國立交通大學

資訊工程學系

博士論文

抗傳輸通道損失與幾何失真之

強韌性浮水印

Robust Watermarking Against Transmission Channel Loss and Geometric Distortion

研究生：戴敏倫

指導教授：李素瑛 教授

周義昌 教授

中華民國九十六年一月

抗傳輸通道損失與幾何失真之強韌性浮水印

Robust Watermarking Against Transmission Channel Loss and Geometric Distortion

研究生：戴敏倫 Student： Miin-Luen Day

指導教授：李素瑛 博士 Advisor： Dr. Suh-Ying Lee

周義昌 博士 Dr. I-Chang Jou

國立交通大學
資訊工程學系
博士論文

A Dissertation
Submitted to Department of Computer Science
College of Computer Science
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Doctor
in

Computer Science

Jan. 2007

Hsinchu, Taiwan, Republic of China

中華民國九十六年一月

抗傳輸通道損失與幾何失真之強韌性浮水印

學生：戴敏倫　　　　　　　　　　　　指導教授：李素瑛 教授

周義昌 教授

國立交通大學資訊工程學系

摘要

隨著網際網路、多媒體與電子商務的盛行，大量數位化資訊的傳送與儲存變的既快速又方便，數位化資訊已逐漸融入我們日常生活中，因此在資訊安全課題上對於個人私密性資料的保護與智慧財產權的保障越來越受到重視。浮水印(資訊隱藏)技術是將某些重要訊息隱藏於文字、聲音、影像或視訊等多媒体資料中，以達到所有權保護、防止盜拷、認證、內容連結(隱藏性標題)與秘密通訊等多種應用。由於浮水印用途廣泛且具潛在商機，其牽涉到的技術包括密碼學、數位信號(影像與聲音)處理、資訊理論與數位通訊各個研究領域，兼具理論與實用價值，因此近十多年來吸引了學術界與產業界眾多人士投入相關的研發。本論文主要目標則在於研發抗傳輸通道損失與幾何失真之強韌性浮水印嵌入與偵測演算法，適於所有權保護、內容連結與秘密通訊之多種應用需求。

在論文的第一部份，我們探討適用於在不可靠之 IP 傳輸網路所需之容錯架構及其浮水印嵌入與抽取演算法，並提出兩種可行之方法。採用的容錯架構為多重描述編碼(Multiple Description Coding, MDC)，首先在浮水印嵌入與抽取演算法提出了索引值餘數量化法(Quantization Index Modulus Modulation, QIMM)，此法經理論分析與大量實驗結果顯示其與目前最尖端的索引值量化法(Quantization Index Modulation, QIM)效能相當。接著我們整合了 QIMM 與 QIM 於 MDC，得到多重描述浮水印(Multiple Description Watermarking, MDW)架構。方法一將浮水印嵌入在任一個子描述 (side description)，其可由接收到之任一個子描述抽取出浮水印。方法二則另提出一更佳之浮水印嵌入與抽取演算法，稱之為多速率格子

索引值量化法(Multi-Rate Lattice Quantization Index Modulation, MRL-QIM)，並將浮水印嵌入在中央描述(central description)，也可由接收到之任一個子描述抽取出浮水印。相較於方法一，方法二之優點在於其所利用之MRL-QIM浮水印嵌入編碼效益較高且浮水印嵌入在中央描述較一般化，彈性較好。

在論文的第二部份，我們探討了幾何不變性數位浮水印之可行方式。首先我們提出結合根據情報先行編碼(informed coding)與Foruier-Mellin轉換之抗旋轉、縮放與平移以及其它多種攻擊之浮水印方法。主要是根據情報先行編碼前處理(informed coding pre-processing)以獲得一最佳浮水印，再嵌入於影像特定區域之Foruier-Mellin幾何不變域。此機制相較於原先未採用根據情報先行編碼前處理之方法，在相同的誤判率(false alarm rate)下大幅提高了偵測率(detection rate)。不同於利用幾何不變域之方法，另一類為利用重新同步(re-synchronization)或自我同步(self-synchronization)之方法，我們也嘗試了多種解決方式並提出一種利用二維條碼之自我同步數位浮水印嵌入與偵測方法。此方法為採用兩階段式(two-stage)之方式，主要為在第一階段採用QR Code (Quick Response Code)二維條碼之編碼方法來編碼數位浮水印訊息(payload)，第二階段利用索引值餘數量化(QIMM)來執行QR Code數位浮水印訊息之嵌入與抽取。由於QR Code二維條碼具有高容量(capacity)、高密度(compact size)與高容錯之特性，有利於數位浮水印訊息之編碼與解碼。另外透過QIMM在高容量訊息嵌入與偵測能力可有效的完成幾何不變性數位浮水印。本方法可有效的將數位浮水印訊息隱藏於影像中並且能抵抗包含旋轉、縮放與平移之幾何攻擊，在可隱藏訊息量、影像保真度(transparency)與強韌性三方面之綜合功效優於現有之方法。

Robust Watermarking Against Transmission Channel Loss and Geometric Distortion

Student: Miin-Luen Day

Advisor: Prof. Suh-Yin Lee
Prof. I-Chang Jou

Department of Computer Science
National Chiao-Tung University

Abstract

Watermarking is a technique to hide data or information imperceptibly within image, audio or video so that valuable contents can be protected. Since the application of watermark is extensive and its market potential is quite promising, and the design of watermarking algorithms implies the integration of many concepts coming from cryptography, digital communication and signal processing. The development of efficient watermarking algorithms has been a very active topic for researchers in this area. We focus on two categories of problems. One is the problem of watermarking for error-prone transmission over unreliable network and the other is the problem of achieving watermark robustness against geometric attacks.

In the first part of the dissertation, we study the problem of watermarking for error-prone transmission over unreliable network and two approaches are proposed. The first approach is to integrate oblivious watermarking techniques (quantization index modulus modulation (QIMM) and QIM) with the multiple description coding (MDC) to get a multiple description watermarking (MDW) framework. In this framework, the watermark embedding is computed in either one description and could

be extracted with the reception of either one or two descriptions. The main drawback of previous mentioned approach is that both the watermark embedding and detection are performed on side description rather than on central description. The other problem is that both QIMM and QIM are quite limited under value-metric attack.

Stimulated by the above-mentioned issues, in the second approach we attempt to find an improvement by studying the problem of watermarking under multiple description diversity transmission from a different perspective; namely, watermark embedding is done in the central description while watermark detection can be done in either central or side description. The merit of watermark embedding done in the central description is that the embedding and detection do not interfere with the MD mechanism. Therefore, this approach is more flexible than the one done in the side description. Furthermore, we propose a blind multi-rate lattice quantization index modulation (MRL-QIM) watermarking technique to boost the effectiveness. As the proposed MRL-QIM encodes two watermark bits into each of the four co-set points of a lattice (multi-rate), the payload (capacity) and robustness of watermark detection will be significantly upgraded. In the mean time, the fidelity of the watermarked image is also preserved.

In the second part of the dissertation, we study the problem of achieving watermark robustness against geometric attacks. This problem has always been a challenging research topic. We firstly propose an RST (rotation, scaling and translation) resilient image watermarking technique using Fourier-Mellin transform and informed coding of watermark message. The watermark is embedded in the geometric invariant Fourier-Mellin domain, and no additional features need to be extracted to form a geometric invariant embedding space. Moreover, by informed watermark coding, our scheme could embed a weak watermark signal (i.e. one that

needs only small perturbations with the host signal) and detect a slightly weaker watermark under the ILPM and the inverse Fourier transform. Secondly, we employ the famous QR Code (Quick Response Code) by first encoding the watermark payload, and then embedding the QR coded watermark into the image spatial domain. Thanks to the characteristic of position detection pattern of QR Code, the self-synchronized QR coded watermark payload can be recovered against geometric distortions once the QR Code is extracted during detection. Experimental results demonstrate that by adopting our approach, the resulting watermark is robust to a variety of combinations of RST attacks while preserving the visual quality of the watermarked image, thereby resolve the unavoidable dilemma faced by the other schemes.

誌 謝

感謝得自於他人的鼓勵、支持與奉獻，讓我能完成博士學業。首先，我要感謝指導教授 李素瑛老師懇切與耐心的指導，李老師對學生總是給予多方包容與體諒，讓我在學業與工作雙方面有較大的彈性空間。同時，我也要感謝共同指導教授 周義昌老師鼓勵並推薦我進修博士學位，承蒙周老師多年來在學業上的指導與工作上的提攜。感謝兩位指導教授的支持與鼓勵，讓我有機會修習並完成博士學位。

此外，我特別要感謝校內論文口試委員 廖弘源與杭學鳴兩位老師對於論文多次的幫忙。廖老師給予我論文撰寫之關鍵性指導與協助，杭老師幫我點出了信號處理領域較深入的問題，他們的協助強化了論文的品質。感謝論文口試委員張隆紋、涂元光與鄭芳炫三位老師給予我論文的寶貴意見，提供我更周詳之思考面向，他們的協助豐富了論文的內容。

感謝 王晃三老師推薦我進修博士學位，老師的身教與言教，讓我銘感於心。感謝 單懷靈博士給予我論文寫作最多的諮詢與幫忙，她的協助讓我能在有限時間內完成每一篇論文。感謝我的弟弟 敏育幫我解決論文格式編排上的所有問題。感謝中華電信研究所的多位長官，特別是劉日煬主任、張慶年主任與呂金寶經理提供在職進修機會，並讓我在工作內容與進修之研究題目間，有較大的關聯性。

感謝 父母的栽培與鼓勵，也要感謝 太太 Amy 的付出，幫忙照顧家庭，僅以此論文獻給我摯愛的家人，願大家分享這份喜悅與榮耀。要感謝的人實在太多了，最後我要感謝 上天讓我有這份機運，碰到這麼多協助我的人。

戴敏倫 謹誌

中華民國九十六年一月

CONTENTS

LIST OF TABLES

LIST OF FIGURES

List of Acronyms

APS	Analog Protection System
CSS	Content Scrambling System
DCT	Discrete Cosine Transform
DFT	Discrete Fourier Transform
FMLR	Frequency Mode Laplacian Removal
FMT	Fourier Mellin Transform
HLQ	Hexagonal Lattice Quantization
ILPM	Inverse Log-Polar Mapping
IP	Internet Protocol
IPR	Intellectual Property Right
LPM	Log-Polar Mapping
MD	Multiple Description
MDC	Multiple Description Coding
MDSQ	Multiple Description Scalar Quantizer
MDW	Multiple Description Watermarking
MRL-QIM	Multi-Rate Lattice Quantization Index Modulation
MSE	Mean Square Error
PSNR	Peak Signal Noise Ratio
QIM	Quantization Index Modulation
QIMM	Quantization Index Modulus Modulation
QR Code	Quick Response Code
RST	Rotation, Scaling and Translation
SMS	Short Message Service
SS-QIM	Spread Spectrum and Quantization Index Modulation

Chapter 1 Introduction

1.1 Overview of Watermarking (Data Hiding)

Data hiding is a science of long history, dating back as far as in ancient Greek time. Since there have always been variety of needs for privacy, people have been making continuous efforts to establish covert communication. With the rapid advance of modern technology together with the prevalence of internet and multimedia content, nowadays people can easily and securely hide messages within text, audio, image or video files and transmit them via covert communication, with only a few clicks. In addition to covert communication, along the way other commercial applications have also proliferated, some of which will most likely change our way of living and consuming. In this section we will begin with introduction and discussion on data hiding, including its core theory, current applications, and key technologies.

Data Hiding is a technique to hide data or information imperceptibly within image, audio or video so that valuable contents can be protected. These hidden messages might not be audible or visible before processing, but they can be displayed with the help of specialized computer algorithms. When applied in different fields, data hiding sometimes carries different names. For the purpose of covert communication, it is called **Steganography**, and is mainly adopted in military application. When applied in industry for copyright protection, copy protection, content authentication (identity authentication) and content linkage (annotations), it is referred as **Digital Watermarking**.

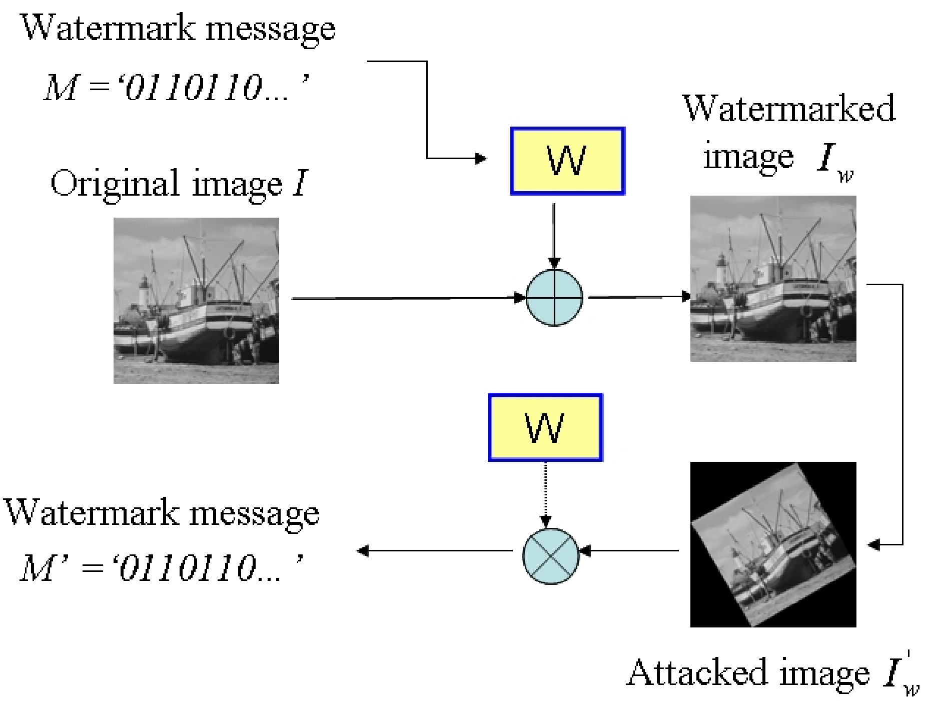


Fig. 1.1: The flow of a typical watermarking system.

Fig. 1.1 shows the flow of a typical watermarking system, which consists of generation, embedding, and extraction of watermark. Details are given below.

(1) Watermark generation (*W*): The hidden message (*M)* is represented by the sequence of 0 and 1, where each digit can be a singleton 0 or 1, combination of multiple -1 and 1, or a sequence (series) of random floating point numbers.

(2) Watermark embedding: The image content is slightly modified (i.e. modulated) based on generated watermark (*W*) that contains hidden message, but it is ensured that the perceptual quality of modified image (I_w) is almost the same as that of the original image (*I*).

(3) Watermark extraction (detection or verification): The message (*M')* is extracted directly from the detected images (I_w or I_w'), or only the

existence of hidden message is detected. For the purpose of authentication, the required watermark should be fragile. That is, the hidden message should disappear ($M' <> M$) upon attempt of tampering or forgery, or the location of modification made can be further detected, while for the purpose of copyright protection, the required watermark should be robust. That is, the hidden message should be recovered ($M' = M$) against intentional or un-intentional attacks.

Although various applications may have different requirements and performance evaluation criteria, the key requirements for watermarking are as follows:

(1) Imperceptibility (or transparency): watermark should be perceptually and statistically invisible.

(2) Robustness: watermark should be resistant to various intentional or un-intentional attacks so as to maintain acceptable visual quality.

(3) Payload (embedding bit rate or capacity): watermark payload may accommodate up to several bits.

(4) Security: watermark discovery and removal by opponent should be extremely hard.

The fulfillment of all these above-mentioned requirements is very challenging, and so far the existing algorithms in the literature are still fragile. However, the application of watermark is extensive and its market potential is quite promising. Moreover the design of watermarking algorithms implies the integration of many concepts coming from cryptography, digital communication and signal processing, therefore the development for efficient watermarking algorithms has been a very active topic for researchers in this area. The main applications are as follows:

(1) Covert communication: With a different design concept from that of traditional cryptography, the very existence of hidden message itself is secret in steganography. Therefore in theory, even an unsophisticated watermark can achieve a high degree of confidentiality. In order to eliminate the abuse of this technology, (e.g. by terrorist, by embedding secret messages in images, MP3 files, or Divx videos and communicating over public communication network,) the research of steganalysis is gaining more and more attention, with the aim of effective analyzing multimedia content over the internet so as to detect existence of hidden messages. This is apparently a rather challenging task.

(2) Copyright protection or copy protection: Many vendors in the related industries have been giving strong support to the research of mechanisms for protecting the IPR of multi-media contents. From Region-coded protection, CSS (Content Scrambling System) encryption, to APS (Analog Protection System) protection, many approaches had been brought up for solving this problem. Despite the vendors' efforts to find better solutions, so far none of them are still vulnerable. Vendors now are having high expectations of watermarking technology as an answer to data ownership protection or anti-piracy. However, due to limited robustness of existing watermark and lack of trusted agency/protocol for proving ownership of copyright, as well as the fragility of current mechanism for copy protection, for the actual commercialized application it still calls for further improvement.

(3) Content authentication (identity authentication): Because hidden watermark makes forgery more difficult, its application in anti-forgery design on ID cards/credentials becomes increasingly important. For example, the

commercial product by Digimarc [1], a leading American vendor in watermark technology, has already been adopted to assist the authentication of driver's license in the U.S. Coupled with the existing barcode and magnetic strip, it will be able to authenticate the driver's licenses better.

(4) Content linkage (annotations): This application mainly integrates the tagging and the multimedia content, and can be applied in different ways. For example, the identity of a patient and his/her sensitive information can be hidden in the medical diagnostic image, so as to avoid external tagging, which may compromise privacy or lead to mislabeling. Another example is when music is played by a mobile phone, if information concerning performer or product is embedded, it will enhance the pleasure of listening and expedite the purchase. The third example is the information retrieval through the snapshot taken by a camera phone. Currently most of 3G mobile phones manufactured in JAPAN have the capability of taking the picture of QR Code (Quick Response Code) [2] on the name-cards or catalogues, and thereafter perform specific functions, such as web-to, phone-to, SMS-to, mail and phonebook registration. If hidden watermarks can be employed to replace the added-on barcode, it will be easier to maintain the integrity of the original design. In a collaborative effort Digimarc [1] and MediaGrid [3] have successfully developed the techniques of watermark detection via camera phone, and the related applications are under extensive field trial and promotion.

There are two commonly used categories of watermarking techniques in the literature: one is spread spectrum approach, and the other is quantization approach. Cox et al. [4] proposed an image watermarking method based on spread spectrum theory, which shows good performance in terms of invisibility and robustness to signal processing operations and common geometric transformations. However, the main drawback of their approach is that both the original image and the watermark are needed in the detection process. On the other hand, the capacity of their watermark is low, since the detector can only tell whether the watermark exists or not. Therefore it is still not convincing enough for the third party to prove the rightful ownership.

Contrast to the low capacity problem inherent in the spread spectrum based watermarking techniques [4-6], the quantization based watermarking techniques [7, 8] normally have relatively high capacity. Chen and Wornell [7] presented a quantization index modulation (QIM) scheme based on the concept of dither modulation, which uses the watermark information as an index to select a dither signal. The dither signal is then added to the host signal, and a least distorted quantizer is then selected from a set of possible quantizers. The dithered host signal is quantized using this selected quantizer and finally the dither signal is subtracted from the quantized signal to form a watermarked value:

$$s(x;m) = Q(x + d(m)) - d(m), \tag{1.1}$$

where x is the host signal, $d(m)$ is the dither signal representing watermark message m (one bit of information), $Q(.)$ denotes the selected quantizer and $s(x;m)$ is corresponding to the host signal embedded with watermark message m. In the detection process, different dither signal representing watermark message is added to the received signal using Eq. (1.1), and the index (m=0 or 1) of dither signal is the extracted watermark information. The detected index $m*$ is chosen so that it gives the minimum distance between the received signal (x') and its closest quantized signal.

$$m* = \arg\min_{m} \| x' - s(x';m) \| . \tag{1.2}$$

1.2 The Considered Problems

Currently, some important problems still remain to be resolved. The embedded watermarks can be vulnerable to attacks of various natures (intentional or unintentional ones). As in the actual applications watermark are susceptible to even the unintentional ones, such as distortion from compression or packet loss during transmission over IP network, distortion from image scanning or snapshot, in this dissertation we will focus on the related problems. More detailed descriptions are given as follows.

(1) Robust watermarking against transmission channel loss: In the literature, watermarking techniques have been extensively discussed, but few of them explored watermarking for Video over IP (VoIP) or wireless transmission. In the first part of the dissertation, we study the problem of watermarking for error-prone transmission over unreliable network and two approaches are proposed. The simulation environment is a traditional two-description transmission channel. The design goal is to embed in one description a watermark, which can be detected from either one of the multiple descriptions.

(2) Robust watermarking against geometric distortion: This problem has always been a challenging research topic. However, due to its numerous potential applications many researchers have been engaged in solving this problem. In the second part of the dissertation, we study the problem of achieving watermark robustness against geometric attacks. One of the goals is to embed a robust watermark into the invariant watermarking space of an image, which is resilient to various geometric attacks. The other goal is to explore

re-synchronization or self-synchronization template to assist watermark detection under geometric distortion.

1.3 Organization of this Dissertation

The rest of this dissertation is organized as follows. In Chapter 2 and 3, we study the problem of watermarking for error-prone transmission over unreliable network. Chapter 2 introduces the approach of multiple description watermarking based on Quantization Index Modulus Modulation. We will then describe an improved approach in Chapter 3. In Chapter 4 and 5, we study the problem of achieving watermark robustness against geometric attacks. In Chapter 4, we propose an RST resilient image watermarking technique using Fourier-Mellin transform and informed coding of watermark message. In Chapter 5, we introduce the proposed self-synchronized QR-coded watermark detection. Chapter 6 offers some conclusions and directions of future work.

Chapter 2 Multiple Description Watermarking Based on Quantization Index Modulus Modulation

In this chapter, we study the problem of watermarking for error-prone transmission over unreliable network. We try to integrate an oblivious quantization index modulus modulation (QIMM) watermarking technique into the multiple description coding (MDC) framework and we call it multiple description watermarking technique (MDW). It is known that the balanced MDC encodes a signal source into multiple bitstreams (descriptions) of equal importance and equal data rate. Consider a traditional two-description case in a packet transmission network. The computation for watermark embedding is performed using either description. Once chosen, the corresponding values to be modulated for the other description are assigned with the same values as the just watermarked description. In the detection process, the embedded watermark could be extracted no matter either one or both descriptions are received. That is to say, the watermark is still detectable from MDW even with 50% packet loss. Furthermore, in the case of 50% packet loss, the resulting watermark from MDW is still robust to a variety of image processing attacks, including DCT based compression (JPEG), DWT based compression (JPEG-2000), Gaussian filtering, sharpening and median filtering. The experimental results confirmed the competitive performance and the effectiveness of the proposed scheme.

2.1 Introduction

In the literature, watermarking techniques have been extensively discussed, but few of them explored watermarking for wireless transmission. The researches that develop robust watermarking techniques which can survive intentional attacks have been extensively explored in the past decade. However, few researchers have put their emphasis on dealing with incidental attacks such as attack caused by packet network or error-prone wireless transmission over unreliable network. In [9], Hartung and Ramme pointed out that as the technology of second-generation and third-generation (3G) mobile networks keep advancing, digital media distribution for mobile E-commerce will eventually evolve into a huge business. Under these circumstances, watermarking applications such as media identification and copy control are getting more and more feasible for mobile E-commerce with the virtue that the identity of a user is known. Aiming at the error prone nature of wireless communications, Checcacci et al. [10] proposed a robust MPEG-4 watermarking technique for video sequences corrupted with errors. Graubard et al. [11] proposed a multiple description framework for oblivious watermarking, which uses one description to embed watermark information and another for referential original image to assist detection. However, it is not a completely blind technique, in the sense that a relationship (defined by watermark key) between the watermarked coefficients in one description and the corresponding coefficients in the referential one will be needed during watermark detection process. And since the embedded watermark cannot be extracted without receiving both descriptions, it is not suitable for error-prone packet transmission network applications.

Multiple description coding (MDC) [12-16] is different from either layered coding or simulcast coding or even error resilient tools described in MPEG-4 [17]. On a wireless multi-hop network or a packet-switched network, there are several parallel channels between the source and the destination and each channel may be temporarily down or suffering from long burst errors. The MDC scheme is designed such that the quality of the decoded signal is acceptable with the receiving of any individual description, and can be further improved as more descriptions are received.

In this Chapter, we propose a multiple description watermarking scheme based on oblivious quantization index modulus modulation (QIMM) watermarking technique together with the MDC framework. Consider a traditional two-description case here. The watermark is embedded in either description and can be extracted even when only one description is received. The reason of proposing the above approach is that we want to make sure a high enough watermark payload can be embedded into an images. In the meanwhile, the proposed MDW (multiple description watermarking) is robust to error-prone transmission and incidental data managing attacks. In the next section, the MDC and the proposed QIMM watermarking technique are introduced, while the proposed MDW technique is presented in Section 2.3. In Section 2.4 experimental results are presented. The concluding remarks are drawn in Section 2.5.

2.2 Multiple Description Coding (MDC) and Quantization Index Modulus Modulation (QIMM)

In this section we describe the components of proposed multiple description watermarking technique. The MDC approach [15, 16] is described first. Then we shall present the proposed QIMM watermarking technique.

2.2.1 The MDC Approach [15, 16]

The MDC-based wavelet based coding was proposed by Survetto et al. [15]. The two-description architecture of MDC [15, 16] is illustrated in Fig. 2.1. The major contribution of the MDC scheme is its capability on receiving satisfactory data quality even if part of the channels is broken. As shown in Fig. 2.1, the quality of a decoded signal is usually acceptable if either receiver 1 or receiver 2 receives the correct signal.

Furthermore, the quality of a received signal can be better if both receivers function normally. The most crucial component of an MDC scheme is its multiple description scalar quantizer. It consists of a scalar quantizer, which quantizes continuous sample values to smaller countable integers, and an index assignment counter. The source input signal $x \in X$ is first scalar quantized to $x_Q \in X_Q$. The function of the index assignment component $f : x_Q \rightarrow (x_1, x_2)$ is to split a quantized coefficient x_Q into two complementary and possibly redundant smaller coefficients $x_1 \in X_1$ and $x_2 \in X_2$, so that each of these two small coefficients only needs lower bit rate to describe and both could be recombined later to recover the original quantized coefficient. That is, with the reception of two description values, a perfect reconstructed value $\hat{x}_0 = x_Q$ can be achieved by using $\hat{x}_0 = f^{-1}(x_1, x_2)$. When only one description value is received, an acceptable estimated value $\hat{x}_d$ ($|\hat{x}_d| \prec |x_Q|$) can be obtained through $\hat{x}_d = f^{-1}(x_d)$, where $d=1, 2$.

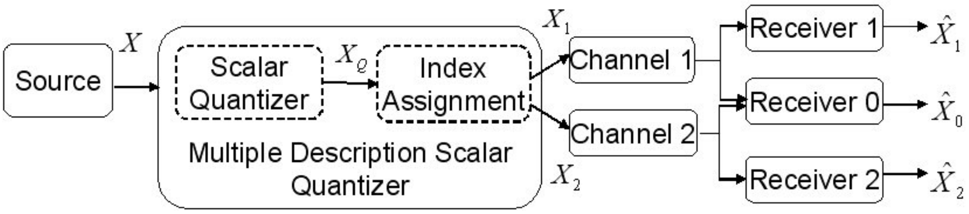


Fig. 2.1: The flowchart of multiple description coding scheme [15, 16].

To better explain the concept, we use an example to discuss the approach. A quantized coefficient x_Q valued 120 is split into the ordered pair $(x_1, x_2) = (39, 40)$, where 39 and 40 are the values assigned to description 1 and 2, respectively. On receiving two descriptions, a perfect recovered value $\hat{x} = \hat{x}_0 = x_Q = 120$ can be achieved by central decoder. When receiving only description 1 for the transmitted value 120, the estimated $\hat{x} = \hat{x}_1$ using $x_1 = 39$ will be 118, while receiving only description 2, the estimated $\hat{x} = \hat{x}_2$ using $x_2 = 40$ will be 121. As can be seen from this example, the central decoder should be more robust against various attacks than the side decoder since the reconstructed value received from central decoder is the same as the watermarked value before transmission. The detailed algorithm for index assignment can be found in [15, 16].

2.2.2 QIMM

In this section, we shall describe in detail how the proposed oblivious quantization index modulus modulation scheme functions. The proposed QIMM approach selects some of wavelet coefficients as the original host signal. Then the index of each quantized coefficient is modulated for embedding one bit of information. The embedding and detection processes are described as follows.

2.2.2.1 The Embedding Process of QIMM

The original host signal $X = \{x_1, x_2, ..., x_n\}$ is first divided by the quantization step size (δ), and a nearest integer index value is obtained by a round function. The quantized index value is then executed with modulo 2 to get the residue with value 0 or 1. If the residue is equal to the watermark message bit, then the watermarked value is the reconstruction point of quantized host signal. Otherwise, the biased (either +1 or −1) quantized index value is used to calculate the watermark value $X^{'} = \{x_1^{'}, x_2^{'}, ..., x_n^{'}\}$. To embed one bit of watermark message *m*, the embedding algorithm consists of the following steps:

Step 1: Take $Q(x_i) = Round(x_i / \delta)$.

Step 2: If $(Q(x_i) \bmod 2) = m$ then

$$x_i^{'} = s(x_i; m) = Q(x_i) * \delta, \tag{2.1}$$

else

$$x_i^{'} = s(x_i; m) = \arg\min_{P(x_i)} \left\| P(x_i) - x_i \right\|, \qquad (2.2)$$

where $P(x_i)$ in Eq. (2.2) is either $(Q(x_i)-1)*\delta,$ or $(Q(x_i)+1)*\delta,$ and $s(x_i;m)$ is the i_{th} host signal embedded with watermark message m. The criterion of selecting either $(Q(x_i)-1)*\delta$ or $(Q(x_i)+1)*\delta$ depends on which one has less distortion with respect to x_i. The one with less distortion is used to reconstruct the watermarked signal $x_i^{'}$. The difference between QIM and QIMM is compared and elaborated as follows.

