1
Introduction

In the digital era, advance development in network technologies and wide availability of internet increases the access to digital media through online services. As a result, sharing of multimedia data such as images, audio and videos has significantly increased and become widespread practice between the end users. At the same time, illegal copying, distribution, unlawful editing of multimedia data becomes an easy task due to the wide availability of sophisticated softwares. As a result, digital forgery and unauthorised use of multimedia data have reached up to a significant level and therefore authenticity and security of multimedia data is one of the major challenge to the information security [Menezes *et al.*, 1996] in today's scenario. The techniques that detect the small change in multimedia data are very important in many practical applications such as forensic investigation, photography artwork and medical database [Swaminathan *et al.*, 2006]. This motivates the research community to develop some standard and robust solutions that allow to examine the safety of exchanged data in terms of authenticity, confidentiality and integrity. Authentication is one of the prime issue in security of multimedia data that allows to determine whether the original multimedia content was altered or not in any way and helps to trace the owner of the multimedia data. Confidentiality is another important issue and plays a vital role in privacy protection of multimedia content. It essentially provides the protection to multimedia data from being accessed by unauthorized parties. In contrast, integrity of multimedia data allows degradation detection with a view to ensure that received multimedia data has not been corrupted or modified by the adversary. Many techniques have been developed to ensure the security of multimedia data by protecting the multimedia data from illegal use. Generally, these techniques can be broadly categorized into following three modalities which includes perceptual image hashing (PIH), encryption and digital watermarking.

The foremost technique namely *Perceptual image hashing (PIH)* [Monga and Evans, 2006] inspired by cryptographic hash function (CHF) [Ahmed and Siyal, 2007] has been introduced for multimedia data authentication which is useful in different image processing applications. Unlike, cryptographic hash, perceptual hash function has different objectives. In general, the cryptographic hashing is bit sensitive, usually key dependent and is used for data integrity. In this, a message is considered to be non-authenticate even if one bit of data has been changed. Thus, the integrity of the data can only be authenticated if all bits are unchanged [Menezes *et al.*, 1996]. This sensitivity is not appropriate for digital image because the content preserving modification such as JPEG compression should be tolerated for image processing application. Due to this sensitivity, the functions such as SHA-1 [Rivest, 1992] or MD-5 [PUB, 2008] are not suitable for the multimedia content authentication. Therefore, the concept of perceptual image hashing is more prominent for multimedia data.

Another important way to secure the multimedia data is through encryption. Encryption techniques [Dang and Chau, 2000] are very important in today's scenario where information security is of primary concern for end users. Multimedia encryption has a variety of applications in medical imaging [Kanso and Ghebleh, 2015], communication [Faragallah *et al.*, 2013], telemedicine [Dagadu and Li, 2018] and multimedia systems [Lian, 2008]. Encryption is used to transform the original data into cipher data which can be decrypted at the later stage to reproduce the multimedia

data. This technique is effectively used to: (1) secure the multimedia data during the transmission, and (2) protect the content during the storage of the multimedia data. Traditionally, many encryption techniques like Advanced Encryption Standard (AES) [Daemen and Rijmen, 2013], Data Encryption Standard (DES) [Coppersmith, 1994], International Data Encryption Standard (IDES) [Zhang *et al.*, 2010] and Rivest-Shamir-Adleman (RSA) [Rivest *et al.*, 1978] have been proposed for textual data encryption [Schneier, 2007]. But, these techniques are not suitable for multimedia encryption due to some inherent features like high correlation between neighbouring pixels and high redundancy. Other important reason is that the size of multimedia data is much large and therefore traditional cryptosystem required greater time to encrypt and decrypt the multimedia data. It mainly used for secure communication and to protect the content of the multimedia data rather than their ownership and copyright protection.

Recently, digital watermarking [Cox *et al.*, 2002] emerges as a prominent solution for aforementioned issues. This technique also address all pertaining issues related to Intellectual Property Right (IPR) and Digital Rights Management (DRM). In digital watermarking, an information carrying signal known as watermark is imperceptibly and robustly embedded into the multimedia data to produce the watermarked data. This watermarked media can be transmitted over the secure or insecure network for the communication purpose as well. At the receiver end, one can extract the embedded information to manifest ownership and copyright protection of multimedia data. During the transmission, watermark may undergo some intentional/unintentional distortions like white Gaussian noise, filtering, geometric attacks and compression. The robust watermarking has the ability to tolerate certain signal processing operations or attacks upto some extent which can occur during the life of watermarked media. Another concept called fragile watermarking is employed to verify the integrity of the multimedia object [Izquierdo, 2006]. Fragile watermarking has the ability to detect the small changes in the multimedia data and locate the corresponding regions, i.e., any alteration in the pixel values by the attacks or modification can be detected. Fragile watermarking is commonly used for the temper detection and to authenticate the multimedia object. Lastly, semi-fragile schemes are designed for the localisation of tempered media [Zhu *et al.*, 2007]. These schemes have the ability to resist unintentional changes or modification caused by common signal processing operations enabling this fact, it is possible to verify the content of the object with some perceptible distortion which may have arised due to transmission or non-malicious operation. However, semi-fragile schemes are used to detect the malicious or intentional attacks rather than validating the original media.

The solutions like image hashing, encryption and digital watermarking are essentially required for comprehensive image security. The basic framework, definition, characteristic and applications of these techniques have been described as follows:

1.1 PERCEPTUAL HASH FUNCTION

An image hash is a small binary string which can be obtained from the appropriate image hash function. It extracts the intrinsic feature of the image and generates the hash value. Robustness and security are two main aspects of a hash function. Robustness implies that the hash value of the perceptually similar image must be approximately same. In contrast, the security of hash function can be obtained by generating the hash value based on some secret key. The used secret key plays a vital role in the security as hash value cannot be easily counterfeited or obtained unbeknown to the correct secret key. In general, there are two types of hashing techniques. First one is cryptographic hashing technique which are useful for non-changing multimedia data such as passwords or files. These type of techniques are very sensitive to the input data, if single bit of data is changed then the corresponding hash value will be completely changed and as a result, data is considered to be non-authentic. On the other hand, content based hashing techniques are

applicable to common signal processing operations such as image scaling, enhancement, cropping, and JPEG compression. These operations alter the pixel values but do not change the perceptual content of the image. A perceptual hash function (PHF) can be designed using a number of key features, which can be described using the following components:
(1) A complete perceptual hash function ϕ. This function takes an input in the form of multimedia object and a secret key to produce a hash value of length ℓ.
(2) A perceptual hash verification function (PHVF) ϕ_V. This function takes two perceptual hash values as the input and compares them to decide whether the perceptual hash corresponds to same hash or different object.
(3) A perceptual hash key generation function (PHKGF) ϕ_{Key}. This function takes an input seed value to generate secret keys to be used in the hashing framework.

1.1.1 General framework for Perceptual Hash Function

Generally, a perceptual hash system consists of three components namely pre-processing, feature extraction and post processing. The first and the last step are not mandatory but essentially required to design a robust and secure hashing system. The security can be further enhanced by incorporating the randomness in any of step using some secret key. A generic framework of hash generation process is depicted in Fig. 1.1.

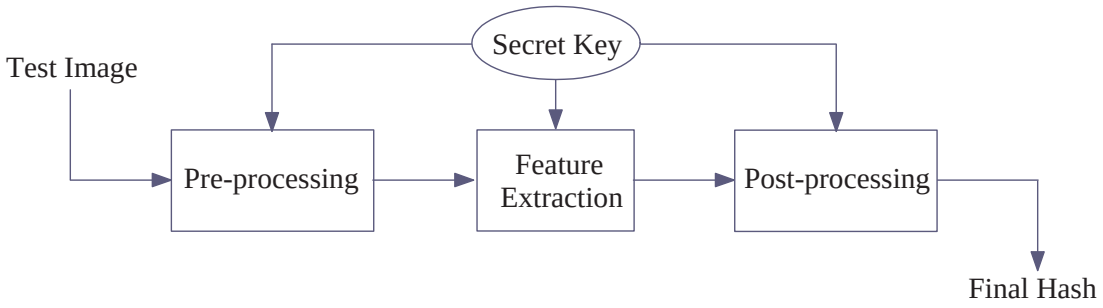

Figure 1.1 : A general framework of image hashing system.

The primary objective of pre-processing is to reduce the sensitivity of the feature extraction against distorted image due to common image processing operations. The pre-processing can be achieved using the image normalization [Hernandez *et al.*, 2011], resizing [Zhao *et al.*, 2013], order statistics filter [Mıhçak and Venkatesan, 2001a], low pass filtering [Neelima and Singh, 2016], down-sampling [Guo and Hatzinakos, 2007], and Gaussian blurring [Xiang *et al.*, 2007]. The pre-processing step essentially provides the base for the robust feature extraction.

In the next level, robust features are extracted from the pre-processed image. The main objective of this step is to estimate the unique features based on the characteristics of the image. These features should be robust to the content preserving operations (CPO) and distinctive enough to authenticate the test image. Moreover, these features should contain discriminative capability for content changing operations (CCO). A comprehensive classification of content preserving operations (CPO) and content changing operations (CCO) is shown in Table 1.1 [Han and Chu, 2010]. Therefore, feature extraction plays a lead role in the hash generation process. The robust features are extracted based on feature points detection [Lv and Wang, 2012; Monga and Evans, 2006], edge detection [Tang *et al.*, 2016], and dimension reduction techniques [Lv and Wang, 2012; Ouyang *et al.*, 2016, 2017; Zhao *et al.*, 2013]. The feature extraction techniques are well-known and publicly available and someone can predict and forge the image hash easily with the help of malicious manipulated or different images. As a result, the security of the authentication system can

be compromised. Therefore, a secret key is employed to incorporate the randomization to prevent these issues. The randomization can be achieved using the random block selection and projection techniques. The outcome of this stage termed as 'intermediate hash'. The intermediate hash may have some redundancies and other issues like image hash compaction, however, these issues can be resolved by applying compaction, binarization and quantization in the post-processing step. In general, error correcting codes (ECC) may be utilized for compaction and binariztion of intermediate hash. Finally, a key based permutation can be applied to enhance the overall security.

Table 1.1 : Content preserving operations and Content changing operations .

