COLOR IMAGE WATERMARKING USING MULTIDIMENSIONAL FOURIER TRANSFORMS

by

Tsz Kin Tsui
B.A.Sc., University of Toronto, Toronto, 2003

A thesis
presented to Ryerson University
in partial fulfillment of the
requirement for the degree of
Master of Applied Science
in the Program of
Electrical and Computer Engineering.

Toronto, Ontario, Canada, 2006

Author's Declaration

I hereby declare that I am the sole author of this thesis.
I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.
Signature

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.
Signature

Instructions on Borrowers

Ryerson University requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.

Abstract

Color Image Watermarking using Multidimensional Fourier Transforms
M.A.Sc., Electrical Engineering, Ryerson University, Toronto, 2006

This thesis presents two vector watermarking schemes that are based on the use of complex and quaternion Fourier transforms and demonstrates, for the first time, how to embed watermarks into the coefficients consistent with our human visual system (HVS). Watermark casting is performed by estimating the Just-Noticeable distortion (JND) of the images, to ensure watermark invisibility.

The first method encodes the chromatic content of a color image as CIE a^*b^* chromaticity coordinates whereas the achromatic content is encoded as CIE L tristimulus value. Color watermarks (yellow and blue) are embedded in the frequency domain of the chromatic channels by using the Spatio Chromatic Discrete Fourier Transform (SCDFT). It first encodes a^* and b^* as complex values, followed by a single discrete Fourier Transform. The most interesting characteristic of the scheme is the possibility of performing watermarking in the frequency domain of chromatic components.

The second method encodes the $L^*a^*b^*$ components of color images and color watermarks are embedded as vectors in the frequency domain of the channels by using the Quaternion Fourier Transform (QFT). The idea is twofold: Robustness is achieved by embedding a color watermark in the coefficient with positive frequency, which spreads it to all components in the spatial domain. On the other hand, invisibility is satisfied by modifying the coefficient with negative frequency, such that the combined effects of the two are insensitive to human eyes.

Acknowledgments

I would like to express my appreciation to my supervisors Dr. Xiao-Ping Zhang, and Dr. Androutsos for their excellent guidance, helpful feedback, and continuous support throughout my research.

I also would like to thank the Department of Electrical and Computer Engineering and the School of Graduate Studies of Ryerson University for their support in terms of departmental facilities, research stipend, and scholarship.

In addition, thanks to Dr. Guan, Dr. Alirezaie, and Dr. Krishnan for becoming the members of the committee for my oral defense.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

THE success of the internet in the 90s has made profound changes to our society and our daily lives. In fact, the process of transforming the world from the analog to digital is accelerated by the internet. Along with powerful software, new electronic devices such as the cellular phone, PDA, laptop computer, have reached customers worldwide, which makes the sharing of digital data possible, and convenient. However, this gives us a big concern - how to secure the data, and avoid improper use by unauthorized people? Statistics show that the music and video industry loses billions of dollars per year, due to illegal copying and downloading of copyrighted materials from the internet. Researchers have proposed an effective way to address this issue - Digital Watermarking. The term watermark is derived from the German term *wassermarke* [1]. It is actually a misnomer, because the element water is not very important in the creation of a mark. The important part is that the mark resembles the effects of water on paper. Traditionally, a watermark is a recognizable image or pattern in paper that appears lighter when viewed by transmitted light. The purpose is to identify the originality of the paper, which is widely used as anticounterfeiting measure on money and other documents. In the digital world, the basic idea is to embed some secret data in a digital content to be protected and "seal" it within the content. Figure 1.1 is a generic watermarking insertion process. Original digital media is inserted into a watermarking system, using a key. The system produces a watermarked work. Figure 1.2 shows a generic watermarking extraction process. Ideally, the extracted watermark

should be identical to the original watermark. However, attackers would try to destroy the watermark by applying some digital image processing techniques such as adding noise. Thus, the robustness of a watermarking system against attacks is a very important issue to deal with.

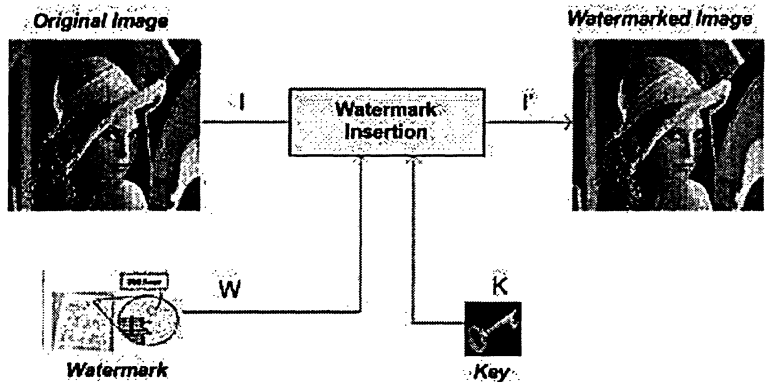


Figure 1.1: Generic watermark insertion

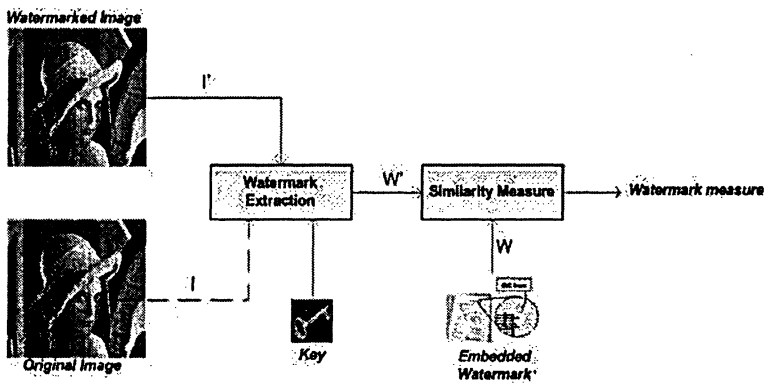


Figure 1.2: Generic watermark extraction

Watermarking techniques can be classified into many categories. They are going to classified according to their **inserted media category, perceptivity, robustness, inserting media type,** and **processing method.**

1.1 Classification on Watermarking Techniques

1.1.1 Inserted media category

Digital work can be in a variety of different formats, specifically text, audio, image, and video. Therefore, watermarking algorithms can be subdivided into *Text Watermarking*, *Image Watermarking*, *Video Watermarking*, and *Audio Watermarking*.

Text Watermarking:

A text document is binary in nature, and composed by characters and spaces, which form a series of text lines. The existing text watermarking algorithms can be subdivided into three categories: *Line-shift coding* - shifts the position of text lines for encoding in the vertical direction; *Word-shift coding* - shifts the position of words in the vertical direction; and *Feature coding* - chooses certain text features and modifies those features. Examples of algorithms marking text documents can be found in [2] [3] [4] and [5]. Among all watermarking methods, text watermarking does not get much attention, due to its limited features, and poor gray-scale information. Nevertheless, it still provides an effective solution to discourage illicit distribution of digital text documents.

Image Watermarking:

Image watermarking algorithms embed information into images, which can be detected and/or extracted later for copyright identification and ownership confirmation. This category is by far the most popular, and will be the main focus of this thesis. The simplest way to watermark an image is to change directly the values of its pixels in the spatial domain. A more advanced way to do it is to insert the watermark in the frequency domain of the image. Since it contains rich gray-scale or color information, existing algorithms usually utilizes the characteristics of the Human Visual System (HVS) to increase the performance of systems. Examples of image watermarking algorithms can be found in hundreds of papers, such as [6], [7], [8], and [9]. In the rest of the thesis, image watermarking will be the main focus.

Video Watermarking:

It is very similar to image watermarking, so it is often treated as an extension of it. The only difference is that it also considers the temporal information and requires real time extraction. A watermark can be added to the raw data (uncompressed data), such as standard uncompressed AVI video, or a compressed bit-stream (MPEG2). More literature can be found in [10], [11].

Audio Watermarking:

Since the rise of the MP3 format, audio watermarking has become a hot issue. Instead of the HVS, this application has to consider the characteristics of the Human Auditory System (HAS) to achieve higher performance. Researches that deals with audio watermarking can be found in [12] [13] [14]. Notice that audio data is one-dimensional, thus, the complexity of audio watermarking systems is usually lower than image systems.

1.1.2 Watermarking Perceptivity

Watermarking techniques can be classified into two types: *Visible*, and *Invisible* watermarks. Invisible watermarks are the most frequently used approach, because it does not degrade the original contents.

Visible Watermarking: It is mainly used with logo and trademark labels, which contain the copyright information about the media. The advantage is that it only requires little processing to do, but it is very fragile and susceptible to attack. Moreover, it degrades the quality of the contents.

Invisible Watermarking: To make the watermarks invisible, the characteristics of the HVS or HAS have to be thoroughly analyzed and modelled. As a result, it requires more computational power and time. The advantage is that it is more difficult for attackers to detect the existence of the watermarks, as a result, it is less vulnerable to attacks.

1.1.3 Robustness

In order for the watermarks to survive under external attacks, robustness is the main issue. Depending on the degree of the robustness, it can be classified as: *Robust Watermarking*, and *Fragile Watermarking.*

Robust Watermarking: The main purpose is to withstand a wide range of possible attacks that the digital work might experience. The Bit Error Rates (BERs) is used to measure the robustness of a given algorithm.

Fragile Watermarking: It is designed to be easily removed or destroyed, if a watermarked image is modified, even it is slightly modified. It is helpful for protection and verification of the original contents.

1.1.4 Inserting watermark type

Depending on the applications, the format of the embedded watermarks can also be different, listed as follows:

Pseudo-noise (PN) sequence: A pseudo-random signal is a deterministic signal that appears to an authorized listener, or observer to be truly random signal. As a result, it's ideal to be used for watermarking because it is statistically difficult to detect. Another advantage of using it is that it is very easy for both embedder and detector to regenerate, as long as the seed of the sequence is available.

Image: The embedder can also generate watermarks in the format of images, such as a binary image, stamp, logo and label. The reason of using this approach is that the extracted watermarks can be visually distinguished, and compared. Different metrics can be used to measure the similarity between the original and extracted watermarks.

1.1.5 Processing method

Watermarking processing methods are classified into two big categories: *spatial domain*, and *frequency domain.*

Spatial domain:

Least Significant Bit (LSB): The earliest work of digital image watermarking schemes can be traced back to the early 90s [6], where the idea is to embed watermarks in the LSB of the pixels. Given an image with $N \times N$ pixels, and each pixel is represented by a 8 bit sequence $(s_1, s_2, ..., s_8)$, the watermarks are embedded in the s_8 of selected pixels of the image. It is easy to implement, and does not generate serious distortion to the image. However, it is not very robust against attacks. For instance, an attacker could simply randomize all LSBs, which effectively destroys the hidden information.

Patchwork [14] [12] [13] : The objective is to answer the question: does this image contain a watermark or not? The process begins by pseudo-randomly choosing n ordered pairs (a_i, b_i) of pixels of the image, followed by the following operation:

$$(a_i, b_i) \leftrightarrow (a_i + 1, b_i - 1) \tag{1.1}$$

The extraction process for the patchwork watermark proceeds by choosing the same n ordered pairs of pixels. The receiver then sums the value $(ai - bi)$ for all n values of i. The watermark is detected if this sum is approximately $2n$, and is not detected otherwise.

Frequency domain:

Compared to spatial domain methods, frequency domain methods are more widely applied. The idea is to embed the watermarks in the spectral coefficients of the image. The most commonly used transforms are the Discrete Cosine Transform (DCT), Discrete Fourier Transform (DFT), Discrete Wavelet Transform (DWT), Discrete Laguerre Transform (DLT) and the Discrete Hadamard Transform (DHT) [15]. The reason for doing watermarking in the frequency domain is that the characteristics of the HVS are better captured by the spectral coefficients. For example, human eyes are more sensitive to low frequency coefficients, and less sensitive to high frequency coefficients [16]. In other words, low frequency coefficients are perceptually significant, which means alternations to those components might cause severe distortion to the original image. On the other hand, high frequency coefficients are considered

insignificant, thus, processing techniques such as compression tend to remove high frequency coefficients aggressively. To get a balance between imperceptibility and robustness, most algorithms embed watermarks in the mid-range frequencies.

DCT-based approach: Cox [11] first proposed an algorithm that inserts the watermarks into the spectral components of the image using techniques analogous to spread spectrum communications [17]. The algorithm is to place the watermarks into a set of frequency components that are perceptually significant. It has been shown that the watermarks are very hard to detect because they consists of relatively weak noise signals. In addition, placing the watermarks in the frequency domain would spread them over all pixels, which increases the robustness, and reliability against unintentional or intentional attack.

Similar algorithms to do watermarking in the DCT domain were proposed in [9] [18]. They incorporate a Just Noticeable Difference (JND) profile [19] to determine the maximum amount of the watermark signal that can be tolerated at each region in the image without degrading its visual quality. The estimation of the JND is usually based on the properties of the HVS, such as spatial masking, and luminance contrast [20].

DWT-based approach: Mike Wang [21] proposed a wavelet-based watermarking algorithm based on the principle of Multi-Threshold Wavelet Codec (MTWC). It searches the significant wavelet coefficients to embed the watermarks in order to increase robustness. Zeng [22] proposed an adaptive watermarking algorithm according to the JND threshold. Kaewkamnerd [23] proposed a wavelet based scheme that employs the characteristics of the HVS to determine the weighting function to adjust the watermark strength.

DLT-based approach: In 2000, Piva [8] proposed a watermarking scheme using the DLT, and DFT together. The DLT is first applied to the RGB components of the image, then watermarking is performed independently in the DFT domain of the KL-transformed bands. Watermark embedding is achieved by modifying the magnitude of mid-frequency DFT coefficients. A Bayesian decision is used to optimally detect the presence of the watermark by the detector. In the same year, Gilani [7] also conducted extensive performance comparisons between DLT-and DCT domain watermarking schemes. He concluded that the quality of

the DLT-domain watermarked images is higher than the corresponding DCT-domain watermarked images, while maintaining the same robustness.

DHT-based approach: Ho [24] and Gilani [25] proposed an algorithm in 2002 that uses the DHT for data embedding. The advantage is that it offers much shorter processing time and easier hardware implementation. Thus, it is optimal for real-time applications.

DFT-based approach: Fourier coefficients have two components, phase and magnitude. Experiments have shown that attacks such as geometric rotation do not modify the phase information of the coefficients. Therefore, many DFT-based algorithms have been proposed that are robustness to translation, rotation, and scaling attacks. Licks [26] algorithm is one example.

Having introduced different watermarking classifications, the proposed schemes in this thesis belong to the categories: **Inserted media category** - image, **Perceptivity** - invisible, **Robustness** - non-fragile, **Inserting watermark type** - image, and **Processing method**: frequency domain.

1.1.6 Applications

Digital image watermarking plays an important role to a multitude of applications. The most significant are.

Copyright Protection: Copyright is protection that covers published and unpublished literary, scientific and artistic works, whatever the form of expression, provided such works are fixed in a tangible or material form. Its significance becomes more important for digital or electronic work because they can be reproduced easily. By embedding some secret signatures in an image, attackers cannot easily remove them without destroying the work itself [27].

Fingerprinting and Traitor Tracing: It is similar to the copyright protection, except that the purpose of embedding the signatures is to trace back unauthorized use of an image to its original recipient [28].

Transaction Tracking: An image is usually produced, copied, modified and distributed multiple times. The embedded watermark records one or more transactions that have taken

place in the history of the image, which makes it possible to retrace these steps.

Media Forensics: The objective is to identify if there has been any illegal processing applied to an image [29].

1.2 Contribution of the thesis

One of the main challenge of the watermarking problem is to achieve a better tradeoff between the robustness and perceptivity. From the engineering's perspective, they are two conflicting requirements that cannot be satisfied at the same time. The robustness can be achieved by increasing the strength of the embedded watermark, but the visible distortion would be increased as well. To deal with this problem effectively, the characteristics of the HVS have to be considered thoroughly in order to increase the perceptivity. Even though many HVS-based watermarking schemes have been invented, only a small portion of them were designed for marking color images, due to a number of reasons: Firstly, an objective and perceptual metric for color distortion is lacking, as a result, it is very hard to quantify how similar two images are. Secondly, a holistic transform that can process the color information of an image in a vector manner has not existed. Thus, the thesis tries to exploit the chromatic information of color images, find appropriate transforms to bring the information to the frequency domain, and be able to embeds watermarks as vectors into the chromatic as well as the luminance channels. The detailed algorithms have been accepted to the ICASSP 2006 [30] , and ACM MM 2006 [31].

1.3 Outline of the thesis

The thesis consists of 5 chapters which are organized as follows:

Chapter 2: **Mathematical Models of Watermarking** Chapter 2 provides a mathematical formulation to model a generic watermarking scheme.

Chapter 3: **Spatio-Chromatic Discrete Fourier Transform** Chapter 3 offers a background on a new transform, Spatio-Chromatic Discrete Fourier Transform, and its applications to various fields in color image processing such as image filtering.

Chapter 4: **Quaternion Fourier Transform** Chapter 4 offers a background on another higher dimensional transform, Quaternion Fourier Transform, and how it can be applied to edge detection and watermarking.

Chapter 5: **New Color Image Watermarking using Multidimensional Transforms** Chapter 5 introduces the main contribution of my work. It is divided into two parts: the first part explains a new color watermarking scheme using the SCDFT, and the second part extends the work using the QFT. Simulation results are also provided to compare their effectiveness.

Chapter 6: **Conclusion and Future Work** Chapter 6 gives a conclusion summarizing the proposed algorithms and hints some directions for further extension of the current work.

Chapter 2

Mathematical Models of Watermarking

FROM a mathematical point of view, watermarking is in essence a form of communication. The objective is to communicate a message from the watermark embedder to the watermark detector. Thus, it is natural to apply the traditional model of a communications system to watermarking, which is called the *Communication-Based Model of Watermarking* [32].

2.1 Communication-based Models of Watermarking

A typical image watermarking framework is illustrated in Figure 2.1. The dash lines imply that the components are optional, depending on the design of the system. The objective is to embed a message, m, in the original cover image, I_o, such that the marked image, I_w, is perceptual identical to I_o. The difference between I_o and I_w is the distortion, D introduced during the embedding process.

The embedded message, m will be extracted from the marked image, I_w, by a detector. Most likely, the marked image has gone through some attacks and processing, denoted as I_{wn}. The difference between the I_w and I_{wn} is called noise, n, represented as D_2. Ideally, the extracted message m_n should be exactly the same as m. The percentage of error between m and m_n is usually measured by a quantity called the Bit Error Rate (BER) [32]. Given two bit sequences, BE and BE', with length N, the BER is defined as:

$$\frac{\sum_i |BE(i) - BE'(i)|}{N} \times 100\% \tag{2.1}$$

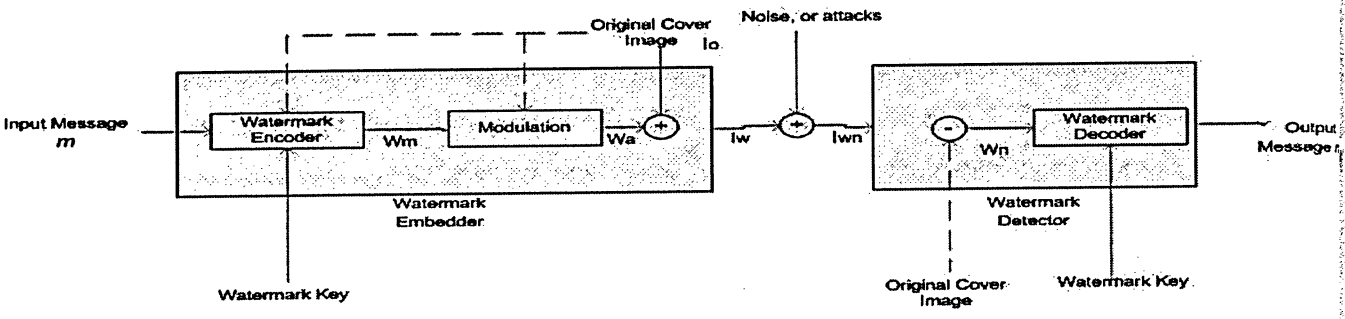


Figure 2.1: A basic watermarking system.

2.2 Basic Message Embedding and Detection

Now let us examine more details of the embedder block in the figure. In general, it consists of three steps.

1. The message is mapped into an *added pattern*, w_a, which usually has the same dimension as the cover image. This could be done with a watermark key.

2. To obtain w_a, a set of predefined *reference patterns* w_{r0}, w_{r1}, w_{r2}... is used to produce a pattern that encodes the underlying message, which is called a *message pattern* or *message mark* w_m.

3. The message pattern is modulated to produce w_a.

4. w_a is added to the original cover image I_o, to produce the watermarked image I_w.

2.2.1 Mapping Messages into Message Vectors

Source Coding is the first process in traditional communications systems, and it maps messages into sequences of symbols. Watermarking systems also need a mechanism to achieve

the same task. There are two common ways to do it, they are *Direct Message Coding*, and *Multi-symbol Message Coding.*

Direct Message Coding is the most straightforward approach [32]. The idea is to assign a unique, predefined message mark to represent each message. Given a set of messages M, and $|M|$ be the number of messages in the set, a set of $|M|$ message marks is defined, denoted by MM, such that for every message in M, there is a corresponding message mark in MM to represent it. The rule of thumb for designing the message marks is to maximize the code separation (i.e., marks that are far apart from one another in marking space), because that prevents the marks from being erroneously decoded as a different message. For a system with only two messages, the best result is achieved by defining one message mark to be the negative of the other.