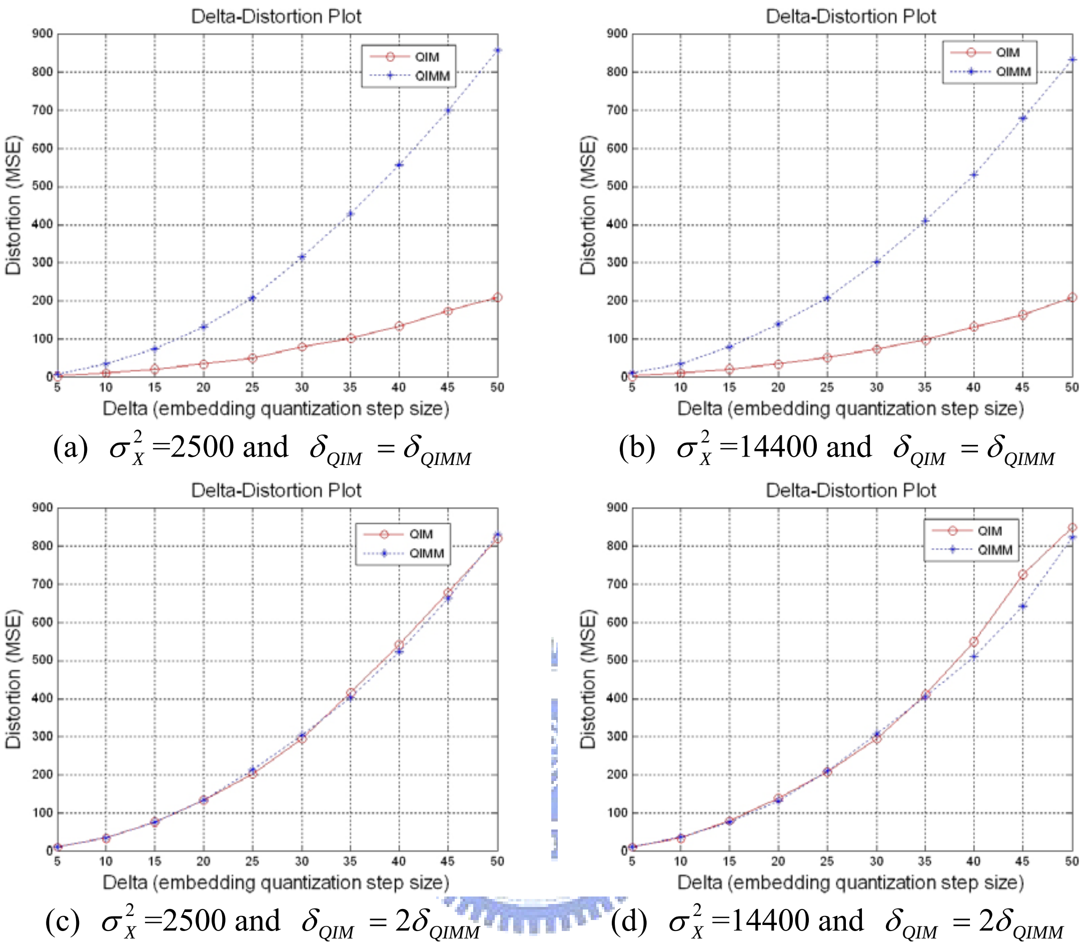
We observe that low embedding distortion (q) leads to low degree of robustness. For QIM embedding with quantization step size δ_{QIM}, the embedding distortion (q) range is $[-\frac{\delta_{QIM}}{2}, \frac{\delta_{QIM}}{2}]$ and the detection robustness range is $[-\frac{\delta_{QIM}}{2}, \frac{\delta_{QIM}}{2}]$ too. If the host signal X is uniform, the mean squared error distortion (MSE) of embedding is the second moment of a random variable uniformly distributed in the interval $[-\frac{\delta_{QIM}}{2}, \frac{\delta_{QIM}}{2}]$:

$$MSE_{QIM} = \frac{1}{\delta_{QIM}} \int_{-\frac{\delta_{QIM}}{2}}^{\frac{\delta_{QIM}}{2}} q^2 dq = \frac{\delta_{QIM}^2}{12} \qquad (2.3)$$

As for QIMM embedding with quantization step size δ_{QIMM}, the embedding distortion (q) range is $[-\delta_{QIMM}, \delta_{QIMM}]$ and the detection robustness range is $[-\delta_{QIMM}, \delta_{QIMM}]$ too. The mean squared distortion (MSE) of embedding is:

$$MSE_{QIMM} = \frac{1}{2\delta_{QIMM}} \int_{-\delta_{QIMM}}^{\delta_{QIMM}} q^2 dq = \frac{\delta_{QIMM}^2}{3} \qquad (2.4)$$

It is expected that by setting $\delta_{QIM} = 2\delta_{QIMM}$, QIM and QIMM should obtain similar embedding distortion and detection robustness.



(a) σ_X^2 =2500 and $\delta_{QIM} = \delta_{QIMM}$ (b) σ_X^2 =14400 and $\delta_{QIM} = \delta_{QIMM}$

(c) σ_X^2 =2500 and $\delta_{QIM} = 2\delta_{QIMM}$ (d) σ_X^2 =14400 and $\delta_{QIM} = 2\delta_{QIMM}$

Fig. 2.2: The Delta-Distortion curve of QIM and QIMM.

To better illustrate the Delta-Distortion relationship of QIM and QIMM, we performed Monte Carlo simulations with host signal X drawn from 1000 samples of a Gaussian zero-mean random variable with variance σ_X^2 ranging from 2500 to 14400. Fig. 2.2(a) and 2.2(b) both show the MSE distortion under various embedding quantization step sizes ranging from 5 to 50 for QIM and QIMM, while Fig. 2.2(c) and 2.2(d) show the MSE distortion under various embedding quantization step sizes ranging from 5 to 50 for QIMM and 10 to 100 for QIM, respectively. As we can see from Fig. 2.2(c) and 2.2(d), to get the same distortion for QIM and QIMM, the embedding quantization step size δ_{QIM} of QIM is almost equal to two times of δ_{QIMM} of QIMM.

2.2.2.2 The Detection Process of QIMM

After receiving the watermarked signal *X'*, the attacked watermarked signal *X"* is also divided by the quantization step size, so that a nearest integer index value is obtained by a round function. The quantized index value is then taken modulo 2 to get the extracted watermark message bit *m**. The detection algorithm consists of the following steps:

Step 1: $Q(x_i^{"}) = Round(x_i^{"} / \delta)$.

Step 2: $m^* = Q(x_i^{"})$ mod 2.

In Section 2.2.1, we have shown that by setting $\delta_{QIM} = 2\delta_{QIMM}$, QIM and QIMM have obtained similar embedding distortion. In this Section, we demonstrate that by setting $\delta_{QIM} = 2\delta_{QIMM}$, QIM and QIMM obtain the similar detection robustness as follows. To evaluate the reliability (robustness) of watermark detection, the correlation ratio ρ was defined as:

$$\rho = \frac{Total\ \ number\ \ of\ \ correctly\ \ detected\ \ bits}{Total\ \ number\ \ of\ \ embedded\ \ bits}, \tag{2.5}$$

A higher value of ρ indicated a more reliable detection. The perfect recognition rate can be achieved when the value of ρ equals 1.

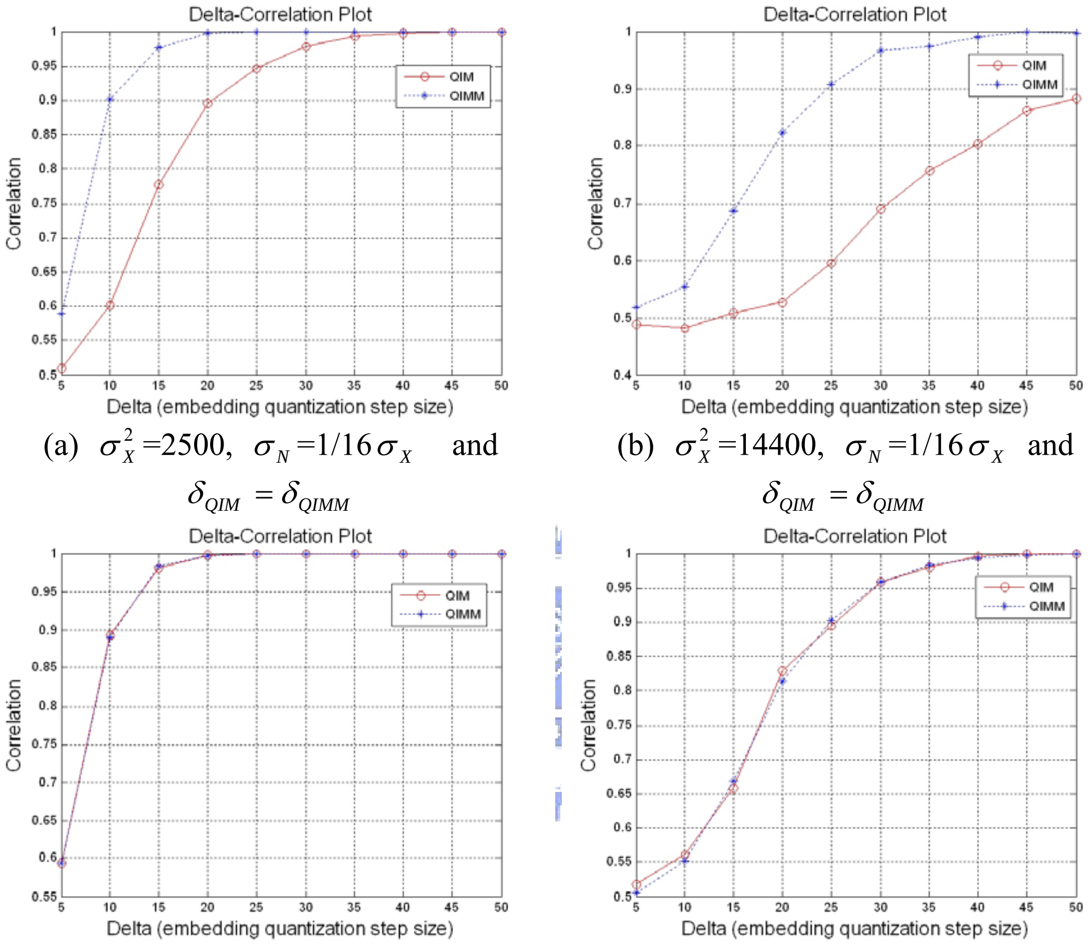
Following the same scenario as in Section 2.2.1, we performed Monte Carlo simulations with host signal X drawn from 1000 samples of a Gaussian zero-mean random variable with variance σ_X^2 ranging from 2500 to 14400. Moreover, a noise signal N drawn from 1000 samples of a Gaussian zero-mean random variable with standard deviation $\sigma_N = \frac{1}{16}\sigma_X$ is employed to simulate the various attacks.

Each sample of signal X was used to embed one bit of watermark information under various embedding quantization step sizes, where totally 1000 bits were embedded for each specific quantization step size. The watermarked signal X' is attacked with noise signal N via $X^{''} = X^{'} + N$ (a similar results can be obtained via $X^{''} = X^{'} - N$) before detection.

As can be seen from Fig. 2.3(a)~(d), smaller embedding quantization step sizes leads to lower degree of robustness for both QIM and QIMM. Fig. 2.3(a) and 2.3(b) both show the correlation ratio under various embedding quantization step sizes ranging from 5 to 50 for QIM and QIMM, while Fig. 2.3(c) and 3(d) show the correlation ratio under various embedding quantization step sizes ranging from 5 to 50 for QIMM and 10 to 100 for QIM, respectively. As we can see from Fig. 2.2(a), 2.2(b), 2.3(a) and 2.3(b), though the MSE of QIM is lower than that of QIMM, the robustness of QIM is inferior to that of QIMM. In contrast, as seen from Fig. 2.2(c), 2.2(d), 2.3(c) and 2.3(d), under the same MSE condition, the robustness of QIM is almost equal to that of QIMM.

As the QIM scheme [7] has been proven to be nearly optimal with respect to the tradeoff among embedding distortion, detection robustness and hiding capacity, we do not expect that QIMM can outperform QIM in the scalar-based case. Rather, we intend to explore this topic from different perspective. Since for watermark embedding based on scalar quantization, focus can not be put solely on distortion introduced by embedding, as the accompanied robustness should also be taken into consideration. Robustness should be compared on the ground of the same distortion.

Furthermore, by understanding QIMM as generalized LSB with delta value larger than 2, the concept is better grasped and more accessible to most readers, and leads to less implementation effort than that of the dithering concept of QIM.



(a) σ_X^2=2500, σ_N=1/16σ_X and $\delta_{QIM} = \delta_{QIMM}$

(b) σ_X^2=14400, σ_N=1/16σ_X and $\delta_{QIM} = \delta_{QIMM}$

(c) σ_X^2=2500, σ_N=1/16σ_X and $\delta_{QIM} = 2\delta_{QIMM}$

(d) σ_X^2=14400, σ_N=1/16σ_X and $\delta_{QIM} = 2\delta_{QIMM}$

Fig. 2.3: The Delta-Correlation curve of QIM and QIMM.

2.3 The Proposed Multiple Description Watermarking (MDW) Scheme

In this section, a multiple description watermarking technique using both MDC and QIMM is described. The design goal of the MDW scheme is to embed in one description a watermark, which can be detected from either one of the multiple descriptions. The advantage of the proposed scheme is two-fold. First, it can increase the detection robustness for error-prone transmission over unreliable network. Second, it is able to increase the capacity while preserving the transparency. This is achieved by modulating the selected coefficients of either description appropriately so that one bit of information can be embedded.

Fig. 2.4 shows the flow of MDW, which is composed of a watermark embedding process and a transmission process. The original image is first transformed into the discrete wavelet domain. The transformed coefficients are then processed by multiple description scalar quantizer (MDSQ) to generate two independent descriptions, X_1 and X_2. Next, a bit (*m*) of the watermark message *M* is embedded in some of the selected coefficients from one of the descriptions using QIMM. During the watermark embedding process, whenever each of the selected coefficients is modulated by the watermarking embedding rule, the corresponding coefficient of the other un-watermarked description (say description 2) is also replaced with the same value. Each of the watermarked bitstream is then sent through one independent channel. The watermarked images (I'_{w1} from side decoder 1, I'_{w2} from side decoder 2 or I'_{w0} from central decoder) could then be obtained by receiving either one description (receiver 1 ($\hat{X}_1$) or receiver 2 ($\hat{X}_2$)) or two descriptions (receiver 0 ($\hat{X}_0$)) and inversing the Discrete Wavelet transforms. In the detection process in Fig. 2.5, the attacked image I''_{wr} (r = 0, 1 or 2) first goes through Discrete Wavelet transform, and some of the selected coefficients are then used to extract the embedded watermark message *M**.

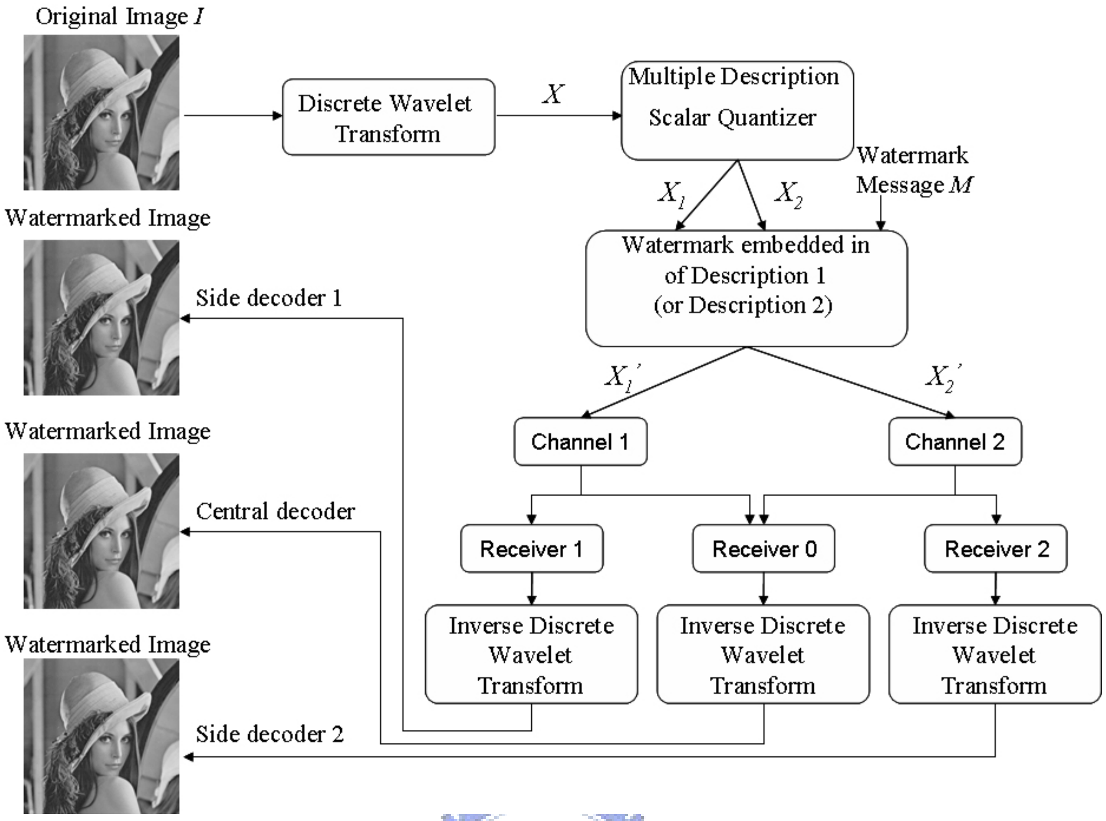

Original Image I
Discrete Wavelet Transform
X
Multiple Description Scalar Quantizer
X_1
X_2
Watermark Message M
Watermark embedded in of Description 1 (or Description 2)
X_1'
X_2'
Channel 1
Channel 2
Receiver 1
Receiver 0
Receiver 2
Inverse Discrete Wavelet Transform
Inverse Discrete Wavelet Transform
Inverse Discrete Wavelet Transform
Watermarked Image
Side decoder 1
Watermarked Image
Central decoder
Watermarked Image
Side decoder 2

Fig. 2.4: The flow of proposed multiple description watermark embedding scheme for error-prone transmission over unreliable network.

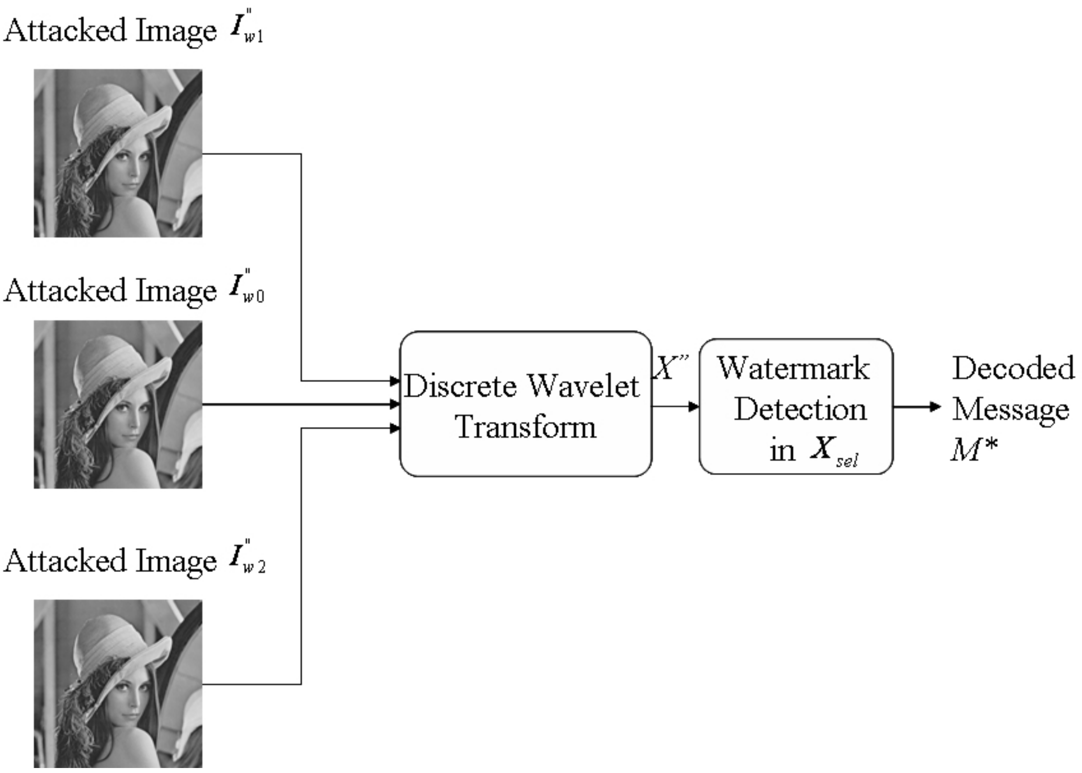

Attacked Image I''_{w1}
Attacked Image I''_{w0}
Attacked Image I''_{w2}
Discrete Wavelet Transform
X''
Watermark Detection in X_{sel}
Decoded Message M^*

Fig. 2.5: The flow of proposed multiple description watermark detection scheme.

The proposed scheme is completely different from that of [11], where two-description design is adopted as well. In contrast to [11], the value pairs of these two descriptions in our scheme are almost with the same value. When the watermark embedding process is executed, one only needs to consider one description. Whenever a coefficient of one description is modulated using watermark embedding rules, the corresponding coefficient of the other description is set to the same value. The time complexity is reduced because watermarking one description implies watermarking another description at the same time. The good characteristics of our proposed MDW results from the design nature of the index assignment function. Moreover, the MDW can detect watermark no matter either one or two descriptions are received. This means in an error-prone packet transmission network, the watermark can still be detected even with 50% packet loss rate.

2.3.1 Embedding and transmission process of MDW

To embed *n* bits of watermark message *M* into image *I*, the algorithm is described as follows:

Step 1: The original image *I* is decomposed into 13-subbands using the 4-level octave band wavelet transform.

Step 2: Each of the subband coefficients are quantized by a uniform scalar quantizer.

Step 3: Two descriptions (X_1, X_2) of the quantized coefficient are created by mapping each quantized coefficient to a pair of numbers by the index assignment component.

Step 4: Select the coefficients on LL band of description 1 (or 2) for watermark embedding, namely $X_{sel} = \{x_1, x_2, ..., x_n\}$.

Step 5: Apply embedding process of QIMM on X_{sel} to embed watermark message *M*.

Step 6: Replace the corresponding coefficients of un-watermarked description 2 (or 1) with the same values as those embedded in description 1 (or 2).

Step 7: Transmit these two watermarked descriptions over network via two different channels.

Step 8: Apply inverse transform to obtain watermarked image I'_{wr} ($r = 0, 1$ or 2) depending on received descriptions $\widehat{X}_r$ ($r = 0, 1$ or 2).

2.3.2 Detection process of MDW

The received watermarked image I'_{wr} ($r = 0, 1$ or 2) could be attacked by intentional or unintentional modifications, leading to attacked image I''_{wr} ($r = 0, 1$ or 2). To extract *n* bits of watermark message *M** from an attacked image, the algorithm is described as follows:

Step 1: The attacked image I''_{wr} is decomposed into 13-subbands using the 4-level octave band wavelet transform.

Step 2: Each of the subband coefficients are quantized by a uniform scalar quantizer.

Step 3: Select the coefficients on LL band of the attacked image for watermark extraction, namely $X''_{sel} = \{x''_1, x''_2, ..., x''_n\}$.

Step 4: Apply detection process of QIMM on X''_{sel} to obtain the extracted watermark message *M**.

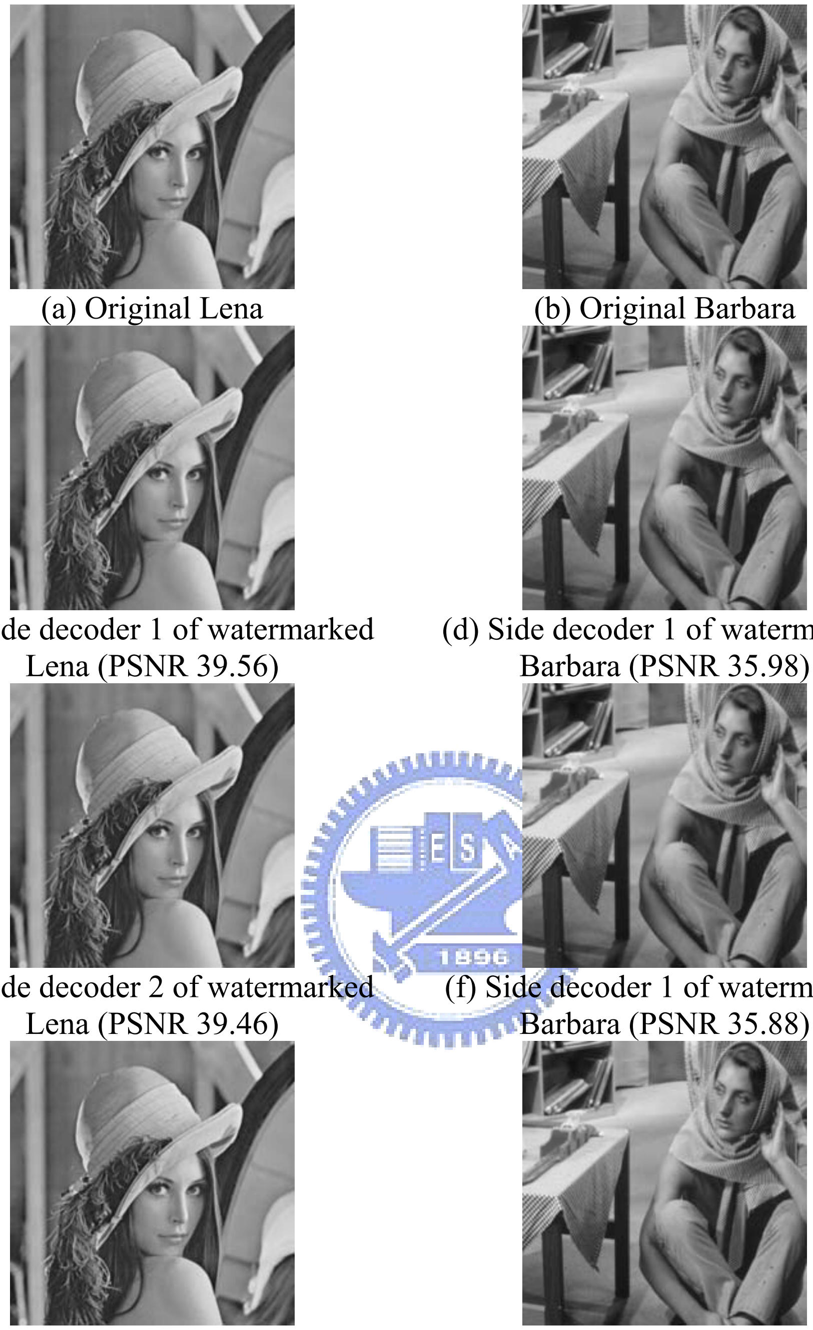
2.4 Experimental Results

To evaluate the effectiveness of the proposed method, one transformed coefficient was used to embed one bit of watermark information, and totally 128 coefficients were used to embed 128 bits of watermark information. Several standard images including “Lena”, “Barbara”, “House” and “Boat” were tested and demonstrated similar performance. To save space, only “Lena” (Fig. 2.6(a)) and “Barbara” (Fig. 2.6(b)) are given here. In order to show the flexibility of our proposed MDW framework and to make comparison with our proposed watermarking technique (QIMM), another state-of-the-art watermark technique QIM [7] detailed in Section 1.1.3 was integrated into the MDW framework with QIMM replaced.

When talking about compression, larger quantization step size will lead to larger distortion MSE (mean square error), meaning smaller PSNR, and hence smaller bit rates is needed. However, when comparing two watermark algorithms, we follow the common practice by fixing two requirements, namely watermark capacity and the transparency (distortion) of watermarked image, and then comparing the robustness. For a fair comparison, the parameter that defined the quantization step was adjusted so that similar PSNR values (in other words, similar distortion) and bit rates could be obtained. The PSNRs of watermarked and un-watermarked “Lena” and “Barbara” for side decoder 1, side decoder 2 and central decoder are illustrated, respectively, in Table 2.1. From our experiments, the degree of PSNR dropped when the quantization step size was increased. A larger quantization step size brought more robustness, but it also introduced more distortion. According to the theoretical and experimental analysis on both QIMM and QIM as well as comparison of their properties in the aspect of embedding distortion and detection robustness as described in Section 2.2, δ_{QIM} was set to 64 and δ_{QIMM} was set to 32 in our setting. The recovered watermarked images by side decoder 1, side decoder 2 and central decoder for QIMM are shown in Fig. 2.6(c), 2.6(e) and 2.6(g) for Lena, respectively, and similarly, in Fig. 2.6(d), 2.6(f) and 2.6(h) for Barbara, respectively. The quality of the pictures recovered from the side decoders was inferior to that recovered from the central decoder, yet still acceptable.

Table 2.1: The PSNRs of un-watermarked and watermarked "Lena" and "Barbara" for proposed QIMM and Chen's QIM

Method	PSNR(dB)					
	Lena			Barbara		
	Side 1	Side 2	Central	Side 1	Side 2	Central
Un-watermark	40.94	40.96	49.46	37.53	37.55	47.11
Our QIMM	39.56	39.46	43.62	35.98	35.88	40.04
Chen's QIM	39.71	39.61	44.05	36.09	35.99	40.31



(a) Original Lena

(b) Original Barbara

(c) Side decoder 1 of watermarked Lena (PSNR 39.56)

(d) Side decoder 1 of watermarked Barbara (PSNR 35.98)

(e) Side decoder 2 of watermarked Lena (PSNR 39.46)

(f) Side decoder 1 of watermarked Barbara (PSNR 35.88)

(g) Central decoder of watermarked Lena (PSNR 43.62)

(h) Central decoder of watermarked Barbara (PSNR 40.04)

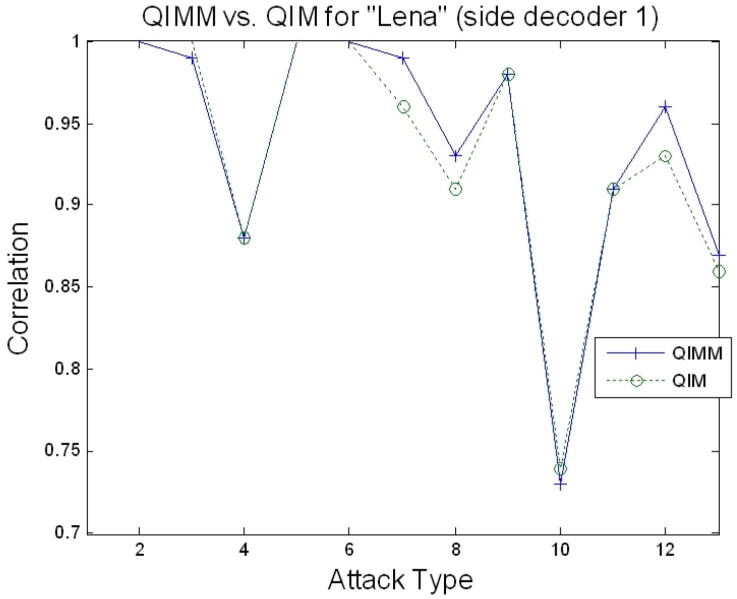
Fig. 2.6: Original and watermarked Lenas and Barbaras.

In addition to the degree of robustness against packet loss, a desirable and fundamental property for a watermarking algorithm is to survive compression attack. In the real-world applications, compression is frequently used to facilitate efficient storage and transmission. Here, we used images compressed by JPEG (low quality factor ranging from 10 to 25) and JPEG-2000 (low bit-rates ranging from 0.125 bpp to 1.0 bpp) to test our algorithm. Moreover, a variety of signal manipulation attacks such as Gaussian filtering, sharpening and median filtering were also introduced to check the feasibility of our approach. Among these attacks, we used JPEG-2000 VM8.0 to compress target images and adopted Stirmark3.1 [18] to manipulate the other attacks. Totally 13 attack types as listed in Table. 2.2 were used in these experiments. Under one description loss combined with each of the 13 attack types, these two methods still have good performance in the MDW framework.

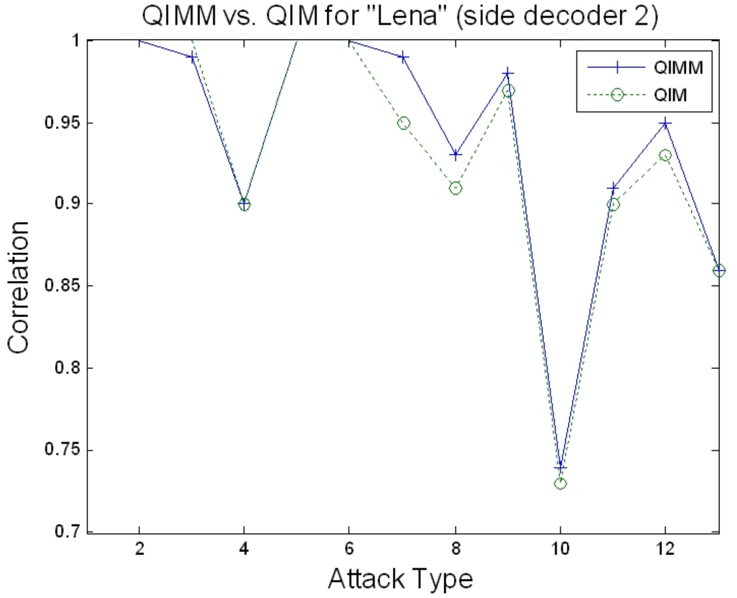
Table 2.2: The tested attack types

Attack Types	
1	JPEG-2000 1.000 bpp
2	JPEG-2000 0.500 bpp
3	JPEG-2000 0.250 bpp
4	JPEG-2000 0.125 bpp
5	JPEG Quality factor Q(%) = 25
6	JPEG Quality factor Q(%) = 20
7	JPEG Quality factor Q(%) = 15
8	JPEG Quality factor Q(%) = 10
9	Gaussian filtering 3x3
10	Sharpening 3x3
11	2x2 Median filtering
12	3x3 Median filtering
13	4x4 Median filtering

The detected correlations ratio from the “Lena” and “Barbara” images against the combined attacks are summarized in Fig. 2.7 and Fig. 2.8, respectively. For some types the detection rate maintains 100% and for other types it degrades. For the “Lena” image, except for the number 10 attack (sharpening 3x3), the correlation ratio ρ were all above 0.85. As to the “Barbara” image, except for the number 13 attack (4x4 median filtering), the correlation ratio ρ were all above 0.7. It is noted that the primary aim of this chapter was to propose a watermarking scheme resilient to packet loss over unreliable network. Therefore, by embedding one bit of information, our scheme uses only one coefficient, and the robustness to these further attacks even with one description loss is an added bonus. It goes without saying, more elaborated schemes which use more coefficients (say one 8x8 block) to embed one bit of information should further improve the detector’s performance. Though this issue is not treated here, it is obvious that our scheme applies in this extension as well.