Content preserving operations	Content changing operations
- Noise addition	- Deletion of image objects
- Filtering operations	- Changing positions of the image objects
- Scaling	- Replacement the objects
- Rotation	- Changing the image background
- Gamma correction	- Changing the image texture
- Contrast adjustment	- Changing the image colors
- Brightness adjustment	
- Compression	
- Transmission error	

1.1.2 Characteristics of Perceptual Hash Function

An ideal perceptual hash function must possess some important properties. The effectiveness of these properties describe the efficiency of the hash functions. The main characteristics of a perceptual hash are illustrated using following notations:

- f: Input image;
- f_{ident}: Perceptually similar version of f under content preserving operations;
- f_{diff}: Perceptually different or maliciously modified version of f ;
- S_k: a Secret key involved in the hash generation;
- $\phi(\cdot)$ = Hash function;
- $Pr(\cdot)$= Probability;
- ε, δ = Parameters s. t. $\varepsilon, \delta \in (0,1)$;

1. *Perceptual robustness:* It refers to the ability that perceptually similar images must have the similar hash values considering the same secret key.

$$Pr(\phi(f, S_k)) \approx Pr(\phi(f_{ident}, S_k)) \geq 1 - \varepsilon, \qquad 0 \leq \varepsilon < 1 \tag{1.1}$$

Generally, digital image may undergo various types of content preserving operations like noisy operations, filtering operations, JPEG compression, rotation, and cropping. The modified images due to these operations are identical in human visual system. The perceptual robustness ensures that similar image generates the similar hash values.

2. *Discrimination:* It refers to the ability that perceptually different images should generate the different hash values.

$$Pr(\phi(f,S_k)) \neq Pr(\phi(f_{diff},S_k)) \geq 1-\delta, \qquad 0 \leq \delta < 1 \tag{1.2}$$

This property signifies that hash value corresponding to perceptually different images must be different.

3. *Unpredictability:* It indicates about equal distribution of the hash value.

$$Pr(\phi(f,S_k) = x) \approx \frac{1}{2^{\ell}}, \qquad \forall x \in \{0,1\}^{\ell} \tag{1.3}$$

where x represents the ℓ-bit binary hash value of image f considering the secret key S_k. For varying secret key, the estimated hash value is almost uniformly distributed to the all ℓ-bit.

4. *Compactness:* It indicates about the size of the hash value. Size of the hash should be much smaller than that of the original image.

$$Size(\phi(f,S_k)) << Size(f) \tag{1.4}$$

The shorter hash value requires smaller storage space and therefore, the image hash should be as small as possible. In addition, this property resolves many problems of large database like simplifying the searching problem.

5. *One-way:* Generally, hash generation processes are non-invertible.

$$\text{f} \rightarrow (\phi(f,S_k)) \tag{1.5}$$

It should be hard to detect the input image from the hash value so that the confidentiality of the input image remains protected.

6. Pairwise independence for visually different inputs f and g:

$$Pr(\phi(f,S_k) = x | (\phi(g,S_k) = y)) \approx Pr(\phi(f,S_k) = x), \qquad \forall x,y \in \{0,1\}^{\ell} \tag{1.6}$$

Further, security is another major concern for hash function, which can be addressed by restricting access control based on some secret key. The incorporation of the secret key in the hash generation process reduces the probability for an attacker to predict the hash value in order to generate the correct hash value. This can be achieved by employing the cryptographic hash function or different randomization process based on some secret key.

1.1.3 Classification of Image Hashing system

Generally, most of the perceptual hashing system (PHS) actively concentrate on the feature extraction stage and are used in different applications based on their actual requirements. The perceptual hashing system can be classified into following categories based on the feature extraction technique:

1. *Image hashing system based on statistical information:* The techniques belonging to this group mainly rely upon the statistical information. These techniques estimate the statistical features in spatial domain by utilizing the image statistics such as variance, histogram and higher order moments. These types of techniques preserve the robustness as the estimated features are more robust against various distortions including compression but show less distinctiveness.

2. *Image hashing system based on invariant information in transformed domain:* The techniques belonging to this group mainly rely on invariant information in transformed domain. These techniques compute the hash features by transforming the multimedia object from spatial to transform domain. The extracted information is robust enough to different kind of content preserving operations.

3. *Image hashing system based on low level features:* The techniques belonging to this group mainly rely upon salient features. These techniques estimate the perceptual hash based on local features such as corner, interest points, edges and blob detection. These types of features contain the local information that are robust to compression and geometric attacks.

4. *Image hashing system based on dimensionality reductions:* The techniques belonging to this group mainly rely upon different dimensionality reduction procedures. The dimensionality reduction refers to a process in which higher dimensionality data is mapped to a lower dimensionality data. In this approach, the perceptual hash is estimated based on dimensionality reductions while maintaining the main characteristics. In most of the cases, a low-rank approximation technique is employed on the multimedia object and retained coefficients are utilized in the hash generation. These type of techniques are robust and sensitive for noise addition and image blurring attacks.

1.1.4 Perceptual Hashing Application

Perceptual hash functions have rich set of applications in multimedia security. The main advantage of perceptual hash function is that the quality of multimedia data is not degraded unlike other available solutions. As a result, the estimated hash value of the original data remains similar after applying the content preserving operations. The perceptual hash functions have ability to identify the ill-legal copies of the original data, since these copies are usually lossy version of the data. The actual application lies in searching the internet for copyright violation, data spam detection and maintaining the different kind of database such as classified files, child pornography, etc. Despite of above applications, the perceptual hash functions have common usages depending upon the requirement as follows:

- *Content Authentication:* The perceptual hash plays key role in the content authentication of the digital data. The digital data may undergo the tempering operations such as unethical-editing, insertion, and deletion of the object due to easily availability of sophisticated softwares. The perceptual hash of possibly malicious manipulated digital data is compared with the original one to authenticate the integrity of the digital data. The outcome can be decided based on hypothesis testing. In other words, if distance between the hash value is less then some predefined threshold value then digital content is considered to be authentic in content authentication. An illustration of hashing in content authentication is depicted in Fig. 1.2.

- *Content Identification:* The main motivation of perceptual hash is to identify a similar object using the valid secret key. For example, if some user uploads some video or image onto some website, then there may be chance that another user may download the image or video without any authorization. This may lead to considerable financial loss to owner of the website. Therefore, primary objective of the hash function is to identify the image or video, even the data has suffered from the insignificant distortions due to transmission, compression or other issues and prevent illegal use of the digital data. Hence, the perceptual hash function can also be used for content protection as well as copyright protection.

- *Efficient indexing and retrieval:* The perceptual hash function has greater ability to manage the large-scale multimedia database and provide excellent performance to the query-retrieval services. This can be achieved through two phases: (1) Database creation, and (2) Content

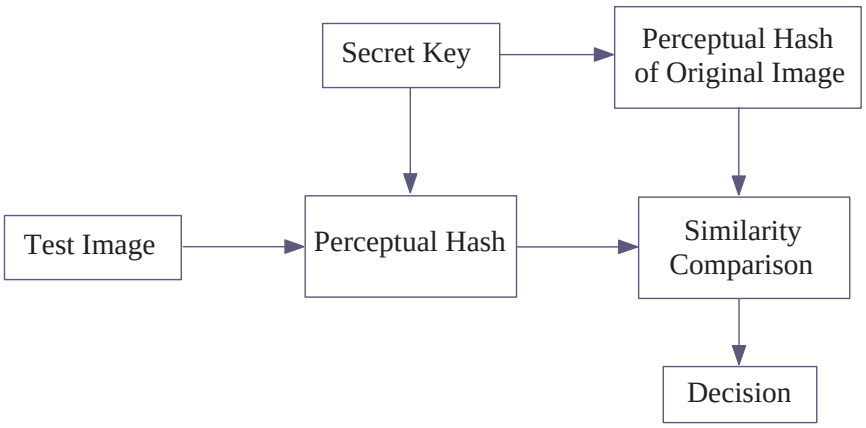

Figure 1.2 : A framework of image authentication using perceptual hash function.

Identification. In database creation phase, the perceptual hash of media object is estimated and stored in database, which can be used in the later stage. In contrast, content identification essentially process unidentified data followed by the perceptual hash computation. The perceptual hash is finally compared with hash value stored in the database. In principle, if the hash value is matched or almost similar, then other details of the media can be obtained from system. An illustration of the same is depicted in Fig. 1.3.

The performance of perceptual hash system can be evaluated using the different objective metrics. These objective metrics help to determine the robustness and discrimination capability of the hash system. Generally, the estimated image hash is stored in the database as an index and used for identification purpose, if a query is received. The outcome is 'similarity score', which plays decisive role in image authentication and identification. Let $H_1 = (h_1, h_2, h_3, \cdots h_n)$ and $H_2 = (g_1, g_2, g_3, \cdots g_n)$ are the image hashes of length L corresponding to image f_1 and f_2. The most popular objective metrics used in the image hashing literature are depicted in Table 1.2.

Table 1.2 : Different objective metrics for perceptual image hashing.

Hamming Distance	$HD(H_1,H_2) = \sum_{i=0}^{n} \lvert H_1(i) \oplus H_2(i) \rvert$
Normalized Hamming Distance	$NHD(H_1,H_2) = \frac{1}{n} \sum_{i=0}^{n} \lvert H_1(i) \oplus H_2(i) \rvert$
Euclidean Distance	$ED(H_1,H_2) = \sqrt{\sum_{i=0}^{n} (H_1(i) - H_2(i))^2}$

1.2 IMAGE ENCRYPTION

Cryptography or Encryption [Boneh and Franklin, 2001] is a fundamental branch of science that deals with encryption and decryption of data. It is widely used for secure communication between the end users. In this approach, a meaningful message (plain-text) is changed into un-meaningful message (cipher-text) such a way that only legitimate recipient can reproduce the original message. This is known as encryption process. On the other hand, the process of transforming the un-meaningful message (cipher-text) into meaningful message (plain-text) is known as decryption process. An illustration of general framework is shown in Fig. 1.4.

