

b. Image enhancement: A digital image once available may need to be modified to make it more convenient for a required application. The enhancement can be based on co-ordinates, known as spatial domain, or signal frequency, known as frequency domain.

The enhancement operations carried out on a digital image has virtually no limits, since it is post-processing part, and many functions like filtering, smoothing, hue and saturation settings, histogram processing, Fourier transforms etc. can be carried out.

1.4 File formats

The images captured by a camera are stored in form of a binary file, with different extensions, depending on the requirement of the application. The basic file types of an image are as below:

TIFF (Tagged Image File Format)

Versatility and compatibility make the TIFF image format the optimum choice for almost any project. It works on both the Mac and PC platforms, supports almost any picture bit depth, and allows various forms of compression.

This flexibility also makes the TIFF format a Pandora's box. There are so many versions and types of compression for the TIFF file format that no current system can decode all of them. Furthermore, there's no way to tell how a TIFF will behave until one attempts to manipulate it.

PICT (Picture File)

PICT is the default file format for any image displayed by a Macintosh. Unfortunately, many people misinterpret this fact to mean that PICT is the best file format to use. PICT images are useful in Macintosh software development, but should be avoided in desktop publishing.

Generally it is avoided to use PICT format in electronic publishing, since PICT images are prone to corruption. Additionally, PageMaker always knocks out PICT images as color separations.

PSD (Photoshop Document)

This is the native Photoshop file format created by Adobe. In this format, we can save multiple alpha channels and paths along with the primary image. However, the import this format is not possible into most desktop publishing applications.

PNG (Portable Network Graphics)

Because of its high compression rate (unsurpassed among ``lossless'' formats -- lossy JPEG is better for photos) PNG is the standard file format that online services use for storing 1- to 8-bit images, succeeding GIF, which is a proprietary format and thus increasingly discarded. Most of the news agencies publishing the images online opt for the PNG format, in order to verify the image forgery.

JPEG (Joint Photographer's Expert Group)

Because of its good image quality and compression, the JPEG file format is becoming increasingly popular in desktop publishing. Developed by the Joint Photographic Experts Group, the JPEG format is expected to become an international standard for encoding digitized photographs. JPEG's major difference from all other current file formats is that it uses Lossy compression.

Even though JPEG is relatively new, a second format already exists, i.e. JFIF being a new TIFF subformat that embeds a JPEG image into a TIFF file. The JFIF file format has yet to become popular due to its complexity and less capability.

TARGA (TGA)

Most common in the video industry, this file format is also used by high-end paint and ray-tracing programs. The TGA format has many variations and supports several types of compression. This file format is used to display AT&T Truevision images and is sometimes supported by DOS applications.

PCX (Personal Computer)

PCX is a straightforward raster file originally available only on the PC. However, PCX is migrating to the Macintosh as more mainstream programs become cross-platform. Aldus PageMaker, Adobe Photoshop, and the newest version of QuarkXPress for the Macintosh now support the PCX file format.

BMP (BitMap)

The BMP file format is available in almost all Windows-based graphics applications, although it is primarily used in Windows application development.

WMF (Windows Metafile)

A Windows Metafile has its good and bad points. Metafiles-composed of a list of calls to the Microsoft Windows graphics drawing library-are both small and flexible, but unless the program that created them is running, they are difficult to display properly.

CGM (Computer Graphics Metafile)

Computer Graphics Metafile is a very flexible vector format that can also save raster information. Unfortunately, it is so flexible that very few applications can use all the types available.

AutoCAD DXF (Drawing)

AutoCAD DXF is an AutoCAD file format that has become a standard for exchanging CAD drawings. However, because the vector information is ASCII encoded, the files can become very large and thus require a lot of memory to read.

Hewlett-Packard GL/2

GL/2 is a HP plotter language, which is often used as an exchange format for graphics.

EPS (Encapsulated PostScript)

It is a vector file relying on the PostScript page description language to draw its image. This format can also contain raster information, even though it is not a raster format. EPS files generally contain a raster graphic as a screen preview-Mac EPS files use a PICT and PC-EPS files use a TIFF graphic. EPS is the only format that supports transparent white in bitmap mode.

1.5 Software and Tools

Due to extensive availability of the image editing software, most of them as a freeware, there is a massive increase in the manipulation of images. A person need not be a professional for the image editing, since the editing tools are easy enough to operate. Let us consider the top five - most popular image editing software used in recent times:

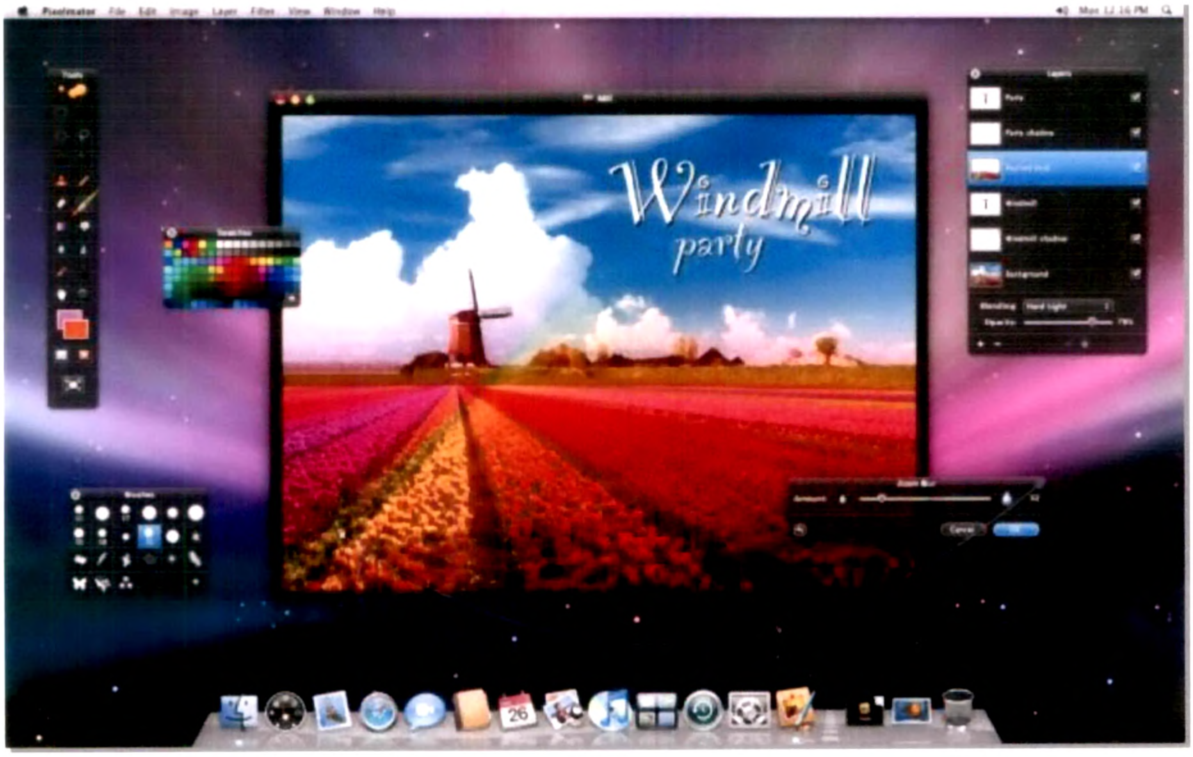


Figure - 1.9: Screenshot of Pixelmator

Pixelmator is a fast and powerful image editing software for the Mac operating system. With its intuitive and beautiful Graphical User Interface (GUI), support for layers to organize the document, a large assortment of painting tools, and simple-to-use photo correction tools are available. Pixelmator is used by Mac users who don't quite need the features of Photoshop.

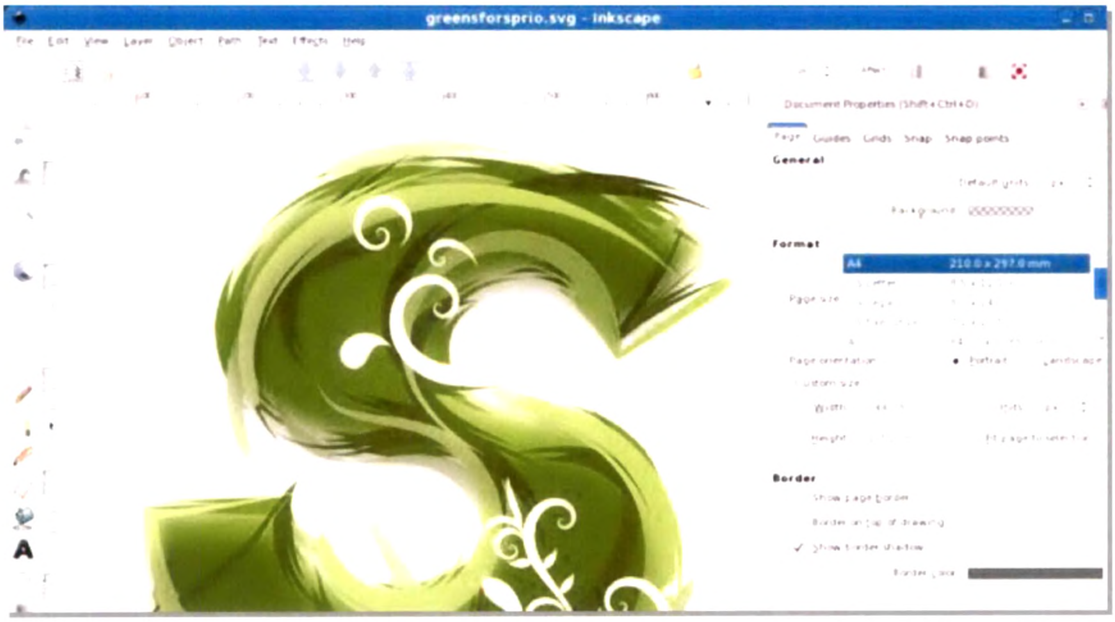

greensforsprio.svg - Inkscape
File Edit View Layer Object Path Text Effects Help
Document Properties (Shift+Ctrl+D)
Page Guides Grids Snap Snap points
General
Format
A4 210.0 x 297.0 mm
Border
A

Figure - 1.10: Screenshot of Inkscape

Inkscape is an open source vector graphics editor much like Adobe Illustrator, CorelDraw, and Xara X. Its default file format is web standards compliant Scalable Vector Graphics (SVG) under World Wide Web Consortium (W3C)'s specifications.

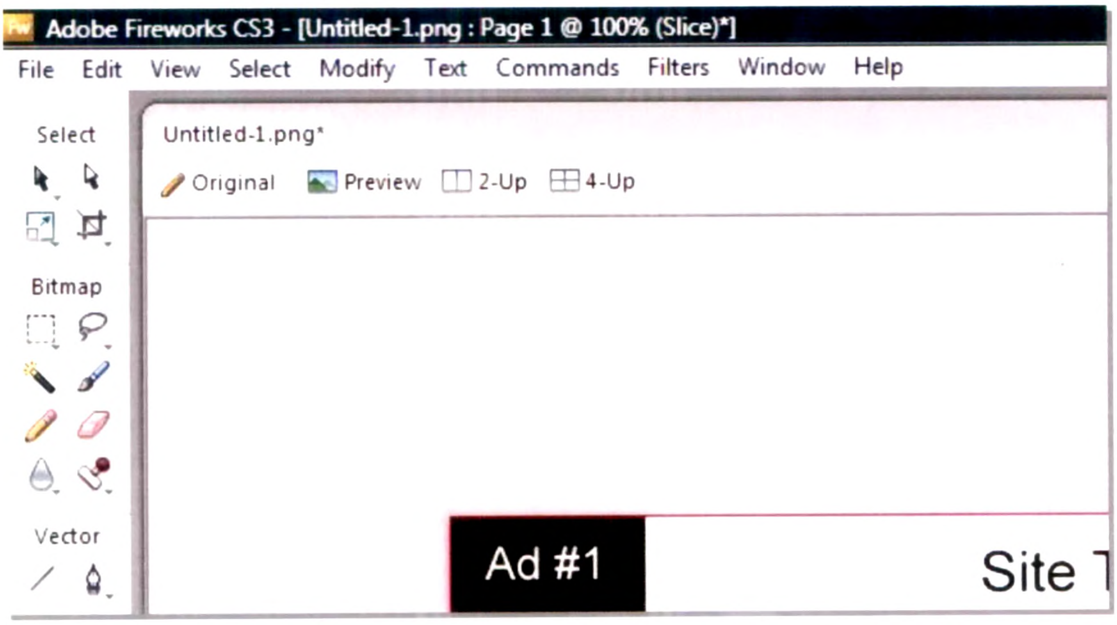

Adobe Fireworks CS3 - [Untitled-1.png : Page 1 @ 100% (Slice)*]
File Edit View Select Modify Text Commands Filters Window Help
Select
Bitmap
Vector
Untitled-1.png*
Original
Preview
2-Up
4-Up
Ad #1
Site 1

Figure - 1.11: Screenshot of Fireworks

Fireworks is Adobe's image editing software for the web designers. It excels in several areas over Photoshop, namely in high-fidelity prototyping of sites and a workspace environment that's optimized for web designers. It is also a raster and vector hybrid, being able to work with raster-based images and vector-based graphics better and more symbiotically than Photoshop.

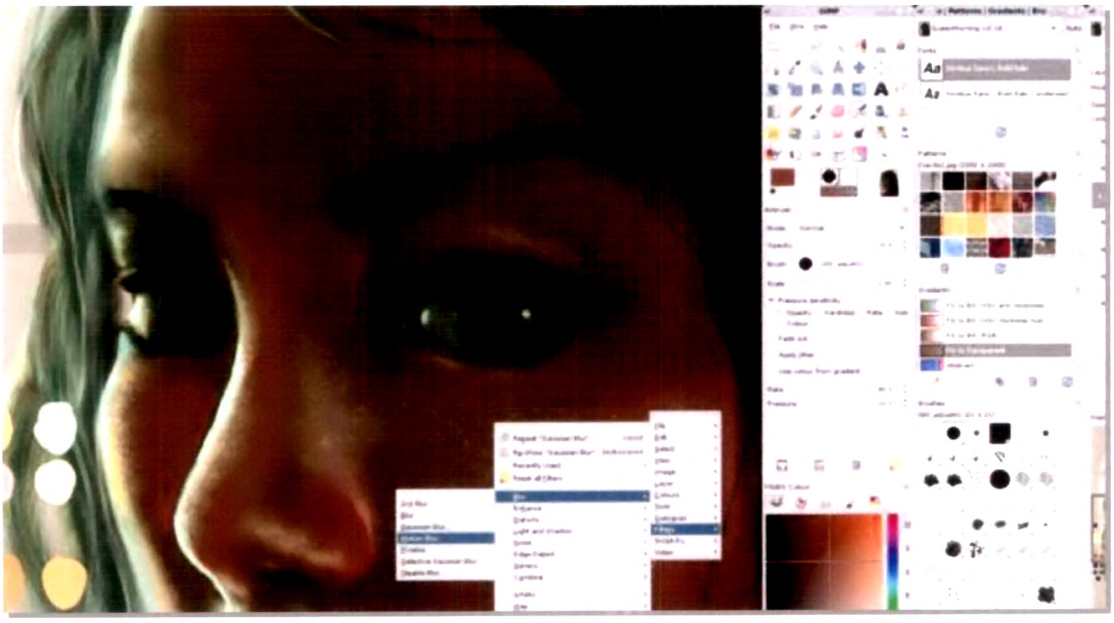


Figure - 1.12: Screenshot of GIMP

GIMP - which stands for the GNU Image Manipulation Program - is a feature-packed and powerful open source image editor that can be used in all major operating systems viz. Linux, Mac and Windows. It has a customizable interface so that one can easily set the view and behavior of GIMP.

It has a huge set of retouching tools that allows performing advanced image retouching and manipulation. The GIMP outputs the work in many common formats like JPG, GIF, PNG, TIFF, and even PSD (Photoshop's native file format).

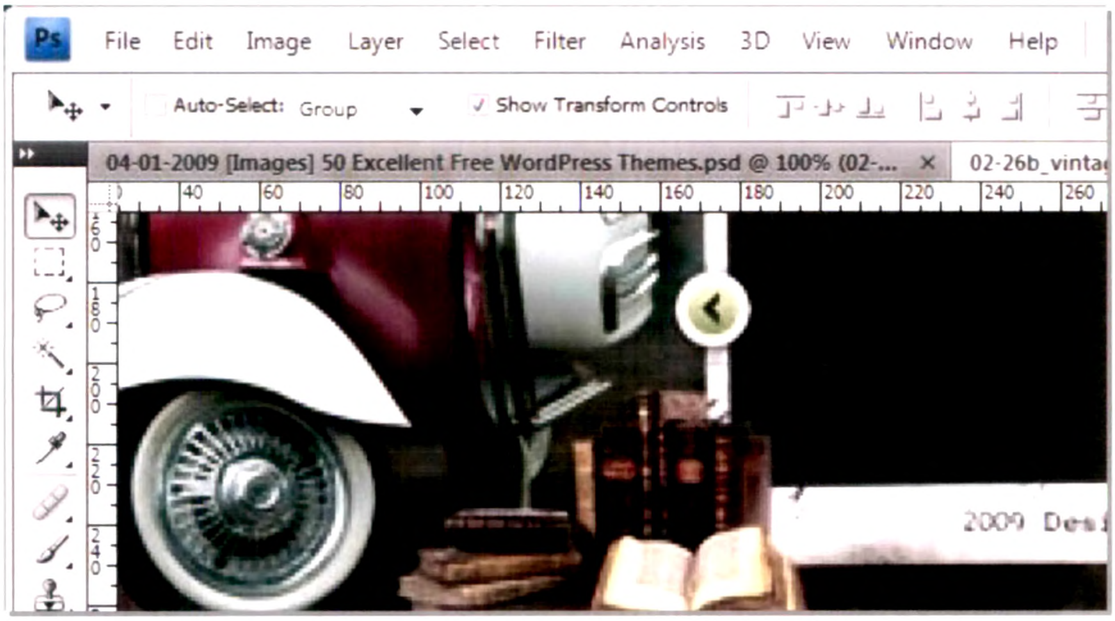


Figure - 1.13: Screenshot of Photoshop

Not surprisingly, Photoshop is the winner by a landslide, garnering over half of all the total votes. Photoshop is what comes to mind when image editing is involved there's very little that can be said about it that hasn't been said already.

With an insurmountable amount of features that help you manipulate and enhance photos as well as create web graphics, all while helping you manage your workflow and image editing environment - Photoshop comes in at numero uno as the best image editing software currently in the market.

In order to keep an eye on such manipulations, the research is conducted, and a variety of detection tools are made available. These tools can act as an investigating agency for the documents which are assumed to be manipulated or tampered.

Although Photoshop being an image editing tool, its developer company-Adobe declared a plans to start rolling out the technology in a number of photo-authentication plug-ins in 2008, for its Photoshop product beginning. Recently, in May 2011, Adobe conducted three-day hands-on workshop covering the use of Photoshop in a forensics environment in areas from basic photo lab applications through latent print comparison, forensic video analysis, and creating court charts. It provided a foundation for the use of imaging in any forensic discipline.

Figure - 1.14: IPS-7300 Image Processing System

One of the examples for forgery detection is the IPS-7300 Machine, which is a full color video processor for the simultaneous display of live and stored images. As shown in Figure - 1.14, it is designed for Questioned Document Examination. It contains functions such as overlay, addition, subtraction and side by side comparison to enable easy comparison of genuine and suspect documents.

1.6 Types of Image Forgery

Any type of modification in a digital image can be referred as Image Forgery, which is also known as Image Manipulation or Image Tampering or Image Doctoring or Image Fraud.

The classification of image forgery is difficult, since there is no limit to the creativity of a person, who manipulates the image for the required cause. The concept of such changes can however be broadly categorized as:

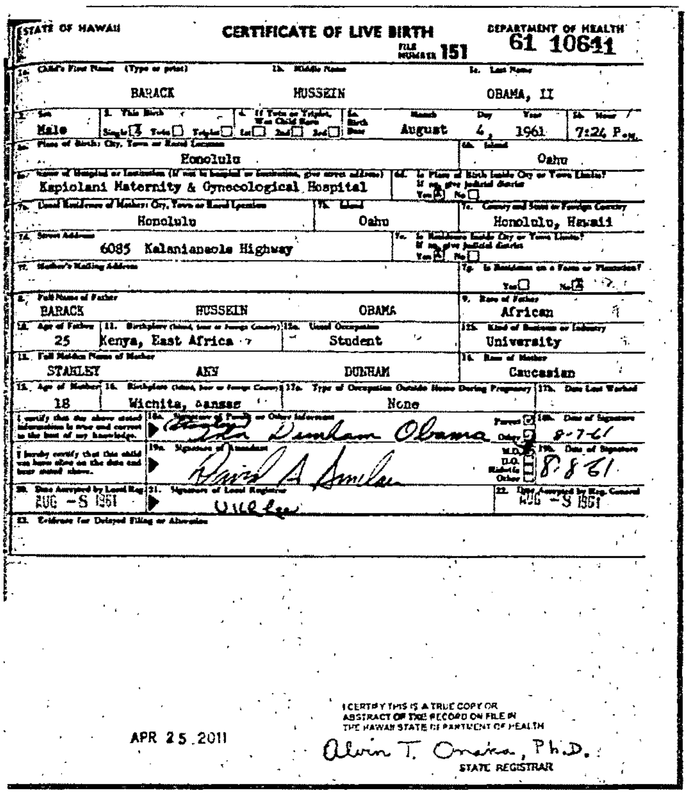
1. Image Retouching: It refers to the feature of an image editing tool, which makes a minor change in the overall impact of an image. Such retouching can be in form of color or contrast change, resolution setting or saturation effect.
2. Image splicing: It refers to the composition of two or more images to describe the situation.
3. Copy-move or Copy-paste: The image may contain a portion which is taken from another image, or some areas of the same image may be copied in the image itself.

Considering the third approach, such forgery can be further classified as:

(a) Elimination forgery: It is the case where a portion of image is hidden by the area which is pasted from the external source.

(b) Enhance forgery: It is the case where a portion of image is repeatedly pasted to show more area or occurrences of the desired ROI (Region of interest).

Since an image can be manipulated in several ways, it is a challenge to determine the type of forgery. It is also a prime concern to determine whether an image is original or forged.



STATE OF HAWAII — CERTIFICATE OF LIVE BIRTH — DEPARTMENT OF HEALTH

FILE NUMBER 151 — 61 10641

1a. Child's First Name (Type or print) BARACK	1b. Middle Name HUSSEIN	1c. Last Name OBAMA, II

2. Sex Male	3. This Birth Single ☒ Twin □ Triplet □	4. If Twin or Triplet, Was Child Born 1st □ 2nd □ 3rd □	5a. Birth Date Month August Day 4, Year 1961	5b. Hour 7:24 P.M.

6a. Place of Birth: City, Town or Rural Location Honolulu	6b. Island Oahu
6c. Name of Hospital or Institution (If not in hospital or institution, give street address) Kapiolani Maternity & Gynecological Hospital	6d. Is Place of Birth Inside City or Town Limits? If no, give judicial district Yes ☒ No □

7a. Usual Residence of Mother: City, Town or Rural Location Honolulu	7b. Island Oahu	7c. County and State or Foreign Country Honolulu, Hawaii
7d. Street Address 6085 Kalanianaole Highway		7e. Is Residence Inside City or Town Limits? If no, give judicial district Yes ☒ No □
7f. Mother's Mailing Address		7g. Is Residence on a Farm or Plantation? Yes □ No ☒

8. Full Name of Father BARACK	HUSSEIN	OBAMA	9. Race of Father African
10. Age of Father 25	11. Birthplace (Island, State or Foreign Country) Kenya, East Africa	12a. Usual Occupation Student	12b. Kind of Business or Industry University
13. Full Maiden Name of Mother STANLEY	ANN	DUNHAM	14. Race of Mother Caucasian
15. Age of Mother 18	16. Birthplace (Island, State or Foreign Country) Wichita, Kansas	17a. Type of Occupation Outside Home During Pregnancy None	17b. Date Last Worked

I certify that the above stated information is true and correct to the best of my knowledge.	18a. Signature of Parent or Other Informant Ann Dunham Obama	Parent ☒ Other □	18b. Date of Signature 8-7-61
I hereby certify that this child was born alive on the date and hour stated above.	19a. Signature of Attendant David A Sinclair	M.D. ☒ D.O. □ Midwife □ Other □	19b. Date of Signature 8-8-61
20. Date Accepted by Local Reg. AUG -8 1961	21. Signature of Local Registrar U K L Lee		22. Date Accepted by Reg. General AUG -8 1961

23. Evidence for Delayed Filing or Alteration

APR 25 2011

I CERTIFY THIS IS A TRUE COPY OR ABSTRACT OF THE RECORD ON FILE IN THE HAWAII STATE DEPARTMENT OF HEALTH

Alvin T. Onaka, Ph.D.
STATE REGISTRAR

Figure - 1.15: Barack Obama's Birth Certificate

Figure - 1.15 is a very interesting example of a document forgery, which claims that a birth certificate of US President Barack Obama is forged, and contains multiple reasons for it to be declared as invalid. A special report is also available in form of News Release of June 13, 2011 as "Final Analysis of President Obama's Certificate of Live Birth" submitted by Douglas Vogt.