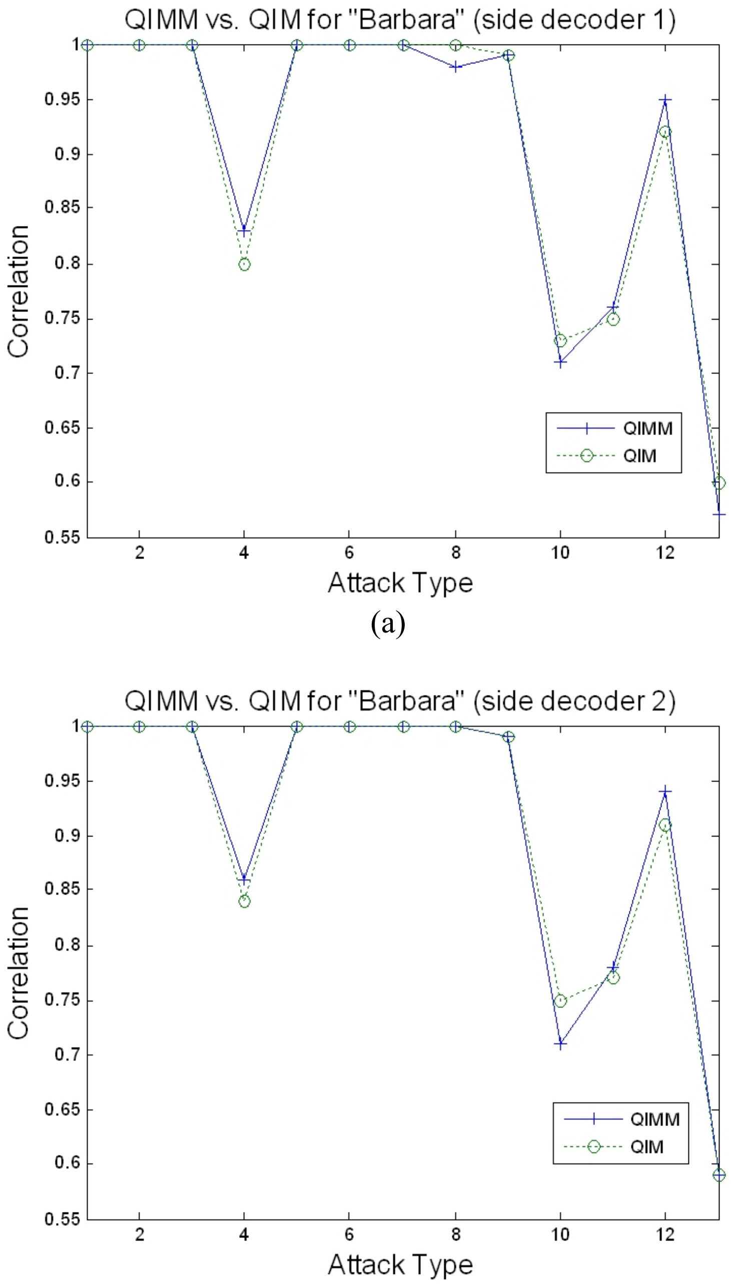


(a)



(b)

Fig. 2.7: The comparison between QIMM and QIM in terms of correlation ratio. (a) QIMM vs. QIM for “Lena” (side decoder 1), (b) QIMM vs. QIM for “Lena” (side decoder 2).



(a)

(b)

Fig. 2.8: The comparison between QIMM and QIM in terms of correlation ratio. (a) QIMM vs. QIM for “Barbara” (side decoder 1), (b) QIMM vs. QIM for “Barbara” (side decoder 2).

2.5 Summary

In this chapter, the theoretical and experimental analysis on both QIMM and QIM are demonstrated. The comparison of their properties in the aspects of embedding distortion and detection robustness is explored. It is verified that by setting $\delta_{QIM} = 2\delta_{QIMM}$, QIM and QIMM obtained similar embedding distortion, as shown by Delta-Distortion curve, and they are competitive in detection robustness, as shown by Delta-Correlation curve. Furthermore, we propose a multiple description watermarking technique which integrates an oblivious QIMM with the MDC framework. The watermark embedding is computed in either description and could be extracted with the reception of either one or two descriptions. Another advantage of our scheme worth mentioning here is the flexibility of the MDW framework. It can be integrated easily with most current watermarking schemes. This flexibility property is demonstrated in our experiments (see Fig. 2.7 and Fig. 2.8), where MDW is integrated with QIM and QIMM, respectively. It is evident that, both of these two methods performed well in our MDW framework. In addition to resilience to packet loss, the performance tradeoff between invisibility and robustness to various attacks shows the usefulness of this proposed approach. In the future, we expect that other MDC approach [12-14] or some error resilient algorithms [17] could be integrated with the more elaborated watermarking schemes. Moreover, as the distortion introduced by losing some of the transmitted descriptions of MD transmission can be viewed as a non-linear value-metric attack [19-22], research of an adaptive hexagonal lattice-based QIM which is more robust to value-metric attack will be investigated in the next chapter.

Chapter 3 Robust MRL-QIM Watermarking Resilient to Multiple Description Transmission Channel

Multiple description (MD) transmission results in non-linear value-metric (value-scaling) distortion in the case when some of the sent descriptions are not received. An acceptable image can still be received in the above situation due to the characteristic of MD. However, it is quite damaging to the traditional quantization based watermarking technique for payload detection. To overcome the problem, a straight forward approach would be increasing the quantization step size in watermark embedding so as to keep distortion maintaining within the tolerant range. However, as a larger quantization step size in watermark embedding would result in worsened watermarked image, it is not feasible to adopt it without further consideration. The proposed MRL-QIM (Multi-Rate Lattice Quantization Index Modulation) encodes two watermark bits into each of the four co-set points of a lattice (so called multi-rate). Compared to traditional vector-based quantization encoding or combined spread spectrum-quantization encoding, it significantly increases the payload (capacity) and enhances the robustness of watermark detection while preserving the fidelity of watermarked image. Comprehensive experiments have confirmed that the overall performance and the effectiveness of the proposed scheme are superior to previous approaches.

3.1 Introduction

We argue that the distortion introduced by losing some of the transmitted descriptions of MD transmission can be viewed as a non-linear value-metric attack (described in Section 3.2.1). While in this situation an acceptable image can still be received due to the characteristic of MD, it is quite damaging to the traditional quantization based watermarking technique (QIM) for payload detection. Some other works address the problem of QIM detection in the presence of value-metric distortion [8, 19-23], yet few papers focus on QIM related watermarking scheme under MD attack.

In Chapter 2, we integrate oblivious watermarking techniques (quantization index modulus modulation (QIMM) and QIM) with the multiple description coding (MDC) to get a multiple description watermarking (MDW) framework. In this framework, the watermark embedding is computed in either description and could be extracted with the reception of either one or two descriptions. The main drawback of the approach is that both the watermark embedding and detection are performed on side description rather than on central description. The other problem is that both QIMM and QIM are quite limited under value-metric attack.

Stimulated by the above-mentioned issues, in this chapter we attempt to find an improvement by studying the problem of watermarking under multiple description diversity transmission from a different perspective; namely, watermark embedding is done in the central description while watermark detection can be done in either central or side description. The merit of watermark embedding done in the central description is that the embedding and detection do not interfere with the MD mechanism. Therefore, this approach is more flexible than the one done in the side description. Furthermore, we propose a blind multi-rate lattice quantization index modulation (MRL-QIM) watermarking technique to boost the effectiveness. As the proposed MRL-QIM encodes two watermark bits into each of the four co-set points of a lattice (multi-rate), with the above design, the payload (capacity) and robustness of watermark detection will be significantly upgraded. In the mean time, the fidelity of the watermarked image is also preserved. In next section, the MD attack channel and the hexagonal lattice quantization (HLQ) algorithm will be introduced. The proposed MRL-QIM watermarking technique will be elaborated in Section 3.3. In Section 3.4 experimental results are presented. The concluding remarks together with future work will be addressed in Section 3.5.

3.2 Multiple Description (MD) Attack Channel and Hexagonal Lattice Quantization (HLQ)

In this section we begin with describing the multiple description attack channel, which introduces possible non-linear value-metric attack for watermark detection through transmission. Then, we shall introduce a hexagonal lattice quantization (HLQ) algorithm which will be adopted in our proposed MRL-QIM watermarking technique.

3.2.1 The MD Attack Channel [15, 16]

The basic two-description architecture of MDC [15, 16] is described in Section 2.2.1. To better explain why the distortion introduced by losing some of the transmitted descriptions of MD transmission can be viewed as a non-linear value-metric attack, we use an example to discuss the approach. A quantized coefficient x_Q valued 392 is split into the ordered pair $(x_1, x_2) = (130, 131)$, where 130 and 131 are the values assigned to description 1 and 2, respectively. Similarly, a quantized coefficient x_Q valued 813 is split into the ordered pair $(x_1, x_2) = (271, 270)$, where 271 and 270 are the values assigned to description 1 and 2, respectively. On receiving only description 1 for the transmitted value 392, the estimated $\hat{x} = \hat{x}_1$ using $x_1 = 130$ will be 391; while receiving only description 2, the estimated $\hat{x} = \hat{x}_2$ using $x_2 = 131$ will be 394. Similarly, on receiving only description 1 for the transmitted value 813, the estimated

$\hat{x} = \hat{x}_1$ using $x_1 = 271$ will be 814; while receiving only description 2, the estimated $\hat{x} = \hat{x}_2$ using $x_2 = 270$ will be 811. As can be seen from this example, this leads to non-linear value-metric distortion. The detailed algorithm for index assignment can be found in [15, 16].

3.2.2 HLQ (Hexagonal Lattice Quantization)

It is well known that quantization error includes overload error and granular error [24]. The overload error can be reduced by vector quantization and the granular error is affected by the size and shape of a quantization region. Since the best shape for a quantization region in two dimensions is a hexagon [25], we adopt an efficient lattice quantization algorithm [26] for hexagonal lattice quantizer. In order to make this algorithm better fit in our watermarking scheme, we employ the concept from nested lattice [27] to implement the software.

3.2.2.1 Lattice Quantization

An *N*-dimensional lattice is a set of points $\Lambda = \{\mathbf{x}\} = \{u_1 a_1 + ... + u_N a_N\}$, where $\mathbf{x}$ is an *N*-dimensional row vector (point) in R^N, $\{a_1,...,a_N\}$ is a set of basis vectors in R^N, and $u_1,...,u_N$ range through all integers. That is,

$$\mathbf{x} = \mathbf{u}\mathbf{A}, \text{where } \mathbf{x} = [x_1, x_2, ..., x_N],\ \mathbf{u} = [u_1, u_2, ..., u_N] \text{ and } \mathbf{A} = \begin{bmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_N \end{bmatrix}.$$

The Voronoi region or the nearest neighbor region for the lattice Λ with respect to $\mathbf{x}'$ is defined as:

$$V(\mathbf{x}') = \{\mathbf{x} \in R^N : \|\mathbf{x}' - \mathbf{x}\| \le \|\mathbf{y} - \mathbf{x}\|, \textit{for all } \mathbf{y} \in \wedge\}.$$

Let Q be the quantization function mapping $\mathbf{x} \in R^n$ to the nearest point $\mathbf{x}'$ in Λ. To quantize $\mathbf{x}$ to $\mathbf{x}'$, we need to find nearest lattice points $\mathbf{x}'$ to $\mathbf{x}$, that is, $Q(\mathbf{x}) = \mathbf{x}'$. Some of the fast algorithms for lattice quantizer have been proposed in the literature [26-28]. We adopt the method proposed in [26] for its simplicity and then make some minor modifications on its format. This minor change makes the lattice quantization feasible for watermark embedding / detection. Our modified quantizer algorithm for two-dimensional hexagonal lattice A_2 in R^2 can be implemented as follows:

Step 1: $\mathbf{x} = \mathbf{uA}$.

Step 2: $\mathbf{u}^* = \mathbf{xA}^{-1}$ (as $\mathbf{u}^*$ might not be integers, $\mathbf{x}^* = \mathbf{u}^*\mathbf{A}$ might not be a lattice point, but a close point to a lattice point $\mathbf{x}'$ defined below).

Step 3: $\mathbf{u}^* = (u_1, u_2)$.

Step 4: Round $\mathbf{u}^*$ to integer point $(u_1, u_2) = (round(u_1), round(u_2))$.

Step 5: Find seven neighbor integer pairs for $\mathbf{u}^*$.

$$\begin{aligned}
&\mathbf{u}(1) = (u_1, u_2),\\
&\mathbf{u}(2) = (u_1 + 1, u_2),\\
&\mathbf{u}(3) = (u_1, u_2 + 1),\\
&\mathbf{u}(4) = (u_1 + 1, u_2 + 1),\\
&\mathbf{u}(5) = (u_1 - 1, u_2),\\
&\mathbf{u}(6) = (u_1, u_2 - 1),\\
&\mathbf{u}(7) = (u_1 - 1, u_2 - 1).
\end{aligned}$$

Step 6: Choose i^*, where i^* is the index i that gives the minimum $\mathbf{u}(i)\mathbf{A} - \mathbf{x}$.

$$i^* = \arg\min_i \left\| \mathbf{u}(i)\mathbf{A} - \mathbf{x} \right\|.$$

Step 7: Get the closest lattice point $\mathbf{x}'$ to $\mathbf{x}$.

$$\mathbf{x}' = \mathbf{u}(i^*)\mathbf{A}.$$

After the process of lattice quantization, the next step is to further partition the constructed lattice into co-sets. With the nested lattice structure, one is able to correctly deal with the payload issue.

3.2.2.2 Nested Lattice

For the purpose of watermarking, the constructed lattice should be further partitioned into several co-sets, where points belonging to different co-set represent different watermark payload. Fig. 3.1(a) and 3.1(b) depict the concept of nested lattice pair (Λ_f, Λ_c) [29], where Λ_f is fine lattice and Λ_c is course lattice. The generating matrix $\mathbf{A_f}$ of Λ_f and $\mathbf{A_c}$ of Λ_c are related by

$$\mathbf{A_c} = \mathbf{J}\mathbf{A_f}.$$

Lattice Λ_f may be decomposed into $|\det \mathbf{J}|$ co-sets, and Λ_f is the union of co-sets, Λ_f and Λ_c are related by

$$\Lambda_f = \bigcup_{k=0}^{|\det \mathbf{J}|-1} \Lambda_k.$$

For example, let $\mathbf{A_f} = \begin{bmatrix} 1 & 0 \\ -1/2 & \sqrt{3}/2 \end{bmatrix}$ and $\mathbf{J} = 2\mathbf{I}_2 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, by mapping $(\xi_1, \xi_2) \in Z^2$ to $(\xi_1 - \frac{1}{2}\xi_2, \frac{\sqrt{3}}{2}\xi_2) \in A_2$, we can therefore use $(\xi_1, \xi_2) \in \{(0,0), (0,1), (1,0), (1,1)\}$ to generate 4 co-set leaders:

$$C = \{\mathbf{d}_0, \mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3\} = \{(\xi_1 - \frac{1}{2}\xi_2, \frac{\sqrt{3}}{2}\xi_2)\} =$$
$$\{(0,0), (\frac{-1}{2}, \frac{\sqrt{3}}{2}), (1,0), (\frac{1}{2}, \frac{\sqrt{3}}{2})\}.$$

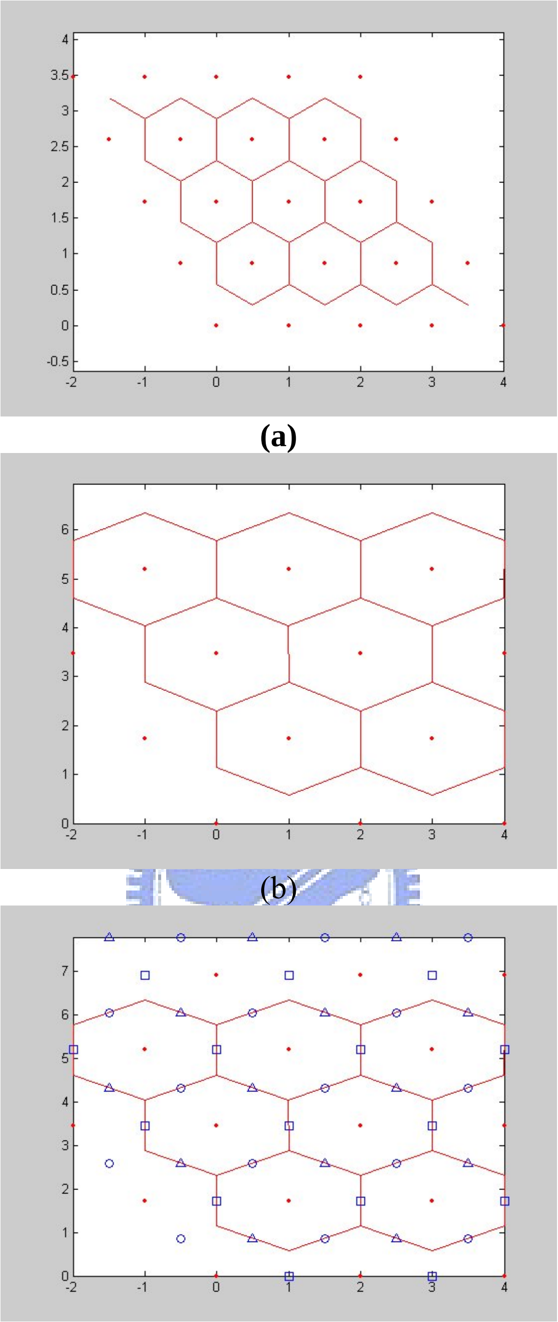
The four co-sets are then defined as follows:

$$\Lambda_0 = \Lambda_c + \mathbf{d}_0,$$
$$\Lambda_1 = \Lambda_c + \mathbf{d}_1,$$
$$\Lambda_2 = \Lambda_c + \mathbf{d}_2,$$
$$\Lambda_3 = \Lambda_c + \mathbf{d}_3,$$

where $\wedge_1$ (marked with ‘o’), $\wedge_2$(marked with ‘□’) and $\wedge_3$(marked with ‘Δ’) are the translated versions of $\wedge_0$ (marked with ‘.’) shifted by $\mathbf{d}_1, \mathbf{d}_2$ and $\mathbf{d}_3$, respectively (see Fig. 3.1 (c)). Note that the generating matrix $\mathbf{A_f}$ of Λ_f and $\mathbf{A_c}$ of $\Lambda_c(\wedge_0)$ are related by $\mathbf{A_c} = 2\mathbf{A_f}$, where the lattice points of $\wedge_f$ are shown in Fig. 3.1 (a), the lattice points of $\wedge_c$ are shown in Fig. 3.1 (b) and the lattice points of $\wedge_k$ (k=0,…,3) are shown in Fig. 3.1 (c).

3.3 The Proposed Multi-Rate Lattice Quantization Index Modulation (MRL-QIM) Watermarking

In this section, a multi-rate lattice quantization index modulation (MRL-QIM) watermarking scheme is described. The design goal of the MRL-QIM scheme is to use multi-rate watermark encoding to increase payload hiding and at the same time to increase the robustness of watermark detection. The advantage of the proposed scheme is two-fold. First, it can increase the detection robustness for error-prone transmission over unreliable network. Second, it is able to increase the watermark capacity while preserving the perceiving transparency. This is achieved by modulating the selected coefficients pair appropriately so that two bits of information can be embedded.



(a)

(b)

(c)

Fig. 3.1: Nested lattice. (a) Fine lattice $\wedge_f$, (b) Coarse lattice $\wedge_c = 2\wedge_f$, (c) Union of coarse lattices $\wedge_k$ (k=0,…,3), where $\wedge_0 = \wedge_c$.

Fig. 3.2(a) shows the flow of MRL-QIM, which is composed of a watermark embedding process and a transmission process. The original image is first transformed into the discrete cosine domain. The transformed coefficients are then grouped into 64 feature bands $\mathbf{X}(i,j)=(x_1,x_2,...,x_c)$, where i = *1,..,8; j =1,..,8* and c is the total coefficient count of each feature band. Next, each of two bits ($\mathbf{m}$) of the watermark message $\mathbf{W}$ is embedded in the selected two coefficient pair (where the selection depends on Key $\mathbf{K}$) using MRL-QIM quantizer as described in Section 3.1. The perturbed coefficients $\mathbf{Y}$ are then processed by multiple description scalar quantizer (MDSQ) to generate two independent descriptions, $\mathbf{Y}_1$ and $\mathbf{Y}_2$, and sent through two independent channels. The watermarked images ($\mathbf{I}'_{w1}$ from side decoder 1, $\mathbf{I}'_{w2}$ from side decoder 2 or $\mathbf{I}'_{w0}$ from central decoder) could then be obtained by receiving either one description (decoder 1 ($\hat{\mathbf{Y}}_1$) or decoder 2 ($\hat{\mathbf{Y}}_2$)) or two descriptions (decoder 0 ($\hat{\mathbf{Y}}_0$)) and inversing the discrete cosine transforms. In the detection process (Fig. 3.2(b)), the received image $\mathbf{I}'_{wr}$ (*r = 0, 1 or 2*) first goes through discrete cosine transform. The DCT coefficients are then grouped into 64 feature bands, $\hat{\mathbf{Y}}_r(i,j)=(x_1,x_2,...,x_c)$, where *i = 1,..,8; j =1,..,8* and c is the total number of coefficients of each feature band. Finally, apply detection process of MRL-QIM on $\hat{\mathbf{Y}}_r$ depending on Key $\mathbf{K}$ to extract the embedded watermark message $\mathbf{W}^*$.

3.3.1 MRL-QIM quantizer

The proposed MRL-QIM quantizer takes two DCT coefficients each time to encode two watermark bits into each of the four co-set points of a lattice (so called multi-rate). For the host signal $\mathbf{X}$, watermark message $\mathbf{W}$ and key $\mathbf{K}$, the embedding function is defined as $\mathbf{Y}=f(\mathbf{X},\mathbf{W},\mathbf{K})$. For a chosen DCT coefficient pair $\mathbf{x}=(x_1,x_2)\in\mathbf{X}$ depending on key $\mathbf{K}$, if $Q(\mathbf{x})$ is used for finding the nearest point of a lattice Λ, then $Q(\mathbf{x}-\mathbf{d}_m)+\mathbf{d}_m$ can be used for finding the nearest point of a co-set $\Lambda + \mathbf{d}_m$. To embed message $\mathbf{m}$ = *00 or 01 or 10 or 11* $\in \mathbf{W}$ into the host signal $\mathbf{x}$, we calculate $\mathbf{y} = Q_0(\mathbf{x} - \mathbf{d}_m) + \mathbf{d}_m$ to replace $\mathbf{x}$ with the watermarked coefficient pair $\mathbf{y} = (y_1, y_2) \in \mathbf{Y}$. For a received signal $\widehat{\mathbf{Y}}_r$ ($r \in \{0,1,2\}$) and key $\mathbf{K}$, the detection function is defined as $\widehat{\mathbf{W}} = f(\widehat{\mathbf{Y}}_r, \mathbf{K})$. To detect watermark message from watermarked signal $\mathbf{y} \in \widehat{\mathbf{Y}}_r$, we calculate $\mathbf{m}^* = \arg\min_m \|\mathbf{y} - Q_0(\mathbf{y} - \mathbf{d}_m) - \mathbf{d}_m\|$ to get the watermark message $\mathbf{m}^*$.

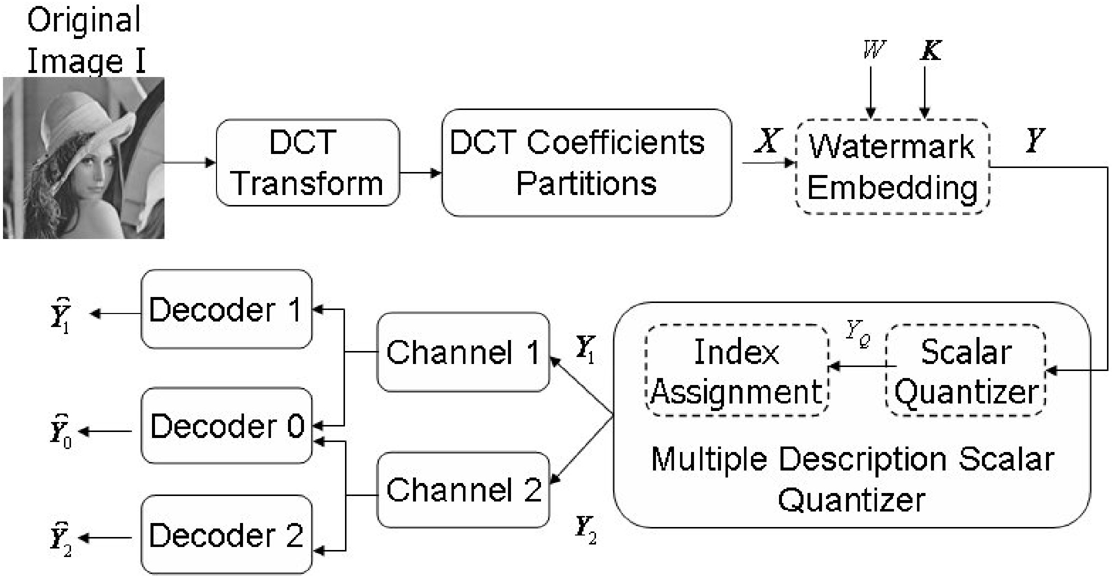


Fig. 3.2(a): Flow of proposed MRL-QIM watermark embedding scheme for error-prone transmission over unreliable network.

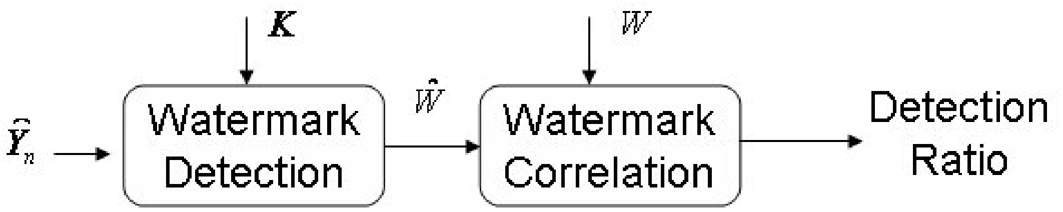


Fig. 3.2(b): Flow of proposed MRL-QIM watermark detection scheme.

For example, as depicted in Fig. 3.3(a), if one wants to embed watermark bits $\mathbf{m}$ = *00*, the original point marked with '+' is quantized to the point marked with '.',

with ‘x’ superimposed on the latter. And similarly, as depicted in Fig. 3.3(b), if one wants to embed watermark bits **m** = 01, the original point marked with ‘+’ is quantized to the point marked with ‘o’, with ‘x’ superimposed on the latter. Fig. 3.3(c) illustrate the case of embedding bits 10, the original point marked with ‘+’ is quantized to that marked with ‘□’ ,with ‘x’ superimposed on the latter. The case of embedding bits 11 is illustrated in Fig. 3.3(d).

3.3.2 The Embedding and Transmission Process of MRL-QIM

To embed *n* bits of watermark message $\mathbf{W}$, the algorithm is described as follows:

Step 1: The original image $\mathbf{I}$ is transformed using 8 by 8 block DCT transform.

Step 2: The DCT coefficients are then grouped into 64 feature bands

$\mathbf{X}(i, j) = (x_1, x_2, ..., x_c)$, where *i = 1,..,8; j =1,..,8* and *c* is the total number of coefficients of each feature band.

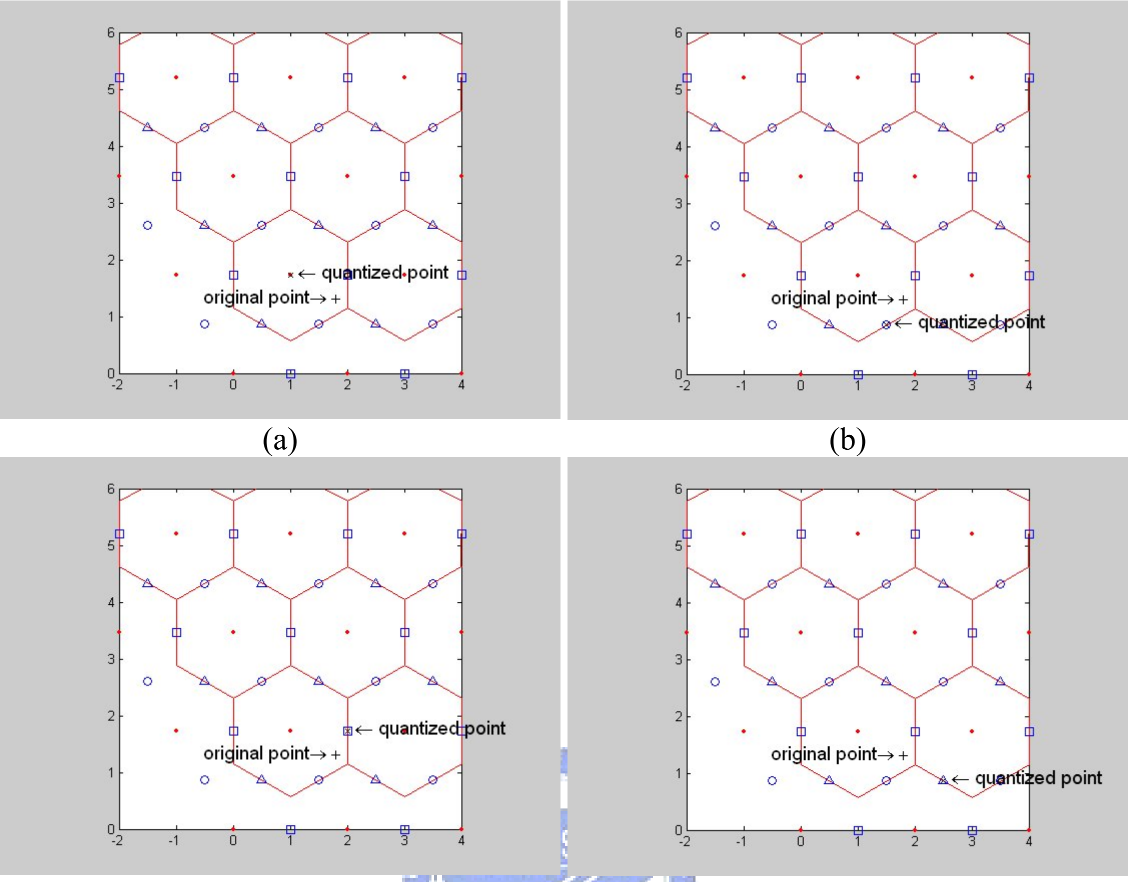
Step 3: Apply embedding process of MRL-QIM on $\mathbf{X}$ depending on Key $\mathbf{K}$ to embed watermark message $\mathbf{W}$ to obtain perturbed coefficients $\mathbf{Y}$.

Step 4: Each of the perturbed coefficients $\mathbf{Y}$ is quantized by a uniform scalar quantizer.