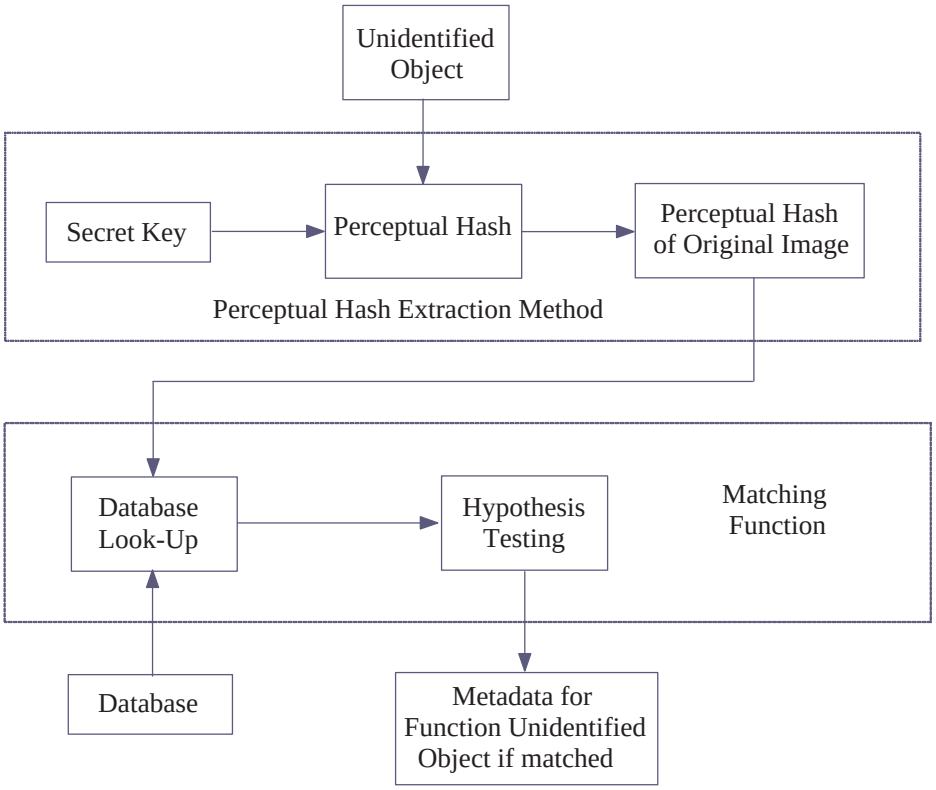

Figure 1.3 : A framework of content identification using perceptual hash function.

1.2.1 Characteristics of cryptographic system
1.2.2 Types of Image Encryption Systems

Generally, a secure cryptographic system can be designed using some secret keys. This is due to the fact that different secret keys produce different cipher-texts. More precisely, the security of a cryptographic system mainly depends upon two things: secrecy of the secret key and robustness of cryptographic algorithm. In literature, there are two types of cryptographic systems: secret key (symmetric), and public key (asymmetric) cryptographic systems. The basic mechanism of secret and public key cryptographic systems is described in Figs. 1.5-1.6. In secret key cryptographic systems, sender and receiver use the same secret key to encrypt and decrypt the data and the respective key remains secret between them. However, a pair of keys are utilized for the same purpose in public key cryptographic systems. A publicly available key is used to encrypt the data whereas a private or a secret key is used to decrypt the data. The main advantage of public key cryptographic systems is that anyone can encrypt the data with the help of the public key but decryption of the data is not possible without actual knowledge of private key. However, it is difficult to estimate the private key from the public key and an authorised person with a private key can decrypt the information.

The primary goal of cryptography is to protect the confidentiality of the data by enabling new features in information security. However, there are some important characteristics which can be described as follows:

- *Confidentiality:* Confidentiality is most important characteristic of the information security. It can be practised during the storage of the data as well as transmission of data. The privacy of the data is preserved against different kind of attacks.

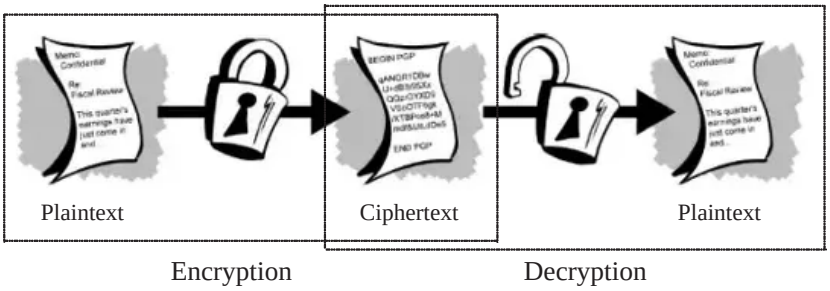

Figure 1.4 : A General framework for image encryption and decryption.

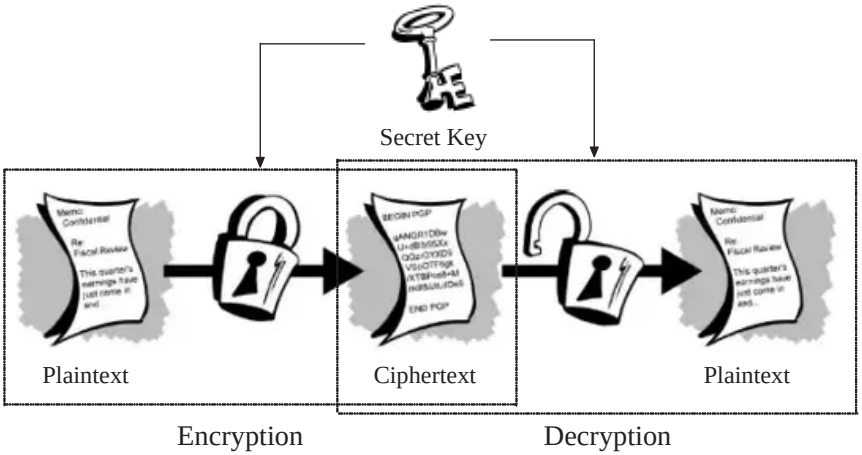

Figure 1.5 : A typical framework of secret key cryptosystem.

- *Data Integrity:* It ensures the content protection from unauthorized modification or alternation of data. This includes the operations such as insertion, deletion and unusual manipulations of data. The data integrity can be assured by detecting the data manipulation from unauthorized entities. This can be accomplished through hash functions.

- *Authentication:* It refers to attributing the identification of information/message and the respective entities involved in the communication, i.e, it provides the authentication not only for sender and receiver in connected oriented communication, but also for transmitting the information itself.

- *Non-repudiation:* Non-repudiation is an assurance that prevents repudiation by sender or receiver of the data. Non-repudiation is a similar characteristic like authentication and its implementations can be shared often with the same primitives. This can be illustrated using the non-repudiation of public key signature in which one specific entity has the ability to produce the signature.

In literature, a variety encryption techniques have been proposed based on conventional cryptographic system such as DES [Coppersmith, 1994], AES [Daemen and Rijmen, 2013], RSA [Rivest *et al.*, 1978], and IDEA [Zhang *et al.*, 2010]. These cryptographic systems can control the unauthorized access of static (not real time) multimedia data. But for real-time streaming, the encryption of

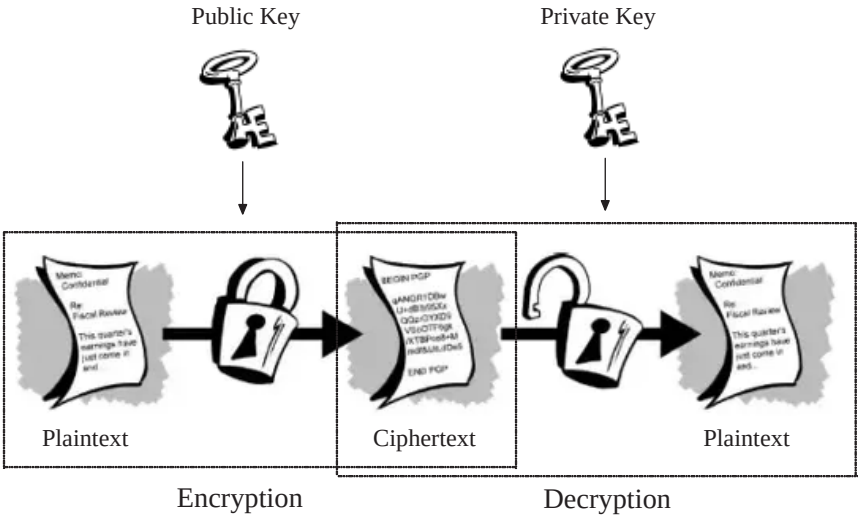

Figure 1.6 : A typical framework public key cryptosystem.

the multimedia is hard to accomplish using these techniques. Moreover, the encrypted multimedia data may be compressed to fulfil the requirement of the channel bandwidth during the transmission. The transmission effect like compression, could not be handled by these standard cryptographic systems and therefore they are not suitable for multimedia encryption. This is mainly due to the fact that the multimedia data is having higher redundancies and large volume. Therefore, efficient multimedia encryption techniques are developed to provide the higher security for data transmission and other real time applications. There are some important issues which need to be considered at the time of designing the multimedia encryption technique. The details of the issues are given as follows:

1. Multimedia encryption and decryption take considerable amount of time in encoding and decoding process due to large size of multimedia files. Therefore, significant portion of multimedia data can be encrypted based on their visual importance [Xiao *et al.*, 2016].

2. There are various standard methods available for the encryption of still images [Chuang and Lin, 1999; Shaohui *et al.*, 2003]. However, applications of these methods on raw images increase the size of bandwidth significantly, which may effect the image communication.

3. Multimedia and text encryption techniques have their own security criteria, depending upon the requirements and applications, but security levels for multimedia data is not high as for the textual data. The security levels should be optimized in a way that the encrypted multimedia can not be compromised during media streaming and limited lifetime.

From the above discussion, it can be observed that the requirement of multimedia data encryption is different than the textual data for a cryptographic system [Furht, 1998, 2004; Lian, 2008]. The multimedia cryptographic system requires perceptual security as well as cryptographic security. The perceptual security means that one can not predict the actual overview and structural information of the original multimedia from the encrypted multimedia data. This fact can be visualized from Fig. 1.7. Figure Fig. 1.7(a,e) show Lena and Cube images whereas Fig. 1.7(b-d) and Fig. 1.7(f-h) show the encrypted version using the raster, Zig-zag and Hilbert scan. From the figure, it can be seen that Fig. 1.7(b,f) reveal the information about the original data whereas from the Fig. 1.7(b,f), it is hard to predict the overview of the original data. This is due to the fact that zig-zag and Hilbert

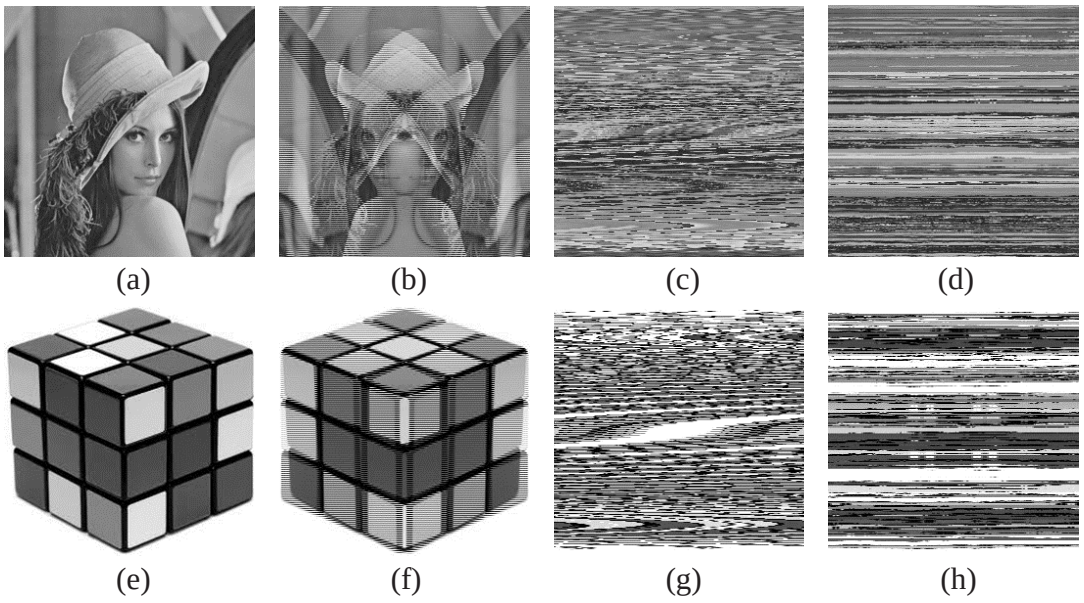

Figure 1.7 : a,e) Original images; b,f) Encrypted Images using raster; c,g) Encrypted Images using Zig-zag scan; d,h) Encrypted Images using hilbert scan.

