# Joint Compression and Watermarking Using Variable-Rate Quantization and its Applications to JPEG

by

Yuhan Zhou

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2008

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

**Abstract**

In digital watermarking, one embeds a watermark into a covertext, in such a way that the resulting watermarked signal is robust to a certain distortion caused by either standard data processing in a friendly environment or malicious attacks in an unfriendly environment. In addition to the robustness, there are two other conflicting requirements a good watermarking system should meet: one is referred as perceptual quality, that is, the distortion incurred to the original signal should be small; and the other is payload, the amount of information embedded (embedding rate) should be as high as possible. To a large extent, digital watermarking is a science and/or art aiming to design watermarking systems meeting these three conflicting requirements. As watermarked signals are highly desired to be compressed in real world applications, we have looked into the design and analysis of joint watermarking and compression (JWC) systems to achieve efficient tradeoffs among the embedding rate, compression rate, distortion and robustness.

Using variable-rate scalar quantization, an optimum encoding and decoding scheme for JWC systems is designed and analyzed to maximize the robustness in the presence of additive Gaussian attacks under constraints on both compression distortion and composite rate. Simulation results show that in comparison with the previous work of designing JWC systems using fixed-rate scalar quantization, optimum JWC systems using variable-rate scalar quantization can achieve better performance in the distortion-to-noise ratio region of practical interest.

Inspired by the good performance of JWC systems, we then investigate its applications in image compression. We look into the design of a joint image compression and blind watermarking system to maximize the compression rate-distortion performance while maintaining baseline JPEG decoder compatibility and satisfying the additional constraints imposed by watermarking. Two watermarking embedding schemes, odd-even watermarking (OEW) and zero-nonzero watermarking (ZNW), have been proposed for the robustness to a class of standard JPEG recompression attacks. To maximize the compression performance, two corresponding alternating algorithms have been developed to jointly optimize run-length coding, Huffman coding and quantization table selection subject to the additional constraints imposed by OEW and ZNW respectively. Both of two algorithms have been demonstrated to have better compression performance than the DQW and DEW algorithms developed in the recent literature. Compared with OEW scheme, the ZNW embedding method sacrifices some payload but earns more robustness against other types of attacks. In particular, the zero-nonzero watermarking scheme can survive a class

of valumetric distortion attacks including additive noise, amplitude changes and recompression for everyday usage.

**Acknowledgements**

First, I would like to express my sincere gratitude to my advisor Professor Enhui Yang, for his technical guidance, helpful insights and invaluable advice. He has not only lead me to this cutting-edged research, but also given me the impetus to be precise in thinking and writing. And I also gratefully acknowledge my another advisor Professor Alexei Kaltchenko for his help and financial support.

I am grateful to Professor Liang-liang Xie and Professor Zhou Wang for being readers of my thesis. I would also like to thank Dr. Guixing Wu and Dr. Longji Wang with Research In Motion for their great help during my thesis writing.

I am deeply indebted to my friends of the Multimedia Communications Laboratory at the University of Waterloo, Dr. Xiang Yu (now with RIM), Dr. Wei Sun (now with Mitsubishi Electric Research Laboratories, Boston), Mr. Jin Meng, Miss Lin Zheng, Miss Jiao Wang, and Mr. Krzysztof Michal Hebel. Their support and discussion are the invaluable resource of my improvement and happiness. I am also deeply committed to my family. I thank my father and mother, for their unselfish love and constant understanding. A few words mention here cannot adequately capture all my appreciation. Still, I wish to thank those professors, who are the instructors of class I have attended, and all the other friends of mine.

## Dedication

This is dedicated to my parents and all of my friends.

# Contents

# List of Figures

# Chapter 1

# Introduction

## 1.1 Digital Watermarking

In the recent decade, new devices and powerful software have made it possible for consumers worldwide to access, create, and manipulate multimedia data. Internet and wireless networks offer ubiquitous channels to deliver and to exchange such multimedia information. However, the potential offered by the information technology era cannot be fully realized without the guarantee on the security and protection of multimedia data. Thus, there is a strong need for techniques to protect the copyright of content owners. Cryptography and digital watermarking are two complementary techniques proposed so far to protect digital content.

Cryptography is the processing of information into an encrypted form for the purpose of secure transmission. Before delivery, the digital content is encrypted by the owner by using a secret key. A corresponding decryption key is provided only to a legitimate receiver. The encrypted content is then transmitted via Internet or other public channels, and it will be meaningless to pirate without the decryption key. At the receiver end, however, once the encrypted content is decrypted, it has no protection anymore.

