CHAPTER 1

INTRODUCTION
CHAPTER -1

Introduction

1.1 Introduction to Image Processing

Human beings are primarily visual creatures who depend on their eyes to gather information around them. Of the five senses that human beings have, sight is we depends upon most. Image processing is an integral of us and we continue process image always. Image processing as a subject involves obtained by camera. Images from a camera are fed into a computer where algorithms are written to process these images. Here, the camera replaces the human eye and the computer does the processing. Hence image processing as an engineering subject is basically manipulation of images by a computer. Processing of digital image involve procedures that are usually expressed in algorithmic from due to which most image processing functions are implemented in software. Any image can be converted into its digitized form, and processing on this image is known as digital image processing.

Digital images

Digital images become an important part of our daily lives due to the rapid growth of Internet, the increasing demand of multimedia and rapid advancement of computer technology contents from people. The up soaring number of image applications facilitates image processing. With the overwhelming diffusion of multimedia contents in every-day life, protecting the authenticity and the integrity of these contents from undesired manipulations has become an increasingly important Authentication of digital documents has aroused great interest due to their wide application areas such as legal documents, certificates, digital books and engineering drawings. In addition, more important documents such as fax, insurance and personal documents are digitized and stored.

Image authentication techniques have recently gained great attention due to their importance for a large number of multimedia applications. Digital images are increasingly transmitted over non-secure channels such as the Internet. Therefore, military, medical and quality control images must be protected against attempts to manipulate them; such manipulations could tamper the decisions based on these images. The risks for security are further exacerbated by the improved possibilities of tampering with media contents such as photos, an ability that would have traditionally required many hours of cumbersome work in a
darkroom and that has become now a simple practice using a computer and some commercial software tools. In the case of images, different versions of the same file might differ from the original because of processing due, for instance, to transcoding or bit stream truncation. Other legitimate, content preserving alterations of the original picture are also possible, when the image is enhanced by means of photo editing tools. This kind of modifications includes, for instance, moderate geometrical transformations or slight brightness/contrast adjustments. In other cases, however, one could tamper with part of the image and possibly affect its semantic content in order to illegally abuse it, e.g. to manipulate public opinion or to influence the verdict of the jury in a criminal trial. Given these premises, it is not surprising that a great deal of attention has been turned to methods able to offer proof of authenticity of an image and, in the case of detected tampering, to identify which kind of attack has been carried out.

The purpose of image authentication is used to ensure the integrity of the contents of the digital image. Therefore, efficient and automatic techniques are desired to identify and verify the contents of digital images. Image authentication is such a promising technique to automatically identify whether a query image is a different one, or a fabrication, or a simple copy of an anchor image. Here, the anchor image is the ground truth image or the original image as an authentication reference, and the query image is the one under suspicion.

1.2 Digital Image Processing

An image may be defined as a two-dimensional function, $f(x, y)$, where $x$ and $y$ are spatial (plane) coordinates, and the amplitude of $f$ at any pair of coordinates $(x, y)$ is called the intensity or gray level of the image at that point.

When $x$, $y$, and the intensity values of $f$ are all finite, discrete quantities, we call the image a digital image. The field of digital image processing refers to processing digital images by means of a digital computer. Note that a digital image is composed of a finite number of elements, each of which has a particular location and value. These elements are called picture elements, image elements, and pixels.

Digital image processing is a subset of the image processing domain wherein the image is converted to an array of small integers, called pixels, representing a physical quantity such as scene radiance, stored in a digital memory, and processed by computer or other digital hardware. Digital image processing, either as enhancement for human observers or performing autonomous analysis, offers advantages in cost, speed, and flexibility, and with
the rapidly falling price and rising performance of personal computers it has become the
dominant method in use. Digital image processing allows one to enhance image features of
interest while attenuating detail irrelevant to a given application, and then extract useful
information about the scene from the enhanced image. Often the raw image is not directly
suitable for this purpose, and must be processed in some way. Such processing is called
image enhancement; processing by an observer to extract information is called image
analysis. Enhancement and analysis are distinguished by their output, images vs. scene
information, and by the challenges faced and methods employed. We divide our discussion of
digital image processing algorithms into image matching and image analysis.

Image matching defines the process of finding out the same or similar image modes from the
target image according to the known image template. With the development of technology,
image matching technique is very important in many applications including modern
spaceflight, military affairs, medicine and industry in latter day information processing.
Because of variant imaging conditions such as illumination more requirements of image
matching are raised.

Image analysis, by contrast, produces information that is much smaller in quantity but much
more highly refined than an image, for example the position and orientation of a template. In
many cases the output is just an accept/reject decision, the smallest quantity of information
but perhaps the highest refinement. Output behaviour and execution speed are generally
difficult and sometimes impossible to characterize. Image analysis algorithms are often a
vendor’s most important intellectual property.

Basically image matching is a subset of Geometrical Pattern Matching system which
describes the pattern matching in image. A well-designed Geometrical Pattern Matching
(GPM) system should be as easy to train as Normalized Correlation (NC) template matching,
yet offer rotation, size, and shading independence. It should be robust under conditions of
low contrast, noise, poor focus, and missing and unexpected features. Pattern recognition
time is application-specific, as is typical of image analysis methods.