Chapter 2

REVIEW OF LITERATURE

⊗ The Forensic Framework and laws

⊗ Digital Fraud detection techniques

⊗ Watermarking Schemes

⊗ Modern Forgery detection techniques

CHAPTER-2: REVIEW OF LITERATURE

2.1 The Forensic Framework and laws

A security of information is vital for every organization, for a better and efficient management. The assets of information have to be protected for a business as well as private data. The processes required for the protection of the data are dependent on the human behavior. A framework for the reference of information security (***Niekerk and Solms, 2010***) is essential, because the employees are considered to be a threat to an organization, either due to their dire intentions, through their negligence, or in most of the cases, due to the lack of knowledge. The reference framework suggests the following:

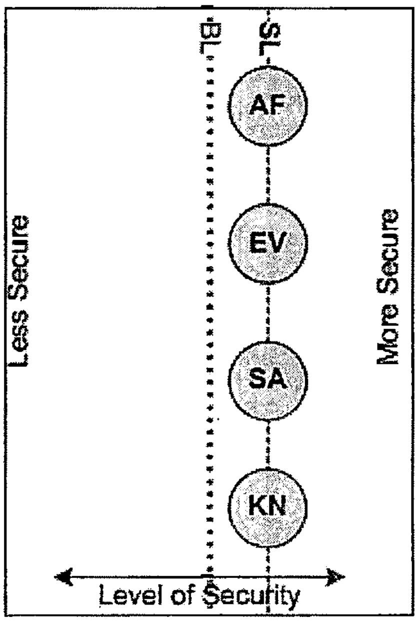


Figure 2.1: Forensic Framework

BL: Minimum Acceptable Baseline - This line indicates what would be an acceptable minimum security baseline; in other words, a culture whose net effect would meet the minimum requirements for some industry standard.

SL: Nett Security Level - This line indicates the actual net effect of the culture on the overall security effort. This line can be seen as the cumulative effect of the four underlying levels of the culture. The net security level (SL) can either be more secure (to the right), less secure (to the left), or just as secure (overlapping) as the minimum acceptable baseline (BL).

AF: Artifacts - This node represents the relative strength of the artifact level (AF) of the culture. If this node is to the left of the minimum acceptable baseline (BL), it indicates that the measurable artifacts are not as secure as they should be. A node to the right of the baseline (BL) would indicate artifacts that are even more secure than the acceptable minimum. A node exactly on the baseline (BL) would indicate artifacts that are just as secure as required by this baseline.

EV: Espoused Values - This node represents the relative strength of the organization's espoused value level (EV). The various policies and procedures comprising this level could be more, less, or just as comprehensive than those recommended as the minimum acceptable baseline.

SA: Shared Tacit Assumptions - This node represents the relative strength of the organization's shared tacit assumption level (SA). The underlying beliefs or values of the employees could be either more, less, or just as in favor of good secure practices as required by the minimum accept able baseline.

KN: Knowledge - This node represents how much knowledge the organization's employees have regarding information security. Employees can be more knowledgeable than a certain minimum level needed to perform their jobs securely, they could be less knowledgeable, or they could have exactly the minimum requisite level of knowledge.

There are ever challenging implications of a data theft present in the jurisdiction and laws in India and many other countries. Such rules come into picture in the recent era, due to the advances in digital technology (***Biswas, 2011***). The Indian Penal Code (IPC), 1860 have specified that "whoever being in any manner entrusted with property, or with any dominion over property, dishonestly misappropriates or converts to his own use that property, or dishonestly uses or disposes of that property in violation of any direction of law prescribing the mode in which such trust is to be discharged, or of any legal contract, express or implied, which he has made touching the discharge of such trust, or willfully suffers any other person so to do".

With an advent of the new devices and software tools for editing the digital documents, there are new rules to be applied after this IPC Act, such as - Computer Misuse Act 1990. This Act punishes unauthorized access to computer material and a person is guilty of an offence if:

(i) he causes a computer to perform any function with intent to secure access to any programmer data held in any computer.

(ii) the access he intends to secure is unauthorized.

(iii) he knows at the time when he causes the computer to perform the function that, that is the case.

The Information Technology Act, 2000 has been amended as well as legislation is introduced before the Rajya Sabha in respect of personal data protection, in the shape of the Personal Data Protection Bill. The laws for other countries with reference to the digital documents are of varying nature.

The security of data can be considered as a third wave of technology ***(Solms, 2000)***. There exists a system for automatically detecting the ways in which images have been copied and edited or manipulated ***(Kennedy and Chang, 2009)***. The conclusion can be drawn upon these manipulation cues to construct probable parent-child relationships between pairs of images, where the child image was derived through a series of visual manipulations on the parent image. Through the detection of these relationships across a plurality of images, we can construct a history of the image, called the visual migration map (VMM), which traces the manipulations applied to the image through past generations. The VMMs are proposed to be applied as part of a larger internet image archaeology system (IIAS), which can process a given set of related images and surface many interesting instances of images from within the set. In particular, the image closest to the "original" photograph might be among the images with the most descendants in the VMM. The images that are most deeply descended from the original may exhibit unique differences and changes in the perspective being conveyed.

The system is evaluated across a set of photographs crawled from the web and it was found that many types of image manipulations can be automatically detected and used to construct plausible VMMs. These maps can then be successfully mined to find interesting instances of images and to suppress uninteresting or redundant ones, leading to a better understanding of how images are used over different times, sources, and contexts.

2.2 Digital Fraud detection techniques

2.2.1 Use of Timestamps

The use of stochastic forensic characteristics is considered to be one of the first steps towards the forgery detection. Such method examines the file system to find out the time at which the files were copied. This technique ***(Bose et al., 2011)*** can be implemented by using the stochastic model of the file system and determining the Mac file timestamps, which are exclusively used to check the copy operation. The advantage of the suggested method is that, the detection is possible at any time in future also. Here, the silent features of an image files are extracted successfully by using the concept of stochastic.

The connectivity ***(Rosenfeld, 1970)*** and adjacency ***(Rosenfeld, 1974)*** in a digital picture is a peculiarity to determine the common elements amongst them. The two simply connected sets that have the same area are IP equivalent ***(Rosenfeld and Nakamura, 2002)***.

The properties of the files can be determined using the following features:

(i) File access: Depending on the type of operation performed on the files, a timestamp always suggests the fact. It determines whether there is an update in the time-stamp of file access, and if it is so, it also identifies the set of files with such case. The copy of all files or a portion of the folder can be found. Comparatively, in case of virus check, only specific files such as executables are updated and searching will result into the change in subsequence of the file name.

(ii) Skipped folders and files: Under certain known conditions, few files and folders cannot be determined. These files or folders may be system (OS) files and executables, hidden files, files names starting with '.' (period), Alternate Data Streams and hidden files of NTFS and 'Thumbs.db' file in case of Windows OS.

(iii) Method of tree traversal: There are possibilities that a recursion method is applied with either depth first approach or breadth first or another method.

(iv) Visit order of sibling files: The default order of visit is according to file system. However, the alphabetical or date or file type may also be used as a criterion. The order is of importance in a situation where a folder consists of a set of files, as well as folders, and priority can start either with a file or folder.

(v) Speed: It refers to the rate at which the files and folders are accessed by the user. It is also dependent on the number of files and the file size. In case of copying a folder, the change in timestamp of files and folders inside it are noted prior to the actual copy operation.

The approach of using timestamps of a file can also be analysed by using attack of pod slurping ***(Kavallaris and Katos, 2010).*** This technique, similar to the one suggested by Grier, also considers the time as a critical factor, and constructs a synthetic metrix from the timestamp of a file system.

In this system, the probability of an unauthorized copy of a file can be determined with an aid of comparison between the rate of file transfer by using USB drive detected through the Windows OS registry ***(Carvey, 2005)*** and the rate of transfer previously known from the last access. From the initial stages of identification, it can be concluded that the rate of file transfer is directly dependent on the model and make of USB device, and thus assists the investigator to determine a leakage of files.

The following conclusions are drawn for the attacks of pod slurping:

- The leakage probability L is negligible when there is no slurping attack. This suggests that false positives would also be negligible.
- The leakage probability L is high (in the range of 0.3 to 1.0), even for a small number of files transferred in the case of a slurping attack.

- The leakage probability L decreases monotonically for files accessed after the attack is complete.

With the suggested methods available for file system verification, any of the existing technique can work depending on the particular type, considering the survey of outlier detection methodologies ***(Hodge and Austin, 2004).*** An intrusion can also be detected based on the host, by using dynamic and static behavioral models ***(Yeung and Ding, 2003).***

2.2.2 Plastic card fraud detection.

Due to modernization in the banking system, the customers are provided with direct access to their accounts for the deposit, withdrawal and many other transactions through credit or debit cards. A framework suggesting the detection of frauds made through such plastic cards is very useful for our study, since it ultimately focuses to the solution strategies.

A model for plastic card fraud detection ***(Fawcett and Provost, 2002)*** systems suggests a useful framework for such fraud.

The hybrid model for fraud detection suggests data-customised approach combines elements of supervised and unsupervised methodologies aiming to compensate for the individual deficiencies of the methods ***(Krivko, 2010).*** It demonstrates the ability of the hybrid model to identify fraudulent activity on the real debit card transaction data. The framework also explores the model's efficiency against that of the existing monitoring system of the collaborating bank, using appropriate performance assessment criteria.

Once the model is constructed one can assess the "status" of each account with every new transaction made. The hybrid model produces a suspiciousness score, which is a real number between 0 and 1 associated with each account which is updated as a new transaction occurs. This is then compared with a threshold value in order to assign account "status" to one of two classes: "suspected to be compromised" and "assumed to be legitimate". The threshold is set up during model training such that it delivers user-specified values of performance measures.

Frauds can be detected by reviewing different statistical methods ***(Bolton and Hand, 2002).***

2.2.3 Biometrics

Biometrics is another security tool most widely accepted for the authorization. This includes the input in form of fingerprints, iris recognition, face detection, ear pattern, voice input and many others.

A novel image hiding approach based on correlation analysis for secure multimodal biometrics proposes methodology based on correlation analysis, which is used to protect the security and integrity of transmitted multimodal biometric images for network-based identification ***(Grier, 2010).*** Compared with existing methods, the correlation between the biometric images and the cover image is first analyzed by partial least squares (PLS) and particle swarm optimization (PSO), aiming to make use of the abundant information of cover image to represent the biometric images. Representing the biometric images using the corresponding content of cover image results in the generation of the residual images with much less energy. Then, considering the human visual system (HVS) model, the residual images as the secret images are embedded into the cover image using middle-significant-bit (MSB) method. Extensive experimental results demonstrate that the proposed approach not only provides good imperceptibility but also resists some common attacks and assures the effectiveness of network-based multimodal biometrics identification.

2.2.4 File Systems

A similar approach is also available to check the self-similarity in file systems ***(Gribble et al., 1998).*** The attacks made on sensitive data by insider can be indentified through SIDD ***(Yali et al., 2009)***: A framework for such detection.

In addition to the timestamps, other properties of files and folders are also useful for determining the history or file forensic. One of the approaches is the use of purpose-built functions and block hashes to enable small block and sub-file forensics ***(Garfinkel et al., 2010).***

There is a growing need for automated techniques and tools that operate on bulk data, and specifically on bulk data at the block level. The reasons for such requirement are as stated below:

- File systems and files may not be recoverable due to damage, media failure, partial overwriting, or the use of an unknown file system.
- There may be insufficient time to read the entire file system, or a need to process data in parallel.
- File contents may be encrypted.
- The trees structure of file systems makes it hard to parallelize many types of forensic operations.

The research work by Garfinkel *et al* introduces an approach for performing small block forensics. Some of this work is based upon block hash calculations, that is, the calculation of cryptographic hashes on individual blocks of data, rather than on entire files. Other work is based on bulk data analysis and the examination of blocks of data for specific features or traits irrespective of file boundaries.

Standardization for forensic corpora is also anticipated to bring science to digital forensic ***(Gerfinkel et al., 2009).*** The statistical analysis can also help the identification of data type ***(Moody and Erbacher, 2008).***

The timestamp of files can also be used with a different perspective, by leaving timing-channel fingerprints ***(Shebaro et al., 2010)*** in hidden service log files.

There are three main reasons why, among the many information channels various log files afford, we focus on only timing channels using the timestamps:

For legal reasons, standardized methods are preferable to ad-hoc methods, because precedents can be established for well-analyzed algorithms for recovering a footprint. This requires that a single method be used for many services, and, while various services log different data that is application-specific, most contain some sort of timestamp.

Anonymization technologies sometimes hide IP addresses, URLs, and other objects in the log file. For example, when Apache is set up as a Tor hidden service using Proxy, the IP address for all logged requests is 127.0.0.1 due to local proxying. Timing information, on the other hand, is typically preserved.

By using exclusively timing and timestamps for leaving the fingerprint, the other channels of information (e.g., the URL of the document being requested) can be reserved for other information that the fingerprinter may want to preserve in the log. For example, proof of the existence of a file on the server at a given time.

By using TCP timestamps, covert messaging can be performed ***(Giffin et al., 2002)***. Wray, J.C. also executed the analysis of covert timing channels in 1991.

The advanced forensic format can also be extended to accommodate multiple data sources, logical evidence, arbitrary information and forensic workflow ***(Cohen et al., 2009)***. A framework for managing and storing digital evidence is suggested, where existing evidence management file formats are first examined and then their strengths and limitations are outlined. The proposed Advanced Forensics Format (AFF4) framework extends these efforts into a universal evidence management system. A forensic investigation framework can also be proposed based on the event ***(Carrier and Spafford, 2004)***. Consequently, advance carving techniques ***(Cohen, 2007)***, as well as specialized tool Pyflag, used as advanced network forensic framework ***(Cohen, 2008)*** are proposed,

Contemporary fingerprint system uses solid flat sensor which requires contact of the finger on a platen surface. This often results in several problems such as image deformation, durability weakening in the sensor, latent fingerprint issues which can lead to forgery and hygienic problems. On the other hand, bio-metric characteristics cannot be changed; therefore, the loss of privacy is permanent if they are ever compromised. Coupled with template protection mechanism, a touch-less fingerprint verification system is further provoked. In this issue, a secure end-to-end touch-less fingerprint verification system is presented ***(Hiewa et al., 2010)***. The fingerprint image captured with a digital camera is first pre-processed via the proposed pre-processing algorithm to reduce the problems appear in the image. Then, Multiple Random Projections-Support Vector Machine (MRP-SVM) is proposed to secure fingerprint template while improving system performance.

An image pattern, as used in many applications such as geoseismic surveys or medical diagnostics, can also be used for improving radiometry of imaging spectrometers by using programmable spectral regions of interest.

Programmable imaging spectrometers can be adjusted to fit specific application requirements that differ from the instrument initial spectral design goals. Sensor spectral characteristics and its signal-to-noise ratio (SNR) can be changed by applying customized online binning patterns ***(Dell'Endice et al., in press)***.

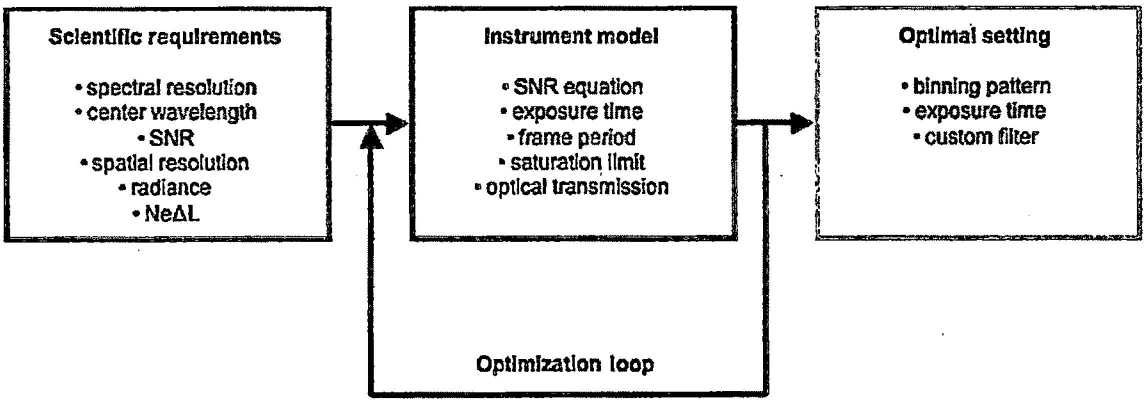


Figure 2.2: BinGO Model

The researchers have devised a software utility that generates application driven spectral binning patterns by using an SNR dependent sensor model. The utility, named BinGO (Binning Pattern Generator and Optimiser), is used to produce predefined binning patterns that either

(a) Allow an existing imaging spectrometer to optimize its spectral characteristics for a specific application.

(b) Allow an existing imaging spectrometer to spectral and/or spatially emulate another instrument.

(c) Design new multispectral or imaging spectrometer missions, which may be space borne, airborne or terrestrial.

The noise distributions for imaging spectrometers can be studied for analysis ***(Nieke et al., 1999).***

When the image forgery is performed, the preservation of connectivity between the pixels becomes questionable issue (Bose et al., 2011). By considering local modification operation on binary images in which black pixel p and a white pixel q are interchanged, the interchange operation can be performed on the current image I to obtain a new image I′, where

$$I'(x) = \begin{matrix} I(p) & \text{if } x = q \\ I(q) & \text{if } x = p \\ I(x) & \text{otherwise.} \end{matrix}$$

The exclusive efforts were made in research in form of a Special issue on image and video retrieval evaluation ***(Hanbury et al., 2010).*** Each of the work in this special issue deals with various aspects of designing suitable evaluation resources to promote research and development in image and video retrieval systems, along with the technologies required to implement such frameworks in practice. The evaluation ideas being used are - Pictorial information retrieval ***(Enser, 1995),*** overview and proposals for performance evaluation in content-based image retrieval ***(Muller et al., 2001)*** and Image retrieval evaluation ***(Smith, 1998).***

The audio signals can also be checked to have the similarity using Recursive Nearest Neighbor Search in a Sparse and Multiscale Domain ***(Sturm and Daudet, 2011)***. To approximate the cosine distance between the signals, pairwise comparisons are made between the elements of localized sparse models built from large and redundant multiscale dictionaries of time-frequency atoms.

Image processing and analysis are critically important for the medical images also. A tumor can be identified from the given image with an aid of wavelet packet transforms and neighborhood rough set ***(Zhang et al., 2010)***.

Tumor classification is an important application domain of gene expression data. Because of its characteristics of high dimensionality and small sample size (SSS), and a great number of redundant genes not related to tumor phenotypes, various feature extraction or gene selection methods have been applied to gene expression data analysis. Wavelet packet transforms (WPT) and neighborhood rough sets (NRS) are effective tools to extract and select features. A novel approach of tumor classification is available, based on WPT and NRS. First the classification features are extracted by WPT and the decision tables are formed, then the attributes of the decision tables are reduced by NRS.

Thirdly, a feature subset with few attributes and high classification ability is obtained. The experimental results on three gene expression datasets demonstrate that the method proposed by Zhang *et al* is effective and feasible.

SVM-Support Vector Machine is a relatively new type of statistic learning theory. It builds up a hyper-plane as the decision surface to maximize the margin of separation between two-class samples.

K-NN is a most common and non parametric method. To classify an unknown sample x, K-NN extracts k closest vectors from the training set using similarity measures, and makes decision for the table of the unknown sample x using the majority class label of the k nearest neighbors. Here, Euclidean distance is used to measure the similarity of samples.

NEC is similar to K-NN, and based on the general idea of estimating the class of unknown sample according to its neighbors, but differing from K-NN, NEC considers a kind of neighbor within a sufficiently small and near area around the sample, in other words, all training samples surrounding the test sample take part in the classification decision process.

The Biomarkers are identified by feature wrappers ***(Xiong et al., 2001)***. Equivalent approaches to this technique are neighborhood classifiers ***(Hu et al., 2008)*** and Neighborhood operator systems and approximations ***(Wu and Zhang, 2002)***.

The survey of almost 300 key theoretical and empirical contributions in the current decade is related to image retrieval and automatic image annotation ***(Datta et al., 2008)***. The significant challenges are involved in the adaptation of existing image retrieval techniques to build systems that can be useful in the real world. In retrospect of what has been achieved so far, along with the inference of what the future may hold for image retrieval research.

Ideal Spatial adaptation can be verified by wavelet shrinkage ***(Donoho and Johnstone, 1994)*** and Multiresolution analysis ***(Cohen et al., 1993)***. The algorithms are devised to contain the optimal results.

2.3 Watermarking Schemes

The digital image can be protected with most popular approach to Watermarking, which contains images with Self-Correcting Capabilities ***(Fridrich and Goljan, 1999)***. There are two techniques for self-embedding an image in itself as a means for protecting the image content. After self-embedding, it is possible to recover portions of the image that have been cropped out, replaced, damaged, or otherwise tampered without accessing the original image. The first method is based on transforming small 8×8 blocks using a DCT, quantizing the coefficients, and carefully encoding them in the least significant bits of other, distant squares. This method provides very high quality of reconstruction but it is very fragile. The quality of the reconstructed image areas is roughly equivalent to a 50% quality JPEG compressed original. The second method uses a principle similar to differential encoding to embed a circular shift of the original image with decreased color depth into the original image. The quality of the reconstructed image gradually degrades with increasing amount of noise in the tampered image. The first technique can also be used as a fragile watermark for image authentication, while the second technique can be classified as a semi-robust watermark. The watermark techniques are proposed with following criteria:

- Slippery New age ***(Walton, 1995)***
- General concept ***(Wolfgang and Delp, 1996)***
- Invisible technique ***(Yeung and Mintzer, 1997)***
- Distortion measurement ***(Zhu et al., 1997)***
- Image integrity and Ownership verification ***(Wong, 1998)***
- Tamper Detection ***(Fridrich, 1998a)***
- Methods for Detecting changes ***(Fridrich, 1998b)***

The watermarking approach is considered to be widely accepted, where semi-fragile watermarking is useful for authenticating JPEG visual content ***(Lin and***

Chang, 2000). In order to practically implement the approach, effective tool like SARI - Self-Authentication-and-Recovery Image Watermarking System ***(Lin and Chang, 2001)*** is accessible. In SARI project, a novel image authentication system is designed, based on semi-fragile watermarking technique. The system, called SARI, can accept quantization-based lossy compression to a determined degree without any false alarm and can sensitively detect and locate malicious manipulations. It is the first system that has such capability in distinguishing malicious attacks from acceptable operations. Furthermore, the corrupted area can be approximately recovered by the information hidden in the image. The amount of information embedded in our SARI system has nearly reached the theoretical maximum zero-error information hiding capacity of digital images. The software prototype includes two parts - the watermark embedder that is freely distributed and the authenticator that can be deployed online as a third-party service or used in the recipient side. This is an example of implementation of concepts in form of software, which is common is our current research.

Reversible fragile watermarking is used for locating tampered blocks in JPEG images ***(Zhang et al., 2010).*** It proposes a fragile watermarking scheme for JPEG images, in which two watermark bits are embedded into each block using a reversible data-hiding scheme. On the receiver side, after attempting to extract the watermark data and to recover the original content, the number of mismatches between the watermark data extracted from the received image and derived from the recovered contents is used to judge whether a block has been tampered. If the content replacement is not serious, we can always identify the blocks containing fake contents and perfectly recover the original information of the remaining blocks. Similarly, the watermarking schemes such as - Secret and Public key ***(Wong and Memon, 2001)*** and Gradient image for improved localization ***(Suthaharan, 2004)*** detects the image forgery.

The Secure hybrid approach against tampering and copy attack ***(Deguillaume et al., 2003)*** is known to be effective method. Imperceptibility and robustness of Genetic watermarking can be verified with the study of the effect DCT and DWT domains ***(Shaamala et al., 2011)***. The more common and easily implemented method, which is also comparable to forgery detection referred in Chapter-4, is a watermarking based on DCT-domain of three RGB color channels ***(El-Fegh et al., 2009)***.

Watermarking using genetic algorithm for the optimization of the tread-off between the watermarking requirements has attacked the attention of researchers; amongst the watermarking requirements, the imperceptibility and robustness is one of the main requirements. The image adaptive watermark is generated based on image features, which allows the sharp detection of microscopic changes to locate modifications in the image ***(Shefali et al., 2007)***. Further, the scheme utilizes the multipurpose watermark consisting of soft authenticator watermark and chrominance watermark, which has been proved fragile to some predefined processing like intentional fabrication of the image or forgery and robust to other incidental attacks caused in the communication channel. The invisible watermarking in an image is implemented, as shown in Figure 2.3.

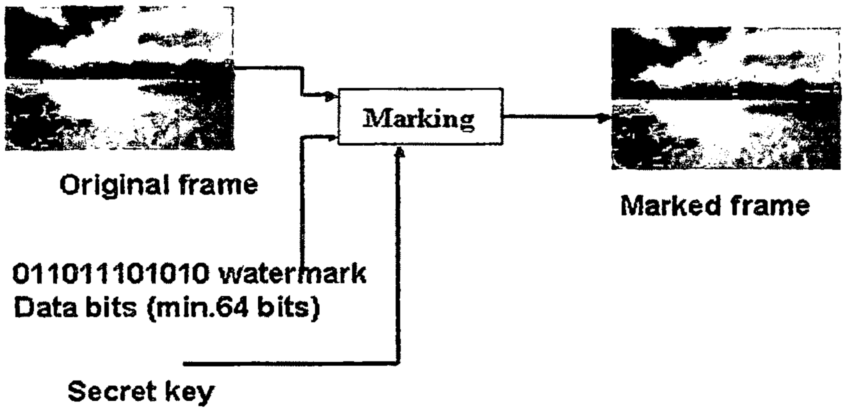


Figure 2.3: Invisible Watermarking in an image.