On the other hand, digital watermarking is a technique that can protect the digital content even after it is decrypted. In digital watermarking, a watermark is embedded into a covertext or host signal (the digital contents to be protected), resulting in a watermarked signal called stegotext which has no visible difference from the covertext. The stegotext is subject to manipulation by a malicious attacker, who produces a forgery. The goal of the attacker is to make the watermark undetectable from the forgery. Careful design of the watermarking system can minimize

the chance that such an attack will be successful.

Three key issues in the design of watermarking schemes are as follows.

- **Payload**.

  This refers to the number of information bits that are embedded in the covertext. This can vary from megabytes of information (for secret communication applications) to as little as a few bits (for copyright protection applications). For instance, DVD players have been proposed that verify the status of only four information bits before recognizing the file as legitimate and playing it. The payload is often normalized by the number of samples of the host signal, resulting in a bit rate $R_w$ per sample of the covertext.

- **Transparency (Fidelity)**.

  In most applications, embedding of information should not cause perceptual degradation of the covertext. Embedded information should be invisible in images and text, and inaudible in speech and audio. For a given application there is a tolerable distortion level, generically denoted as $D_1$.

- **Robustness**.

  Although an attacker could possibly introduce distortion (e.g., common signal processing operations such as compression, filtering, noise addition, desynchronization, cropping, insertions, mosaicing, and collage.) into the stegotext and thus create a forgery, the hidden message should still be detectable. The watermark embedding schemes are commonly designed to survive a certain level of distortion, generically denoted as $D_2$.

Because of its applications to areas such as copyright protection, broadcast monitoring and fingerprinting, digital watermarking has been studied extensively during the past a few years. The best tradeoff among the embedding rate, distortion, and robustness was investigated recently from an information-theoretic perspective. Specifically, in [5], Moulin and O'Sullivan introduced an information-theoretic model of the watermarking game and determined upper and lower bounds on the information embedding capacity for both public and private watermarking. In [6], information rates were investigated for Gaussian host signals and the squared-error distortion measure. In [10], Chen and Wornell showed that a coding strategy called distortion-compensated quantization index modulation (DC-QIM) can achieve the capacity for several scenarios when the statistics of the attack channel is known.

A lot of practical watermarking schemes were also designed and tested empirically (see, for instance, [1] [3] and the references therein). Among them are two most popular approaches to watermarking problem proposed so far, that is, spread-spectrum watermarking proposed in [27] and quantization based watermarking proposed in [10]. In spread-spectrum watermarking, the watermark information is embedded by linearly combining the host signal with a small pseudo-noise signal that is modulated by embedded watermark. Although this approach has been received considerable attention in the literature, it is limited by the interference from the host signal when the host signal is not available at the watermark decoder, which is typical in most of the watermarking applications. In quantization-based watermarking, the watermark information is conveyed in the choice of different quantizers. This approach has the advantage of rejecting the host signal interference, therefore, it has a higher information embedding rate than spread spectrum watermarking and is useful in a digital watermarking system where the watermark decoder can not access to the host signal.

## 1.2 Research Problems and Motivations

Since in most applications, watermarked signals will be likely stored and/or transmitted in compressed format, another aspect of the watermarking problem is that of joint information embedding and lossy compression, where quantization and entropy coding of the stegotext are carried out as an integral part of the watermarking scheme. In contrast with a vast amount of research in digital watermarking, there are only a few research works in the domain of joint watermarking and compression. Specifically, some ad hoc JWC algorithms were proposed for applications in images, audio, and video [29] [22] [30]. A set of efficient practical schemes for joint watermarking and compression (JWC) are proposed by Wu and Yang in [12]. The schemes of JWC are based on creating disjoint codebooks representing different watermarks by using fixed-rate different scalar quantizers and aim at maximizing the robustness of the embedding in the presence of additive Gaussian attacks, under constraints on the quantization distortion. Yet, another possible implementation of such practical schemes is the one proposed in [13], which uses modulated lattice vector quantization (MLVQ), based on dither modulation and lattice vector quantization. Though it has been shown that the MLVQ scheme has good performances, due to the high complexity of vector quantization, this approach has its disadvantage in real applications.