GPM is capable of much higher pose accuracy than any template based method, as much as
an order of magnitude better when orientation and size vary. GPM is also capable of
providing detailed data on differences between a trained pattern and a template being
inspected. This difference data is also rotation, size, and shading independent. Simply
subtracting the template from an image and looking for differences does not work in practice,
since the variation in gray-scale due to ordinary and acceptable conditions can be as great as that due to defects. This is particularly true along edges, where slight (i.e. sub-pixel) mix-registration of template and image can give rise to large variation in grayscale. Variation in illumination and surface reflectance can also give rise to differences that are not defects, as can noise. A practical method of template comparison for inspection uses a combination of enhancement and analysis steps to distinguish shading variation due to defects from that due to ordinary conditions:

- A pattern recognition step (e.g. GPM) determines the relative pose of the template and image.

A digital re-sampling step uses the pose to achieve precise alignment of template to image.

- A pixel mapping step using histogram specification compensates for variations in illumination and surface reflectance.

- The absolute difference of the template and image is computed.

- A threshold is used to mark pixels that may correspond to defects. Each pixel has a separate threshold, with pixels near edges having a higher threshold because their gray-scale is more uncertain.

A blob analysis or morphology step is used to identify those clusters of marked pixels that correspond to true defects.

Image processing is a growth field covering a wide range of techniques for the manipulation of digital images with the low cost and high availability of cameras, more companies than ever before are able to utilize image processing software and reliable hardware in their research, production and quality control environments, as well as in their products. This technical supplement briefly describes some of the principle techniques available for image processing and how they may be used. Until recently, image acquisition and analysis was a costly procedure requiring specialist camera equipment, frame stores and high specification computers with processing accelerators. Modern computer speeds, cameras and a dramatic increase in storage capacity mean the cost of hardware has been significantly reduced, but with increased quality. A typical hardware and software solution for image processing is shown below:
Overview of Processing Methods

Image processing is a general term most often used to describe digital image manipulation in all forms, though it can be dissected into three main areas:

- Encoding
- Transformation
- Analysis

1.2.1 Encoding

Describes the methods by which an image can be represented as a series of binary digits. They say that an image is worth a thousand words and in terms of storage space that can be a conservative estimate. Encoding concentrates on reducing the size of the data that represents an image, for transmission or storage. As images are information rich, they can take a large
number of binary digits to describe them. For this reason, the main aim of encoding is size reduction, equating to higher speed of transmission and lower requirements for storage. Lossless methods concentrate on compact the binary data using encoding algorithms—the most commonly used example is WinZip. Good levels of compression can be achieved in this way, with no loss of clarity when the image is decompressed.

1.2.2 Transformation

It is the process of altering the image to make it more suitable for the intended purpose. Many people use software programs such as Photoshop to remove noise and correct light imbalances in their digital photos. Operations like these can also be performed in an automated environment to prepare an image for presentation or further analysis. When processing an image, there are a variety of methods available to get the desired results. Broadly they can be broken down into Histogram Operations, Arithmetic and Logical Operations, Convolution Operations, Derivative Operations and Morphological Operations.

Histogram Operations allow the balance of color or intensity within an image to be remapped. A histogram of a number of pixels against intensity is produced for an image (this describes, in graphical form, the probability distribution function (PDF) of the image). A function f is then generated to describe the mapping between the actual PDF of the image and the desired PDF. Each pixel is transformed by f to generate a new image. Using this method, contrast can be ‘stretched’ to increase the definition of features within an image. Another common application of Histogram operations is that of Equalization. This application sets all intensity levels to have an equal distribution and is used to compensate for differential lighting conditions on images before comparison.

Arithmetic and Logical Operations allow the combination of images by performing simple calculations on the values of corresponding pixels within two images. The resultant value of the calculation becomes the value of the pixel in the new image. The simplest operations of this sort are binary calculations with black and white images.

Convolution Operations are very powerful whilst being essentially simple, giving a means to remove noise and smooth an image. Convolution produces a new image based on the result of a convolution filter (also known as a kernel) being moved over the source image. The filter is a grid that is smaller than the source image. At the centre of the filter is the target pixel; the output value (O) of the pixel at the target position is the sum of the multiples of the filter values (K) by their corresponding source pixel values (I).
Or more generally:

By adjusting the values of the coefficients in the filter, it is possible to smooth the image and reduce the noise contained within it to different degrees.

Morphological Operations are used to pre-process images for recognition and classification. Dilation and Erosion are basic operations used to ‘grow’ or ‘shrink’ entities within an image. They are used as the basis for the more useful operations Opening and Closing. Opening is used to separate joined entities within an image, Closing is used to join entities within an image. These two operations are very important in pre-processing for image analysis as they can be used to differentiate between entities. The construction of the Opening and Closing filters requires structuring elements to be defined. The structuring elements tend to be application specific and require a priori knowledge of the problem domain meaning there is no ‘perfect’ solution for all situations, rather requiring experimental development for each problem type. Additional morphological operations such as skeletoning and seeding can be used to define the underlying structures in images, providing a form of analysis in their own right.