A hierarchical digital watermarking method can also be used for image tamper detection and recovery ***(Lin et al., 2005)***. The method is efficient as it only uses simple operations such as parity check and comparison between average intensities. It is effective because the detection is based on a hierarchical structure so that the accuracy of tamper localization can be ensured. That is, if a tampered block is not detected in level-1 inspection, it will be detected in level-2 or level-3 inspection with a probability of nearly 1. This method is also very storage effective, as it only requires a secret key and a public chaotic mixing algorithm to recover a tampered image. The experimental results demonstrate that the precision of tamper detection and localization is 99.6% and 100% after level-2 and level-3 inspection, respectively. The tamper recovery rate is better than 93% for a less than half tampered image. The method is not only as simple and as effective in tamper detection and localization, it also provides with the capability of tamper recovery by trading off the quality of the watermarked images about 5 dB.

Watermarking scheme can detect the tamper as well as provide recovery mechanism ***(Lin et al., 2007)***. The main goal is to detect and recover the tampered region accurately. In addition, the proposed method has robustness to resist the attacks of JPEG compression and cropping.

Watermarking techniques can be classified as robust, semi-fragile and fragile. Robust watermarks are designed to survive intentional (malicious) and unintentional (non-malicious) modifications of the watermarked image, Semi-fragile watermarks are layout for detecting any unauthorized alteration, and allowing in the same time some image processing operations. On the contrary, a watermarking technique that cannot robust against noise or attacks is called fragile technique. Fragile watermarking techniques are concerned with complete integrity verification. Furthermore, watermarking techniques can be classified as blind and non-blind, Blind watermarking techniques do not require access to the original un-watermarked data of image, video, audio, etc. to recover the watermark. In contrast, non-blind watermarking technique requires the original data needed for extraction of the watermarked. In general, the non-blind scheme is more robust than the blind watermark as it is obvious that the watermark can be extracted easily by knowing the un-watermarked data.

2.4 Modern Forgery Detection Techniques

2.4.1 Data Hiding

The major challenge to the processing of digital images is retaining the image into its original form. The method used for adaptive reversible data hiding scheme based on integer transform ***(Peng et al., 2012)*** is suitable for high capacity.

It is based on integer transform and adaptive embedding, and proposes a new reversible data hiding algorithm. By tuning parameter in integer transform, data can be embedded adaptively according to the block type determined by the pre-estimated distortion. In this way, it avoids large distortion generated by noisy blocks and can embed more data into smooth blocks, and thus image quality is improved compared with the previous integer-transform-based algorithm. For future work, reversible data hiding can be designed in a more meaningful way, to further enhance capacity-distortion performance while keeping low computational complexity. Moreover, investigating practically and theoretically, the maximum embedding capacity with reversibility for natural image is also an interesting problem.

As there are several methods for information hiding ***(Petitcolas, 1999),*** a steganography method is useful for images by pixel-value differencing ***(Wu and Tsai, 2003)***. The reversible data embedding is implemented by using difference expansion ***(Tian, 2003)***.

The software tools, which successfully implement the ideas, are considered to be more powerful for the image processing or forensic detections. One of such example is DART, software to analyse root system architecture and development from captured images ***(Bot et al., 2010).***

Image analysis is used in numerous studies of root system architecture (RSA). To date, fully automatic procedures have not been good enough to completely replace alternative manual methods. DART (Data Analysis of Root Tracings) is freeware based on human vision to identify roots, particularly across time-series. Each root is described by a series of ordered links encapsulating specific information and is connected to other roots. The population of links constitutes the RSA. DART creates a comprehensive dataset ready for individual or global analyses and this can display root growth sequences along time. It exemplifies individual tomato root growth response to shortfall in solar radiation and analyses the global distribution of the inter-root branching distances.

Similar to DART, EZ-RHIZO is integrated software for the fast and accurate measurement of root system architecture ***(Armengaud, 2008)***.

The Learning approach is useful in form of one of the several ideas of Artificial Intelligence. Based on this fact, Sparsity-based Image Denoising ***(Dong et al., 2011)*** can be determined via Dictionary Learning and Structural Clustering. Image denoising can be carried out with a non-local algorithm ***(Buades et al., 2005)***, while image restoration is carried with non-local sparse models ***(Mairal et al., 2009)***.

The concept presents a variational framework for unifying the two views and propose a new denoising algorithm built upon clustering-based sparse representation (CSR). Inspired by the success of l1-optimization, it formulates a double-header l1-optimization problem where the regularization involves both dictionary learning and structural structuring.

Considering the geoseismic applications for digital images, the analysis of images aid us in our study to check the similarities. The process of Co-registration and correlation of aerial photographs for ground deformation measurements ***(Ayoub et al., 2009)*** is one of such approach.

The technique requires the digitization of the film based photographs with a high spatial and radiometric resolution scanner. Digital photography is not considered in this study as aerial photography archives are mainly film based.

However, the technique described by them could be used with digital frame cameras as well.

2.4.2 Geographical images

The triggered aseismic fault slip from nearby earthquakes may be static or dynamic ***(Du et al., 2003)***. The optimal imagery is used for monitoring of earth surface dynamics ***(Leprince et al., 2008)***.

De-Noising with the traditional orthogonal, maximally-decimated wavelet transform sometimes exhibits visual artifacts; we attribute some of these, for example, Gibbs phenomena in the neighborhood of discontinuities to the lack of translation invariance of the wavelet basis. One method to suppress such artifacts, termed cycle spinning ***(Coifman and Donoho, 1998)***, is to average out the translation dependence. For a range of shifts, one shifts the data (right or left as the case may be), De-Noises the shifted data, and then unshifts the de-noised data.

Prior to comparison, images are co-registered through their ortho-rectification on a common reference system. Cumulative uncertainties on both the acquisition parameters and topography lead to mis-registrations between the ortho-rectified images to be compared. The co-registration is therefore improved by optimizing the acquisition parameters of the second image as slave, with respect to the first ortho-rectified image as master.

Ortho-rectified and precisely co-registered images are then correlated using a sliding window. At each step, horizontal offsets along the East or West and North or South directions are measured and stored.

The proposal of determining the noise from an image is exceptionally good criterion to aid forensic. The blind image forensic approach is applied ***(Mahdian and Saic, 2009)*** by using noise inconsistencies.

2.4.3 Blind Methods

Existing digital forgery detection methods are divided into active and passive blind approaches. The blind approach is regarded as the new direction and interest in this field has rapidly increased over the last few years. In contrast to active approaches, blind approaches do not need any explicit priori information about the image. They work in the absence of any digital watermark or signature. Blind approaches have not yet been thoroughly researched by many.

The forgeries can be reduced by using writer-independent off-line signature verification through ensemble of classifiers ***(Bertolini et al., 2010).*** In this work, two important issues of off-line signature verification are addressed. The first one regards feature extraction. A new graphometric feature set is developed, that considers the curvature of the most important segments, perceptually speaking, of the signature. The idea is to simulate the shape of the signature by using Bezier curves and then extract features from these curves. The second important aspect is the use of an ensemble of classifiers based on graphometric features to improve the reliability of the classification, hence reducing the false acceptance.

The statistical properties are another set of data useful in forensic.

The distance between histograms of image ***(Cha and Srihari, 2002)*** exposes many of the forensic characteristics.

Pondering the digital technology, the digital documents or files, similar to digital images, also carry the same signal properties. A signature of an individual saved in form of digital document, can be verified offline by using fuzzy modeling ***(Hanmandlua et al., 2005)***. Such models are based on decisions to be derived in form of probability of forgery. The computer verification of Handwritten system is performed through multi-resolution ***(Hunt and Qi, 1995)*** approach.

The neighboring color analysis can be performed through Steganalysis for palette-based images using generalized difference image and color correlogram ***(Zhao et al., 2011).*** In this notion, an attempt was made to propose a novel blind steganalysis algorithm for palette-based images. First, the generalized difference images between adjacent pixels were constructed, and then the moments of characteristic functions of histograms of difference images were extracted as features. Second, in order to measure the dependencies of neighboring colors, color correlogram technique is used to capture the global distribution of local spatial correlation of colors. The center of mass of the characteristic function of color correlogram and the absolute moments of autocorrelogram were extracted. Total of 13 dimension features were classified with machine learning technique. Number of experiments on several existing GIF steganography algorithms indicated that the proposed scheme is effective and gets good performance, especially when the embedding rate is not less than 20%. Experimental results also show that the average accuracy of our proposed scheme for different GIF steganography algorithms outperforms Lyu's algorithm more than 20%. It also showed that the proposed scheme achieved similar performance with Fridrich's scheme and higher accuracies comparing to Du's algorithm and biologically inspired features.

A generalized Benford's law ***(Jolion, 2001)*** is useful for JPEG coefficients and its applications in image forensics ***(Fu et al., 2007)***. The statistical distribution of image DCT coefficients ***(Eggerton and Srinath, 1986)*** is also functional to image forensic.

Forensic Detection of Image Tampering is conducted by using Intrinsic Statistical Fingerprints in Histograms ***(Stamm and Liu, 2010)***. To test the performance of global contrast enhancement detection algorithm, database of 341 unaltered images consisting was used, which contains many different subjects and images captured under varying light conditions. These images were taken with several different cameras and range in size from 1500 × 1000 pixels to 2592 × 1944 pixels. To simplify the testing process, it used the green color layer of each of these images to form a set of unaltered grayscale images. Next, a set of contrast enhanced grayscale images as created by applying the power law transformation.

2.4.4 Similar Block Matching

The similarity in image blocks can be determined with Sequential Straightforward Clustering for Local Image Block Matching ***(Sekeh et al., 2011).*** The idea concentrates on computational time and proposes a local block matching algorithm based on block clustering to enhance time complexity. Time complexity of the proposed algorithm is formulated and effects of two parameter, block size and number of cluster, on efficiency of this algorithm are considered. The experimental results and mathematical analysis demonstrate that this algorithm is more cost-effective than lexicographically algorithms in time complexity issue when the image is complex.

It suggests a novel statistical model based on Benford's law for the probability distributions of the first digits of the block-DCT and quantized JPEG coefficients is presented. A parametric logarithmic law, i.e., the generalized Benford's law, is formulated. Furthermore, some potential applications of this model in image forensics are discussed in this paper, which include the detection of JPEG compression for images in bitmap format, the estimation of JPEG compression Q-factor for JPEG compressed bitmap image, and the detection of double compressed JPEG image. The results of our extensive experiments demonstrate the effectiveness of the proposed statistical model.

A simple method to detect image tampering operations that involve sharpness or bluriness adjustment is available ***(Sutchu et al., 2007)***. The approach is based on the assumption that if a digital image undergoes a copy-paste type of forgery, average sharpness or blurriness value of the forged region is expected to be different as compared to the non-tampered parts of the image. The method of estimating sharpness value of an image is based on the regularity properties of wavelet transform coefficients which involves measuring the decay of wavelet transform coefficients across scales. The preliminary results show that the estimated sharpness scores can be used to identify tampered areas of the image.

The Tamper detection techniques are based on artifacts created by Color Filter Array (CFA) processing in most digital cameras ***(Dirik and Memon, 2009)***. The techniques are based on computing a single feature and a simple threshold based classifier. The efficacy of the approach was tested over thousands of authentic, tampered, and computer generated images. Experimental results demonstrate reasonably low error rates.

The steady improvement in image and video editing techniques has enabled people to synthesize realistic images or videos conveniently. Some legal issues may occur when a doctored image cannot be distinguished from a real one by visual examination. Realizing that it might be impossible to develop a method that is universal for all kinds of images and JPEG is the most frequently used image format, we propose an approach that can detect doctored JPEG images and further locate the doctored parts, by examining the double quantization effect hidden among the DCT coefficients ***(He et al., 2006)***.

In order to check the aboriginality and integrity of a digital photograph, a blind forensics scheme for detecting blur manipulation exists. A cost-effective local blur estimator is designed to measure the blurriness of each pixel along a doubted edge. Consistency metric of such a blurriness sequence is constructed based on the deviation from its linear fitting. Then the metric is used as evidence for identifying blur operation. Experimental results both on synthetic and natural images have shown the efficiency of our proposed blur forensics scheme ***(Cao et al., 2009)***.

The biggest problem and challenge in digital image forgery is how to ensure that intellectual assets in digital form are authentic and to tampered and their entire contents are authentic and consistent, the provenance of consistency, integrity authenticity (CIA) can only assure the digital intellectual assets origin and originality ***(Math and Tripathi, 2010)***.

Chinese Remainder Theorem (CRT)-based technique for digital watermarking in the Discrete Cosine Transform (DCT) domain is robust to several common attacks ***(Patra et al., 2010)***. It compared the performance of the proposed technique with recently proposed Singular Value Decomposition (SVD)-based ***(Chung et al., 2007)*** and spatial CRT-based watermarking schemes. Experimental results have shown that the technique successfully makes the watermark perceptually invisible and has better robustness to common image manipulation techniques such as JPEG compression, brightening and sharpening effects compared to the spatial domain-based CRT scheme. The scheme is able to achieve a Tamper Assessment Function (TAF) value of less than 10% when the watermarked image undergoes JPEG compression between a range of 50 to 70%, where-as, the spatial CRT-based scheme produce TAF value of more than 35% and the SVD-based scheme produces TAF value between 10 to 40% depending on the host image, for the same range of compression.

Digital signature methods offer an interesting alternative to classical watermarking techniques, in so far there is no longer a limitation in terms of capacity, nor a problem of robustness, thus offering better localisation of the manipulated areas, better quality reconstruction, and a limited risk of false alarms. Moreover, there is already a high level of expertise in the area of community security. However, the major drawback of these techniques is that the image alone is not self-sufficient. Therefore, the benefits of watermarking are reduced and it becomes necessary to be able to guarantee the authenticity of the image or signature pair. Moreover digital signature methods are not very practical to use with multimedia documents. Finally, future developments should not exclude methods based on the combination of robust watermarking and external signature methods. Watermarking would just be an identifier which would allow a trusted user access to the registered signature ***(Rey and Dugelay, 2002).***

A solution to the real world problem of Digital Document Forgery is through a 1D hash algorithm coupled with 2D iFFT (irreversible Fast Fourier Transform) by encryption of digital documents in the 2D spatial domain ***(Cheddad et al., 2009).*** Further by applying an imperceptible information hiding technique we can add another security layer which is resistant to noise and to a certain extent JPEG compression.

Data hiding based on the similarity between neighboring pixels with reversibility ***(Li et al., 2010)*** is the technique which recovers the original image from a stego-image without distortion, once the hidden data are extracted. A natural image usually contains several smooth areas. The difference between two adjacent pixels has a high probability of being a small value. Therefore, this study proposed a novel reversible data hiding method, Adjacent Pixel Difference (APD), which employs the histogram of the pixel difference sequence to increase the embedding capacity. Experimental results reveal that APD achieves a high embedded capacity and still maintains a high stego-image quality. Furthermore, the stego-image quality and embedded capacity of the APD method outperform other existing methods.

Image Manipulation can also be detected with Binary Similarity Measures ***(Bayram et al., 2005).*** It proposed a method for digital image forensics, based on Binary Similarity Measures between bit planes used as features. It contains the design of several classifiers to test the tampered or un-tampered status of the images. The performance results in detecting and differentiating a host of attacks were encouraging as it is able to discriminate a doctored image from its original, with a reasonable accuracy. The methods are accessed with vis-à-vis the closest competitor image forensic detector. This method outperform Farid's detector especially in contrast enhancement and brightness adjustment attacks. On the other hand, while it gives better performance at stronger levels of manipulations, Farid outperforms this technique at weaker levels. In this respect, the two schemes seem to be complementary; hence fusion of forensic detectors at feature level or decision level must be envisioned.

A lossless data, like file formats other than JPEG, can be embedded by joint neighboring coding ***(Chang et al., 2009)***. Similarly, the biometric templates can be securely transmitted with hiding and secure content ***(Khan et al., 2007)***. Considering an iris as input for verification, the pattern can be authenticated based on the user-specific feature ***(Qi et al., 2008)***.

The image tampering can be detected by use of Blind Deconvolution ***(Swaminathan et al., 2007)***. The proposed method is based on the observation that many tampering operations can be approximated as a combination of linear and non-linear components.

A classifier design approach proposes a framework ***(Avcıbaş et al., 2004)***, where the following results are derived:

Table 2.1: The performance of the classifiers

Image Alternation	False Positive	False Negative	Accuracy
Scaling -10%	0%	0%	100%
Rotation, 5 Degree	0%	0%	100%
Brightness Adjustment	1%	1%	99%
Histogram Equilisation	0%	5%	97.5%
Mixed Process	0%	0%	100%

The correlation between the bit planes as well the binary texture characteristics within the bit planes will differ between an original and a doctored image. This change in the intrinsic characteristics of the image can be monitored via the quantal-spatial moments of the bit planes. These so-called Binary Similarity Measures are used as features in classifier design ***(Bayram et al., 2005)***. It has been shown that the linear classifiers based on BSM features can detect with satisfactory reliability most of the image doctoring executed via Photoshop tool. Consequently, the image manipulation is exposed through feature selection ***(Bayram et al., 2006)***. The feature selection process was implemented with the sequential forward floating search, □ SFFS method. The SFFS method analyzes the features in ensembles and can eliminate redundant ones. The floating search method ***(Pudil et al., 2003)*** claims that the best feature set is constructed by adding to or removing from the current set of features until no more performance improvement is possible. The SFFS procedure can be described as follows:

- Choose from the set of K features the best two features; i.e., the pair yielding the best classification result.
- Add the most significant feature from those remaining, where the selection is made on the basis of the feature that contributes most to the classification result when all are considered together.
- Determine the least significant feature from the selected set by conditionally removing features one by one, while checking to see if the removal of any one improves or reduces the classification result. If it improves, remove this feature and go to step 3, otherwise do not remove this feature and go to step 2.
- Stop when the number of selected features equals the number of features required.

Hàny Farid is considered to be the foremost person in developing many algorithms and concepts related to digital image forensic. As an overview of all techniques suggested for image forensic, following are the methods surveyed ***(Farid, 2009)***:

Pixel-Based : The legal system routinely relies on a range of forensic analysis ranging from forensic identification (Deoxyribonucleic acid -DNA or fingerprint) to forensic odontology (teeth), forensic entomology (insects), and forensic geology (soil).

Cloning: Two computationally efficient algorithms have been developed to detect cloned image regions, viz. DCT and PCA.

Resampling: This process requires resampling the original image onto a new sampling lattice, introducing specific periodic correlations between neighboring pixels.

Splicing: A common form of photographic manipulation is the digital splicing of two or more images into a single composite. When performed carefully, the border between the spliced regions can be visually imperceptible.

Statistical: The statistical model is composed of the first four statistical moments of each wavelet subband and higher-order statistics that capture the correlations between the various subbands.

Format based: The unique properties of lossy compression such as JPEG can be exploited for forensic analysis.

JPEG Quantization: Given a three-channel color image (RGB), the standard JPEG compression scheme proceeds as follows: The RGB image is first converted into luminance/chrominance space (YCbCr). The two chrominance channels (CbCr) are typically subsampled by a factor of two relative to the luminance channel (Y). Each channel is then partitioned into 8x8 pixel blocks. These values are converted from unsigned to signed integers. Each block is converted to frequency space using a 2-D discrete cosine transform (DCT). Depending on the specific frequency and channel, each DCT coefficient is then quantized by a quantized amount q. This stage is the primary source of compression.

Double JPEG: At a minimum, any digital manipulation requires that an image be loaded into a photo-editing software program and resaved. Since most images are stored in the JPEG format, it is likely that both the original and manipulated images are stored in this format. In this scenario, the manipulated image is compressed twice.

JPEG Blocking: It characterizes the blocking artifacts using pixel value differences within and across block boundaries.

Camera based: Inconsistencies in the artifacts of a camera can then be used as evidence of tampering.

Chromatic Aberration: In an ideal imaging system, light passes through the lens and is focused to a single point on the sensor. Optical systems, however, deviate from such ideal models in that they fail to perfectly focus light of all wavelengths. Specifically, lateral chromatic aberration manifests itself as a spatial shift in the locations where light of different wavelengths reaches the sensor.

Color Filter Array: Most CFAs employ three color filters (red, green, and blue) placed atop each sensor element. Since only a single color sample is recorded at each pixel location, the other two color samples must be estimated from the neighboring samples in order to obtain a three-channel color image.

Camera Response: Because most digital camera sensors are very nearly linear, there should be a linear relationship between the amount of light measured by each sensor element and the corresponding final pixel value.

Sensor noise: As a digital image moves from the camera sensor to the computer memory, it undergoes a series of processing steps, including quantization, white balancing, demosaicking, color correction, gamma correction, filtering and, usually, JPEG compression. This processing introduces a distinct signature into the image.

Physics based: Differences in lighting across an image can then be used as evidence of tampering.

Light Direction and environment: The required 3-D surface normals are determined by leveraging a 3-D model of the human eye.

Metric measurements: Several tools from projective geometry allow the rectification of planar surfaces and under certain conditions, the ability to make real-world measurements from a planar surface. A study performed by Farid is as below:

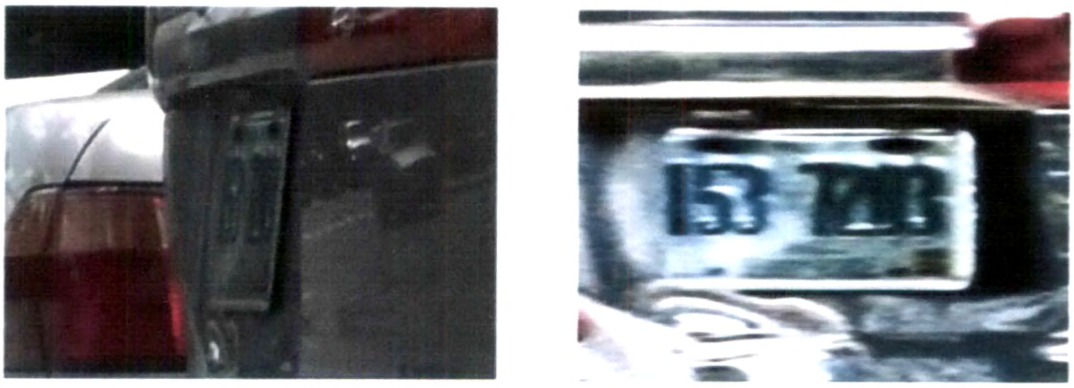


Figure 2.4: (a) Number plate not legible (b) Result of planar rectification

There exist some basic tamper detection methods ***(Fridrich, 1998)***. He provides a comprehensive overview of steganography techniques for tamper detection and authentication of digital images. The techniques are divided into several categories according to their ability to identify changes. Fragile watermarks can detect changes to every pixel and provide accurate information about the image integrity. However, it is not possible to distinguish small, innocuous changes due to common image processing operation from malicious changes, such as feature removal or addition. Semi-fragile watermarks are more robust and allow "authentication with a degree". It is possible to set a threshold in those techniques so that images after high quality JPEG compression, or contrast or brightness adjustment will still be considered authentic to a high degree. In the third category, it provides techniques that attempt to authenticate image features. Such techniques are even more robust and enable robust distinction between innocuous and malicious modifications at the expense of losing the sensitivity to small changes and sometimes the ability to localize modifications.

There are many ways to categorize the image tampering based on various points of view, where the blind methods can also be used ***(Mahdian and Saic, 2008a).*** Generally, we can say that the most often operations in photo manipulation are:

- Deleting or hiding a region in the image.
- Adding a new object into the image.
- Misrepresenting the image information.

In blind methods, as they are regarded as a new direction and in contrast to active methods, they work in absence of any protecting techniques and without using any prior information about the image or the camera that took the image. To detect the traces of tampering, blind methods use the image function and the fact that forgeries can bring into the image specific detectable changes (e.g., statistical changes).

The main drawback of existing methods is highly limited usability and reliability. This is mainly caused by the complexity of the problem and the blind character of approaches. But it should be noted that the area of Blind Methods ***(Mahdian and Saic, 2008b)*** is growing rapidly and results obtained promise a significant improvement in forgery detection in the never ending competition between image forgery creators and image forgery detectors.

A copy-move forgery in a digital image can be detected through few of the below mentioned methodologies also:

- Wavelet representation ***(Mallat, 1989)***
- Digital watermarking ***(Yeung, 1998)***
- Multi stage region merging ***(Brox et al., 2001)***
- Gradient vector diffusion ***(Yu and Bajaj, 2002)***
- Blur moment invariants ***(Mahdian and Saic, 2007)***

A successful approach is proposed to detect the forgery by analyzing the Color Filter Array values ***(Popescu and Farid, 2005)*** or Higher Order Wavelet statistics ***(Farid and Lyu, 2003).***

It quantifies the specific correlations introduced by CFA interpolation, and describes how these correlations, or lack thereof, can be automatically detected in any portion of an image. It shows the efficacy of this approach in revealing traces of digital tampering in lossless and lossy compressed color images interpolated with several different CFA algorithms.