Inspired by the approach of designing JWC systems using fixed-rate scalar quantization, we raise the following questions:

- Can we get more efficient joint watermarking and compression schemes if we use variable-rate scalar quantization (VRSQ) instead of fixed-rate scalar quantization?
- How can we implement VRSQ in the JWC system design for real world applications in order to get efficient tradeoffs among payload, transparency, compression rate and robustness meanwhile the designed watermark encoder is compatible with the decoders in current multimedia compression standards?

In this thesis, we will look into how to address these problems as described in the following paragraphs.

(1) *JWC using variable-rate scalar quantization*: Since it has been shown that JWC systems using fixed-rate scalar quantization have great advantage over separately designed watermarking systems [12], we want to further improve the JWC system performance by using variable-rate scalar quantization. We show that by using variable-rate scalar quantization, a potential distortion-to-noise ratio (DNR) gain can be obtained when considering decoding bit error probability in the presence of additive white Gaussian noise (AWGN) attacks. An alternating algorithm is also developed to implement this scheme with low complexity.

(2) *Joint image compression and blind watermarking with baseline JPEG decoder compatible*: Inspired by the advantage of designing JWC using VRSQ, we go one step further to investigate more efficient ways to embed watermark information associated with an image invisibly into compressed bit streams. In this work, we propose two innovative joint compression and blind watermarking methods to hide the data or similar type of information invisibly into a compressed image with high payload. The resulting data can be attacked by legitimate signal processing for everyday usage in the decompressed domain. Later, the hidden information can be extracted using a watermark decoder whenever necessary. We focus on embedding watermarks into JPEG compressed bit streams, due to the wide applications of the JPEG standard. It is shown that both of our proposed watermark embedding algorithms achieve better rate-distortion performance than the DQW algorithm [23] and the DEW algorithm [21] when the same information embedding rate and JPEG recompression attacks are considered. In particular, the second proposed algorithm, zero-nonzero watermarking (ZNW), also achieves good robustness against

other types of valumetric distortion attacks including additive Gaussian noise and amplitude scaling in everyday usage.

## 1.3 Thesis Organization and Contributions

The rest of the thesis is organized as follows. In Chapter 2, we first give a brief review of digital watermarking and joint compression and digital watermarking from the information-theoretic point of view. Then, some of the correlative theoretic results are stated therein. In Chapter 3, we first review JWC system design using fixed-rate scalar quantization in [12], and then a more efficient JWC scheme using variable-rate scalar quantization with an alternating algorithm is proposed. The experiment results in the case of AWGN attacks and a comparison with the performance in literature are reported thereafter to show that better performance can be obtained by using variable-rate scalar quantization in designing JWC systems. In Chapter 4, two new joint JPEG compression and blind watermarking schemes are proposed after reviewing the previous works in the literature. Experiment results and comparisons with the DQW and DEW algorithms are reported therein. We summarize the whole thesis and discuss open problems that arise from the presented research in Chapter 5.

## 1.4 Notation

Throughout the thesis, the following notations are adopted. We use capital letters to denote random variable, lowercase letters for its realization, and script letters for its alphabet. For instance, $X$ is a random variable over its alphabet $\mathcal{X}$ and $x \in \mathcal{X}$ is a realization. We use $p_X(x)$ to denote the probability distribution of a discrete random variables $X$ taking values over its alphabet $\mathcal{X}$, and also to denote the probability density function of a continuous random variable $X$. If there is no ambiguity, sometimes the subscript in $p_X(x)$ is omitted and we write $p(x)$ instead. Similarly, $X^n = (X_1, X_2, \ldots, X_n)$ denotes a random vector taking values over $\mathcal{X}^n$, and $x^n = (x_1, x_2, \ldots, x_n)$ is a realization. Furthermore, $\mathbb{E}$ denotes the expectation operator, $H(X)$ is the entropy of $X$, and $I(X;Y)$ denotes the mutual information between $X$ and $Y$.

# Chapter 2

# Joint Compression and Digital Watermarking: Information-Theoretic Viewpoint Review

In this chapter, the standard model of digital watermarking is introduced first from an information theoretic viewpoint. Then, the main problem on joint compression and watermarking is formulated and the correlative results are stated.

## 2.1 Information-Theoretic Review of Digital Watermarking

From an information theoretic viewpoint, a digital watermarking system can be modeled as a communication system with side information at the watermark transmitter, as depicted in Fig. 2.1. In this model, $M$ is the message to be embedded and it is uniformly distributed over the message set and is to be reliably transmitted to the decoder. The host data are a sequence $S^N = (S_1, S_2, \ldots, S_N)$ of independent and identically distributed (i.i.d.) samples drawn from $p(s)$. The composite data set $X^N$ is subject to attacks embodied by the channel $A(y|x)$.