1.2.3 Analysis

It allows conclusions to be drawn from an image. Image analysis can take the form of identification of features, statistical analysis, classification, measurement or a combination of these. The methods used to analyze images vary greatly. The Analysis that can be performed on an image is dependent on the type of image and the information it contains. In the communication, the useful information in an image can be thought of extra information as the Noise. To perform accurate analysis of the Image it is important to remember that the image has been changed when carrying out pre-processing, therefore measurements recorded during the analysis may have to be adjusted to compensate.

Region Identification provides an automated way to identify regions within an image by correlation or criteria matching. Correlation involves comparing collections of points on two images for similarities; a correlation of one would be a complete match. Setting up such automated systems often requires a significant ‘training set’ of data so that thresholds and compensation values can be set. Region identification has been used successfully in several fields, including tissue typing in medical images and environmental surveying from satellite images.
Counting and Measurement are widely used within the biological community to perform automated measurements of cell size and number of cells. Techniques such as Opening and Closing are used to identify the individual cells before a series of measurements are taken. By identifying the cells beforehand the problem becomes quite simple and measurements can be taken from a ‘mask’ of the cell. The proximal axis, total area, boundary size and orientation can all be determined from a simple binary mask of the cell. Additional measurement accuracy can be gained by considering an intensity mask to overcome the quantization effect of using a digital image. An intensity mask considers the pixels neighboring those on the edge of an entity and assigns a proportional value based on the intensity. Using an intensity mask is often problematic to set up, as the relationship between the intensity and the size is non-linear and direction dependant due to lighting. The problem can become much more complex in environments when overlapping occurs and segmentation is non-trivial.

Object Classification is often used in combination with region identification to classify templates identified within an image. In simple cases the templates can be quite different, such as differentiating circles from squares-more complex tasks require the identification of defective goods or classification of very similar templates. Classification often involves identifying the unique features of a class and subjecting the image to analysis to look for those features, although there is another way. Neural networks (Artificial Intelligence) have been used to perform classification without specific identification of unique features but rather by ‘training’ with examples.

1.3 Type of Images

When using digital equipment to capture, store, modify and view photographic images, they must first be converted to a set of numbers in a process called digitization or scanning [3]. Computers are very good at storing and manipulating numbers, so once your image has been digitized you can use your computer to archive, examine, alter, display, transmit, or print your photographs in an incredible variety of ways.

Color pixels are 3-vectors (this is a fact of human physiology, not physics). Several representations, called color spaces, are commonly used for representing color. The simplest to produce is the red, green, blue space (RGB), although hue, intensity, saturation (HIS) may be more useful for image analysis. For the lower quality single-sensor cameras, the luminance, chroma1, chroma2 space (YCC) is sometimes used.
Monochrome pixels are usually 8 bits (256 gray levels), although 10-bit and 12-bit devices are sometimes used. Video signals tend to be noisy, however, and careful engineering is required to get more than 8 useful bits out of the signal. Furthermore, robust image analysis algorithms do not rely on photometric accuracy, so unless the application calls for accurate measurements of scene radiance, there is usually little or no benefit beyond 8bits. Wide dynamic range is more useful than photometric accuracy, but it is usually best achieved by using a logarithmic response than by going to more bits.

1.3.1 Pixels

Digital images are composed of pixels (short for picture elements). Each pixel represents the color (or gray level for black and white photos) at a single point in the image, so a pixel is like a tiny dot of a particular color. By measuring the color of an image at a large number of points, we can create a digital approximation of the image from which a copy of the original can be reconstructed. Pixels are a little like grain particles in a conventional photographic image, but

![Figure 1.2 An Image pixel](image)

arranged in a regular pattern of rows and columns and store information somewhat differently. A digital image is a rectangular array of pixels sometimes called a bitmap.

1.4 Types of Digital Images

For photographic purposes, there are two important types of digital images—color and black and white. Color images are made up of colored pixels while black and white images are made of pixels in different shades of gray.

1.4.1 Gray Scale Images

A gray scale image is made up of pixels each of which holds a single number corresponding to the gray level of the image at a particular location. These gray levels span the full range from black to white in a series of very fine steps, normally 256 different grays. Since the eye
can barely distinguish about 200 different gray levels, this is enough to give the illusion of a stepless tonal scale.

### 1.4.2 Color Images

A color image is made up of pixels each of which holds three numbers corresponding to the red, green, and blue levels of the image at a particular location. Red, green, and blue (sometimes referred to as RGB) are the primary colors for mixing light these so-called additive primary colors are different from the subtractive primary colors used for mixing paints (cyan, magenta, and yellow). Any color can be created by mixing the correct amounts of red, green, and blue light. Assuming 256 levels for each primary, each color pixel can be stored in three bytes (24 bits) of memory. This corresponds to roughly 16.7 million different possible colors. Note that for images of the same size, a black and white version will use three times less memory than a color version.

### 1.4.3 Binary Images

Binary images use only a single bit to represent each pixel. Since a bit can only exist in two states on or off, every pixel in a binary image must be one of two colors, usually black or white. This inability to represent intermediate shades of gray is what limits their usefulness in dealing with photographic images.

