Chapter 2

Principles and Models of Digital Watermarking

2.1 Introduction

This chapter introduces some fundamental definitions and the main elements of a digital watermarking system. It further develops a generic model for a digital watermarking system. Concepts and functional modules are presented to describe an abstract scheme of digital watermarking methods. A natural starting point in the discussion of information embedding systems is to develop mathematical models that suitably describe information embedding applications in medical images. Such models facilitate a precise consideration of the issues involved in the design and performance evaluation. This generic model provides a common basis to understand the watermarking systems, described in the latter chapters of the thesis, and to evaluate their performance.

2.2 The Image Watermarking Problem

Image watermarking is the process that imperceptibly embeds a watermark $W$ into a host image, $I$, to form the watermarked image, $I_w$. The watermark, $W$ is generally a vector containing either pseudo-random bits or pseudo-random samples from a probability distribution. Much like a real paper watermark, $W$ is physically bound to $I_w$ and can be detected to make an assertion about $I_w$ [54]. The watermark $W$ should be recoverable from the watermarked image $I_w$ even if $I_w$ is altered by one or more image processing procedures such as filtering, geometric distortions, re-sampling or by malicious attacks (doctoring). The application context of a watermarking scheme determines which image processing transformations of $I_w$ should not hinder watermark detection. For example, watermarking schemes used for image authentication are
required to withstand distortions of $I_w$ introduced by file format conversion, compression, and re-sampling. However, such schemes must reject $I_w$ if $I_w$ is perceivably different from $I$ due to filtering, “doctoring,” or geometric distortions, since the purpose of authentication watermarking is to detect significant modifications of the image.

### 2.3 Watermarking Issues

Each watermarking application has its own needs that determine the required attributes of the watermarking system and techniques for watermark embedding and detection. Many forms of robustness against distortion and attack, visibility of the embedded mark, payload of the watermark, immunity of the detector to false alarms, and security are important attributes of real world systems. These issues have been the focus of intense study in this rapidly evolving research area [55]. Many challenges still exist, because digital watermarking is inherently a multidisciplinary research field comprising information and communication theory, decision and detection theory, signal processing and cryptography and cryptographic protocols. Each of these areas deals with a particular aspect of the digital watermarking problem. Generally speaking, information and communication theoretic methods deal with the data embedding (encoder) side of the problem.

### 2.4 Design of a Watermarking System

A watermarking system can be considered much like a communication system consisting of three main elements: a transmitter, a communication channel, and a receiver. To be more specific, the embedding of the to-be-hidden information within the host signal plays the role of data transmission; any processing applied to the host data after information concealment, along with the interaction between the concealed data and the host data itself, represents the transmission through a communication channel; the recovery (decoding) of the hidden information from the host data acts the part of the receiver [56]. Table 2.1 presents a comparative description of the communication and watermarking systems.
<table>
<thead>
<tr>
<th>Communication System</th>
<th>Watermarking System</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information</td>
<td>Watermark</td>
</tr>
<tr>
<td>Communication channel</td>
<td>Host signal (such as image, video)</td>
</tr>
<tr>
<td>Power constraint on transmitted signal due to physical limitations</td>
<td>Power constraint on watermark due to audio/visual quality limitations</td>
</tr>
<tr>
<td>Interference</td>
<td>Host signal and watermark attacks</td>
</tr>
<tr>
<td>Side information at transmitter and/or receiver</td>
<td>Knowledge of host signal, watermarking parameters such as key at the encoder and/or decoder</td>
</tr>
<tr>
<td>Channel capacity</td>
<td>Watermarking capacity</td>
</tr>
</tbody>
</table>

By following the communication analogy a generic watermarking model can be depicted as in Fig. 2.1. The watermark generator takes a watermark signal, \( B \) and a key, \( K \) to generate the watermark, \( W \). The encoder then embeds \( W \) into \( I \) to produce a watermarked image, \( I_w \). The watermarked image is distributed and potentially distorted by image processing operations and by targeted watermark attacks. In the last stage of the model, the decoder attempts to ascertain the presence of the original watermark in the distorted image.