Direct Message Coding is simple, and effective, but it does not work well if $|M|$ is large. The detector must compare the distance of the decoded mark with all marks in the set MM, and choose the most probable one. The computational complexity grows rapidly for large $|M|$. This problem can be solved by using Multi-symbol Message Coding [32]. The idea is to represent each message with a sequence of symbols, drawn from an alphabet, A. Each symbol in a sequence can then be embedded and detected independently with its own reference mark. If each message is represented with a sequence of R separate symbols, drawn from the alphabet $|A|$, at most $|A|^L$ messages using this method can be represented. Detecting these messages requires only a fraction of the processing required with direct message coding. For example, suppose there is an alphabet of $|A| = 3$ symbols and a message length $R = 7$. This can represent up to $|A|^R = 2187$ different messages. The detector only requires comparison with each of the three reference marks to detect each of the seven symbols, resulting in 21 comparisons in total, which is much less than the 2187 required in a direct message coding method.

2.2.2 Error Correction Coding

Depending on the design of the message marks, they might not always have good code separation, especially when $|M|$ is large using the direct message coding. The problem motivates the introduction of error correction codes [32].

The idea of error correction coding is to increase the length of the symbol sequences by adding extra information in order to increase the probability of correct decoding. Sequences that correspond to messages are called *codewords*, where the rest are referred to *corrupted codewords*. There are many error correction codes available, such as Repetition Codes, Hamming Codes, Turbo Codes, and Trellis Codes [32].

2.2.3 Modulation

After the encoding step, each message mark is modulated and/or scaled by a factor s, yielding a modulated watermark w_a before adding to the original image I_o. The purpose of doing that is to ensure the embedding distortion will be small enough to be imperceptible, and at the same time, strong enough to be robust against attacks. The value of s can be determined dynamically with or without using the information provided by the original image.

2.2.4 Correlation-based Detection

To detect the embedded message, the detector has to detect the presence of the signal w_m in the received, possibly corrupted image. Depending on the applications the original image I_o may or may not be available to the detector. In the watermarking system shown in Figure 2.1, if there is a link connecting the original cover image to the detector, the system is called an *informed detection system*, otherwise, it is a *blind detection system*. The detection methodologies are different for both types. The underlying reason is because in an informed detecting system, the addition of the image in the embedder can be exactly cancelled out by its subtraction in the detector, the only interference is caused by the noise signal, or external attacks. In a blind detection system, the entire unwatermarked image is unknown, therefore, the watermarked image might be corrupted by the addition of a single noise pattern formed

by the combination of the original image and the noise signal. Under this circumstance, the linear correlation of the received image and the message pattern w_m must be calculated, which is defined by:

$$LC(w_m, I_{wn}) = \frac{1}{X.Y} w_m . I_{wn} = \frac{1}{X.Y} \sum_{i,j} w_{m(i,j)} . I_{wn(i,j)}, \tag{2.2}$$

where $w_{m(i,j)}$ and $I_{wn(i,j)}$ represent pixel values at location i,j in w_m and I_{wn}; X and Y represent the image dimensions. A threshold can be used to determine if the image contains the watermark or not.

2.3 Watermarking with Side Information

2.3.1 Informed Embedding

The embedding components of the watermarking systems described so far do not use the information provided by the original image I_o. The values of s is independently selected of I_o. Obviously the embedder has access to I_o, thus, this kind of scheme does not take full advantage of that information. To improve the watermark embedding performance, a new kind of system, called *informed embedding*, uses the information about I_o to generate s.

An informed embedding watermarking system can be casted as an optimization problem: Given I_o, select the strength s to maximize or minimize a parameter or property, such as fidelity, imperceptibility, or robustness. For example, the objective of the proposed algorithms in chapter 5 is to improve the robustness of the watermark while maintaining imperceptibility. This can be done by taking advantage of imperfections of the Human Visual System (HVS) and its inability to recognize all the changes equally. The characteristics of HVS such as frequency masking, luminance masking have been captured in many models [22] [19]. The details will be discussed in chapter 5.

2.3.2 Informed Coding

In the previous sections, the reference pattern, w_r is generated independently from the cover image I_o. This approach can also be improved by utilizing the side information about I_o during the watermark encoding and creation process to choose between several available alternative watermarks, and select the one that will cause the least distortion of I_o. This technique is called *informed coding.* This topic is, however, outside the scope of this thesis. The interested reader is directed to [32] for details.

2.4 Attacks

The most important requirement of an embedded watermark is that it must be able to survive under external attacks. Possible types of attacks are *image compression*, *geometric transformations*, *image enhancements* and *information reduction.*

• **Image Compression** - For images, JPEG is the most popular format to achieve image compression. Since it quantizes the high-frequency coefficients aggressively, the watermarks have to be very strong if the insertion process was done by modifying those coefficients.

• **Geometric transformations** - They include rotation, translation, sheering, or resizing of an image.

• **Image Enhancements** - Sharpening, color correction/calibration, contrast modifications, and histogram equalization belong to this category.

• **Information Reduction** - Cropping, color gamut reduction Image filtering and the introduction of noise.

The proposed algorithms given in Chapter 5 of this paper were not designed to optimally deal with a specific attack. They try to protect against attacks by maximizing the strength s of the watermark, so that it can survive under any kind of external interference.

2.5 Evaluation of Watermarking Systems

To evaluate the performance of a watermarking system, some common metrics are given here:

Fidelity: The embedded watermarks should be invisible in the cover images. In mathematical sense, a function D to measure the distortion between I_o and I_w is defined. The most commonly used metric is the mean square error (MSE) function [32], which is defined as:

$$D_{mse}(I_o, I_w) = \frac{1}{N^2} \sum_{i}^{N} \sum_{j}^{N} (I_o(i,j) - I_w(i,j))^2, \tag{2.3}$$

where i, j are the spatial coordinates, and N is the size of the images. MSE is often used as a rough test of a watermark system's fidelity, because it is known to provide a poor estimate of the true fidelity. Therefore, an improved metric will be defined in chapter 5 that takes the luminance as well as the chrominance information into consideration.

Data Payload: It refers to the amount of information (bits) embedded in the images, and is often normalized by the number of samples of the host images.

Robustness: It measures the ability of the watermark to survive against common signal processing operations such as compression, filtering, histogram equalization and noise addition. The Bit Error Rate (BER) is often used to measure the robustness of a watermarking system.

Chapter 3

Spatio-Chromatic Image Processing (SCIP)

THE traditional methods for dealing with color images simply process the channels separately, which does not consider the correlations between the color channels. In recent years, researchers have developed some multidimensional transforms that can holistically process the images in the frequency domain. The Spatio-Chromatic Discrete Fourier Transform (SCDFT), proposed by McCabe in 1997 [33], is the one that is important in my thesis. The color watermarking schemes presented in this thesis are built on top of his work. Thus, it is vital to fully understand the concept and its applications.

3.1 Theory of Spatiochromatic Image Processing

The idea of spatio-chromatic image processing is that it introduces a new two-dimensional transform, SCDFT, to process color images holistically. It is actually derived from the Discrete Fourier Transform (DFT). The new idea is that it analyzes the chromatic information of color images via the DFT, and provides a new framework to interpret its spectral coefficients.

3.1.1 Discrete Fourier Transform (DFT) and its spectral interpretation

The DFT is defined as:

$$F(P,Q) = \sum_{p=-\frac{N}{2}+1}^{\frac{N}{2}} \sum_{q=-\frac{N}{2}+1}^{\frac{N}{2}} f(p,q) e^{\frac{(-2\pi j(Pp+Qq)}{N})}, \tag{3.1}$$

and the inverse DFT is defined as:

$$f(p,q) = \sum_{P=-\frac{N}{2}+1}^{\frac{N}{2}} \sum_{Q=-\frac{N}{2}+1}^{\frac{N}{2}} F(P,Q) e^{\frac{2\pi j(Pp+Qq)}{N})}, \tag{3.2}$$

where (p,q) are image spatial coordinates, (P,Q) are spatial frequencies, and $f(p,q)$ is an arbitrary discrete signal. In general, $f(p,q)$ is a complex quantity.

McCabe uses the concept of *phasors* to interpret the DFT coefficients. A phasor is simply a rotating vector in the complex plane centered at the origin and described by a magnitude and angle (phase). Consider the one-dimensional case where $f(x) \leftrightarrow F(k)$, the frequency coefficients of the DFT for input data contain positive and negative complex-valued frequency pairs,

$$F(k) = Re(k) + jIm(k), \tag{3.3}$$

and

$$F(-k) = Re(-k) + jIm(-k). \tag{3.4}$$

For the fundamental frequency $k = 1$, it can be interpreted as a clockwise rotation about the origin, starting at a position given by $\phi_+ = tan^{-1}(\frac{Im(k)}{Re(k)})$, for one pass along the original data from $x = 0$ to $x = N - 1$, with the increment of $\frac{2\pi}{N}$. Similarly, for a component with negative frequency, the path is counterclockwise, and the initial angle is $\phi_- = tan^{-1}(\frac{Im(-k)}{Re(-k)})$. The magnitude of the vectors are defined as:

$$||F(k)|| = \sqrt{Re(k)^2 + Im(k)^2} \tag{3.5}$$

and

$$||F(-k)|| = \sqrt{Re(-k)^2 + Im(-k)^2}. \tag{3.6}$$

The vector sum of the two components produce the resultant path around the origin. For increasing frequency values of k, the number of revolutions about the origin increases: two revolutions take place for $k = 2$, three revolutions for $k = 3$, and so on. With the permission by the author of [33], Figure 3.1 is reproduced to illustrate the concept by showing three different cases of $F(k)$ and $F(-k)$.

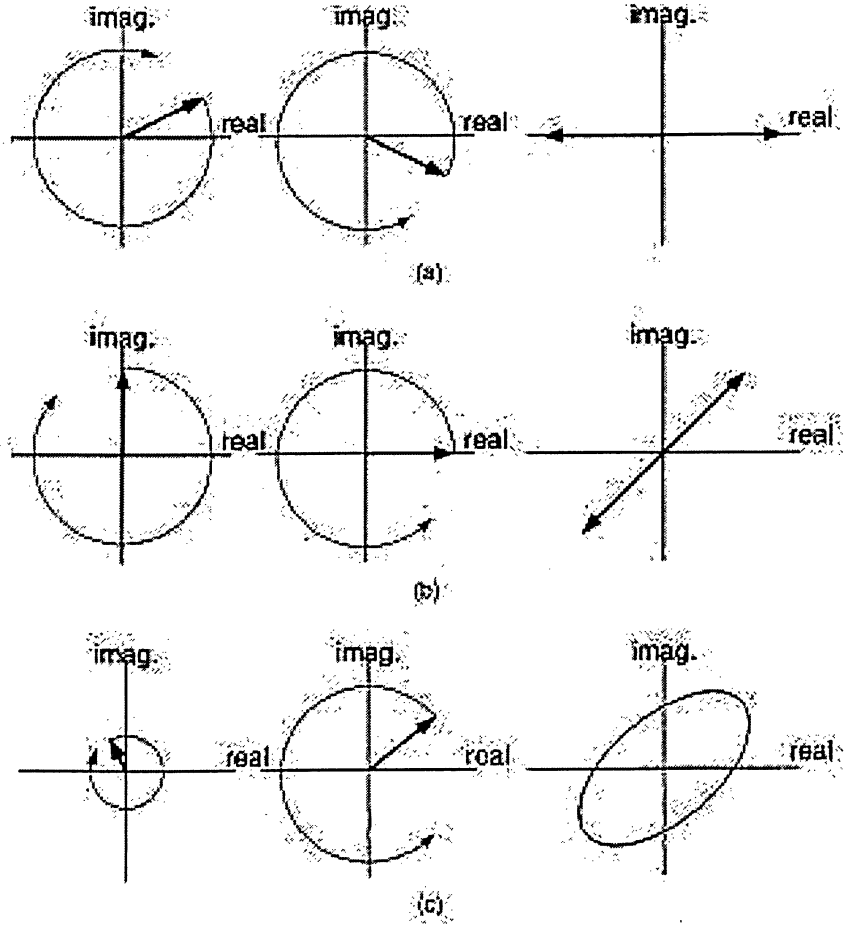


Figure 3.1: Phasor diagram representation for a complex-input DFT showing the two component phasors with different magnitudes and phases. The resultant path is shown on the right. (a) Real input, (b) complex input: same magnitude, different phase, (c) complex input, different magnitude and phase.

Case 1: $F(k)$ and $F(-k)$ are real values, then at every increment, the imaginary components of the coefficients cancel each other, and the resultant path will be a straight line

varying along the real axis, as shown in figure 3.1a.

Case 2: $F(k)$ and $F(-k)$ are complex values with the same magnitude but different initial phase, then the resultant path will also has a straight-line motion, but the angle of the motion will be a combination of the two component phases. The phase of the resultant path is $\phi_+ + \phi_-$, as shown in figure 3.1b.

Case 3: The most general case where $F(k)$ and $F(-k)$ do not have any symmetry. The resultant path may take the shape of a circle, an ellipse, or a straight line, as shown in figure 3.1c.

The concept can be interpreted from the mathematical point of view. Consider a one-dimensional case of Eq. (3.2), which is defined as:

$$\begin{aligned} f(x) &= \sum_{k=-\frac{N}{2}+1}^{\frac{N}{2}} F(k)e^{(\frac{2\pi jkx}{N})} \\ &= F(-\frac{N}{2}+1)e^{(\frac{2\pi j(-\frac{N}{2}+1)x}{N})} + F(-\frac{N}{2}+2)e^{(\frac{2\pi j(-\frac{N}{2}+2)x}{N})} + ... + F(0) + \\ &\quad F(\frac{N}{2}-2)e^{(\frac{2\pi j(\frac{N}{2}-2)x}{N})} + ... + F(\frac{N}{2}-1)e^{(\frac{2\pi j(\frac{N}{2}-1)x}{N})} + F(\frac{N}{2})e^{(\frac{2\pi j(\frac{N}{2})x}{N})}. \end{aligned} \quad (3.7)$$

In general, $F(k)$ is a complex quantity. Let $|M(k)|$ be its magnitude, and $e^{j\theta(k)}$ be its phase. Consider only the fundamental frequency $k = \pm 1$, and assume $F(k) = 0$ for all k's other than $k = \pm 1$. Then Eq. (3.7) can be reformulated as:

$$f(x) = |M(1)|e^{j\theta(1)}e^{(\frac{2\pi jx}{N})} + |M(-1)|e^{j\theta(-1)}e^{(-\frac{2\pi jx}{N})}. \quad (3.8)$$

From Eq. (3.8), there are two phasors rotating in the opposite direction. Initially when $x = 0$, $e^{(\frac{2\pi j0}{N})} = 1$, the positive phasor with magnitude $|M(1)|$ starts at angle $\theta(1)$. At each increment of x, the phasor moves to a position with angle $\theta(1) + x\frac{2\pi j}{N}$, until it reaches the angle $\theta(1) + (N-1)\frac{2\pi j}{N}$. The negative phasor behaves similarly, except that it is moving in the negative direction. The trajectory looks like the phasors are rotating about the origin, as shown in Figure 3.1. For the fundamental frequency $k = 2, 3$ ..., the only difference is the

phasors increment by an angle of $\frac{4\pi j}{N}$, $\frac{6\pi j}{N}$... at each increment of x. With this model, $f(x)$ can be viewed as superpositions of a series of phasors with different fundamental frequencies.

3.1.2 Relationship between the DFT and SCDFT

The SCDFT only uses the chromatic components, for example, the CIE a^*b^* chromaticity space (as shown in Figure 3.2) of a color image and analyzes the contents via the two-dimensional DFT, defined as:

$$A(P,Q) + jB(P,Q) = \sum_{p=-\frac{N}{2}+1}^{\frac{N}{2}} \sum_{q=-\frac{N}{2}+1}^{\frac{N}{2}} [a^*(p,q) + jb^*(p,q)]e^{\frac{(-2\pi j(Pp+Qq)}{N})}, \tag{3.9}$$

, and the inverse SCDFT is:

$$a(p,q) + jb(p,q) = \sum_{P=-\frac{N}{2}+1}^{\frac{N}{2}} \sum_{Q=-\frac{N}{2}+1}^{\frac{N}{2}} [A(p,q) + jB(p,q)]e^{\frac{(2\pi j(Pp+Qq)}{N})}, \tag{3.10}$$

where N x N is the size of the image, (p,q) are the image spatial coordinates, (P,Q) are the frequency coordinates, and (A,B) are the real and imaginary values occurring at the spatial frequency (P,Q), which determines how the colors in the image change spatially.

Using the concept of phasors, the SCDFT coefficients can be interpreted as rainbow gratings rotating along the circular path containing the a^* and b^* components, as shown in Figure 3.2. One important observation is that the magnitude of the SCDFT coefficient, $A(P,Q)+jB(P,Q)$, can be used to represent the saturation of the color, and the phase (angle) represent its hue. Figures 3.3 and 3.4 show examples of basis functions produced when either the negative or the positive frequency component is zero, generating an oscillation through all colors. The rainbow gratings in the figures oscillate in the same direction, but different phases. The initial phases are 0, $-\frac{\pi}{2}$, $-\pi$, and $\frac{\pi}{2}$ respectively.

Notice that the the rainbow gratings might oscillate with different spatial frequencies. Figures 3.3 and 3.4 oscillate one revolution in one cycle, whereas figure 3.5 shows a grating with different frequency.

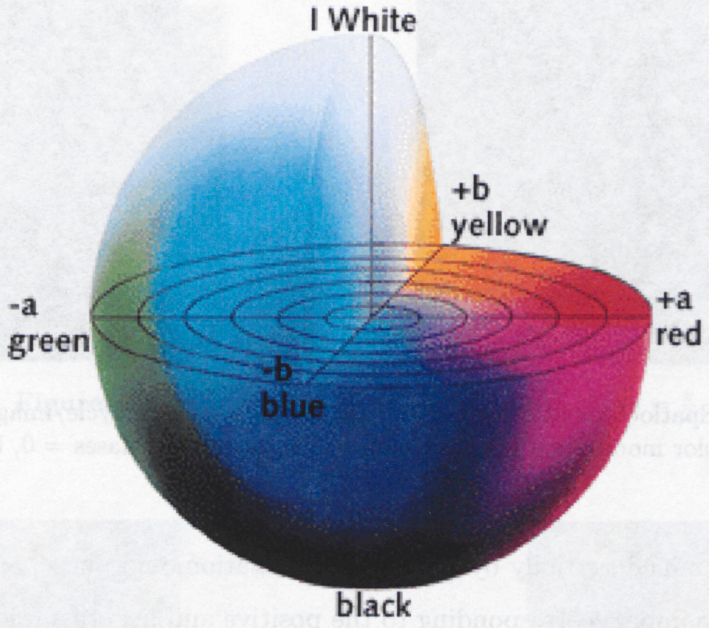


Figure 3.2: La^*b^* color space, www.sapdesignguild.org

One interesting thing about McCabe's theory is that it allows us to generate spatial frequencies with modulations operating in either the red-green or the yellow-blue color opponency paths, if $A(P,Q)$, $B(P,Q)$, $A(-P,-Q)$,$B(-P,-Q)$ satisfy the following conditions:

- For variation along red-green, $A(P,Q) = k$, $B(P,Q) = 0$, $A(-P,-Q) = k$,$B(-P,-Q) = 0$, such that the resultant path only varies along the real axis
- For variation along yellow-blue, $A(P,Q) = 0$, $B(P,Q) = k$, $A(-P,-Q) = 0$,$B(-P,-Q) = k$, which makes the resultant path varies along the imaginary axis.

The theory of the color opponency paths is useful for the proposed algorithms in Chapter 5. The reason is that by modifying the SCDFT coefficients in certain ways, a watermark with yellow-blue variation can be generated. Since human eyes are not sensitive to the colors, the imperceptibility can be improved. In conclusion, the chromatic contents of a color image can be perceived as superpositions of a set of spatial-frequency gratings, each of

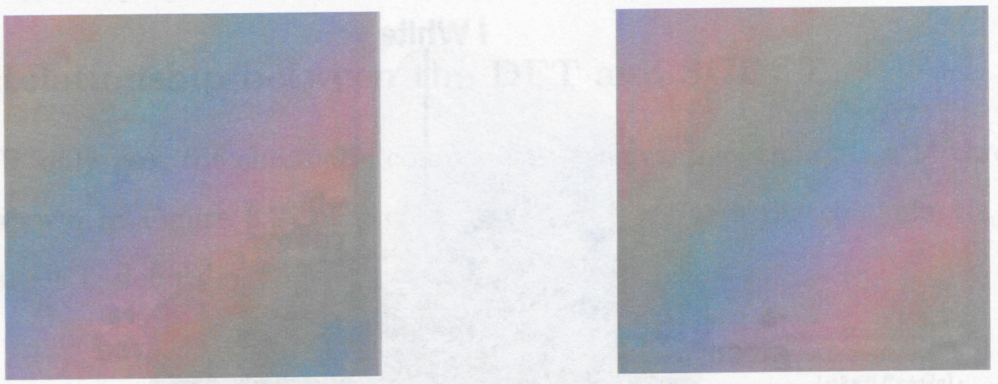


Figure 3.3: Spatiochromatic frequency gratings defined at 1 cycle/image in x and y directions and circular color modulation paths in color space. a) Initial phases = 0, b) Initial phase = $-\frac{\pi}{2}$

which is determined spatially by frequency, orientation, and phase; and chromaticity by the two complex numbers corresponding to the positive and negative frequency values.