The information hider and the attacker are subjected to distortion constraints between the covertext and watermarked signals. We define a distortion function for

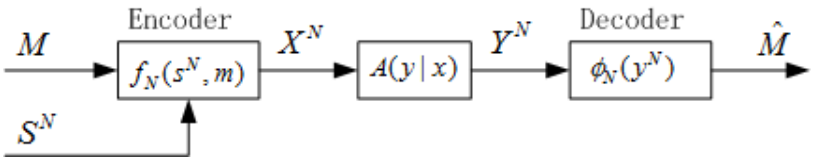


Figure 2.1: Formulation of information hiding as a communication problem.

the information hider as a nonnegative function $d_1 : \mathcal{S} \times \mathcal{X} \to \mathbb{R}_+$. The distortion function for the attacker is defined as a nonnegative function $d_2 : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+$.

The distortion function for the information hider is bounded and the distortion functions $d_i$, $i \in \{1, 2\}$ are extended to per-symbol distortions on $N$-tuples by

$$d_i^N(x^N, y^N) = \frac{1}{N} \sum_{k=1}^{N} d_i(x_k, y_k).$$

Without ambiguity, the subscript $N$ in $d^N$ is omitted in this chapter.

**Definition 2.1.1.** A length-$N$ watermarking code subject to distortion $D_1$ is a triple $(\mathcal{M}, f_N, \phi_N)$, where

- $\mathcal{M}$ is the message set of cardinality $|\mathcal{M}|$;
- $f_N : \mathcal{S}^N \times \mathcal{M} \to \mathcal{X}^N$ is the encoder mapping a covertext sequence $s^N$ and a watermark message $m$ to a sequence $x^N$. This mapping is subject to the distortion constraint $\mathbb{E}d_1(s^N, f_N(s^N, m)) \leq D_1$ and the sequence $x^N = f_N(s^n, m)$ is called a stegotext;
- $\phi_N$: $\mathcal{Y}^N \to \mathcal{M}$, $\hat{m} = \phi_N(y^N)$ is the watermark decoder mapping the received forgery sequence $y^N$ to a decoded message $m$.

If the watermark decoder can access to the covertext, then it is called a **private** decoder otherwise it is called a **public** decoder. We only consider the public watermarking decoder in this chapter.

**Definition 2.1.2.** An attack channel with memory, subject to distortion $D_2$ , is a sequence of conditional pmfs $A^N(y^N|x^N)$ from $\mathcal{X}^N$ to $\mathcal{Y}^N$, such that $\mathbb{E}d_2(x^N, y^N) \leq D_2$. Denote this class of attack channels by $\mathcal{A}^N(D_2)$.

Moreover, $R = \frac{1}{N} \log |\mathcal{M}|$ is called its watermark embedding rate. Given a watermarking encoder and watermarking decoder pair $(f_N, \phi_N)$, the error probability of watermarking is defined by $P_e = Pr\{\hat{M} \neq M\}$.

**Definition 2.1.3.** A rate $R$ is achievable for distortion $D_1$ and for a class of attack channels $\{\mathcal{A}^N,\ N \geq 1\}$ , if there is a sequence of codes $(\mathcal{M}, f_N, \phi_N)$, subject to distortion $D_1$, with rate $R$ such that $\sup_{A^N \in \mathcal{A}^N} P_e(A^N) \to 0$ as $N \to \infty$.

Consider an auxiliary random variable $U$ defined over a finite set $\mathcal{U}$ of cardinality $|\mathcal{U}| \leq |\mathcal{X}||\mathcal{S}| + 1$. When the attack channel $A(y|x)$ is a fixed known one, the information hiding capacity is given by [5]

$$C = \max_{p_{X,U|S}} I(U;Y) - I(U;S) \tag{2.1}$$

where the sequence $x^N$ satisfies the distortion constraint $\mathbb{E}d_1(s^N, x^N) \leq D_1$. In the more general case, watermark embedding can be thought of as a game between two players, the information hider (including watermark encoder and decoder) and the attacker, in cases where the attack channel is not fixed and known. The first player tries to maximize a payoff function (e.g., achievable rate), and the second one tries to minimize it. The information available to each player critically determines the value of the game. In our scenario, we assume that the information hider chooses the encoder $f_N$ and the attacker is able to learn $f_N$ and choose the attack channel $A_{Y^N|X^N}(y^N|x^N)$ accordingly. We also assume that the decoder knows the attack channel $A_{Y^N|X^N}(y^N|x^N)$ and chooses $\phi_N$ accordingly. These assumptions may be too optimistic. In [6] [8] a conservative approach for the watermark encoder and the decoder is to assume that they are unable to know $A_{Y^N|X^N}(y^N|x^N)$, but the attacker is able to find out both $f_N$ and $\phi_N$ and design the attack channel accordingly.