Step 5: Two descriptions $(\mathbf{Y}_1, \mathbf{Y}_2)$ of the quantized coefficient are created by mapping each quantized coefficient of $\mathbf{Y}_Q$ to a pair of numbers by the index assignment component.

Step 6: Transmit these two watermarked descriptions over network via two different channels.

Step 7: Apply an inverse transform to obtain a watermarked image $\mathbf{I}'_{wr}$ *(r = 0, 1 or 2)* depending on received descriptions $\hat{\mathbf{Y}}_r$ *(r = 0, 1 or 2).*



(a) (b)

(c) (d)

Fig. 3.3: Examples of multi-rate lattice watermark embedding. (a) For embedding watermark bits 00, the original point marked with '+' is quantized to the point marked with '.' ,with 'x' superimposed on the latter, (b) For embedding watermark bits 01, the original point marked with '+' is quantized to the point marked with 'o' ,with 'x' superimposed on the latter, (c) For embedding watermark bits 10, the original point marked with '+' is quantized to the point marked with '□' ,with 'x' superimposed on the latter, (d) For embedding watermark bits 11, the original point marked with '+' is quantized to the point marked with 'Δ' ,with 'x' superimposed on the latter.

3.3.3 The Detection Process of MRL-QIM

To extract *n* bits of watermark message $\mathbf{W}^*$, the algorithm is described as follows:

Step 1: The received image $\mathbf{I}_{wr}^{'}$ is transformed using 8 by 8 block DCT transform.

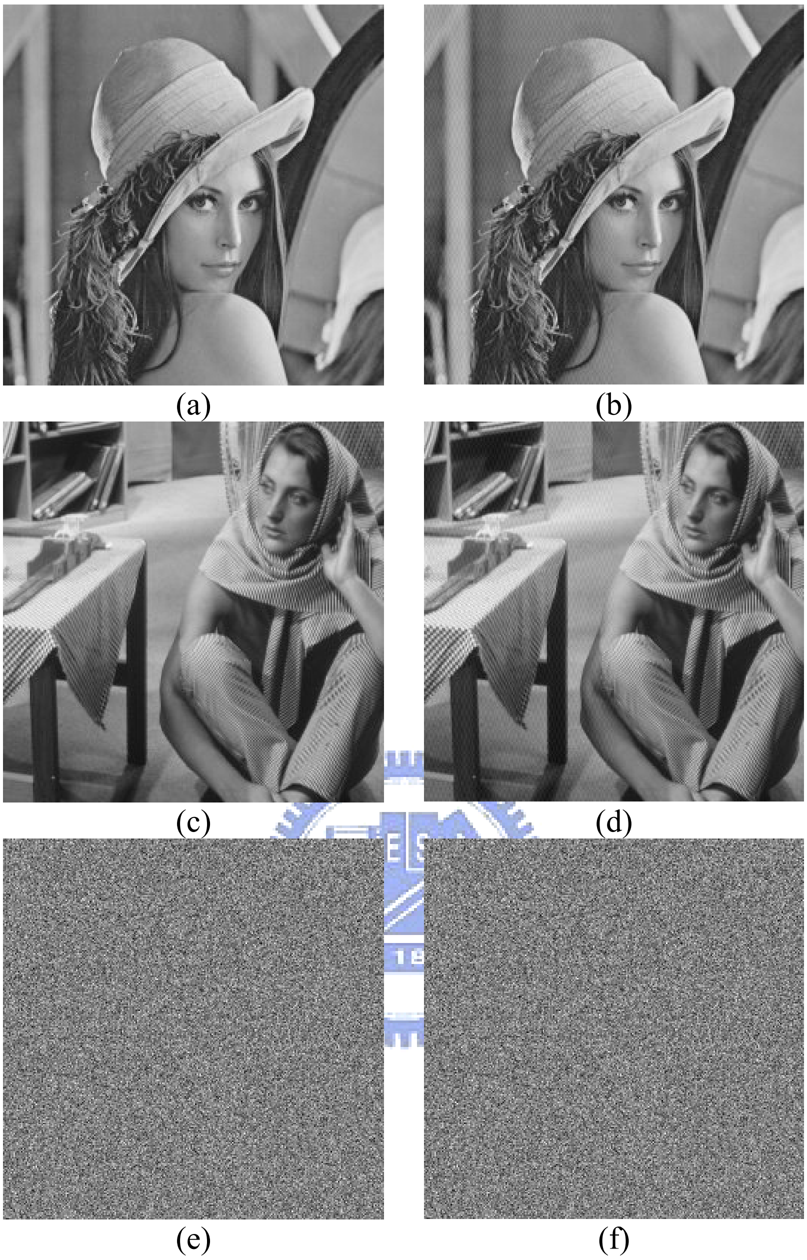
Step 2: The DCT coefficients are then grouped into 64 feature

bands $\widehat{\mathbf{Y}}_r(i, j) = (x_1, x_2, ..., x_c)$, *r = 0, 1 or 2; i = 1,..,8; j =1,..,8* and *c* is the total number of coefficients of each feature band.

Step 3: Apply detection process of MRL-QIM on $\widehat{\mathbf{Y}}_r$ *(r = 0, 1 or 2)* depending on Key $\mathbf{K}$ to get the extracted watermark message $\mathbf{W}^*$.

3.4 Experimental Results

To evaluate the effectiveness of the proposed method, experimental simulations on both of the Monte Carlo simulated Gaussian images and several real images (Lena, Barbara, House and boat) were performed. To save space, only "Lena" and "Barbara" as well as the average of detection ratios of Gaussian images are given here. For each run of Monte Carlo simulation, host signals X drawn from 256 x 256 samples of a Gaussian zero-mean random variable were generated, each having standard deviation σ_X ranging from 10 to 100 with step size 10. All these Gaussian simulated data were then normalized to the range of 0~255 to simulate Gaussian gray level images. We performed 100 times on each of the above simulations using different seeds, so that totally 1000 (10 x 100) Gaussian images were employed to obtain the average detection ratios. In order to further demonstrate the effectiveness of our proposed MRL-QIM, the state-of-the-art watermark technique QIM [7] and a combined spread spectrum and QIM (SS-QIM) [21, 30] were simulated for comparisons. The SS-QIM scheme utilizes spread spectrum approach, in which a watermark strength weighting parameter α is needed, to obtain a correlation value. This correlation value will then be quantized based on specified embedding quantization step size δ and watermark bit (0 or 1) to produce a watermark value.



(a) (b) (c) (d) (e) (f)

Fig. 3.4: (a) Original Lena (256 x 256 with a gray scaled level), (b) Watermarked Lena (PSNR 41.1 dB), (c) Original Barbara (256 x 256 with a gray scaled level), (d) Watermarked Barbara (PSNR 41.1 dB), (e) One sample of original Gaussian images (256 x 256 with a gray scaled level), (f) Watermarked Gaussian image (PSNR 41.7 dB).

For MRL-QIM scheme, two DCT transformed coefficients were simultaneously used to embed two bits of watermark information, and totally 1024 coefficients were used to embed 1024 bits of watermark information. For traditional vector QIM scheme, two DCT transformed coefficients were formed as a vector to embed one bit of watermark information, and totally 2048 coefficients were used to embed 1024 bits of watermark information. As for SS-QIM, an algorithm adopted from [30] was implemented and the trellis error correction coding module was removed to make fair comparisons. Note that the detection rate can be further enhanced by employing error correction coding module for all these three schemes.

From our experiments, the degree of PSNR dropped depending on the embedding quantization step size. A larger quantization step size brought more robustness, but it also introduced more distortion. We follow the common practice by fixing two requirements, namely watermark capacity and the transparency (distortion) of watermarked image, and then comparing the robustness. To make the comparison fair, the parameters that defined the embedding quantization step size or watermark strength weighting parameter were adjusted so that similar PSNR values (about 41 dB) for these three schemes could be obtained. In our setting, the embedding quantization step size δ was set to 28, 44 and 7 for MRL-QIM, vector-QIM and SS-QIM, respectively. And for SS-QIM, the other watermark strength weighting parameter α was set to 0.9. The original and watermarked images for MRL-QIM were shown in Fig. 3.4(a) and 3.4(b) (Lena), Fig. 3.4(c) and 3.4(d) (Barbara) and Fig. 3.4(e) and 3.4(f) (one sample of Gaussian images), respectively.

To evaluate the reliability of watermark detection, the detection ratio ρ was defined as:

$$\rho = \frac{Total\ \ number\ \ of\ \ correctly\ \ \det ected\ \ bits}{Total\ \ number\ \ of\ \ embedded\ \ bits}. \tag{3.1}$$

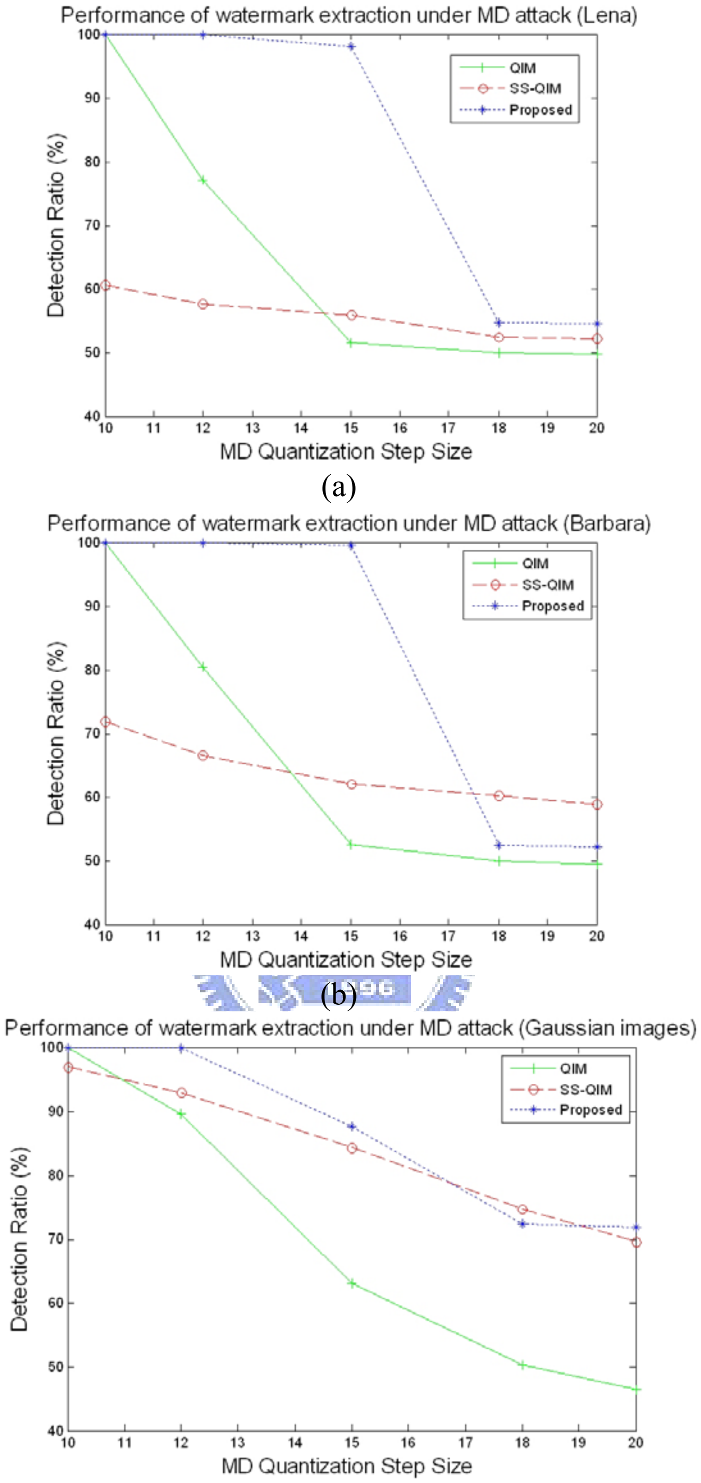
A higher value of ρ indicated a more reliable detection. The perfect recognition rate could be achieved when the value of ρ equals 1.

In addition to the degree of robustness against packet loss, a desirable and fundamental property for a watermarking algorithm is to survive compression attack. In the real-world applications, compression is frequently used to facilitate efficient storage and transmission. Here, we used images compressed by JPEG (low quality factor ranging from 60 to 80) as test images.

The performance of the detection on receiving only description 1 (similar results can be obtained via description 2) against the MD attack over various transmission rates is evaluated first. The detection ratios of “Lena”, “Barbara” and 1000 Gaussian images are depicted in Fig. 3.5(a), 3.5(b) and 3.5(c), respectively. It is clear that the proposed MRL-QIM outperformed traditional vector QIM for all testing images over all transmission rates, and MRAL-QIM performed better than SS-QIM except in the case of high-rate transmission (MD quantization step size = 18) for Gaussian and “Barbara” images.

As for JPEG compression attack, the detection ratios of “Lena”, “Barbara” and 1000 Gaussian images are shown in Fig. 3.6(a), 3.6(b) and 3.6(c), respectively. The performance of proposed MRL-QIM is still better than traditional vector QIM and superior to SS-QIM for most of the compression rates, except at JPEG Quality = 60 for “Lena” and “Barbara”.

Note that, even though the detection capability of SS-QIM approach is not so good as the other two schemes when under weaker attacks, it has a smoother decay in detection ratios than pure quantization based (vector QIM and MRL-QIM) ones when the attacks become stronger. Based on the experimental result, we speculate that the possible combination SS-MRL-QIM would be a very interesting topic worth further investigation.

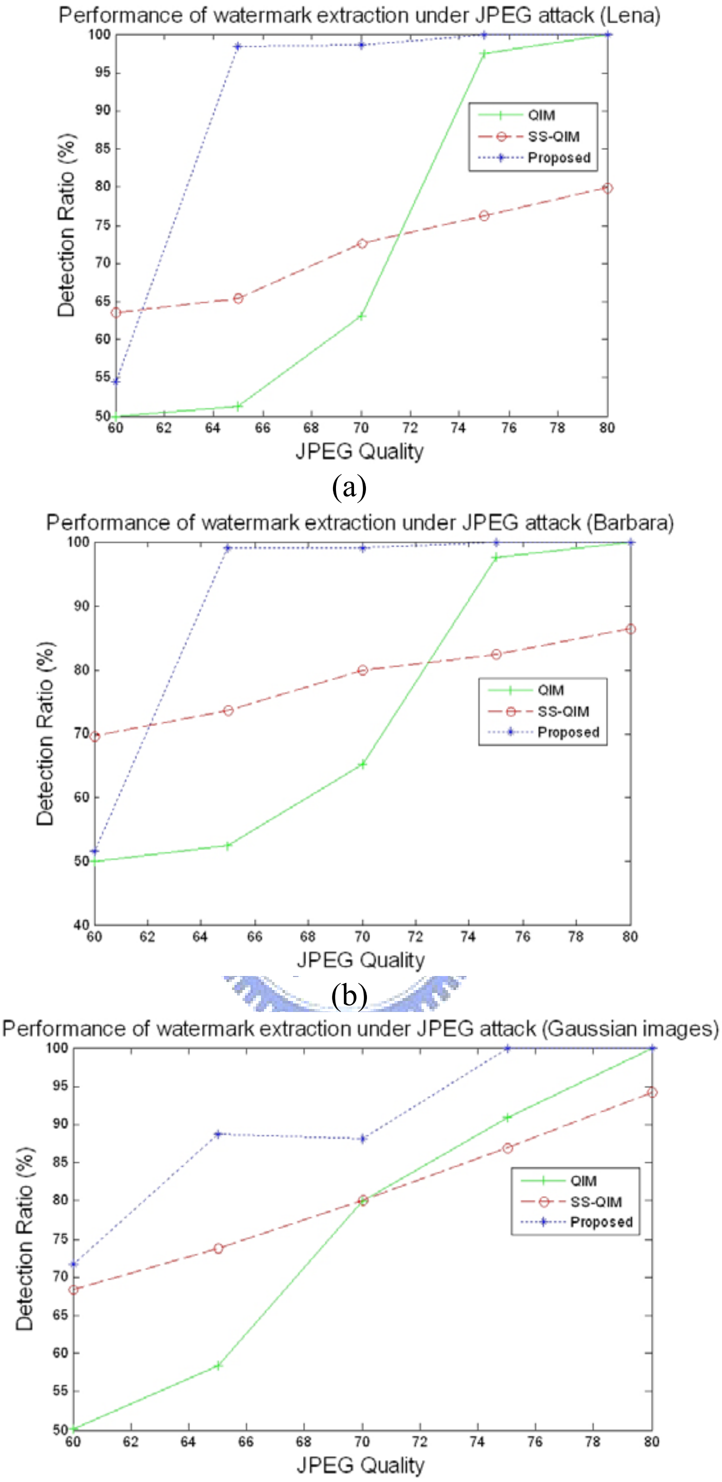


(a)

(b)

(c)

Fig. 3.5: The comparison in terms of detection ratios among QIM, SS-QIM and proposed MRL-QIM against various MD transmission rates. (a) “Lena”, (b) “Barbara”, (c) the average of 1000 Gaussian images.



(a)

(b)

(c)

Fig. 3.6: The comparison in terms of detection ratios among QIM, SS-QIM and proposed MRL-QIM against JPEG compression. (a) “Lena”, (b) “Barbara”, (c) the average of 1000 Gaussian images.

3.5 Summary

We have presented in this chapter an MRL-QIM watermarking scheme that is robust to non-linear value-metric distortion introduced by MD transmission. For a traditional balanced two-description case in a packet transmission network, the embedded watermark can be extracted with the reception of either one or two descriptions. The experimental result shows that the proposed MRL-QIM outperforms traditional vector QIM overall and performs better than SS-QIM in the case of high-rate transmission. Furthermore, in the case of compression attack, the performance of proposed MRL-QIM still performs better than traditional vector QIM and superior to SS-QIM for most of the compression rates. In the future, we expect to seek other transforms and statistic models to further enhance the robustness and increase the watermark payload while preserving the visual quality of the transmitted image. Furthermore, the steganography security against statistical steganalysis should also be addressed to enhance the security for reliable transmission.

Chapter 4 The Detection of Weak Watermark Signal in the Fourier-Mellin Domain

We propose a rotation, scaling and translation (RST) resilient blind image watermarking technique by using Fourier-Mellin transform and informed coding, where watermark detection does not require the existence of the original image. Since the inverse Log-Polar mapping (ILPM) could severely destroy the embedded watermark, it seems that a very strong watermark should be employed to survive the self-destruction process from ILPM, but in turn this leads to a poorly watermarked image. In order to get around this dilemma, a local optimal watermark is selected (informed coding) for embedding to avoid using heavy watermark in the Fourier-Mellin invariant domain and by doing so the survived weak watermark signal can be reliably detected. The transparency is further preserved by transforming (Log-Polar mapping) only the specified circular area of the image while keeping other portions of the image intact. Experimental results demonstrate that the resulting watermark is robust to a variety of image processing attacks and validate the effectiveness of the proposed scheme.

4.1 Introduction

Achieving watermark robustness against geometric attacks has always been a challenging research topic. Due to its numerous potential applications many researchers have been engaged in solving this problem [31]. There are primarily two main categories of watermarking techniques for resolving geometric attacks in the literature. One approach is adopting resynchronization template, and the other approach is using invariant features. For resynchronization approach [32-38], either an additional template or a self-reference pattern is utilized for detecting synchronization. For invariant features approach [39-42], features from geometric invariant domain are selected for watermark embedding. Some of the methods [32-38] attempt to find out particular feature peaks by performing auto-correlation or cross-correlation in the Fourier transform domain, and then the geometric offsets are estimated for resynchronization. Others [39-42] make use of the property of circular shift invariance of Fourier transform accompanied with another transform (such as Log-Polar) or additional skills to derive geometric invariance.

Pereira and Pun [34] extract some local peaks of Fourier spectrum as feature points, whose Cartesian coordinates are then transformed into polar ones. The feature points in the polar form are rearranged into several equally spaced bins according to their angles, and thus the feature points in the same bin form a radial line. Two lines are randomly selected to act as a resynchronization template, where each line with seven points is sufficient for reliably recovering the geometric offset. Kutter [33] embeds the self-reference pattern several times at horizontally and vertically shifted locations for recovering affine transformations. The above two template based approaches are vulnerable to template attack [44]. Solachidis and Pitas [41] propose to watermark on a ring in the Fourier spectrum. The watermark is detectable for small rotations by using identically valued watermark in the *S* ring sectors, but further searching is needed for arbitrary rotation angles. O'Ruanaidh and Pun [42] first suggest a watermarking method based on the Fourier-Mellin transform. They notice some practical implementation difficulties for watermarking in this invariant Fourier-Mellin domain. One of the key problems is the self-destruction of the embedded watermark during inverse Log-Polar transform phase. They invent a scheme that images to be watermarked avoid passing through the lossy inverse

Log-Polar mapping (ILPM), and at the same time allowing only the watermark itself to pass through this lossy process. Another solution based on Fourier-Mellin transform [43], employs an iterative method to estimate changes from the Log-Polar map and then iterates the embedding on the DFT (Discrete Fourier Transform) coefficients, hence images to be watermarked avoid passing through the lossy inverse Log-Polar mapping. However, their method needs further searching except for some particular rotation angles.

The goal of this chapter is aiming at embedding a robust watermark into the invariant watermarking space of an image, which is resilient to various geometric attacks. In the next section, the DFT properties and Log-Polar Mapping (LPM) are overviewed. Section 4.3 discusses the invariant watermarking algorithm using informed coding. In Section 4.4 experimental results are presented and analyzed. Finally, in Section 4.5 conclusions as well as future work are given.

4.2 Fourier Mellin Transform (FMT)

The Fourier-Mellin transform [45-46] is widely used in the pattern recognition and computer vision community. For recognition applications, it is the most desired to pursue the invariant features. However, for watermarking applications, it is essential to further carry out the nearly perfect reconstruction of a watermarked image with no dramatic loss of embedded watermark for reliable detection. The two components (DFT and Log-Polar mapping) of Fourier-Mellin transform which build our invariant watermark embedding space are described as follows.

4.2.1 DFT and its Properties [47]

The Fourier transform is an infinite linear combination of dilated cosine and sine waves whose statistical properties are invariant over time delay (translation).

4.2.1.1 DFT of image

The DFT for an image *f (x, y)* of size *N* x *N* is defined as follows:

$$F(u,\ v) = \frac{1}{N^2}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,\ y)\ e^{-j2\pi(ux/N+vy/N)}, \tag{4.1}$$

where $0 \le u \prec N, 0 \le v \prec N$.

The inverse transform is

$$f(x,\ y) = \sum_{u=0}^{N-1}\sum_{v=0}^{N-1} F(u,\ v)\ e^{j2\pi(ux/N+vy/N)}, \tag{4.2}$$

where $0 \le x \prec N, 0 \le y \prec N$.

There are two parts for the DFT transformed image: the real part *R (u, v)*, and the imaginary part *I (u, v)*. The Fourier spectrum (magnitude) and phase (angle) representation are defined as follows:

$$M = \ |F(u,v)| = \sqrt{R^2(u,v) + I^2(u,v)}. \tag{4.3}$$

$$\varphi(u,v) = \tan^{-1}\frac{I(u,v)}{R(u,v)}. \tag{4.4}$$

4.2.1.2 Fourier Transform Circular Shift Invariant Properties

If image *f₂* is a translated replica of *f₁* with translation *(x₁, y₁)*, i.e.,

$$f_2(x,\ y) = f_1(x - x_1,\ y - y_1). \tag{4.5}$$

The corresponding Fourier transforms F_1 and F_2 are related by

$$F_2(u,\ v) = F_1(u,\ v) * e^{-j2\pi(ux_1+vy_1)}. \tag{4.6}$$

The magnitudes of F_1 and F_2 are the same, and only the phase representations differ.

4.2.2 Log-Polar Mapping [46]

For a point (x, y) in Cartesian coordinates, its corresponding Polar form is (ρ,θ) with the relationship of $x = \rho\cos\theta$ and $y = \rho\sin\theta$, where $\rho \in R$ and $0 \le \theta \prec 2\pi$. The point in Polar coordinates is then transformed into the point (μ,θ) in Log-Polar coordinates by taking the logarithm of the scale ρ ($\mu = \log(\rho)$). For image *f(x, y)* in Log-Polar form, the scaling is converted to shifts along the radial axis μ and rotation is converted to shifts along the angle axis θ.

For example, if image f_2 is a replica of f_1 scaled by *s* with the described matrix

$$\begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix}, \tag{4.7}$$

these two images will then be related by

$$f_2(x,\ y) = f_1(sx,\ sy). \tag{4.8}$$

Note the Fourier scaling property that scaling f_1 in spatial domain by a scale *s* results in inverse scaling F_1 in frequency domain. Let M_1 and M_2 be the magnitudes of F_1 and F_2. The Fourier transforms of f_1 and f_2 will be related by

$$M_2(u,\ v) = \frac{1}{s^2} M_1(u/s,\ v/s)\ . \tag{4.9}$$

Taking the multiplicative factor ($1/s^2$) as constant *c*, their Fourier magnitude spectra in polar representation are related by

$$M_2(\rho,\ \theta) = cM_1(\rho/s,\ \theta). \tag{4.10}$$

Taking the logarithmic scale in Eq. (4.10), the scaling is converted to shifts along the *x*-axis, i.e.

$$M_2(\log\rho,\ \theta) = cM_1(\log\rho - \log s,\ \theta). \tag{4.11}$$

On the other hand, for an image f_1 rotated θ_0 as image f_2 with the described matrix

$$\begin{bmatrix} \cos\theta_0 & -\sin\theta_0 \\ \sin\theta_0 & \cos\theta_0 \end{bmatrix}, \tag{4.12}$$

these two images will be related by

$$f_2(x,\ y) = f_1(x\cos\theta_0 + y\sin\theta_0,\ -x\sin\theta_0 + y\cos\theta_0). \tag{4.13}$$

Due to the Fourier rotation property that rotating *f₁* in spatial domain by an angle θ_0 rotates F_1 in frequency domain by the same angle, we have

$$M_2(u,\ v) = M_1(u\cos\theta_0 + v\sin\theta_0,\ -u\sin\theta_0 + v\cos\theta_0), \tag{4.14}$$

and in their log polar coordinates

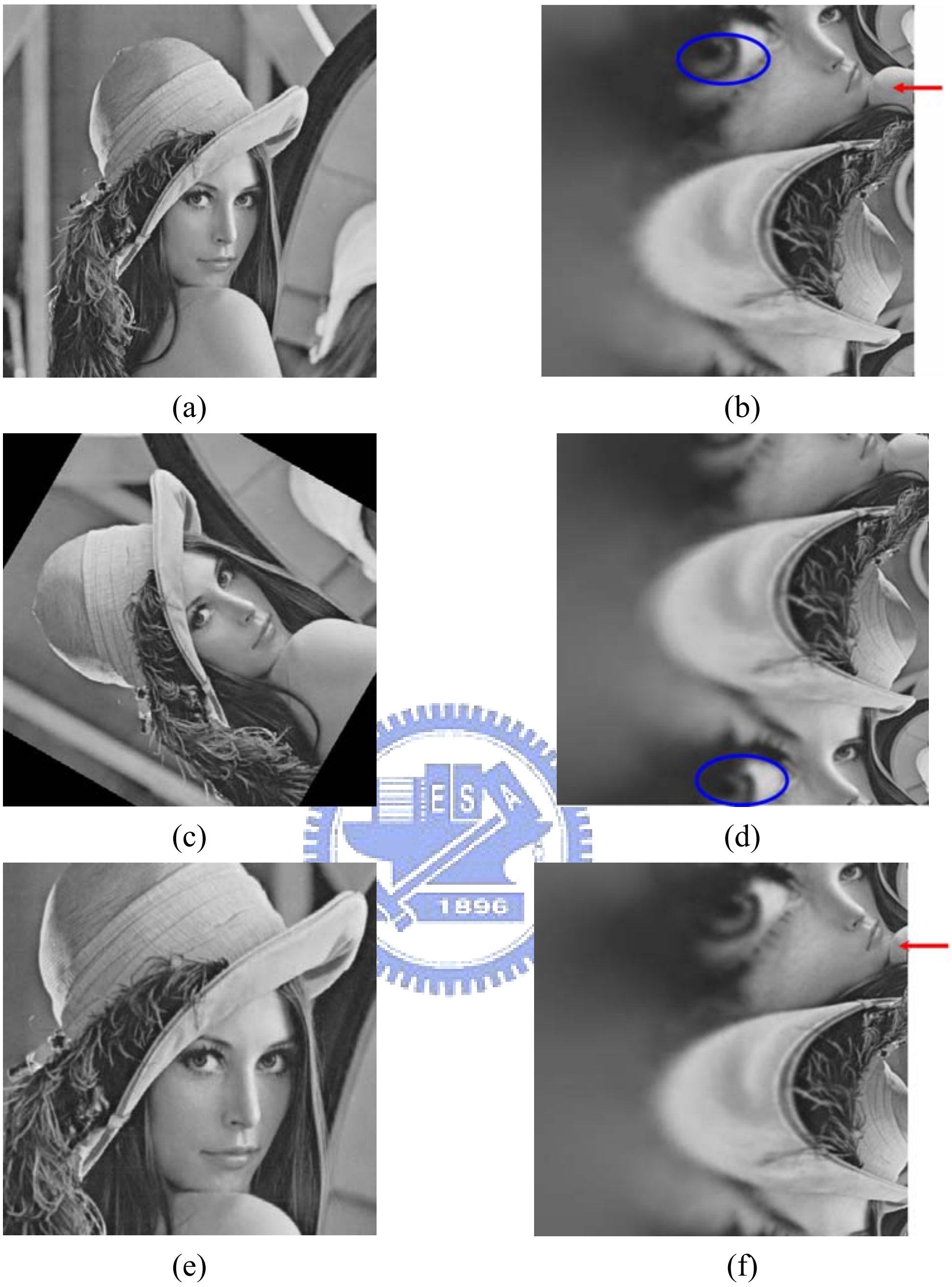
$$M_2(\rho,\ \theta) = M_1(\rho,\ \theta - \theta_0)\,. \tag{4.15}$$

Combining the rotations, scaling and translations in one step, their Fourier magnitudes in the log polar form are related by

$$M_2(\log\rho,\ \theta) = M_1(\log\rho - \log s,\ \theta - \theta_0)\,. \tag{4.16}$$

After converting the rotations and scaling in spatial domain into translations problem in frequency domain, further taking the Fourier transform of this Log-Polar map will lead to the so-called RST invariant Fourier-Mellin domain [45-46]

To better explain the characteristic of Log-Polar mapping, we use an example to demonstrate the effect through rotation and scaling of image. Fig. 4.1(b) is the Log-Polar image of original “Lena”. We can see that rotation is represented as vertical shift in the Log-Polar image, which is marked by the eclipse in Fig. 4.1(d); whereas scaling is represented as horizontal shift in the Log-Polar image, which is indicated by the arrow marker in Fig. 4.1(f).



(a)

(b)

(c)

(d)

(e)

(f)

Fig. 4.1: The characteristic of Log-Polar mapping. (a) “Lena”, (b) Log-Polar map of “Lena”, (c) “Lena” rotated 60^{o}, (d) Log-Polar map of rotated “Lena”, (e) “Lena” scaled 1.5, (f) Log-Polar map of scaled “Lena”.