Abstractly, a watermarking scheme is a seven-tuple \((B, I, W, K, G, E, D)\) where

1. \( I \) represents the set of original or host images.
2. \( W \) is the set of all watermarks \( w \) such that \( \exists i \in B; k \in K; W = G(i, k) \)
3. \( K \) is the set of numbers called watermark keys.
4. \( G \) is the algorithm that generates \( W \) using \( K \) and \( B \):
   \[
   G: I \times K \to W, w = G(i, k)
   \]
5. \( E \) denotes the encoding algorithm that embeds a watermark \( W \) in an image \( I \) with some strength \( \alpha \):
   \[
   E: I \times W \times \mathbb{R} \to I, I_w = E(i, w, \alpha)
   \]
6. \( D \) is the decoding algorithm that detects whether a watermark \( W \) is present in an image.
   \[
   D: I \times W \to \{0,1\}, D(i,w) = \begin{cases} 1 & \text{if } w \text{ exists in } i \\ 0 & \text{otherwise} \end{cases}
   \]
Fig. 2.1: A generic model for digital watermarking system

As seen in Fig. 2.1, a digital watermarking system consists of two main components: watermark embedder and watermark detector. The watermark code $B$ represents the very input of the chain. Then, $B$ is transformed in a watermark signal $W$ (optionally $B = W$), which is embedded into the host image $I$, thus producing the watermarked asset $I_w$. Due to possible attacks, $I_w$ is transformed into $I_{w'}$. Finally the decoder/detector recovers the hidden information from $I_{w'}$. Note that embedding and watermark recovery may require the knowledge of a secret key $K$, and that recovery may benefit from the knowledge of the original, non-marked asset $I$. 
2.4.1 Design of the Watermark Signal

In many watermarking systems, the information message $B$ is not embedded directly within the host signal. On the contrary, before insertion, vector $B$ is transformed into a watermark signal $W = \{w_1, w_2, \ldots, w_n\}$ which is more suitable for embedding. $B$ also may be left as it is, thus leading to a scheme in which the watermark code is directly inserted within $I$. In this case the watermark signal $W$ coincides with the watermark message $B$. Before transforming the watermark code into the watermark signal, $B$ may be channel-coded to increase robustness against possible attacks. Channel coding greatly improves the performance of any watermarking system [60]. Typically, the watermark signal depends on the key $K$ and watermark information $B$:

$$W = f_0(B, K)$$  \hspace{1cm} (2.1)

Possibly, it may also depend on the host data $I$ that is embedded into:

$$W = f_0(B, K, I)$$  \hspace{1cm} (2.2)

2.4.2 Watermark Embedder

The embedder combines the digital data, $I$, and hidden information $W$ representing the watermark to be added to $I$. The output of the embedder is the watermarked $I_w$, which is perceptually identical to $I$ but with the embedded watermark $W$. The attacker box in Fig. 2.1 represents malicious or accidental attacks that intend to change the watermarked data. Depending on the application, the goal of the attacker may be the modification of $I_w$ to make the detection of the watermark $W$ impossible (copyright infringement), or to covertly corrupt some sensitive content in $I$ for violating its integrity (authentication).

The watermark encoding algorithm, $E$, embeds a watermark either in the spatial or in a transform domain of an image. Spatial domain watermarking schemes are generally more computationally efficient than the schemes operating in transform domains. However, transform domain watermarks are far more resistant to various image processing attacks [28]. As a result, almost all recently proposed watermarking schemes operate in a transform domain. A generic transform domain encoder $E(i, w, \alpha)$ works as follows:
After selecting the appropriate transform domain, the next step performed by the embedder is to insert the hidden information in the content. There are two widely used methods for information embedding, modulation and quantization.