When creating a digital composite of, for example, two people standing side-by-side, it is often difficult to match the lighting conditions from the individual photographs. Lighting inconsistencies ***(Johnson and Farid, 2005)*** can therefore be a useful tool for revealing traces of digital tampering. Borrowing and extending tools from the field of computer vision, we describe how the direction of a point light source can be estimated from only a single image. We show the efficacy of this approach in real-world settings.

In such approach, the direction of projected light source can be estimated automatically ***(Nillius and Eklundh, 2001)*** in order to verify if the lighting conditions are consistent in an image or not.

Several statistical techniques are suggested for detecting traces of digital tampering ***(Popescu and Farid, 2004a)*** in the absence of any digital watermark or signature. In particular, statistical correlations are quantified that result from specific forms of digital tampering, and detection schemes are devised to reveal these correlations. The forgeries can be exposed by detection of duplicated regions ***(Popescu and Farid, 2004b).***

A bibliography on all blind methods for image fakery detection ***(Mahdian and Saic, 2010)*** suggests the following factors as criterion:

- Near Duplicated image regions
- Interpolation and geometric transformations
- Image Splicing
- Computer Graphics and paintings
- JPEG and Compression properties

- Lighting
- CFA and inter pixel correlation
- Local noise
- Chromatic Aberration
- Blur and Sharpening
- Projective Geometry

Common sense reasoning (CSR) is a unique approach to detect the false captioning ***(Lee et al., 2006).*** Many previous signal-processing techniques are concerned about finding forgery through simple transformation (e.g. resizing, rotating, or scaling), yet little attention is given to examining the semantic content of an image, which is the main issue in recent image forgeries. Here, we present a complete workflow for finding the anomalies within images by combining the methods known in computer graphics and artificial intelligence.

CSR finds perceptually meaningful regions using an image segmentation technique and classify these regions based on image statistics. Consequently, it uses AI common-sense reasoning techniques to find ambiguities and anomalies within an image as well as performs reasoning across a corpus of images to identify a semantically based candidate list of potential fraudulent images. Our method introduces a novel framework for forensic reasoning, which allows detection of image tampering. This method works on an assumption of photo forgery performed on the following techniques:

- Deletion of details: removing scene elements.
- Insertion of details: adding scene elements.
- Photomontage: combining multiple images.
- False captioning: misrepresenting image content.

Copy-move forgery is a specific type of image tampering where a part of the image is copied and pasted on another part generally to conceal unwanted portions of the image. Hence, the goal in detection of copy-move forgeries is to detect image areas that are same or extremely similar. Several methods exist ***(Bayram at al., 2008)*** to achieve this goal. These methods in general use block-matching procedures, which first divide the image into overlapping blocks and extract features from each block, assuming similar blocks will yield similar features. Later, a matching step takes place where the aim is to find the duplicated blocks based on their feature vectors.

A forgery detection decision is made only if similar features are detected within the same distance of features associated to connected blocks. It examines several different block based features proposed for this purpose in relation to their time complexity and robustness to common processing scaling up or down, compression, and rotation.

The Digital Forgery in the images is possible to be exposed for JPEG and Bitmap files ***(Sunderrajan, 2009)***, as follows:

- Use of DCT coefficient quantization for JPG image.
- Sum of squared distances for BMP image.

The Copy-Paste forgery can be detected using the concept of BAG - Block Artifact Grid extraction ***(Li et al., 2008)***. The forensics approach is to locate the BAG firstly, and then check whether the BAG mismatches or not. Once a BAG mismatch is affirmed, then the image can be authenticated as doctored.

The Copy-move Image Forgery can be detected based on DWT-PCA ***(Zimba and Xingming, 2011)***. According to this criterion, an improved algorithm was proposed based on Discrete Wavelet Transform (DWT) and Principal Component Analysis Eigen value Decomposition (PCA-EVD) to detect such cloning forgery. Furthermore, for academic purposes and via a simplified, toy image we demonstrate how such algorithm works in detecting cloning forgery. Experimental results show that the proposed scheme accurately detects such specific image manipulations as long as the copied region is not rotated or scaled.

Similar to DART suggested earlier, specialized software RepFinder is useful for finding Approximately Repeated Scene Elements for Image Editing ***(Cheng et al., 2010)***. It proposes framework where user scribbles are used to guide detection and extraction of such repeated elements. This detection process, which is based on a boundary band method, robustly extracts the repetitions along with their deformations. The algorithm only considers the shape of the elements, and ignores similarity based on color, texture, etc. Subsequently, it uses topological sorting to establish a partial depth ordering of overlapping repeated instances. Missing parts on occluded instances are completed using information from other instances. The extracted repeated instances can then be seamlessly edited and manipulated for a variety of high level tasks that are otherwise difficult to perform. It demonstrates the versatility of framework on a large set of inputs of varying complexity, showing applications to image rearrangement, edit transfer, deformation propagation, and instance replacement.

```
1. Initialize B = {p}, the set of all boundary pixels;
2. Initialize queue Q = /0;
3. For each boundary pixel p ∈ B do
   mp = tangent vector of the boundary image, at p;
   push all neighbors q of p onto Q for which q ∈ B;
   end for
4. Initialize i = |B| * b,
   where |B| is the size of set B
   and b is a width parameter;
5. while i = 0 and Q = /0 do
   Pop p from Q;
   mp = 0, num = 0;
        for each already-initialized neighbor q of p do
             if |mp + mq| > |mp - mq| then
                  mp = mp + mq;
             else
                  mp = mp - mq;
             end if
        num = num + 1;
        end for
   mp = mp/num;
   Push each uninitialized neighbor of p
   in the band onto Q;
   i = i - 1;
   end while
6. For each remaining uninitialized pixel p,
   set mp = 0;
```

Algorithm 2.1: Building the BBM from a boundary image.

A comparison study on Copy-Cover Image Forgery Detection ***(Shih and Yuan, 2010)*** describes and compares the techniques of copy-cover image forgery detection. It is organized as follows. It suggests four copy-cover detection methods, including Principal Component Analysis (PCA), Discrete Cosine Transform (DCT), spatial domain, and statistical domain. It compares the four copy-cover detection methods, and provides the effectiveness and sensitivity under variant additive noises and lossy Joint Photographic Experts Group (JPEG) compressions.

The duplicate regions of a digital image can be identified by using SURF ***(Shivakumar and Baboo, 2011)***. The task of finding point correspondence between two images of an object or same scene is part of many computer vision applications. Recently Herbert Bay et al. proposed fast detectors and descriptors, called SURF - Speeded Up Robust Features. SURF's detector and descriptor is said to be faster and at same time robust to noise, detection displacements and geometric and photometric deformations.

With the existing DCT algorithms for copy-paste detection, the time complexity can be reduced ***(Khan and Kulkarni, 2010)***. The algorithm proposed for the same, is as mentioned in Figure-2.5.

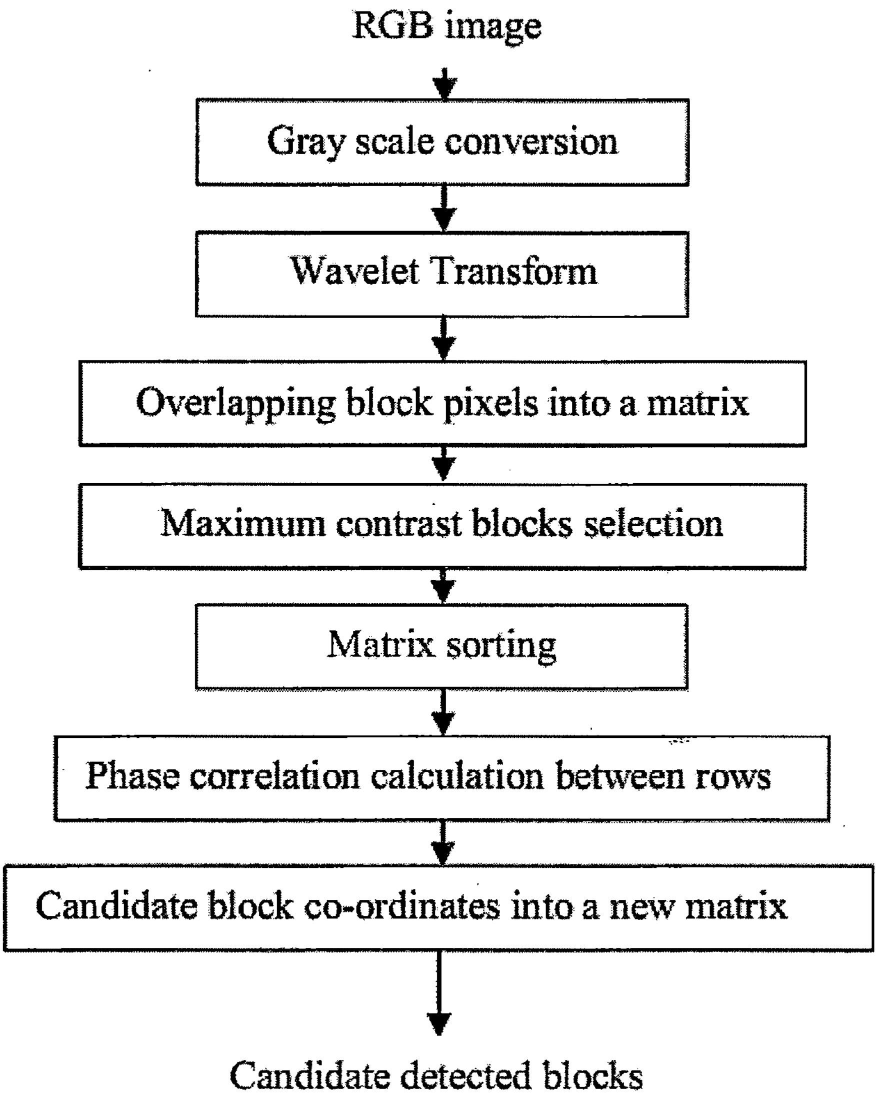

RGB image
Gray scale conversion
Wavelet Transform
Overlapping block pixels into a matrix
Maximum contrast blocks selection
Matrix sorting
Phase correlation calculation between rows
Candidate block co-ordinates into a new matrix
Candidate detected blocks

Figure 2.5: Algorithm to detect Reference and Match Blocks

Many image processing applications make use of MATLAB for analysis of images used by the system.

The image processing capabilities of MATLAB are very useful for all types of Image forgery detection. The photogrammatic mapping carried out using MATLAB tool, which can also be used in traffic surveillance ***(Madeira et al., 2010)***.

Open-source software including an easy-to-use graphical user interface (GUI) has been developed for processing, modeling and mapping of gravity and magnetic data. The program, called Potensoft ***(Arisoy and Dikmen, 2011)***, is a set of functions written in MATLAB. The most common application of Potensoft is spatial and frequency domain filtering of gravity and magnetic data. The GUI helps the user easily change all the required parameters. One of the major advantages of the program is to display the input and processed maps in a preview window, thereby allowing the user to track the results during the ongoing process. Source codes can be modified depending on the users' goals. It represents the main features of the program and its capabilities are demonstrated by means of illustrative examples.

One of such applications being geographical data is Seascorr ***(Meko et al., 2011)***. It is a MATLAB program for identifying the seasonal climate signal in an annual tree-ring time series, uses is identification of the monthly or seasonal climate signal in an annual time series of indices of ring width.

All the copy-paste detection methods discussed here are useful in particular case of forgery, and no common framework or widely used suggested method is available.

Chapter 3

METHODOLOGY

- ⊗ Copy-Paste region detection
- ⊗ Suspected Partition detection

CHAPTER-3: METHODOLOGY

3.1 Copy-Paste Detection

The prime objective of the study was to device a program or software, which takes as an input the digital image file, assumed to be forged, and determines or spots the areas of forgery from it.

The modern familiar programming techniques used to perform different operations on image files are:

1. Java Development kit (JDK 1.2)
2. MATLAB® 7.5.0.342 (R2007b)

The basic programming operations required on a given suspected image file available in JPEG or BMP or PNG format are as follows:

a. Recording each pixel value of an image in form of an array data structure.

b. Comparison of every picture element (pixel) with all other elements, in order to determine the similarity between the set of pixels, if exists.

c. Determining the statistics of similar pixels found in the same image.

d. Identifying the extent of image forgery, and spot out the areas of forgery within an image.

In order to test the methodology suggested, 10 sample images with different details of scenes are selected. These sample images act as an input to the program, whereas the execution of program resembles the processing part. In a whole, the following notion is used to solve the problem of image forgery detection.

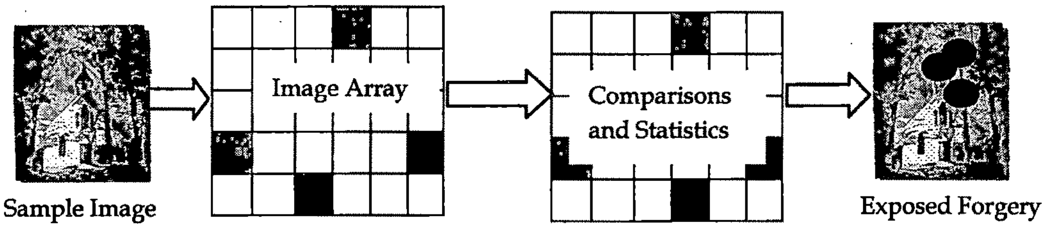


Figure 3.1: Notion applied towards Copy-Paste region Detection.

Based on the above requisite, the program using Java was developed as below:

```
import java.awt.*;
import java.applet.*;
import java.awt.image.*;
public class forgtest extends Applet
{
Image ori, nw; // ori is test image, nw is resultant image.
int pxl[] = new int [15000]; // 100x150 pixel image.
PixelGrabber p=new PixelGrabber (ori,0,0,300,50,pxi,0,300);
     // Image is stored into array
     try { p.grabPixels(); }
     catch( InterruptedException e) { }
     for (int i=0; i<15000; i++)  // Compare every value.
     {    int p = pxl[i];
          for (int j=0; j<15000; j++)
          {    int q = pxl[j]; // Check if they are same.
               if (q==p) pxl[i] = 0;  // Make BLACK.
     }
nw = createImage (new MemoryImageSource(300,100,pxl,0,300);
} // Applet designed.
     public void paint (Graphics g)
     {    g.drawImage( ori, 10, 10, this);  // Original.
          g.drawImage(nw, 400, 10, this); // Result.
     }
```

However, by using Java source code mentioned above, the image forgery was not identified as per the desired results.

Hence, a program using MATLAB 7.2 was devised, considering the below mentioned advantages of MATLAB:

Recording of the processing used
MATLAB is a general purpose programming language. When it is used to process images one generally writes function files, or script files to perform the operations. These files form a formal record of the processing used and ensures that the final results can be tested and replicated by others should the need arise.

Access to implementation details
MATLAB provides many functions for image processing and other tasks. Most of these functions are written in the MATLAB language and are publicly readable as plain text files. Thus the implementation details of these functions are accessible and open to scrutiny. The defense can examine the processing used in

complete detail, and any challenges raised can be responded to in an informed way by the prosecution. This makes MATLAB very different from applications, such as Photoshop.

Some MATLAB functions cannot be viewed. These are generally lower level functions that are computationally expensive and are hence provided as 'built-in' functions running as native code. These functions are heavily used and tested and can be relied on with considerable confidence.

Numerical accuracy

Another advantage of MATLAB is that it allows one to ensure maximal numerical precision in the final result.

In general, image files store data to 8 bit precision. This corresponds to a range of integer values from 0-255. A pixel in a colour image may be represented by three 8 bit numbers, each representing the red, green and blue components as an integer value between 0 and 255. Typically this is ample precision for representing normal images.

However as soon as one reads this image data into memory and starts to process it is very easy to generate values that lie outside the range 0-255. For example, to double the contrast of an image one multiplies the intensity values by 2. An image value of 200 will become 400 and numerical overflow will result. How this is dealt with will vary between image processing programs. Some may truncate the results to an integer in the range 0-255, others may perform the mathematical operations in floating point arithmetic and then rescale the final results to an integer in the range 0-255.

It is here that numerical precision, and hence image fidelity, may be lost. Some image processing algorithms result in some pixel values with very large magnitudes (positive or negative). Typically these large values occur at points in the image where intensity discontinuities occur, the edges of the image are common sources of this problem. When this image with widely varying values is rescaled to integers in the range 0-255 much of this range may be used just to represent the few pixels with the large values. The bulk of the image data may then have to be represented within a small range of integer values, say from 0-50. Clearly this represents a considerable loss of image information. If another process is then applied to this image the problems can then accumulate. Trying to establish the extent of this problem, if any, is hard if one is using proprietary software.

Being a general programming language it is possible to have complete control of the precision with which one represents data in MATLAB. An image can be read into memory and the data cast into double precision floating point values. All image processing steps can then be performed in double precision floating point arithmetic, and at no intermediate stage does one need to rescale the results to integers in the range 0-255. Only at the final point when the image is to be displayed and/or written to file does it need to be rescaled. Here one can use histogram truncation to eliminate extreme pixel values so that the bulk of the image data is properly represented.

Advanced algorithms

MATLAB is a scientific programming language and provides strong mathematical and numerical support for the implementation of advanced algorithms. It is for this reason that MATLAB is widely used by the image processing and computer vision community. New algorithms are very likely to be implemented first in MATLAB, indeed they may only be available in MATLAB.

The following algorithm is applied for detection of Copy-Paste regions in an image:

Algorithm 3.1: Detection of Copy-Paste regions

```
Step 1.   Initialize the following parameters of an image.
□         I: Image array of dimension M x N comprising of color value.
          P: An M x N Boolean value array indicating whether the positioned
             pixel comparison is completed or not.
          x and y: Current pixel position coordinates.
          a and b: Likely to be similar pixel coordinates.
          c: Similarity color coefficient, set to zero in case of same color.
Step 2.   Initialize $P_{x,y}$ = false,
          indicating the comparison operation not performed.
Step 3.   Process Step-4 and Step-5 for x, a = 1 to M and y, b = 1 to N
Step 4.   If ( $I_{x,y} = I_{a,b} \pm c$) and ($I_{x+1,y} = I_{a+1,b} \pm c$) and ($P_{x,y}$ = true)
          then
          $I_{a,b}$ = BLACK to set the detected similar pixels with desired color.
Step 5.   Set $P_{x,y}$ = true, indicating that comparison is over.
```

The code developed to implement this algorithm, for Copy-Paste detection in an image using MATLAB is as shown below:

Program 3.1: Implementation of algorithm 3.1 in MATLAB

```
% Identification of Copy-Paste Area in BMP image

i=imread('D:\RESEARCH\PhD\t1.png','png');
% Save Bitmap 24 bit format in Array I
z=zeros(220,220);
% To check that positioned pixel is already processed or
not
for x=1:95
     for y=1:40     % Compare each pixel with every other.
          for a=x:95
               for b=1:40
     if (z(a,b)==0)  % If not processed then ..
     if ( (i(x,y,1)==i(a,b,1)) &&(i(x,y,2)==i(a,b,2))
     &&(i(x,y,3)==i(a,b,3))&&(~ ((x==a)&&(y==b))))

% Checked that - (i) Pixel value of adjacent is same
%              - (ii) Its not compared with itself
     i(x,y,1)=0;  % Make all 3 values 0 ,
     i(x,y,2)=0;
     % Indicating that same pixels are DETECTED and Black.
     i(x,y,3)=0;
     %i(a,b,1)=100;
     %i(a,b,2)=100; % i(a,b) represents another set of
copy.
     %i(a,b,3)=100;
     end % of inner if
     end % of outer if
     end % of var. b
     end % of var. a
z(x,y)=1; % Checking is over for selected pixel.
end % of var.y
end % of var.x
imshow(i)  % Display detected portions, with black.
```

3.2 Suspected Partition Detection

In an experiment to detect the suspected partitions of an image, an algorithm is devised to spot such forgery or manipulated areas. The logical sequence for implementation of an idea is exhibited in figure-3.2.

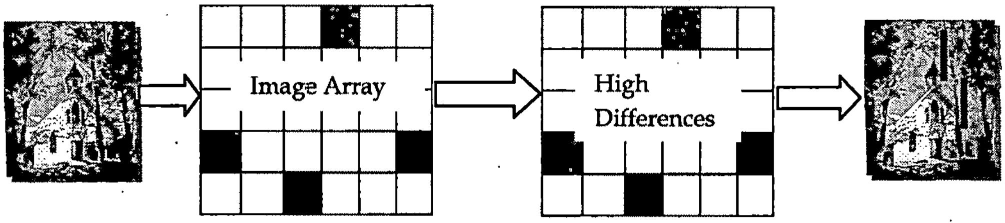
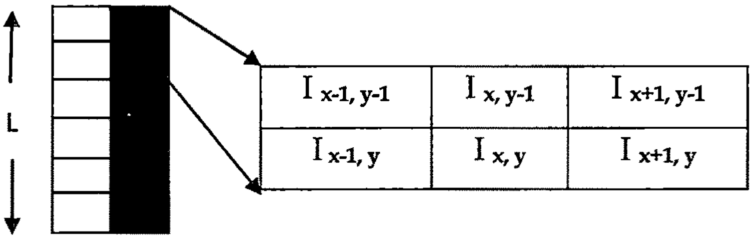


Figure 3.2: Notion used in suspected partition detection.

When the copy-paste operation is assumed not to be performed in the same image, we can assume that region can be suspected to be from some other image. In such cases, there are chances of the traces at the position at which the paste operation is carried out. Such locations can be easily detected, since there are high changes in neighboring pixel values, compared to the other positions in an image. The changes can be detected in horizontal (as in figure-3.3) or vertical axis.



$I_{x-1,y-1}$	$I_{x,y-1}$	$I_{x+1,y-1}$
$I_{x-1,y}$	$I_{x,y}$	$I_{x+1,y}$

Figure 3.3: Pixel positions in an image.

When a portion of image is copied from another image, we can assume to identify a vertical or horizontal line of reference, of length L, where all neighboring pixels carry high amount of difference in color value i.e. $I_{x,y} - I_{x+1,y} = F$. Based on this assumption, algorithm 3.2 is devised.

Algorithm 3.2: Detection of Suspected Partitions.

```
Step 1.   Let a and b are height and width of an RGB image,
          I is the image pixel value representation matrix of size a x b.
          F is the determining factor for the pixel difference.
Step 2.   Determine the difference between two pixels.
          For each value of i = 1 to a, and j = 1 to b
          Check for the following conditions to be evaluated to true:
                a) $abs ( I_{a,b} - I_{a,b+1} ) > F$
                b) $abs ( I_{a,b} - I_{a,b-1} ) < F$
                c) $abs ( I_{a-1,b} - I_{a-1,b+1} ) > F$
                d) $abs ( I_{a-1,b-1} - I_{a-1,b} ) < F$
Step 3.   If all above four conditions (a, b, c and d) are satisfied, it indicates
          that two pixels along the same axis contain the split color values
          (highly different) or assumed to be forgery prone.
Step 4.   Considering the required level of detection L being no. of columns
          or height of forgery, Set partition line to indicate black color.
          $I_{a,b} = 0$, where x = a - (L/2) to a + (L/2).
```

The algorithm 3.2 is implemented using equivalent MATLAB code, as below:

Program 3.2: Implementation of algorithm 3.2 in MATLAB

```
% Detection of a major variation in image row/column due to
% forgery.

% Considering im2 as merged image.
% Find difference between pixel values of same image.

i1 = imread('d:\sample1.bmp','bmp');
x = 1;
for a = 2:109
  for b = 2:109
    % Next Column is different
    nc1 = abs(i1(a,b,1) - i1(a,b+1,1));
    nc2 = abs(i1(a,b,2) - i1(a,b+1,2));
    nc3 = abs(i1(a,b,3) - i1(a,b+1,3));
    % Previous column is similar
    pc1 = abs(i1(a,b,1) - i1(a,b-1,1));
    pc2 = abs(i1(a,b,2) - i1(a,b-1,2));
    pc3 = abs(i1(a,b,3) - i1(a,b-1,3));
```

```
      % Above next is different
     an1 = abs(i1(a-1,b,1) - i1(a-1,b+1,1));
     an2 = abs(i1(a-1,b,2) - i1(a-1,b+1,2));
     an3 = abs(i1(a-1,b,3) - i1(a-1,b+1,3));
     % Above Previous is similar
     ap1 = abs(i1(a-1,b-1,1) - i1(a-1,b,1));
     ap2 = abs(i1(a-1,b-1,2) - i1(a-1,b,2));
     ap3 = abs(i1(a-1,b-1,3) - i1(a-1,b,3));

if (nc1>5||nc2>5||nc3>5)&&(an1>5||an2>5||an3>5)&&
   (pc1<5 && pc2<5 && pc3<5 && ap1<5 && ap2<5 && ap3<5 )
        i1(a,b,1)=0;i1(a-1,b,1)=0;  %Pixel and its above.
        i1(a,b,2)=0; i1(a-1,b,2)=0;
        i1(a,b,3)=0; i1(a-1,b,3)=0;
    end
  end
end

i2 = imread('Merge.bmp','bmp');
% Check for 7 consecutive occurrences, for L = 7
for a = 4:109
   for b= 1:109
     if (i1(a,b)==0 && i1(a-2,b)==0 && i1(a-1,b)==0 &&
i1(a+1,b)==0 && i1(a+2,b)==0 && i1(a-3,b)==0 &&
i1(a+3,b)==0)
        i2(a-3,b,1)=0; i2(a-3,b,2)=0;i2(a-3,b,3)=0;
        i2(a-2,b,1)=0;i2(a-2,b,2)=0;i2(a-2,b,3)=0;
        i2(a-1,b,1)=0;i2(a-1,b,2)=0;i2(a-1,b,3)=0;
        i2(a,b,1)=0;i2(a,b,2)=0;i2(a,b,3)=0; % Set to BLACK
        i2(a+1,b,1)=0; i2(a+1,b,2)=0; i2(a+1,b,3)=0;
        i2(a+2,b,1)=0;i2(a+2,b,2)=0;i2(a+2,b,3)=0;
        i2(a+3,b,1)=0;i2(a+3,b,2)=0;i2(a+3,b,3)=0;
     end % of IF statement
   end  % Inner loop
end  % Outer loop
imshow(i2);
```

Chapter 4

RESULTS AND DISCUSSION

- ⊗ Copy-Paste region detection
- ⊗ Suspected Partition detection
- ⊗ Benefits

CHAPTER-4: Results and Discussion

4.1 Copy-Paste detection

The detection method devised in MATLAB, to detect the Copy-Paste Regions (CPR) within the same image, fetches an image input and results into the same image, with BLACK pixel values representing the area of forgery or CPR.