**Definition 2.1.4.** A memoryless covert channel subject to distortion $D_1$ is a conditional distribution $Q_{X,U|S}(x, u|s)$ from $\mathcal{S}$ to $\mathcal{X} \times \mathcal{U}$ such that

$$\sum_{x,s,u} d_1(s, x) Q_{X,U|S}(x, u|s) P(s) \leq D_1 \tag{2.2}$$

The class $\mathcal{Q}$ is the set of all memoryless covert channels subject to distortion $D_1$. The class $\mathcal{A}(Q, D_2)$ is the set of all memoryless attack channels subject to distortion $D_2$ under covert channels from the class $\mathcal{Q}$. An expression for the information-hiding capacity is derived in terms of optimal covert and attack channels in [5] [1].

**Theorem 2.1.1.** *Assume that for any $N \geq 1$, the attacker knows $f_N$, and the decoder knows both $f_N$ and the attack channel. A rate $R$ is achievable for distortion $D_1$ and attacks in the class $\{\mathcal{A}(f_N)\}$ if and only if $R < C$, where*

[1] In [5], authors did not succeed to prove the converse part of the theorem 2.1.1. however, the conclusion of this theorem is well accepted to be correct.

$$C = \max_{Q_{X,U|S}(x,u|s)\in\mathcal{Q}} \min_{A_{Y|X}(y|x)} \{I(U;Y) - I(U;S)\} \tag{2.3}$$

*and $U$ is a random variable defined over an alphabet $\mathcal{U}$ of cardinality $|\mathcal{U}| \leq |\mathcal{X}||\mathcal{S}| + 1$, and the random variables $U$, $S$, $X$, $Y$ are jointly distributed as $P_{U,S,X,Y}(u,s,x,y) = P(s)Q_{X,U|S}(x,u|s)A_{Y|X}(y|x)$, i.e. $(U,S) \to X \to Y$ forms a Markov chain.*

A particular interesting case is also studied in [5] and [6] i.e. watermarking in memoryless attack channels with Gaussian covertext. Consider the case of a Gaussian $S$ and the squared-error distortion measure $d(x,y) \triangleq d_1(x,y) = d_2(x,y) = (x-y)^2$. Here $\mathcal{S} = \mathcal{X} = \mathcal{Y} = \mathbb{R}$, and $S \sim \mathcal{N}(0,\sigma^2)$. The class of attack channels is $\mathcal{A}(Q, D_2)$. And we have the following theorem for Gaussian case [5]

**Theorem 2.1.2.** *Let $\mathcal{S} = \mathcal{X} = \mathcal{Y} = \mathbb{R}$ and $d(x,y) = (x-y)^2$ be the squared-error distortion measure. Assume that $D_2 < (\sigma + \sqrt{D_1})^2$. Let $a$ be the maximizer of the expression*

$$f(a) = \frac{[(2a-1)\sigma^2 - D_2 + D_1][D_1 - (a-1)^2\sigma^2]}{[D_1 + (2a-1)\sigma^2]D_2}$$

*in the interval $(a_{\inf}, 1 + \sqrt{D_1}/\sigma)$, where*

$$a_{\inf} = \max\left(1, \frac{\sigma^2 + D_2 - D_1}{2\sigma^2}\right).$$

*Then we have the following.*

*(a) If $S$ has Gaussian distribution with zero mean and variance $\sigma^2$, the embedding-capacity is given by*

$$C = \frac{1}{2}\log\left(1 + \frac{[(2a-1)\sigma^2 - D_2 + D_1][D_1 - (a-1)^2\sigma^2]}{[D_1 + (2a-1)\sigma^2]D_2}\right). \tag{2.4}$$