scan disturbed the adjacent relation between the neighbouring pixels whereas raster scan is unable to accomplish the similar task. In general, the perceptual security can be analyzed based on objective and subjective evaluation as follows:

1. **Objective Evaluation:** The evaluation using objective metric is one of the best suitable method to measure the performance of a system using simulation. There are number of objective metrics available to assess the performance but some of the quality metrics are very popular among them and are widely used for encrypted multimedia. The metrics such as signal-to-noise ratio (SNR) and peak signal to noise ratio are frequently used quality metrics but PSNR got more attention by research community and can be defined as follows:

$$PSNR = 10\log_{10}\left(\frac{Max_f^2}{RMSE^2}\right) \tag{1.7}$$

$$RMSE = \sqrt{\frac{1}{MN}\sum_{i,j}\left[f_o(i,j) - f_e(i,j)\right]} \tag{1.8}$$

where f_o and f_e denote the original and encrypted multimedia respectively. The PSNR value indicates the characteristics of the encryption technique. The lower PSNR value describes the better encryption quality of multimedia data.

2. **Subjective Evaluation:** According to Shannon model "communication theory of secrecy system", the strength of encryption algorithm can be evaluated using the statistical analysis which essentially defines the strength of the perceptual security. The statistical analysis of the encrypted media provides the information regarding the merit and demerit in security with respect to statistical attacks. This can be achieved by histogram and correlation analysis of adjacent pixels of the encrypted multimedia data.

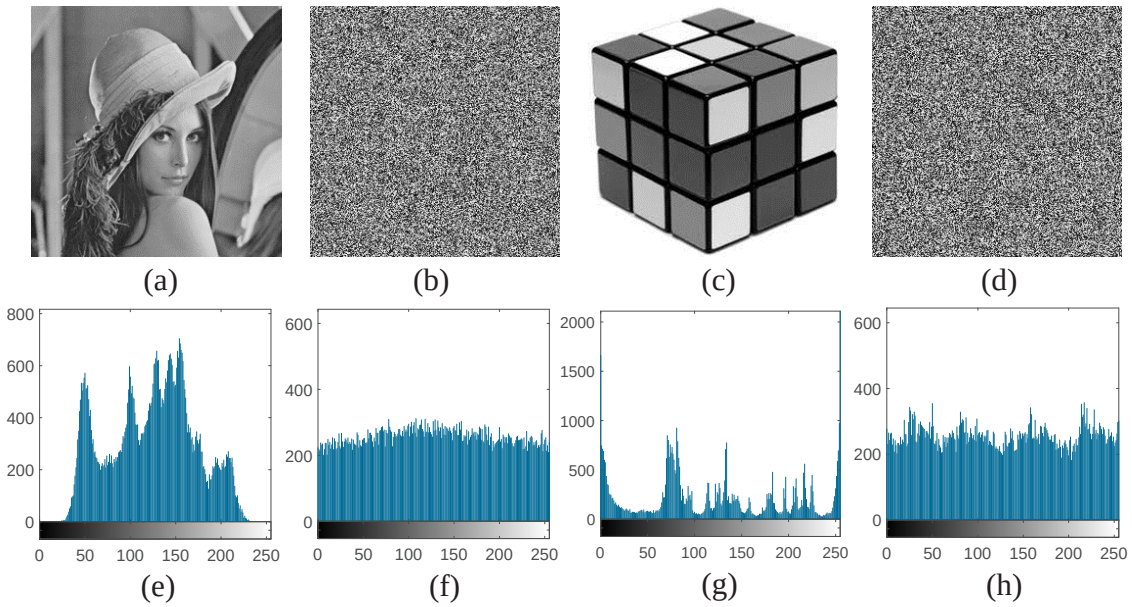

Figure 1.8 : Histograms of the original images and their encrypted versions.

- *Histogram Analysis:* The histogram analysis [Brunelli and Mich, 2001] is used to compare the histogram of the original and encrypted data. It essentially describes confusion and diffusion property in encrypted data. The histogram of original and encrypted images are depicted in Fig. 1.8. From the figure, it can be seen that the histogram of the encrypted images is almost uniform and preserves the significant difference from the histogram of original image.

- *Correlation Analysis:* The correlation analysis plays a decisive role to measure the similarity between the images. For multimedia applications, it determines the similarity between the original and encrypted data. There are three ways to analyze the correlation between the pixels. In initial step, the correlation among the adjacent pixels are computed in horizontal and vertical directions. In addition, the correlation between the adjacent pixels are also determined in diagonal direction to measure the randomness introduced by the algorithm. The correlation coefficients [Hardoon *et al.*, 2004] can be calculated as follows:

$$CC(\rho) = \frac{cov(x,y)}{\sqrt{var(x)}\sqrt{var(y)}} \tag{1.9}$$

$$cov(x,y) = E[(x - E[x])(y - E[y])] \tag{1.10}$$

where x and y denote two adjacent pixels present in the image.

1.3 WATERMARKING SYSTEM

Recently, digital watermarking emerges as a prominent solution to overcome all security issues of multimedia data. [Katzenbeisser and Petitcolas, 2000]. Basically, watermarking used to describe the conveyed information in a hiding manner [Cox and Miller, 2002]. The concept of watermarking has been used in various forms and can be traced back to the 13^{th} century. Watermarks were made on the paper using the thin translucent layer and wire to specify the brand and producer name as the registered trade-mark for the product in Italy. The watermarks began to be used for different prospective during 16^{th} century and by 18^{th} century, they had been used as

anti-counterfeiting measures on currencies and other documents. Thereafter, digital watermarking gained the considerable attention and researchers actively contribute intense research work in this area in past few years.

Digital watermarking is a process in which an ownership sign or information is imperceptibly embedded into multimedia signal and can be extracted at later stage for verification or authentication purposes. The embedded information into the signal is called a digital watermark and the signal carrying the information is know as the host signal. The watermarked signal is usually stored or transmitted between the different channel for communication purposes. During the transmission, if the watermarked signal is unmodified then watermark information is still present in the signal and can be extracted. For a robust watermarking algorithm, the watermark information can also be recovered even the watermarked signal is modified strongly. The quality of original and extracted watermark signal should be perceptually similar. The perceptual transparency can be ensured using some perceptibility criterion which can be adaptive or fixed. The watermark information remains confined with the watermarked signal to protect the ownership. If watermarked signal is copied then copied version also contains the embedded information present in the signal. Generally, an ideal watermarking algorithm has good perceptual quality as well as robustness.

1.3.1 General Framework of Watermarking System

Generally, a watermarking system [Arnold *et al.*, 2003; Cox *et al.*, 2002; Muharemagic and Furht, 2004] comprise of two components: (1) Watermark embedder, and (2) Watermark extractor/detector. The watermark embedder requires three inputs that include original host media, watermark and a secret key to produce the watermarked media. The embedder inserts a machine-readable watermark into the host media by modifying the media object which is generally controlled by a private key. The private key is assigned to ensure the security during the complete procedure. In contrast, secret key and watermarked media are produced as the output, which can be utilized for extraction purposes. From the watermarked media, the embedding of watermark can be confirmed by the watermark detector or extractor. However, the terms 'extractor' and 'detector' do not have same meaning. The former is employed to extract the embedding information from the watermarked media, whereas the latter defines the existence of the watermark in the media. A typical watermarking system can be observed from Fig. 1.9.

In embedding process, encoder combines the host media (H), watermark code (W), a secret key (S) and an embedding algorithm to create the watermarked media (H_W). Mathematically, the embedding process can be expressed as follows:

$$H_W = E_M\Big(H, W, [S]\Big) \tag{1.11}$$

where $E_M(\cdot)$ is the embedding function and $[\cdot]$ indicates the optional nature of the parameters and can be used to extend the watermark embedder. The considering key (S) may be disclosed privately or publically. In extraction process, the watermarked media, secret key, original media and extraction algorithm is required to extract the watermark (W_{ext}). Mathematically, the extraction process can be illustrated as given below:

$$W_{ext} = E_X\Big(H_W, [H], [S]\Big) \tag{1.12}$$

where $E_X(\cdot)$ represents the extraction function. In detection process, detector is used to confirm the existence of watermark and can be summarized as:

$$D\Big(H_W, [H], W, [S]\Big) = \begin{cases} 0 & \text{There is no watermark} \\ 1 & \text{There is a watermark} \end{cases} \tag{1.13}$$