3.2 Applications of the SCIP

The theory of SCIP can extend the current techniques for processing greyscale images to color images in an intrinsic manner. This section introduces how to apply the theory to *chromatic image filtering.*

3.2.1 Spatio-Chromatic Image Filtering

Image filtering in the frequency domain is an important tool to manipulate images. The reason for doing so is because it is computationally faster to perform two-dimensional Fourier transforms of the images, and multiply with a filter than to perform a convolution in the image (spatial) domain. For greyscale images, many good low-pass, high pass, and bandpass filters have been carefully designed, and their effects to images have also been well examined. However, color filters in the frequency domain are still at the early stage of development. Using the theory of SCIP can help to solve this problem. There are three type of filters for consideration:

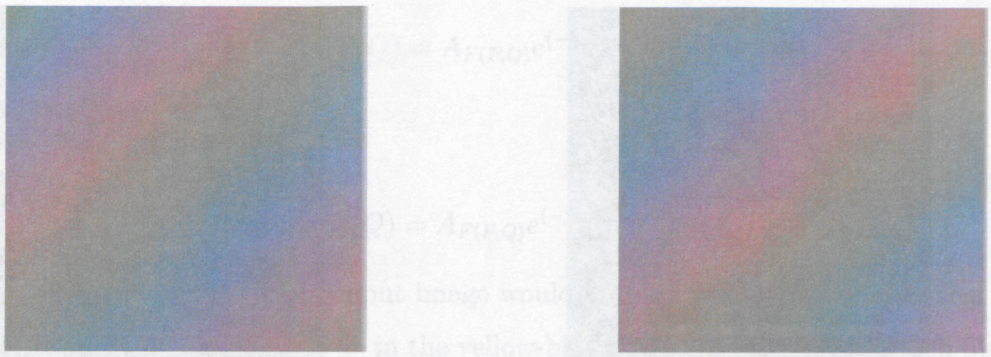


Figure 3.4: c) Initial phases $= -\pi$, d) Initial phase $= \frac{\pi}{2}$

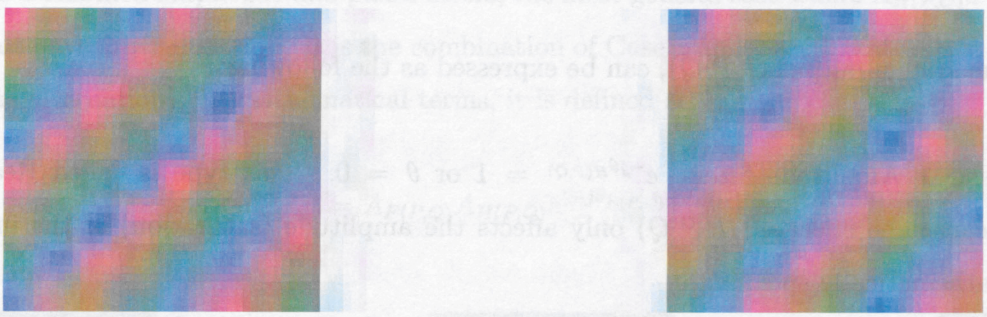


Figure 3.5: Spatiochromatic frequency gratings defined at 2 cycles/image in x and y directions and circular color modulation paths in color space. a) Initial phases $= 0$, b) Initial phase $= -\frac{\pi}{2}$

- Amplitude filters,
- Phase filters,
- Combined amplitude and phase filters.

Consider a system with an input image

$$F(P,Q) = A_{F(P,Q)}e^{-j\theta_{F(P,Q)}}, \tag{3.11}$$

with an impulse response

$$H(P,Q) = A_{H(P,Q)}e^{-j\theta_{H(P,Q)}}, \tag{3.12}$$

where (P,Q) is a spatial frequency, A is an amplitude quantity, and θ represents the phase.

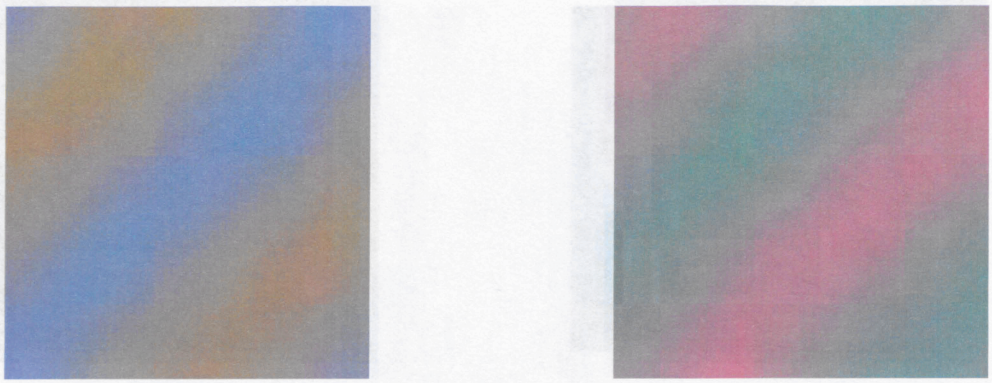


Figure 3.6: Identical spectral frequencies with linear paths in color space [see Fig 3.1b] differing only in starting (phase) values.

The system output $G(P,Q)$, can be expressed as the following:

Case 1 Amplitude filters, $e^{-j\theta_{H(P,Q)}} = 1$ or $\theta = 0$. This type is called Real Filtering because the effect of $H(P,Q)$ only affects the amplitude (saturation) of the input image. Mathematically speaking,

$$G(P,Q) = A_{F(P,Q)}e^{-j\theta_{F(P,Q)}}A_{H(P,Q)}e^{-j\theta_{H(P,Q)}} =$$

$$A_{F(P,Q)}A_{H(P,Q)}e^{-j\theta_{F(P,Q)+0}} = A_{F(P,Q)}A_{H(P,Q)}e^{-j\theta_{F(P,Q)}}, \quad (3.13)$$

therefore, the hue of the output $G(P,Q)$ is the same as $F(P,Q)$, and the saturation of the image is amplified (saturation scaling), as shown in Figure (3.7).

Case 2 Phase filters, $A_{H(P,Q)} = 1$ and $e^{-j\theta_{H(P,Q)}}$ only has the imaginary component. In that case, only $\theta = \frac{\pi}{2}$ or $\frac{3\pi}{2}$ would satisfy the requirement, because $e^{-j\frac{\pi}{2}} = -j$ and $e^{-j\frac{3\pi}{2}} = j$. θ with other values would generate real components. This type is called Imaginary Filtering because the filter introduces the imaginary factor, j, to the output. The output can be formulated as:

$$G(P,Q) = A_{F(P,Q)} e^{(-j\theta_{F(P,Q)} + \frac{\pi}{2})}, \tag{3.14}$$

or

$$G(P,Q) = A_{F(P,Q)} e^{(-j\theta_{F(P,Q)} + \frac{3\pi}{2})}. \tag{3.15}$$

As a result, the phase of the input image would be rotated by $\frac{\pi}{2}$ or $\frac{3\pi}{2}$. In other words, the red-green plane is rotated to lie in the yellow-blue plane, and the yellow-green plane lies in the red-green plane (hue alternation), as shown in Figure (3.8).

Case 3 Combined amplitude and phase filters, the most general case where $A_{H(P,Q)} e^{-j\theta_{H(P,Q)}}$ does not have any restriction. It is the combination of Case 1 and Case 2 (saturation scaling and hue alternation). In mathematical terms, it is defined as:

$$G(P,Q) = A_{F(P,Q)} A_{H(P,Q)} e^{(-j\theta_{F(P,Q)} + \theta_{H(P,Q)})}. \tag{3.16}$$

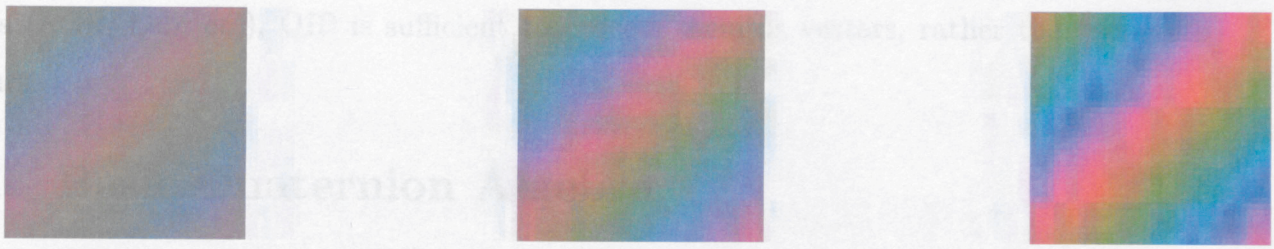


Figure 3.7: Saturation Scaling: a) Original image, b) Saturated image with scaling factor 2, c) Saturated image with scaling factor 3

3.3 Chapter Summary

The theory of the SCDFT is significant because it is the first comprehensive method to provide a framework to describe how the chromatic information of color images vary in the

Figure 3.8: Hue alternation: a) Original image, b) Image with hue alternation

frequency domain. It has many applications such as color filtering, and experiments have been shown that it provides satisfactory results. In Chapter 5 of the thesis, this theory will be applied to color image watermarking. For the first time, the characteristics of the HVS to colors will be considered holistically in the SCDFT domain in order to improve the robustness of the embedded watermark.

Chapter 4

Quaternion Image Processing (QIP)

QUATERNION Image Processing (QIP), proposed by Sangwine and Ell [34], is a new framework to exploit further information inherited among the channels of color images. Extensive research on QIP started in 1996, and the objective is to generalize different techniques from signal and image processing, such as frequency domain filtering, correlation, and compression, to handle color images as vector images [35]. Compared to Spatio-Chromatic Image Processing, which can only handle data with two components, QIP can process data with up to four components. Since most color images only have three components (RGB, Lab, etc), QIP is sufficient to process them as vectors, rather than separate scalars.

4.1 Basic Quaternion Algebra

Before discussing QIP in detail, it is important to understand the basis of quaternion algebra. A quaternion number has one real and three orthogonal imaginary components. It is usually expressed in the form,

$$Q = a + ib + jc + kd, \tag{4.1}$$

where a, b, c, and d are scalar coefficients, and i, j, and k are imaginary operators that have the following characteristics:

$$i^2 = j^2 = k^2 = ijk = -1;$$

$$ij = k,\ jk = i,\ ki = j;$$
$$ji = -k,\ kj = -i,\ ik = -j.$$

Given a quaternion Q, its conjugate is $Q = a - ib - jc - kd$, and its magnitude is $\sqrt{a^2 + b^2 + c^2 + d^2}$. It is sometimes useful to express Q into two parts: vector and scalar,

$$Q = S(Q) + \mathbf{V}(Q), \tag{4.2}$$

where $S(Q) = a$ and $\mathbf{V}(Q) = ib + jc + kd$. Note that multiplication of quaternions is not commutative, that is, $Q1 \times Q2 \neq Q2 \times Q1$.

Parallel and Perpendicular Decomposition: Given a pure quaternion u, and a second pure unit quaternion v, u may be decomposed into components parallel and perpendicular to v as follows:

$$u_{\perp} = \frac{1}{2}(u + vuv), \tag{4.3}$$

where $u_{\perp}$ is perpendicular to v, and

$$u_{\parallel} = \frac{1}{2}(u - vuv), \tag{4.4}$$

where $u_{\parallel}$ is parallel to v.

4.2 Applications of the QIP

Quaternion image processing usually encodes the red, green, and blue components of a color image as follows:

$$F(m,n) = 0 + R(m,n)i + G(m,n)j + B(m,n)k, \tag{4.5}$$

where $R(m,n)$, $G(m,n)$, and $B(m,n)$ are the three color components in the RGB color space. The following subsections explain two common applications using the quaternion representation: *Color Image Edge Detection*, and *Color Image Watermarking*.

4.2.1 Color Image Edge Detection

Edge detection is a very common approach for detecting meaningful discontinuities in grey level images [35]. Many approaches implement the first-and second derivatives for detecting edges in an image. The idea is to convolve the image with a filter or mask, such that the edges can be extracted. Common filters are Roberts, Prewitt, and Sobel [35]. However, techniques for detecting chromatic edges using the quaternion algebra have not been matured yet. Thus, researchers proposed a new quaternion convolution, so the filtering techniques for grey-scale images can be applied to quaternion images. The convolution is defined as:

$$F'(m,n) = \sum_{x=-X}^{X} \sum_{y=-Y}^{Y} h_L(x,y) F(m - XmodM, n - ymodN) h_R(x,y), \tag{4.6}$$

where $F'(m,n)$ is the filtered image, $F(m,n)$ is the original image, $h_L(x,y)$ and $h_R(x,y)$ are left and right quaternion masks. The reason for having two masks is because quaternion multiplication is not commutative and different filters would result from application of a given mask on the left, or on the right, of the image pixels. Thus, the general case of a quaternion-based filter requires both masks. For more details regarding quaternion edge detection, please refer to [35].

4.2.2 Color Image Watermarking

The discrete Quaternion Fourier Transform (QFT) was first presented by Sangwine and Ell. There are at least two types of QFT [36] [37] [38].

Type 1:

$$F(u,v) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} e^{-i2\pi(\frac{mu}{M})} F(m,n) e^{-j2\pi(\frac{nv}{N})}, \tag{4.7}$$

and the inverse QFT is formulated as:

$$F(m,n) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} e^{i2\pi(\frac{mu}{M})} F(u,v) e^{j2\pi(\frac{nv}{N})}, \tag{4.8}$$

where M and N are the dimensions of the image.

Type 2:

$$F(u,v) = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} e^{-\mu 2\pi(\frac{mu}{M}+\frac{nv}{N})} F(m,n), \tag{4.9}$$

and the inverse QFT is:

$$F(m,n) = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} e^{\mu 2\pi(\frac{mu}{M}+\frac{nv}{N})} F(u,v),. \tag{4.10}$$

The difference between them is that Type 2 depends on the definition of μ. To be valid, μ has to be a unit pure quaternion (i.e. $\mu^2 = -1$ and the real part of $\mu = 0$).

Bas and Bihan [39] proposed a quaternion image watermarking algorithm using **Type 2**. Their approach was to embed a watermark into $F(u,v)_{||}$. The achieve this, the value of μ has to be found such that the perceptual impact after embedding is minimal. First of all, empirical tests were performed to find values of μ that lead to minor visible distortion. The tests were performed by adding gratings on 8×8 of natural images and evaluating the distortion on CRT and LCD monitors. The value of μ offering the smallest distortions was kept, which was called μ_{perc}. For example, $\mu_{perc} = (-2j + 8k)/\sqrt{68}$ gives the smallest distortion for the Lenna image. Secondly, the authors even improved the distortion by considering the mean color of the image block, defined as:

$$\mu_{im} = \frac{E_b[R]i + E_b[G]j + E_b[B]k}{\sqrt{E_b^2[R] + E_b^2[G] + E_b^2[B]}}, \tag{4.11}$$

where $E_b[X]$ represents the mean of the component X in a block b of the image. By considering both μ_{perc} and μ_{im}, the authors developed a better estimate $\mu_{perc'}$ [39], and used it for the QFT:

$$F(u,v) = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} e^{-\mu_{perc'} 2\pi(\frac{mu}{M}+\frac{nv}{N})} F(m,n). \tag{4.12}$$

and the watermark is inserted into $F(u,v)_{||}$ after the decomposition using eq. (4.4).

4.3 Chapter Summary

This chapter gives a detail description of how QIP works and its applications to color image edge detection, and watermarking. Since it can handle data up to four components, more applications on the vector image processing are expected to be invented and worth researching on.

Chapter 5

New Color Image Watermarking using Multidimensional Fourier Transforms

IN recent years, watermark researchers have realized the importance of the chromatic information of color images, and extensive research has been conducted. For examples, Reed [40] *et. al.* proposed a system that takes advantage of the low sensitivity of the HVS to high frequency changes along the yellow-blue axis, to place most of the watermark in the yellow component of the image. Bas *et. al.* proposed a digital color image watermarking scheme using hypercomplex numbers representation and the Quaternion Fourier Transform [39]. The following sections explain two new color image watermarking schemes based on the theory about the SCDFT introduced in Chapter 3.

5.1 Motivation and Inspiration

Color is an important content in color images. Most digital images have three components such as the red, green, and blue components in the RGB color space or the Y, u, v components in the Yuv domain. The luminance/chromaince models, such as the Yuv, are more frequently used in digital image processing, because they model human perception of color more closely than the RGB model. However, as mentioned earlier, most existing watermarking algorithms for color images simply mark the luminance component only, ignoring others. Even though it contains the most significant information, nevertheless, the chrominance components should also be considered as well. The primary question is how we can take advantage of the

color contents in order to increase the robustness, or decrease the visible distortion. The proposed idea was inspired by the fact that *colors can be added or subtracted to produce other colors.* There are three primary colors (red, green, blue), and different colors can be obtained by mixing the primary colors with the proper intensity. One characteristics of the HVS is that it is more sensitive to certain colors such as red, and less sensitive to colors such as blue [16]. If the watermark is embedded in the less sensitive area, then the strength of it can be increased and still be undetected by human eyes. However, this creates a problem. Many compression techniques, such as JPEG, eliminate the insensitive chromatic components aggressively, which might easily destroy the embedded watermark. On the other hand, if the watermark is embedded in the highly sensitive area, then the strength of it must be low in order to be invisible. The fundamental good is to try and achieve both. That is, to embed in the highly sensitive area but keep the visible distortion negligible. Even though robustness and imperceptibility are conflicting requirements, how could a better tradeoff be achieved?

Intuitively, if a watermark is embedded in the low-frequency SCDFT coefficient $F(u, v)$, using the theory introduced in Chapter 3, the effect is the same as adding a rainbow grating to the image, as demonstrated in Figure 5.1.

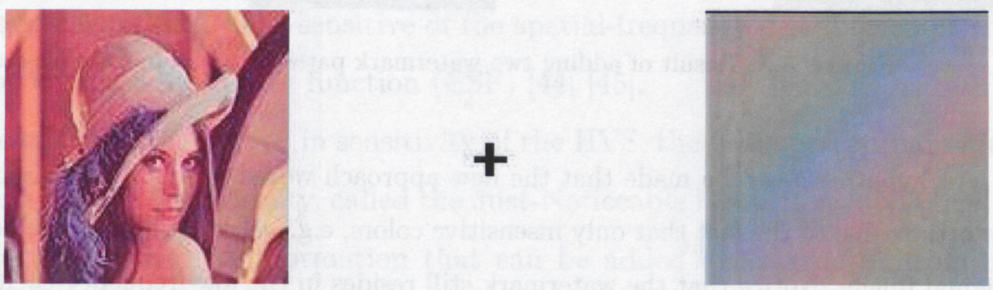


Figure 5.1: Adding a watermark pattern to an image. a) Original image, b) Watermark pattern.

Since the watermark pattern is a rainbow grating that contains sensitive colors like red and green, the watermark after embedding would be easily detected by human eyes. In order to decrease the visible distortion, the proposed algorithm adds one more step: adding

a **compensation pattern** in the negative frequency $F(-u, -v)$. The overall procedure is summarized in Figure 5.2. The main advantage of doing this is that the summation (addition) of the colors of the two patterns provides an interesting new pattern, that only has **two insensitive colors - yellow and blue**, as illustrated in Figure 5.3.

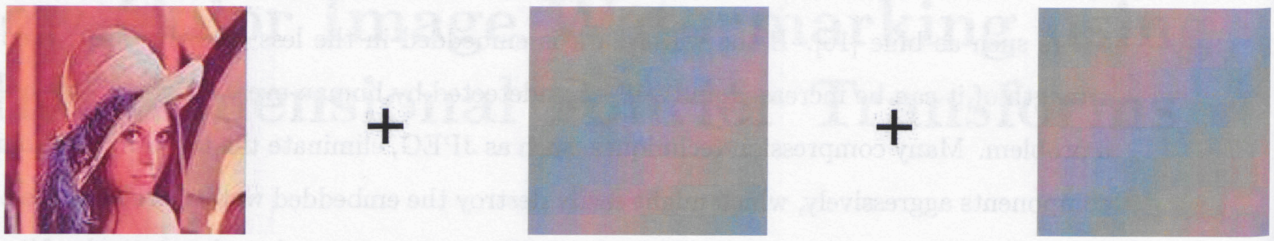


Figure 5.2: Adding two watermark patterns to an image. a) Original image, b) Watermark pattern, c) Compensation pattern

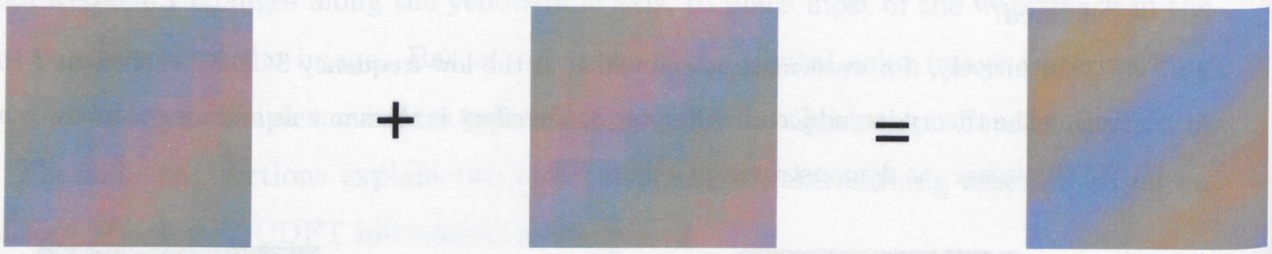


Figure 5.3: Result of adding two watermark patterns with opposite phases.