The table-4.1 represents 3 sample BMP images and their data, used to verify the results.

Table 4.1: Results of 3 samples BMP files.

Image name	*Colors*	*CPR Shape*	*False positives*
Sample1.bmp	Uniform	Rectangle	730
Sample2.bmp	Normal	Polygon	103
Sample3.bmp	Random	Complex	12

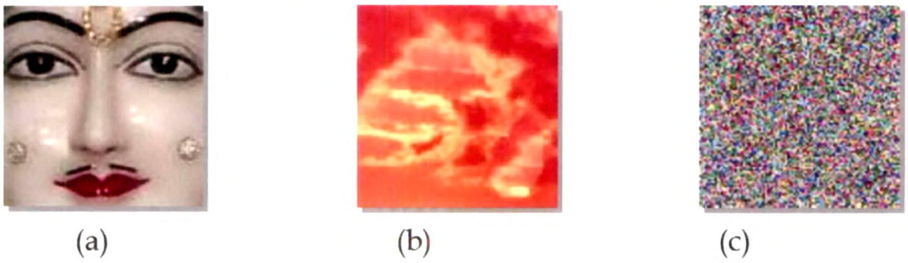


Figure 4.1: (a) Uniform (b) Normal and (c) Random - color value images.

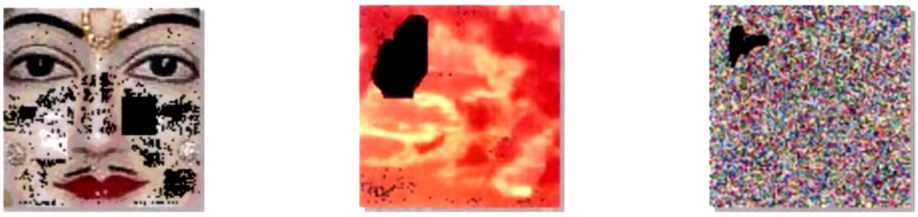


Figure 4.2: CPR detections for images (a), (b) and (c) respectively.

In addition to the above three cases, other sample BMP images were taken as an input, with different sizes and types of data, and different shapes of CPR were created, to verify the detection of the same. The database comprises of 10 such sample files, along with their results shown in table-4.2.

Table-4.2: Results of the image sample database.

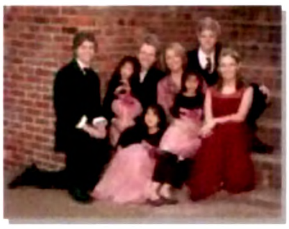
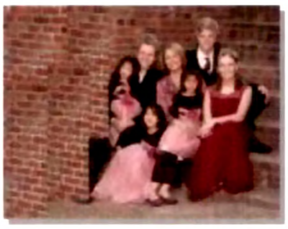
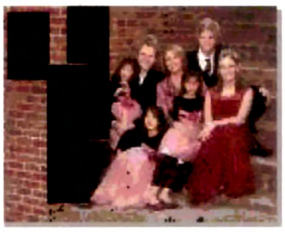
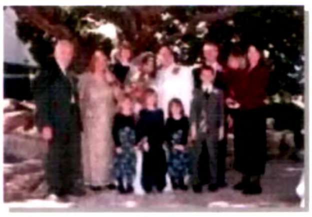
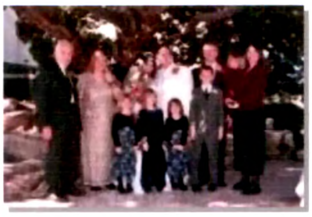
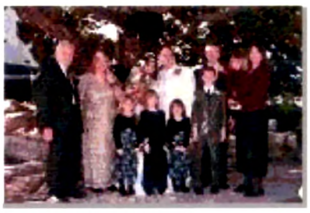
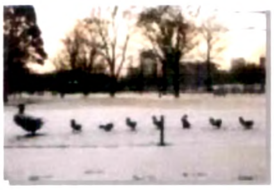
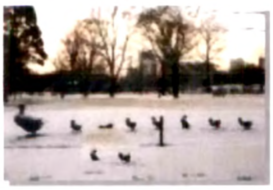
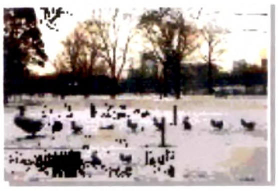
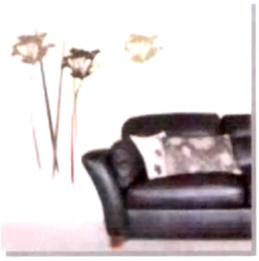
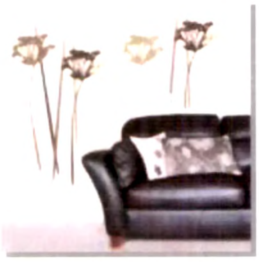
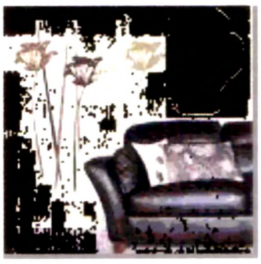
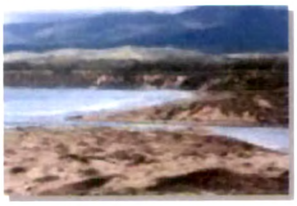
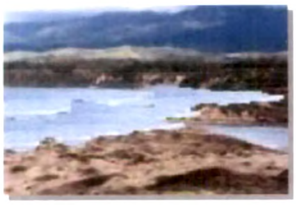
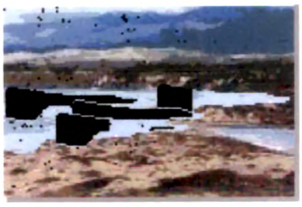
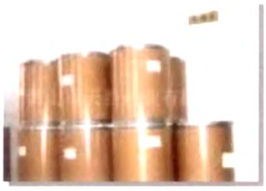
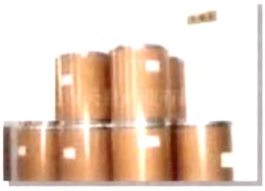
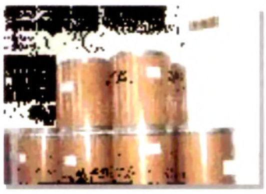
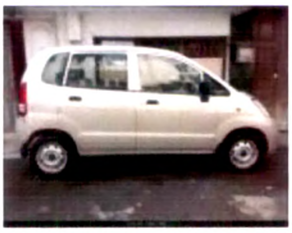
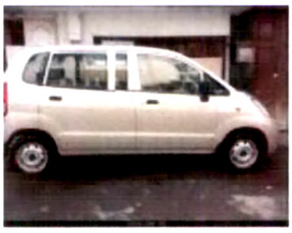
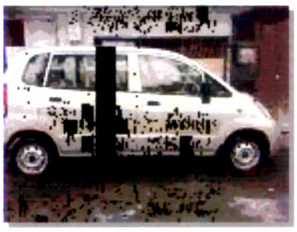
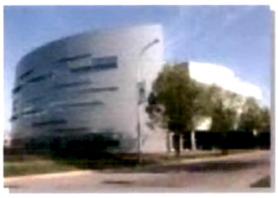
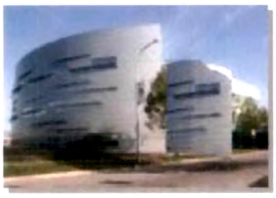
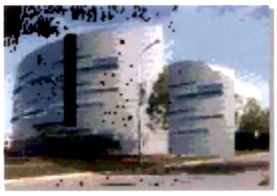
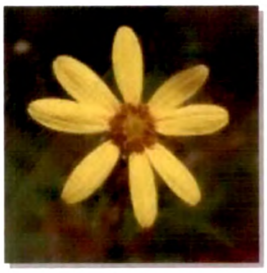
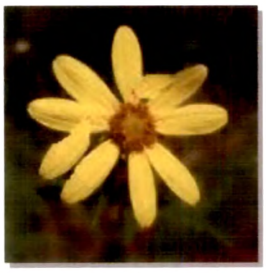
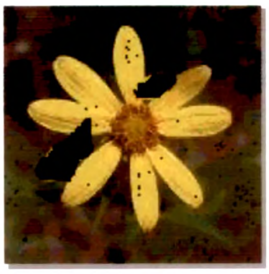
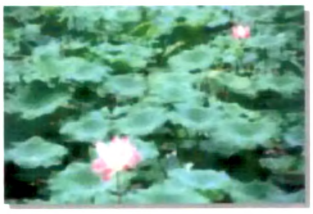
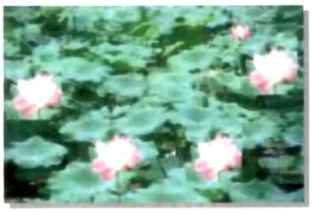
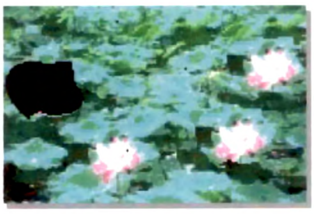
Image	*Original image*	*CPR formation*	*CPR Detection*
s1			
s2			
s3			
s4			
s5			

Table-4.2: Results of the image sample database.

Image	*Original image*	*CPR formation*	*CPR Detection*
s6			
s7			
s8			
s9			
s10			

Based on the above mentioned results, the extra BLACK pixels were also identified, which are not considered to be the part of CPRs. These values, referred as false positives (Table-4.3) for our results, are relatively very few.

Table 4.3: False positive results.

Name	*Height (px)*	*Width (px)*	*Total Pixels*	*False Positives*	*Accuracy*
s1	139	106	14734	48	99.67
s2	150	100	15000	57	99.62
s3	130	87	11310	300	97.35
s4	124	124	15376	1060	93.11
s5	143	95	13585	203	98.51
s6	127	91	11557	1417	87.74
s7	143	110	15730	528	96.64
s8	133	90	11970	519	95.66
s9	127	127	16129	511	96.83
s10	150	97	14550	15	99.90

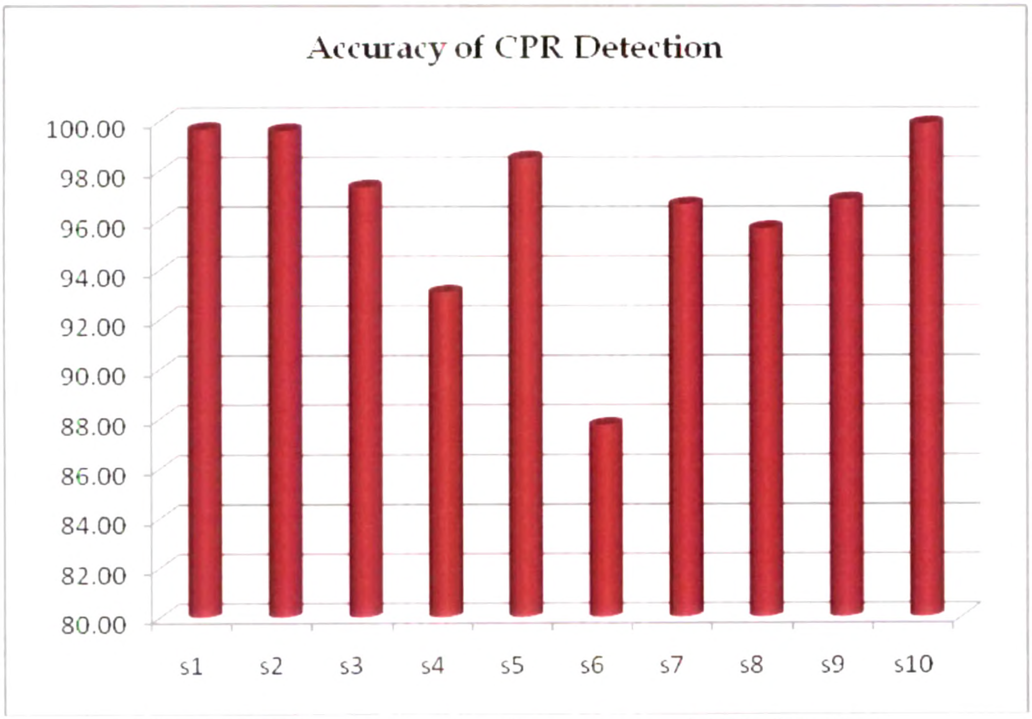


Figure 4.3: Graph showing the accuracy of CPR detection.

Thus, we can conclude from fig.-4.3 that the accuracy of CPRs found into BMP images range from 87.74 % (worst case) to 99.9 % (best case).

The similar results are also achieved for PNG file format.

4.2 Suspected Partition Detection (SPD)

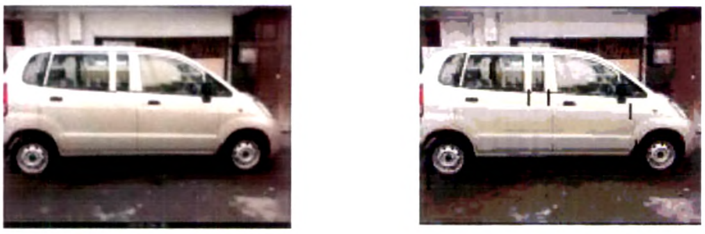
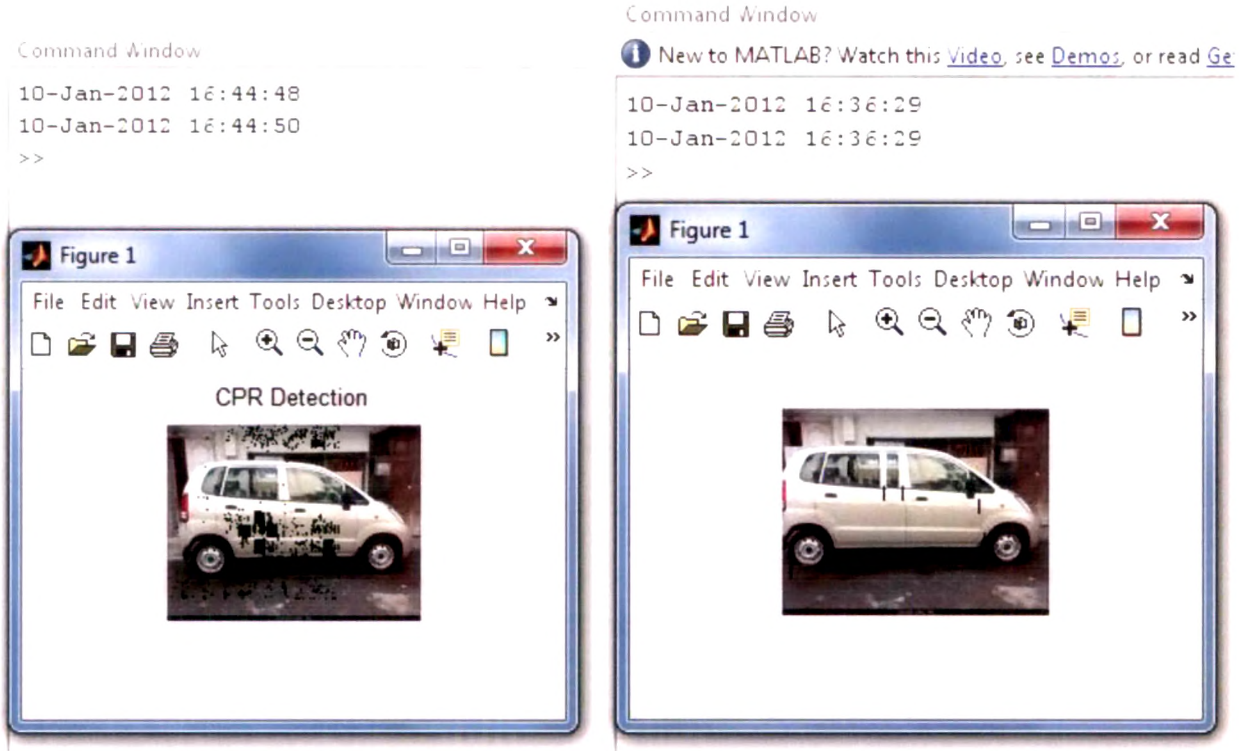


Figure 4.4: (a) Forged image (b) Partition detection

The suspected partitions are clearly visible in form of vertical BLACK lines near the door, which confirms that the spotted axis is pasted from another image.

While the CPR detection takes 2 to 5 seconds' time, these results are achieved in less than 1 second, which is the same image used in table-4.2.



Command Window

```
10-Jan-2012 16:44:48
10-Jan-2012 16:44:50
>>
```

Command Window

New to MATLAB? Watch this Video, see Demos, or read Ge

```
10-Jan-2012 16:36:29
10-Jan-2012 16:36:29
>>
```

Figure 1
File Edit View Insert Tools Desktop Window Help
CPR Detection
Figure 1
File Edit View Insert Tools Desktop Window Help

Figure 4.5 : (a) Timing of 2 sec for CPR (b) Timing of less than 1 sec for SPD

4.3 Benefits

With the results achieved in CPR and SPD, the following benefits of these methods are identified:

1. Since both the approaches are using straight-forward pixel checking, they act as a first step towards forgery detection for the suspected images.
2. The results are successful for BMP as well as PNG file format, considering the fact that most of the press and media publication impart compulsion on PNG format, rather than any other lossy formats like JPG.
3. With the availability of other methods for JPG or BMP forgery detection, these methods can supplement to the existing programs or algorithms.
4. A drastic reduction of false positives is the most important criteria of the programs designed.

Chapter 5

FUTURE SCOPE

Future Scope

Forensic investigation is a complicated science with its own history, implications and future. It is not sufficient merely to consider it a branch of criminology, or the study of cyber criminal behavior, or research into the relationship between the causes of tech related crime and social policies. For cyber criminals, their knowledge and their crimes are bound together. The possible suspects are rich in knowledge and technical skills. They have mastered the technology better than the technology's creators, and they know how to use technology against technology.

A multidisciplinary approach is required to fully foresee the future of forensics. It requires a team of specialists from different disciplines within the IT industry and related industrial and social segments such as telecom and law. However, in this article the author looks at the future of forensics based on his knowledge and experience in this field.

Forensics for Governments

Forensics at the governmental level will be more complicated in the future. Governments will need to turn more to their national security organisations to hunt down cyber criminals. In addition, they will need to invent anti-forensic tools and methods to keep their activities and information assets secret.

Cyberspace security and computer related technologies will be a real challenge for governments. The platforms and protocols for computer related technologies may have both domestic and international uses. Therefore, it will be difficult for governments to reach an agreement for international cyber security policies.

At the same time, some countries are the technology owners and this intellectual property ownership will give them an advantage compared to other countries without such a privilege. The technology ownership issue will force the other countries to utilise the open source platforms to develop their own customised operating systems and software.

Forensics for Corporates

Currently a few companies have dominated the forensic markets. These are the pioneers in forensics and analysis. They have the tools and the solutions for cyber forensic investigation. They train law enforcement agencies to use their tools and solutions and some of them even have special tools just for governmental use.

There are also many small companies with one or two consultant partners who are either retired law enforcement officers or former IT professionals from Fortune 500 companies. These people use their contacts and credentials to achieve some market share. However, in the future, forensics at the corporate level will be diversified to education and certain specialties and products. It will be difficult for small companies to build a team with the right core competencies. In addition, due to security clearance requirements and national security interests, most of these companies will only practice in their country of origin.

Furthermore, information security standards such as ISO27001 and ITIL will be implemented more in medium to enterprise size companies. Realistically, only these companies can afford the cost of compliance implementation. Therefore, it will be necessary for them to have proper incident response procedures and the corresponding forensic investigation capabilities. These companies may well have their own forensic investigation units.

Forensics in Professional Institutions

Forensics is a new battleground for professional institutions. Currently, there is no real internationally recognised authority to govern forensics practices, regulations and certification. Therefore, professional institutions are offering forensic investigation training programs, certifications and conferences. Currently, some of these institutions are forming alliances (as trade and training partners) to achieve their sales targets. In the future, it is likely that these institutions will start to attack each other to gain market share.

Forensics in Universities

It is sad to note that more and more often information technology advances are coming from industry rather than universities. Within IT, a few companies dominate the industry and therefore the innovations. It will be the same for forensics; the companies with market share have the money for research and development. The main issue with academic institutions is their approach, which is slow and traditional compared to the faster speed of development and implementation found in industry.

Furthermore, the training programs in universities are not aligned with the current job market and industry needs. The university students have a lack of practical knowledge compared to the IT professionals who are in the industry (and possibly without academic studies). This is the major reason why students choose further training to achieve professional certification and so distinguish themselves from other graduates.

Forensics in the Media

There will be more magazines, websites and blogs specialising in forensics and analysis. They will be the voice of the industry with the power to review, promote and criticise books, products, solutions and training programs. They will sell advertising and help vendors sell their products. Whoever has more marketing budget and better relations will be the most successful in the forensics industry. Nevertheless, there will be one or two magazines and websites that will remain independent, but they will find it difficult to survive in such a tough market.

Forensics and Technical Trends

The market will be divided to four main segments with specialised service providers for each segment. The segments are: Microsoft Windows related products, UNIX & Linux related products, Apple related products and computer network & telecom related products.

The solution providers will create more comprehensive tools and solutions to gain better market share. The technology will transform their solutions into a set of tools for non-IT professionals. It will also try to make their tools web based, for remote forensic investigations.

The open source community will be active for the UNIX & Linux platforms to accrue required legislation to accredit the open source tools in the various countries and judicial systems.

Apple created a giant market for those who want to develop Apple device related tools and solutions. This will be a new era for the professionals who are working in forensics.

Cloud computing, cellular networks, WiMax and virtualization will be the other areas of the interest for study and product development. It is obvious that everything is merging towards IT and cyberspace plays an important role in the near future. This will lead governments and authorities to pursue other methods of intelligence gathering, such as web and data mining, to protect their interests.

This will lead to the biggest privacy issue in history. All the data communication, of all users, will be logged at the carrier level. Then the authorities will use data mining tools to identify suspicious behavior of a particular user or users in their own or an allies' territory. All this information will be saved in massive databases and then the commercial, financial and personal information, in addition to the communication records and social behaviors, will be linked together.

All this will ultimately lead to a new chapter in the history of forensics, namely Applied Artificial Intelligence in Forensics.

The scope for CPR and SPD

With reference to the detection methods of study, there is a scope in following areas:

1. JPG Detection: Though JPG is lossy method, the algorithm can be modified to obtain the similar detections is JPG files as that in BMP files.
2. Optimization: The images used in the research work are of around 100x100 pixels, consuming about 2 to 4 second for detections, due to pixel comparisons in terms of thousands. The programs can be modified to eliminate unneeded comparisons.
3. Platform independence: The system requirement for the methods used here comprises of Windows OS based MATLAB. The common program such as Java can be useful in order to avoid specific OS requirement.

REFERENCES

REFERENCES

Arısoy, M.O. and Dikmen, U. (2011) Potensoft: MATLAB-based software for potential field data processing, modeling and mapping. *Computers & Geosciences* 37: 935-942.

Armengaud, P., Zambaux, K., Hills, A., Sulpice, R., Pattison, R. J., Blatt, M. R., Amtmann, A. (2008) EZ-RHIZO: integrated software for the fast and accurate measurement of root system architecture. *Plant Journal* 57: 945–956.

Avcıbaş, I., Bayram, S., Memon, N., Sankur, B. and Ramkumar M. (2004) A Classifier Design For Detecting Image Manipulations. *ICIP 2004* 4: 2645-2648.

Ayoub, F., Leprince, S. and Avouac, J. P. (2009) Co-registration and correlation of aerial photographs for ground deformation measurements. *ISPRS Journal of Photogrammetry and Remote Sensing* 64: 551-560.

Bayram, S., Avcıbas, I., Sankur, B. and Memon N. (2005) Image Manipulation Detection With Binary Similarity Measures. *EUSIPCO 2005.*

Bayram, S., Avcıbas, I., Sankur, B. and Memon N. (2006) Image manipulation detection. *Journal of Electronic Imaging* 15(4): 411002-411017.