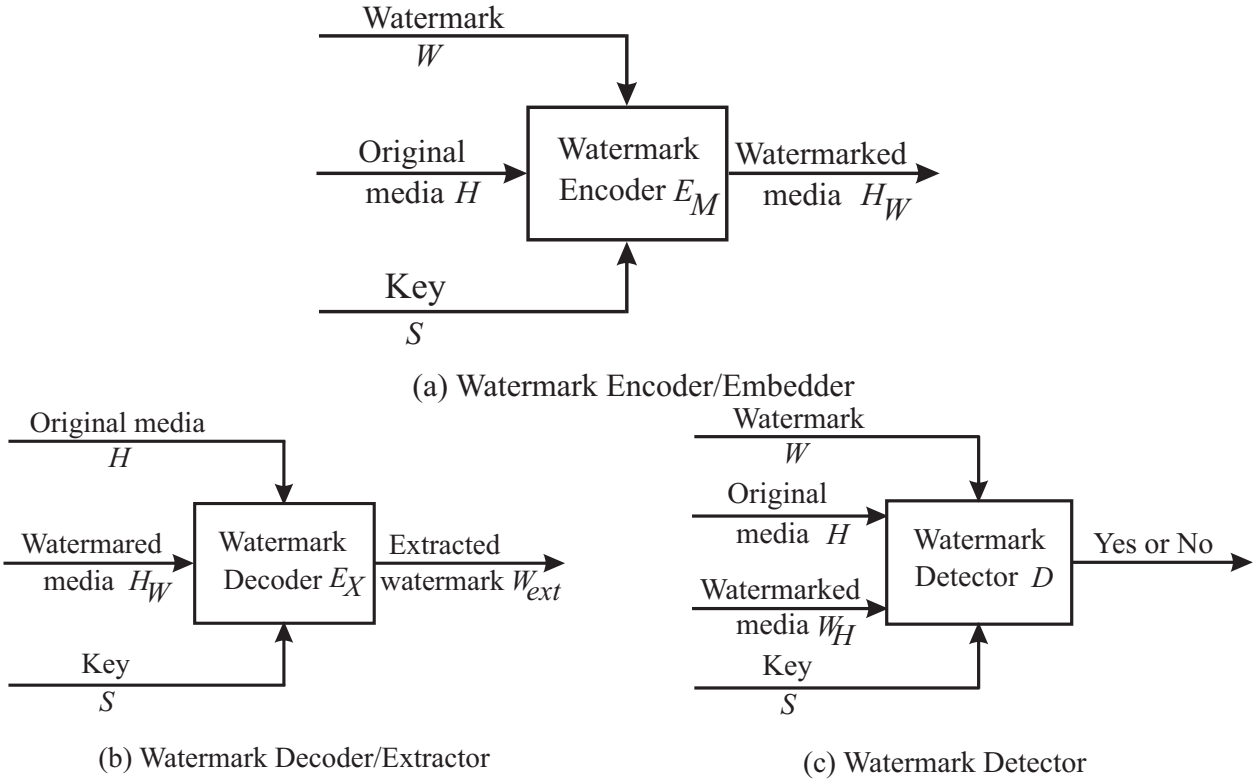

Figure 1.9 : A general framework for watermarking system.

where $D(\cdot)$ denotes the detection function.

1.3.2 Characteristic of Watermarking System

An ideal watermarking system should posses some important characteristics. They mainly depend upon the requirement of the intended application. The main characteristics of the watermarking system can be summarized as follows:

1. *Robustness:* The term 'robustness' refers to resilience ability of the embedded watermark against removal by standard data processing operations. A watermark is considered to be robust if it can survive under distortion introduced by intentional or unintentional attacks. Typically, it is almost impossible to design a perfect watermarking system which is robust against all potential attacks. This property mainly depends on the working medium (image, audio, video) and the type of application. Therefore, it may be possible that a watermarking system may require resistance only for some of the attacks.

2. *Imperceptibility:* Imperceptibility of the watermark can be described in terms of perceptual transparency. The perceptual transparency is the degree of invisibility of the watermark in the watermarked signal, i.e., a watermark is said to be truly imperceptible to human eye if no perceptual difference can be detected between watermarked and the original media.

3. *Capacity:* Capacity is referred as maximum amount of information that can be hided in the host media without noticeably reducing perceptual quality. Generally, a technique with higher capacity is more desirable because, more data information can be embedded in the host multimedia.

4. *Security:* Watermark security refers to the ability to resist against the hostile attacks. The hostile attacks indicate the process that can destroy the requirement of the watermark. There are different kind of hostile attacks and hence each watermark application requires its own type of security. The security of a watermarking system can be compromised, if an intruder have raw information about the secret key and corresponding system. In this scenario, the attacker will have the details about the specified watermark embedding locations or watermark bits with respective frequencies. The prediction of the secret key becomes possible by analyzing the similarity between the characteristics of a set of watermarked media.

The above characteristics generally oppose each other and hence a trade-off is essential between them. A possible trade-off is illustrated in Fig. 1.10. Usually, imperceptibility is the basic and most important requirement for a watermarking system. This is directly connected to robustness and capacity. The amount of embedding information may lead to some artifacts or less imperceptibility between watermarked and original multimedia. However, improvement in imperceptibility may lead to less robustness and conversely. Both of the factors depend on the capacity, i.e., if the capacity is increased then robustness might be increased but imperceptibility will be compromised. Therefore, these constraints can be optimized significantly for an efficient watermarking system.

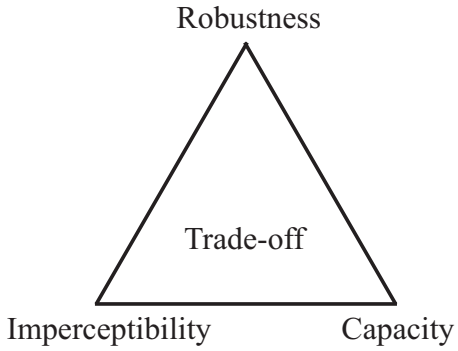

Figure 1.10 : Trade-off between the different characteristics of watermarking system.

1.3.3 Watermarking Application

Digital watermarking have diverse applications in the field of image security [Podilchuk and Delp, 2001]. The watermarking techniques have different requirements depending upon the target applications. The main application of watermarking includes copyright protection [Lin and Chen, 2000; Wang and Lin, 2004], fingerprinting [Alattar *et al.*, 2006; Cox *et al.*, 2000; Kirovski *et al.*, 2002], authentication [Kundur and Hatzinakos, 1999; Lu and Liao, 2001; Wu and Liu, 1998], copy control [Miwa *et al.*, 2001; Morito *et al.*, 2001; Petitcolas, 2003] and broadcast monitoring [De Strycker *et al.*, 2000; Depovere *et al.*, 1999; Kalker *et al.*, 1999]. The details of the watermarking applications are summarized as follows:

1. *Copyright Protection:* Copyright protection is one of the leading application of digital watermarking. It helps in the identification of copyright owner and protects the property rights in content distribution. For this purpose, the metadata that contains the copyright information (watermark) is embedded into cover image to protect the right of ownership. The detection of watermark must be possible despite of the several image processing operations including geometric operation s and, image compression. The successful watermark extraction or detection can identify the ownership.

2. *Fingerprinting:* The main objective of fingerprinting is to trace the source of the illegal copies

and locate the origin of the piracy. This can be achieved by assigning a unique information to each individual copy of digital content. If same information is embedded in all the copies then it may arise the problem. It is due to the fact that if one of the legal customer among all the customers is selling the content illegally then the tracing of that customer who is redistributing copies illegally becomes very difficult. This problem can be resolved by customization of each distributed copy for each legal recipient. For example, a serial number related to customer identity can be embedded into data. It enables the intellectual properties rights for creator to identify the legal customer who has breached the licence agreement by distributing data illegally to other parties.

3. *Authentication:* Due to advance development in computer technology and freely availability of sophisticated multimedia software, the manipulation of multimedia data becomes easier and authenticity of the originality data becomes a challenging task. Data authentication is required in the forensic investigation where integrity of the evidence has vital importance and needs to be protected. Digital watermarking is one of finest solution for these issues. The main advantage of watermarking is that embedded signature or watermark stays with image data and cannot be destroyed easily. In addition, it does not require the additional space to store the watermark information. The authenticity of the data can be determined using the robust and fragile watermarking system. A watermarking system can be designed to make the watermark sensitive to intentional attacks and robust to common signal processing noise such as compression and channel noise. The modification due to these operations will not affect the integrity of the data. On the other hand, a fragile watermark is used to determine whether the content is altered or not and essentially provides the altered locations..

4. *Copy control:* The goal of the copy control application is to prevent the customer to make the illegal copies of copyrighted content. This can be achieved by using a watermark that directly controls digital recording device to prevent the copying form copyright content where the watermark can be considered as the copy prohibit bit. The watermark detector application in the recording device uses this bit to determine whether the content is copyrighted or not. The watermark contains the copy control information which includes copy control not asserted, copy-no more and copy never. The detector in the camera act according to copy control information.

5. *Broadcast monitoring:* Digital watermarking technology is highly valuable to content provider and broadcaster. The broadcast monitoring system aims to identify with granular precision when and where the content is broadcast which can be recognized by the embedded watermark.By embedding the watermark in video, audio, or commercial advertisement at the time of production or broadcast, an automated system can monitor and verify whether the content is broadcasted according to the contract. There are various organizations working in the fields of broadcast monitoring. Their primary responsibility is to make sure that they receive the airtime for broadcasting as they purchased in the agreement.

1.3.4 Attacks on Watermarking System:

Generally, an ideal watermarking system should be robust enough against different kinds of attacks. An attack is a type of processing that diminish watermark detection or information conveyed by the watermark in the communication process. The processed data is considered as attacked watermarked data. There are various type of attacks and can be broadly classified into two categories: (1) Intentional; and (2) Unintentional attacks. In intentional attacks, the watermark information is illegally extracted, removed or destroyed by the attacker in different situations. However, in the unintentional attacks, the attacker do not modify or remove the watermark unlike the intentional attacks. There may be other various reasons like compression, transmission delay that

affect or destroy the watermark information. These attacks can be further classified into different groups based on several other criteria described in the literature [Kutter and Petitcolas, 1999; Kutter *et al.*, 2000; Licks and Jordan, 2005; Petitcolas, 2000; Petitcolas *et al.*, 1998; Voloshynovskiy *et al.*, 2001]. A brief classification of attacks are summarized as follows:

1. *Active Attacks:* In these type of attacks, the attacker/intruder attempts to remove or insert the unauthenticated watermark to arise the disturbance in the watermarking system. The main goal is to alter the information to make watermark undetectable. An attack is considered to be successful, if the watermark is unreadable or watermark detection is not possible anymore. These attacks are critical for the application such as copyright protection, owner identification and fingerprinting.

    - *Noise Addition:* The addition of noise with a given distribution such as Gaussian, uniform degrade the quality of the image as well as the hidden watermark information. The additive noise may be introduced due to transmission error or Analog-to-digital (A/D) converter. However, an attacker may use a image dependent mask with some unnoticeable power for noise addition. As a result, it will increase the threshold value at which the correlation response is detected.

    - *Compression:* This operation lies under the category of un-intentional attack and is widely used in many multimedia applications. There are various compression techniques available, but JPEG compression is popular among all. In JPEG compression, quantization and entropy encoding are two main steps but information loss occur during the quantization process and hence the same occurs in the watermark.

    - *Geometrical Attacks:* The geometric attacks are not aimed to remove the embedded watermark, but intends to either distort it or disable it in the detection process. The detector can recover the information under the situations where perfect synchronization is established. However, the computation cost of this synchronization may be too higher for the practical. The main geometric operation includes translation, flipping, shearing, cropping and rotation of the multimedia data.

    - *Enhancement Attack:* This includes the variety of operations such as high and low pass filtering, sharpening, histogram equalization, gamma correction and contrast adjustment. The distortion introduced by the low pass filter does not considerably degrade the quality of watermarked image but can have significant impact on the performance, since spectral component are non-negligible in the spread spectrum based watermark.