An hypothesis can be made that the new approach would produce a much lower visible distortion, due to the fact that only insensitive colors, e.g. yellow and blue, are added to the original image. Notice that the watermark still resides in the low-frequency coefficient. The purpose of adding the compensation mark in the negative coefficient is merely to decrease (i.e. compensate) the visible distortion generated by adding the watermark in the positive frequency coefficient.

From the decoder's perspective, both approaches are exactly the same. The decoder would only look at the positive coefficient for watermark extraction, ignoring the compen-

sation mark. In fact, the compensation mark would be most likely destroyed by external attacks. In other words, with everything remaining the same, the proposed approach generates lower distortion under the same watermark strength (i.e. higher imperceptibility with the same robustness). Having introduced the inspiration and logic, the technical details will now be given, including the discussion relevant for designing a comprehensive watermarking scheme, including the characteristics of the HVS, color space, and methodologies.

5.2 Human Visual System (HVS)

The HVS has been mentioned many times in the previous sections, but a clear description of it has not be given. In fact, this topic has been researched for many years [41][42], and a single mathematical model that incorporates all masking effects is still unavailable. Nevertheless, a simpler model that considers some common masking effects related to image processing have been proposed. The masking effects are:

Luminance Masking: It is a visual phenomenon that the HVS is less sensitive to brighter areas, and more sensitive to darker areas [41] [42].

Texture Masking: The stronger the texture in the background, the lower the visibility of the embedded signal [43] [41].

Frequency Masking: The sensitive of the spatial-frequency distribution of an image is modelled by contrast sensitivity function (CSF) [44] [45].

Because of the inconsistency in sensitivity of the HVS, there are perceptual redundancies inherited in an image. A quantity, called the Just-Noticeable Distortion (JND) [19] indicates the limit of the amount of information that can be added or subtracted from the image without sacrificing its visual quality. It is ideal for watermarking, because it gives an upper bound to the strength of the embedded watermark, such that the watermarked image would look the same as the original. To satisfy the invisibility requirement, the new algorithms proposed in this thesis use the following characteristics

- The eye is particularly sensitive to changes in image hue;
- The eye is less sensitive to yellow-blue component [16],

and the JND model proposed by Chou in 2003 [46].

5.3 Color Spaces

A color model is an abstract mathematical model describing the way colors can be represented as tuples of numbers, typically as three or four values or color components. The most commonly used model is called the RGB model. It is composed of three values (e.g. $[255, 255, 255]$), representing the intensity of the colors Red, Green, and Blue. When displaying on a computer monitor, it uses additive color mixing to produce the total color. However, this color space has a big disadvantage: adequate means for measuring color similarity is not available. In the context of watermarking, an objective metric to measure the perceptual distortion between the original image and watermarked image is definitely required. A common metric often used to measure the difference between two images is the MSE, as described in Eq. (2.3). When applying to RGB images, it can be extended to:

$$D_{mse,RGB}(I_o, I_w) = \frac{1}{N^2}\sum_{i}^{N}\sum_{j}^{N}(\frac{R(i,j)+G(i,j)+B(i,j)}{3})^2, \tag{5.1}$$

where $R(i,j)$, $G(i,j)$, $B(i,j)$ are the spatial difference of the corresponding components of the images. However, this metric is not good to achieve this goal. Thus, another color space, La^*b^*, was chosen for the color space used for the proposed algorithms.

CIE $L^*a^*b^*$ (CIELAB) is the most complete color model used conventionally to describe all the colors visible to the human eye [47]. It was developed for this specific purpose by the International Commission on Illumination (Commission Internationale d'Eclairage, hence its CIE initialism). The parameter L represents the lightness of the color (ranges from 0 to 100), whereas a^* and b^* are chromatic information. Figure 3.2 shows the distribution of colors in this color space. The equations to convert between RGB and La^*b^* can be formulated as:

$$L = 116 \times f(\frac{Y}{Y_n}) - 16 \tag{5.2}$$

$$a^* = 500 \times [f(\frac{X}{X_n}) - f(\frac{Y}{Y_n})] \tag{5.3}$$

$$b^* = 200 \times [f(\frac{Y}{Y_n}) - f(\frac{Z}{Z_n})] \tag{5.4}$$

and

$$f(t) = t^{\frac{1}{3}}, t > 0.008856 \tag{5.5}$$

$$f(t) = 7.777t + \frac{11}{116}, otherwise, \tag{5.6}$$

where (X, Y, Z) and (X_n, Y_n, Z_n) are CIE XYZ tristimulus values. Moreover, La^*b^* is also a uniform color space used to measure the color difference $\triangle E$ at each pixel, which is defined as:

$$\triangle E = \sqrt{\triangle L^2 + (\triangle a^*)^2 + (\triangle b^*)^2}. \tag{5.7}$$

Experiments have shown that $\triangle E < 1$ is not detectable by humans, and $\triangle E < 3$ is not apparent. In C.H. Chou's paper [46], $\triangle E = 3$ is called the *Uniform Just-Noticeable Color Difference (UJNCD)*, and he derived a JND model, which he called the *Non-Uniform Just-Noticeable Color Difference (NUJNCD)* after considering some masking effects. His model produces a value $JND(L, a^*, b^*)$ for every pixel of a color image in the La^*b^* color space. As long as the distortion is less than this value, human eyes would not be able to detect the difference. This thesis uses his model to bound the distortion generated by embedding the watermark. The next two sections introduce the algorithms using two multidimensional transforms - SCDFT and QFT.

5.4 SCDFT image watermarking

Given an image of size $N \times N$, the image is divided into blocks of size 8×8, followed by taking the SCDFT of each block. A pair of marks $w(k_1, k_2)$ and $w(-k_1, -k_2)$ are embedded

into the coefficients $f(k_1, k_2)$ and $f(-k_1, -k_2)$ in the following way:

$$f'(k_1, k_2) = f(k_1, k_2) \pm \alpha w(k_1, k_2), \tag{5.8}$$

$$f'(-k_1, -k_2) = f(-k_1, -k_2) \pm \alpha w(-k_1, -k_2), \tag{5.9}$$

where α is the watermark strength to be maximized. To apply the two characteristics of the HVS as mentioned in Section 5.2, some constraints are imposed to the values $w(k_1, k_2)$ and $w(-k_1, -k_2)$. Let

$$f(k_1, k_2) = f_{rp} + jf_{ip}, \tag{5.10}$$

$$w(k_1, k_2) = w_{rp} + jw_{ip}, \tag{5.11}$$

$$w(-k_1, -k_2) = w_{rn} + jw_{in}, \tag{5.12}$$

where the subscript $r \equiv real$, $i \equiv imaginary$, $p \equiv positive$, $n \equiv negative$. From Section 5.2, human eyes are particularly sensitive to changes in image hue, and less sensitive to yellow-blue component. Therefore, the watermark should be designed to satisfy the following conditions:

1. Keep the hue (angle) unchanged:

Similar to most watermarking schemes, the watermark $w(k_1, k_2)$ is adjusted by a scaling factor α, before being added to the coefficient $f(k_1, k_2)$. The equation is formulated as:

$$f'(k_1, k_2) = f_{rp} + jf_{ip} \pm \alpha(w_{rp} + jw_{ip}), \tag{5.13}$$

where $f'(k_1, k_2)$ is the new coefficient. In general, $f(k_1, k_2)$ is a complex quantity that can be represented by its magnitude and angle. As described in Chapter 3, the angle of a SCDFT coefficient is interpreted as the hue of a color corresponding to the chromaticity diagram in Figure (3.2). Our goal is to make the angle of $f'(k_1, k_2)$ be the same as $f(k_1, k_2)$, which is $tan^{-1}(\frac{f_{ip} \pm \alpha w_{ip}}{f_{rp} \pm \alpha w_{rp}}) = tan^{-1}(\frac{f_{ip}}{f_{rp}})$. To achieve this requirement, if $w_{rp} = p \times f_{rp}$ and $w_{ip} = p \times f_{ip}$, where p is a constant, then

$$tan^{-1}(\frac{f_{ip} \pm \alpha w_{ip}}{f_{rp} \pm \alpha w_{rp}}) = tan^{-1}(\frac{f_{ip}(1 \pm \alpha p)}{f_{rp}(1 \pm \alpha p)}) = tan^{-1}(\frac{f_{ip}}{f_{rp}}). \tag{5.14}$$

Thus, the requirement to keep the hue unchanged is fulfilled if $p = 1$, and we get

$$w_{rp} = f_{rp}, w_{ip} = f_{ip}. \tag{5.15}$$

2. Make the overall effect be yellow-blue:

According to McCabe's theory described in Chapter 3, if the summation of the real parts of the positive and negative frequency coefficients is zero, then the resultant path would vary along the yellow-blue axis of the chromaticity diagram. Setting

$$w_{rn} = -w_{rp}, w_{in} = w_{ip} \tag{5.16}$$

would satisfy this condition because $w_{rn} = -w_{rp}$ cancels out the real part of the watermarks, and only the imaginary part remains. Substituting Eq. (5.16) and Eq. (5.15) into Eq. (5.8) and (5.9) gives the result:

$$f'(k_1, k_2) = f(k_1, k_2) \pm \alpha(f_{rp} + jf_{ip}), \tag{5.17}$$

$$f'(-k_1, -k_2) = f(-k_1, -k_2) \pm \alpha(-f_{rp} + jf_{ip}). \tag{5.18}$$

To ensure invisibility of the watermarks, a distortion metric $D(n, m, k_1, k_2)$ is defined to measure the difference in the spatial domain generated by modifying the coefficients $f(k_1, k_2)$ and $f(-k_1, -k_2)$ in the frequency domain, formulated as:

$$\begin{aligned} D(n, m, k_1, k_2) = \frac{1}{N^2}((f(k_1, k_2) - f'(k_1, k_2))e^{(\frac{2\pi jk_1n}{N})} \\ e^{(\frac{2\pi jk_2m}{N})} + (f(-k_1, -k_2) - f'(-k_1, -k_2))e^{(\frac{-2\pi jk_1n}{N})}e^{(\frac{-2\pi jk_2m}{N})}). \end{aligned} \tag{5.19}$$

Substituting Eq. (5.17) and (5.18) into Eq. (5.19) gives

$$\begin{aligned} D(n, m, k_1, k_2) = \frac{1}{N^2}(-\alpha(f_{rp} + jf_{ip})e^{(\frac{2\pi jk_1n}{N})}e^{(\frac{2\pi jk_2m}{N})} \\ - \alpha(-f_{rp} + jf_{ip})e^{(\frac{-2\pi jk_1n}{N})}e^{(\frac{-2\pi jk_2m}{N})}). \end{aligned} \tag{5.20}$$

Let $\theta_1 = \frac{2\pi k_1 n}{N}$, and $\theta_2 = \frac{2\pi k_2 m}{N}$, noting that $e^x = \cos x + j \sin x$ and with some algebraic manipulations, Eq. (5.20) can be simplified to:

$$D(n,m,k_1,k_2) = \frac{1}{N^2}(-j\alpha(2f_{rp}(\sin\theta_2\cos\theta_1 + \sin\theta_1\cos\theta_2) + 2f_{ip}(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2))). \quad (5.21)$$

To maximize the scaling factor α before being detected by human eyes, the parameter $NUJNCD$ as a upper bound of the $D(n,m,k_1,k_2)$ is used, that is, $D(n,m,k_1,k_2) < NUJNCD$ for every spatial coordinates n and m. Taking the square of both sides gives:

$$\|D(n,m,k_1,k_2)\|^2 = \frac{\alpha^2}{N^4}(2f_{rp}(\sin\theta_2\cos\theta_1 + \sin\theta_1\cos\theta_2) + 2f_{ip}(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2))^2 \leq (NUJNCD)^2. \quad (5.22)$$

Depending on the perceptual requirement of the applications, two methods are available to determine α.

- Make the **average** distortions of a block less than $NUJNCD$ (loose condition).
- Make the distortion of **every** pixel less than $NUJNCD$ (tight condition).

Approach 1: Calculate α such that the average distortions of a block are less than $NUJNCD$

$$\|D(k_1,k_2)\|^2 = \alpha^2 \sum_{n=0}^{N-1}\sum_{m=0}^{N-1} \frac{1}{N^4}(2f_{rp}(\sin\theta_2\cos\theta_1 + \sin\theta_1\cos\theta_2) + 2f_{ip}(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2))^2 = N^2(NUJNCD)^2. \quad (5.23)$$

Let $\mathrm{x} = \cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2$ and $\mathrm{y} = \sin\theta_2\cos\theta_1 + \sin\theta_1\cos\theta_2$,

$$\alpha = \sqrt{\frac{N^6(NUJNCD)^2}{\sum_{n=0}^{N-1}\sum_{m=0}^{N-1}(2f_{rp}y + 2f_{ip}x)^2}}. \quad (5.24)$$

Approach 2: Recursively decrease α until the distortions of all pixels are smaller than $NUJNCD$. An initial α is assigned, and then $\|D(n, m, k_1, k_2)\|^2$ is calculated for every pixel of the block. If the distortion of some pixels exceed $(NUJNCD)^2$, α is decreased until $\|D(n, m, k_1, k_2)\|^2 < (NUJNCD)^2 \; \forall \; n, m$.

In general, approach 1 should be used when the robustness of the watermarks is critical while the invisibility constraint can be relaxed. On the other hand, approach 2 is good for cases where the invisibility constraint is more important.

To extract the embedded watermarks, the original image, and the watermarked (possibly corrupted) image are needed. $f(k_1, k_2) + \alpha(f_{rp} + jf_{ip})$ and $f(k_1, k_2) - \alpha(f_{rp} + jf_{ip})$ are calculated from the original image, followed by a distance comparison to the coefficient extracted from the watermarked image. If the extracted coefficient is closer to $f(k_1, k_2) + \alpha(f_{rp} + jf_{ip})$, we assume bit 1 was embedded, otherwise, bit 0 was embedded. The flow diagram is illustrated in Figure (5.4).

5.5 QFT Image Watermarking

The SCDFT scheme is attractive because it utilizes chromatic information to improve the imperceptibility of the watermarked image, but it ignores the luminance component. Attacks such as JPEG compression and color to greyscale conversion would destroy the watermark. To solve the problems, a new color watermarking algorithm using the QFT is introduced.

McCabe provided an interpretation of the SCDFT coefficients, which was extended to image watermarking. Unfortunately, a similar interpretation to the QFT coefficients is not available. The first question that comes to mind is: How does embedding watermarks in the QFT coefficients affect the image $F(m, n)$ in the spatial domain? Before answering this question, a general framework is needed to interpret the quaternion spectral points of color images. To achieve this goal, McCabe's interpretation will be extended to the QFT domain.

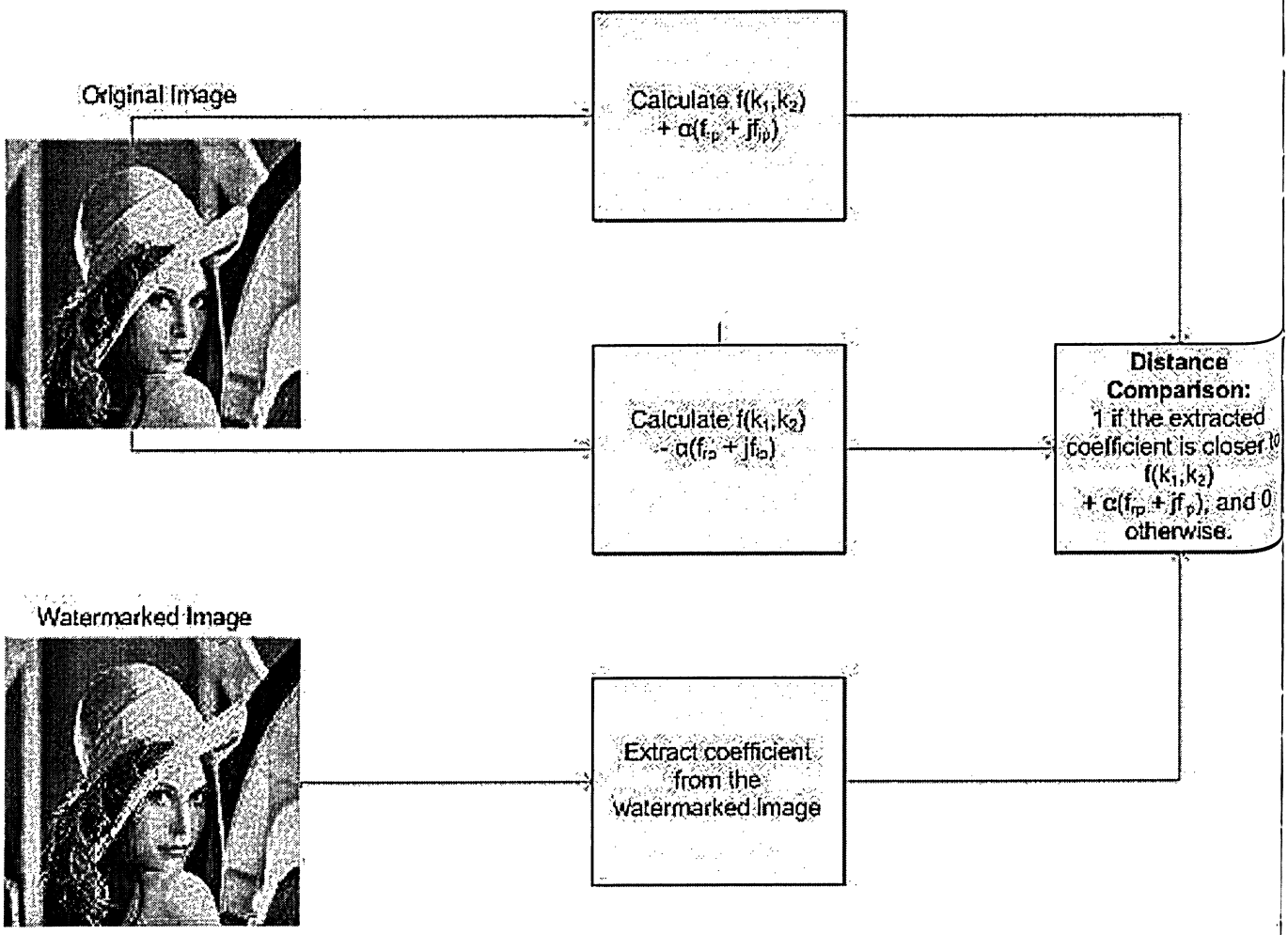


Figure 5.4: Flow diagram showing the extraction procedure using the SCDFT scheme.

5.5.1 Spectrum Interpretation

Having introduced the concept of phasors in Chapter 3, let's expand Eq. (4.8). Let $\theta_1 = 2\pi(\frac{mu}{M})$, $\theta_2 = 2\pi(\frac{nv}{N})$, $F(u,v) = a_1 + b_1 i + c_1 j + d_1 k$ and $F(-u,-v) = a_2 + b_2 i + c_2 j + d_2 k$,

$$e^{i\theta_1}F(u,v)e^{j\theta_2} =$$

$$a_1\cos\theta_1\cos\theta_2 - c_1\cos\theta_1\sin\theta_2 - b_1\sin\theta_1\cos\theta_2 +$$
$$d_1\sin\theta_1\sin\theta_2 + i(b_1\cos\theta_1\cos\theta_2 - d_1\cos\theta_1\sin\theta_2 +$$
$$a_1\sin\theta_1\cos\theta_2 - c_1\sin\theta_1\sin\theta_2) + j(a_1\cos\theta_1\sin\theta_2 +$$
$$c_1\cos\theta_1\cos\theta_2 - b_1\sin\theta_1\sin\theta_2 - d_1\sin\theta_1\cos\theta_2) +$$
$$k(b_1\cos\theta_1\sin\theta_2 + d_1\cos\theta_1\cos\theta_2 + a_1\sin\theta_1\sin\theta_2 +$$
$$c_1\sin\theta_1\cos\theta_2), \quad (5.25)$$

$$e^{-i\theta_1}F(-u,-v)e^{-j\theta_2} =$$

$$a_2\cos\theta_1\cos\theta_2 + c_2\cos\theta_1\sin\theta_2 + b_2\sin\theta_1\cos\theta_2 +$$
$$d_2\sin\theta_1\sin\theta_2 + i(b_2\cos\theta_1\cos\theta_2 + d_2\cos\theta_1\sin\theta_2 -$$
$$a_2\sin\theta_1\cos\theta_2 - c_2\sin\theta_1\sin\theta_2) + j(-a_2\cos\theta_1\sin\theta_2 +$$
$$c_2\cos\theta_1\cos\theta_2 - b_2\sin\theta_1\sin\theta_2 + d_2\sin\theta_1\cos\theta_2) +$$
$$k(-b_2\cos\theta_1\sin\theta_2 + d_2\cos\theta_1\cos\theta_2 + a_2\sin\theta_1\sin\theta_2 -$$
$$c_2\sin\theta_1\cos\theta_2), \quad (5.26)$$

Considerable simplifications can be made to Eq. (5.25) and Eq. (5.26), and an interpretation of the QFT coefficients can be drawn if the parameters a_1, a_2, b_1, b_2 ..... are assigned in certain ways. There are four cases of interest. Let $F(\pm u, \pm v)$ denote $F(u,v) = F(-u,-v)$,

Case 1: $F(\pm u, \pm v) = dk$, then the summation of (5.25), and (5.26) gives

$$2d\sin\theta_1\sin\theta_2 + k2d\cos\theta_1\cos\theta_2, \quad (5.27)$$

(5.27) shows that if $F(\pm u, \pm v) = dk$, then the combined effect of the positive and negative coefficients is the same as a cosine function with magnitude of $2d$ varying along the k

component (the real component of the image is ignored).