*and the optimal covert channel is given by $X = aS + Z$ and $U = \alpha S + Z$, where $Z \sim \mathcal{N}(0, D_1 - (a-1)^2)\sigma^2$ is independent of $S$. The optimal attack channel $A(y|x)$ is the Gaussian test channel given by*

$$A^*(y|x) = \mathcal{N}(\beta^{-1}x, \beta^{-1}D_2)$$

*where $\beta = \frac{(2a-1)\sigma^2 + D_1}{(2a-1)\sigma^2 - D_2 + D_1}$ and $\alpha = \frac{D_1 - (a-1)^2\sigma^2}{D_1 - (a-1)^2\sigma^2 + \beta D_2}$.*

*(b) If $S$ is non-Gaussian with zero mean and variance $\sigma^2$, (2.4) is the upper bound on embedding capacity.*

## 2.2 Joint Lossy Compression and Watermarking

Another aspect of the watermarking problem is that of joint lossy compression and watermarking. The problem is as follows: there is a set of messages to be embedded in the covertext meanwhile the composite signal is compressed subject to some distortion constraint. The embedded message must be reliably decodable without access to the original host data, either directly from the stegotext or from its forgery. Although the compression of the composite sequence can be lossless, the entire process must be lossy since the reconstruction of the covertext from stegotext cannot be perfect after the watermark embedding.

The difference between this model and the model presented in Fig. 2.1 is the compression of the stegotext $X^N$. The watermark encoder, in this setting, conveys the covertext $S^N$ and the message $m$ through an encoding function $f_N$, by producing the watermarked signal $X^N = f_N(S^N, m)$. Here, the stegotext $X^N$ is entropy-coded, i.e., compressed in a blockwise manner using the optimum lossless code and the corresponding watermarked signal rate should not exceed a prescribed value $R_c$. The compressed watermarked signal is sent to the decoder. A simple way to express it is that we add a constraint to the original model in Fig. 2.1, i.e.

$$\frac{H(f_N(S^N, m))}{N} \leq R_c. \tag{2.5}$$

In this case, the Nash equilibrium of the game between the watermark embedder and the attacker has not been found yet. However, two interesting cases, when both of the covertext and the attack channel are discrete memoryless and both of them are Gaussian, have been considered in [9] and [7] respectively. We refer them as Discrete Memoryless Case and Gaussian Case respectively.

### 2.2.1 Discrete Memoryless Case

Let $\Omega$ denote the set of all triples $(U, S, X)$ of random variables taking values in the finite sets $\mathcal{U}$, $\mathcal{S}$, $\mathcal{X}$, where $\mathcal{U}$ is an arbitrary finite alphabet of size $|\mathcal{U}| \leq |\mathcal{S}||\mathcal{X}| + 1$, and the joint probability distribution of $(U, S, X)$, $P_{U,S,X}(u, s, x)$, is such that the marginal distribution of $S$ is $P_S(\cdot)$, and $\mathbb{E}d_1(s^N, x^N) \leq D_1$. For any triple $(U, S, X)$, there exists a related quadruple $(U, S, X, Y)$, with $Y$ taking values in $\mathcal{Y}$, such that

$$P_{U,S,X,Y}(u, s, x, y) = P_{U,S,X}(u, s, x)P_{Y|X}(y|x).$$

where $P_{Y|X}(y|x)$ is a transition probability of the discrete stationary memoryless attack channel. Then the following theorem is obtained in [9].

**Theorem 2.2.1.** *Let $R(D)$ be the rate distortion function for source $P_s(\cdot)$. The information hiding capacity for a discrete memoryless covertext $\mathcal{S}$, a memoryless attack channel $A_{Y^N|X^N}(y^N|x^N)$ and $R_c \geq R(D_1)$ is given by*

$$C(R_c, D_1) = \max_{(u,s,x)\in\Omega} \min\{I(U;Y) - I(U;S), R_c - I(S;U,X)\}. \tag{2.6}$$

An alternative coding scheme to Gel'fand and Pinsker's coding scheme [4] was then proposed, which takes into account the compression. This coding scheme utilized the classical random coding technique in information theory [2] and it is listed as follows.