2. *Passive Attacks:* In these type of attacks, attacker tries to determine the presence of the watermark without damaging or removal of the watermark. The basic information about the presence of the watermark is a point of greater interest in covert communication, where security against passive attacks is highly required.

3. *Collusion Attacks:* Collusion Attacks are special case of the active attacks. These attacks have the similar goal but contain the methodology difference. In this attack, the attacker needs to construct a watermark copy from the several watermarked data without actual knowledge of the watermark. However, the accessibility of the multiple copies with same data remains a challenging task for the attacker.

4. *Forgery Attacks:* In forgery attacks, attacker tries to embed a new watermark instead of removing the embedded watermark. This allows him to manipulate the secured data according to their choice and then, re-implant a new key to replace the older one. As a result, the process of making the corrupt image seems genuine.

1.3.5 Classification of Watermarking Techniques

Watermarking techniques can be categorised in different ways as per underlying classification criteria . The most common classification of watermarking techniques have been illustrated in Fig. 1.11 which shows that classification are based on numerous factors that include working domain, type of data, human perception, robustness and extraction strategies. Among these, watermarking techniques based on human perception and working domain have received considerable attention by the researchers in past few decades. Considering to working domain, spatial domain techniques are simple but not robust when compared to frequency domain techniques. The main reason behind this fact is that, when inverse process is performed on the image, the embedding information is scattered irregularly over the entire image, which makes the attacker difficult to read or modify. Based on human perception, watermark is either visible or invisible to viewers. In visible watermarking, the embedded watermark is perceptually visible to the viewer whereas in invisible watermarking, watermark is embedded in a manner, such that it cannot be perceptually noticed. An illustration is shown in the Fig. 1.12.

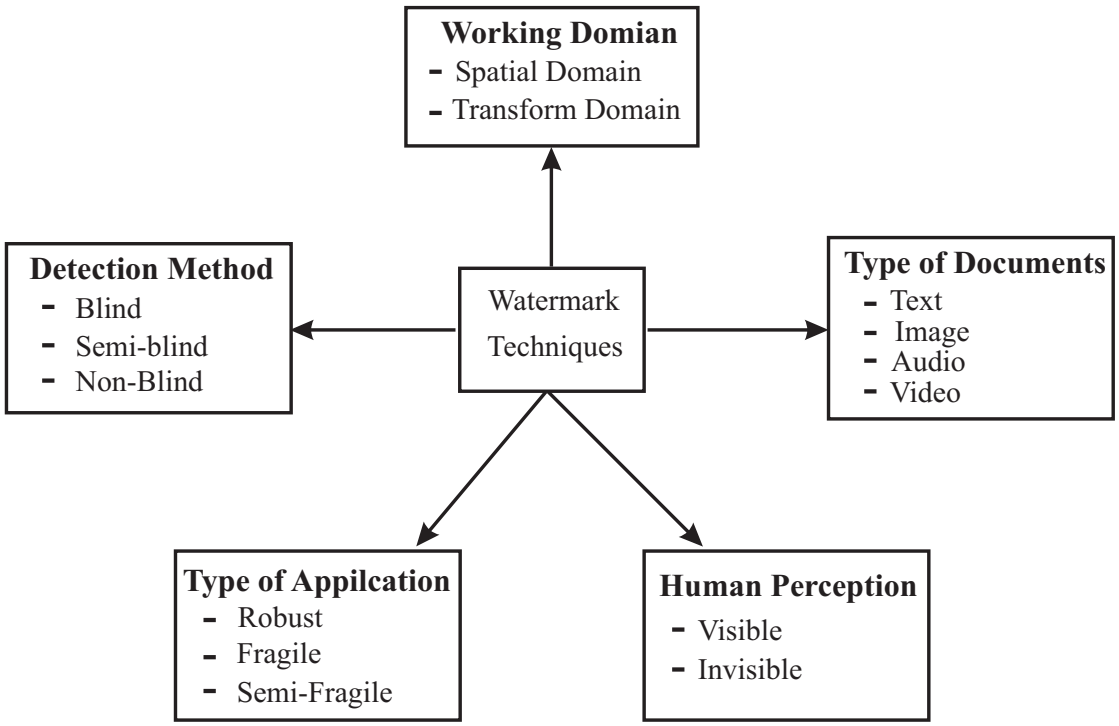

Figure 1.11 : Classification of watermarking techniques.

In addition, watermarking techniques can be divided based on robustness of feature into three groups namely robust, fragile and semi-fragile. Firstly, the robust watermarking techniques are mainly designed for copyright protection and identification of ownership [Craver *et al.*, 1998; Liu and Tan, 2002; Qiao and Nahrstedt, 1998; Zeng and Liu, 1999]. In these type of techniques, user embeds the watermark or proof of ownership into the host media with some perceptible distortion. This watermarked media can be transmitted over the secure or insecure network for the communication purpose. At the other end, receiver can extract the embedding information through the extraction process. During the transmission, watermark may undergo some attacks like white Gaussian noise, filtering operation, geometric attacks or data compression. The robust watermarking has ability to tolerate certain signal processing operations or attacks upto some extent which can occur during the life of watermarked media. Secondly, fragile watermarking is employed to verify the integrity of the multimedia object [Lu *et al.*, 2003; Zhang and Wang, 2007]. This watermarking has the ability to detect the small changes in the original multimedia object and locate the

Figure 1.12 : Experimental Images: a,e) Original images; b,d) Watermark images; e,g) Invisible watermarked images; d,h) Visible watermarked images.

corresponding regions, i.e., any alteration in the pixel values by the attacks or modification can be detected. Fragile watermarking is commonly used for the temper detection and to authenticate the multimedia object. Lastly, semi-fragile schemes are designed for the localisation of tempered media [Qi and Xin, 2011; Sun *et al.*, 2002]. The scheme is employed for content authentication of media object. The scheme has the ability to resist unintentional changes or modification caused by common signal processing operations. Due to this fact, it is possible to verify the content of the object with some perceptible distortion which may arise due to transmission or non-malicious operation. However, semi-fragile schemes detect the malicious or intentional attacks rather than validating the original media.

The extraction of watermark is carried out to obtain an estimate of embedded watermark using a private or public keys. Therefore, watermarking techniques can be further divided into three groups based on information required in the extraction process.

1. *Non-blind Extraction:* This scheme also referred as the private watermarking scheme. These schemes require the original media and extraction key to extract the watermark.

2. *Semi-blind Extraction:* This type of scheme is usually called as the semi-private watermarking scheme. The extraction process requires the information about the watermark and the related keys to extract the watermark.

3. *Blind Extraction:* The scheme is considered as the public watermarking scheme. In extraction process, the information about the original media and watermark is not required. The knowledge of extraction key is sufficient to extract the watermark.

1.4 LITERATURE SURVEY

In recent years, perceptual image hashing got considerable attention by the research community and several hashing schemes have been proposed. Despite of numerous hashing schemes,

finding an ideal scheme remains a challenge. Most of the schemes mainly focused on feature extraction stage, because robust feature extraction can maintain the relevant information after content preserving operations and simultaneously detect content-changing manipulations. However, there are some schemes which address the security of the perceptual hashing system. In [Schneider and Chang, 1996], authors proposed a perceptual hash function based on intensity histogram of the image blocks, where the spatial information of each blocks are computed and public key encrypted to generate the final hash value. For verification the hash is decrypted first and then histogram intensity are utilized for similarity estimation. The main drawback of the scheme is that it can be easily attacked while maintaining the same histogram intensity. In [Venkatesan *et al.*, 2000], authors construct a perceptual hash function using random tiling on low frequency-band in wavelet decomposition. The average of the coarse sub-band and variance are estimated followed by the probabilistic quantization. The quantized statistics vector are sent as an input to error-correcting code [Blahut, 1983] to generate the final image hash. Generally, Statistical properties of wavelet subbands are robust against content preserving operations (CPO) but less sensitive to content changing operations (CCO). This scheme is robust to geometric attacks but fails to detect the small change in the image content. A similar work is proposed using k-mean clustering, where statistics of the each region such as mean, variance and higher order moments are used to derive the hash value. This scheme have similar limitation as described in [Schneider and Chang, 1996]. Image histogram can also effectively used in estimation of statistical features due to invariance of pixel positions in the image. In [Xiang *et al.*, 2007], author proposed a perceptual hashing scheme using the histogram shape and utilize it in the digital watermarking [Xiang *et al.*, 2008]. It is investigated that the number of pixels among different bins of histogram are invariant to common geometric operations. The input image is pre-processed using Gaussian low pass filter and the population among different bins are compared to generate a binary string. Further, a key based permutation is applied to estimate final hash value. The scheme shows good robust against geometric operation as well other challenging operations image shearing and wrapping. However, the local structure of the respective image is lost in the estimated image hash.

Another approaches for perceptual hashing function is based on low level features. Several techniques have been proposed based on features such as interest points, edge and blobs present in the images. The core idea is based on robust low level feature extraction which is useful in the hash generation and therefore gains popularity in the recent years. In [Monga and Evans, 2004, 2006], a framework for perceptual hashing have been presented based on feature points. In this scheme, perceptually significant features are extracted on the basis of wavelet based feature detector which is characterized by the visual system [Hubel and Wiesel, 1965]. Then an iterative algorithm [Mıhçak and Venkatesan, 2001a] preserve the good invariance to visually insignificant perturbations is employed to obtain the final hash value. In second approach, a key based random tiling is used in advance for random hash generation. The limitation of the scheme is that feature point detection may not be achieved in case of smooth texture. An extending work is presented in [Monga, 2005] where they try to approximate the geometric distortion using the affine transform. As a result, the approximation increases the robust against geometric distortion but this decreases the efficiency. In [Roy and Sun, 2007], a hashing scheme is presented specially for image tempering detection. The scheme is divided into two parts: the first part aiming the authentication and utilizes the scale invariant feature transform (SIFT) which shows the excellent performance against geometric transformations. Second part focused on the tempering localization and utilized edge direction of the histogram. The efficiency and performance is analyzed in terms of robustness and collision resistance in [Roy *et al.*, 2008].