4.3 Proposed informed FMT watermarking algorithm

It is a reasonable speculation that, among the enormous watermark patterns, there would always exist some watermark patterns more suitable for certain specific media content, while others might not so appropriate. For the correlation based spread spectrum watermarking technique [48], if the host signal is *V* and watermark signal is *W,* then the embedding rule for watermarked signal *V'* can be simplified as

$$V' = V + W, \tag{4.17}$$

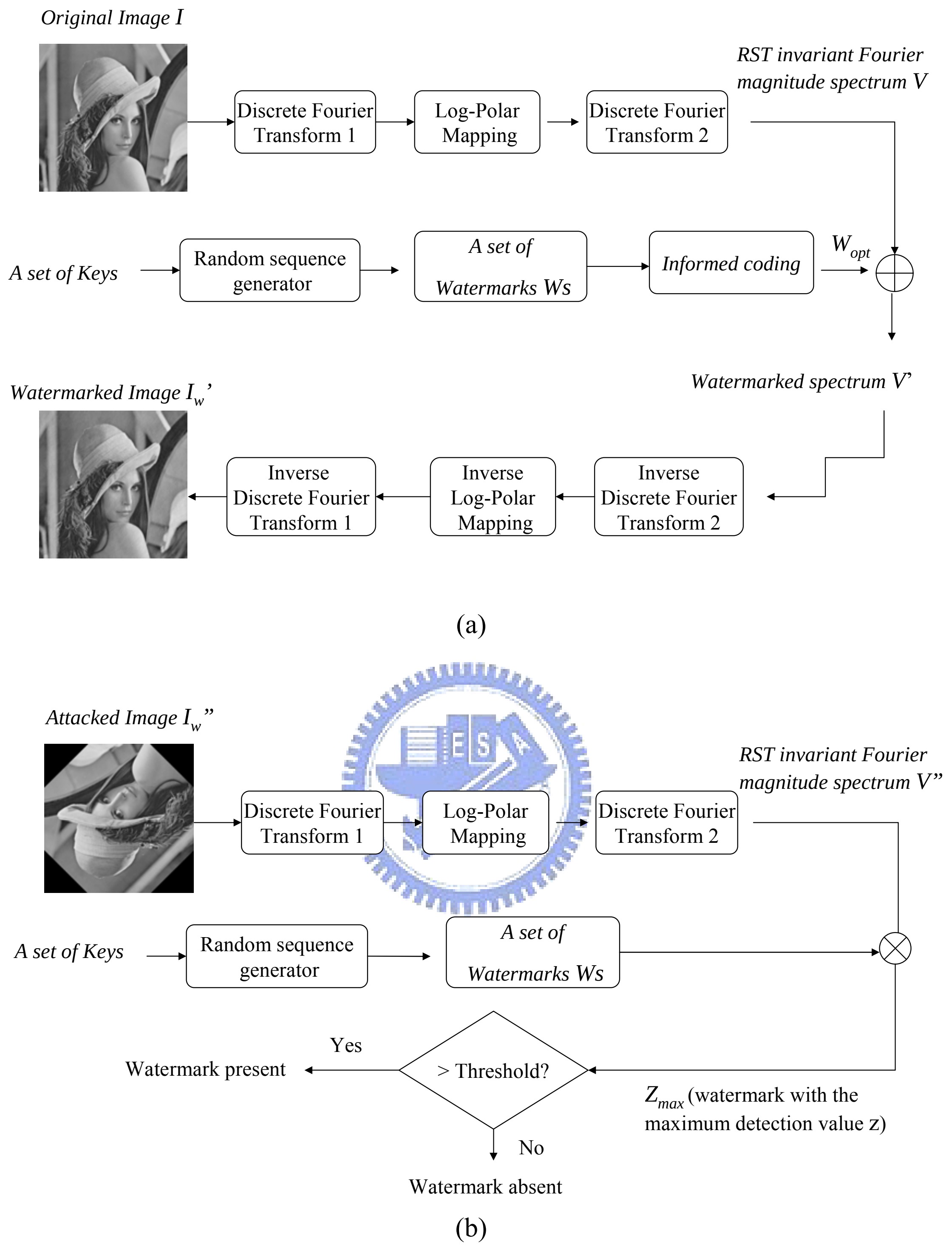
and the detection rule can be written as

$$z = <V',W> = <V,W> + <W,W>, \tag{4.18}$$

where <*V'*, *W*> is the inner product of *V'* and *W*, and *z* is the correlation value.

The detection objective is to make *z* as high as possible. To make z higher, some authors [49] try to find robust watermarking space to make the term <*V*, *W*> higher for fixed *W*. We argue that there might exist a case that a different watermark *W'* (not the originally embedded watermark) has a higher value of <*V*, *W'*> + <*W'*, *W'*> leading to higher *z*, thus false alarm can be increased. We propose the scheme which employs an informed coding approach [21], in which a robust watermark (W_{opt}) is selected from numerous watermarks by maximizing <*V'*, *W*> prior to embedding.

In this section, a geometric invariant watermarking technique using Fourier-Mellin transform and informed coding is proposed. The main ideas of the proposed scheme as depicted in Fig. 4.2(a) and 4.2(b) are to increase the detection robustness and to increase transparency. To this end, a robust watermark is chosen prior to embedding to maximize detection robustness. Furthermore, only the circular parts of an image with radius no more than half the image row size are LPM transformed, and other pixels of the image remain intact. This is because the interpolation distortions by LPM /ILPM increase with larger radius. In the embedding process in Fig. 4.2(a), the original image first goes through Fourier transform, the magnitudes of the coefficients within a specified circular area (whose radius is no more than half the image row size) of the image are then re-sampled in Log-Polar coordinates; Fourier transform is again applied on the re-sampled coordinates to derive Fourier magnitude which is RST invariant. The robust watermark chosen in the informed coding stage is then embedded into the RST invariant Fourier magnitude. The watermarked image could then be obtained by inversing the transforms of the above processes. In the detection process in Fig. 4.2(b), the attacked image first goes through Fourier transform, the magnitudes of the coefficients are then re-sampled in Log-Polar coordinates and Fourier transform is again applied on the re-sampled coordinates to derive RST invariant Fourier magnitude. Once the feature vector from the middle frequency band of this RST invariant domain is selected, we then compute the correlation coefficients between the feature vector and each of predefined reference watermarks. Based on the theory of hypothesis testing, the watermark is present if the resulting maximum correlation value computed is greater than a specified threshold. Otherwise the watermark is absent.


Original Image I
Discrete Fourier Transform 1
Log-Polar Mapping
Discrete Fourier Transform 2
RST invariant Fourier magnitude spectrum V
A set of Keys
Random sequence generator
A set of Watermarks Ws
Informed coding
W_{opt}
Watermarked spectrum V'
Watermarked Image I_w'
Inverse Discrete Fourier Transform 1
Inverse Log-Polar Mapping
Inverse Discrete Fourier Transform 2
Attacked Image I_w"
Discrete Fourier Transform 1
Log-Polar Mapping
Discrete Fourier Transform 2
RST invariant Fourier magnitude spectrum V"
A set of Keys
Random sequence generator
A set of Watermarks Ws
Yes
Watermark present
> Threshold?
Z_{max} (watermark with the maximum detection value z)
No
Watermark absent

(a)

(b)

Fig. 4.2: (a) The flow of proposed watermark embedding scheme, (b) The flow of proposed watermark detection scheme.

4.3.1 Watermark Generation

The watermark $W = \{x_1, x_2, \ldots, x_k\}$ consists of a pseudo-random sequence of k real numbers drawn from independent identically normal distributions and determined by a secret key. There are totally N_w watermarks generated. Prior to watermarking embedding, these N_w watermarks are used as reference watermarks at informed coding stage to find the one that has local optimal correlation with the host image. On the other hand, these N_w watermarks are again applied for detection to distinguish the one with maximum correlation.

4.3.2 Informed FMT Watermark Embedding

The embedding algorithm consists of the following steps:

Step 1: Take the magnitude of the DFT coefficients of the original image I of size N by N.

Step 2: Perform the Log-Polar mapping of the log magnitude obtained from step (1). Note that the maximal radius ρ is half of the image row size, i.e., $N / 2$.

Step 3: Sample uniformly along the log scale and θ axis to obtain a Log-Polar sampled spectrum image of size N_ρ by N_θ.

Step 4: Take the magnitude of the DFT coefficients of these samples. Select k coefficients residing in the middle frequency band to form feature vector V.

Step 5: Informed coding: select the local optimal watermark pattern from a set of candidate watermark patterns (say N_w =*1000*). Embed each W_i into V using Eq. (4.19). Then compute the correlation of each V_i' with each corresponding W_i. Identify the one, say W_{opt}, which has the local optimal correlation among these N_w computed correlation values.

$$V_i' = V(1 + \alpha W_i), \tag{4.19}$$

$$W_{opt} = \arg\max_{W_i} W_i . V_i^{'} , \tag{4.20}$$

where $i = 1,\ldots,\ N_w$ and α is embedding strength.

Step 6: Embed the selected W_{opt} obtained into the feature vector V using Eq. (4.21).

$$V' = V(1 + \alpha W_{opt}) \tag{4.21}$$

Step 7: Take the inverse DFT

Step 8: Take ILPM. Note that as in step (2) the maximal radius ρ is *N / 2*. Some spectrum pixels are not inversely mapped, and these missed pixels are substituted by the original corresponding spectrum pixels obtained in step (1).

Step 9: Take the inverse DFT to get watermarked image $I_w^{'}$

4.3.3 Informed FMT Watermark Detection

The detection algorithm consists of the following steps:

Step 1: Take the magnitudes of the DFT coefficients of the investigated image $I_w^{''}$ ($I_w^{'}$ being attacked) of size *N* by *N*.

Step 2: Perform the Log-Polar mapping of the log magnitude obtained from step (1). Note that the maximal radius ρ is half of the image row size, i.e., *N* / 2.

Step 3: Sample uniformly along the log scale and θ axis to obtain a Log-Polar sampled spectrum image of size N_ρ by N_θ.

Step 4: Take the magnitude of the DFT coefficients of these samples. Select *k* coefficients residing in the middle frequency band to form feature vector *V”*.

Step 5: Compute the correlation coefficients z_i between the feature vector *V”* and each of the N_w predefined reference watermarks, W_1, W_2,.., W_{Nw}.

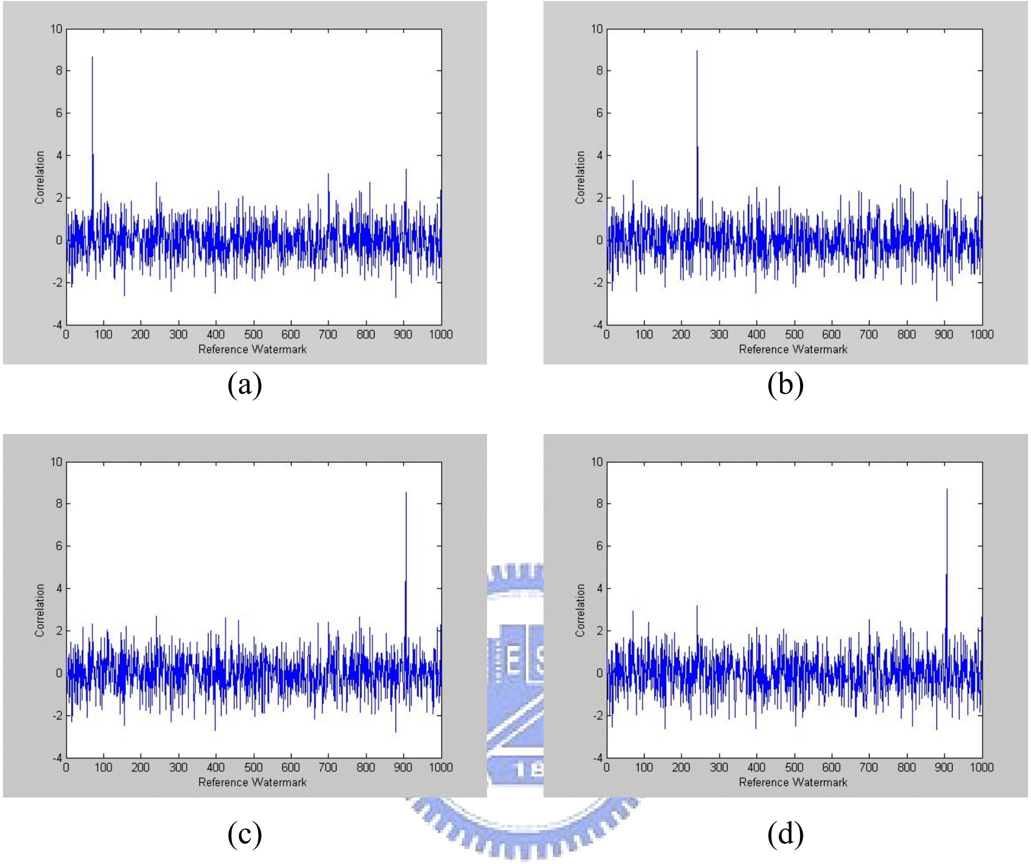
$$z_i = \frac{W_i \cdot V''}{k} \text{ , where } i = 1, \ldots, N_w \tag{4.22}$$

Step 6: The watermark is present if the resulting maximum correlation value (called z_{max}) computed is greater than a specific threshold. Otherwise the watermark is absent.

4.4 Experimental Results

To evaluate the effectiveness of the proposed method, four standard test images of size 512 x 512 including “Lena”, “Barbara”, “Goldhill” and “Boat” are used as host signals to embed watermark information. The parameters used in these experiments are $N_\rho = 512, N_\theta = 512, Nw = 1000, k = 45000$ and $\alpha = 0.4$. The watermark strength α determines the tradeoff between the robustness and distortion. The PSNRs of watermarked “Lena”, “Barbara”, “Goldhill” and “Boat” are 34.91, 32.57, 34.25 and 33.33, respectively. More discussions on image fidelity are presented in Section 4.1. The watermark which is most correlated with the watermarked image are detected with peak correlation values 8.67, 8.94, 8.53 and 8.69, respectively as depicted in Fig. 4.3. Their corresponding watermark indexes are 71, 241, 906 and 906, which are the same watermark indexes as originally embedded.

In addition to compression attack, some variety of signal processing attacks such as RST, cropping, asymmetric row and column removal, frequency mode Laplacian removal (FMLR) [50], Gaussian filtering, sharpening, median filtering, noise adding, histogram equalization, hybrid manipulations and Stirmark random bending are also tested. Among them, the JPEG-2000 attacked images are generated using JPEG-2000 VM8.0, and others are generated using Stirmark3.1 [18] and PaintShopPro7.0. The watermarked images using our proposed scheme survive well for all the abovementioned attack types except for the random bending attack. Main issues regarding performance evaluation of the proposed scheme are discussed in the following.



(a) (b)

(c) (d)

Fig. 4.3: (a) Detection response for watermarked “Lena” (Peak value 8.67 for watermark index 71), (b) Detection response for watermarked “Barbara” (Peak value 8.94 for watermark index 241), (c) Detection response for watermarked “Goldhill” (Peak value 8.53 for watermark index 906), (d) Detection response for watermarked “Boat” (Peak value 8.69 for watermark index 906).

4.4.1 Fidelity

Since the degree of PSNR drop depends mostly on the LPM and ILPM, the PSNR quality of the watermarked image is now constrained by the LPM/ILPM transform. We observe that for original “Lena” with no watermarking the PSNR of the reconstructed image through the two rounds of DFT transform is 56.00 dB, while the PSNR of reconstructed image through the two rounds of DFT transform with additional LPM in between is 36.00 dB. In contrast, the PSNR of watermarked image through the two rounds of DFT transform with LPM is 34.91 dB and its quality loss of 1.09 dB is rather insignificant compared to the quality loss of 20.00 dB through LPM. Note that the PSNR quality of watermarked image still meets the acceptable quality standard (near 35.00 dB).

4.4.2 Payload

The current system utilizes all of the selected DFT coefficients to determine if a watermark exists, which gives one bit of information. The possible extension is to first tile the input image or to partition the DFT coefficients so that more information can be embedded. A more advanced approach is by developing lattice codes [21] for exploring the multiple-bit watermark capacity.

4.4.3 Probability of False Positive

For watermark detection process, there are two types of error: one is false negative error (miss detection, false rejection), and the other is false positive error (false alarm). It is essential for a copyright protection system to have low false rejection probability (1-P_D, where P_D represents correct detection rate or recognition rate) for fixed false positive error detection probability (P_F).

The test images are obtained by manipulating each of the four tested standard images with various combinations of rotation (0° to 359° with step-size 18°), scaling (0.8 to 1.2 with step-size 0.05) and translation (0 to 15 with step-size 5) to generate 720 images for each and there are 2880 in total. For each of these 2880 watermarked images, the peak correlation value of the image correlated with 1,000 watermarks is obtained and collected. Yet for original un-watermarked images, all the 2880 x 1000 correlation values for all these 2880 images correlated with 1,000 watermarks are collected.

Fig. 4.4 illustrates the distributions of these detection values for watermarked and un-watermarked images. For watermarked images with local optimal watermark selection, the detections values (using Eq. (4.21)) are ranging from 3.36 to 8.95 with mean 4.58 and variance 0.70. While without informed coding, if a worst watermark (the one that is the most uncorrelated with the host signal) is selected, the detections values are ranging from 3.06 to 4.22 with mean 3.60 and variance 0.07. The detection values for un-watermarked images are ranging from –4.22 to 4.29 with mean –0.01 and variance 1.05. It is easy to perceive the contrast between Fig. 4.4(a) and Fig. 4.4(b). With informed watermark coding, the detection values are higher to distinguish from that of un-watermarked. With the detection thresholds setting between 3.9 and 4.5, the probability of false positive error rate and the probability of recognition rate could be obtained to evaluate the performance of the system. More watermark patterns and images could be utilized to guarantee even more accuracy of these ratings.

4.4.4 Robustness

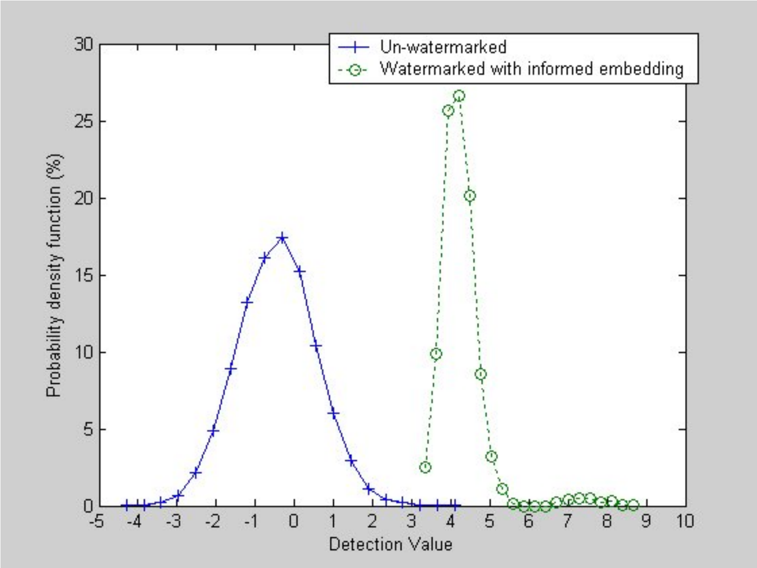
Even though the proposed algorithm originally focuses on the geometric attacks, however the performance of robustness to general attacks such as lossy compression, filtering, hybrid attacks is also very good. Table 4.1 lists all the tested attacked types. The correlation ratios of embedded “Lena” under these various attack types including some geometric distorted ones are summarized in Fig. 4.5. The correlation values are all above the threshold $\tau = 4.5$. Note that there is no false alarm reported for all the tested 720 images by setting threshold $\tau = 4.5$. More extensive tests on the robustness for geometric attacks are again evaluated using the abovementioned 4 sets of 720 images. The probability of recognition rate (P_D) with/without local optimal watermark selection and the false positive error detection probability (P_F) for various thresholds are shown in Fig. 4.6. The performance of proposed invariant watermarking with informed coding is far better than that without using informed coding. The detection performance without informed coding for “Goldhill” (Fig. 4.6(c)) and “Boat” (Fig. 4.6(d)) is even worse, i.e., detection rate is too low while false alarm rate is too high. More detailed discussions concerning geometric attacks are given below:

4.4.4.1 Translation

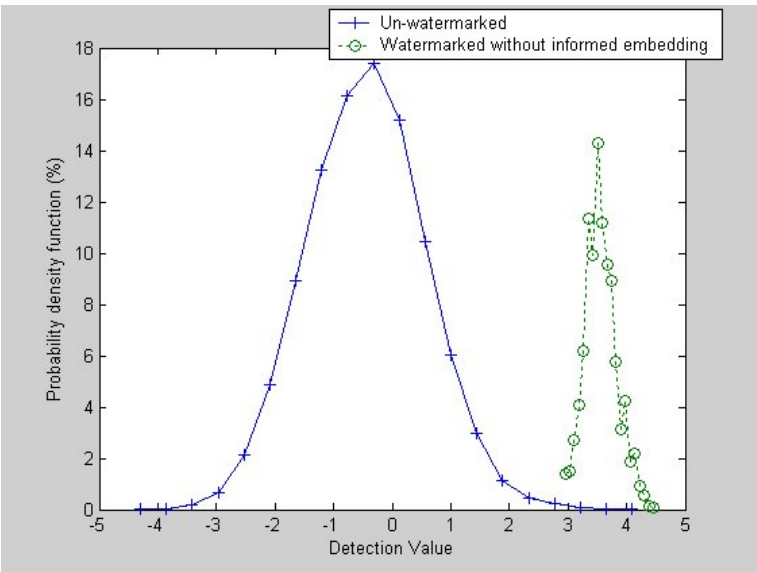
For translation attack, the distortion is resolved by the first round DFT transform. Various translations ranging from (0, 0) to (200, 200) are tested in this experiment. The detection response for this type of attack is quite good. When translated (200, 200), with 40% of the image pixels missed and the other 60% position shifted, the detection value (4.53) is still above the threshold $\tau = 4.5$. Note that for “Lena” image, there is no false alarm reported for all the tested 720 images by setting threshold $\tau = 4.5$.

4.4.4.2 Rotation

For rotation attack, the distortion is resolved by the LPM with successive second round DFT transform. Note again that the rotation here, being a translation along the radial axis μ in the polar coordinate, can be overcome by a DFT transform. For those attacked images whose media content information are lossless (rotation angle is 90^{o}, 180^{o}, 270^{o} or 360^{o}), the peak detection values (8.58 for rotation 90^{o}) are well above the threshold $\tau = 4.5$. Various rotation angles ranging from 0^{o} to 359^{o} are tested in this experiment. The algorithm performs well for arbitrary rotation angles.



(a)



(b)

Fig. 4.4: (a) Distributions of correlation values for watermarked / un-watermarked images with informed coding, (b) Distributions of correlation values for watermarked / un-watermarked images without informed coding.

Table 4.1: The tested attack types.

1. Translation (50, 50)	29. Cropping 10% and scaling to size 512 x 512 (stirmark3.1)
2. Translation (100, 100)	30. Cropping 5% and scaling to size 512 x 512 (stirmark3.1)
3. Translation (150, 150)	31. Cropping 1% and scaling to size 512 x 512 (stirmark3.1)
4. Translation (200, 200)	32. 1_row_1_col removal and scaling to size 512 x 512 (stirmark3.1)
5. Translation (75, 150)	33. 1_row_5_col removal and scaling to size 512 x 512 (stirmark3.1)
6. Translation (150, 75)	34. 5_row_1_col removal and scaling to size 512 x 512 (stirmark3.1)
7. Rotation 45°	35. FMLR (stirmark3.1)
8. Rotation 90°	36. JPEG Quality factor Q(%) = 30 (stirmark3.1)
9. Rotation 170°	37. JPEG Quality factor Q(%) = 40 (stirmark3.1)
10. Rotation 185°	38. JPEG Quality factor Q(%) = 50 (stirmark3.1)
11. Rotation 250°	39. JPEG Quality factor Q(%) = 60 (stirmark3.1)
12. Rotation 290°	40. JPEG Quality factor Q(%) = 70 (stirmark3.1)
13. Rotation 345°	41. JPEG Quality factor Q(%) = 80 (stirmark3.1)
14. Scale 1.05	42. JPEG Quality factor Q(%) = 90 (stirmark3.1)
15. Scale 1.10	43. JPEG-2000 0.25 bpp
16. Scale 1.20	44. JPEG-2000 0.50 bpp
17. Scale 1.25	45. JPEG-2000 1.00 bpp
18. Scale 0.90	46. Gaussian filtering 3x3 (stirmark3.1)
19. Scale 0.80	47. Sharpening 3x3 (stirmark3.1)
20. Scale 0.70	48. 2x2 Median filtering (stirmark3.1)
21. Scale 0.60	49. 3x3 Median filtering (stirmark3.1)
22. Scale 0.50	50. 4x4 Median filtering (stirmark3.1)
23. Rotation 2° and scaling to size 512 x 512 (stirmark3.1)	51. Random noise adding 2%
24. Rotation 5° and scaling to size 512 x 512 (stirmark3.1)	52. Random noise adding 4%
25. Scale 1.1 and Rotation 45°	53. Uniform noise adding 10%
26. Scale 1.1 and Rotation 190°	54. Uniform noise adding 15%
27. Translation (30, 30), Scale 1.1 and Rotation 30°	55. Histogram equalization
28. Translation (20, 20), Scale 0.9 and Rotation 45°	56. Hybrid

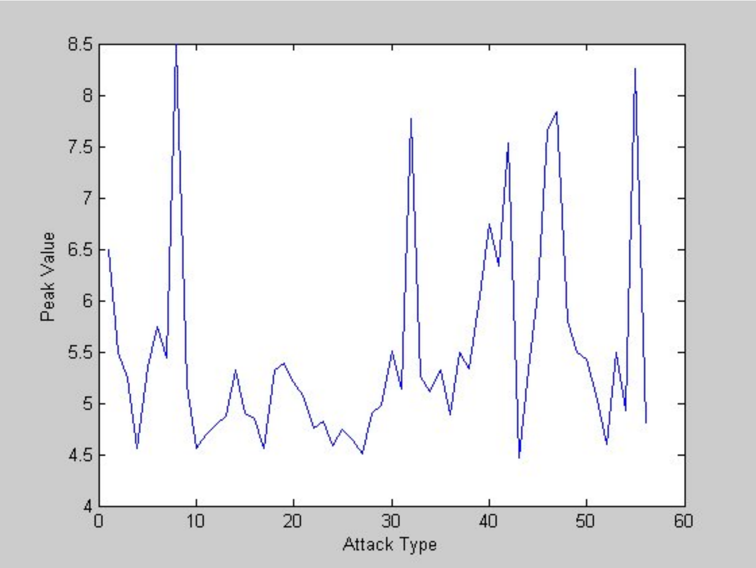
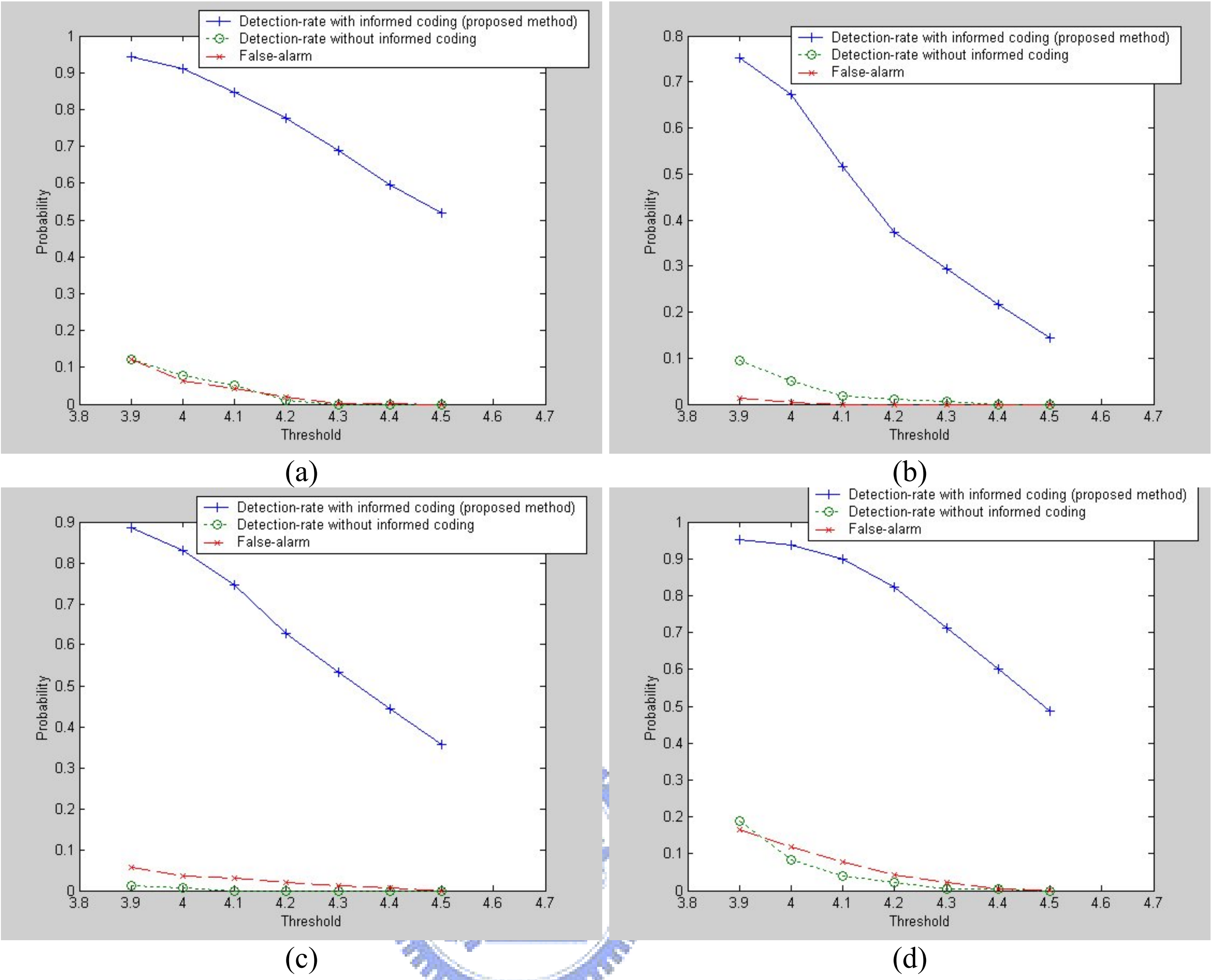


Fig. 4.5: Peak correlation values of various attack types as shown in Table 4.1 (All are above the threshold $\tau = 4.5$. Note that for "Lena" image, there is no false alarm reported for all the tested 720 images by setting threshold $\tau = 4.5$).

4.4.4.3 Scaling

For scaling attack, the distortion is resolved by the LPM with successive second round DFT transform. Note again that the scaling here, being a translation along the angle axis θ in the polar coordinate, can be overcome by a DFT transform. Various scaling ratio ranging from 0.5 to 1.5 are tested in this experiment. Better detection performance is achievable when scaling ratio is ranging from 0.8 to 1.2. Too large a scaling ratio has serious effects on the loss of media content, and thus probably deters the reliable detection.



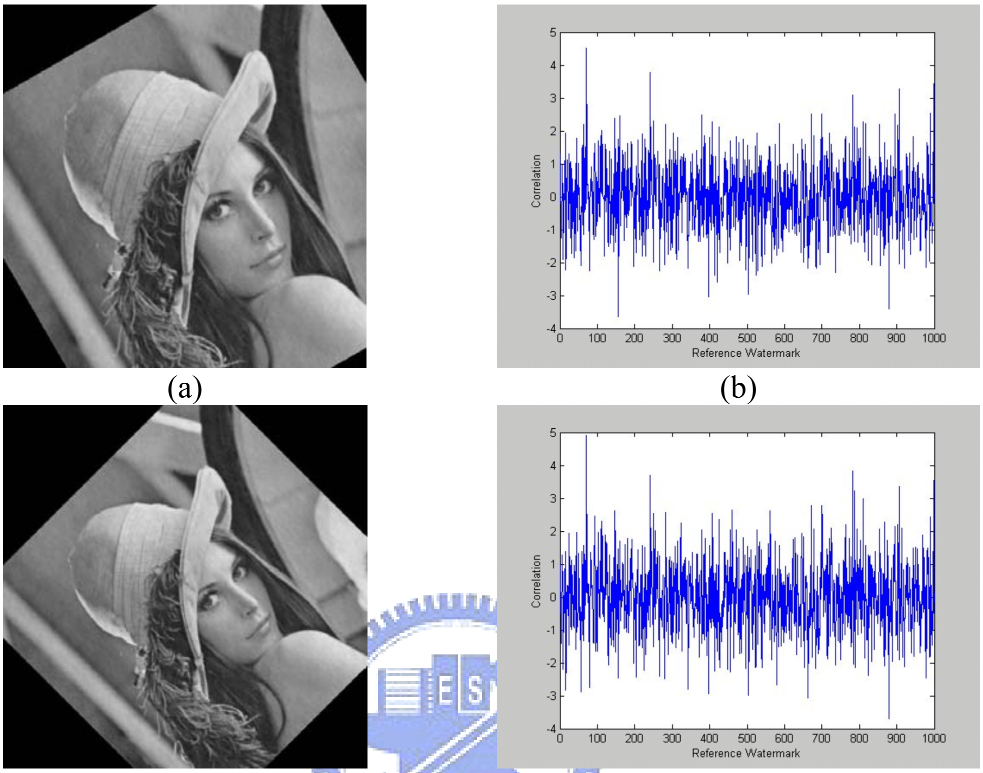
(a) (b)

(c) (d)

Fig. 4.6: The comparison of detection probability with/without informed coding and false positive error detection. (a) “Lena”, (b) “Barbara”, (c) “Goldhill”, (d) “Boat”.