### 1.4.4 Indexed Color Images

Some color images are created using a limited palette of colors, typically 256 different colors. These images are referred to as indexed color images because the data for each pixel consists of a palette index indicating which of the colors in the palette applies to that pixel. There are several problems with using indexed color to represent photographic images. First, if the image contains more different colors than are in the palette, techniques such as dithering must be applied to represent the missing colors and this degrades the image. Second, combining two indexed color images that use different palettes or even retouching part of a single indexed color image creates problems because of the limited number of available colors.

### 1.4.5 Continuous Tone vs. Halftone Images

There are many different technologies used to reproduce photographic images. Some printing technologies such as laser and ink jet printers work by creating individual dots, each of which is one of a small number of solid colors. For example, a black and white laser printer can only print black or white pixels it cannot produce gray dots. To create the illusion of a
photographic image, these tiny dots must be clustered in different proportions to reproduce the different colors and gray levels in the image. This dot clustering process is called half toning or dithering and it can be done in many different ways.

![Continuous tone image, ordered dither halftone, error diffusion halftone](image)

**Figure 1.3  Halftone image**

Devices such as computer monitors, dye sublimation printers, and film recorders can create dots of any color without half toning, and these are called continuous tone printers because they can reproduce a continuous range of gray levels or colors.

### 1.5 Color Management

The process of getting an image to look the same between two or more different media or devices is called color management, and there are many different color management systems available today. Unfortunately, most are complex, expensive, and not available for a full range of devices.

#### 1.5.1 Hue

The hue of a color identifies what is commonly called “color”. For example, all reds have a similar hue value whether they are light, dark, intense, or pastel.

#### 1.5.2 Saturation

The saturation of a color identifies how pure or intense the color is. A fully saturated color is deep and brilliant as the saturation decreases; the color gets paler and more washed out until it eventually fades to neutral.

#### 1.5.3 Brightness

The brightness of a color identifies how light or dark the color is. Any color whose brightness is zero is black, regardless of its hue or saturation. There are different schemes for specifying a color's brightness and depending on which one is used; the results of lightening a color can vary considerably.
1.5.4 Luminance

The luminance of a color is a measure of its perceived brightness. The computation of luminance takes into account the fact that the human eye is far more sensitive to certain colors (like yellow-green) than to others (like blue).

1.5.5 Chrominance

Chrominance is a complementary concept to luminance. If you think of how a television signal works, there are two components—a black and white image which represents the luminance and a color signal which contains the chrominance information. Chrominance is a 2-dimensional color space that represents hue and saturation, independent of brightness.

1.6 Color Spaces

A color space is a mathematical system for representing colors. Since it takes at least three independent measurements to determine a color, most color spaces are three-dimensional. Many different color spaces have been created over the years in an effort to categorize the full gamut of possible colors according to different characteristics. Picture Window uses three different color spaces:

1.6.1 RGB (Red Green Blue)

Most computer monitors work by specifying colors according to their red, green, and blue components. These three values define a 3-dimensional color space call the RGB color space. The RGB color space can be visualized as a cube with red varying along one axis, green varying along the second, and blue varying along the third. Every color that can be created by mixing red, green, and blue light is located somewhere within the cube. The following images show the outside of the RGB cube viewed from two different directions.

The eight corners of the cube correspond to the three primary colors (Red, Green, and Blue), the three secondary colors (Cyan, Magenta, and Yellow) and black and white. All the different neutral grays are located on the diagonal of the cube that connects the black and the white vertices.

1.6.2 HSV (Hue Saturation Value)

The HSV color space attempts to characterize colors according to their hue, saturation, and value (brightness). This color space is based on a so-called hexcone model which can be visualized as a prism with a hexagon on one end that tapers down to a single point at the
other. The hexagonal face of the prism is derived by looking at the RGB cube centered on its white corner. The cube, when viewed from this angle, looks like a hexagon with white in the center and the primary and secondary colors making up the six vertices of the hexagon. This color hexagon is the one Picture Window uses in its color picker to display the brightest possible versions of all possible colors based on their hue and saturation. Successive cross sections of the HSV hex cone as it narrows to its vertex are illustrated below showing how the colors get darker and darker, eventually reaching black.

1.6.3 HSL (Hue Saturation Lightness)

The HSL color space (also sometimes called HSB) attempts to characterize colors according to their hue, saturation, and lightness (brightness). This color space is based on a double hex cone model which consists of a hexagon in the middle that converges down to a point at each end. Like the HSV color space, the HSL space goes to black at one end, but unlike HSV, it tends toward white at the opposite end. The most saturated colors appear in the middle. Note that unlike in the HSL color space; this central cross-section has 50% gray in the center and not white.

1.7 Introduction of Image Authentication:

In today's commercial environment, establishing a framework for the authentication of computer-based information requires a familiarity with concepts and professional skills from both the legal and computer security fields. Combining these two disciplines is not an easy task. Concepts from the information security field often correspond only loosely to concepts from the legal field, even in situations where the terminology is similar. For example, from the information security point of view, "digital signature" means the result of applying to specific information certain specific technical processes described below. The historical legal concept of "signature" is very broad. It recognizes any mark made with the intention of authenticating the marked document whether it is images [59].