Case 2: $F(\pm u, \pm v) = bi$, then the summation of Eq. (5.25), and Eq. (5.26) gives

$$i2b\cos\theta_1\cos\theta_2 - j2b\sin\theta_1\sin\theta_2, \tag{5.28}$$

(5.28) shows that if $F(\pm u, \pm v) = bi$, then the combined effects of the positive and negative coefficients is the same as a circle with diameter $2b$, starting at angle 0, rotating in the negative direction on a complex plane.

Case 3: $F(\pm u, \pm v) = cj$, then the summation of (5.25), and (5.26) gives

$$-i2c\sin\theta_1\sin\theta_2 + j2c\cos\theta_1\cos\theta_2 \tag{5.29}$$

(5.29) shows that if $F(\pm u, \pm v) = cj$, then the combined effects of the positive and negative coefficients is the same as a circle with diameter $2c$, starting at angle $\pi/2$, rotating in the negative direction on a complex plane.

Case 4: $F(\pm u, \pm v) = bi + cj + dk$ can be interpreted as the combinations of case 1, 2 and 3.

If we encode a color image as $F(m,n) = a^*(m,n)i + b^*(m,n)j + L^*(m,n)k$, using the interpretation from above, some conclusions can be made:

Case 1: $F(\pm u, \pm v) = dk$ can be interpreted as adding a cosine function in the luminance component of the image, as shown in Figure (5.5).

Case 2 and 3: $F(\pm u, \pm v) = bi$ or $F(\pm u, \pm v) = cj$ enables us to use the interpretations provided by McCabe, as in the SCDFT domain, to add a function, in the form of a rainbow grating in the chromatic components of the image, as shown in Figure (5.6) and Figure (5.7).

Case 4: $F(\pm u, \pm v) = bi + cj + dk$ can be interpreted as a cosine grating embedded in the luminance component and two rainbow gratings rotating with the same direction and different initial angles in the chrominance components.

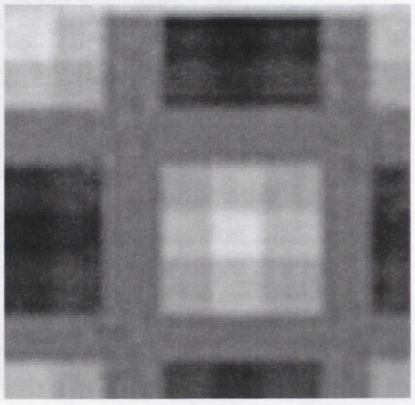


Figure 5.5: Adding a luminance pattern to an image. a) Original image b) $W(\pm u, \pm v) = L^*k$.

5.5.2 New Adaptive Color Watermarking Algorithm

The interpretation of the QFT coefficients is very important to predict the possible change made to the image after watermarking. In the context of digital image watermarking, the goal is to design a watermark $W(u, v)$, with the scaling factor, α as big as possible, such that the visible distortion to the image is perceptually minimal. Mathematically speaking, the watermarked image $WF(m, n)$ can be formulated as:

$$WF(m,n) = IQFT(F(u,v) + \alpha W(u,v)) = IQFT(F(u,v)) + IQFT(\alpha W(u,v))$$
$$= F(m,n) + W(m,n). \quad (5.30)$$

Let $W(\pm u, \pm v) = b_3 i + c_3 j + d_3 k$. b_3, c_3, and d_3 can be designed according to the guidelines given in the previous section. The component b_3 would add the watermark in the luminance component, where c_3 and d_3 are responsible for the chromatic components. In order to make the overall effect varies along the yellow-blue axis, as we did in the SCDFT scheme, some constraints must be imposed to the values b_3, c_3, and d_3. From Eq. (5.28) and Eq. (5.29), if $b_3 = c_3$, the summation of Eq. (5.28) and Eq. (5.29) becomes

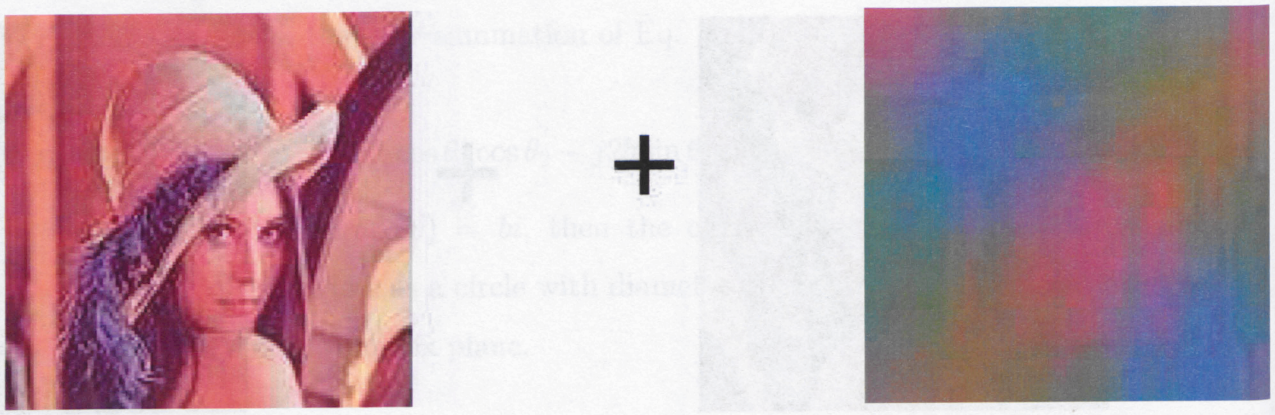


Figure 5.6: Adding a color pattern to an image. a) Original image b) $W(\pm u, \pm v) = a^* i$.

$$i2b_3(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) + j2b_3(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2)$$
$$= 2b_3(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2)(i+j), \quad (5.31)$$

as shown in Figure (5.8), and the overall effect would vary along the 45 degree axis. It means the color of the overall effect would be red and blue. To satisfy the requirement of varying along the yellow-blue axis, we have to transform the a^*b^* by 45 degrees before taking the QFT of the image $F(m,n)$, such that it would align with the $+b$ and $-b$ axis and oscillate between yellow and blue.

Since the sensitivity of human eyes to the yellow-blue component is approximately $\frac{1}{5}$ compared to the luminance component [16]; set $d_3 = \frac{1}{5}b_3$. As a result,

$$W(\pm u, \pm v) = b_3 i + b_3 j + \frac{1}{5}b_3 k. \quad (5.32)$$

To ensure invisibility of the watermark, a distortion metric $D(n,m,u,v)$ is defined to measure the difference in the spatial domain generated by adding the coefficients $W(u,v)$ and

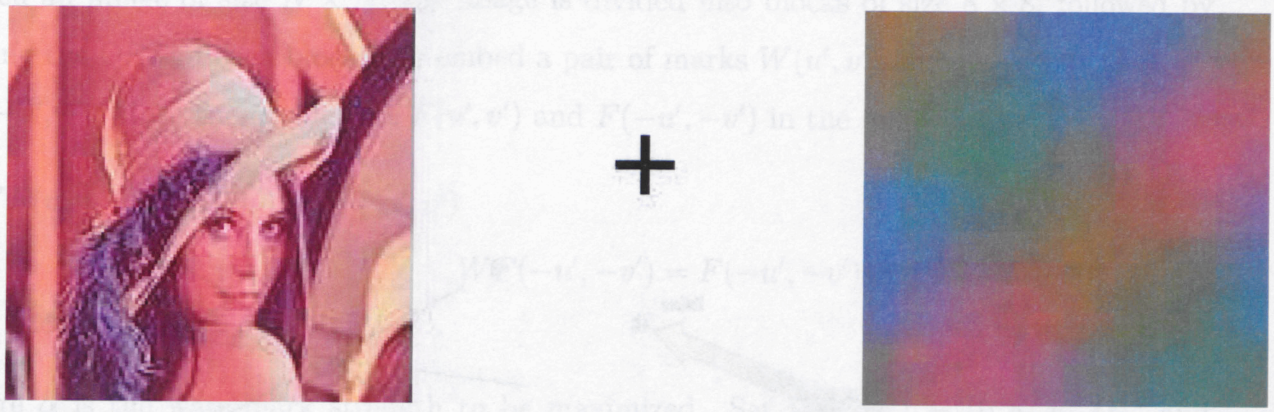


Figure 5.7: Adding a color pattern to an image. a) Original image b) $W(\pm u, \pm v) = b^* j$.

$W(-u, -v)$ in the frequency domain, defined as:

$$D(n,m,u,v) = \frac{\alpha}{N^2}(i2b_3(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) + j2b_3\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2 + k\frac{2}{5}b_3\cos\theta_1\cos\theta_2) \leq NUJNCD. \quad (5.33)$$

Again, watermark invisibility is achieved if the distortion of the watermarked image is smaller than the quantity $NUJNCD$, $\|D(n,m,u,v)\|^2 \leq (NUJNCD)^2\ \forall n,m$, that is,

$$\|D(n,m,u,v)\|^2 = \frac{\alpha^2}{N^4}(8(b_3(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2))^2 + \frac{4}{25}(b_3\cos\theta_1\cos\theta_2)^2) \leq (NUJNCD)^2. \quad (5.34)$$

The parameter α is an important factor, because the larger it is, the more robust the watermarked image will be. Depending on the perceptual requirement of the applications, two options are available to maximize α. One is to make the **average** distortions of a block less than $NUJNCD$ (loose condition), and the other one is to make the distortion of **every** pixel less than $NUJNCD$ (tight condition):

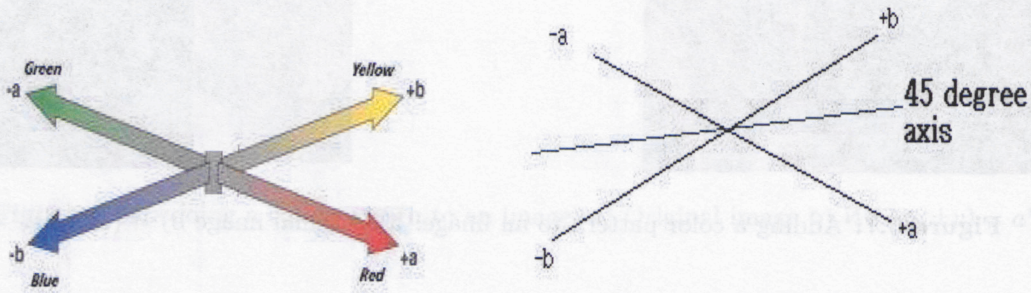


Figure 5.8: a) Lab chromaticity diagram, b) $W(\pm u, \pm v) = b_3 i + b_3 j$.

Approach 1: Calculate α such that the average distortion of a block is less than $NUJNCD$

$$\|D(u,v)\|^2 = \frac{\alpha^2}{N^4} \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} 8(b_3(\cos\theta_1 \cos\theta_2 - \sin\theta_1 \sin\theta_2))^2 + \frac{4}{25}(b_3 \cos\theta_1 \cos\theta_2)^2 = N^2(NUJNCD)^2. \quad (5.35)$$

Let x $= \cos\theta_1 \cos\theta_2 - \sin\theta_1 \sin\theta_2$ and y $= \cos\theta_1 \cos\theta_2$,

$$\alpha = \sqrt{\frac{N^6(NUJNCD)^2}{\sum_{n=0}^{N-1} \sum_{m=0}^{N-1} 8(b_3 x)^2 + \frac{4}{25}(b_3 y)^2}}. \quad (5.36)$$

Approach 2: Recursively decrease α until the distortions of all pixels are smaller than $NUJNCD$. An initial α is assigned, and calculate $\|D(n,m,u,v)\|^2$ for every pixel of the block. If some pixel distortion exceeds $(NUJNCD)^2$, we decrease α until $\|D(n,m,u,v)\|^2 < (NUJNCD)^2 \forall\ n,m$.

5.5.3 Embedding and Decoding Procedures

Given an image of size $N \times N$, the image is divided into blocks of size 8×8, followed by taking the QFT of each block. We embed a pair of marks $W(u', v')$ and $W(-u', -v')$ into the coefficients $F(u', v')$ and $F(-u', -v')$ in the following way:

$$WF(u', v') = F(u', v') \pm \alpha W(u', v')$$
$$WF(-u', -v') = F(-u', -v') \pm \alpha W(-u', -v'), \quad (5.37)$$

where α is the watermark strength to be maximized. Set $W(u', v') = (0, b_3, b_3, \frac{1}{5}b_3)$ and $W(-u', -v') = (0, b_3, b_3, \frac{1}{5}b_3)$ to generate watermarks that vary along the luminance and yellow-blue component. Notice that the real component is ignored when converting the watermarked image back to the La^*b^* color space, therefore, the real component $Re(m, n)$ is kept, as a key when extracting the watermark.

To decode the embedded watermark $W(u, v)$, the original image $F(m, n)$, the watermarked (possibly corrupted) image $WF(m, n)$, and $Re(m, n)$ are needed. First, the image is encoded to quaternion format, and replace the real component of $WF(m, n)$ with $Re(m, n)$, followed by taking the QFT of both images. Then $F(u', v') + \alpha W(u', v')$ and $F(u', v') - \alpha W(u', v')$ are calculated from the original image, followed by a distance comparison to the coefficient extracted from $WF(u', v')$. If the extracted coefficient is closer to $F(u', v') + \alpha W(u', v')$, bit 1 was embedded, otherwise, bit 0 was embedded. The flow diagram is shown in Figure (5.9).

5.6 Simulation Results

In order to prove the algorithms and theories described above indeed increases the robustness of the watermarked images against attacks, a series of experiments have been conducted by applying the attacks described in Chapter 2 to four images - the Lenna, Mandrill, and Deer images with size 512×480. The attacks include:

- Image Compression - JPEG.

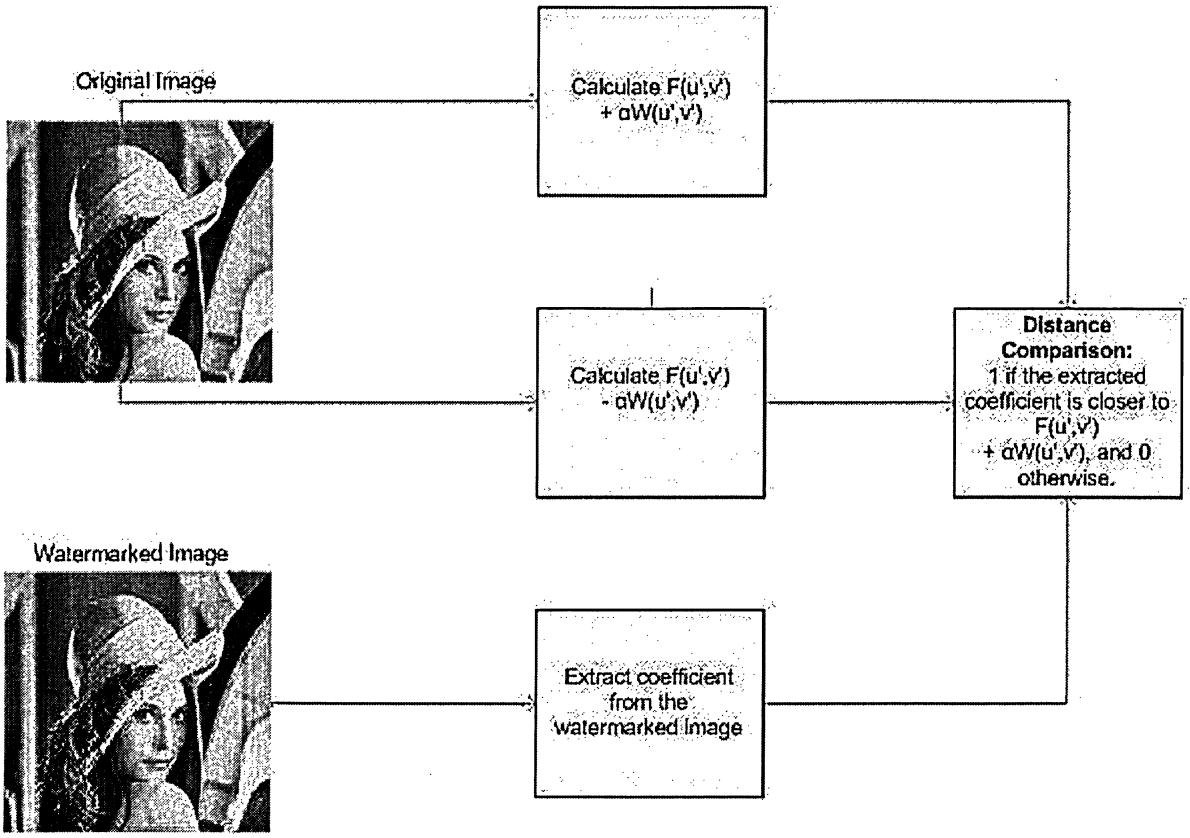


Figure 5.9: Flow diagram showing the extraction procedure using the QFT scheme.

- Geometric transformations - Resizing.
- Image Enhancement - Sharpening, and histogram equalization.
- Color Manipulation - Color to greyscale conversion
- Information Reduction - Gaussian noise.

One bit of information was embedded into each block imperceptibly using the above watermarking techniques on $AC(2,2)$ and $AC(6,6)$. The original Lenna, Mandrill, and Deer images are shown in Figures (5.10), (5.11) and (5.12), and 3840 bits of information were embedded into each image. Figure (5.13), (5.14), and (5.15) show the watermarked images using the SCDFT approach, where the watermarked images using the QFT approach are shown in Figure (5.16) , (5.17) and (5.18).

In order to test the robustness of the algorithms, some common attacks were applied, and the results were compared to two well-known watermarking algorithms - the DCT based and AC estimation based methods [21]. The PSNRs and the number of bits embedded in the Lenna image using the SCDFT, QFT, DCT, and AC algorithms are: (30.1027, 3840 bits), (30.9283, 3840 bits), (30.4064, 3840 bits) and (30.1081, 416 bits) respectively. Notice that the PSNRs are adjusted to be about the same in order to reflect the power of the embedded watermarks is also the same. The watermarked images using the AC estimation, and DCT are shown in Figure (5.19), (5.20), (5.21), (5.22), (5.23), and (5.24). Tables 5.1, 5.2 and 5.3 show the Bit Error Rates(BERs) against histogram equalization, sharpening, blurring, resizing, and color to greyscale conversion. The results for the robustness against JPEG compression and Gaussian noise are shown in Figure (5.28),(5.29), (5.30), (5.31), (5.32), and (5.33).

In order to show that the proposed schemes also work for images with poor texture, same attacks were also applied to the Blue image, as shown in Figure (5.34). The watermarked images using the SCDFT, QFT, DCT, and AC estimation are shown in Figures (5.35), (5.36), (5.37), and (5.38).

5.6.1 Discussions

From the measurements, several observations can be made:

- **SCDFT approach**: This scheme provides satisfactory results for most attacks. In addition, the compensation mark does reduce the visible distortion. Notice that the watermarked images using other schemes (i.e. Figures (5.19), (5.20), (5.21), (5.22), (5.23), and (5.24)) have perceptual artifacts under the same PSNR, where the SCDFT scheme does not. However, it is NOT robust against JPEG compression, especially when the compression factor is low (compression factor = compressed size / uncompressed size). The underlying reason is that the SCDFT algorithm only modifies the chromatic components of a color image, and JPEG quantizes them aggressively for high compression, which corrupts the watermark as the compression factor goes down. In addition, the watermark is completely destroyed if the

image is converted to greyscale. The underlying reason is because the watermark is resided in the chromatic components.

- **QFT approach**: The most successful scheme, which effectively solves the problem that the SCDFT has. The underlying reason is that the watermark is spread evenly among the luminance and chrominance components of the images. Also, the compensation mark makes the watermark being undetected even the strength of it is high. Even though it cannot completely recover the embedded messages for low compression factor, the BER is significant smaller than the SCDFT approach. Compared to the SCDFT approach, this scheme significantly reduces the BER rate for converting color to greyscale images by approximately 30%, because the watermark is embedded in the luminance as well as chrominance components. One small advantage of it is that the computational complexity is high for quaternion algebras.

- **AC estimation method**: It provides very good results for most attacks, and the biggest advantage of it is that it is a blind watermarking scheme. However, it has two major drawbacks - The maximum number of bits that can be embedded is much smaller, and the visible distortion is detectable at the same PSNR.