Another popular approach for image hashing technique is based on dimensionality reduction. Dimensionality reduction is a process where, higher dimension dataset is converted into relatively lower dimension dataset while preserving the properties of original data [Fodor, 2002]. Generally, the hashing approaches based on dimensionality reduction retain the coefficients in the

low-rank approximation of the image. In [Kozat *et al.*, 2004] authors developed a hashing scheme based on singular value decomposition (SVD). The intermediate features are extracted using SVD and a secondary image is constructed. Further, SVD is again applied on the secondary image to generate the hash value. The SVD based techniques capture the robustness against geometric and other common operations. The similar concept is used in the hashing technique [Monga and Mihçak, 2007] based on non-negative matrix factorization. The NMF decomposition is applied twice due to its non-negativity constraints and then pseudo randomization is used to generate the final hash string. This scheme obtains the higher robustness against CPOs and reducing the misclassification under CCOs. In [Lv and Wang, 2008] authors proposed a perceptual hashing scheme based on fast Johnson-Lindenstrauss transform (FJLT). The FJLT shares characteristics using the random projections and requires less complexity to accomplish the task. The scheme achieve robustness comparable to NMF scheme [Monga and Mihçak, 2007] but lower computation cost. In [Hernandez *et al.*, 2011], authors proposed a hashing scheme based on SVD and image normalization. Firstly, image normalization is employed on the input image and a features vector is then constructed based on random tiling and SVD. The resulting vector is quantized and a binary sequence is obtained using gray level coding. The binary sequence is compressed using Reed-Muller decoder [Mıhçak and Venkatesan, 2001b] to generate the final hash value. In [Tang *et al.*, 2014b] authors proposed a hashing technique based on improved NMF and ring partitions. The ring partition creates rotation invariance matrix and then NMF is applied to produce the final hash value.

There are several other perceptual hashing systems developed based on invariant properties using different transformations. In these type of approaches, the input image is transformed from spatial domain to transformed domain such as discrete cosine transform (DCT) [Jie, 2013; Tang *et al.*, 2014a], discrete Fourier transform (DFT) [Ouyang *et al.*, 2015; Qin *et al.*, 2013], discrete wavelet transform (DWT)[Guo and Hatzinakos, 2007; Singh and Bhatnagar, 2017b] etc. These techniques extract invariant properties in different to construct image hash. In the early work [Fridrich, 2000], authors extract image hash by projecting image coefficient onto the pseudo-random patterns. The final hash sequence is then utilized to generate the random watermark sequences which mainly depends on the sensitivity of the secret key. In a similar work [Fridrich and Goljan, 2000], a perceptual hashing system based on DCT is presented, where each DCT block is projected on zero-mean random patterns. The image hash mainly depends on the low frequency DCT coefficients and its resistance under CPOs. In [Mıhçak and Venkatesan, 2001a], a wavelet based image hashing technique have been discussed, where an iterative approach is employed to binarize the approximation sub-band of the considering image using DWT. The stable features are preserved and unstable features are discarded during the iterations. In [Swaminathan *et al.*, 2006], a content based image hashing is developed based on Fourier-Mellin transform. The image hash is based on the controlled randomization rotation invariance of the Fourier transform. The scheme shows good robustness to CPOs but not able to detect the local malicious modification in the image, because Fourier transform did not provide the frequency localization. In a more recent development, robust features are exploited using radon transform in the perceptual hashing schemes [De Roover *et al.*, 2005; Lefebvre *et al.*, 2003; Lei *et al.*, 2011]. In [Lei *et al.*, 2011], authors presented a radon based image hashing technique, where input image is transformed using radon transform and then moment based features are extracted. Further, DFT is employed on the moment features followed by normalization on the selected DFT coefficients to generate final hash sequence. The proposed work shows the good robustness against geometric and other operations including JPEG compression. In [Guo and Hatzinakos, 2007], authors proposed a hashing framework by combining the DWT and radon transform. DWT offer good frequency localization whereas radon transform achieves shift/rotation invariance. Due to combined effect, the algorithm can detect local malicious manipulation and offer excellent robustness against content-preserving operations.

In recent development, there are various moment based perceptual hashing approaches [Ouyang *et al.*, 2017; Zhao *et al.*, 2013] have been presented based on different novel ideas. In

[Ouyang *et al.*, 2017] author proposed a perceptual hash function by using the local and global features. Global features are extracted by computing the zernika moments (ZM) while local features are based on the salient regions of image. However, its performance mainly depends on the accuracy of the saliency detection. A similar approach has been discussed in [Ouyang *et al.*, 2016]. In [Wang *et al.*, 2015] authors proposed a scheme based on scale-invariant feature transform (SIFT). They extract the feature with the combination of SIFT and then produce the hash value with the help of the discrete cosine transform. In [Lv and Wang, 2012] authors proposed a shape-contexts-based image hashing scheme by utilizing the local features. They employed the scale-invariant feature transform (SIFT) and incorporate with Harris corner detector. In [Tang *et al.*, 2019] authors proposed a scheme based on tensor decomposition (SVD). This scheme is robust against the geometric distortion but vulnerable to large geometric transformation and brightness change. In [Tang *et al.*, 2016] author proposed a hashing scheme based on ring partition and invariant vector distance. The statistical features are extracted from the image ring in uniform color space.

Image encryption is another important solution, which protects the confidentiality of the data during the transmission and has wide range of applications in communication, military, medical imaging etc. The modern clinical diagnosis has entered into the digital age due to substantial proliferation in medical research and technology. This paradigm shift leads to fast and better accuracy in the diagnosis and the treatment of patients. Most of the medical imaging sensor provides the data in the form of digital images such as computed tomography (CT), ultrasound, X-ray, magnetic resonance image (MRI) and positron emission tomography (PET). These medical images contain important, confidential and sensitive information related to the patient's health condition. The main problem arises during the storage and communication of these images for various purposes. These diagnostics images are generally stored and transmitted across several public channels through internet for various applications such as denoising, compression, segmentation and data hiding [Kamran and Farooq, 2012; Phophalia *et al.*, 2014; Wu *et al.*, 2015]. However, due to lack of inherent security measures, these medical images may be exposed to the serious threats like illegal manipulation, privacy leakage and data integrity [Li and Lo, 2011; Zhang *et al.*, 2015a]. Therefore, the security of medical images is a crucial issue and needs to be addressed.

In the last decade, various methodologies have been developed to protect the medical images such as encryption, hashing, stenography and watermarking. Among these, encryption is the most suitable methodology to protect the integrity of the data. However, the conventional encryption techniques such as Advanced Encryption Standard (AES), Data Encryption Standard (DES) and International Data Encryption Standard (IDES) are not suitable for the encryption of medical images due to the integral features of the medical images such as high correlation between neighbouring pixels and redundancy [Menezes *et al.*, 1996; Sadourny and Conan, 2003]. More precisely, these techniques are developed to ensure the security of textual data and the requirements of medical image encryption are different from the textual data as they have large data size, local structure, bulk data capacity and modalities of amplitude-spectra [Kanso and Ghebleh, 2015; Shi *et al.*, 2017]. Hence, the stressed motive of this work is to fulfil the requirement of robust encryption technique for medical images with considerably high security.

In medical image encryption, the original image is converted into cipher image by changing the pixel values in such a way that the original image becomes apparently meaningless such that it should not reveal the important information contained in the original medical image. However, an authorized person can reproduce the original image using the decryption process for different purposes. Recently, various encryption techniques have been proposed [Cao *et al.*, 2017; Dzwonkowski *et al.*, 2015; Laiphrakpam and Khumanthem, 2017; Li *et al.*, 2009, 2013] in the literature. In [Zhang and Xiao, 2014], authors proposed a new encryption algorithm based on block diffusion and bit-level permutation of a matrix. The input image is divided into non-overlapping blocks and each block is transformed into a 3-D binary matrix. Further, this matrix is multiplied by a rotation

matrix to perform the permutation. In [Hua *et al.*, 2015], authors design an encryption technique using the sine logistic modulation map (SLMM), which is derived from the sine and logistic map. Firstly, pixels position in the images are changed using chaotic magic transform (CMT), then the input image is encrypted with the combination of CMT and SLMM. In [Chen and Hu, 2017], authors proposed an adaptive encryption technique for medical images by utilizing the improved chaotic map to overcome the drawbacks of the existing techniques. They employed sine chaos map for image scrambling to the input image, then scrambled image coefficients are partitioned into 2×2 sub-blocks which can be adaptively encrypted using hyper chaotic system. In [Laiphrakpam and Khumanthem, 2017], authors proposed a medical encryption technique based improved Elgamal encryption. They removed the calculation that essentially encodes the plaintext message into elliptic curve coordinates. The technique seep up the execution and resolved the data expansion issue. In [Hua *et al.*, 2018], authors presented encryption technique for medical images using adaptive pixel diffusion. In this process, firstly the surrounding of the image is changed by inserting the random data, then adaptive pixel diffusion and image scrambling method are employed to shuffle the neighbouring pixels randomly. Simultaneously, the inserted data is spread over the entire image. In [Ping *et al.*, 2018], authors proposed an encryption technique based on discrete Henon map, where the two point diffusion process significantly improved the performance. Also, the intermingle of permutation and substitution reduce the scanning time of the image. Recently, several DNA approach based encryption techniques have been proposed to protect the digital images [Enayatifar *et al.*, 2017; Wu *et al.*, 2018]. In [Wu *et al.*, 2018], authors developed an encryption technique based on DNA encoding and 2D Henneon-sine map. Firstly, a Henneon-sine map is derived by combining the Henneon and sine map, then image pixels of input image are diffused by DNA encoding followed by permutation using Henneon-sine map. In [Enayatifar *et al.*, 2017], authors proposed an encryption scheme based on 3-D logistic map and DNA. The permutation of image pixels are achieved by utilizing DNA sequence and 1-dimension logistic map whereas the association of second and third dimension with DNA operator alter the image pixels. Among these techniques, most of them are falling short to cater the requirements having vulnerability to different malicious attacks and providing a lower degree of security. To increase the security, medical image encryption techniques have been presented based on the chaos theory in [Hsiao and Lee, 2015; Zhang *et al.*, 2015b]. However, these techniques are not completely secure due to poor key management procedure [Jain *et al.*, 2006; Uludag *et al.*, 2004]. In general, the security of an encryption technique is mainly dependent on the associated keys such that if these keys are copied or stolen then unauthorized entity can access the data. On the other hand, if these keys are destroyed or corrupted then accessing the data would not be possible and hence the key management associated with security framework must be secure and untraceable [Giri *et al.*, 2016].

These issues can be addressed by employing a perfect key management system which is simple, unique, untraceable and non-repudiation. One way to achieve this is to use the biometrics [Hao *et al.*, 2006; Uludag and Jain, 2003] of the patients/owners. Biometrics refers to biological or behavioral characteristics that uniquely identify an individual [Jain *et al.*, 2007]. Some of the well-known characteristics are facial structure, fingerprint, irises, palm prints, voice and gait. However, behavioral characteristics cannot be used for the identification purposes as they can be imitated. Therefore, the biological characteristics inherited in the biometrics can be used in the key management to strengthen the security of the encryption technique [Chen and Chandran, 2007; Gaddam and Lal, 2010].