Bayram, S., Sencar, H. T. and Memon, N. (2008) A Survey of Copy-Move Forgery Detection Techniques. *IEEE Workshop.*

Bertolini, D., Oliveira, L. S., Justino, E. and Sabourin, R. (2010) Reducing forgeries in writer-independent off-line signature verification through ensemble of classifiers. *Pattern Recognition* 43: 387-396.

Biswas, T. K. (2011) Data and information theft in e-commerce, jurisdictional challenges, related issues and response of Indian laws. *Computer Law & Security Review* 27: 385-393.

Bolton, R. J. and Hand, D. J. (2002) Statistical fraud detection: A review. *Statistical Science* 17(3): 235–255.

Bose, P., Dujmovic, V., Hurtado, F. and Morin P. (2011) Connectivity-preserving transformations of binary images. *Computer Vision and Image Understanding* doi:10.1016/j.cviu.2007.06.003

Bot, J. L., Serra, V., Fabre, J., Draye, X. and Pagès, S. A. (2010) DART: a software to analyse root system architecture and development from captured images. *Plant Soil* 326: 261-273.

Brox, T., Farin, D. and de With, P. H. N. (2001) Multi-stage region merging for image segmentation. In: *Proceedings of the 22nd Symposium on Information Theory in the Benelux.*

Buades, A., Coll, B., Morel, J. M. (2005) A non-local algorithm for image denoising. *CVPR* 2: 60-65.

Cao, G., Zhao, Y. and Ni, R. (2010) Edge-based Blur Metric for Tamper Detection. *Journal of Information Hiding and Multimedia Signal Processing* 1(1): 20-27.

Carrier, B. D., Spafford, E. H. (2004) An event-based digital forensic investigation framework. *Digital forensics research workshop.*

Carvey, H. (2005) The Windows registry as a forensic resource. *Digital Investigation* 2(3): 201–205.

Cha, S. H. and Srihari, S. N. (2002) On measuring the distance between histograms. *Pattern Recognition* 35: 1355–1370.

Chang, C. C., Kieu, T. D., Wu, W.C. (2009) A lossless data embedding technique by joint neighboring coding. *Pattern Recognition* 42: 1597–1603.

Cheddad, A., Condell, J., Curran, K. and McKevitt P. (2009) A Secure and Improved Self-Embedding Algorithm to Combat Digital Document Forgery. *Signal Processing* 89(12): 2324-2332.

Cheng, M. M., Zhang, F. L., Mitra, N. J., Huang, X. and Hu, S. (2010) RepFinder: Finding Approximately Repeated Scene Elements for Image Editing. *ACM Proceedings* 29(4): 1-8.

Chung, K. L., Yang, W. N., Huang, Y. H., Wu, S.T. and Hsu, Y. C. (2007) On SVD-based watermarking algorithm. *Appl. Math. Comput.* 188(1):54-57.

Cohen, A., Daubechies, I., Jawerth, B., Vial, P. (1993) Multiresolution analysis, wavelets and fast algorithms on an interval. *Comptes Rendus Acad. Sci. Paris* 417-421

Cohen, M. (2007) Advanced carving techniques. *Digital Investigation* 4(3–4): 119–128.

Cohen, M. (2008) PYFLAG: an advanced network forensic framework. In: *Proceedings of the 2008 digital forensics research workshop.*

Cohen, M., Garfinkel, S. and Schatz, B. (2009) Extending the advanced forensic format to accommodate multiple data sources, logical evidence, arbitrary information and forensic workflow. *Digital Investigation* 6: S57-S68.

Coifman, R. R. and Donoho, D. L. (1998) Translation-Invariant De-Noising. *Signal Processing* 3414-3420.

Datta, R., Joshi, D., Li, J. and Wang, J. Z. (2008) Image Retrieval: Ideas, Influences, and Trends of the New Age. *ACM Computing Surveys* 40(2): 1-60.

Deguillaume, F., Voloshynovskiy, S. and Pun, T. (2003) Secure hybrid robust watermarking resistant against tampering and copy attack. *Signal Processing* 83(10): 2122-2170

Dell'Endice, F., Nieke, J., Koetz, B., Schaepman, M. E. and Itten K. (2011) Improving radiometry of imaging spectrometers by using programmable spectral regions of interest. *ISPRS Journal of Photogrammetry and Remote Sensing* doi:10.1016/j.isprsjprs.2009.05.007

Dirik, A. E. and Memon, N. (2009) Image Tamper Detection Based on Demosaicing Artifacts. *Image Processing* 1497-1500.

Dong, W., Li, X., Zhang, L. L. and Shi, G. (2011) Sparsity-based Image Denoising via Dictionary Learning and Structural Clustering. *Computer Vision and Pattern Recognition* 457-464.

Donoho, D. and Johnstone, I. (1994) Ideal spatial adaptation by wavelet shrinkage. *Biometrika* 3: 425-455.

Du, W., Sykes, L., Shaw, B., Scholz, C. (2003) Triggered aseismic fault slip from nearby earthquakes, static or dynamic effect? *Journal of Geophysical Research* - Solid Earth 108 (B2): 2131.

Eggerton, J. D. and Srinath, M. D. (1986) Statistical distribution of image DCT coefficients. *Computer and Electrical Engineering* 12: 137-145

El-Fegh, I., Mustafa, D. and Zub, Z. S. (2009) Color image watermarking based on the DCT-domain of three RGB color channels. In: *Proceedings of the 10th WSEAS international conference on evolutionary computing* 36-39.

Enser, P. G. B. (1995) Pictorial information retrieval. *Journal of Documentation* 51(2): 126–170.

Farid H. (2009) Image Forgery Detection - A survey. *IEEE Signal Processing Magazine* 16-25.

Farid, H. and Lyu, S. (2003) Higher-order wavelet statistics and their application to digital forensics. *IEEE Workshop on Statistical Analysis in Computer Vision.*

Fawcett, T. and Provost, F. (2002) Fraud detection. *Handbook of Knowledge Discovery and Data Mining* 726–731.

Fridrich, J. (1998a) Image Watermarking for Tamper Detection. In: *Proceedings of ICIP.*

Fridrich, J. (1998b) Methods for Detecting Changes in Digital images. *ISPACS.*

Fridrich, J. (1999) Methods for Tamper Detection in Digital Images. *ACM Workshop on Multimedia and Security* 19-23.

Fridrich, J. and Goljan, M. (1999) Images with Self-Correcting Capabilities. *Image Processing* 3: 792-796.

Fridrich, J., Soukal, D. and Lukas, J. (2003) Detection of Copy-Move Forgery in Digital Images. *DFRWS 2003.*

Fu, D., Shi, Y. Q. and Su, W. (2007) A generalized Benford's law for JPEG coefficients and its applications in image forensics. *SPIE Proceedings* 6505: L1-11.

Garfinkel, S. L., Paul, F., Vassil, R. and George D. (2009) Bringing science to digital forensics with standardized forensic corpora. Digital *Forensic Research Workshop (DFRWS).*

Garfinkel, S., Nelson, A., White, D. and Roussev, V. (2010) Using purpose-built functions and block hashes to enable small block and sub-file forensics. *Digital Investigation* 7: S13-S23.

Giffin, J., Greenstadt, R., Litwack, P. and Tibbetts, R. (2002) Covert messaging through TCP timestamps. *Workshop on privacy enhancing technologies* 194-208.

Gribble, S. D., Singh, M. G., Drew, R., Brewer, E. A., Gibson, T. J., Miller, T. L. (1998) Self-similarity in file Systems. *SIGMETRICS Perform. Eval Rev* 26(1): 141-150.

Grier, J. (2011) Detecting data theft using stochastic forensics. *Digital Investigation* 8: 71-77.

Hanbury, A., Müller, H. and Clough P. (2010) Special issue on image and video retrieval evaluation. *Computer Vision and Image Understanding* 114: 409-410.

Hanmandlua, M., Yusofb, M. and Madasuc, V. K. (2005) Off-line signature verification and forgery detection using fuzzy modeling. *Pattern Recognition* 38(3): 341-356.

He, J., Lin, Z., Wang, L. and Tang, X. (2006) Detecting Doctored JPEG Images Via DCT Coefficient Analysis. *ECCV 2006* 3: 423-435.

Hiewa, B. Y., Teoh, A. B. J. and Yin, O. S. (2010) A secure digital camera based fingerprint verification system. *Journal of Visual Communication and Image Research* 21: 219-231.

Hodge, V. and Austin, J. (2004) A survey of outlier detection methodologies. *Artificial Intelligence Review* 22(2): 85–126.

Hu, Q., Yu, D. and Xie, Z. (2008) Neighborhood classifiers. *Expert Systems with Applications* 34: 866–876.

Hunt, R. and Qi, Y. (1995) A multi-resolution approach to computer verification of handwritten system. *IEEE Transactions on Image Processing* 4: 870-874.

Johnson, M. K. and Farid, H. (2005) Exposing Digital Forgeries by Detecting Inconsistencies in Lighting. *ACM Multimedia and Security Workshop.*

Jolion, J. M. (2001) Images and Benford's law. Journal *of Mathematical Imaging and Vision* 14: 73–81.

Kavallaris, T. and Katos, V. (2010) On the detection of pod slurping attacks. *Computers & Security* 29: 680-685.

Kennedy, L. and Chang, S. (2009) Internet Image Archaeology: Automatically Tracing the Manipulation History of Photographs on the Web. *Patent* IR M08-087.

Khan, M. K., Jiashu, Z. and Lei, T. (2007) Chaotic secure content-based hidden transmission of biometric templates. *Chaos, Solitons and Fractal* 32: 1749–1759.

Khan, S. and Kulkarni, A. (2010) Reduced Time Complexity for Detection of Copy-Move Forgery Using Discrete Wavelet Transform. *International Journal of Computer Applications* 6(7): 31-36.

Krivko, M. (2010) A hybrid model for plastic card fraud detection systems. *Expert Systems with Applications* 37: 6070-6076.

Lee, S., Shamma, D. A. and Gooch, B. (2006) Detecting false captioning using common-sense reasoning. *Digital Investigation* 3S: S65-S70.

Leprince, S., Berthier, E., Ayoub, F., Delacourt, C. and Avouac, J. (2008) Monitoring earth surface dynamics with optical imagery. *Eos, Transactions, AGU* 89(1): 1-2

Li, W., Yuan, Y. and Yu, N. (2008) Detecting Copy-Paste Forgery of JPEG Image Via Block Artifact Grid Extraction. *LNLA Conference* 1006-1011.

Li, Y. , Yeh, M. and Chang, C. (2010) Data hiding based on the similarity between neighboring pixels with reversibility. *Digital Signal Processing* 20: 1116–1128.

Lin, C. and Chang, S. (2000) Semi-Fragile Watermarking for Authenticating JPEG Visual content. *SPIE Security and Watermarking of Multimedia Content II.*

Lin, C. and Chang, S. (2001) SARI: Self-Authentication-and-Recovery Image Watermarking System. *ACM Multimedia 2001.*

Lin, S. D. and Kuo, Y. (2007) An image watermarking Scheme with Tamper Detection and Recovery. *International Journal of Innovative Computing, Information and Control* 3/6(A): 1379-1387.

Lin, P. L., Hsieh, C. and Huang, P. (2005) A hierarchical digital watermarking method for image tamper detection and recovery. *Pattern Recognition* 38: 2519-2529.

Lukas, J., Fridrich, J. and Goljan M. (2006) Detecting Digital Image Forgeries Using Sensor Pattern Noise. *SPIE Electronic Imaging.*

Madeira, S., Gonçalves, J. and Bastos L. (2010) Photogrammetric Mapping and Measuring Application Using MATLAB. *Computer & Geosciences* 36(6): 699-706.

Mahdian, B. and Saic, S. (2007) Detection of copy-move forgery using a method based on blur moment invariants. *Forensic Science International* 171(2-3): 180-189.

Mahdian, B. and Saic, S. (2008a) Blind authentication using periodic properties of interpolation. *IEEE Transactions on Information Forensics and Security* 3(3): 529–538.

Mahdian, B., and Saic, S. (2008b) Blind Methods for Detecting Image Fakery. *Security Technology* 280-286.

Mahdian, B. and Saic, S. (2009) Using noise inconsistencies for blind image forensics. *Image and Vision Computing* 27: 1497-1503.

Mahdian, B. and Saic, S. (2010) A bibliography on blind methods for identifying image forgery. *Signal Processing: Image Communication* 25: 389-399.

Mairal, J., Bach, F., Ponce, J., Sapiro, G. and Zisserman, A. (2009) Non-local sparse models for image restoration. *IEEE 12th International Conference on Computer Vision* 2272–2279.

Mallat, S. G. (1989) A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11(7): 674–693.

Math, S. and Tripathi, R. C. (2010) Digital Forgeries: Problems and Challenges. *International Journal of Computer Applications* 5(12): 9-12.

Meko, D. M., Touchan, R. and Anchukaitis, K. J. (2011) Seascorr: A MATLAB program for identifying the seasonal climate signal in an annual tree-ring time series. *Computers & Geosciences* 37: 1234-1241.

Moody, S. J., Erbacher, R. F. (2008) Statistical analysis for data type identification. *IEEE International Workshop on Systematic Approaches to Digital Forensic Engineering* 41-54

Müller, H., Müller, W., Squire, D. M., Maillet, M. and Pun, T. (2001) Performance evaluation in content-based image retrieval: overview and proposals. *Pattern Recognition Letters* 22(5): 593–601.

Nieke, J., Solbrig, M., Neumann, A (1999) Noise contributions for imaging spectrometers. *Applied Optics* 38 (24): 5191- 5194.

Niekerk, J.F.V. and Solms, R. V. (2010) Information security culture: A management perspective. *Computers & Security* 29: 476-486.

Nillius, P. and Eklundh, J. O. (2001) Automatic estimation of the projected light source direction. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition.*

Patra, J. C., Phua, J. E. and Bornand, C. (2010) A novel DCT domain CRT-based watermarking scheme for image authentication surviving JPEG compression. *Digital Signal Processing* 20: 1597-1611.

Peng, F., Li, X. and Yang B. (2012) Adaptive reversible data hiding scheme based on integer transform. *Signal Processing* 92: 54-62.

Petitcolas, F. A. P., Anderson, R. J. and Kuhn, M. G. (1999) Information hiding—a survey. In: *Proceedings of IEEE* 1062–1078.

Popescu, A. C. and Farid, H. (2004a) Statistical Tools for Digital Forensics. *Dartmouth College, Hanover.*

Popescu, A. C. and Farid, H. (2004b) Exposing Digital Forgeries by Detecting Duplicated Image Regions. *Darmouth College.*

Popescu, A. C. and Farid, H. (2005) Exposing Digital Forgeries in Color Filter Array Interpolated Images. *Signal Processing* 3948-3959.

Pudil, P., Novovičová, J. and Kittler, J. (2003) Floating search methods in feature selection. *Pattern Recognition Letters* 2(11): 1119-1125.

Qi, M., Lu, Y. H., Li, J. S., Li, X. L., Kong, J. (2008) User-Specific Iris Authentication Based on Feature Selection. *CSSE* 1040–1043.

Qi, M., Lu, Y., Du, N., Zhang, Y., Wang, C. and Kong, J. (2010) A novel image hiding approach based on correlation analysis for secure multimodal biometrics. *Journal of Network and Computer Applications* 33: 247-257.

Redi, J. A., Taktak, W. and Dugelay, J. (2011) Digital image forensics: a booklet for beginners. *Multimed Tools Appl* 51:133–162.

Rey, C. and Dugelay, J. (2002) A Survey of Watermarking Algorithms for Image Authentication. EURASIP *Journal on Applied Signal Processing* 6: 613-621.

Rosenfeld, A. (1970) Connectivity in digital pictures. *Journal of the ACM* 17(1): 146–160.

Rosenfeld, A. (1974) Adjacency in digital pictures. *Information and Control* 26: 24–33.

Rosenfeld, A., Nakamura (2002) Two simply connected sets that have the same area are IP-equivalent. *Pattern Recognition* 35(2): 537-541.

Sekeh, M. A., Maarof, M. A., Rohani, M. F. and Motiei, M. (2011) Sequential Straightforward Clustering for Local Image Block Matching. *World Academy of Science, Engineering and Technology* 74: 775-779.

Shaamala, A., Abdullah, S. M. and Manaf, A. A. (2011) Study of the effect DCT and DWT domains on the imperceptibility and robustness of Genetic watermarking. International *Journal of Computer Science Issues* 8(5/2): 220-225.

Shebaro, B., Gonzalez, F. P. and Crandall J. R. (2010) Leaving timing-channel fingerprints in hidden service log files. *Digital Investigation* 7: S104-S113.

Shefali, S., Deshpande, S. M. and Tamhankar, S. G. (2007) Attack Detection through Image Adaptive Self Embedding Watermarking. *World Academy of Science, Engineering and Technology* 29: 298-304.

Shih, F.Y. and Yuan, Y. (2010) A Comparison Study on Copy-Cover Image Forgery Detection. *The Open Artificial Intelligence Journal* 4: 49-54.

Shivakumar, B. L., and Baboo, S. (2011) Detection of Region Duplication Forgery in Digital Images Using SURF. *International Journal of Computer Science Issues* 8/4(1): 199-205.

Smith, J. R. (1998) Image retrieval evaluation. In: *Proceedings of the IEEE Workshop on Content-based Access of Image and Video Libraries* 112–113.

Stamm, M.C. and Liu, K. J. R. (2010) Forensic Detection of Image Tampering Using Intrinsic Statistical Fingerprints in Histograms. *Information Forensics and Security* 5(3): 492-506.

Sturm, B.L. and Daudet, L. (2011) Recursive Nearest Neighbor Search in a Sparse and Multiscale Domain for Comparing Audio Signals. *EURASIP Journal on Applied Signal Processing.*

Sunderrajan, S. (2009) Exposing Digital Forgeries in JPEG and Bitmap Images. *University of California, Santa Barbara.*

Sutcu, Y., Coskun, B., Sencar, H. T. and Memon, N. (2007) Tamper Detection Based on Regularity of Wavelet Transform Coefficients. *Image Processing* 397-400.

Suthaharan, S. (2004) Fragile image watermarking using a gradient image for improved localization and security. *Pattern Recognition Letters* 25: 1893-1903.

Swaminathan, A., Wu, M. and Liu K. J. R. (2007) Image Tampering Identification Using Blind Deconvolution. *ICIP 2006* 2309-2312.

Tian, J. (2003) Reversible data embedding using a difference expansion. *IEEE Trans. Circuits Syst. Video Technol.* 13(8): 890–896.

Von Solms B. (2000) Information security - the third wave? *Computers and Security* 19(7): 615–620.

Walton, S. (1995) Information Authentication for a Slippery New Age. *Dr. Dobbs Journal* 20(4): 18–26.

Wolfgang, R. B. and Delp, E. J. (1996) A Watermark for Digital Images. In: *Proceedings of IEEE Int. Conf. on Image Processing* 3: 219–222.

Wong, P. (1998) A Watermark for Image Integrity and Ownership Verification. In: *Proceedings of IS&T PIC.*

Wong, P. and Memon, N. (2001) Secret and public key image watermarking schemes for image authentication and verification. *IEEE Trans. on Image Processing* 10(10): 1593:1601.

Wray, J. C (1991) An analysis of covert timing channels. *IEEE symposium on security and privacy* 2-7.

Wu, D. C. and Tsai, W. H. (2003) A steganographic method for images by pixel-value differencing. *Pattern Recognition Letters* 24: 1613–1626.

Wu, W. Z. and Zhang, W. X. (2002) Neighborhood operator systems and approximations. *Information Sciences* 144: 201–217.

Xiong, M. M., Fang, X. Z. and Zhao, J. Y. (2001) Biomarker identification by feature wrappers. *Genome Research* 1787–1887.

Yali, L., Cherita, C., Ken, C., Rennie, A., Mukherjee, B. and Ghosal, D. (2009) SIDD: a framework for detecting sensitive data exfiltration by an insider attack. *Hawaii International Conference on Systems Science* 1-10

Yeung, D. and Ding, Y. (2003) Host-based intrusion detection using dynamic and static behavioral models. *Pattern Recognition* 36(1): 229–243.

Yeung, M. and Mintzer, F. (1997) An Invisible Watermarking Technique for Image Verificaton. In: *Proceedings of ICIP.*

Yeung, M. M. (1998) Digital watermarking. *ACM Communications* 41(7): 30–33.

Yu, Z. and Bajaj, C. (2002) Image segmentation using gradient vector diffusion and region merging. *International Conference on Pattern Recognition* 2: 20941.

Zhang, S., Huang, D. and Wang S. (2010) A method of tumor classification based on wavelet packet transforms and neighborhood rough set. *Computers in Biology and Medicine* 40: 430-437.

Zhang, X., Wang, S., Qian, Z. and Feng G. (2010) Reversible fragile watermarking for locating tampered blocks in JPEG images. *Signal Processing* 90: 3026-3036.

Zhao, H., Wang, H. and Khan M. K. (2011) Stegnalysis for palette-based images using generalized difference image and color correlogram. *Signal Processing* 91: 2595-2605.

Zhu, B., Swanson, M. D. and Tewfik, A. (1997) Transparent Robust Authentication and Distortion Measurement Technique for Images. In: *Proceedings of ICIP.*

Zimba, M. and Xingming, S. (2011) DWT-PCA (EVD) Based Copy-move Image Forgery Detection. *International Journal of Digital Content Technology and its Applications* 5: 251-258.

ooooooooooooooooooooooooooooooooooo

PUBLICATIONS

International Journal of Computational Intelligence Research
ISSN 0973-1873 Volume 5, Number 2 (2009), pp. 123–126
© Research India Publications
http://www.ripublication.com/ijcir.htm

Statistical Analysis of Fingerprint Pattern

Brijesh Jajal[1] and Vipul Desai[2]

[1]Department of Computer Science, ARIBAS, New Vidyanagar, Gujarat, India.
[2]Charutar Vidya Mandal, Vallabh Vidyanagar, Gujarat, India.
[1]E-mail: jajalbr@yahoo.com and [2]E-mail: drvdesai@rediffmail.com

Abstract

Fingerprints are the ridge and furrow patterns on the tip of the finger and have been used extensively for personal identification of people. A number of approaches to fingerprint classification have been developed. Some of the earliest approaches did not make use of the rich information in the ridge structures and exclusively depended on the orientation field information.

Due to the limited amount of information present in the minutiae-based representation, it is desirable to explore alternative representations of fingerprints. To assess the performance limitations of popular minutiae-based fingerprint verification system, we theoretically estimate the probability of a false correspondence between two fingerprints from different fingers based on the minutiae representation of fingerprints. As a result, the most successful approaches need to use reliable structural/syntactic pattern recognition methods and statistical methods. This paper performs the comparison of a particular fingerprint configuration using different models. For a fair comparison, the differences between minutiae types are not considered.

Keywords: Fingerprint statistics, Fingerprint analysis

Introduction

Every human contains a God gifted pattern of fingerprint, which is always different for two individuals even if they are identical twins. A fingerprint is an impression of the friction ridges of all parts of the finger tip. A friction ridge is a raised portion of the epidermis on the palm or digits (fingers and toes), consisting of one or more connected ridge units of the skin. Fingerprints are one of the most proficient biometric technologies and are considered legitimate proof of evidence in courts of law all over the world. Fingerprints are, therefore, used in forensic divisions worldwide for criminal investigations.

The patterns of a fingerprint are generally classified based on its minutiae [3],[7] as depicted in Fig. 1.

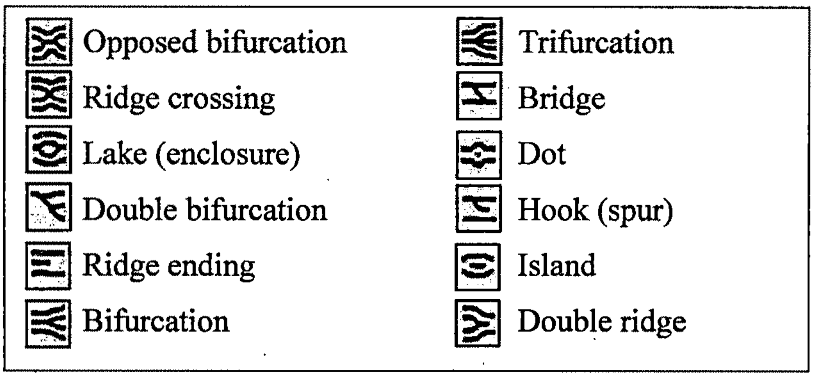


Figure 1: Minutiae based classification of fingerprints.

Experiments and results

The aim of current study is to determine the individuality of a given fingerprint sample [2], and statistically compare it with the other samples. In order to analyse, the four major criterions are Ridge area, Minutiae density, distance between neighbouring minutiae and Ridge wavelength.

Ridge area

This is a novel approach to fingerprint statistics, which determines the global perspective for the comparison. The ridge area for a sample fingerprint can be formulated as follows:

$$A_r = \sum r_p / t_p,$$

where r_p represents the total no. of pixels represented in form of a foreground ridge, and t_p denotes the total no. of pixels of an image. An advantage of a single resultant value of this approach is that the comparison criteria can be set to required level of accuracy. The results have shown that the value of A_r ranges from 0.4 to 0.6.