In the field of Data Communication whether images or texts, the top priority issues is security. Classical cryptography is one of the ways to secure plain text messages. Cryptography addresses the necessary elements for secure communication namely privacy, confidentiality, key exchange, authentication, and non-repudiation but reveals the fact that communication is happening. Authentication takes cryptography a step to determine that the received data is authenticated or not.
1.8 History of Authentication:

It is probably not surprising that the inventors of writing, the Sumerians, were the inventors of an authentication mechanism. The Sumerians used intricate seals, applied into their clay cuneiform tablets using rollers, to authenticate their writings. Seals continued to be used as the primary authentication mechanism until recent times [58].

Use of signatures is recorded in the Talmud (fourth century), complete with security procedures to prevent the alteration of documents after they are signed. The Talmud even describes use of a form of "signature card" by witnesses to deeds. The practice of authenticating documents by affixing handwritten signatures began to be used within the Roman Empire in the year AD 439, during the rule of Valentinian III. The subscription a short handwritten sentence at the end of a document stating that the signer "subscribed" to the document was first used for authenticating wills. The practice of affixing signatures to documents spread rapidly from this initial usage, and the form of signatures (a hand-written representation of one’s own name) remained essentially unchanged for over 1,400 years. It is from this Roman usage of signatures that the practice obtained its significance in Western legal tradition. Now we have to discuss about the need of Authentication.

1.9 Need of Authentication:

New ways of verification are being developed daily to daily life of humans. Biometrics and other methods keep getting formulated and incorporated into the information technology industry. One interesting biometric authentication mechanism developed by a leading Japanese biometric company has found a way to get your DNA. You sign a document and it is digitally scanned, this document is then can be scanned in the future to verify its authenticity. Identity should be verified whenever there is doubt of the 3rd party being whom they say or when there is personal information at risk. Personal information like credit card details and banking information should be kept safe using digital certification as one of the security layers. Some banking institutions require that a user verifies his/her identity by validating identification credentials using a digital certificate. Important e-mail can also use digital signatures that verify the e-mail is from the originating sender and that it has not been tampered with. On many occasions users are unsure if they are dealing with reputable suppliers of institutions. Digital certification gives the user a sense of legitimacy and formalizes the process. It ensures that the company that the user is dealing with has a registration with a trusted authority and that the transaction is guaranteed to be done with the
intended parties. Now we will define the basic components of Digital Signature i.e. Encryption, Decryption and Hashing [3].

1.10 Basic Purpose of Encryption:

The purpose of this document is to provide a high level overview of encryption, and some of the standard techniques in which it is used to protect information. This includes such topics as encryption types, hashing, email, data transfer, remote access, key management, and securing portable devices like laptops, blackberry's, smart phones etc.

This document does not intend to recommend or promote any particular product or technology. Where specific brands or products are mentioned, they are used as examples only. There are many solutions available, and this makes no attempt to recommend any one solution over another [3, 4].

1.10.1 Encryption:

Encryption is simply defined as the process of transforming information (referred to as plaintext) using an algorithm (called cipher) to make it unreadable to anyone except those possessing special knowledge, usually referred to as a key."

1.10.2 Asymmetric vs. Symmetric Encryption:

"Are there different kinds of encryption? And how do I know if they are truly safe?" There are a few things about encryption that are important to understand. There are many books and other publications on the subject of encryption that go into much more detail, but here are the basic things that everyone should understand [5]:

1.10.3 Symmetric Encryption

Cryptographic solutions fall into one of two types. Symmetric encryption requires that each party who wants to decipher the encoded message (cipher text) must have the secret key (or password). So if person X wanted to send person Y his grandmother’s secret recipe for chocolate chip cookies, he could encrypt the message and use the key to do so. Now for Y to decrypt that message, she would need to know the password (or “key”). This works well when the intended recipients are few or when there are only a small number of secret messages that will be delivered over time. However, the management of secret keys for each message becomes complicated real fast especially as the number of intended recipients grows. This can become a key management [5].
1.10.4 Asymmetric Encryption:

Not long ago, some real smart people from MIT developed a system known as Public Key encryption. Simply described, each person has two keys. One is a secret (or private) key, and the other is a Public key. Public keys can be stored in a public database for anyone to see. If X wants to send Y a secret message, all he needs to do is to encrypt that message with Y Public key and send her the message. The trick is that only that Y’s Private key can decrypt the message. Now Y and X do not need to worry about sending the private keys to anyone, because they only need to know each other's Public keys, and can keep their private keys to themselves.

Additionally, X can encrypt a message with his Private Key knowing that anyone can decrypt it with his Public key. Now why would X want to do this? Because anyone who used X’s Public key to decrypt the message would know that only X could have encrypted it (remember that X used his Private Key that only he knows). This process is referred to as “Digitally Signing” and it is used to prove the origin of the message. In this manner, a digital signature is much more reliable than a written signature which can often be easily forged [5].

1.10.5 Hashing:

Another clever use of encryption is to create a unique fixed length string of characters from a selected text (such as the entire document). This string is called a HASH. If anything at all changes within that document (or file) then the HASH is completely different. This process is used to verify the INTEGRITY of the file. Using this hashing process, and by comparing the hashes of the original and received messages or files, Y could tell immediately if anything had been altered within the message.