1. Modulation is performed by one of the three formulae proposed by:

\[
I_W = I + W \\
I_W = I(1 + \alpha W') \\
I_W = Ie^{\alpha W}
\]

where \(\alpha\) is a strength factor for the modulation [57]. Transform domain watermarking schemes can be classified into additive, multiplicative, and quantization-based [58], depending on how their encoders compute the \(I_W\).

2. Quantization embeds information by quantizing \(I\) using a quantization function \(q\) that provides different values to different embedded message values \(W\). The embedding strength is defined by the minimal distance between two adjacent values in \(W\). Quantization based encoders [59] generate the watermarked coefficients \(I_W\) using some non-linear function such as in Eqn. 2.5

Fig. 2.2: Watermark embedding via. invertible feature extraction

In watermark embedding, or watermark casting, an embedding function, \(E\) takes the host asset \(I\), the watermark signal \(W\), and, possibly, a key \(K\), and generates the watermarked asset, \(I_W\)

\[
E(I, W, K) = I_W
\]
The definition of $E$ usually goes through the selection of a set of asset features, called host features that are modified according to the watermark signal. By letting the host features be denoted by $\mathcal{F}(I) = f_I = \{f_1, f_2, ..., f_n\} \in F^n$, watermark embedding amounts to the definition of an insertion operator $\oplus$ which transforms $\mathcal{F}(I)$ into the set of watermarked features $\mathcal{F}(I_W)$ i.e.,

$$\mathcal{F}(I_W) = \mathcal{F}(E(I,W,K)) = \mathcal{F}(I) \oplus W \quad (2.7)$$

In general, $m \neq n$, that is the cardinality of the host feature set need not be equal to the watermark signal length. Though Eqns. 2.6 and 2.7 basically describe the same process, namely watermark casting within $A$, they tend to view the embedding problem from two different perspectives. According to Eqn. 2.6, embedding is more naturally achieved by operating on the host asset, i.e., $E$ modifies $I$ so that when the feature extraction function, $\mathcal{F}$ is applied to $I_w$, the desired set of features $f_{iw} = \{f_{iw,1}, f_{iw,2}, ..., f_{iw,n}\}$ is obtained. Eqn. 2.7 tends to describe the watermarking process as a direct modification of $f_I$ through the embedding operator $\oplus$. According to this formulation, the watermark embedding process assumes the form shown in Fig. 2.2. First the host feature set is extracted from $I$, then the operator $\oplus$ is applied producing $f_{iw}$; finally the extraction procedure is inverted to obtain $I_W$:

$$I_W = \mathcal{F}^{-1}(f_{iw}) \quad (2.8)$$

The necessity of ensuring the invertibility of $\mathcal{F}^{-1}$ may be relaxed by allowing $\mathcal{F}^{-1}$ to exploit the knowledge of $I$ to obtain $I_W$, that is (weak invertibility)

$$I_W = \mathcal{F}^{-1}(f_{iw}, I) \quad (2.9)$$

### 2.4.3 Watermark Concealment

The main concern of the embedding part of any data hiding system is to make the hidden data imperceptible. This task can be achieved either implicitly, by properly choosing a set of host features and an embedding rule, or explicitly, by introducing a concealment step after watermark embedding [60]. To this aim, the properties of the human senses must be carefully studied, since imperceptibility ultimately relies on the imperfections of such senses. Thereby, still image and video watermarking will rely on the characteristics of the HVS, whereas audio watermarking will exploit the properties
of the Human Auditory System (HAS). A detailed description of the important phenomena underlying the HVS is given in chapter 3, where a novel watermarking algorithm is developed exploiting HVS features of DWT.