Minutiae Density

It is a factor determining the no. of minutiae points found in a given sample. The fingerprint images of 360 x 360 pixels are considered as a sample, with 1 mm^2 represented by 19.7 pixels. The density is calculated by means of important bifurcations found in an image of fingerprint. The results of Table 1 are carried out based on fingerprints of different patterns (Fig.1).

The results of Table 1 illustrate that the mean density attained in current study are at par with the other sources. However the high value of standard deviation indicates that the fingerprint image samples are of varying types. Another major reason for variation in S.D. value is that the images are used without applying image enhancement algorithm [6].

Table 1: Statistical values obtained for fingerprint samples.

Data Set	Source	Sample Size	Mean Density	S.D.
1	Dankmeijer *et al* [5]	1000	0.190	0.007
2	Raymond Thai [1]	30	0.204	0.029
3	Current study	30	0.200	0.049

Distance between neighbouring minutiae

Subsequent to determining the minutiae points, a distance between two minutiae can now be analysed, where Delaunay triangulation method can be used to create a triangular grid for the scattered minutiae points. This method can be accessed in MATLAB via the delaunay function. The triangulation function is performed for specific types of minutiae, such as loop type and arch type (Fig.2).

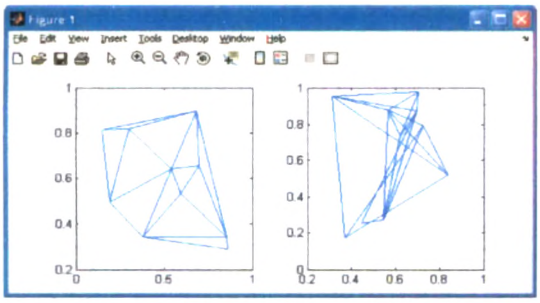


Figure 2: Triangulation pattern for (a) Loop type (b) Arch type

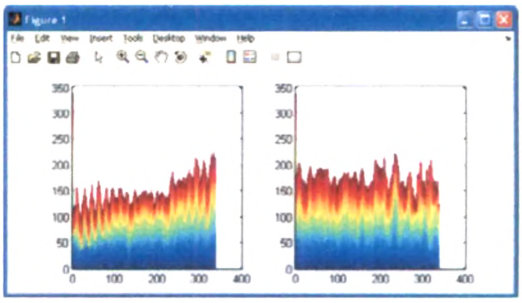


Figure 3: Ridge wavelength for (a) Loop type (b) Arch type

Ridge Wavelength

A fingerprint image can be segmented to foreground data, which represents the ridge characteristics. As a more accurate measure in testing the normality of the data, the normal probability plot via the MATLAB function normplot is used. The observed ridge wavelength data does not quite follow a normal distribution (Fig. 3.)

Future Scope

The statistical experiments used in this paper can be performed on a larger sample size, and a full analysis of the observed results can be conducted. The further study into the statistical theory of fingerprint minutiae can be carried out. In particular, the Tu and Hartley approach [4] can be investigated to determine the number of degrees of freedom within a fingerprint population [1].

The basic parameters specified in the paper can act as a foundation to further analyze the statistical parameters of a fingerprint, which adds a more uniqueness and precision in comparison.

References

[1] R. Thai, Report, The University of Western Australia, 2003.

[2] S. Pankanti, S. Prabhakar, and A.K. Jain, "On the individuality of fingerprints", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 8, 2002, pp. 1010–1025.

[3] A.K. Jain, A. Ross, and S. Prabhakar, "Fingerprint Matching Using Minutiae And Texture Features", Proc. of Int'l Conference on Image Processing (ICIP), Thessaloniki, Greece, Oct 2001, pp.282-285.

[4] Tu, P., and Hartley, R. , "Statistical significance as an aid to system performance evaluation", ECCV (2) 2000, vol. 85, pp. 366–378.

[5] A.K. Jain and S. Prabhakar, "A Multichannel Approach to Fingerprint Classification", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, no. 4, April 1999, pp. 348-359.

[6] L. Hong, Y. Wan, and A.K. Jain, "Fingerprint image enhancement: Algorithm and performance algorithm", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 8, May 1998, pp. 777–789.

[7] M. Chong, T. Ngee, L. Jun, and R. Gay, "Geometric framework for fingerprint image classification", Pattern Recognition, vol.30(9), 1997, pp. 1475–1488.

PRAJÑĀ - Journal of Pure and Applied Sciences
Vol. 18: 104-107, 2009 ISSN 0975 - 2595

SENSING IMAGE FORGERY USING MATLAB

Brijesh Jajal+ and Vipul Desai *

Ashok & Rita Patel Biotechnology Institute (ARIBAS), New Vallabh Vidyanagar, Gujarat - 388 121
Management Advisor, Charutar Vidya Mandal (CVM), Vallabh Vidyanagar, Gujarat - 388120

ABSTRACT

Due to availability of immense number of image editing software, a layman can easily perform editing operations into a digital image. Among the major operations, Copy-paste of one image area into another is considered to be the common example of manipulating image. A digital image consists of storage representation of RGB (Red-Green-Blue) color format. Usually an image is assumed to be acquired in 24 bits Bitmap format. The procedure devised here senses the same regions within an image, assuming the test image to contain such forgery. The suggested procedure can be considered as a tool to image forensics. This paper discusses the procedure successfully implemented using Matlab.

***Keywords*: Image forensic, image tampering, image doctoring, detection of copy-paste regions.**

INTRODUCTION

In July 2009, a picture essay in The New York Times (NYTimes) Sunday Magazine entitled "Ruins of the Second Gilded Age", by Edgar Martins, showed large housing construction projects that were halted due to the housing market collapse. After discovering the photo manipulations, the Times posted the following on their website. "After a reader discovered that the photos were digitally altered, Editors later confronted the photographer and determined that most of the images did not wholly reflect the reality they purported to show. Had the editors known that the photographs had been digitally manipulated, they would not have published the picture essay, which has been removed from their website."

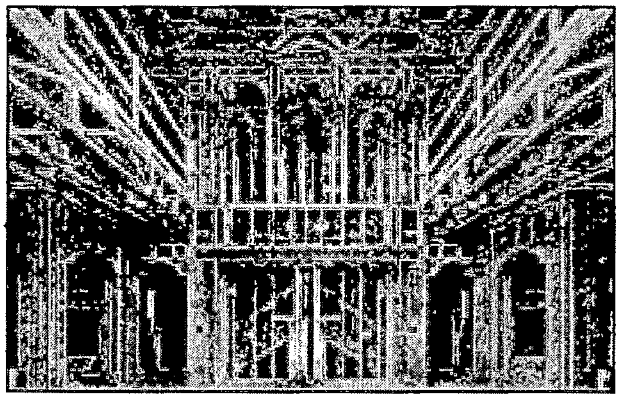


Figure 1: The image forgery by photographer of NYTimes.

The given photograph (Fig. 1.) is an example of forgery into the original image by pasting a specific area of another digital image. Since the image is represented in RGB format [2], such manipulations contain a clear partition line indicating the major changes into pixel values [3] of an image. The detection of such partitions can expose the forgery.

There are few methods available for the detection of such digital image forgery. The lighting condition of two separate photographs can detect the forgery with the help of estimation of a point light source [5]. However, this method might not work when the object does not have a compatible surface. The other detection methods are based on digital camera properties, like watermarking [4] and sensor pattern noise [8]. The Copy-Move detection methods [9], [7] are limited to one particular case of forgeries. This paper introduces the forgery detection with a general perspective, by using color data of an image.

We refer to a new approach for detecting the image portions which are copied from another image. This method can supplement to the CFA (Color Filter Array) interpolation technique [6]. This research paper illustrates a source code in Matlab, which senses the image forgery which is carried out using a Copy-Paste operation.

METHODOLOGY

A test image contains the image forgery, where the portion of a same image and copied and pasted into another part of an image. A raster scan [2] of an image senses the image points having the same color values. Since such forgery is assumed to contain a group of pixels instead of an individual

* Corresponding author – jajalbr@yahoo.com

value, the comparison of a left/right co-ordinate value is also done. The double comparison also reduces the false positive values. A 4-connected method for flood-filling a graphical object considers the neighbours on all 4 sides, while in a current study; 1-connected approach [3] is applied.

A Matlab Source Code for Sensing image forgery is as follows:

The above algorithm executes to compare each and every pixel of an image with another one having exactly the same value. Once the same pixels are determined, they are also checked to have same value of neighbouring pixels. This check is with reference to the assumption that a wide area is copied, rather than a single pixel.

On determination of such set of same pixels, they are assumed for forgery and an algorithm changes the pixel value to zero, representing black color.

```
% Matlab code to sense the copy-paste regions into image file
i=imread('sample1.bmp');

z=zeros(100,100); %  To check that pixel is already processed or not

for x=1:98

  for y=1:98

    for a=x:98

      for b=y:98

        if (z(a,b)==0)  % If not processed then ..

          if ( (i(x,y,1)==i(a,b,1)) && (i(x,y,2)==i(a,b,2))&&(i(x,y,3)==i(a,b,3))&&

          (i(x+1,y,1)==i(a+1,b,1)) && (i(x+1,y,2)==i(a+1,b,2))&&(i(x+1,y,3)==i(a+1,b,3))&&

                                        (~ ((x==a)&&(y==b))))

                                        % Check that pixel with its adjacent (x+1) is also same

                                              i(x,y,1)=0;  % Make all 3 values 0 for black.

                                              i(x,y,2)=0;

                                              i(x,y,3)=0;  % i(a,b) represents another set of copy.

        end  % inner if over

      end  % outer if over

    end  % b loop

   end  % a loop

   z(x,y)=1; % Now pixel checking is over with all other pixels of image.

 end % y loop – image width

end % x loop – image height

imshow(i)  % Displaying resultant sensed areas.
```

n image manipulation or forgery is assumed to e carried out with the two major intentions:

I. Copying the background portion to hide some information, considered as eliminate forgery.
II. Copying the portion to show it multiple times, considered as enhance forgery.

The Fig. 2. represents an original family photo, which is tampered to remove the unwanted person on left side (Fig. 3). The execution of a procedure considers the forgery image as an input and results into sensing the area which is tampered (Fig.7). In order to verify the procedure for enhance forgery, another image (Fig.4) is tampered by the user (Fig.5), which also senses the forgery area successfully (Fig. 8).

ANALYSIS

From the above results, it can be easily predicted that a second run of the same algorithm may be carried out in order to reduce the false positives, which is found to be less than 10% in the results.

ure 4: River photo

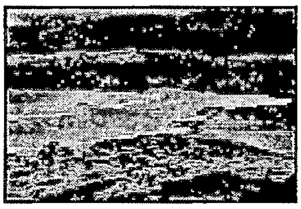
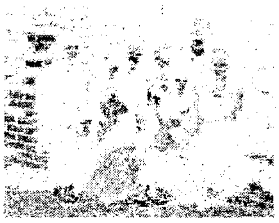


Figure 5: Enhance Forgery



ure 2: Family photo

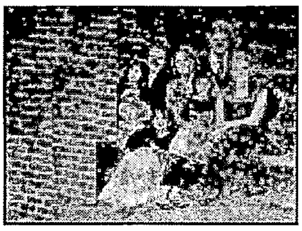


Figure 3: Eliminate Forgery

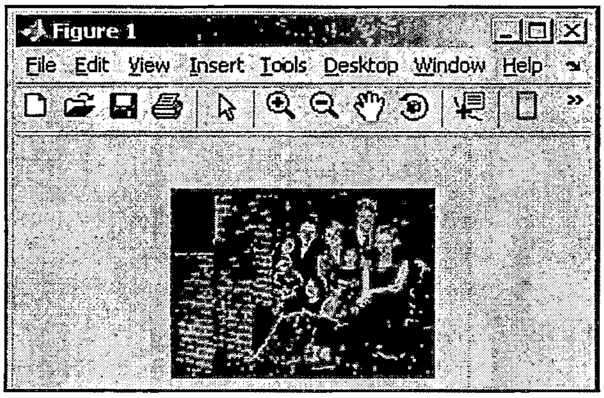


Figure 7: Sensing Eliminate Forgery

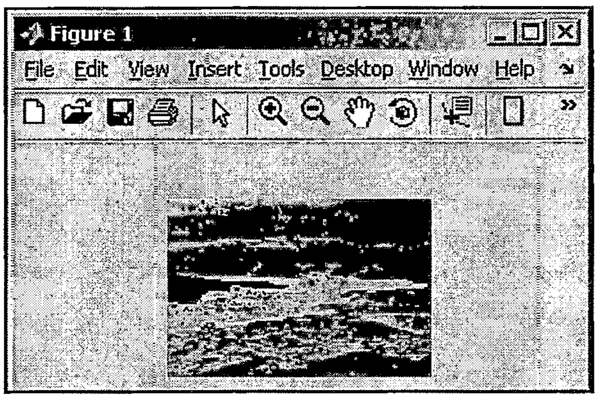


Figure 8: Sensing Enhance forgery

Table 1: Image samples with corresponding false positives

Image RGB Distribution	False positives	Percentage
Uniform 100x100 (Worst Case)	730	7.30
Normal 100x100	103	1.03
Random 100x100 (Best case)	12	0.12

Table-1 shows the results of using 1-connected approach. From the above results, it can be easily predicted that if the image comprises the similar areas within, the false positives are naturally ought to be more. However, the moderate images are found to be almost near to the normal distribution.

CONCLUSION

This paper is an effort to determine the image forgery at preliminary level, with a straight forward imperative. The method can be utilized in conjunction with the other existing techniques, for better results and cross verifications. The problem of detection of digital forgeries is a complex one with no universally applicable solution.

ACKNOWLEDGEMENT

The authors convey sincere gratitude to Dr. Darshan Chokshi, Professor and Head, P. G. Department of Computer Science for his helpful tips and guidance. The appreciation is also due to Dr. Paresh Virparia for being a source of inspiration to the research work.

FERENCES

1] *Website*: http://www.hackerfactor.com

2] Gonzalez, R. C. and Wood, R. E. (2005) *Digital Image Processing*, 2nd edn. Addison Wesley Publishing Company.

3] Hern, D. and Baker, M. P. (1996) *Computer Graphics*, 2nd edn. Prentice-Hall, India.

[4] Katzenbeisser, S. and Petitcolas, F. (2000) *Information Techniques for Steganography and Digital Watermarking*, Artec House.

[5] Johnson, M. K. and Farid, H. (2005) *Exposing Digital Forgeries by Detecting Inconsistencies in Lighting*, ACM Multimedia and Security Workshop '05 New York, New York, USA.

[6] Popescu, A. C. and Farid, H. (2005) *Exposing Digital Forgeries in Color Filter Array Interpolated Images*, IEEE Transactions on Signal Processing, **53**(10), pp.3948-3959.

[7] Popescu, A. C. and Farid, H. (2004) *Exposing Digital Forgeries by Detecting Duplicated Image Regions*, Technical Report, TR2004-515, Darmouth College, Computer Science 2004.

[8] Lukáš, J., Fridrich, J. and Goljan, M. (2006) Detecting Digital Image Forgeries Using Sensor Pattern Noise, In *Proceedings of the SPIE*, **6072**, pp. 362-372.

[9] Fridrich, J., Soukal, D. and Lukáš, J. (2003) Detection of Copy-Move Forgery in Digital Images, In *Proceedings of Digital Forensics Research Workshop.*

entification of Copy–Paste Regions In Digital Image

jesh Jajal[1], Vipul Desai[2]

shok & Rita Patel Institute of Integrated Study and Research in Biotechnology and Allied
ences (ARIBAS), New Vidyanagar, Gujarat, India

harutar Vidya Mandal, Vallabh Vidyanagar, Gujarat, India

eived 22 October 2008; accepted 28 September 2010

TRACT: In recent technocrat world, an alteration of a digital ge is more ubiquitous amongst techno-savvy professionals, which also been proved recurrent even for laymen. This has been popu- ed on account of a lucid access of different types of image-editing are. Copying a particular region from a digital image to selective tion within the same image is one of the good citations of image toring. Usually, the bitmap pictures are represented in the form of -channel color image, in which the algorithm identifies the similar s with an assumption of image acquisition in 24 bits Bitmap for- . In the present article, an exclusive procedure was applied to pro- e an output image, pinpointing the copy–paste area with more n 90% accuracy. The resultant image was depicting a forgery oper- n presumed to be performed, which determined two areas of simi- . A novel approach of 1-connected graph was applied, as the for- is not believed to be done in the form of a petite point-like area. ally, the forgery area was exposed with an aid of discerning color e, commonly as a black color for an apparent visibility of an ge. The present application will be a tool in image forensics that be applicable to identify the copy–paste regions in a single bitmap ge. This article refers to a new approach for detecting the image ions which are copied from another image. Besides, the present estigation discusses an algorithm effectively implemented to deter- e the areas formed by copy–paste operation in an image.

010 Wiley Periodicals, Inc. Int J Imaging Syst Technol, 20, 367–369, 2010; lished online in Wiley Online Library (wileyonlinelibrary.com). DOI 10. 2/ima.20257

y words: image doctoring; image forensic; image tampering; y–paste detection

INTRODUCTION

e modification of a digital image in terms of copy–paste opera- n may generate misinterpretation and misconception too amongst age viewers, which leads to dire consequences thereby. In recent es, the latest example in such cases has been reported (http:// .hackerfactor.com).

The present example is exemplified in case of one of the provin- of Gulf viz. Iran. Recently, Iran released a photo of the test. The photo claimed to show four missiles in flight. However, the photo was clearly doctored. The real photo shows three missiles and a truck into it. For the doctored picture, someone copy cloned one of the missiles and the smoke cloud therein. This fake photo was aired by many news organizations.

Although Figure 1 depicts the single manipulated photo with one extra missile; it actually represents a larger distortion of an image. The image formed due to copy–paste operation does not reveal its originality. In controversy to the news of launching nine missiles, the photo exhibits only three being launched.

The given photograph is one of the appropriate examples of forgery into the original image by pasting a specific area of the same image (Fig. 1). As the image is represented in Red-Green-Blue (RGB) format, such manipulations contain similarity in pixel values of an image. The detection of such similarity can expose the forgery. The method of comparing lighting condition of two separate photographs can detect the tampering with the help of estimation of a point light source (Johnson and Farid, 2005). However, this method might not work when the object does not have a compatible surface. Another detection method is based on digital camera properties like watermarking (Katzenbeisser and Petitcolas, 2000) and sensor pattern noise (Lukáš et al., 2006), whereas the copy–paste forgery detection method in this article deals with the image color data. The copy–move detection methods are limited to one particular case of forgeries, for a portion of an image copied and pasted into the same image (Popescu and Farid, 2004; Fridrich et al., 2003). This article refers to a new approach for detecting the image portions which are copied from another image. The method used here can be utilized as a good supplement to the color filter array interpolation technique (Popescu and Farid, 2005). This research paper illustrates an algorithm implemented using MATLAB® (The MathWorks, Inc., 2007), which senses the image forgery performed by means of a copy–paste operation.

II. METHOD

The input of an algorithm is a digital image, which was assumed to be forged using the copy–paste operation. Each pixel value of an image was compared with the residual pixels; performing the comparison in form of Raster Scan Technique (Gonzalez and Wood,

Correspondence to: Brijesh Jajal; e-mail: jajalbr@yahoo.com

ure 1. Forged image of Iran test. [Color figure can be viewed in online issue, which is available at wileyonlinelibrary.com.]

)5). To reduce the false-positives, the algorithm performs a comison with the neighboring pixels too. Consequently, an image a was detected to be forged only if the pixel and its neighbor ler consideration were identical to the set of other two neighbor- pixels.

In the present method, a basic idea was analogous to a 4-con- :ted approach for flood filling of the graphical objects. As only a gle neighbor has been considered in the present investigation, it ıld be termed as a 1-connected approach henceforth (Hern and ςer, 1996).

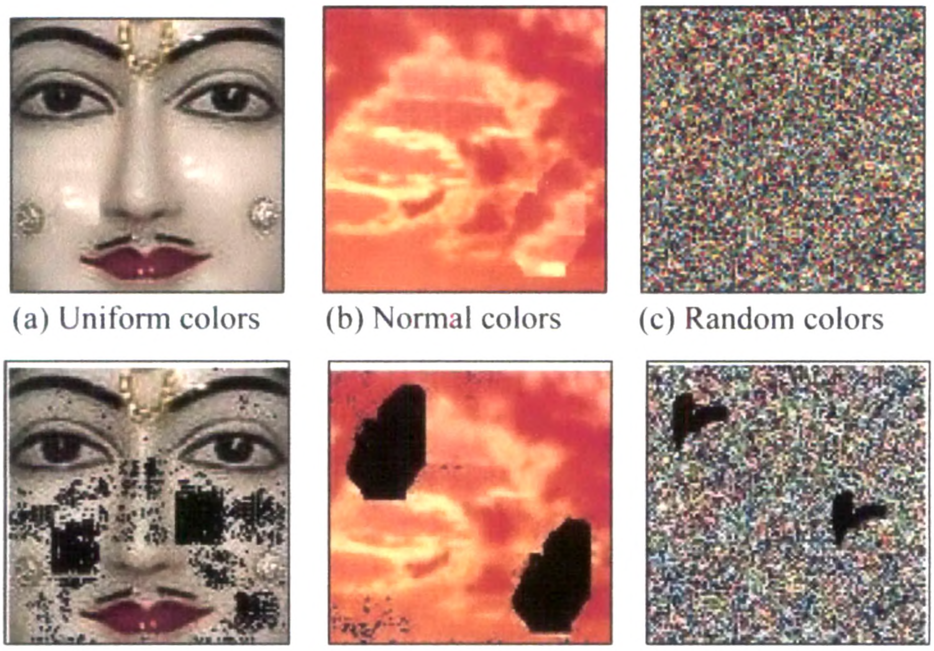
;orithm for Detection of Copy–Paste Regions

```
1. Assume the following parameters of an image.
    • $I$: Image array of dimension $M \times N$ comprising of its corresponding color value
    • $P$: An $M \times N$ Boolean value array indicating whether the positioned pixel comparison is completed or not
    • $x$ and $y$: Current pixel position coordinates
    • $a$ and $b$: Likely to be similar pixel coordinates
    • $c$: Similarity color coefficient, set to zero in case of exactly same color
2. Initialize $P_{x,y}$ = False, indicating the comparison operation not performed.
3. Process step 4 and step 5 for $x, a = 1,\ldots, M$ and $y, b = 1,\ldots, N$
4. If $(I_{x,y} = I_{a,b} \pm c)$ and $(I_{x+1,y} = I_{a+1,b} \pm c)$ and $(P_{x,y}$ = True) then
    • $I_{a,b}$ = BLACK to set the detected similar pixels with desired color.
5. Set $P_{x,y}$ = True, indicating that comparison is over.
```

In accordance with an aforementioned algorithm, the initial false values of P signify the expected comparisons between the pixel color values within the image array (I). This algorithm was implemented using MATLAB® by setting the color coefficient (c) to 5. This condition illustrates that the positive or negative differences of 5 in color values of pixel were identified as same. The similarity values of I are then equated to BLACK color, which eventually spots the forgery. Thus, the final status of an image array (I) exposes the identical areas within the same image during rendering.

III. RESULTS AND DISCUSSION

To implement the algorithm for the detection of forged regions in an image, original images (a, b, c) were chosen, which were considered as an input image array ($I_{x,y}$) of an algorithm. As a result of an application of an algorithm, the resultant corresponding output images (d, e, f) were obtained with dark blotches represented by copy–paste regions (Fig. 2). In the images under consideration, a uniform distribution of colors indicates high chances of false-



(a) Uniform colors (b) Normal colors (c) Random colors

(d) CPR in Uniform (e) CPR in Normal (f) CPR in Random

ure 2. Detection of copy–paste regions (CPRs). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

e I. Image samples with corresponding false-positives.

age RGB Distribution	False-Positives	Percentage
rm 100 × 100 (worst case)	730	7.30
al 100 × 100 (average case)	103	1.03
om 100 × 100 (best case)	12	0.10

tives; while it is reverse in case of random colors, as all the pix- e different.

e results of an implementation of algorithm were derived g the proclaimed 1-connected approach. From the results ired, it can be easily predicted that if the image comprises the lar areas within itself, the false-positives are naturally ought to ore (Table I). A second run of the same algorithm is suggestive educe the false-positives. However, the moderate images are d to be almost near to the normal distribution.

comparison of the results is done by considering 20 sample ges with the block artifact method (Li et al., 2008), where 10 of ixels were found to be false-positives, and 10 of 33 pixels were ely detected in case of statistical artifact (Suderrajan, 2009; le II).

e suggested method in statistical artifacts fundamentally uses k artifact grid mismatch, and hence, proves to be an enhanced ell as more efficient tool compared with the later. The method for the consideration in the present study is solely a distinct, ch directly determines the similar image areas for forgery opera- s. It is able to produce the best results in comparison to the ting methods. Furthermore, the false-positive values for current k approximates to the average of the three classified image s (Table I).

CONCLUSIONS

s article is an extensive attempt to determine the image forgery roundwork with an unobtrusive imperative. In fact, the same orithm can also be used to determine the pixels with assumed ilarity color coefficient value as per the applications and types orgery. At the end of such procedure, using the suggested algo- m, the copied and pasted areas remain indistinct. However, com- n sense logic can be effective to draw out the precise separation, the area being pasted can be considered to be used for the pur- 'e of hiding the original information. Any intended alteration in image such as blurring, contrast settings, or image balance uld be avoided to have precise results using this method. The sent method can be utilized in conjunction with the other exist- tools and techniques for better results and cross verifications of age forgery.