1. *Code book generation*

   For each message $m$, generate $2^{NR_0}$ codewords

   $$U^N(m,j) \in \{u^N(m,1), \ldots, u^N(m, 2^{NR_0})\},$$

   i.i.d. according to the distribution $P_U(\cdot)$. For each codeword $u^N(m,j)$, generate $2^{NR_x}$ composite sequences $X^N(m,j,k) \in \{u^N(m,j,1), \ldots, u^N(m,j,2^{NR_x})\}$ i.i.d. according to the distribution $P_{X|U}(\cdot|\cdot)$. Let

   $$\mathcal{C}(m,j) = \{u^N(m,j,1), \ldots, u^N(m,j,2^{NR_x})\}.$$

2. *Encoding/Embedding*

   Given the watermark message $m$ and the state sequence $s^N$, the encoder seeks a codeword in bin $m$ that is jointly typical with $s^N$, say $u^N(m,j)$. The first composite sequence found in $\mathcal{C}(m,j)$ that is jointly typical with $(s^N, u^N(m,j))$, say $x^N(m,j,k)$, is chosen for transmission. If there exist more than one such sequence, the described above process is applied to the first matching $u^N(m,j)$ found in a bin's list. If no such $u^N(m,j)$ exists declare an encoding error.

3. *Decoding*

   The decoder finds $\hat{m}$ and $\hat{j}$ such that $u^N(\hat{m},\hat{j})$ is jointly typical with channel output sequence $y^N$. If there exist more than one such pair $(\hat{m},\hat{j})$, or no such pair exits at all, declare a decoding error. The probability of encoding failure

goes to zero as long as $R_0 \geq I(U;S)$ and $R_x \geq I(S;X|U)$, and the probability of decoding failure goes to zero as long as $R_w + R_0 \leq I(U;Y)$. Thus, the overall probability of error goes to zero as long as $R_w \leq I(U;Y) - I(U;S)$ and $R_x \geq I(S;X,U)$. Now, since the compression procedure applied to the composite sequences is lossless, it satisfies $R_c \geq R_w + Rx \geq R_w + I(S;U,X)$. Therefore, $R_w \leq \min\{I(U;Y) - I(U;S), R_c - I(S;U,X)\}$.

### 2.2.2 Gaussian Case

In this case, we assume both of the covertext and the attack channel are Gaussian as shown in Fig. 2.2. No closed-form expressions for the rate region of watermarking embedding rate $R_w$ versus composite rate $R_c$ have been found yet. In [7], Karakos and Papamarcou established the achievable rate region in the terms of the relations between the composite rate, the embedding rate, and the prescribed distortion constraint for the private decoder case and it can serve as an outer bound of the Gaussian case when the watermark decoder is public. It is stated as follows.

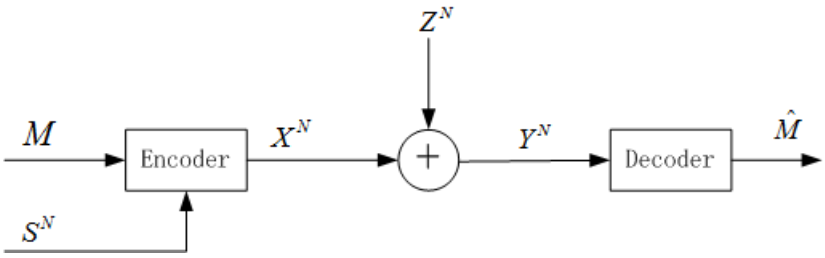


Figure 2.2: Gaussian joint compression and watermarking model.

**Theorem 2.2.2.** *Assume covertext $S^N$ is i.i.d. Gaussian with zero mean and variance $\sigma_s^2$ and the attack is additive i.i.d Gaussian noise with zero mean and variance $D_2$. A private, continuous alphabet joint watermarking and compression code $(2^{nR_c}, 2^{nR_w}, n)$ satisfies requirements*

$$\frac{1}{N}\mathbb{E} \parallel S^N - X^N \parallel^2 \leq D_1$$

*and*

$$Pr\{\hat{M} \neq M\} \to 0 \text{ as } N \to \infty,$$

*respectively, if and only if $(R_c, R_w) \in R_{D_1,D_2}$ where $R_{D_1,D_2}$ is defined as*

$$R_{D_1,D_2} = \left\{ \begin{array}{c} (R_c, R_w) : R_c \geq \left[\frac{1}{2}\log(\frac{\sigma_s^2}{D_1})\right]^+, \\ R_w \leq \max_{\gamma \in [\sigma_s^2, 2^{2R_c}]} \min\{R_c - \frac{1}{2}\log(\gamma), \frac{1}{2}\log(1 + \frac{P_w(\gamma)}{D_2})\} \end{array} \right\} \quad (2.7)$$

*where*

$$P_w(\gamma) = \frac{\gamma(\sigma_s^2 + D_1) - 2\sigma_s^2 + 2\sqrt{\sigma_s^2(\gamma D_1 - \sigma_s^2)(\gamma - 1)}}{\gamma} \qquad \sigma_s^2 \geq D_1.$$

Specifically, we investigate the relationship between watermarking and composite rates in the presence of additive memoryless Gaussian noise, for the quantization index modulation (QIM) watermark embedding system which is widely used in real applications.