### 2.4.4 Watermark Decoding

The receiver part (detector) of the watermarking system may assume two different forms. According to the scheme reported in Fig. 2.3a, the watermark detector reads $I'_W$ and a watermark code $W'$, and decides whether $I'_W$ contains $W'$ or not. The detector may require that the secret key $K$, used to embed the watermark, is known. In addition, the detector may perform its task by comparing the watermarked asset $I'_W$ with the original, nonmarked, asset $I$, or it may not need to know $I$ to take its decision. In the latter case the detector is classified as blind, whereas in the former case it is said to be non-blind.

![Diagram](image)

**Fig. 2.3:** With detectable watermarking (a) the detector just verifies the presence of a given watermark within the host asset. With readable watermarking (b) the prior knowledge of $W'$ is not necessary.

The decoder may work as in Fig. 2.3b. In this case the watermark code, $W'$ is not known in advance, the aim of the receiver just being that of extracting $W'$ from $I'_W$. As before, the extraction may require that the original asset $I$ and the secret key $K$ are known. The two different schemes given in Fig. 2.3 lead to a distinction between algorithms embedding a mark that can be read as well as those inserting a code that can only be detected. In the former case, the bits contained in the watermark can be read without knowing them in advance (Fig. 2.3b). In the latter case, one can only verify if a
given code is present in the document, i.e. the watermark can only be revealed if its content is known.

There are two common ways in which blind watermark decoders decide whether a given watermark $W$ is present in the image $I'_W$:

1. Extract the watermark $w'$ from the transform coefficients $\tilde{v}_1, \tilde{v}_2, ..., \tilde{v}_m$ of $I'_W$ and test if $c(w, w') > t$, where $c$ is the comparison function and $t$ is the detection threshold. The form of $W$ determines which comparison function $c$ is used. For example, if $W \in \{0,1\}^s$, $c$ might compute the percent bitwise similarity between $W$ and $W'$: $c(W, W') = 1 - \frac{1}{s} \sum_{i=1}^{s} W_i \oplus W_i'$. The extract and compare method of detection is usually employed by quantization-based watermarking schemes.

2. Compute a correlation measure $Z$ between $W$ and transform coefficients $\tilde{v}_1, \tilde{v}_2, ..., \tilde{v}_m$ and test if $z$ is greater than the detection threshold $t$. This form of detection is commonly used with additive or multiplicative schemes with the threshold set to $\frac{\kappa}{sm} \sum_{i=1}^{m} |\tilde{v}_i|$, where the role of the detector is to recover the watermark in the corrupted watermarked data $I'_W$ (robust watermarking) or to detect integrity violation (fragile watermarking).

So far we have presented a generic model for a watermarking system. Based on this model different types of watermarking techniques can be developed. In general the various techniques may be classified as shown in Fig. 2.4. In the following section, we give a brief account of these various types of watermarking techniques.

### 2.5 Domain-Based Classification

#### 2.5.1 Spatial Domain Watermarking

Spatial domain watermarking methods are easy to implement, but often fail under image processing attacks. Also the fidelity of the original image data can be severely degraded since the watermark logo is directly applied on the pixel value. Spatial domain methods are also not good at perceptual tuning and they are relatively weak to filtering, compression, and intensity adjustments [61]. On the other hand, transform-domain methods could easily exploit perceptual tuning to achieve good transparency. Often times, they are robust to many digital data manipulations, but weak to the geometrical
transforms. Nevertheless, due to the good transparency and/or strong robustness to various attacks, the transform-domain information hiding is especially promising.

![Fig. 2.4: Classification of watermarking techniques](image)

### 2.5.2 Frequency Domain Watermarking

As pointed out above, frequency domain watermarking methods are relatively more robust than the spatial domain watermarking schemes. The frequency domain of the image is viewed as a communication channel, and watermark is viewed as a signal that is transmitted through it. Attacks and unintentional signal distortions are thus treated as noise that the immersed signal must be immune to. Generally discrete cosine transform (DCT), fast Fourier transform (FFT) and DWT are used as the methods of data transformation. The main strength offered by transform domain techniques is that they can take advantage of special properties of alternate domains to address the limitations of pixel-based methods or to support additional features [62]. For instance, designing a watermarking scheme in the DCT domain leads to better implementation compatibility with popular video coding algorithms such as MPEG. Generally, the main
drawback of transform domain methods is their higher computational requirement. A brief account of two methods viz., DCT and DWT are presented next.