Table II. Comparison of efficiency with the conventional methods.

Method Used	False-Positive Percentage
Block artifact grid extraction	37.04
Statistical artifacts	31.44
Current study	03.08

ACKNOWLEDGMENTS

The authors acknowledge their sincere thanks to Mr. Ramji Makwana for inspiring them to explore the possibilities and avenues of solutions. They cannot forget the whole-hearted support of Dr. Darshan Choksi, Head, Post Graduate Department of Computer Science, Sardar Patel University, Vallabh Vidyanagar, for his constant encouragement and motivation throughout the tenure of the work.

REFERENCES

J. Fridrich, D. Soukal, and J. Lukáš, Detection of copy-move forgery in digital images, Proceedings of Digital Forensics Research Workshop, Cleveland, USA, August 2003.

R.C. Gonzalez and R.E. Wood, Digital image processing, 2nd ed., Addison Wesley Publishing Company, USA, 2005.

D. Hearn and M.P. Baker, Computer graphics, 2nd ed., Prentice-Hall, India, 1996.

M.K. Johnson and H. Farid, Exposing digital forgeries by detecting inconsistencies in lighting, ACM Multimedia and Security Workshop '05, ACM Digital Library, New York, USA, 2005.

S. Katzenbeisser and F. Petitcolas, Information techniques for steganography and digital watermarking, Artec House, UK, 2000.

W. Li, Y. Yuan, and N. Yu, Detecting copy-paste forgery of JPEG image via block artifact grid extraction, The 2008 International Workshop on Local and Non-Local Approximation in Image Processing, Tampere International Center for Signal Processing, LNLA, Switzerland, 2008, pp. 121–126.

J. Lukáš, J. Fridrich, and M. Goljan, Detecting digital image forgeries using sensor pattern noise, Proceedings of the SPIE digital Library, Vol. 6072, 2006, pp. 362–372.

A.C. Popescu and H. Farid, Exposing digital forgeries by detecting duplicated image regions, Technical Report, TR2004-515, Darmouth College, Computer Science, 2004.

A.C. Popescu and H. Farid, Exposing digital forgeries in color filter array interpolated images, IEEE Trans Signal Process 53(2005), 3948–3959.

S. Suderrajan, Exposing digital forgeries in JPEG and bitmap images, Project submitted for degree - MS (ECE), University of California, USA, March 2009.

The MathWorks, Inc. MATLAB® and Simulink®, The MathWorks, Inc., USA, 2007.

Detection of major suspected partition of image forgery in a digital image

Brijesh R. Jajal, Vipul Desai,

***Abstract*—The altering of digital images for a specific purpose not new. It has been around since the early days of otography. The concepts have moved into the digital world by e of digital cameras and the availability of digital image iting software. The ease of use of such software, which does not uire any special skills, makes image manipulation easy to hieve. In most of the cases of such image alteration, also termed image doctoring, the image can be misinterpreted by roducing a portion of another image into the original one. ith an aid to the suggested methodology presented into this per, a forensic analyst can identify a lucid discrimination tween the two separate parts of an image. An RGB (Red Green ue) color value image under test is analyzed, resulting into the ntification of partitions where the image doctoring is rformed. The determined partitions appear in form of distinct rtical black values and locate the forgery area.**

***Index Terms*—Image Processing, Image Color Analysis, Image rensic, Image Forgery Detection**

I. Introduction

N June 2009, the photo of a group of members of the Scottish National Party (SNP) appeared in a SNP wsletter. The two photos in the background of Scottish gends William Wallace and Robert the Bruce were digitally serted, replacing royal portraits of the Queen and the Duke Edinburgh. Local SNP councillor Cecil Meiklejohn blamed "overly enthusiastic" local party member for removing the yal portraits.

The above mentioned case (Figure 1.) is an example of rgery into the original image by pasting a specific area of other digital image. Such manipulations contain a clear 'tion line indicating the major changes into pixel values of image. The detection of such partitions can expose the rgery.

The existing forgery detection method of watermarking [4] ecific to camera properties, where as another method uses timation of a point light source [5] from an image and its ade. The image doctoring can also be identified through olor Filter Array (CFA) interpolation technique [6]. Nevertheless, considering a unique aspect of a type of forgery, any one or more such techniques can be applied to expose the Copy Paste areas [7], consider sensor pattern noise of an image [8]. The Copy-Move detection [9] suggested by Fridrich, Saukal and Lukas and Block Artifact Grid Extraction method [10] are limited to one particular case of forgeries, for a portion of an image copied and pasted into the same image. This paper refers to a new approach for targeting the image partitions where a major change is presumed, and thus reduces the time and effort to determine the whole image portion, as compared to the Copy Paste detection [11]. The paper illustrates an algorithm implemented using MATLAB® (Release 2007b, The MathWorks Inc., U.S.A.), which senses the image forgery performed by means of a Copy-Paste operation [12] and exhibits the suspected partitions.

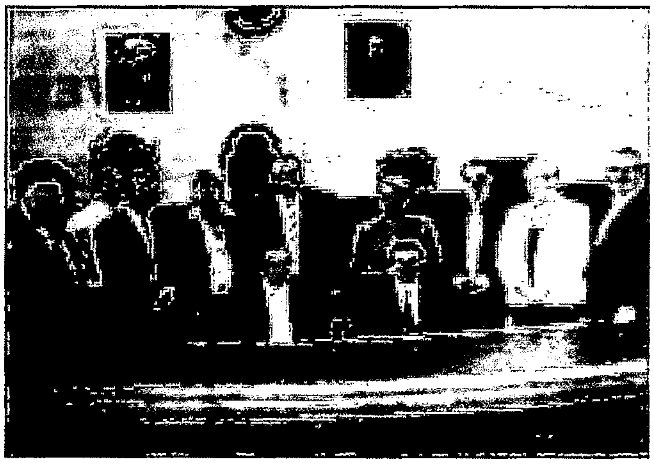


Fig. 1 Scottish National Party photograph

II. Method

In an experiment to detect the suspected partitions of an image, an algorithm is devised to spot such forgery or manipulated areas. The logical sequence for implementation of an idea is exhibited as below:

Step 1. Let a and b are height and width of an RGB image and I is the image pixel value representation matrix of size aXb. F is the determining factor for the pixel difference.

Step 2. Determine the difference between two pixels.

For each value of i = 1 .. a, and j = 1 .. b

Check for the following conditions to be evaluated to true:

a) abs (I a, b – I a, b+1) > F

b) abs (I a, b – I a, b-1) < F

c) abs (I a-1,b – I a-1, b+1) > F

B. R. Jajal is with the Ashok & Rita Patel Institute of Integrated Study & esearch in Biotechnology and Allied Sciences, ARIBAS, New Vidyanagar, ujarat, India (phone: 91-2692-645801; fax: 91-2692-229189; e-mail: jalbr@ yahoo.com).

V. Desai, was in Oil Exploration Research. He is now with the Charutar idya Mandal, Vallabh Vidyanagar, Gujarat, India (e-mail: rvdesai@rediffmail.com).

78-1-6123-3002-0

d) $abs(I_{a-1,b-1} - I_{a-1,b}) < F$

Step 3. If all above four conditions (a, b, c and d) are 'sfied, it indicates that two pixels along the same axis ntain the split color values (highly different) or assumed to forgery prone.

Step 4. Considering the required level of detection L being . of columns or height of forgery, Set partition line to icate black color.

$I_{a,b} = 0$,

where x = a - (L/2), ... , a + (L/2) are the current pixel sitions under study.

A distinct vertical black lining appears into a digital image a consequence of execution of the above algorithm, viz. eation of an input image array I; analysis of the major anges from column B to column C, as compared to the anges from column A to column B (Table 1); spotting the spected areas; and exemplifying vertical forgery partitions.

III. Results

The suggested algorithm is implemented into an image of a (Fig.2), which is prone to the image forgery. The resultant age (Fig.3) shows a clear set of partition lines of black lor, indicating that the region through a given vertical line ntains forgery.

It reveals that the portion of a door is copied and pasted in der to make a car look longer. The lesser value of F in an gorithm results into more precise partition detection, but ore number of false positives.

TABLE I
IMAGE PIXEL POSITIONS

Column A	Column B	Column C
$I_{a-1,b-1}$	$I_{a-1,b}$	$I_{a-1,b+1}$
$I_{a,b-1}$	$I_{a,b}$	$I_{a,b+1}$

IV. Conclusion

Considering the image manipulation or image forgery being complex process, there are bound to be the forgeries with a ariety of nature. The algorithm suggested for vertical line xposing such operation, can also be employed with other aight line techniques.

This paper is an effort to determine the image forgery at reliminary level, with a straight forward imperative. The ethod can be utilized in conjunction with the other existing chniques, for better results and cross verifications.

Acknowledgment

The authors wish to appreciate the concern of Dr. C.L. atel, Chairman, Charutar Vidya Mandal and Dr. D. B. hoksi, Director, Post Graduate Department of Computer cience, Sardar Patel University, during the tenure of work. e acclaim is also due to Dr. Pradip Patel and Dr. V.S. Patel or their kind gestures.

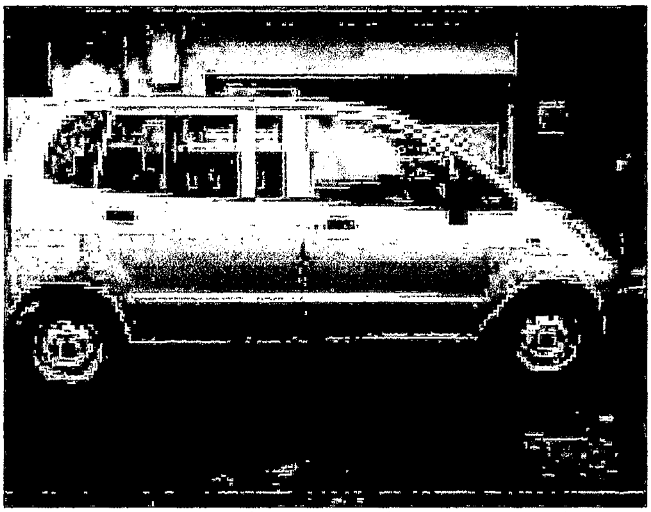
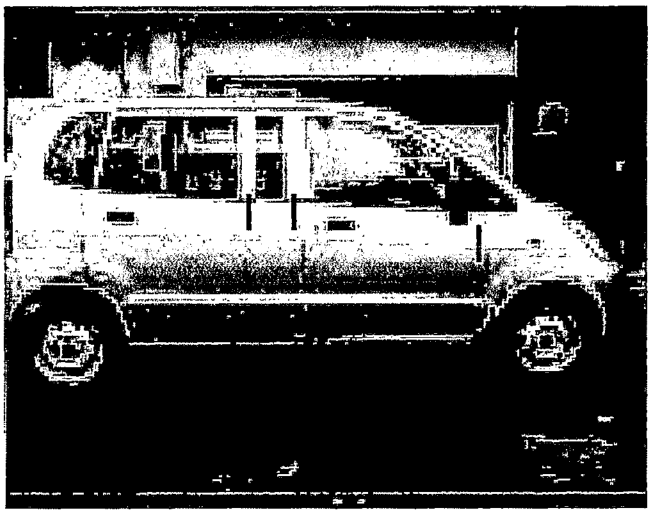


Fig. 2 Forgery prone image



· Fig. 3 Detection of Forgery Partitions

References

[1] Online Source, Digital forgery news, Available: http://news.scotsman.com/politics

[2] R.C. Gonzalez, R.A.Wood, Digital Image Processing, Addison Wesley, 2005.

[3] D. Hern, M. Baker, Computer Graphics, Prentice-Hall, 1996.

[4] S. Katzenbeisser, F. Petitcolas, Information Techniques for Steganography and Digital Watermarking, Artec House, 2000

[5] M.K. Johnson, H. Farid, "Exposing Digital Forgeries by Detecting Inconsistencies in Lighting", ACM Multimedia and Security Workshop '05 New York, New York, USA, 2005.

[6] A.C. Popescu, H. Farid, "Exposing Digital Forgeries in Color Filter Array Interpolated Images", IEEE Transactions on Signal Processing, vol.53(10), pp. 3948-3959, 2005.

[7] A.C. Popescu, H. Farid, "Exposing Digital Forgeries by Detecting Duplicated Image Regions", Technical Report, TR2004-515, Darmouth College, Computer Science 2004.

[8] J. Lukáš, J. Fridrich, M. Goljan, "Detecting Digital Image Forgeries Using Sensor Pattern Noise", in Proceedings of the SPIE, Volume 6072, pp. 362-372, 2006.

[9] J. Fridrich, D. Soukal, J. Lukáš, "Detection of Copy-Move Forgery in Digital Images", in Proceedings of Digital Forensics Research Workshop, 2003.

[10] W. Li, Y. Yuan, N. Yu, "Detecting Copy-Paste Forgery of JPEG Image via Block Artifact Grid Extraction", The 2008 International Workshop

on Local and Non-Local Approximation in Image Processing, LNLA 2008, Switzerland, pp. 121-126.

B.R. Jajal, V. Desai, "Identification of Copy Paste regions in a digital image", PRAJNA Journal of Pure and Applied Sciences, Sardar Patel University, vol.17, pp. 104-107, 2009.

S. Suderrajan, "Exposing Digital Forgeries in JPEG and Bitmap images", Report prepared for University of California, 2009.

esh R. Jajal is a Life Member of Computer Society of India, born in kundla village of Gujarat, on 26th September, 1975. He graduated in ar Patel University and completed MCA in 1999. The author's major y area is image processing methods and determination of digital image ery.

e worked as Lecturer and Head in Computer Science, at Anand cantile College, Anand for 5 years. Currently he works as Assistant essor in Computer Science at ARIBAS, New Vidyanagar, Gujarat, India. research work is accepted in Int. J. of Imaging Systems and Technology 0), Wiley Blackwell Publishers, U.K.

. Jajal received special contribution award in 2008 for his efforts in site design and development of the institute.

ISSN No. 0976-5697

Volume 1, No. 5, Sept-Oct 2011

International Journal of Advanced Research in Computer Science

TECHNICAL NOTE

Available Online at www.ijarcs.info

Technical review of Digital Camera features and High Dynamic Range Challenges

Brijesh Jajal
Asst. Professor: Computer Science
ARIBAS, New Vidyanagar, India
jajalbr@yahoo.com

Vipul Desai
Management Advisor
Charutar Vidya mandal, Vallabh Vidyanagar, India
drvdesai@rediffmail.com

: The inevitable use of existing Digital Cameras has enabled the layman to use them in a professional manner. The present era demands the ce of camera users with robust methods available, and adaptation of a technology to suit their requisite.

s: digital camera; HDR; camera properties; camera features; camera parameters

I. INTRODUCTION

trends and technology of this century have onalised the representation of an image – be it a aph of small dew drops or that of a mega cture. The digital cameras available today have ubiquitous with various salient features.

extensive use of digital camera has imposed the al to have the technical know-how of such devices. erstanding of camera properties can surely aid a apher to achieve better results. The current article has ght to analyze the existing digital camera attributes t the potential technologies.

II. PATH TRAVERSED BY CAMERAS

annes Kepler invented camera and coined the term a obscura". A Latin word "camera" means vaulted r/room, where as "obscura" means darkened r/room.

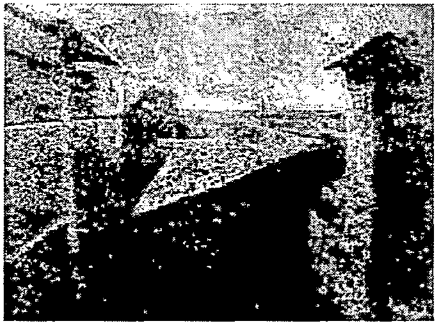
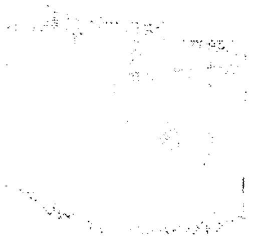


Figure 1. First photograph



re 2. Roll Film camera

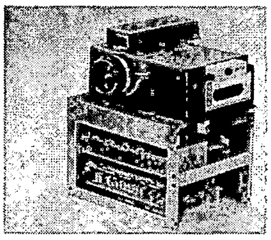


Figure 3. First Digicam

The world's first photograph was taken by Joseph Niepce [1] in 1826, developed using a plate and petroleum product, and named as "heliograph" (Fig. 1). The first roll film camera was invented by Peter Houston (Fig. 2) in 1881. Consequently Eastman patented it for camera and its parts [2]. As known to most users, a roll film contains a film strip, with frame numbering, being wounded forward after every exposure. This strip is 0.025 mm thick and it is made up of celluloid and gelatin. The grains contained in it are photo sensitive, especially for Red, Green and Blue colors. The cost of such camera was $1, when there was a typical salary of $2 per day.

The digital camera was commercially available in 1981 in form of Sony's MAVICA- Magnetic Video Camera (Fig. 3) to be the first one. Subsequently the companies like Kodak, Nikon, Apple and Casio launched their digital cameras.

A Digital Camera functions based on three major aspects –viz. lens, shutter and aperture, with an aim to control focusing of an object of interest, duration of light exposure and the amount of light required respectively. The image capturing also includes various calculations, image storage, format conversion, etc. which are done through the electronics (ICs) present into it. The cost of digital camera available today, ranges from $380-$10,000.

III. ROLL FILM V/S DIGITAL CAMERA

The roll-film camera is used even today since its onset, in view of the fact that it has few advantages over a digital camera [3]. The camera cost factor being a major one, a compulsory photo print is also a plus. The roll film photos can be viewed anywhere without requiring computer screens or TV sets.

However, the digital camera is now replacing the film technology, since it is considered better in many ways. The advantages of digital technology are:

a. No need of roll purchase and tricky insertion.
b. No roll-development charge and time involved.
c. Lower weight compared to film camera.
d. The availability of video recording feature.
e. Memory storage capacity of hundreds of photos.

mmediate removal of unwanted photos.
vailability of special effects and functions.
uilt-in battery to circumvent separate charger.
will now discuss the potential features of digital available currently.

IV. BASIC PARAMETERS

digital camera has a unique feature as per the ion, ranging from very simple functions understood an to the professional settings. The most common l specifications are discussed below.

a pixels: Each portion of an image is stored in form al values, generally in RGB (Red, Green, Blue) indicating the color value of that position. An image on is thus measured in Mega pixels (MP) as being o 1 million pixels [4]. The commonly available capacities are 2 MP, 4 MP, 6 MP, 7.2 MP, 8.1 MP .

not always true that higher the MP, better is the uality, because the lens quality also matters [5]. The value is to be decided by the user, since the higher ue indicates more memory storage required for each Usually 5 MP to 8 MP camera is good enough, if only all prints (5x7 inch) or screen displays (15") are in the output form.

ital and Optical zoom: Most of the cameras provide zoom, which is a bit of illusion. The digital zoom s the post processing part, which reveals that an be viewed by zooming to a given factor like 2X. cal zoom is the "real" zoom feature, where a lens has ity to detect a distant scene clearly [6], [7] without ear to it.

s: The lens property is measured in form of its focal measured in millimeters. It represents the distance n the optical center of a lens and its focal plane. The e focal length, the wider is an angle of view. The d camera has a typical focal length of 50mm to while a professional lens has 80mm to 500mm focal

mory cards: A removable memory card in a camera is tual place where all photos are stored. They are ly in form of flat Smart Media, Compact Flash or Stick type of cards. The latest cards have Secure (SD) feature, available in sizes of SD, Mini-SD and SD.

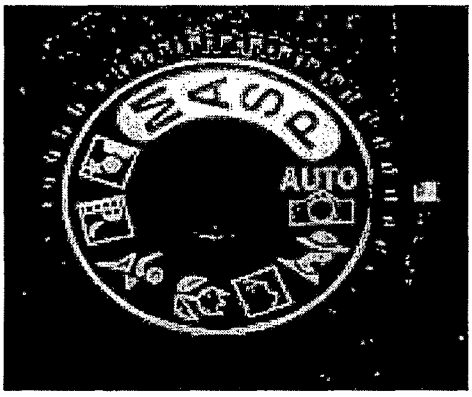


Figure 4. Capture modes

Table I. Applications of Camera modes

Sr .No.	Situation	Mode
1	A family posing outdoors during daytime.	Auto
2	A fashion enthusiast showing tattoo on his hand.	Macro
3	An interesting sign board found on the road.	Aperture
4	Popular car model moving on the road.	Shutter
5	The children blowing crackers at New year eve.	Night
6	A sunrise with different color effects.	Manual

Image capture modes: During the image acquisition, a user is able to select one of the different modes available in a camera (Fig.4.), depending on the type of scene one needs to capture.

a. Auto Mode: The automatic or default settings of a camera are used, which results into a good image quality under normal situation; irrespective of the area to be covered, light conditions, type of surface, etc.
b. Shutter Priority Mode: This feature focuses on the speed of shutter, which is useful in image acquisition of fast moving objects.
c. Aperture Priority Mode: It is useful to focus on a stationary object at a fixed depth, whereas Shutter speed is tuned accordingly.
d. Manual Mode: The multiple factors can be simultaneously adjusted by the user, through manual mode.
e. Macro Mode: The objects which are very close to the camera can be captured using macro mode (typically few inches).
f. Night Mode: A low light conditions, the image production is difficult. This mode uses lower shutter speed in order to capture maximum information and light flash to get the image details without reflection and blur effects.

The features provided by camera manufacturers are generally condition specific. In order to have a glimpse of such conditions, Table 1 represents the applications of such features.

V. ADVANCE FEATURES

The professional camera differs to the commonly available ones in form of the special features that they provide [8]. The subsequent lines will now narrate such pro features.

Red-eye removal: Due to the biological properties, the human or animal eyes appear red with the flash exposure in low light condition. This novel feature removes such defect of an eye colour [9].

Image stabilization: A moving object can be captured by using aperture mode, but the movement of camera itself will result into bad quality image. This feature allows the vibrations in a camera to some extent, tolerating camera shaking by hands or absence of camera stands.

Face detection: The colour and texture features of human face can be identified by using this feature. Hence, the can be focused to face and nearby area, resulting to uman face images.

and DSLR: SLR stands for Single Lens Reflex used essional photographers. Basically the design is such t entering the lens is reflected by a mirror up into the der, allowing the photographer to see exactly what l picture will look like. DSLR - Digital SLR is just a erized version of the mechanical SLR.

VI. FUTURISTIC CAMERAS

Figure.5. Fujifilm G-Shot with 3D image capture

Figure 6. Samsung DualView with 12 MP Sensor

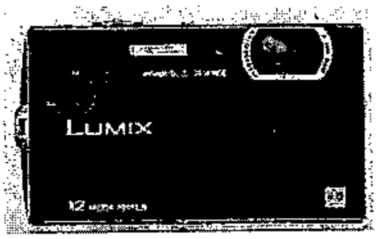


Figure 7. Panasonic FP8 with 27 scene modes

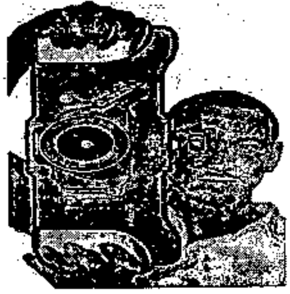


Figure 8. Seitz 160 MP

The frontiers in a digital camera technology suggest creative ideas every day. This results into special camera for fun, such as 3D camera compatible for 3D viewing (Fig.5),two-sided LCD preview for self photo (Fig.6.), light weight camera with multiple modes (Fig.7.) and high MP (160 MP) camera worth $36,000 with mini computer storage instead of memory card (Fig. 8).

VII. ACKNOWLEDGMENT

The authors wish to sincerely thank Dr. C. L. Patel for developing a constant thrive for study and research. An acknowledgement is also due to Dr. Darshan Choksi and Dr. Paresh Virparia for provoking ideas and implementation.

VIII. REFERENCES

[1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955.

[2] E. Raum, "The History of the camera," Chicago, Heinemann Library,2000.

[3] R. White, "Discovering Old cameras 1839-1939," 3rd ed. Pembrokeshire, CIT Printing Services, 2001.

[4] M.Galer, and L.Horvat, "Digital imaging," 3rd ed. Burlington, Focal Press, 2005.

[5] J. Bidner, "Digital Camera Basics: Getting the Most from Your Digital Camera," New York, Silver Pixel, 2001.

[6] S. Kelby, "The Digital Photography Book," Berkeley, CA, Peachpit Press, 2010.

[7] B. Long, "Complete Digital Photography," 4th ed. Massachusetts: Charles River Media, Inc, 2005.

[8] D. D. Busch, "Digital photography all-in-one desk reference for dummies," 3rd ed., New Jersey, Wiley publishing, Inc., 2006.

[9] M. Perkins, "Step by step Digital Landscape Photography. New York: Amherst Media, Inc., 2005.

[10] Jack and S. Dafahl, "Advanced Digital Camera Techniques," New York, Amherst Media, Inc., 2010.

[11] R. Lukac, "Computational Photography: Methods and Applications." Florida, CRC Press, 2010.

[12] D. Johnson, "How to do everything with your digital camera," 4th ed., California, McGraw Hill, 2008.

[13] R. Szeliski, "Computer Vision: Algorithms and Applications," 1st ed., Washington, Springer, 2010.