There are several watermarking taxonomies that have been reported in the existing literature. Among them, watermarking taxonomies based on working domain are most popular and represented by spatial and transform domain. In spatial domain approaches [Bas and Furon, 2013; Bender *et al.*, 1996; Chen *et al.*, 2016; Liu and Tan, 2002; Nikolaidis and Pitas, 1996; Pérez-Freire and Pérez-González, 2009; Pitas and Kaskalis, 1995; Singh and Bhatnagar, 2017a; Van Schyndel *et al.*, 1994; Voyatzis and Pitas, 1996], embedding of watermark directly changes the gray values of the

image. These techniques are easy to implement and require low computational cost, but less robust against several type of distortions. In contrast, the transform domain approaches [Barni *et al.*, 2001; Cox *et al.*, 1997; Daren *et al.*, 2001; Ganic *et al.*, 2003, 2005; Huang *et al.*, 2000; Lin *et al.*, 2008; Singh and Bhatnagar, 2018b, 2019; Tao *et al.*, 2014; Zhao *et al.*, 2004; Zhu *et al.*, 1999] are more robust. This is due to the fact that watermarking information is embedded into the transformed coefficients and the inverse process distributes it irregularly over the entire image. This makes difficult to modify or remove the watermark. A range of transform domain techniques have been proposed in the literature based on discrete Fourier transform (DFT) [Cox *et al.*, 1997], discrete cosine transform [Huang *et al.*, 2000; Tao *et al.*, 2014], discrete wavelet transform (DWT) [Daren *et al.*, 2001; Singh and Bhatnagar, 2019] and others [Chouhan *et al.*, 2011; Jha *et al.*, 2013, 2014; Singh and Bhatnagar, 2018a; Sun and Lei, 2008; Wu and Qiu, 2006].

In spatial domain, the most common watermarking approaches are based on least significant bit (LSB) using PN sequence. In LSB techniques, the watermark is embedded by changing the least significant bit of the host data. The change introduced in the pixel values due to LSB modification is visually less significant which ensures the invisibility of the watermark. In [Van Schyndel *et al.*, 1994], authors proposed two LBS based techniques. Firstly, they replace the least significant bit by PN sequence to modify the pixel values whereas a PN sequence is added in LSB of the pixel values in the second technique. In [Bender *et al.*, 1996], authors developed a statistical scheme based on spread spectrum principle. They select m pairs (u, v) randomly in an image, then increase the value u by one unit and decrease v in the same amount simultaneously. In [Nikolaidis and Pitas, 1996] presented a binary watermarking scheme in which they modified the intensity levels at randomly selected location by adding a small positive quantity. In detection process, the mean value of marked pixels is compared with unmarked pixels to verify the presence of the watermark. In [Voyatzis and Pitas, 1996], authors utilized the toral automorphism to generate a chaotic sequence and hide the watermark at different locations given by chaotic sequence which are dense in spatial domain. In [Pitas and Kaskalis, 1995], authors proposed a watermarking scheme in which the host image is randomly split into two parts of equal size. For watermark embedding, the mean value of one of the selected part is increased by a factor and authors investigate security of spread spectrum communication based watermarking theoretically and practically. In [Liu and Tan, 2002], authors presented a robust watermarking scheme based on singular value decomposition (SVD). In this method, singular values (SVs) of cover and watermark images are estimated and modified in order to embed the watermark. The modified singular values and known components are combined to produce the watermarked image. A new security measure based on the effective key length is presented in [Bas and Furon, 2013], where the author estimated the effective key length for spread spectrum communication (SSC) and improved SSC. In [Chen *et al.*, 2016], authors reported an optimal watermarking scheme based on integrated quantization embedding. The watermark is embedded into low frequency component of the discrete wavelet transform (DWT) using amplitude quantization to ensure the robustness. In [Singh and Bhatnagar, 2017a], authors presented efficient watermarking scheme using the spread spectrum principle. They embedded a binary watermark using the PN sequence generated from linear feedback shift register (LFSR). In general, the LSB based approaches that modify the image data using a PN sequence of fixed magnitude are highly vulnerable to signal processing attacks. The main reason behind this weakness is due to limited magnitude of the embedded noise which ensures the invisibility of the watermark.

Generally, transform domain approaches are more resilient to signal processing operations in comparison to spatial domain. In [Cox *et al.*, 1997], authors proposed a watermarking scheme using the spread spectrum principle. The watermark embedding is carried out using first ℓ highest magnitude coefficients of DFT, whereas comparing the DFT coefficients of the watermarked and original image to extract the watermark. In [Barni *et al.*, 2001], authors reported a watermarking scheme in DCT domain wherein watermark is embedded in DCT coefficients considering a sequence of real numbers. An adaptive watermarking scheme [Huang *et al.*, 2000] is proposed by

utilizing luminance and texture in DCT domain. The watermark is embedded in DC coefficients, due to larger perceptually capacity than AC coefficients which ultimately improves the robustness. A watermarking scheme [Lu *et al.*, 2007] is presented based on sub-sampling and difference correlation. The Sub-sampling of host image produces the four sub-images and a binary watermark is embedded into DCT domain of the two sub-images using the secret key. In [Etemad *et al.*, 2018], authors developed a watermarking scheme in Hadamard domain. In this method, the watermark information is stored redundantly in the selected bit planes. In [Di Martino and Sessa, 2018], authors presented a fragile watermarking scheme based on binary fuzzy relation. They applied F-transform on the original image and then watermark is applied for temper detection. Several other watermarking schemes based on block-based DCT can be found in [Tao *et al.*, 2014].

In past few years, a new mathematical transform namely discrete wavelet transform is frequently used in watermarking [Daren *et al.*, 2001; Ganic *et al.*, 2003, 2005; Lin *et al.*, 2008; Serdean *et al.*, 2003; Singh and Bhatnagar, 2019; Zhu *et al.*, 1999]. DWT has several advantages over other frequency domain approaches such as multiresolution property, better energy compaction and effective scale space approximation. In [Zhu *et al.*, 1999], authors presented a unified approach for digital watermarking using multi-level discrete wavelet transform. This method adds the watermark using a Gaussian distributed random vector in high frequency band. In [Ganic *et al.*, 2003], authors reported a watermarking scheme in which the payload factor is determined by considering the SVs of the cover and watermark image. In [Ganic *et al.*, 2005], authors proposed an optimal watermarking scheme utilizing DWT and SVD. The SVD is performed to each wavelet sub-bands of the images and modifies the singular values to produce the watermarked image. Authors design a blind watermarking scheme [Lin *et al.*, 2008] based on the significant differences of wavelet coefficients. In extraction process, maximum wavelet coefficients are quantized and the energy difference between the significant difference are used to estimate the extracted watermark. The embedding in frequency domain can be further extended to the multiple frequency domain. The use of multiple frequency domain for embedding purpose can have the combined advantages derived from each domain. A watermarking scheme [Iovane *et al.*, 2011] proposed based on wavelet multiresolution analysis (WMA) and face biometrics. They construct a marker for watermark embedding using scale invariant feature transform (SIFT). In [Liu *et al.*, 2018c], authors proposed a watermarking scheme based on logistic map and RSA encryption where embedding is carried out based on discrete wavelet transform and singular value decomposition. In [Liu *et al.*, 2018b], authors developed a blind dual watermarking technique for color images. They embedded the watermark in YCbCr space utilizing DWT. In [Zhao *et al.*, 2004], authors proposed a watermarking scheme in dual domain comprising of DCT and DWT, where DCT transform is utilized for watermark generation and embedded in the DWT domain. In [Amirgholipour and Naghsh-Nilchi, 2009], authors proposed an optimal watermarking scheme based on dual domain using DWT and DCT. However, the scheme is less robust against geometric modifications.

Recently, a new approach based on stochastic resonance (SR) has been used to determine the existence of the watermark [Wu and Qiu, 2006]. The core idea involves the detection of the watermark signal from the possible attacked watermarked image and therefore SR phenomena is employed to optimize the authenticity of the watermark in extraction process. Thereafter, numerous watermarking schemes have been proposed by the incorporation of stochastic resonance with different philosophies [Chouhan *et al.*, 2011; Jha *et al.*, 2013, 2014; Singh and Bhatnagar, 2018a; Sun and Lei, 2008]. In [Sun and Lei, 2008], authors design a new watermarking system based on aperiodic signal processor and DCT, where embedding is done in DCT coefficients and aperiodic signal processor employed to detect the binary pulse amplitude modulation (PAM) signals. A more efficient watermarking technique has been developed based on dynamic stochastic resonance (DSR) [Chouhan *et al.*, 2011; Jha *et al.*, 2014]. In former technique, a binary watermark is embedded in DCT domain using pseudo random sequence whereas multilevel DWT and DCT are considered in later one. In both the cases, dynamic stochastic resonance (DSR) based detector is employed to

improve the authenticity of the watermark. In [Jha *et al.*, 2013] author embedded a gray scale logo by modifying the singular values whereas the incorporation of DSR and singular values extract the watermark. The major drawback of the algorithm is that it has limited robustness and suffer from false positive problem. In [Singh and Bhatnagar, 2018a], authors embedded the watermark in integer DCT domain and remove the falsification problem which essentially resolves the existing problem.

1.5 THESIS ORGANIZATION

The primary objective of this research work is to provide the robust algorithmic framework for image security. Our aim is to design secure and efficient systems which include image hashing, encryption and digital watermarking. The thesis work is comprised of eight chapters and can be summarized as below.

- **Chapter 2** provides the basic overview of mathematical preliminaries which are essentially required to accomplish the desired work. The mathematical preliminaries such as KAZE features, singular value decomposition, Log-polar mapping and integer DCT domain are briefly described.

- **Chapter 3** presents a chaos based robust and secure hashing framework. The technique utilizes the statistical features in the hash generation based on discrete cosine transform. The technique is robust against various operations and has ability to authenticate the maliciously modified image. In addition, a new robust reference hashing system is presented in **Chapter 4**. This technique combines the local and global features to estimate the image hash. The main advantage of the technique is the high sensitivity against small change in the image content and therefore plays a vital role in image authentication as well as classification.

- **Chapter 5** provides a biometric inspired security framework for medical images. This work incorporates the biometric features into the encryption to protect the confidentiality of the data. The technique is highly secure and can protect the content of image data.

- **Chapter 6** proposes a robust watermarking framework based on lifting wavelet transform. A binary watermark is embedded for copyright protection of a gray scale image. A blind watermark detection process is then employed in order to detect the watermark. **Chapter 7** offers a new watermarking system in integer DCT domain. The scheme incorporates DSR phenomena in extraction process which enhances the overall robustness of the watermarking system. The proposed technique rectifies the false-positive detection problem of SVD-based watermarking.

- **Chapter 8** presents the concluding remarks of the thesis. The research outcome and novel contribution are highlighted. Finally, future scope is also discussed at the end.