- **DCT approach**: Even it still provides satisfactory results, but the BERs of the attacks are generally higher. The underlying reason is that most existing DCT-based algorithms do not embed watermark in every single block of an image. They selectively pick the regions that do not generate visible distortion (for example, highly textured regions) for embedding, thus decreasing the data payload, but increasing the accuracy of extraction. In the experiments done in this thesis, in order to show that the proposed algorithms (i.e. SCDFT and QFT) can accept watermark in all regions (due to the compensation mark), this constraint was removed, and the watermarks were embedded in every block of the image. In addition, since this algorithm only embeds the watermark in the luminance component, ignoring the masking effects of the chromatic components, the imperceptibility of it is not as good. However, it only requires limited computational power due to the high availability of fast algorithms to compute the DCT and DFT.

Table 5.1: Lenna image: BERs against the attacks.

Attack	SCDFT	QFT	DCT	AC
No attack	0%	0%	0%	0%
Histogram equalization	0%	0%	3%	0%
Sharpening	0%	0%	4%	0.24%
Blurring	0%	0%	3%	0%
Resizing by factor 2	6.58%	0%	4.87%	0%
Color to greyscale conversion	49.96%	30.14%	0%	0%

Table 5.2: Mandrill image: BERs against the attacks.

Attack	SCDFT	QFT	DCT	AC
No attack	0%	0%	0%	0%
Histogram equalization	0%	0%	3%	0%
Sharpening	0%	0%	3%	0%
Blurring	0%	0%	3%	0%
Resizing by factor 2	10.6%	0%	6.22%	0%
Color to greyscale conversion	50.16%	28.86%	0%	0%

Table 5.3: Deer image: BERs against the attacks.

Attack	SCDFT	QFT	DCT	AC
No attack	0%	0%	0%	0%
Histogram equalization	0%	0%	4%	0%
Sharpening	0%	0%	3.13%	0%
Blurring	0%	0%	3.28%	0%
Resizing by factor 2	5.81%	0%	5.13%	0%
Color to greyscale conversion	48.88%	32.14%	0%	0%

Figure 5.10: Original Lenna image.

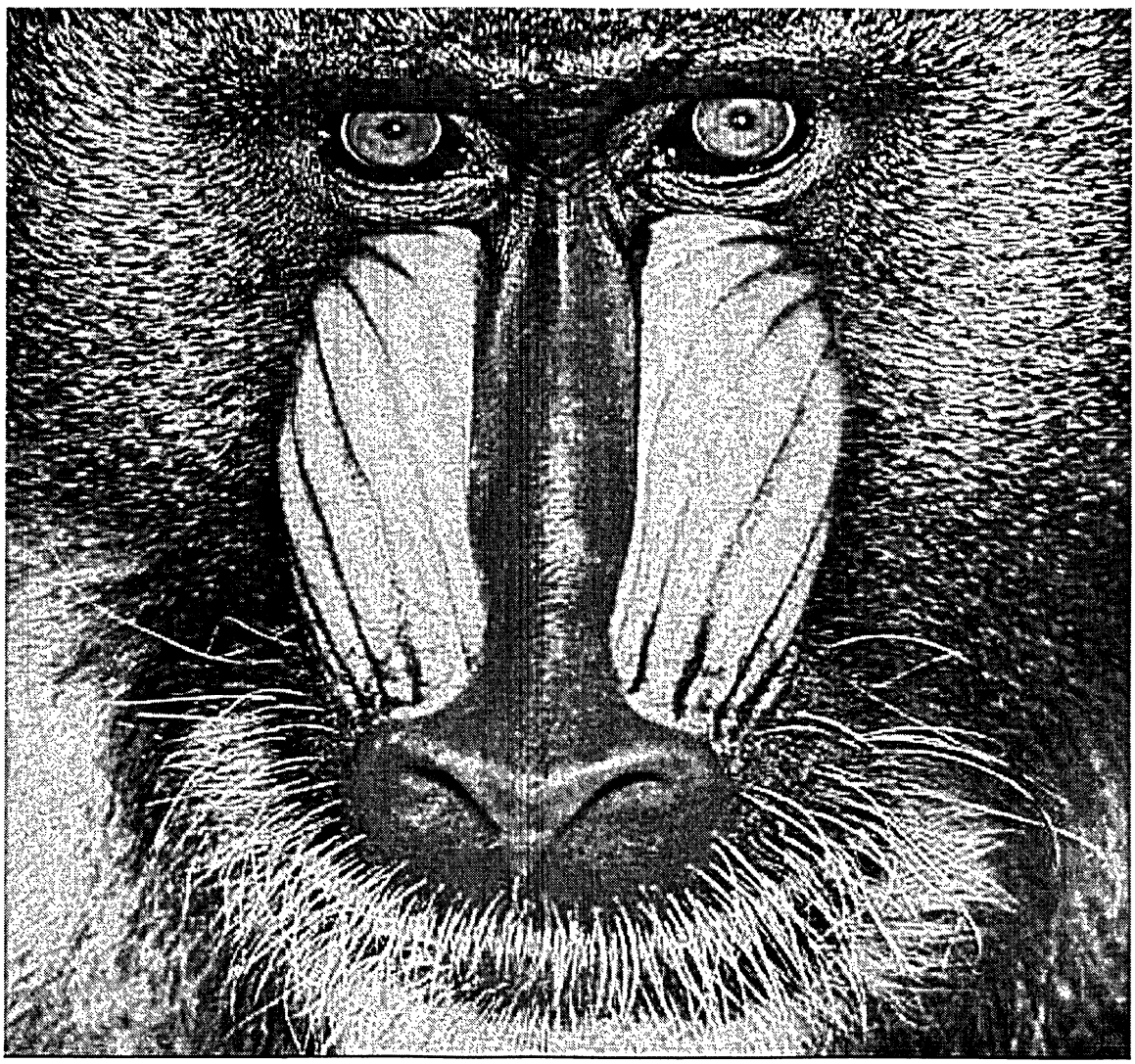


Figure 5.11: Original Mandrill image.

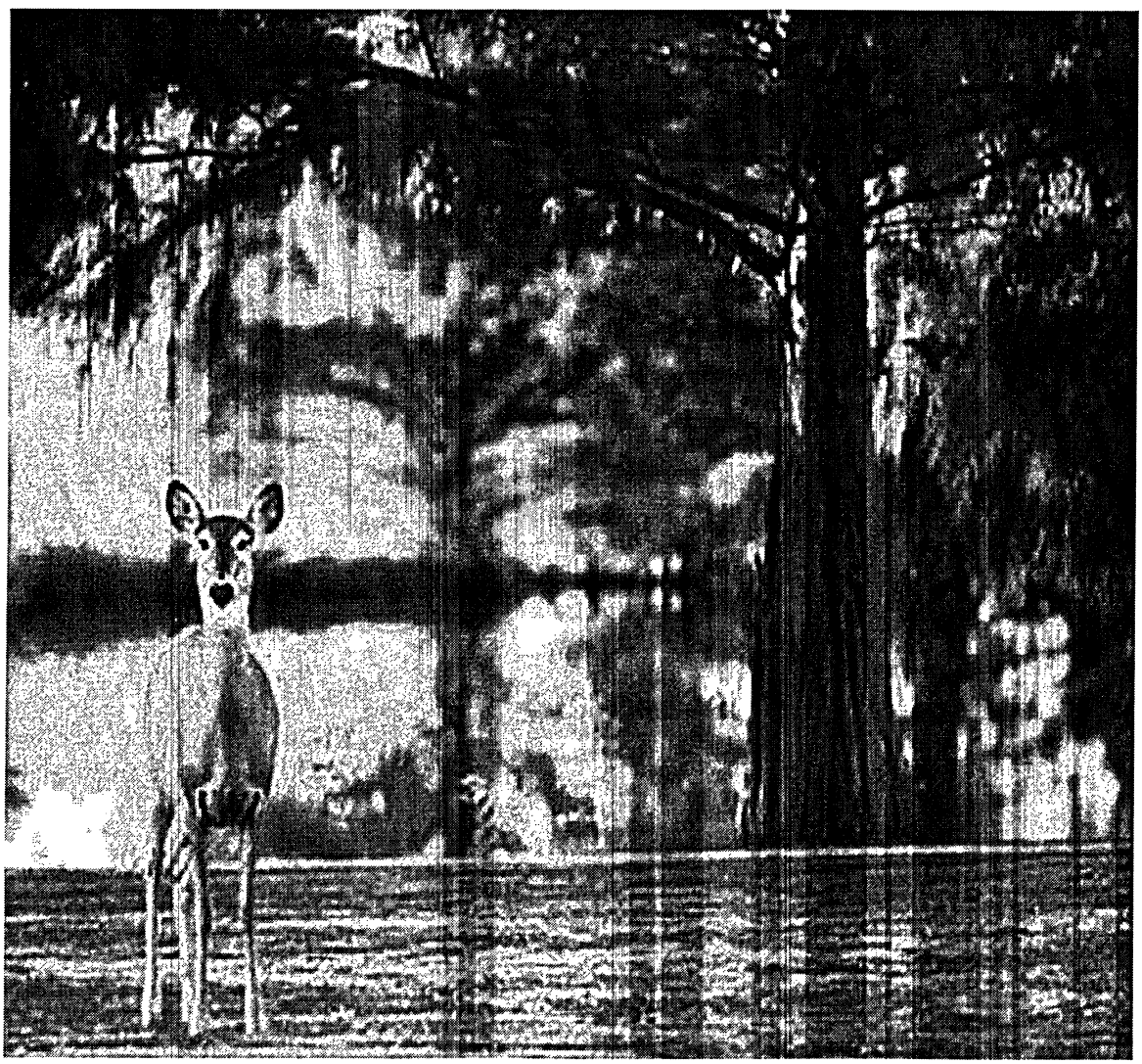


Figure 5.12: Original Deer image.

Figure 5.13: Watermarked Lenna image using the SCDFT.

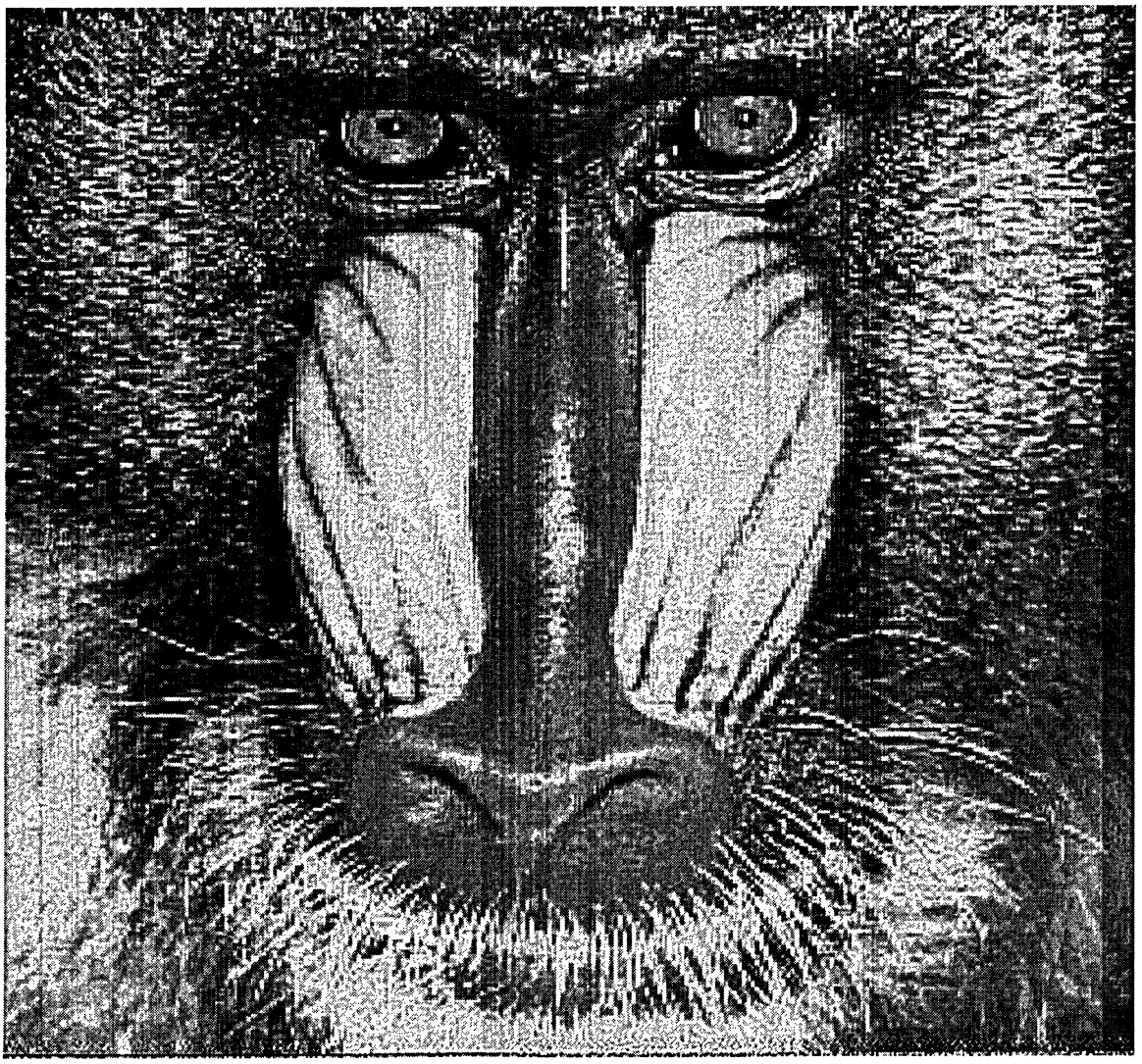


Figure 5.14: Watermarked Mandrill image using the SCDFT.

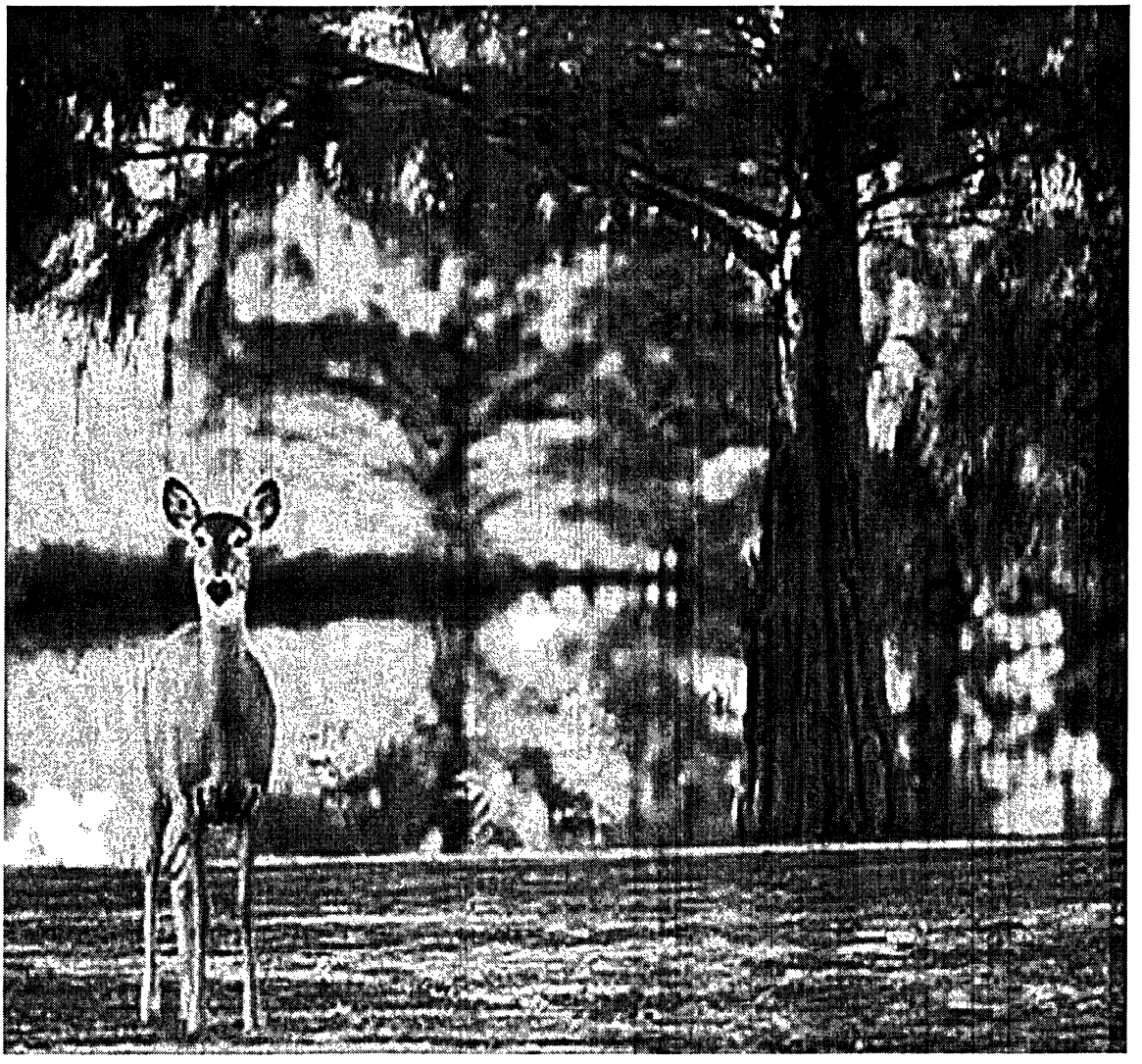


Figure 5.15: Watermarked Deer image using the SCDFT.

Figure 5.16: Watermarked Lenna image using the QFT.

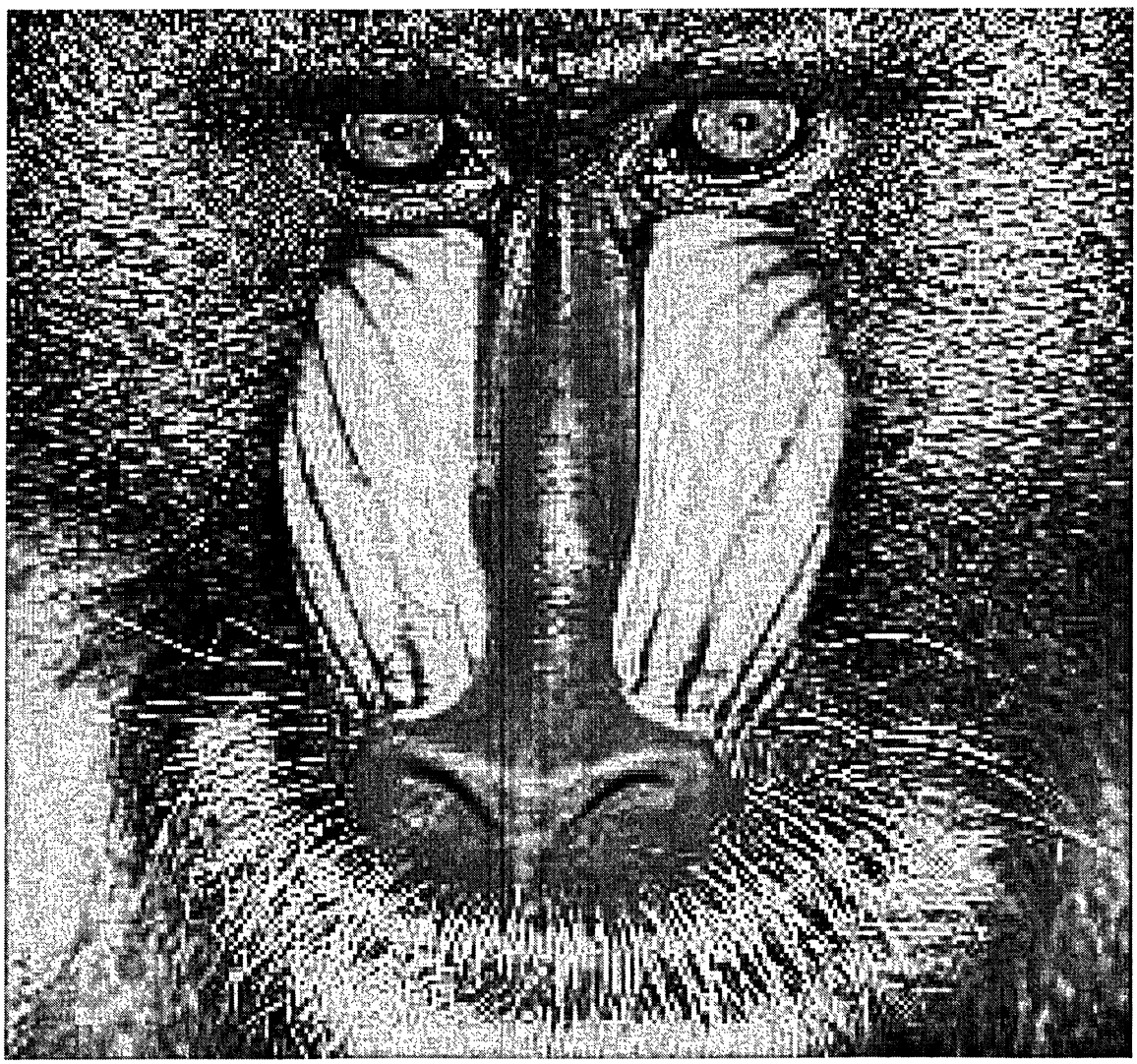


Figure 5.17: Watermarked Mandrill image using the QFT.