A hash is a one way process and is impossible to reverse. That means their are no possible way could you create all of the data in the Library of Congress with only the hash that was created. Because of this, hashes can be posted publicly, or delivered without any concern of privacy or security short-comings. In fact, many people will post the hashes for their applications on the same page as the download so you can tell if the application has been altered in anyway. As you can see, encryption is very useful in ensuring the confidentiality and integrity of data as well as proving the point of origin (which serves to prove non-repudiation) [5].
1.11 Digital Signature:

1.11.1 Reasons for using digital security:

- It insures by means of verification and validation that the user is whom he/she claims to be. This is done by combine the users credential to the digital certificate and in turn this method uses one point of authentication.

- Digital certificates insure data Integrity giving the user piece of mind that the message or transaction has not been accidentally or maliciously altered. This is done cryptographically.

- Digital certificates ensure confidentiality and ensure that messages can only be read by authorized intended recipients.

- Digital certificates also verify date and time so that senders or recipients can not dispute if the message was actually sent or received.

1.11.2 Components of a digital signature:

- Your public key: This is the part that anyone can get a copy of and is part of the verification system.

- Your name and e-mail address: This is necessary for contact information purposes and to enable the viewer to identify the details.

- Expiration date of the public key: This part of the signature is used to set a shelf life and to ensure that in the event of prolonged abuse of a signature eventually the signature is reset.

- Name of the company: This section identifies the company that the signature belongs too.

- Serial number of the Digital ID: This part is a unique number that is bundled to the signature for tracking ad extra identification reasons.

- Digital signature of the Certification Authority: This is a signature that is issued by the authority that issues the certificates.

1.12 Signatures and Its Law:

A signature is not part of the substance of a transaction, but rather of its representation or form. Signing writings serve the following general purposes:
• **Evidence:** A signature authenticates writing by identifying the signer with the signed document. When the signer makes a mark in a distinctive manner, the writing becomes attributable to the signer.

• **Ceremony:** The act of signing a document calls to the signer's attention the legal significance of the signer's act, and thereby helps prevent "inconsiderate engagements.

• **Approval:** In certain contexts defined by law or custom, a signature expresses the signer's approval or authorization of the writing, or the signer's intention that it has legal effect.

• **Efficiency and logistics:** A signature on a written document often imparts a sense of clarity and finality to the transaction and may lessen the subsequent need to inquire beyond the face of a document. Negotiable instruments, for example, rely upon formal requirements, including a signature, for their ability to change hands with ease, rapidity, and minimal interruption.

The formal requirements for legal transactions, including the need for signatures, vary in different legal systems, and also vary with the passage of time. There is also variance in the legal consequences of failure to cast the transaction in a required form. The statute of frauds of the common law tradition, for example, does not render a transaction invalid for lack of a "writing signed by the party to be charged," but rather makes it unenforceable in court, a distinction which has caused the practical application of the statute to be greatly limited in case law.

### 1.12.1 Methods for Digital Signature:

**Method I:**

Digital signatures are created and verified by cryptography, that concerns itself with transforming messages into seemingly unintelligible forms and back again. Digital signatures use what is known as "public key cryptography," which employs an algorithm using two different but mathematically related "keys;" one for creating a digital signature or transforming data into a seemingly unintelligible form, and another key for verifying a digital signature or returning the message to its original form. Computer equipment and software utilizing two such keys are often collectively termed an "asymmetric cryptosystem."

The complementary keys of an asymmetric cryptosystem for digital signatures are arbitrarily termed the private key, which is known only to the signer and used to create the digital
signature, and the public key, which is ordinarily more widely known and is used by a relying party to verify the digital signature. If many people need to verify the signer's digital signatures, the public key must be available or distributed to all of them, perhaps by publication in an on-line repository or directory where it is easily accessible. Although the keys of the pair are mathematically related, if the asymmetric cryptosystem has been designed and implemented securely it is "computationally infeasible to derive the private key from knowledge of the public key. Thus, although many people may know the public key of a given signer and use it to verify that signer's signatures, they cannot discover that signer's private key and use it to forge digital signatures. This is sometimes referred to as the principle of "irreversibility" [5, 6].

Method II:

Another fundamental process, termed a "hash function," is used in both creating and verifying a digital signature. A hash function is an algorithm which creates a digital representation or "fingerprint" in the form of a "hash value" or "hash result" of a standard length which is usually much smaller than the message but nevertheless substantially unique to it. Any change to the message invariably produces a different hash result when the same hash function is used. In the case of a secure hash function, sometimes termed a "one-way hash function," it is computationally infeasible to derive the original message from knowledge of its hash value. Hash functions therefore enable the software for creating digital signatures to operate on smaller and predictable amounts of data, while still providing robust evidentiary correlation to the original message content, thereby efficiently providing assurance that there has been no modification of the message since it was digitally signed.

Thus, use of digital signatures usually involves two processes, one performed by the signer and the other by the receiver of the digital signature:

- Digital signature creation uses a hash result derived from and unique to both the signed message and a given private key. For the hash result to be secure, there must be only a negligible possibility that the same digital signature could be created by the combination of any other message or private key.