4.4.4.4 Combined Rotation, Scaling and Translation

In the case of combined RST attack, the distortion is resolved by the first round DFT transform with successive LPM and second round DFT transform. Various combinations of rotation (0^{o} to 359^{o} with step-size 18^{o}), scaling (0.8 to 1.2 with step-size 0.5) and translation (0 to 15 with step-size 5) are tested in this experiment. Fig. 4.7 shows the combined RST attacked images and their detection values for two specific geometric attacks listed in Table. 4.1. With this type of single geometric attack (translation, scaling and rotation occurs alone), the detection is easy. However, in the case of combined attack, the detection is more difficult since distortion increases dramatically.



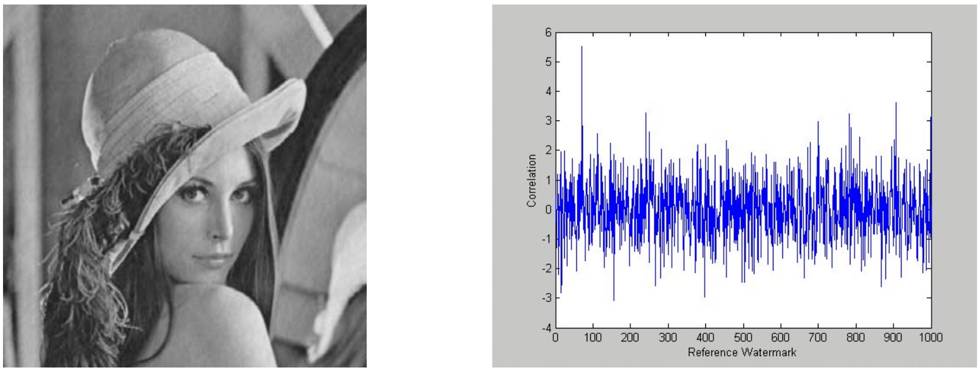
(a) (b)

(c) (d)

Fig. 4.7: (a) Image attacked by translation (30, 30), rotation 30^o, scale 1.1, (b) Detection response (Peak value 4.51), (c) Image attacked by translation (20, 20), rotation 45^o, scale 0.9, (d) Detection response (Peak value 4.90).

4.4.4.5 Cropping

The cropped image is resized to its original size before detection. The affecting factors are pixels shifting and scaling. Various cropping ratio ranging from 0% to 15% are tested in this experiment. Fig. 4.8 shows the attacked image and its detection value. In the case of cropping, the missed portions are at the surroundings of the image. Note that the cropping attack is also a geometric distorted type and hence is difficult to detect. However, our proposed algorithm performs well for cropping below 10%.



(a) (b)

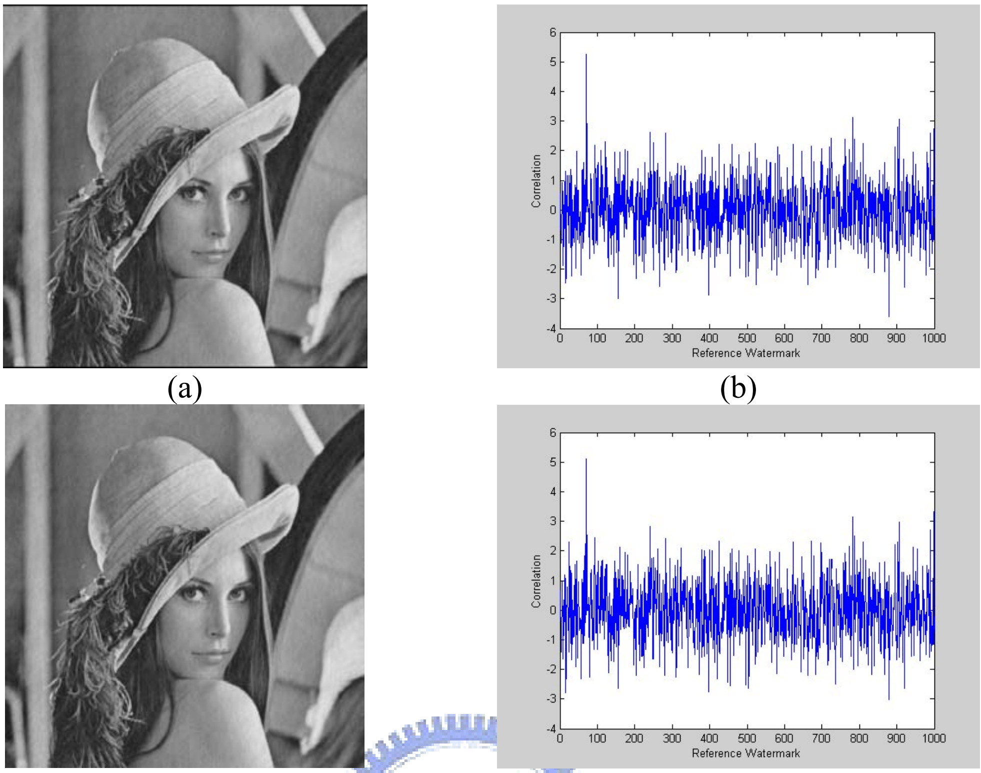
Fig. 4.8: (a) Image attacked by cropping 5%, (b) Detection response (Peak value 5.51).

4.4.4.6 Symmetric row and column removal

In the case of asymmetric row and column removal, some pixels at specified row and column are removed from the image, and the distortions are due to clipping and scaling. The row and column removed image is also resized to its original size before detection. The tested images are generated by Stirmark3.1 with maximum 5 rows or 5 columns removed. Fig. 4.9 shows the attacked images and their detection values. Note that the cropping attack is also a geometric distorted type and hence is difficult to detect. Our proposed algorithm also performs well for removing either 5 rows or 5 columns.

4.4.4.7 Hybrid

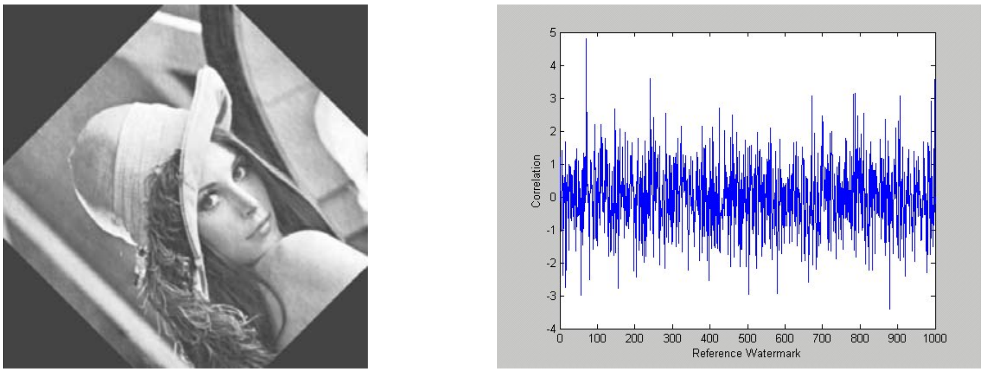
The PaintShopPro7.0 is employed to further manipulate the attacked image in Fig. 4.9(c). The image is translated by the offset (20, 20), rotated by 45^{o}, scaled by 0.9, and furthermore attacked by Gaussian blur, sharpening, and finally histogram equalization. The hybrid attacked image and its detection value are displayed in Fig. 4.10. Even in this critical case our informed coding based spread spectrum watermark is still quite robust subjected to this attack.



(a) (b)

(c) (d)

Fig. 4.9: (a) Image attacked by removing 1 row and 5 columns, (b) Detection response (Peak value 5.26), (c) Image attacked by removing 5 rows and 1 column, (d) Detection response (Peak value 5.11).



(a) (b)

Fig. 4.10: (a) Hybrid attacked image, (b) Detection response (Peak value 4.80).

4.5 Summary

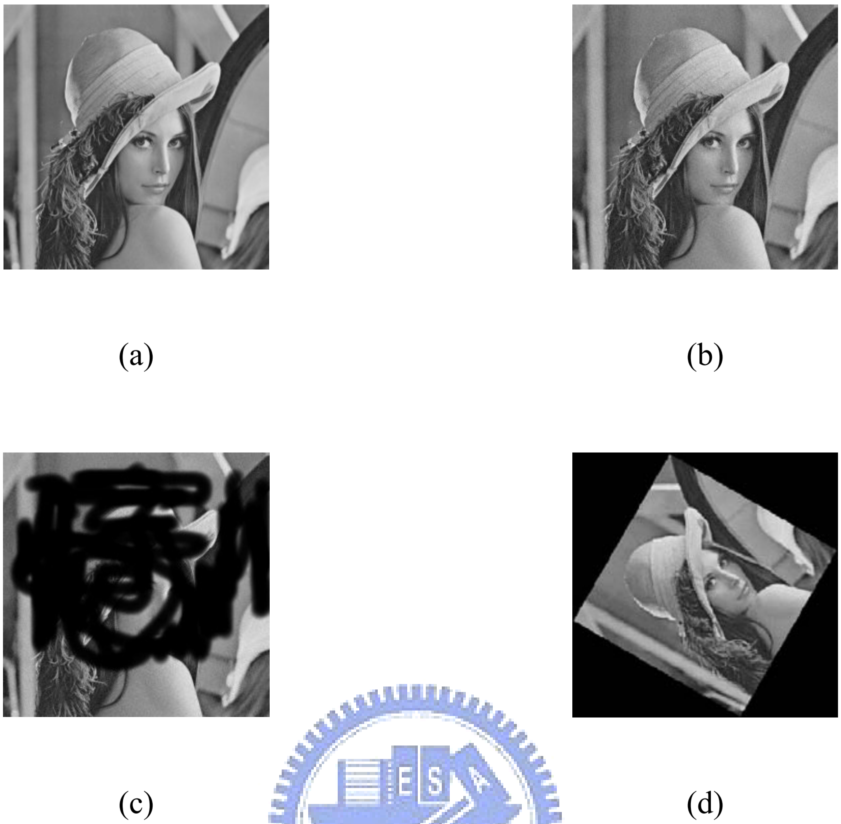
We propose in this chapter an RST resilient image watermarking technique using Fourier-Mellin transform and informed coding of watermark message. The watermark is embedded in the geometric invariant Fourier-Mellin domain, and no additional features need to be extracted to form a geometric invariant embedding space. Moreover, by informed watermark coding, our scheme could embed a weak watermark signal (i.e. one that needs only small perturbations with the host signal) and detect a slightly weaker watermark under the ILPM and the inverse Fourier transform. The main contributions of the proposed scheme are: (1) to increase the detection robustness, a robust watermark via informed coding is chosen prior to embedding, (2) to increase transparency, only the circular parts of an image with radius no more than half the image row size are LPM transformed, and other pixels of the image remain intact. Experimental results demonstrate that the resulting watermark is robust to a variety of image processing attacks, including RST, cropping, asymmetric row and column removal, frequency mode Laplacian removal (FMLR), DCT based compression (JPEG), DWT based compression (JPEG-2000), Gaussian filtering, sharpening, median filtering, noise adding, histogram equalization and hybrid manipulations. As most of the quality degradation of watermarked images in such approach is due to the LPM and ILPM, finding a more precise algorithm on the LPM and ILPM would be a very interesting topic for further research. To increase robustness, a robust watermarking space could be combined with our proposed robust watermark approach to further enhance the robustness. Furthermore, the structured dirty-paper codes [21] such as lattice codes should be further developed for exploring the multiple-bit watermark capacity.

Chapter 5 Self-Synchronized QR-Coded Watermark Detection

The previous geometric invariant watermarking schemes by adopting either resynchronization templates or invariant features suffer from both limited payload and reliable detection. The payload, which is typically embedded in 256 x 256 sized cover image, ranges from as small as one bit up to at most 64 or 128 bits. In addition, due to the difficulties of either accurately estimating the geometric distortion or finding perfect invariant features, the reliabilities of these schemes are hence decreased. To overcome the problems of limited capacity and reduced reliability, we employ the famous QR Code (Quick Response Code) by first encoding the watermark payload, and then embedding the QR coded watermark into the image spatial domain. Thanks to the characteristic of position detection pattern of QR Code, the self-synchronized QR coded watermark payload can be recovered against geometric distortions once the QR Code is extracted during detection. Extensive experimental results demonstrate that by adopting our approach, the resulting watermark is robust to a variety of combinations of RST (rotation, scaling and translation) attacks while preserving the visual quality of the watermarked image, thereby resolve the unavoidable dilemma faced by the other schemes.

5.1 Introduction

Although a great deal of effort has been made on the study of geometric invariant watermarking, the reported performance is not yet convincing for use in practical applications, such as print-and-scan or print-and-snapshot. For example, one of the best state-of-the-art schemes [37] claimed to have a very good result which has a payload of 64 bits only with the required PSNR of about 38 dB. However, the reliability was not confirmed by any publicly accessible software. Meanwhile, the watermark software “PictureMarc” [38, 1] developed by the Digimarc Corporation becomes one of the most successful products on watermarking which has been adopted by many commercial software packages. To verify the robustness of PictureMarc, we used PictureMarc Demo plug-in attached in PhotoImpact software. We embedded copyright year information “2006”, which is no more than 16 bits, with 3-levels of watermark strength (durability 2, 3 and 4, respectively; the more the durability, the more the robustness but less the visibility) into “Lena” image to obtain 3 watermarked images. The robustness to common image processing attacks (filtering, compression, blurring and rotation) is quite impressive. However, for combined rotation and scaling, the performance is very poor for embedding durability 2, and is unreliable for embedding durability 3 and 4. Fig. 5.1(a) and 5.1(b) show the original and watermarked “Lena”, while Fig. 5.1(c) and 5.1(d) illustrate the strength and weakness of PictureMarc detection. It is evident that PicutreMarc is quite fragile to the combined rotation and scaling attack.



(a) (b)

(c) (d)

Fig. 5.1: The strength and weakness of PictureMarc by DigiMarc Cooperation: (a) original “Lena” (256 x 256), (b) watermarked “Lena” with message “2006” embedded, the PSNR was 34 dB using durability 4 (less visible but most durable), (c) robust to graffiti attack, (d) fragile to combined rotation and scaling attack.

Having recognized that the combined RST attacks remain challenging, this chapter aims at improving the robustness against combined RST attacks while preserving the picture quality of watermarked image. Most of the self-synchronized watermarking schemes rely on the accuracy of the re-synchronization peaks detection to recover geometric transformations. And then employ spread-spectrum based scheme for watermark detection. However, the payload by adopting this approach is quite limited and the detection accuracy is un-reliable [51, 52]. Since payload capacity of watermark embedding based on quantization is rather high [7, 53], we intend to explore this topic from a different perspective, specifically, by combining the strength of quantization based approach and QR Code [2] (one kind of two-dimension barcode). While in the watermark area, a solution having both the characteristic of geometric invariant and proper capacity is not yet within reach, in the area of barcode detection solutions with the above-mentioned characteristics are well developed. We therefore are motivated to introduce the concept of QR Code to shed a light on our problem. Section 5.2 discusses the proposed QR Coded watermark generation, embedding, and detection algorithm. In Section 5.3 experimental results are presented and analyzed. Finally, in Section 5.4 summary as well as future work are given.

5.2. Proposed QR Coded watermarking algorithm

As the characteristics of QR Code are proven outstanding [2], our intuition was that some of its properties can be explored and even employed to develop a better watermark scheme. As the design of QR Code is quite intricate, in our scheme we developed only a simplified version and propose a two-stage watermarking scheme to meet our requirements as shown in Fig. 5.2. In stage 1, the watermark message is first coded into the simplified QR Code. Stage 2 then embeds (or detects) the self-synchronized QR coded watermark to resist geometric distortions. For embedding and detection, we adopt QIMM (Quantization Index Modulus Modulation) [53], which has been both theoretically proved and empirically verified to be highly competitive against QIM (Quantization Index Modulation) in detection robustness under the same MSE (Mean Square Error) condition. The details are described as follows:

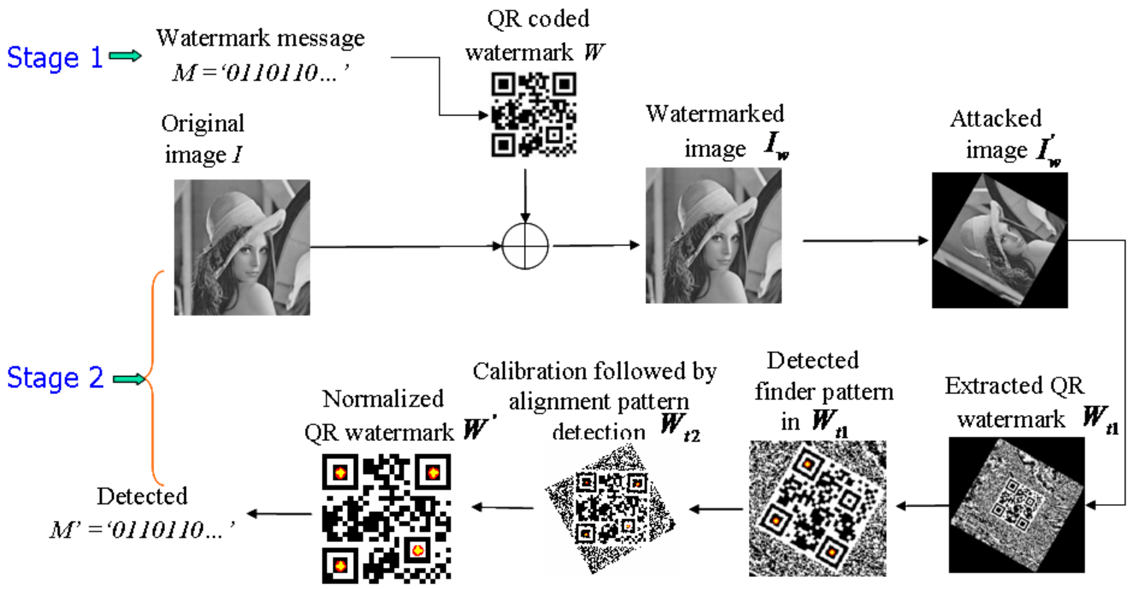
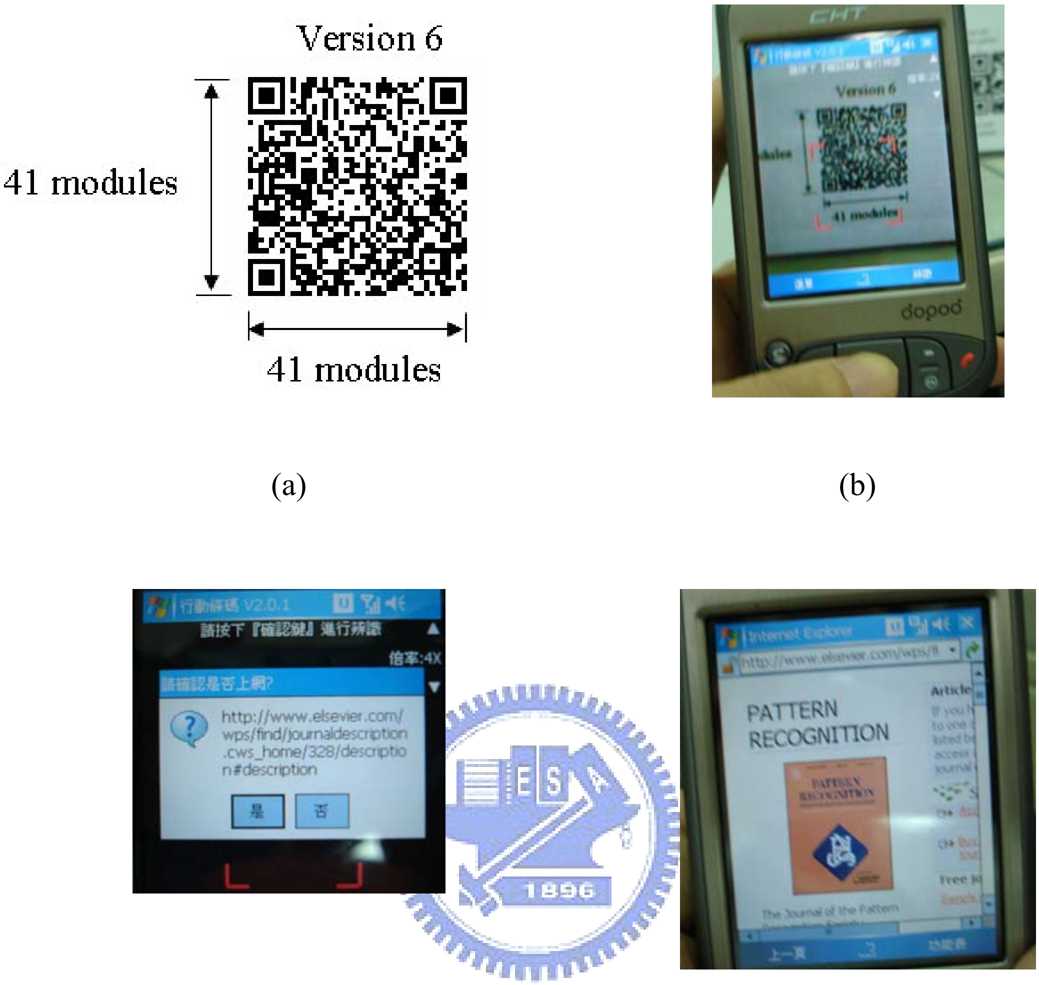


Fig. 5.2: The flow of proposed watermark embedding / detection scheme.

5.2.1 Concept of QR Code

QR Code [2] (developed by Denso Co.) is an open two-dimensional symbology consisting of nominally black and white square cells (modules). The features of QR Code lie in its storage of high capacity of information, error-correction ability, 360 degree (omni-directional) reading, and high speed reading. The encoding capacity and reading ability depend very much on the configuration of QR Code version, module size and error correction level. The versions of QR code defined in its original specification [2] range from Version 1 (21 x 21 modules) to Version 40 (177 x 177 modules), in which bigger version indicates larger encoding capacity. To increase the code readability, either a higher error correction level should be set, or each cell should contain several image pixels at the expense of code area. Several kinds of messages can be encoded into QR Code to perform specific functionality, such as web-to, phone-to, SMS-to, mail and phonebook registration. Fig. 5.3(a) shows one sample of QR Code which encodes the message

“//www.elsevier.com/wps/find/journaldescription.cws_home/328/description#description”. A mobile-phone equipped with QR Code recognizer can take a snapshot of the code (Fig. 5.3(b)); and once the code is recognized, the recognizer will ask the user whether to connect the web (Fig. 5.3(c)); upon confirmation the screen will be directed to the web page of The Journal of the Pattern Recognition Society (Fig. 5.3(d)). By capturing the QR code, it facilitates reader’s access to the web page as it does not require typing the URL.



(a) (b)

(c) (d)

Fig. 5.3: (a) One sample of QR Code which encodes the URL "http://www.elsevier.com/wps/find/journaldescription.cws_home/328/description#description", (b) Take a snapshot of the code, (c) Recognized result and ask for confirmation, (d) Directed to the Pattern Recognition web page.

5.2.2 QR Coded Watermark

We employ both finder pattern and alignment pattern of QR Code symbol in designing our watermark pattern. Fig. 5.4(a) shows the simplified QR coded watermark *W*, which is a binary image in 25 x 25 modules, consisting of 3 identical position detection patterns (so-called finder pattern), 1 alignment pattern and 384 bits watermark payload. The module size defined in QR Code specification is the vertical or horizontal millimeter size of each printed code cell. However, for our purpose we re-define module size, say *p*, to be the vertical as well as horizontal pixels number of each cell. Therefore, for module size *p*, the size of *W* is 25 x *p* x 25 x *p* pixels. The finder pattern is designed in a way so that it has very low probability to appear elsewhere in the symbol other than at the specified 3 locations that are used to define the location and orientation of the symbol. Fig. 5.4(b) shows the structure of position detection pattern, with size 7 x 7 modules, and with the characteristic of a dark-light-dark-light-dark (b-w-b-w-b) sequence with relative widths of each dark or light element in the ratios 1:1:3:1:1, independent of the direction it is scanned. The alignment pattern (Fig. 5.4(c)) is similar to that of position detection pattern but with the size of 5 x 5 modules, and with the characteristic of a b-w-b-w-b sequence in the ratios 1:1:1:1:1. The alignment pattern accompanied with the 3 position detection patterns are used as reference points for watermark symbol calibration and normalization.

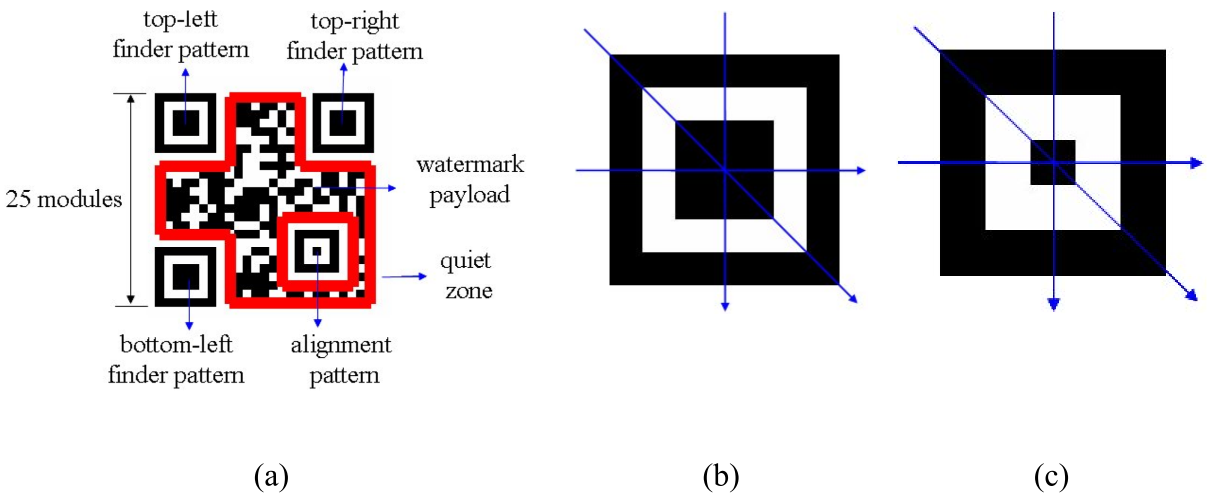


Fig. 5.4: (a) Simplified QR code, (b) Characteristic of position detection pattern: module width in each position detection pattern is constructed of a dark-light-dark-light-dark sequence with relative widths of each element in the ratios 1:1:3:1:1, (c) Characteristic of alignment pattern: module width in alignment pattern is constructed of a dark-light-dark-light-dark sequence with relative widths of each element in the ratios 1:1:1:1:1.

5.2.3 QR Coded Watermark Embedding

As we employ additional 4 light corner bars (so called quiet zone) surrounding the 25 x 25 modules sized watermark to efficiently find the location of finder pattern, we use central 27 x 27 modules of the image as host signal $X = \{x_1, x_2, ..., x_n\}$ for embedding, and where *n* equals to 27 x *p* x 27 x *p* and *p* is the module size. The embedding algorithm consists of the following steps:

Step 1: Each element of X is first divided by the quantization step size (δ), and a nearest integer index value is obtained by a round function. The quantized index value is then executed with modulo 2 to get the residue with value 0 or 1.

If the residue is equal to the watermark message bit, then the watermarked value is the reconstruction point of quantized host signal. Otherwise, the biased (either +1 or –1) quantized index value is used to calculate the watermarked value $X^{'} = \{x_1^{'}, x_2^{'}, ..., x_n^{'}\}$.

To embed one bit of watermark payload *m*, the embedding algorithm consists of the following steps:

a) Take $Q(x_i) = Round(x_i / \delta)$.

b) If $(Q(x_i) \bmod 2) = m$ then

$$x_i^{'} = s(x_i; m) = Q(x_i) * \delta, \qquad (5.1)$$

else

$$x_i^{'} = s(x_i; m) = \arg\min_{P(x_i)} \|(P(x_i) - x_i\|, \qquad (5.2)$$

where $P(x_i)$ in Eq. (5.2) is either $(Q(x_i) - 1) * \delta$, or $(Q(x_i) + 1) * \delta$, and $s(x_i;m)$ is the i_{th} host signal embedded with watermark message *m*. The criterion of selecting either $(Q(x_i) - 1) * \delta$ or $(Q(x_i) + 1) * \delta$ depends on which one has less distortion with respect to x_i. The one with less distortion is used to reconstruct the watermarked signal $x_i^{'}$.

Step 2: Replace the corresponding 27 x 27 modules of *I* with $X^{'}$ to obtain the watermarked image I_w.

5.2.4 QR Coded Watermark Detection

The detection algorithm consists of the following steps:

Step 1: Extract the QR coded watermark W_{t1} (Fig. 5.5(b)) from attacked image $I_w^{'}$ (Fig. 5.5(a)): All the elements of the attacked watermarked image $I_w^{'} = \{x_1^{''}, x_2^{''}, ..., x_r^{''}\}$, where *r* equals 256 x 256, are used for detection.

Each element of $I_w^{'}$ is also divided by the quantization step size, so that a nearest integer index value is obtained by a round function. The quantized index value is then taken modulo 2 to obtain the QR coded watermark bit m^*. To extract one bit of watermark payload m^*, the detection algorithm consists of the following two steps:

a) $Q(x_i^{"}) = Round(x_i^{"} / \delta)$.

b) $m^* = Q(x_i^{"}) \bmod 2$.

Step 2: Detect finder pattern in W_{t1} as shown in Fig. 5.5(c) by finding the intersections of row-wise and column-wise line segments which are of the ratio 1:1:3:1:1 for b-w-b-w-b sequence.

Step 3: Calibrate W_{t1} to obtain W_{t2} and then detect alignment pattern as shown in Fig. 5.5(d). Finally, 4 reference points including 3 center points of top-left, top-right and bottom-left finder pattern and one center point of bottom- right alignment pattern are obtained.

Step 4: Obtain normalized QR watermark $W^{'}$ as shown in Fig. 5.6(a). This is achieved by projective transforming [54] the 4 reference point pairs between $W^{'}$and W_{t2} as follows.