Figure 5.18: Watermarked Deer image using the QFT.

Figure 5.19: Watermarked Lenna image using the AC estimation.

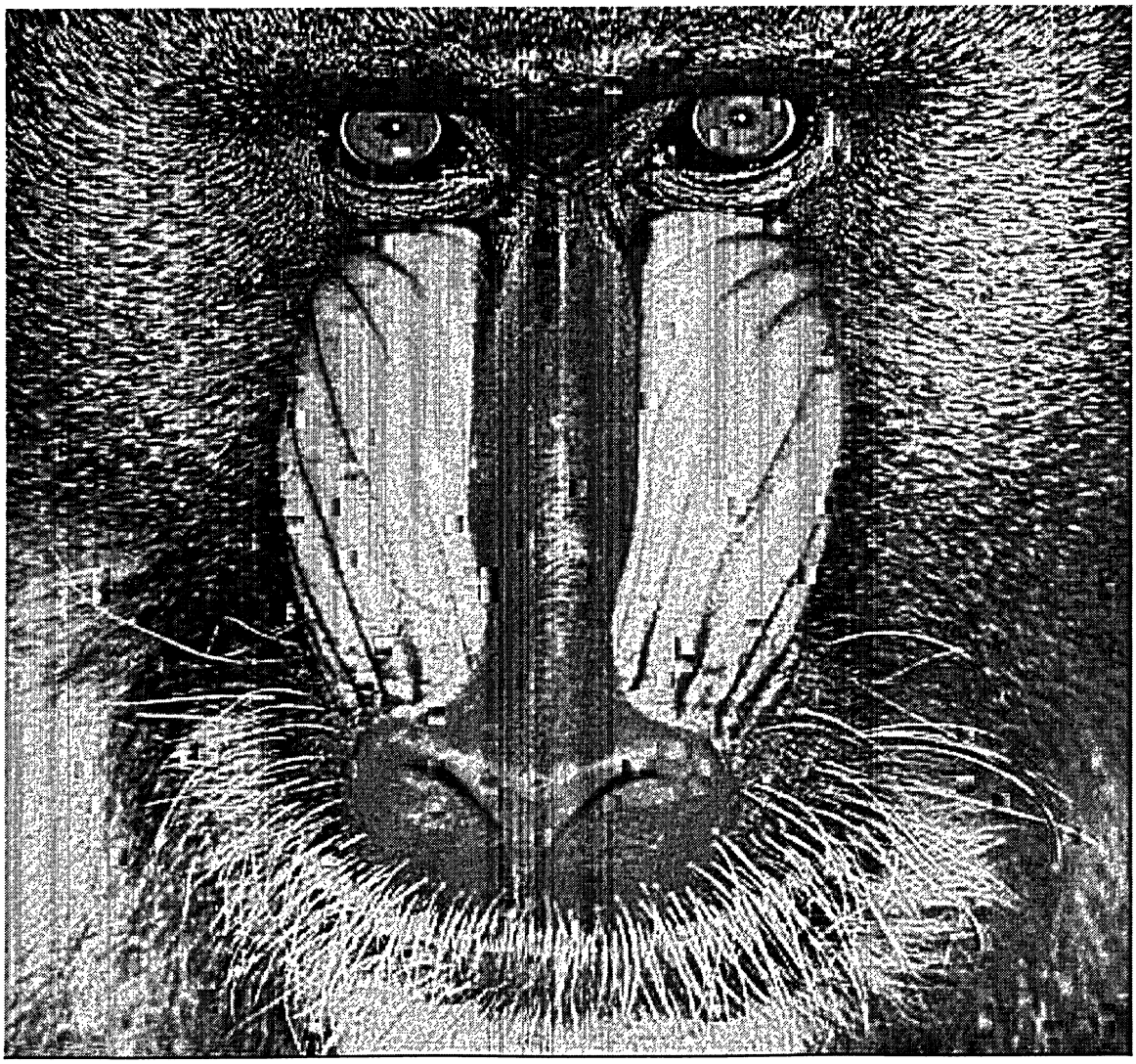


Figure 5.20: Watermarked Mandrill image using the AC estimation.

Figure 5.21: Watermarked Deer image using the AC estimation.

Figure 5.22: Watermarked Lenna image using the DCT.

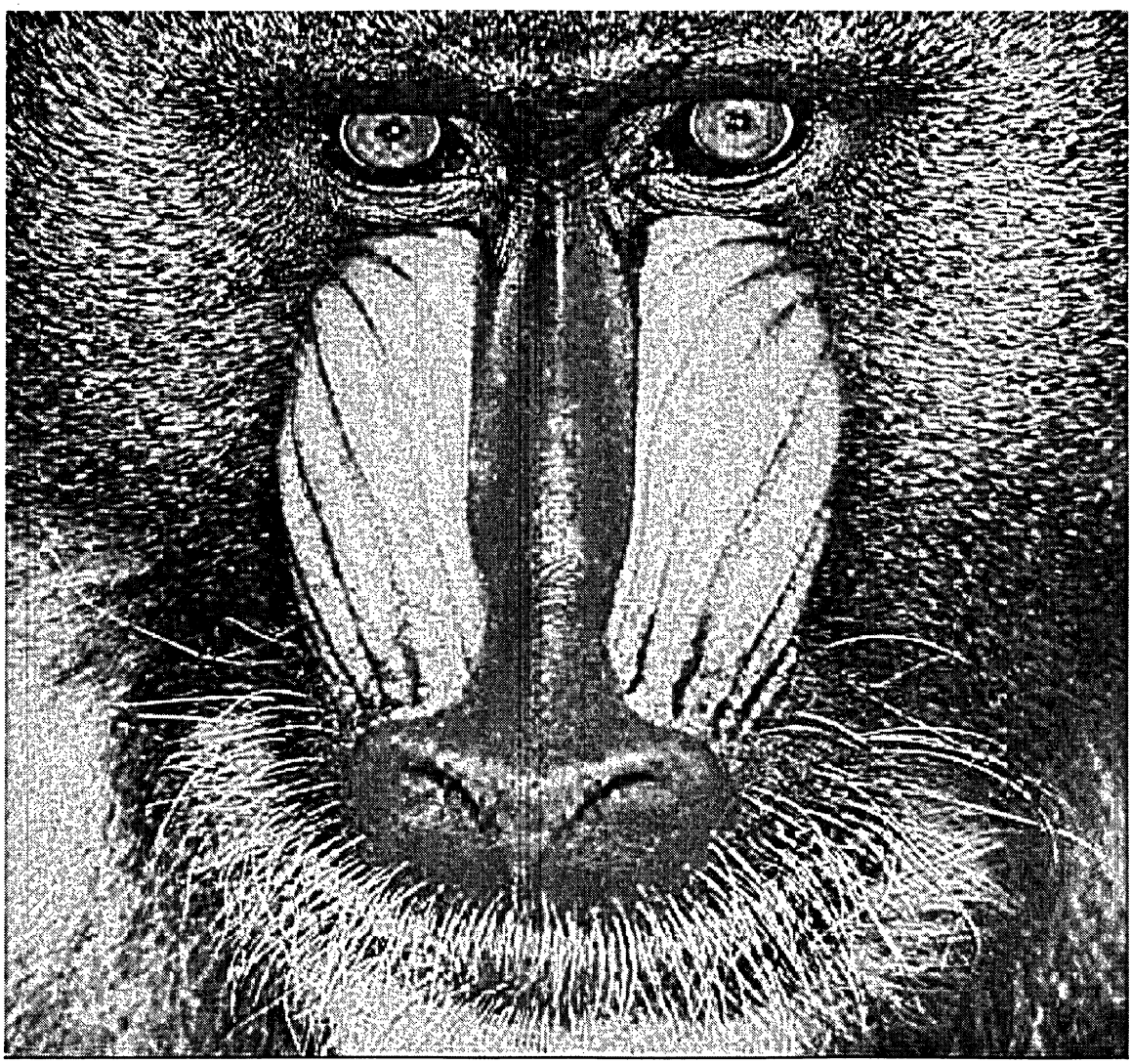


Figure 5.23: Watermarked Mandrill image using the DCT.

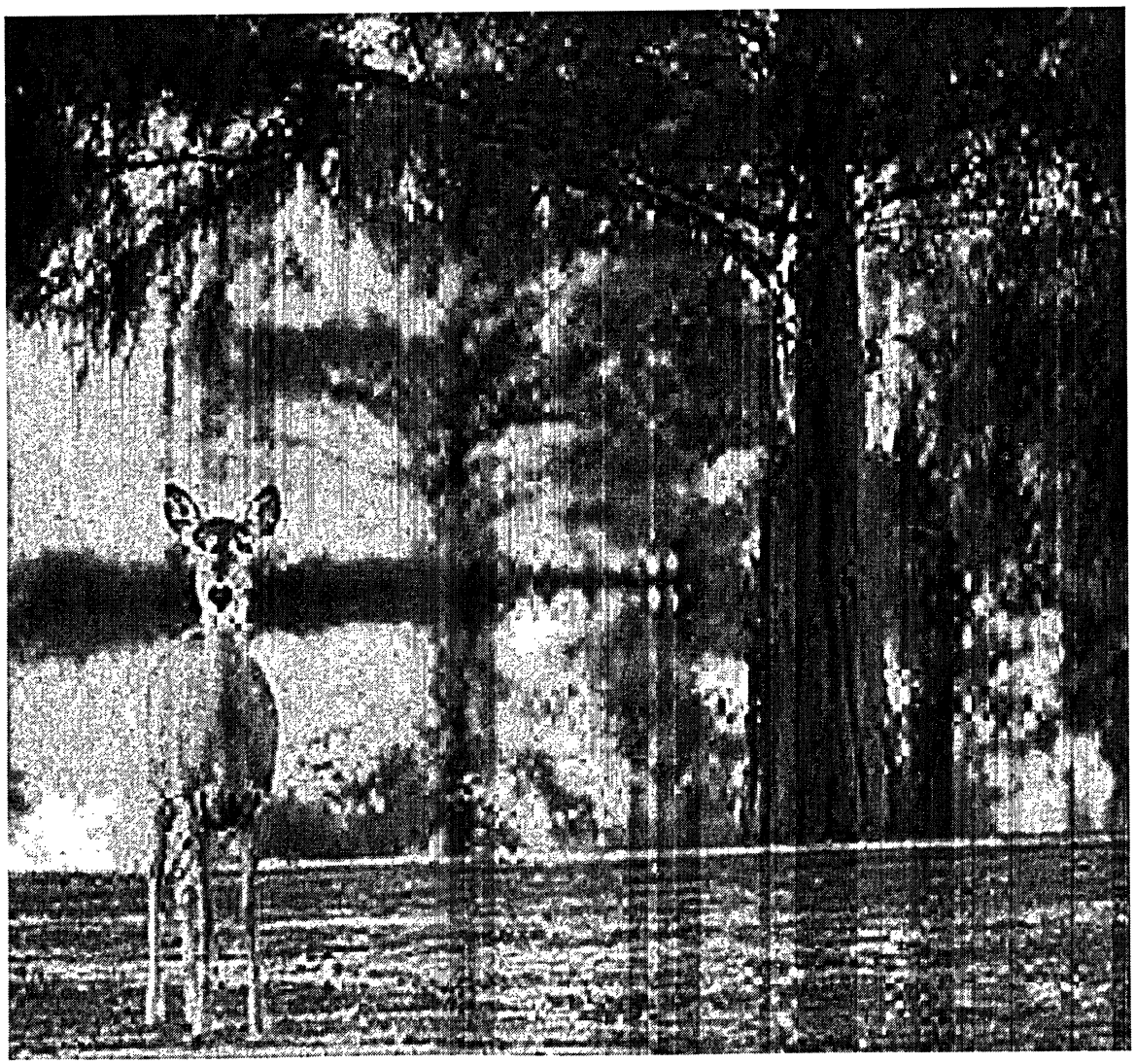


Figure 5.24: Watermarked Deer image using the DCT.

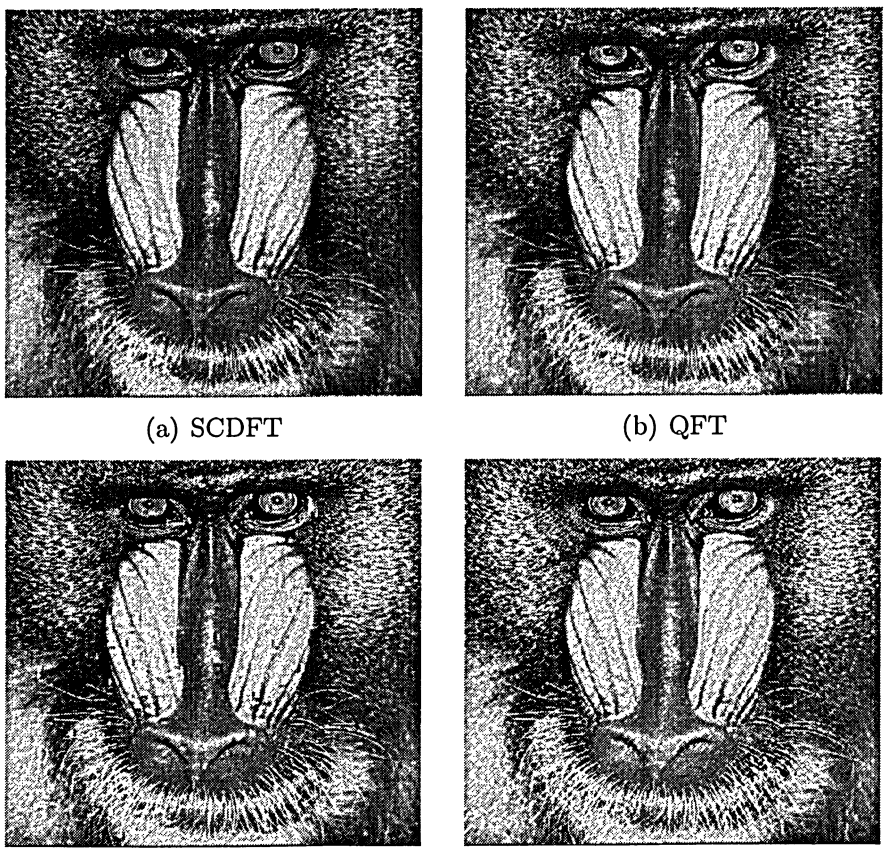
(a) SCDFT (b) QFT

(c) AC (d) DCT

Figure 5.25: Side-by-Side view of the Lenna watermarked images.

5.7 Chapter Summary

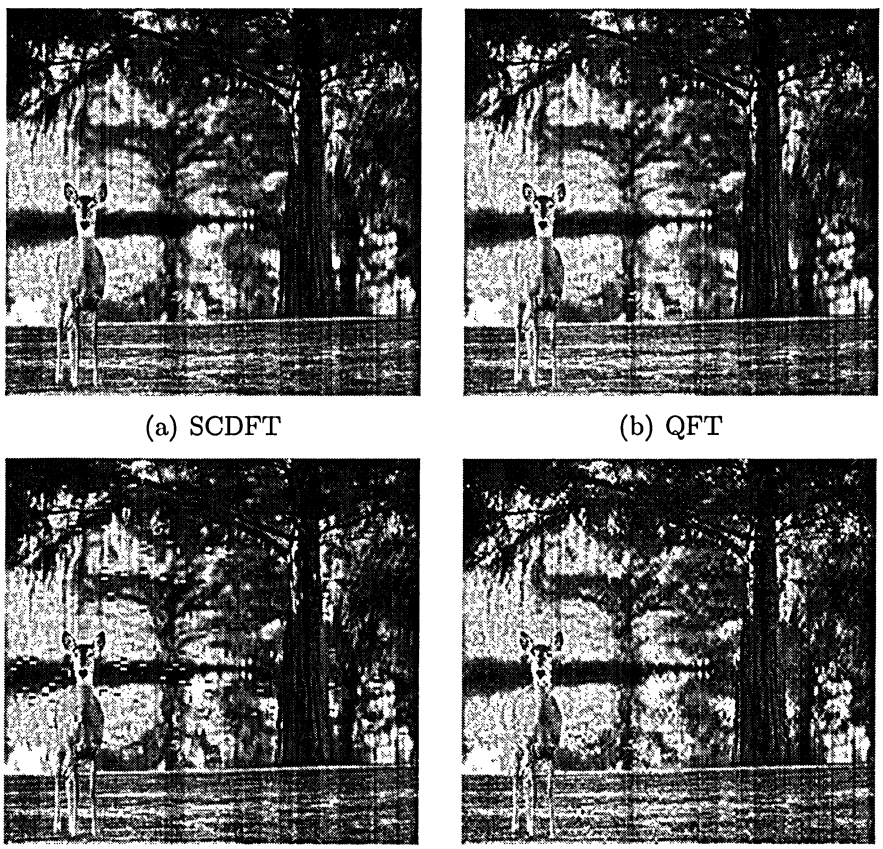
This chapter gives a detail description of two proposed algorithms using the SCDFT and QFT. Experimental results are also provided to compare their effectiveness against other existing algorithms.



(a) SCDFT (b) QFT

(c) AC (d) DCT

Figure 5.26: Side-by-Side view of the Mandrill watermarked images.



(a) SCDFT (b) QFT

(c) AC (d) DCT

Figure 5.27: Side-by-Side view of the Deer watermarked images.

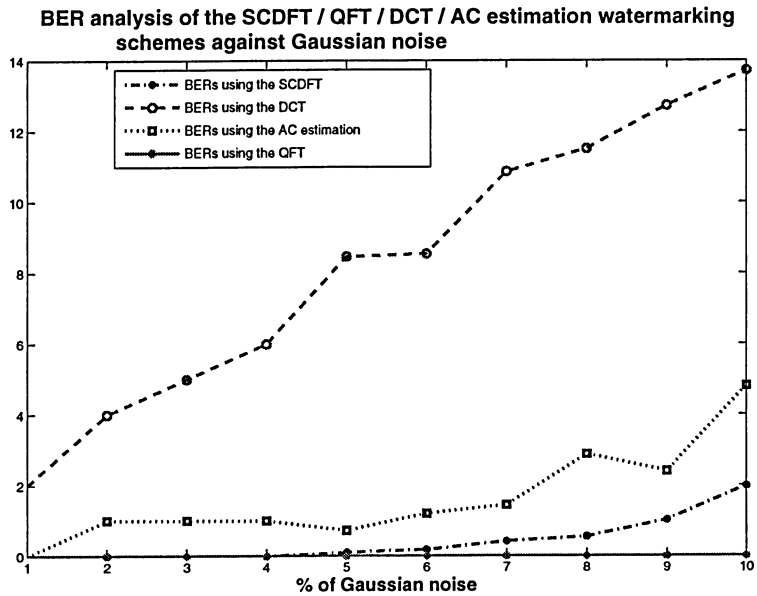


Figure 5.28: Lenna image: BERs against Gaussian noise.

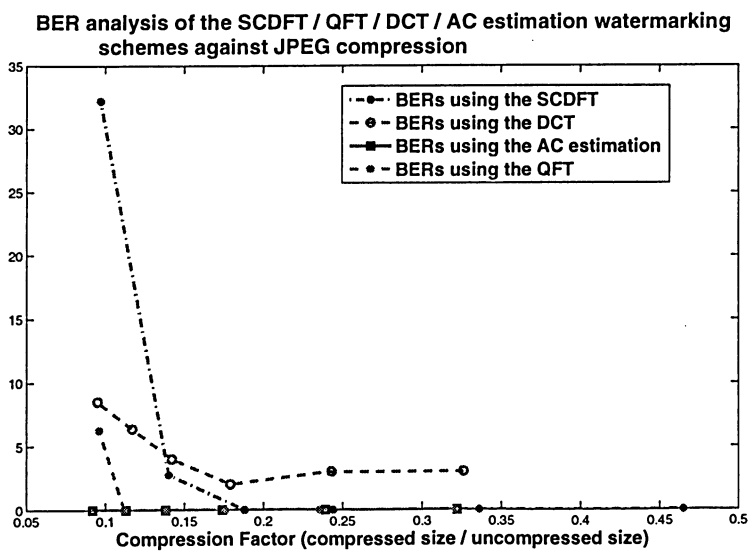


Figure 5.29: Lenna image: BERs against JPEG compression.

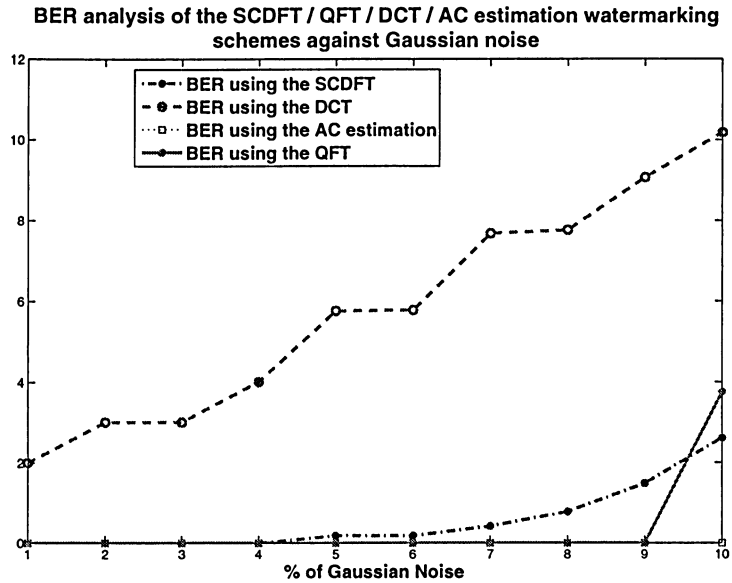


Figure 5.30: Mandrill image: BERs against Gaussian noise.