### 2.5.3 DCT Watermarking

The classic and popular domain for image processing is that of the DCT. The DCT allows an image to be broken up into different frequency bands, making it much easier to embed watermarking information into the middle frequency bands of an image. The middle frequency bands are chosen such that they avoid the most visual important parts of the image (low frequencies) without over-exposing themselves to removal through compression and noise attacks (high frequencies) [63]. In DCT domain one can have a 2-D watermark signal $W$, which is embedded in the middle band frequency of $8 \times 8$ DCT block. The $8 \times 8$ DCT coefficients $F(u,v)$ are modulated according to the following equation:

$$I_{W,x,y} = \begin{cases} I_{x,y}(u,v) + k \times W_{x,y}(u,v) & \text{if } u,v \in F_M \\ I_{x,y}(u,v) & \text{if else} \end{cases}$$ (2.7)

Here $F_M$ denotes the middle band frequency coefficients, $k$ the gain factor, $(x,y)$ the spatial domain location of an $8 \times 8$ pixel block and $(u,v)$ the DCT coefficients in the corresponding $8 \times 8$ DCT block.

### 2.5.4 DWT Watermarking

Another possible domain for watermark embedding is that of the wavelet domain. The DWT separates an image into a lower resolution approximation image (LL) as well as horizontal (HL), vertical (LH) and diagonal (HH) detail components. Embedding the watermark in the low frequency band retains the fidelity of the image, whereas embedding the watermark in the high frequency band increases the robustness. Wavelet transform has the excellent properties to minimize the data loss in the frequency transformation of images, to reduce noise and bias generation in medical images, and to provide extra robustness against irregular attacks. A detailed analysis of the various other factors that point to the advantages of DWT for digital image watermarking is presented in chapter 3 (section 3.3.4).
2.6 Human Perception-Based Classification

2.6.1 Robust Watermark

Robust watermarking requires that the watermark must be resistant against non-malicious manipulations. Application fields of robust watermarking include all the situations in which it is unlikely that some one purposely manipulates the host data with the intention to remove the watermark [28]. At the same time, the normal use of data may comprise several kinds of manipulations which must not damage the hidden data. In copyright protection applications, the adoption of robust watermarking may use a copyright protection protocol in which all the involved actors are not interested in removing the watermark.

2.6.2 Semi-fragile Watermark

In some applications robustness is not a major requirement, mainly because the host signal is not intended to undergo any manipulations, except a very few limited number of minor modifications such as moderate lossy compression, or quality enhancement. This is the case, for example, of data labeling for improved archival retrieval, in which the hidden data are only needed to retrieve the host data from an archive, and thereby it can be discarded once the data have been correctly accessed. It is likely, though, that data are archived in compressed format, and that the watermark is embedded prior to compression. In this case, the watermark needs to be robust against lossy coding. In general, one can say that a watermark is semi-fragile if it survives only a limited, well-specified, set of manipulations leaving the quality of the host document virtually intact [54].

2.6.3 Fragile Watermark

A watermark is said to be fragile, if the information hidden within the host data is lost or irremediably altered as soon as any modification is applied to the host signal. Such a loss of information may be global, i.e. no part of the watermark can be recovered, or local, i.e. only part of the watermark is damaged. The main application of fragile
watermarking is data authentication, where watermark loss or alteration is taken as evidence that data are tampered with, whereas the recovery of the information contained within the data is used to demonstrate data origin [54].