- Digital signature verification is the process of checking the digital signature by reference to the original message and a given public key, thereby determining whether the digital signature was created for that same message using the private key that corresponds to the referenced public key [6, 7].
1.13 The Creation of a Digital Signature:

In the simplest terms a digital signature is a stream of bits appended to a document. The purpose of a digital signature is to provide assurance about the origin of the message and the integrity of the message contents. When a message with a digital signature is transmitted and received, the following parties are involved:

- the signer who signs the document;
- the verifier who receives the signed document and verifies the signature;
- the arbitrator who arbitrates any disputes between the signer and the verifier if there is a disagreement on the validity of the digital signature.

Digitally signing a document begins with producing a summary of the document using mathematical functions known as hash functions. Some examples are Message Digest-5 (MD5), Secure Hash Algorithm-1 (SHA-1) and Message Digest-160 [10]. The output of a hash function, a document summary called the hash, always has the same number of bits e.g. 128 for MD5 and 160 for SHA-1, regardless of the length of the input document. It is obvious that different documents will produce different hashes. It is considered virtually impossible to have an identical hash even from two similar documents. The hash function is encrypted by the signer using his/her private key and forms the digital signature of the encrypted document.

The verifier receives both the document and the signature, calculates the summary of the document using the same hash function used by the signer. The signature is decrypted using the signer’s public key. The last step is to compare the decrypted summary with the one previously computed by the verifier from the document. If the two summaries are identical then the signature has been verified. The verifier is now sure of the identity of the signer and that the data was not been modified.
The figure below shows the signing process again in steps.

Let us suppose that Alice is the signer and Bob the verifier:

- $Y$ calculates the summary of the document, the hash;
- $Y$ encrypts the summary with her own private key to create the digital signature;
- $Y$ sends the digital signature and the document to $X$, the verifier;
- $X$ calculates the summary of the document, the hash;
- $X$ decrypts the digital signature with Alice’s public key and obtains a summary;
- $X$ compares the two summaries he has made;
- if they are equal $X$ is sure that the document was not modified and that $Y$ really did sign the document herself.

*Figure 1.4: A brief about digital signature*
1.14 Digital Signature Standard (DSS):

The standard specifies a Digital Signature appropriate for applications requiring a digital signature rather than written signature. The digital signature is a pair of large numbers represented in a computer as strings of binary digits. The digital signature is computed using a set of rules (i.e., the Digital Signature Algorithm (DSA)) and a set of parameters such that the identity of the signatory and integrity of the data can be verified. The DSA provides the capability to generate and verify signatures. Signature generation makes use of a private key to generate a digital signature. Signature verification makes use of a public key which corresponds to, but is not the same as, the private key. Each user possesses a private and public key pair. Public keys are assumed to be known to the public in general. Private keys are never shared. Anyone can verify the signature of a user by employing that user's public key. Signature generation can be performed only by the possessor of the user's private key. A hash function is used in the signature generation process to obtain a condensed version of data, called a message digest. The message digest is then input to the DSA to generate the digital signature. The digital signature is sent to the intended verifier along with the signed data (often called the message). The verifier of the message and signature verifies the signature by using the sender's public key. The same hash function must also be used in the verification process. The hash function is specified in a separate standard, the Secure Hash Standard (SHS). Similar procedures may be used to generate and verify signatures for stored as well as transmitted data. For use of digital signature we need some keys, their management, storage, usages etc [7, 8].

1.14.1 Key Management:

The proper management of cryptographic keys is essential to the effective use of cryptography for security. Ultimately, the security of information protected by cryptography directly depends on the protection afforded the keys. All keys need to be protected against modification, and keys used for data encryption and digital signatures need to be protected against unauthorized disclosure. All users and administrators need to be aware of their liabilities and responsibilities with respect to properly managing cryptographic keys, and need to understand the importance of keeping their cryptographic keys secure at all times. Managers and administrators should be prepared for the possibility of cryptographic key compromise and should have proper controls in place to adequately manage this type of situation. Whenever possible, organizations should utilize centrally stored and managed
cryptographic encryption keys. User generated and managed encryption keys require much greater vigilance of key lifecycle management procedures, as the organization could lose access to data if something were to happen to an employee or they were to leave the organization without divulging the keys used to encrypt sensitive or mission critical company data [8].

Passwords are generally used to protect (encrypt) cryptographic keys used for data encryption and digital signatures. Passwords are also keys, but are typically much simpler than the mathematically generated cryptographic keys, and sometimes the cryptographic encryption key is derived from a user entered password. As a result, data encryption systems are only as strong as their weakest link, which is usually the password used to protect the cryptographic keys. However, for data encryption, one can generally think of two types of applications, applications which encrypt data in transit using short-term keys and applications which encrypt data at rest using long-term keys. As one would expect, the key lifecycle management requirements will be different for short-term and long-term keys.

1.14.2 Key Generation:

The generation of keys is the most sensitive of all the cryptographic functions, and any inadequacies in the implementation of the key generation function can seriously undermine the integrity of the system. Security measures should be in place to prevent unauthorized disclosure, modification, or replacement of keys. Depending on the key management implemented, some applications will locally generate keys and other applications will generate and distribution keys from a central authority. In either case, company employees must maintain complete control of generated keys from the time of generation and throughout their lifecycle [8, 9].