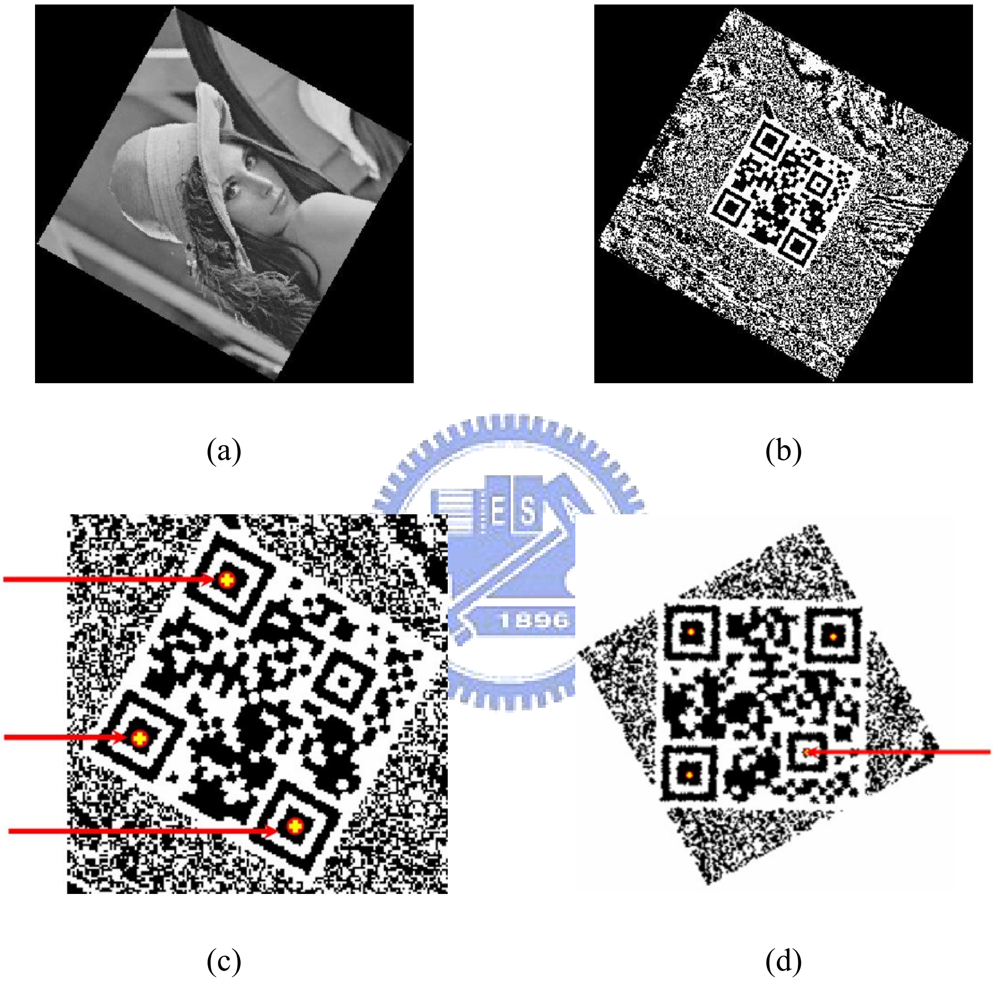
Given four pairs of reference points (u_i, v_i) in $W^{'}$and (x_i, y_i) in W_{t2}, where $i = 1,...,4$. We have the equations:

$$x = \frac{au + bv + c}{gu + hv + 1}, \qquad y = \frac{du + ev + f}{gu + hv + 1}. \tag{5.3}$$

After solving the parameters a to h by substituting the above relationship in the four corresponding pairs, all the pixels positions in $W^{'}$can be obtained by mapping the corresponding points in W_{t2}. As compared to that of original

QR coded watermark (Fig. 5.6(b)), though one bit of the bottom-left finder pattern was wrong in decoded QR watermark (Fig. 5.6(a)), all the watermark payload bits were correct.

Step 5: Extract the embedded 384 watermark message bits from $W^{'}$.

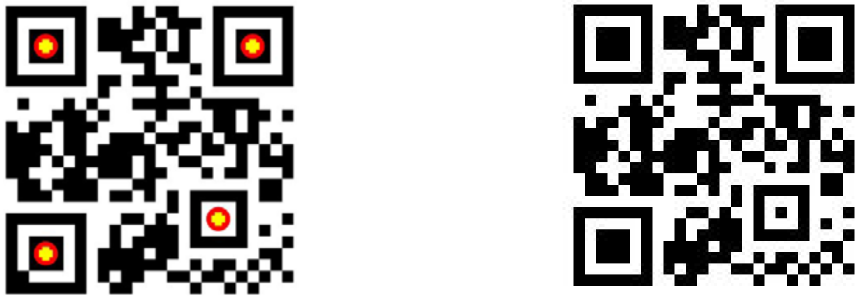

(a)

(b)

(c)

(d)

Fig. 5.5: The calibration process of extracted QR watermark: (a) attacked image, (b) extracted QR watermark, (c) detected finder pattern, and (d) calibration followed by alignment pattern detection.



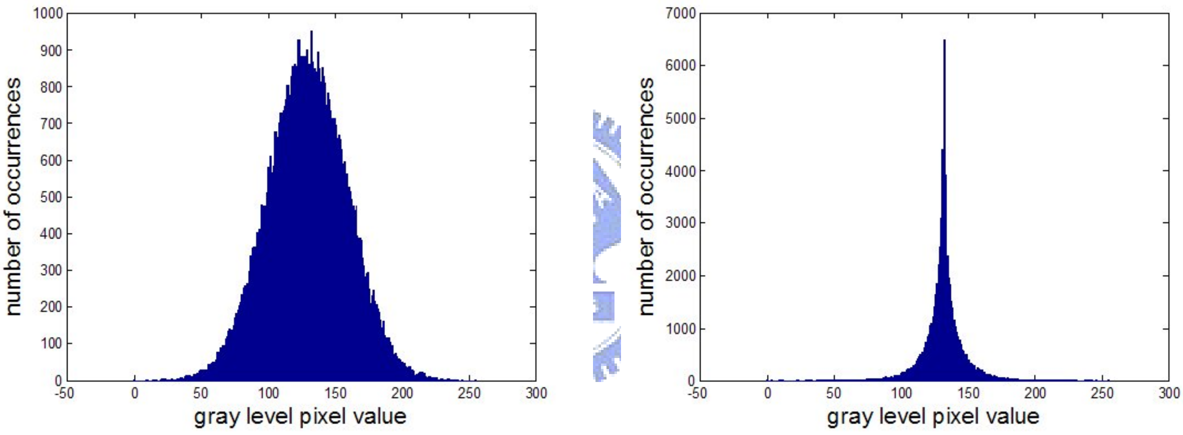
(a) (b)

Fig. 5.6: Recovered and original QR watermark: (a) normalized QR decoded watermark, (b) original QR coded watermark.

5.3 Experimental Results

To evaluate the effectiveness of the proposed method, experimental simulations on two real images (Lena, Barbara) as well as both of the Monte Carlo simulated Gaussian and non-Gaussian images were utilized as the host signals to embed the QR coded watermark. For each run of Monte Carlo Gaussian simulation, host signals X drawn from 256 x 256 samples of a Gaussian zero-mean random variable were generated, each having standard deviation σ_X (we will later denote it by $N(0,\sigma_X)$.), ranging from 10 to 100 with step size 10. As for each run of Monte Carlo non-Gaussian simulation, Gaussian Scale Mixture [55-59] host signals X is defined as $\sqrt{Z}.U$, where $\sqrt{Z}$ is a hidden positive multiplier vector and is approximated by absolute Gaussian vector $|N(0,\sigma_Z)|$; whereas U is approximated by Gaussian vector $N(0,\sigma_U)$. The reason for employing Gaussian Scale Mixture to model non-Gaussian data is that this model is proved to be effective in modeling natural scene statistics [55-59].

All these Gaussian and non-Gaussian simulated data were then normalized to the range of 0~255 to simulate Gaussian and non-Gaussian gray level images. Fig. 5.7(a) and 5.7(b) show both the histograms of simulated gray level Gaussian and non-Gaussian images. Although the researchers [55-59] have shown that the Gaussian Scale Mixture performed better by modeling wavelet coefficients, here we simply employ Gaussian Scale Mixture to modeling the non-Gaussian statistics of spatial pixels, and the result is quite satisfactory. Indeed we can see that the first is of a bell shape while the second is a heavy-tailed one, which is non-Gaussian.



(a) (b)

Fig. 5.7: Two types of simulated data: (a) Histogram of simulated gray level Gaussian image, (b) Histogram of simulated gray level non-Gaussian image.

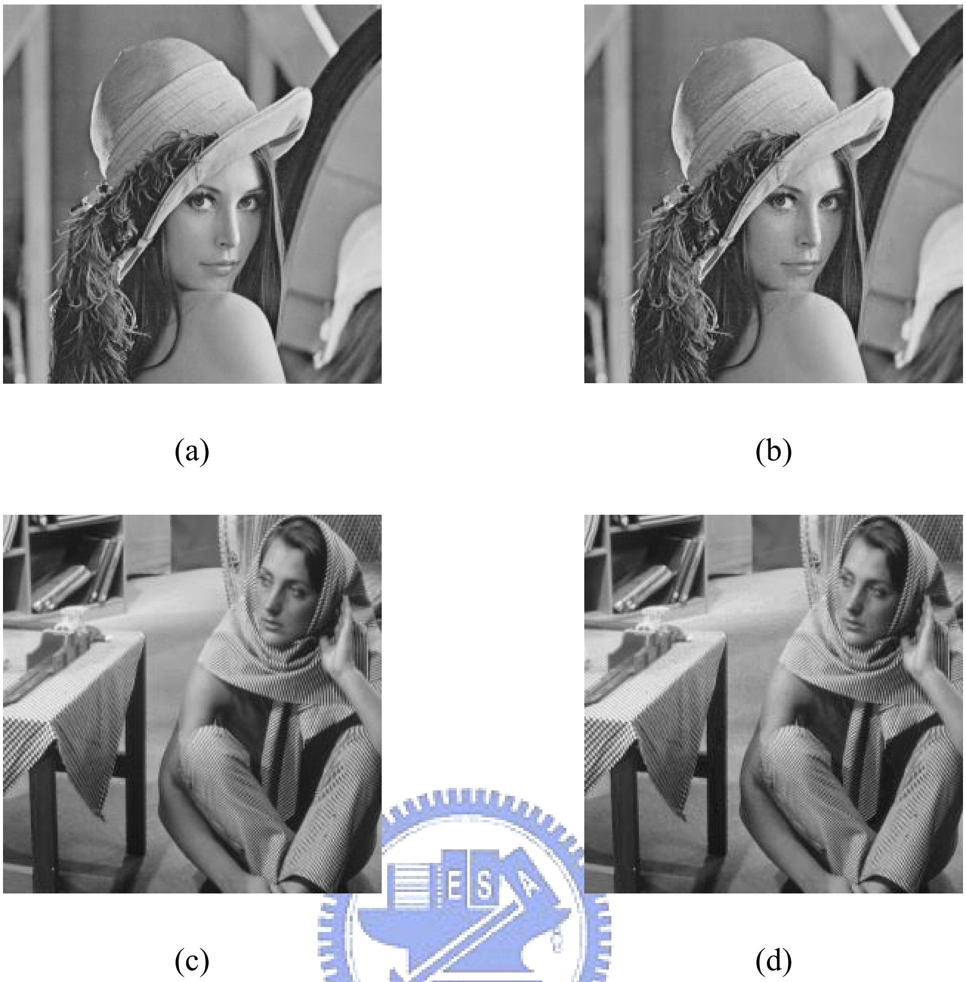
As we embedded the watermark into the central part of the image, it is evident that the robustness of detection was un-affected if the corresponding embedding region of image existed, no matter how the image was translated or cropped. To test the robustness of rotation and scaling, the attacked images were obtained by manipulating the watermarked image with various combinations of rotation (20^o to

360^{o} with step-size 20^{o}) and scaling (0.5 to 1.5 with step-size 0.1) to generate 198 images in total. For "Lena" and "Barbara" images, we further performed 10 times on each of the above simulations using different QR coded watermarks, so that totally 1980 test images for each group (Lena, Barbara, Gaussian, Non-Gaussian) were employed to obtain the average detection ratios.

From our experiments, the degree of PSNR dropped when the module size *p* increased. A larger module size brought more robustness, but it also introduced more distortion. In our setting, the module size *p* of QR coded watermark was set to 4, 5 and 6, respectively, to compare the robustness. Meanwhile the embedding quantization step size δ was set to 10 to ensure the transparency of watermarked image. The average PSNR of each group of 10 watermarked images with module size *p* = 4, 5, and 6 were about 46.5 dB, 44.4 dB and 42.9 dB, respectively. The details were summarized in Table 5.1. The original and some samples of watermarked images were shown in Fig. 5.8 and Fig. 5.9. The watermarked "Lena" and "Barbara" embedded with module size *p* = 6 were shown in Fig. 5.8(b) and 5.8(d), respectively. While one sample of the watermarked Gaussian and non-Gaussian image embedded with module size *p* = 6 were shown in Fig. 5.9(b) and 5.9(d), respectively. Even for such strong watermark embedding, the watermarked images are nearly indistinguishable from their original ones.

Table 5.1: The average PSNR of each group of watermarked images (dB)

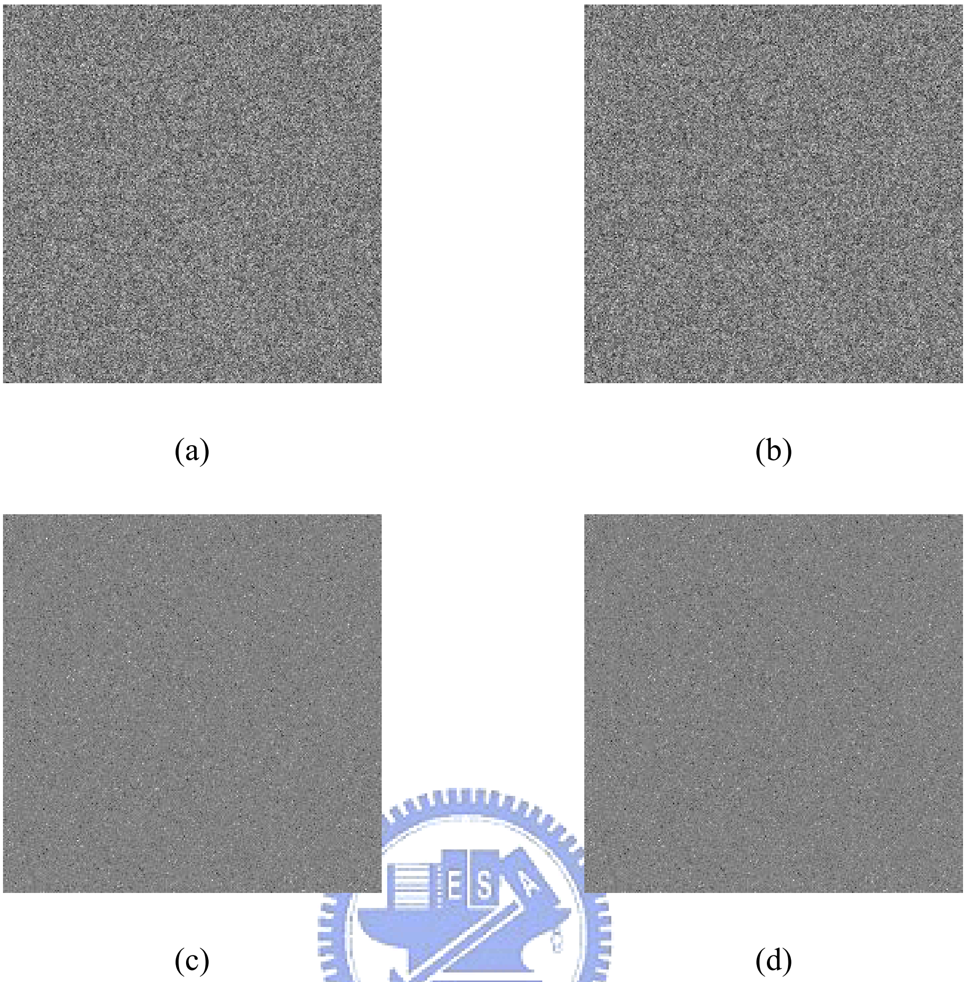
Images\Module size	p = 4	p = 5	p = 6
Lena (10 images)	46.5	44.5	42.9
Barbara (10 images)	46.5	44.5	43.0
Gaussian (10 images)	46.5	44.4	42.9
Non-Gaussian (10 images)	46.3	44.3	42.8



(a) (b)

(c) (d)

Fig. 5.8: Original and watermarked images: (a) original “Lena” (256 x 256), (b) watermarked “Lena” with 384 bits embedded, the PSNR was 42.9 dB using module sized 6 (less visible but more durable), (c) original “Barbara” (256 x 256), and (d) watermarked “Barbara” with 384 bits embedded, the PSNR was 43.0 dB using module sized 6 (less visible but most durable).



(a) (b)

(c) (d)

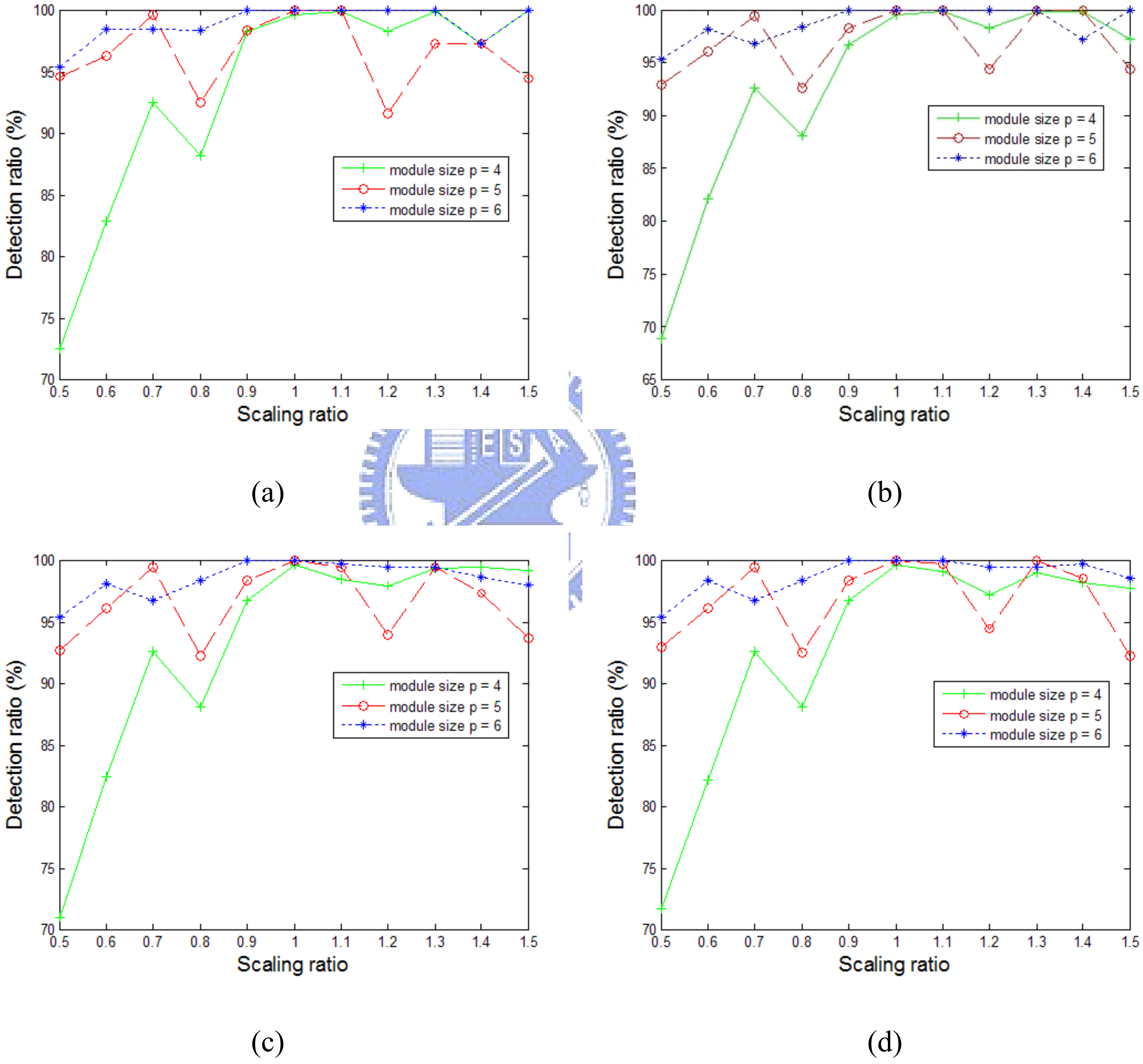
Fig. 5.9: Original and watermarked simulated images: (a) one sample of original Gaussian images (256 x 256), (b) watermarked Gaussian image with 384 bits embedded, the PSNR was 42.9 dB using module sized 6, (c) one sample of original non-Gaussian images (256 x 256), and (d) watermarked non-Gaussian image with 384 bits embedded, the PSNR was 42.8 dB using module sized 6.

To evaluate the reliability of watermark detection, for total number of correctly detected bits N_C and total number of embedded bits N_T , the detection ratio ρ was defined as $\rho = \frac{N_C}{N_T} \times 100\%$. A higher value of ρ indicated a more reliable detection. The perfect recognition rate could be achieved when the value of ρ equals 100%.

Table 5.2 summarizes the average detection ratios of each group of 1980 attacked images of module size p = 4, 5, and 6, in which the ratios were about 93.8%, 97.2% and 98.8%, respectively. Two sets of statistics were further collected to differentiate the robustness against scaling and rotation: one for various rotation angles under specific scale and the other for various scales for specific rotation angle. Fig. 5.10(a) ~ 5.10(d) illustrates the distributions of average detection values to each specific scale ratio and Fig. 5.11(a) ~ 5.11(d) to each specific angle for these 1980 tested images. From Fig. 5.10, we can see that the detection is less affected by up-scaling than by downscaling. This is because up-scaling the image can be viewed as increasing the module size of QR coded watermark, which is beneficial to the detection. However, since the interpolation of pixel through up-scaling might still introduce small error, this process will slightly worsen the detection. As for downscaling, specifically for downscaling ratio equal to 0.5 when p = 4, in which case the detection ratio declines dramatically and can be as small as 70%. We note that this situation can be overcome by using larger module size QR coded watermark, e.g., module size p = 6, in such case the ratio can be boosted to 95%. When the scaling is lager than 0.7, we see that the detection ratios were all above 85%, which presents a lesser decline compared to the result when scaling is 0.5. On the other hand, the detection ratios were all above 80% for almost all of the tested rotation angles. In summary, it is clear from aforementioned analysis that the detection is more robust to rotation than downscaling.

Table 5.2: The average detection ratio of each group of images (%)

Images\Module size	p = 4	p = 5	p = 6
Lena (1980 images)	95.8	98.8	99.5
Barbara (1980 images)	93.3	96.8	98.6
Gaussian (1980 images)	93.2	96.5	98.6
Non-Gaussian (1980 images)	92.9	96.7	98.5



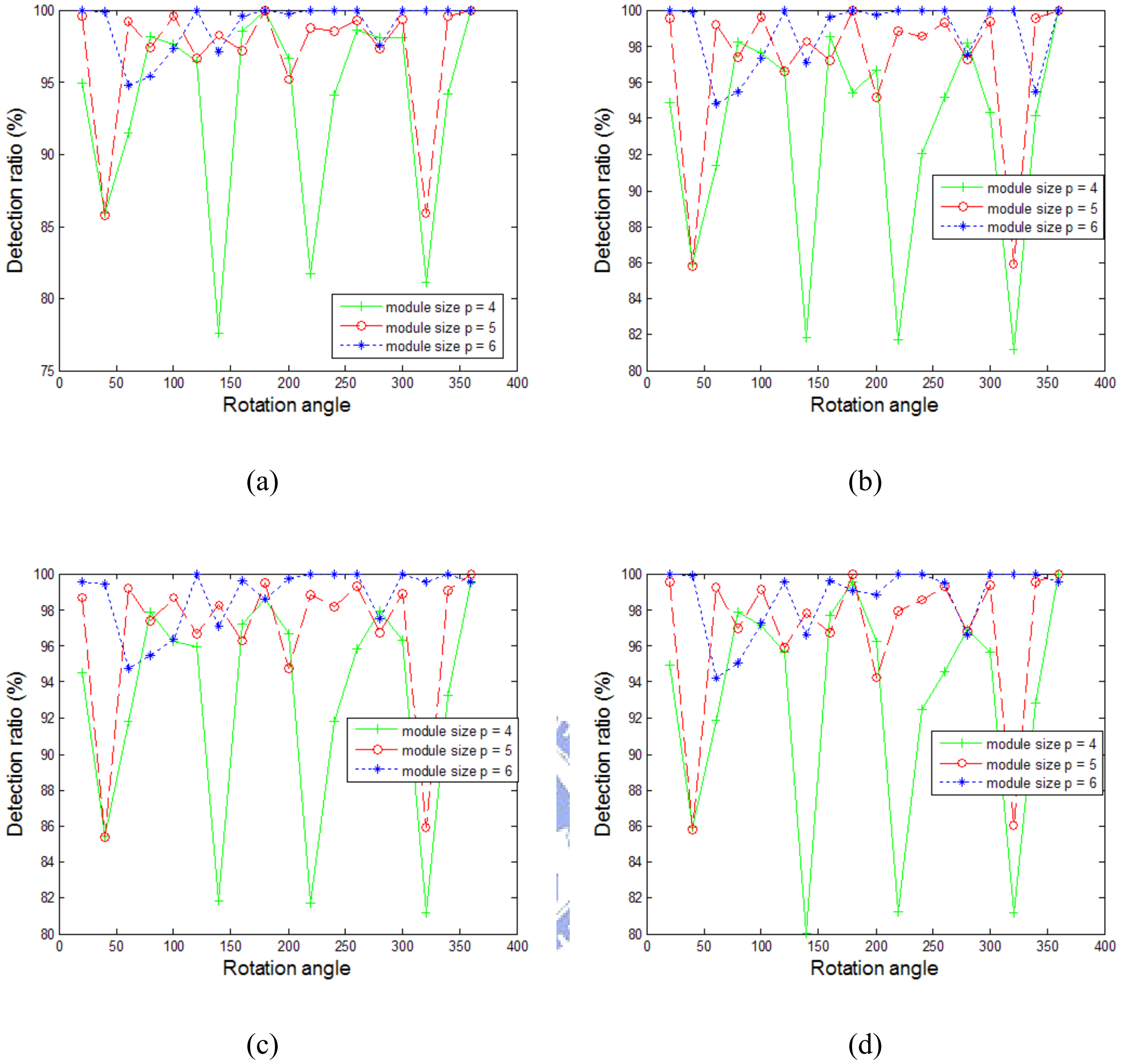
(a)

(b)

(c)

(d)

Fig. 5.10: The distributions of average detection ratio (1980 images) to each specific scaling ratio: (a) “Lena”, (b) “Barbara”, (c) Gaussian images, and (d) non-Gaussian images.



(a)

(b)

(c)

(d)

Fig. 5.11: The distributions of average detection ratio (1980 images) to each specific rotation angle: (a) “Lena”, (b) “Barbara”, (c) Gaussian images, and (d) non-Gaussian images.

We observed that the major cause for false detection is the failure to identify the three finder patterns. Therefore, one way to enhance the detection ratio is by focusing on the improvement of the finder pattern detection. On the other hand, the detection ratio can also be further enhanced by employing error correction coding module (such as BCH or Reed-Solomon) in our scheme.

Because of the recently increasing demands for mobile phone camera based barcode/watermark applications [60-62], the technique for barcode detection has evolved rapidly and commercially acceptable products are generated. However, fulfilling the requirements of reliable watermark detection remains more challenging, since barcode is a visible pattern whereas watermark is an invisible pattern and is interleaved with another signal. A more general purpose robust watermarking method which is capable of real-time detection for low-cost camera-equipped mobile phones remains an interesting topic for further study.

5.4 Summary

We propose in this chapter a new approach for geometric invariant watermarking technique by embedding QR coded watermark in spatial domain of image coupled with detecting the self-synchronized watermark to combat the geometric attacks. The scheme explores the features of QR Code for watermark self-synchronization. In addition, it finds a viable choice for the robust QR coded watermark embedding which will maximize the detection while preserving the visual quality of watermarked image. Experimental results demonstrate that by adopting our approach, the resulting self-synchronized watermark can be accurately detected under a variety of combined RST attacks. In the future, we expect to explore further to combat other attacks, especially noising, blurring and compression which commonly occur in mobile phone camera based applications. To this end, the watermark embedding in frequency domain instead of spatial domain might serve a good alternative. Also, designing a better finder pattern for self-synchronized watermark embedding by innovating or improving some other existing barcode technology (Visual Code [63], Data Matrix Code [64], MaxiCode [65]) might be worth exploring.

Chapter 6 Conclusions and Future Work

6.1 Achievements

In this dissertation, we focus on two categories of problems. One is the problem of watermarking for error-prone transmission over unreliable network and the other is the problem of achieving watermark robustness against geometric attacks. We summarize some major contributions as follows:

(1) The theoretical and experimental analysis on both QIMM and QIM are demonstrated. The comparison of their properties in the aspects of embedding distortion and detection robustness is explored. It is verified that QIM and QIMM obtain similar embedding distortion, as shown by Delta-Distortion curve, and they are competitive in detection robustness, as shown by Delta-Correlation curve.

(2) We propose a multiple description watermarking technique which integrates an oblivious QIMM with the MDC framework. The watermark embedding is computed in either description and could be extracted with the reception of either one or two descriptions. Another advantage of our scheme worth mentioning here is the flexibility of the MDW framework. It can be integrated easily with most current watermarking schemes.

(3) We have presented an MRL-QIM watermarking scheme that is robust to non-linear value-metric distortion introduced by MD transmission. For a traditional balanced two-description case in a packet transmission network, the embedded watermark can be extracted with the reception of either one or two descriptions. The experimental result shows that the proposed MRL-QIM outperforms traditional vector QIM overall and performs better than SS-QIM in the case of high-rate transmission.

Furthermore, in the case of compression attack, the performance of proposed MRL-QIM still performs better than traditional vector QIM and is superior to SS-QIM for most of the compression rates.

(4) For geometric invariant watermarking via invariant features, the advantages of the presented approach are twofold. The first is to increase the detection robustness, a robust watermark via informed coding is chosen prior to embedding. The second is to increase transparency, which is done by LPM transforming only the circular parts of an image with radius no more than half the image row size while keeping other pixels of the image intact. Experimental results demonstrate that the resulting watermark is robust to a variety of image processing attacks (including RST attacks).

(5) For geometric invariant watermarking via re-synchronization, we propose a new approach [Appendix A] for geometric invariant watermarking technique by embedding sinusoidal signals individually in each of the selected sub-bands of dual-tree complex wavelet domain (DT-CWT), and then detecting the re-synchronization peaks by using the accumulated embedding sub-bands. The proposed scheme explores the feasibility of taking DT-CWT as transform domain for watermark re-synchronization. In addition, it finds a viable choice for the robust DT-CWT sub-bands for embedding, which will maximize the detection response peaks while preserving the visual quality of watermarked image.

(6) For geometric invariant watermarking via self-synchronization, we explore the features of QR Code for watermark self-synchronization and find a viable choice for the robust QR coded watermark embedding which will maximize the detection while preserving the visual quality of watermarked image. Experimental results demonstrate that by adopting our approach, the resulting self-synchronized watermark can be accurately detected under a variety of combined RST attacks.

6.2 Future Work

Although the practicality of watermarking applications on copyright protection and copy protection has been highly questioned and challenged [72, 73], yet its application in covert communication provides high assurance of confidentiality. Its various applications such as in the case of driver's license authentication and content linkage have also become more and more mature. However, we believe that there are still some problems and possible directions of research worthy to be further investigated in the future, as described below:

(1) For the problem of watermarking for error-prone transmission over unreliable network, we expect that other MDC approach or some error resilient algorithms could be integrated with the more elaborated watermarking schemes. Moreover, we expect to seek other transforms and statistic models to further enhance the robustness and increase the watermark payload while preserving the visual quality of the transmitted images. Furthermore, the steganography security against statistical steganalysis should also be addressed to enhance the security for reliable transmission. We believe that the results of these works will make it possible the watermarking of multimedia content for mobile E-commerce applications.

(2) For the problem of achieving watermark robustness against geometric attacks, we expect to explore further to combat other attacks, especially noising, blurring and compression which commonly occur in mobile phone camera based applications. To this end, the watermark embedding in frequency domain instead of spatial domain might serve a good alternative. Also, designing a better finder pattern for self-synchronized watermark embedding by innovating or improving some other existing barcode technology might be worth exploring.

(3) In contrast to the aforementioned problems, which are robustness-oriented, there is another kind of application aiming at answering the urgent demand for the difficulty of video quality metric in real-time communication [74-76] owing to lack of reference video on receiver's side. In this aspect, some researchers suggest employing hidden fragile watermark [77, 78] to assist video quality evaluation. The key concept is to embed fragile watermark into video before transmission, as the hidden fragile watermark will be distorted under various bit rate transmission or packet loss. Through the characteristic of fragile watermark, the evaluation of correlation ratio between the extracted and referenced watermark can reflect the quality of received video to certain degree. This fact is consistent with the results from direct evaluation between the received and referenced video.