2.7 Watermark Requirements

Different applications pose different requirements for watermarking techniques. The various requirements can be categorized the following way:

1. **Fidelity / Imperceptibility** is the measure of similarity between the original and the watermarked content. The similarity can be defined as the distance \( d(I, I_w) \), where the distance is some measure of the distortion introduced by the embedder. A watermark is said to have high fidelity if the degradation it causes is very difficult for a viewer to perceive. The insertion of the watermark in the image has to be performed in such a way that the resulting modifications are not visible to the human eye. This is a very delicate and complex issue because it depends on a variety of interacting factors [25].

2. **Robustness** expresses the resilience of the detector for detecting the watermark in a corrupted data \( I'_w \). It can be measured as the probability of identifying the watermark, \( p(W' = W | I'_w) \). It should be impossible to manipulate the watermark by intentional or unintentional operations.

3. **Fragility** is a complex concept characterizing the ability of the detector to detect the alteration of the watermarked content. It requires maximum robustness against non-malicious attacks: \( \max p(W' = W | I'_w) \), but minimal tolerance for malicious attacks: \( \min p(W' = W | \| I'_w - I_w \| > N_C) \), where \( I'_w \) is the watermarked data by modified malicious attacks that need to be detected.

4. **Capacity** is a fundamental property of any watermarking algorithm, which very often determines whether a technique can be profitably used in a given context or not. Capacity requirements struggle against two other important requirements that is watermark imperceptibility and watermark robustness.

5. The fifth requirement is the *computational cost*. Different applications require the embedders and detectors to work at different speeds. In broadcast monitoring, both embedders and detectors must work in real time so they need
to be fairly fast and should have low computational complexity. On the other hand, a detector for identifying the patient details will be valuable even if it takes more time to find a watermark, as it can guarantee the integrity of the medical image. Such a detector is important enough that the user will be willing to wait.

There is an inherent trade-off between many of these attributes which play a different role in a real-world application. Usually robustness, imperceptibility and watermark capacity have to be traded against each other. There is no general solution for satisfying these conflicting requirements using a single framework.

At the detector, robustness, false positive rate and speed often compete with each other. During detector design, these attributes must be balanced with the desire for speed to meet the application requirements. For example, robustness to geometric distortions can be achieved at the cost of reduced speed. Application requirements also influence the mode of data acquisition at the detector. The data acquisition device often determines the choice of watermarking technology and its capabilities.

The first task of application design is to determine the product requirements and use them to prioritize the various watermarking attributes. Practical applications necessarily involve contradictory constraints and requirements that have to be traded against one another to achieve the intended goals and satisfy the user’s needs.

### 2.8 Multiple Watermarks

In some cases the possibility of inserting more than one watermark may arise. For example, a copyright protection scheme, may require two watermarks: one with the identity of the author of the work and the other indicating the name of the authorized consumer. Of course, algorithms enabling multiple watermark embedding must grant that all the watermarks are correctly read by the decoder. In addition, the insertion of several watermarks should not deteriorate the quality of the host data. Though necessary in many cases, the possibility of inserting more than one watermark must be carefully considered by system designers, since it may produce some ambiguities in the interpretation of the information hidden within the protected piece of work [64, 65].
2.9 Challenges in Medical Image Watermarking

With the increase in the transmission of medical images across multiple user systems, the concern for image security has raised. Proper data integrity, coupled with a method of maintaining strong information binding for teleradiology data, is paramount to ensure this requirement. Digital watermarking offers a possible solution to this problem. Hiding watermark information in medical image data files is one solution for enhancing security and privacy protection of data. But, medical image watermarking demands stringent quality assurance [22, 23]. Typically, watermarks embedded in medical images must not cause any visual artifacts that may affect the interpretation by medical doctors [24 - 27]. Also, patient information embedded must be detected and recovered in an accurate manner. Watermark information must be robust enough to withstand intentional as well as non-intentional attacks. At the same time, the embedding of watermark should not degrade the asset image. Any change to image content must be done in a highly controlled and reliable way so that the clinical decision-making is not compromised [23]. Such a high fidelity requirement renders watermarking in medical images more challenging than multimedia applications where perceptual distortion may be tolerated.