1.14.3 Key Usage:

The use of cryptographic keys in applications may require special handling and protection practices to ensure the confidentiality and integrity of the keys. You will need to establish procedures to ensure the protection of sensitive cryptographic keys, and use of physical devices may be necessary to protect the keys. Authentication timeout features should be used to protect cryptographic keys from compromise or misuse and key recovery mechanisms should be implemented for cryptographic keys that are used to encrypt sensitive data [8,9].
1.14.4 Key Storage:

The proper storage of cryptographic keys is critical to ensure the confidentiality and integrity of keys used by end users, administrators, and automated applications. You will need to ensure that appropriate storage methods are used whenever storing cryptographic keys. The two most common ways to store cryptographic keys are software storage and hardware storage. Procedures are necessary to manage software stored cryptographic keys due to greater susceptibility from key exposure; if using passwords, strong passwords should be chosen for providing access to the keys. There should also be adequate access controls for the encrypted key data file, to prevent theft or misuse. If storing cryptographic keys in hardware storage, you will need procedures to provide for the initialization, distribution, replacement, and destruction (as necessary) for the hardware token [8, 9].

1.14.5 Key Archival:

As part of the key lifecycle, archival of cryptographic keys is vital to ensuring that any data encrypted with those keys is always available to authorized parties. Sensitive user or centrally managed cryptographic keys should be archived for time periods which align with the data which has been encrypted by those keys. This will ensure that encrypted data can always be decrypted for as long as the encrypted data exists in the company’s data archives. For keys associated with digital signatures, key archival times will determine the timeline for signature validation [8, 9].

1.14.6 Key Destruction

As part of the key lifecycle, expiration is an important part of key management. Reasonable lifetimes for keys should be determined, again aligned with the data encrypted with those keys. For keys associated with digital signatures, procedures are needed for the deactivation/revocation of keys so that data digitally signed prior to a compromise date (or date of loss) can be verified [8, 9].

1.15 Conclusion:

In this chapter we introduced the concept of data authentication used to protect intellectual property rights and rightful ownership. We presented required criteria for a digital signature scheme to be successful. We also identified areas e.g. biometric where data authentication can be applied as well as other areas also. We finally addressed the issue of whether the presence of a digital signature can prove ownership and concluded that this is only possible
through the use of a higher, controlling governing body where all original media can be registered.

A final conclusion of authentication can be successfully employed if the value of the digital media warrants the added expense. If not, it is an exercise in futility, and digital signature relies on the protection afforded a private signature key by the signer, and the procedures implemented by a Certification Authority. Digital signatures must be applied by a computer commanded by the signer. Also this chapter contains a brief discussion about the key management, key storage, key usages and their application for data authentication.

This chapter also gives the introduction to the field of image processing and its applications. In the first part image processing steps or methods are explain than the chapter explains about the digital image processing and gives a introduction about it. Entering into the field of digital image first of all the most important part of any image through which images are created pixels are explained. Than types of digital images such as gray scale images, color images, binary images, indexed color images etc are introduced. After this color management of the digital images such as hue, saturation, brightness, luminance, and chrominance are discussed. Color spaces which is the most important part of any digital image is introduce which are of three types that is used by the digital images are RGB (red green blue), HSV (hue saturation value), HSL (hue saturation lightness). The different applications of digital image is discussed in the next section applications are medical imaging, forensic, criminology, medical, remote sensing, military, and biology. After a complete introduction to the digital image processing and its concepts and introduction to image authentication, its techniques and its need is talk about in the next segment. The chapter ends with the problem statement and organization of the complete thesis is given.

1.16 Organization of the Thesis

In this report, the complete work regarding the thesis is described in seven chapters. The first chapter has been described above.

Outline of Dissertation:

The rest of this dissertation is organized as follows:

Chapter 2-Literature Survey: This chapter describes prior art and some important work done in the research areas of image authentication and basic introduction of low density parity check codes. The purpose of this chapter is to allow the reader to place the above
contributions in the context of previous work done in these areas and incorporates literature assessment in the field of Image Processing and Image authentication.

**Chapter 3-** Slepian-Wolf Coding: This chapter presents a framework for Slepian-Wolf coding on binary data. Also it contains the Encoding and Decoding using low density parity check codes and elaborates about Low Density Parity Check (LDPC).

**Chapter 4-** Image Authentication Technique: This chapter presents the proposed methodology used for image authentication. The aim of this chapter is the authentication of images from client to server side. Also it contains image authentication methods based on Slepian-Wolf coding and contains the proposed work, proposed algorithm and the tool.

**Chapter 5-** Experimental Result & Analysis: This chapter presents experimental results and analysis of image authentication based on Slepian-Wolf coding and simulates the result.

**Chapter 6-** Conclusions and Future Work: This chapter presents the conclusion and result discussion presented in this thesis. The chapter summaries the research contributions in this thesis and suggests future areas for research work in this field and concludes the thesis with future work. Thus a complete overview of performed work is categorized.