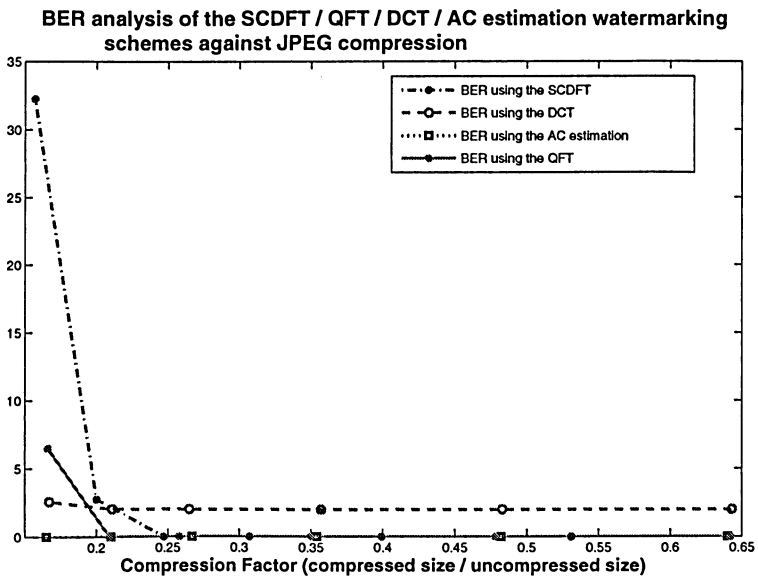


Figure 5.31: Mandrill image: BERs against JPEG compression.

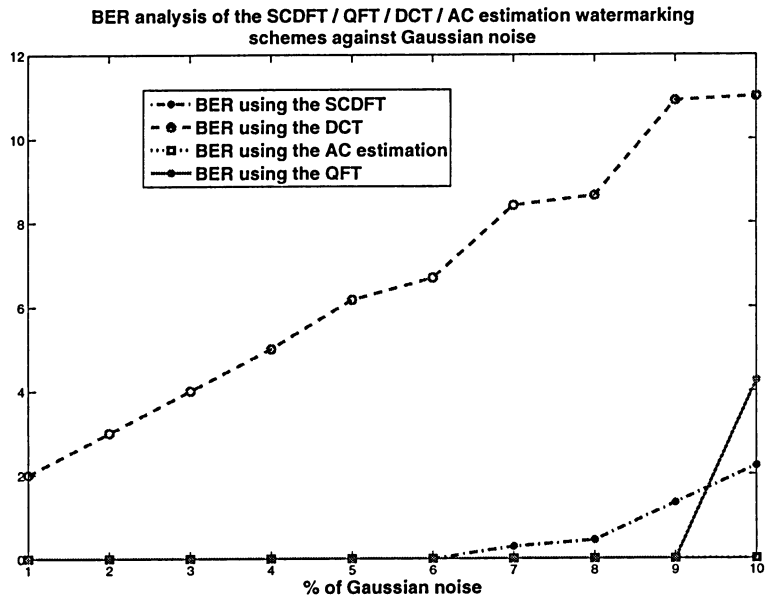


Figure 5.32: Deer image: BERs against Gaussian noise.

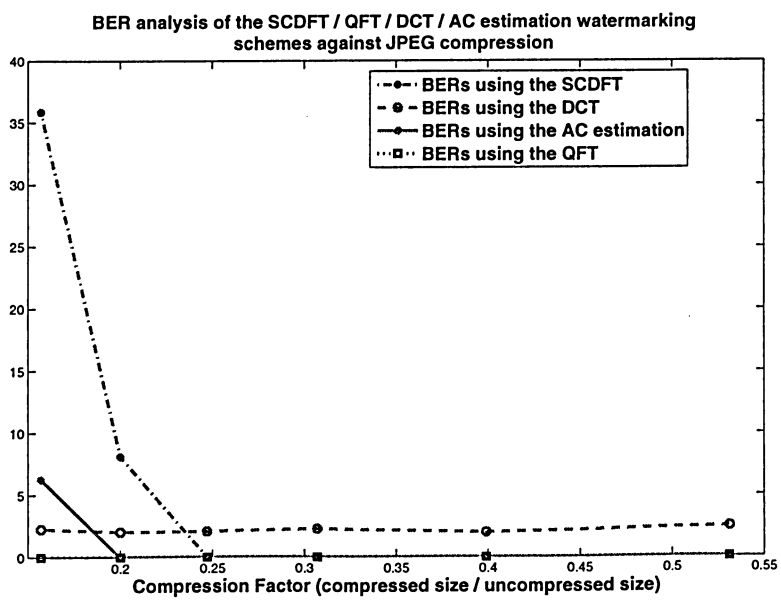


Figure 5.33: Deer image: BERs against JPEG compression.

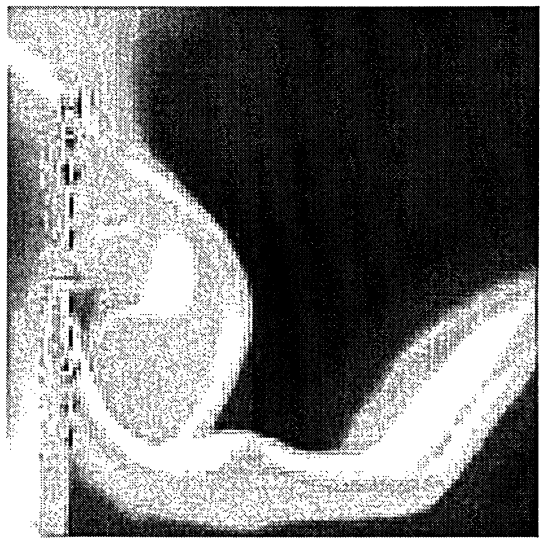


Figure 5.34: Original Blue image.

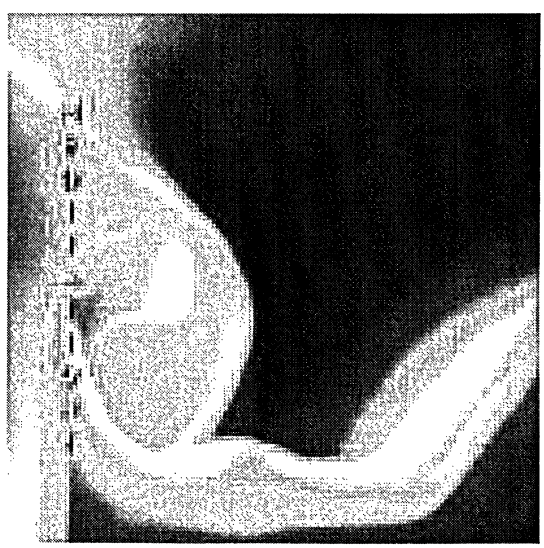


Figure 5.35: Watermarked Blue image using the SCDFT.

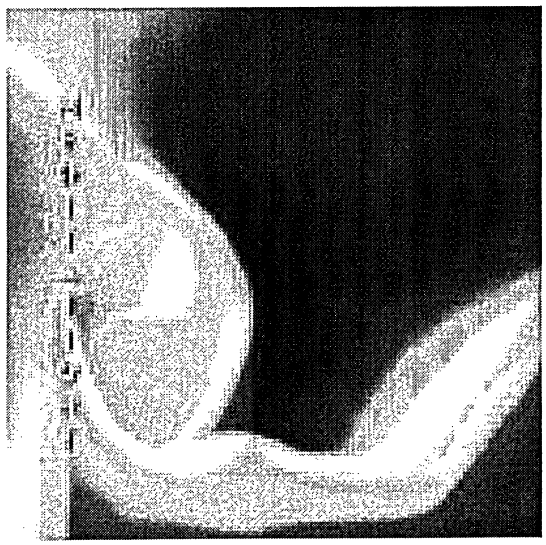


Figure 5.36: Watermarked Blue image using the QFT.

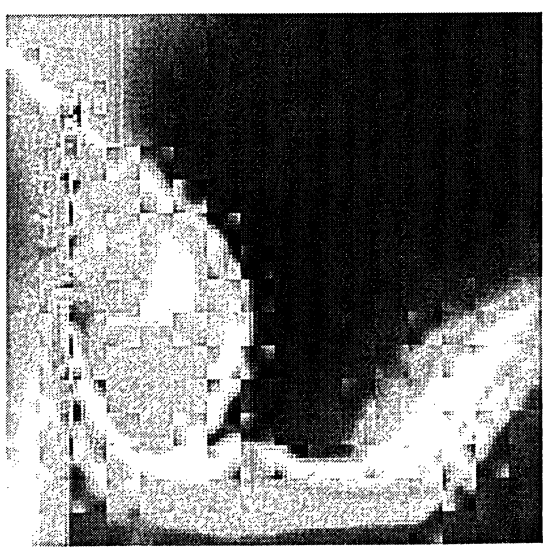


Figure 5.37: Watermarked Blue image using the DCT

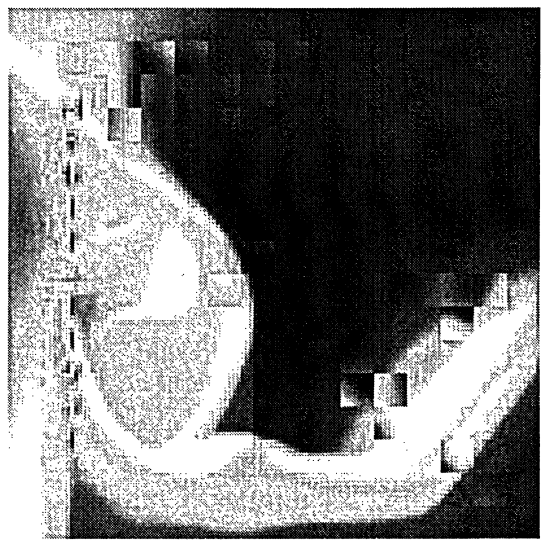


Figure 5.38: Watermarked Blue image using the AC

Table 5.4: Blue image: BERs against the attacks.

Attack	SCDFT	QFT	DCT	AC
No attack	0%	0%	0%	0%
Histogram equalization	0%	0%	5.47%	0%
Sharpening	0%	0%	6.64%	0%
Blurring	0%	0%	5.08%	0%
Resizing by factor 2	8.11%	0%	7.42%	0%
Color to greyscale conversion	48.86%	31.25%	0%	0%
JPEG compression	14.84%	10.38%	13.44%	0%
10% Gaussian noise	5.84%	0%	15.58%	4.17%

Chapter 6

Conclusions and Future Work

WE have investigated a new approach for embedding watermarks into all components (luminance and chrominance) of color images. Chapter 1 gives a general overview of the classification on watermarking techniques. Chapter 2 provides a mathematical model of watermarking by breaking down the problem into several building blocks. Solutions and further references to tackle the blocks are also provided. Chapter 3 and Chapter 4 provides some background information about the SCIP and QIP processing techniques and their applications, where Chapter 5 introduces two new color watermarking schemes using the SCDFT and QFT. The idea is inspired by the fact that colors can be added and subtracted to produce other colors. The technique explores the interpretation of the complex components of the frequency components using the two transforms, and utilizes the characteristic that human eyes are insensitive to yellow and blue to develop two embedding algorithms. By embedding two marks (i.e. watermark and compensation mark) in the coefficients, a composite mark that only contains insensitive colors is able to produce. Results demonstrate the proposed algorithms are robust to many digital signal processing operations and external attacks, because the strength of the watermark is maximized, whereas the imperceptibility is also maximized due to the reason that the compensation mark reduces the perceptual distortion generated by adding the watermark. Future work can be extended to:

- Crossed masking effects between the luminance and chrominance components - The interaction between the luminance and chrominance components can be further analyzed

[48] in order to improve the imperceptibility of the watermarked images.

- The proposed algorithms can be extended such that watermarks can be embedded into multiple frequency coefficients.
- More understanding to the QFT spectral coefficients is helpful to provide a more sophisticated watermarking algorithm.
- A better estimation of the upper bound $JNCD$ can be calculated by incorporating more characteristics of the HVS.
- The theories might be applicable to other similar fields such as color image coding, and color image compression.

Watermarking algorithms considering both luminance and chromatic components of color images in a holistic manner are expected to be the future trend of development.

Bibliography

[1] J. Simpson and E. Weiner, *Oxford English Dictionary* (Oxford University Press, USA, 2000).

[2] D. Huang and H. Yan, Interword distance changes represented by sine waves forwatermarking text images, in *IEEE Transactions on Circuits and Systems for Video Technology* Vol. 11, pp. 1237–1245, Sydney, 2001.

[3] J. Brassil, N. M. S. Low, and L. O'Gorman, Electrical marking and identification techniques to discourage document copying, in *IEEE Journal on Selected Areas in Communications* Vol. 13, pp. 1495–1504, 1995.

[4] S. H. Low and N. Maxemchuk, Performance comparison of two text marking methods, in *IEEE Journal on Selected Areas in Communications* Vol. 16, pp. 561–572, 1998.

[5] S. H. Low, N. Maxemchuk, J. Brassil, and L.O'Gorman, Document marking and identification using both line and word shifting, in *Proc. Infoncom'95*, pp. 853–860, Boston, MA, 1995.

[6] K. Tanaka, Y. Nakamura, and M. K., Embedding secret information into a dithered multi-level image, in *Proceedings of the IEEE MILCOM International Conference*, pp. 216–220, 1990.

[7] S. Gilani and A. Skodras, Dlt-based digital image watermarking, in *Proceedings of the First IEEE Balkan Conference on Communications, Circuits and Systems*, Istanbul, Turkey, 2000.

[8] A. Piva, L. Boccardi, F. B. M. Barni, V. Cappellini, and A. D. Rosa, Watermarking through color image bands decorrelation, in *IEEE International Conference on Multimedia and Expo ICME 2000*, pp. 1283–1286, New York City, NY, USA, 2000.

[9] A. Piva, F. Bartolinin, V. Cappellini, and M. Barni, Exploiting the cross-correlation of rgb-channels for robustwatermarking of color images, in *Proceedings of the International Conference on Image Processing, ICIP 99* Vol. 1, pp. 306–310, Kobe, Japan, 1999.

[10] C. I. P. R. B. Wolfgang and E. J. Delp, Perceptual watermarks for digital images and video, in *Proceedings of the IEEE* Vol. 87, pp. 1108–1126, 1999.

[11] I. Cox, J. Kilian, T. Leighton, and T. Shamoon, Secure spread spectrum watermarking for images, audio and video, in *Proc. IEEE Int. Con5 on Image Processing* Vol. 111, pp. 243–246, Lausanne, Switzerland, 1996.

[12] C. N and T. I, Increasing robustness of patchwork audio watermarking algorithm using attack characterization, in *Proc. 2004 IEEE International Symposium on Consumer Electronics*, pp. 3–6, UK, 2004.

[13] C. N and S. T, Robust audio watermarking in wavelet domain using frequency hopping and patchwork method, in *Proc. 3rd International Symposium on Image and Signal Processing and Analysis*, pp. 251–255, Rome, Italy, 2003.

[14] I.-K. Yeo and H. J. Kim, Modified patchwork algorithm: A novel audio watermarking scheme, in *International Conference on Information Technology: Coding and Computing (ITCC '01)*, p. 237, 2001.

[15] khalid Sayood, *Introduction to Data Compression* (Morgan Kaufmann, USA, 2000).

[16] B. A. Wandell, Foundations of vision, in *Sinauer Associates*, p. 328, Sunderland, Massachusetts, 1995.

[17] I. J. Cox, J. Killian, F. T. Leighton, and T. Shamoon, Secure spread spectrum watermarking for multimedia, in *IEEE Trans. Image Process* Vol. 6, pp. 1673–1687, 1997.

[18] N. Ahmidi and R. Safabakhsh, A novel dct-based approach for secure color image watermarking, in *Proceedings of the International Conference on Information Technology.*

[19] C.-H. Chou and Y.-C. Li, A perceptually tuned subband image coder based on the measure of just-noticeable distortion profile, in *IEEE Trans. Circuits and Systems for Video Technology* Vol. 5, pp. 467–476, 1995.

[20] L. Ghouti and A. Bouridane, A just-noticeable distortion (jnd) profile for balanced multiwavelets, in *EUSIPCO2005 - European Signal Processing Conference*, 2005.

[21] Y. Wang and A. Pearmain, Ac estimation-based image watermarking method, in *5th European workshop on Image Analysis and Multimedia*, Lisbon, Portugal, 2004.

[22] C. Podilchuk and W. Zeng, Image-adaptive watermarking using visual models, in *IEEE Journal on Selected Areas in Communications* Vol. 10, pp. 525–540.

[23] K. R. N. Kaewkamnerd, Wavelet based image adaptive watermarking scheme, in *Electronics Letters* Vol. 36, pp. 518–526, 2000.

[24] A. Ho, J. Shen, and S. Tan, Robust digital image-in-image watermarking algorithm using the fast hadamard transform, in *Proc. SPIE 4793*, 2002.

[25] S.A.M.Gilani, I.Kostopoulos, and A.N.Skodras, Color image-adaptive watermarking, in *14th International Conference on Digital Signal Processing*, pp. 721 – 724, Santorini, Greece, 2002.

[26] V. Licks and R. Jordan, On digital image watermarking robust to geometric transformations, 2000.

[27] J. Bloom *et al.*, Copyright protection for digital video, in *Proc. IEEE Special Issue on Identification and Protection of Multimedia Information*, pp. 1267–1276, 1999.

[28] D. Boneh and J. Shaw, Collusion-secure fingerprinting for digital data, in *IEEE Trans. Inf. Theory* Vol. 5, pp. 1897–1905, 1998.

[29] D. Kundur and D. Hatzinakos, Digital watermarking for telltale tamper proofing and authentication, in *Proceedings of the IEEE (USA)* Vol. 87, p. 1167V1180, 1999.

[30] X. Z. T.K. Tsui and D. Androutsos, Color image watermarking using the spatio-chromatic fourier transform, in *ICASSP*, pp. 305–308, Toulouse, France, 2006.

[31] X. Z. T.K. Tsui and D. Androutsos, Quaternion image watermarking using the spatio-chromatic fourier coefficients analysis, in *ACM*, Santa Barbara, USA, 2006.

[32] I. J. Cox, M. L. Miller, and J. A. Bloom, *Digital Watermarking* (Academic Press, USA, 2002).

[33] A. McCabe and T. Caelli, Theory of spatiochromatic image encoding and feature extraction, in *J. of the Optical Society of America A*, pp. 1744–1754, 2000.

[34] S. J. Sangwine, The problem of defining the fourier transform of a colour image, in *ICIP'98*, pp. 171–175, Chicago, USA, 1998.

[35] S. J. Sangwine, Colour image edge detector based on quaternion convolution, in *Electron. Letter*, pp. 969–971, 1998.

[36] T. Bulow, Hypercomplex spectral signal representations for the processing and analysis of images, in *Ph.D. Thesis, Christian-Albrechts Univ. of Kiel*, 1999.

[37] S. C. Pei, J. J. Ding, and J. Chang, Efficient implementation of quaternion fourier transform convolution, and correlation by 2-d fft, in *IEEE Trans. Signal Processing*, pp. 2787–2797, 2001.

[38] S. J. Sangwine and T. A. Ell, The discrete fourier transforms of a colour image, in *Image Processing Mathematical Methods, Algorithms and Applications*, pp. 430–411, 2000.

[39] P. Bas, N. L. Bihan, and J. Chassery, Color watermarking using quaternion fourier transform, in *ICASSP*, pp. 521–524, Hong Kong SAR, China, 2003.

[40] A. Reed and B.Hannigan, Adaptive color watermarking, in *SPIE*, 2002.

[41] Y. Shi and H. Sun, Image and video compression for multimedia engineering: Fundamentals, algorithms and standards, in *Boca Raton*, 1999.

[42] B. Watson, Dct quantization matrices visually optimized for individual images, in *SPIE: Human Vision, Visual Processing and Digital Display IV* Vol. 1913, pp. 202–216, 1993.

[43] J. J. N. Jayant and R. Safranek, Signal compression based on model of human perception, in *Proc. IEEE* Vol. 81, pp. 1385–1421, 1993.

[44] M. D. Swanson, M. Kobayashi, and A. H. Tewfik, Multimedia data embedding and watermarking technologies, in *Proc. IEEE* Vol. 86, pp. 1064–1087, 1998.

[45] B.Watson, Efficiency of a model human image code, in *J. Opt. Soc. Am* Vol. 4, p. 2401V2416, 1987.

[46] C.-H. Chou and K.-C. Liu, An oblivious and robust watermarking scheme using perceptual model, in *Conf. Video Processing and Multimedia Communications*, pp. 713–720, 2003.

[47] R. G. Kuehni, *Color space and its divisions : color order from antiquity to the present* (J. Wiley, USA, 2003).

[48] Y. Meng and L. Guo, Color image coding by utilizing the crossed masking, in *ICASSP*, pp. 389– 392, 2005.