From a mathematical point of view, digital watermarking is a constraint optimization problem. The goal is to maximize the watermark energy under the minimum visibility constraint. This optimization is non-trivial since it requires an accurate description of the human visual system in order to evaluate the visibility of the watermark. Although several vision models have been proposed in the past, most of them are very complex and often very specific to a given application. To ensure imperceptibility of the modification, caused by watermarking embedding, a perceptibility criteria determined by the HVS must be used. As a consequence of the required imperceptibility, the individual samples that are used for embedding can only be modified by an amount small compared to their averaged amplitude.

At the embedder, the main tradeoffs are between visibility, capacity, robustness, and speed. The degree to which human intervention in the embedding process is permitted, impacts both speed of embedding and visibility. Capacity is always in conflict with robustness. Watermark strength (energy of the embedded signal) also
affects both visibility and robustness. The ability to automatically adapt visibility according to media characteristics without sacrificing robustness (or some other set of attributes) is the foremost goal in embedder design [66]. The problem becomes more challenging because medical images of different modalities have different noise characters. Even images of same modality under different environments differ in their noise characteristics. Not all watermarking methods are suitable for all image types and all applications. It is of interest to develop a knowledge-based approach for automated image-adaptive watermarking.

Like most other engineering problems, the design of suitable embedding strength involves multiple, often conflicting, design criteria and specifications. Finding an optimal embedding strength is, therefore, not a simple task. Consequently, there is a need for optimization-based methods that can be used to obtain optimal embedding strength that would satisfy the conflicting requirements. Ideally, the optimization method should lead to the global optimum of the objective function.

2.10 Scope and Aim of the Thesis

With increase in the amount of digital visual data (image, video and 3D object) being sent on the Internet, their security becomes increasingly important for many applications, e.g., confidential transmission, video surveillance, military and medical applications. Especially, medical image protection and authentication are becoming increasingly important in an e-Health environment where images are readily distributed over electronic networks.

A widely accepted fact in generic image watermarking is that not all watermarking methods are suitable for all image types and all applications. Most medical image watermarking research focuses on developing watermarking systems that preserve image fidelity and/or robustness, under typical non-medical image degradation processes (for example, data communication losses). Each application has its own special requirements with regard to robustness, security, imperceptibility and the amount of data to be embedded. The three main technical challenges faced by watermarking algorithms are fidelity, robustness and security. Fig. 2.5 shows how the embedding strength influences the conflict between the basic requirements of
watermarking such as robustness and fidelity. Similarly, capacity also is in conflict with either robustness or imperceptibility (or both). It is thereby mandatory that a good trade-off is found between these various requirements, depending on the application at hand.

![Diagram showing tradeoff between robustness, embedding strength, and PSNR](image)

**Fig. 2.5: Tradeoff between the basic requirements for a digital watermark**

In order to simultaneously address these diverse medical issues, multiple watermarks with varying characteristics and requirements need to be jointly embedded in a single image. These watermarks should be extractable independently from each other, as they are intended to be used for different purposes and at different levels of the health information management chain.

Given the range of medical image types that exist, as well as the need to protect images without undue loss of data, and the knowledge that not all watermarking methods are suitable for all image types, an important gap in medical image watermarking research has been identified. If medical images are to be protected appropriately as they travel from one site to another, a suitably tailored watermarking scheme must be selected for each image type. Solutions should be found for specific medical image types, or for typical degradation processes arising from typical medical
uses, involving image manipulations. There is no universal watermarking technique that satisfies all requirements of all applications. Consequently, each watermarking techniques should be designed within the context of the entire system in which it is to be used. Given its unique problems, medical image watermarking requires nonconventional approach. In this thesis, multiobjective GA is employed within wavelet framework for medical image watermarking.