Appendix A Watermark Re-synchronization using Sinusoidal Signals in DT-CWT Domain

Embedding sinusoidal signals or tiles patterns into image in the spatial domain to form some peaks is an effective technique for geometric invariant image watermark detection. However, there are two drawbacks in these spatial domain based schemes: one is poor picture quality of resulting watermarked image and the other is weak peaks visibility which is hard to detect. The previous works suffer from requiring a very strong watermark embedding to ease re-synchronization peaks finding, which in turn leads to a poorly watermarked image. To overcome this problem, we explore embedding sinusoidal signals individually in each of the selected sub-bands of dual-tree complex wavelet transform domain (DT-CWT), and then detecting the re-synchronization peaks by using the accumulated embedding sub-bands. Experimental results demonstrate that by adopting our approach, the resulting re-synchronization peaks are robust to rotation, scaling and translation attacks while preserving the visual quality of the watermarked image, thereby resolve the unavoidable dilemma faced by the other schemes.

A.1 Introduction

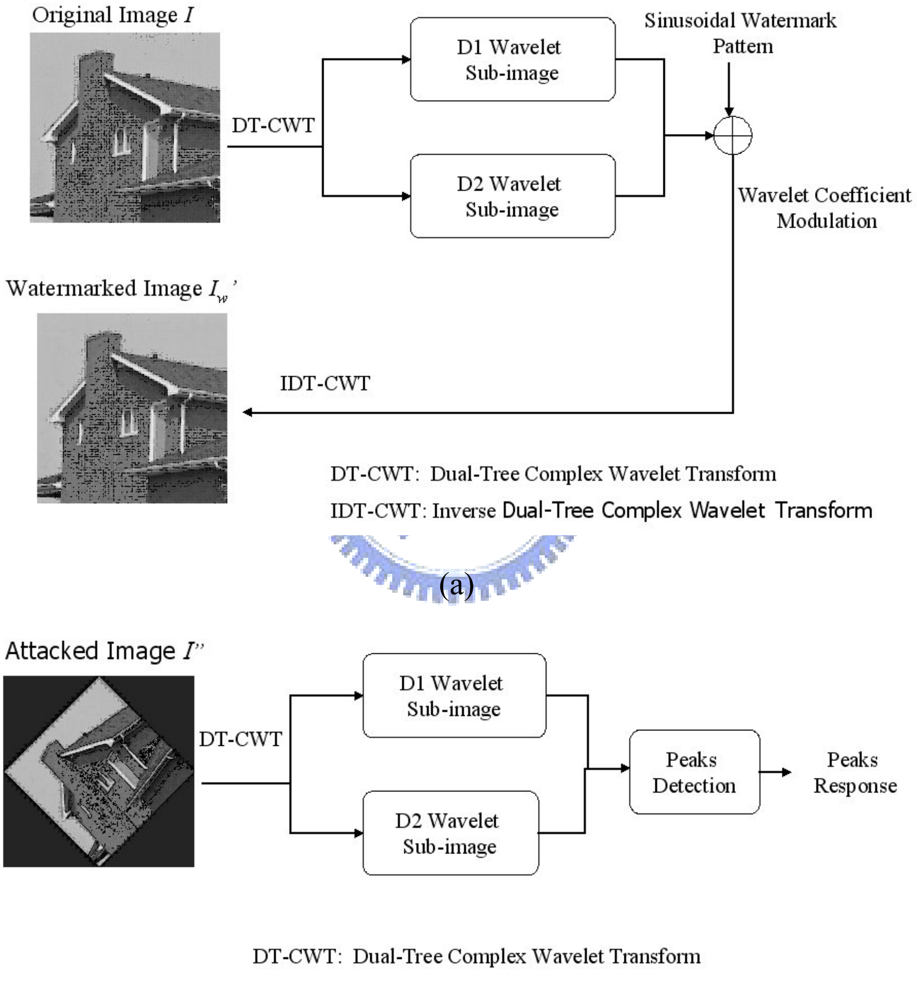
For re-synchronization approach, Kutter [33] embeds the self-reference pattern several times at horizontally and vertically shifted locations for recovering affine transformations. Fleet and Heeger [66] embeds sinusoidal signals into color image instead grey one to avoid artifact of watermarked image.

It is well recognized that the conventional discrete wavelet transform (DWT) is not suitable for geometric invariant watermark detection due to their lacking of shift-invariance property. Kingsbury [67] proposed the dual-tree complex wavelet transform (DT-CWT) which has both the advantages of being approximately shift-invariant and having additional directionalities (+15, +45, +75, -15, 145 and -75) compared to 3 directionalities (H, V, D) for traditional DWT. There are many successful applications by using DT-CWT, such as motion estimation [68], texture classification [69] and de-noising [70] for image or video applications. Although Loo and Kingsbury [71] had some works on DT-CWT watermarking, however none of their works dealt with the re-synchronization watermark embedding / detection; instead they use the original image to assist geometric invariant watermark detection.

The goal of this appendix is aiming at improving the robustness of re-synchronization peaks while preserving the picture quality of watermarked image. Section A.2 discusses the proposed watermark re-synchronization embedding / detection algorithm. In Section A.3 experimental results are presented and analyzed. Finally, in Section A.4 summary as well as future work are given.

A.2 Proposed algorithm

The main idea of the proposed scheme in Fig. A.1(a) and A.1(b) is to choose the robust DT-CWT sub-bands for embedding which will maximize the detection response peaks while preserving the visual quality of watermarked image at the same time. In the embedding process in Fig. A.1(a), the original image first goes through 2-level DT-CWT decomposed into 14-subbands (Fig. A.2). Sub-bands of each level are then grouped into two sub-images (D1: +15, +45 and +75, D2: -15, -45 and -75) based on their directionalities. In our preliminary implementation, only level-2 sub-images are considered for watermarking. The sub-bands of D1 and D2 are chosen to be modulated individually with the watermark pattern, and then are inverse-transformed to obtain the watermarked image. Depending on the respective detection robustness and the visual quality of watermarked image, one group of the sub-bands (either out of D1 or out of D2) will be eventually selected for embedding. In the detection process in Fig. A.1(b), the attacked image first goes through 2-level DT-CWT decompositions, the sub-bands chosen out of D1 and D2 are then accumulated for peaks detection.


Original Image I
DT-CWT
D1 Wavelet Sub-image
D2 Wavelet Sub-image
Sinusoidal Watermark Pattern
Wavelet Coefficient Modulation
Watermarked Image I_w'
IDT-CWT
DT-CWT: Dual-Tree Complex Wavelet Transform
IDT-CWT: Inverse Dual-Tree Complex Wavelet Transform
Attacked Image I''
DT-CWT
D1 Wavelet Sub-image
D2 Wavelet Sub-image
Peaks Detection
Peaks Response
DT-CWT: Dual-Tree Complex Wavelet Transform

(a)

(b)

Fig. A.1: (a) The flow of proposed watermark embedding scheme, (b) The flow of proposed watermark detection scheme.

.

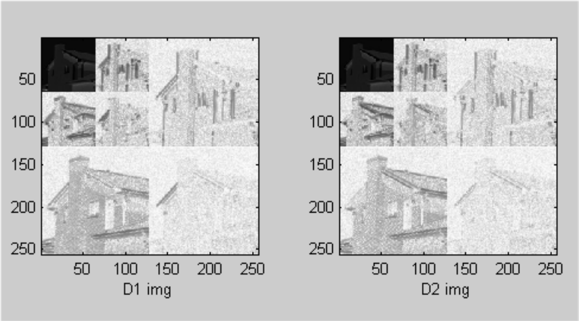


Fig. A.2: 2-levels of DT-CWT decomposition

A.2.1 Sinusoidal Signals as Watermark Pattern

The watermark *W* consists of three sinusoidal signals, with frequency 16 cycles per unit length along x-axis, 32 cycles per unit length along y-axis, and 16/32 cycles per unit length along x/y axis. Fig. A.3 shows the sinusoidal watermark pattern and its peak response in the frequency domain. Ideally there should be totally 6 peaks (except for the center point) for these three sinusoidal signals, with 2 peaks for each sinusoid due to its symmetry in the frequency domain.

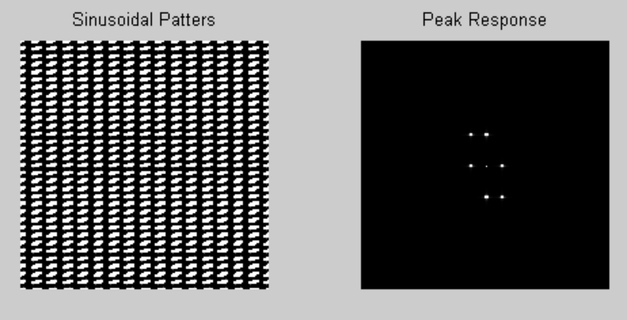


Fig. A.3: The sinusoidal watermark pattern and its peaks response in the frequency domain.

A.2.2 Watermark Embedding

The embedding algorithm consists of the following steps:

Step 1: The original image *I* is decomposed into 14-subbands using the 2-level DT-CWT.

Step 2: The sub-bands of each level are then grouped into two sub-images (D1 and D2) based on their directionalities.

Step 3: The sub-bands of *D1* and *D2* are chosen to be modulated individually with the watermark pattern *W*, and then are inverse-transformed to obtain the watermarked image I_w'.

Step 4: Depending on their detection robustness and the visual quality of watermarked image, either sub-bands of *D1* or *D2* are eventually selected for embedding

A.2.3 Watermark Detection

The detection algorithm consists of the following steps:

Step 1: The attacked image *I*" (being attacked from I_w') is decomposed into 14-subbands using the 2-level DT-CWT.

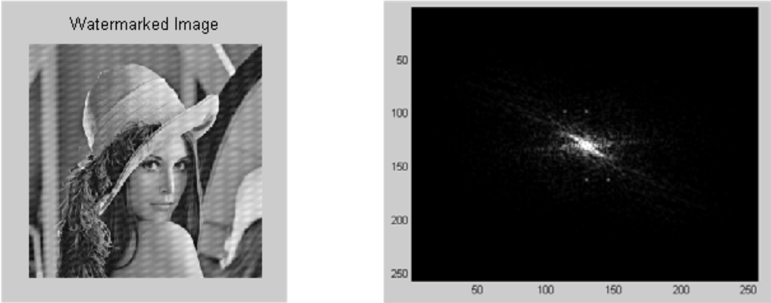
Step 2: The sub-bands of each level are then grouped into two sub-images (*D1* and *D2*) based on their directionalities.

Step 3: The sub-bands of *D1* and *D2* are then accumulated for peaks detection.

A.3 Experimental Results

To evaluate the effectiveness of the proposed method, three standard test images of size 256 x 256 including “Lena”, “Barbara” and “Boat” are utilized as host signals to embed sinusoidal watermark patterns. To save space for the publication, only the results of “Lena” image are shown here, other images show the similar results as well. Main issues regarding performance evaluation of the proposed method are discussed in the following.

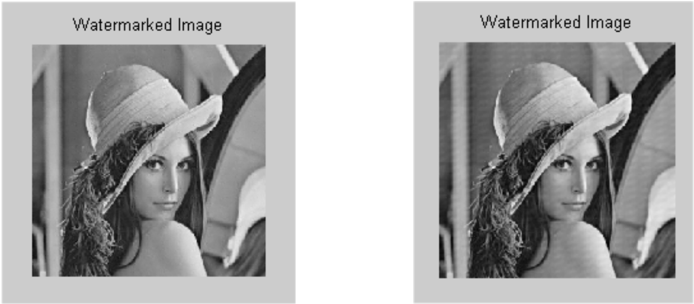
The traditional spatial domain based scheme requires a very strong watermark embedding to ease re-synchronization peaks finding, which in turn leads to a poorly watermarked image. Fig. A.4 shows the watermarked image with PSNR 30dB and its corresponding peaks response. Note that the distortion caused by such strong embedding is quite visually significant.



(a) Watermarked image (b) Peaks response

Fig. A.4: Watermarked image (PSNR 30dB) and its peak responses in the frequency domain of conventional scheme

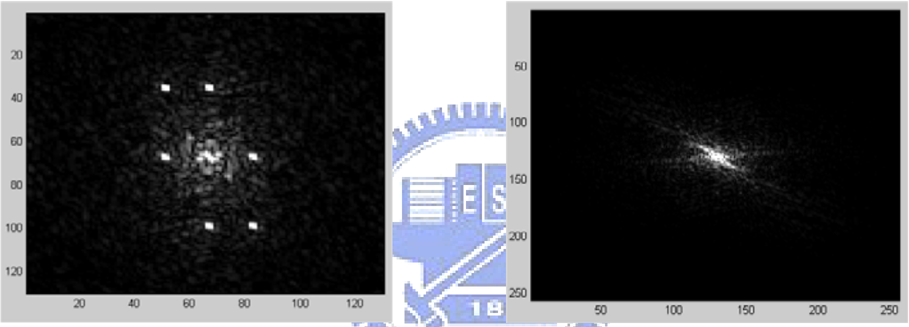
For the fairness of comparison, the weighting parameter of watermark strength of our proposed approach is adapted to obtain the similar PSNR values (about 40dB) as that of the conventional spatial domain based scheme. Fig. A.5(a) and A.5(b) show the watermarked image by adopting our proposed scheme and conventional one respectively, and their corresponding peak responses are shown in Fig. A.6(a) and A.6(b). We can see that the picture quality is very good for our scheme. However, there exists some small pattern artifact in the case of conventional scheme. The visibility of 6 peak responses is rather conspicuous for our proposed scheme. However, we cannot distinguish the 6 peak responses for conventional scheme. It is evident that our proposed scheme is superior to the approach of the conventional one.



(a) Watermarked image of proposed scheme (PSNR 40.5 dB)

(b) Watermarked image of conventional scheme (PSNR 40.3 dB)

Fig. A.5: For the fairness of comparison, the watermarked images are adapted to obtain the similar PSNR values to test the robustness of peaks response



(a) Proposed scheme

(b) Conventional scheme

Fig. A.6: The peak responses of our proposed scheme show far better effect than the approach of the conventional one.

To test the robustness of rotation, scaling and translation (RST), the watermarked image is manipulated with various combinations of RST transformations to generate attacked images. The peak responses are still visible under these various RST attacks of our proposed scheme. Fig. A.7 shows one of the image attacked by combinations of rotation 30°, scaling 1.8 and translation (20, 10). We can see from Fig. A.8 that the 6 peak responses of our proposed scheme are rotated, scaled and translated by the same amounts as that of the attacked image. Our proposed scheme performs still far better than those of the conventional one.

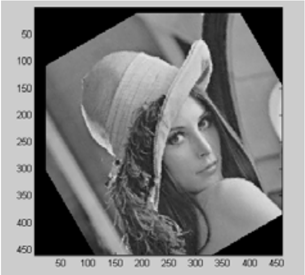
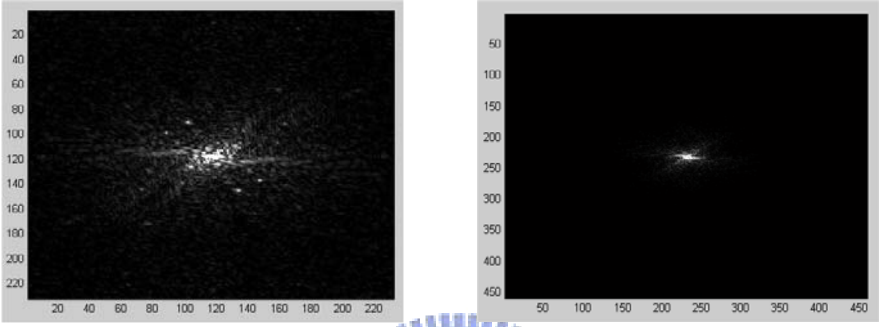


Fig. A.7: Image attacked by combinations of rotation 30^{o}, scaling 1.8 and translation (20, 10)



(a) Proposed scheme (b) Conventional scheme

Fig. A.8: In the case of attack by combinations of RST, the peak responses of our proposed scheme still perform far better than those of the conventional one.

A.4 Summary

We propose a new approach for geometric invariant watermarking technique by embedding sinusoidal signals individually in each of the selected sub-bands of dual-tree complex wavelet domain (DT-CWT), and then detecting the re-synchronization peaks by using the accumulated embedding sub-bands. The main contributions of the proposed scheme are: (1) exploring the feasibility of taking DT-CWT as transform domain for watermark re-synchronization; (2) finding a viable choice for the robust DT-CWT sub-bands for embedding which will maximize the detection response peaks while preserving the visual quality of watermarked image. Experimental results demonstrate that by adopting our approach, the resulting re-synchronization peaks are robust to rotation, scaling and translation attacks while preserving the visual quality of the watermarked image, thereby resolve the unavoidable dilemma faced by the other schemes.

In the future, we expect to see more than 2-level decomposition of DT-CWT being employed, in which case the analysis of robust sub-bands for embedding will become increasingly complicated. On the other hand, further improvement could be made on the peaks finding algorithm to systematically find out the peaks location, so that real watermark payload (other than re-synchronization watermark patterns) embedding/detection could be achieved.

Bibliography

[1] Digimarc Co., PictureMarc, http://www.digimarc.com.

[2] ISO/IEC 18004:2000. Information technology – Automatic identification and data capture techniques – Bar code symbology – QR Code, 2000.

[3] MediaGrid Co., Leadia Pix, http://www.mediagrid.co.jp/.

[4] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1673-1687, Dec. 1997.

[5] C. S. Lu, S. K. Huang, C. J. Sze, and H. Y. M. Liao, "Cocktail watermarking for digital image protection," *IEEE Transactions on Multimedia*, vol. 2, no. 4, pp. 209-224, Dec. 2000.

[6] H. S. Malvar and D. A. F. Florêncio, "Improved spread spectrum: a new modulation technique for robust watermarking," *IEEE Transactions on Signal Processing*, vol. 51, no. 4, pp. 898-905, Apr. 2003.

[7] B. Chen and G. W. Wornell, "Quantization index modulation: a class of provably good methods for digital watermarking and information embedding," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1423-1443, May 2001.

[8] J. J. Eggers, R. Bäuml, R. Tzschoppe, and B. Girod, "Scalar costa scheme for information embedding," *IEEE Transactions on Image Processing*, vol. 51, no. 4, pp. 1003-1019, Apr. 2003.

[9] F. Hartung and F. Ramme, "Digital rights management and watermarking of multimedia content for M-commerce applications," *IEEE Communications Magazine*, vol. 38, no. 11, pp. 78-84, Nov. 2000.

[10] N. Checcacci, M. Barni, F. Bartolini, and S. Basagni, "Robust video watermarking for wireless multimedia communications," in *Proceedings of IEEE Wireless Communications and Networking Conference*, vol. 3, pp. 1530-1535, 2000.

[11] B. Graubard, R. Chandramouli, and C. Richmond, "A multiple description framework for oblivious watermarking," in *Proceeding of SPIE: Security and Watermarking of Multimedia Contents III*, vol. 4314, pp. 585-593, 2001.

[12] Y. Wang, M. T. Orchard, V. A. Vaishampayan, and A. R. Reibman, "Multiple description coding using pairwise correlating transforms," *IEEE Transactions on Image Processing*, vol. 10, no. 3, pp. 351-366, Mar. 2001.

[13] V. K. Goyal, "Multiple description coding: compression meets the network," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 74-93, Sep. 2001.

[14] Y. Wang, A. R. Reibman, and S. Lin, "Multiple description coding for video delivery," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 57-70, Jan. 2005.

[15] S. D. Servetto, K. Ramchandran, V. A. Vaishampayan, and K. Nahrstedt, "Multiple description wavelet based image coding," *IEEE Transactions on Image Processing*, vol. 9, no. 5, pp. 813-826, May 2000.

[16] V. A. Vaishampayan, "Design of multiple description scalar quantizers," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 821-834, May 1993.

[17] Y. Wang, S. Wenger, J. Wen, and A. K. Katsaggelos, "Error resilient video coding techniques," *IEEE Signal Processing Magazine*, vol. 17, no. 4, pp. 61-82, Jul. 2000.

[18] M. Kutter and F. A. P. Petitcolas, "Fair evaluation methods for image watermarking systems," *Journal of Electronic Imaging*, vol. 9, no. 4, pp. 445-455, Oct. 2000.

[19] F. P. González, C. Mosquera, M. Barni, and A. Abrardo, “Rational dither modulation: a high-rate data-hiding method robust to gain attacks,” *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3960-3975, Oct. 2005.

[20] P. Bas, “A quantization watermarking technique robust to linear and non-linear valumetric distortion using a fractal set of floating quantizers,” in *Proceedings of Information Hiding Workshop*, vol. 3727, pp. 106-117, 2005.

[21] M. L. Miller, G. J. Doerr, and I. J. Cox, “Applying informed coding and informed embedding to design a robust, high capacity watermark,” *IEEE Transactions on Image Processing*, vol. 13, no. 6, pp. 792–807, Jun. 2004.

[22] A. Abrardo and M. Barni, “Informed watermarking by means of orthogonal and quasi-orthogonal dirty paper coding,” *IEEE Transactions on Signal Processing*, vol. 53, no. 2, pp. 824–833, Feb. 2005.

[23] J. Oostveen, T. Kalker, and M. Staring, “Adaptive quantization watermarking,” in *Proceedings of SPIE: Security, Steganography, and Watermarking of Multimedia Contents VI,* vol. 5306, pp. 296-303, 2004.

[24] K. Sayood, “Introduction to data compression,” *Morgan Kaufmann Publishers*., 1996.

[25] D. J. Newman, “The hexagon theorem,” *IEEE Transactions on Information Theory*, vol. 28, no. 4, pp. 137-139, Mar. 1982.

[26] K. Sayood and S. J. Blankenau, “A fast quantization algorithm for lattice quantizer design,” *ICASSP*, vol. 2, pp. 1168-1171, 1988.

[27] J. H. Conway and N. J. A. Sloane, “Fast quantizing and decoding algorithms for lattice quantizers and codes,” *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 227-232, Mar. 1982.

[28] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, “Closest point search in lattices,” *IEEE Transactions on information Theory*, vol. 48, no. 8, pp. 2201-2214, Aug. 2002.

[29] R. Zamir, S. Shamai, and U. Erez, “Nested linear/lattice codes for structured multiterminal binning,” *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1250-1276, Jun. 2002.

[30] I. J. Cox, M. L. Miller, and J. A. Bloom, “Watermarking: principles & practice,” *The Morgan Kaufmann Series in Multimedia and Information Systems*, 2001.

[31] V. Licks and R. Jordan, “Geometric attacks on image watermarking systems,” *IEEE Transactions on Multimedia*, vol. 12, no. 3, pp. 68-78, Jul.-Sep. 2005.

[32] P. Dong, J. Brankov, N. Galatsanos, Y. Yang, and F. Davoine, “Digital watermarking robust to geometric distortion,” *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2140-2150, Dec. 2005.

[33] M. Kutter, “Watermarking resistance to translation, rotation, and scaling,” *Proc. SPIE Multimedia Systems Applications*, vol. 3528, pp. 423-432, 1998.

[34] S. Pereira and T. Pun, “Robust template matching for affine resistant image watermarks,” *IEEE Transactions on Image Processing*, vol. 9, no. 6, pp. 1123-1129, Jun. 2000.

[35] D. Delannay and B. Macq, “Watermarking relying on cover signal content to hide synchronization marks,” *IEEE Trans. on Information Forensics and Security*, vol. 1, no. 1, pp. 87-101, Mar. 2006.

[36] C.H. Lee and H.K. Lee, “Improved autocorrelation function based watermarking with side information,” *Journal of Electronic Imaging*, vol. 14, no. 1, pp. 1-13, Jan.-Mar. 2005.

[37] F. Deguillaume and S. Voloshynovskiy, T. Pun, “A method for the estimation and recovering of general affine transforms in digital watermarking applications,” In *IS&T/SPIE's 14th Annual Symposium, Electronic Imaging*, vol. 4675, pp. 313-322, 2002.

[38] G.B. Rhoads, Methods for surveying dissemination of proprietary empirical data, *U.S. Patent* No. 5862260, 1999.

[39] P. Bas, J. M. Chassery, and B. Macq, “Geometrically invariant watermarking using feature points,” *IEEE Transactions on Image Processing*, vol. 11, no. 9, pp. 1014-1028, Sep. 2002.

[40] C. W. Tang and H. M. Hang, “A feature based robust digital image watermarking scheme,” *IEEE Transactions on Signal Processing*, vol. 51, no. 4, pp. 950-959, Apr. 2003.

[41] V. Solachidis and I. Pitas, “Circularly symmetric watermark embedding in 2-D DFT Domain,” *IEEE Transactions on Image Processing*, vol. 10, no. 11, pp. 1741-1753, Nov. 2001.

[42] J. J. K. O’Ruanaidh and T. Pun, “Rotation, scale and translation invariant spread spectrum digital image watermarking,” *Signal Processing*, vol. 66, no. 3, pp. 303-317, May 1998.

[43] C. Y. Lin, M. Wu, J. A. Bloom, I.J. Cox, M. L. Miller, and Y. M. Lui, “Rotation, scale, and translation resilient watermarking for images,” *IEEE Transactions on Image Processing*, vol. 10, no. 5, pp. 767-782, May 2001.

[44] A. Herrigel, S. Voloshynovskiy, and Y. Rytsar, “The watermark template attack,” *Proc. SPIE Electronic Imaging: Security and Watermarking of Multimedia Content III*, vol. 4314, pp. 394-405, 2001.

[45] Q. S. Chen, Image registration and its applications in medical imaging, *Ph.D. Thesis*, Free University Brussels (VUB), 1993.

[46] B. S. Reddy and B. N. Chatterji, “An FFT-based technique for translation, rotation, and scale-invariant image registration,” *IEEE Transactions on Image Processing*, vol. 5, no. 8, pp. 1266-1271, Aug. 1996.

[47] R. C. Gonzalez and R. E. Woods, Digital image processing. Upper Saddle River, New Jersey: *Prentice-Hall, Inc.*, 2002.

[48] M. Barni, F. Bartolini, V. Cappellini, and A. Piva, “A DCT-domain system for robust image watermarking,” *Signal Processing*, vol. 66, no. 3, pp. 357-372, May 1998.

[49] T. S. Liang and J. J. Rodriguez, “Robust watermarking using robust coefficients,” *Proc. SPIE Electronic Imaging: Security and Watermarking of Multimedia Contents II*, vol. 3971, pp. 326-335, 2000.

[50] R. Barnett and D. E. Pearson, “Frequency mode LR attack operator for digitally watermarked images,” *IEE Electronic Letters*, vol. 34, no. 19, pp. 1837-1839, Sep. 1998.

[51] M. L. Day, S. Y. Lee, and I. C. Jou, “DT-CWT domain based self-synchronization watermarking technique,” *17th IPPR Conference on Computer Vision, Graphics and Image Processing*, Taiwan, in CD-ROM, 2004.

[52] M. L. Day, S. Y. Lee, and I. C. Jou, “Watermark re-synchronization using sinusoidal signals in DT-CWT Domain,” *Fifth IEEE Pacific Rim Conference on Multimedia, Tokyo, Japan, Springer LNCS 3333*, pp. 394-401, 2004.

[53] M. L. Day, S. Y. Lee, and I. C. Jou, “Multiple description watermarking based on quantization index modulus modulation,” to appear in *Journal of Information Science and Engineering*.

[54] P. Heckbert, Fundamentals of texture mapping and Image warping, *Master's Thesis*, University of California at Berkeley, Computer Science Division, EECS Department, 1989.

[55] J. Portilla, V. Strela, M. Wainwright, and E. P. Simoncelli, “Image denoising using scale mixtures of Gaussians in the wavelet domain,” *IEEE Trans. on Image Processing*, vol. 12, no. 11, pp. 1338-1351, Nov. 2003.

[56] H. R. Sheikh, A. C. Bovik, and G. D. Veciana, “An information fidelity criterion for image quality assessment using natural scene statistics,” *IEEE Trans. on Image Processing*, vol. 14, no. 12, pp. 2117-2128, Dec. 2005.

[57] H. R. Sheikh, A. C. Bovik, and L. Cormack, “No-reference quality assessment using natural scene statistics: JPEG2000,” *IEEE Trans. on Image Processing*, vol. 14, no. 11, pp. 1918-1927, Nov. 2005.

[58] H. R. Sheikh and A. C. Bovik, “Image information and visual quality,” *IEEE Trans. on Image Processing*, vol. 15, no. 2, pp. 430-444, Feb. 2006.

[59] Z. Wang, G. Wu, H. R. Sheikh, E. P. Simoncelli, E. H. Yang, and A. C. Bovik, “Quality-aware images,” *IEEE Trans. on Image Processing*, vol. 15, no. 6, pp. 1680-1689, Jun. 2006.

[60] E. Ohbuchi, H. Hanaizumi, and L. A. Hock, “Barcode readers using the camera device in mobile phones,” *Proc. International Conference on Cyberworlds*, vol. 00, pp. 260-265, 2004.

[61] J. Stach, T. Brundage, B. Hannigan, B. Bradley, T. Kirk, and H. Brunk, “On the use of web cameras for watermark detection,” *Proc. SPIE, Security and Watermarking of Multimedia Contents IV*, vol. 4675, pp. 611-620, 2002.

[62] T. Nakamura, A. Katayama, M. Yamamuro, and N. Sonehara, “Fast watermark detection scheme for camera-equipped cellular phone,” *ACM Proceedings of the*

3rd international conference on Mobile and ubiquitous multimedia, vol. 83, pp. 101-108, 2004.

[63] M. Rohs, “Real-world interaction with camera-phones,” *2nd International Symposium on Ubiquitous Computing Systems*, vol. 11, pp. 39-48, 2004.

[64] ISO/IEC 16022 International Symbology Specification – Datamatrix, 2000.

[65] http://en.wikipedia.org/wiki/MaxiCode.

[66] D. J. Fleet and D. J. Heeger, “Embedding invisible information in color images,” in *IEEE Int. Conf. Image Processing, ICIP’97*, vol. 1, pp. 532-535, 1997.

[67] N. Kingsbury, “Image processing with complex wavelets,” *Phil. Trans. Royal Society London*. A, vol. 357, pp. 2543-2560, Sep. 1999.

[68] J. Magarey and N. Kingsbury, “Motion estimation using a complex-valued wavelet transform,” *IEEE Transactions on Signal Processing*, vol. 46, no. 4, pp. 1069-1084, Apr. 1998.

[69] P. D. Rivaz, Complex wavelet based image analysis and synthesis, *Ph.D Thesis*, University of Cambridge, Oct., 2000.

[70] L. Sendur and I. W. Selesnick, “Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency,” *IEEE Transactions on Signal Processing*, vol. 50, no. 11, pp. 2744-2756, Nov. 2002.

[71] P. Loo, Digital watermarking using complex wavelet, *Ph.D Thesis*, University of Cambridge, March, 2002.

[72] C. Herley, “Why watermarking is nonsense,” *IEEE Signal Processing Magazine*, vol. 19, no. 5, pp. 10-11, Sep. 2002.

[73] P. Moulin, “Comment on ‘Why watermarking is nonsense’,” *IEEE Signal Processing Magazine*, vol. 20, no. 6, pp. 57-59, Nov. 2003.

[74] M. L. Day, Y. M. Du, J. C. Chen, and C. N. Chang, “A study on MOD video quality metrics,” to appear in *TL Journal*, vol. 37, no. 4, Aug. 2007 (in Chinese).

[75] “Stream PQoS metric evaluation,” available at http://www.genista.com.

[76] S. Winker, Vision models and quality metrics for image processing applications, *Ph. D thesis*, EPFL, Switzerland, 2000.

[77] P. Campisi, M. Carli, G. Giunta, and A. Neri, “Blind quality assessment system for multimedia communications using tracing watermarking,” *IEEE Transactions on Signal Processing*, vol. 51, no. 4, pp.996-1001, Apr. 2003.

[78] M. L. Day, I. C. Jou, and S. Y. Lee, “Video quality metric for real time communication via fragile watermark,” *Workshop on Consumer Electronics (WCE)*, Tainan, Taiwan, R.O.C., in CD-ROM, 2003 (in Chinese).