A few specific aims of the thesis are as follows:

- To develop a novel algorithm for robust public key blind watermarking in wavelet domain for secure transmission of medical images in teleradiology, and to investigate the interplay between the conflicting parameters of fidelity, robustness and the watermark payload.

- To develop a novel multiobjective GA-based image adaptive watermarking algorithm to achieve an optimal trade-off between the conflicting goals of fidelity for diagnostics and robustness for security using a multi objective function, for the first time, for public-key, blind medical image watermarking in wavelet domain.

- To develop a novel, color image hybrid watermarking algorithm for simultaneously achieving the two objectives of security and integrity in medical image transmission, for the first time, by designing a multi gene GA where one gene codes for the embedding strength of the robust watermark and another codes for the number of bits for fragile watermark embedding, based on the design strategy of wavelet subband selection to ensure the orthogonality of the robust and fragile watermarks.

- To develop joint, image-adaptive compression and watermarking algorithm to enhance the potential of watermarking in teleradiology, by exploiting the localization property of wavelet transform and using multi-objective GA to achieve optimal trade-off between transmission time and security.

- To develop a user-friendly graphical user interface for PC-based implementation of these algorithms, and to evaluate the performance of
these algorithms qualitatively as well as quantitatively using real medical images from different modalities, so that the challenges and tradeoffs in developing watermarking applications can be best understood.

2.11 Overview of the Thesis

The rest of the thesis is structured as follows: In Chapter 3, a robust blind public key watermarking algorithm is developed for image security in teleradiology, based on a generic model for digital watermarking in wavelet domain. A novel design strategy integrating symmetric algorithm, key concealing, partial selection and encryption of embedding data using pseudo random sequence is proposed to limit the scope of possible invertible, copy and cryptographic attacks. Both subjective and objective methods are used to evaluate the watermarked image quality and the robustness of the watermark in medical images of three different modalities. The proposed method is autonomous and it ensures the integrity, because the key is spread over the whole image. The algorithm is designed to take advantage of the combination of the benefits using lifting wavelets and a pseudo-random permutation to disperse the spatial relationship of watermark, making it hard for the hacker to remove the watermark.

An innovative algorithm for optimal embedding strength selection to achieve conflicting goals of fidelity and robustness in blind medical image watermarking in wavelet domain is developed in Chapter 4. A multi-parameter fitness function for optimal embedding strength selection is developed, for the first time, for public-key, blind medical image watermarking in wavelet domain. This multiobjective GA-based image adaptive watermarking algorithm optimizes the embedding strength to achieve an optimal trade-off between the conflicting goals of fidelity for diagnostics and robustness for security, using a multi objective function. The fidelity parameter, PSNR and the robustness parameter, NC are used to design the multiobjective fitness function.

In Chapter 5, a novel color image hybrid watermarking algorithm that combines robust and fragile watermarks is presented to achieve simultaneously the two objectives of security and integrity in medical image transmission. To address this problem, for the first time, a multi gene approach, one gene coding for the embedding strength of robust watermark and another coding for the number of bits for fragile watermark embedding,
is developed in this chapter. This multi-objective, multi-gene, hybrid watermarking scheme is evaluated for its potential to preserve security and integrity in fMRI, color echocardiograms and fundus images. The novelty of the algorithm lies in the optimal selection of the watermarking parameters. The design strategy of the hybrid watermarking scheme uses wavelet subband selection to ensure the orthogonality of the robust and semi-fragile watermarks.

Teleradiology applications require the compression of medical images for their fast transfer from one location to other. Chapter 6 attempts to investigate, for the first time, the application of GA in achieving an optimal compression ratio for dual watermarking of color medical images without degrading the image quality. Two image processing applications, compression and watermarking are coupled to enhance the potential of watermarking in teleradiology, by exploiting the localization property of wavelet transform using multi objective GA.

The thesis is concluded in Chapter 7 with a summary highlighting the outcome of the thesis and pointing out possible future enhancements.