- Regular QIM [10], where no knowledge of the covertext is available at the decoder (public scenario).

In the context of QIM for Gaussian case, the attack channel is none other than AWGN channel and the auxiliary sequences $U^N$ are the source codewords themselves. Therefore, in the review of the rate region in (2.6), we have the convert channel given as $U = X$, which leads to the following relationships:

$$\begin{aligned} R_c &= I(X;Y) = \frac{1}{2}\log\left(1 + \frac{P_X}{D_2}\right) \\ R_w &= I(Y;X) - I(S;X) \end{aligned}$$

where $P_X$ is the variance of stegotext $X^N$ and $D_2$ is the variance of additive noise.

Therefore, we can obtain the rate region for this scenario as follows

$$R_w = \left[R_c - \frac{1}{2}\log\left(\frac{\sigma_s^2 D_2(2^{2R_c} - 1)}{\sigma_s^2 D_2(2^{2R_c} - 1) - \frac{1}{4}(\sigma_s^2 + D_2(2^{2R_c} - 1) - D_1)^2}\right)\right]^+$$

where $\sigma_s^2$ is the variance of the covertext.

A numerical result with $\sigma_s^2 = 1$, $D_1 = 0.5$ and $D_2 = 0.25$ is shown in Fig. 2.3. Compared with the outer bound given by (2.7), which is the straight line in the figure, we can see that there is a huge gap between the rate region of the QIM joint compression and watermarking scheme and its outer bound. Finding the optimal convert channel, i.e., the optimal auxiliary variable $U$ for Gaussian covertexts and attack channels is now still an open problem.

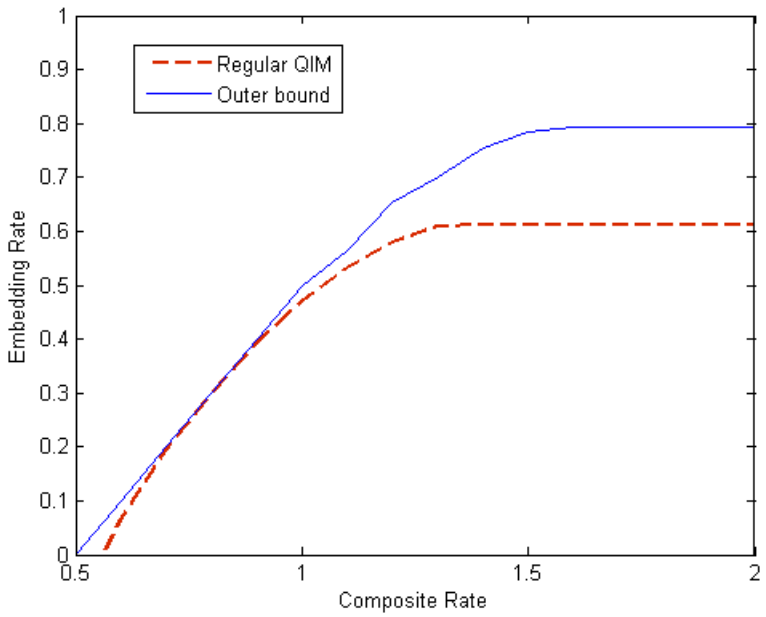


Figure 2.3: Achievable rate region for public QIM and private additive Gaussian case is its outer bound.

## 2.3 Chapter Summary

In this chapter, we briefly reviewed the digital watermarking and joint watermarking and compression model from information-theoretical point of view. Basically, watermark embedding can be viewed as a game between two cooperative players (the watermark encoder and watermark decoder) and an opponent (the attacker). When there is no rate constraint on the stegotext, it has been found that both of the optimal convert channel and attack channel are memoryless which give the saddlepoints of the game. If there is a rate constraint on the stegotext, which gives the joint watermarking and compression scenario, the rate region of the embedding rate vs. composite rate for discrete memoryless attack channels and covertext sources has been obtained. The rate region of public QIM when both of the covertext and the attack channel are Gaussian has also been stated therein.

# Chapter 3

# Joint Watermarking and Compression Using Variable-Rate Scalar Quantization

## 3.1 Introduction

In most applications, watermarked signals will be likely stored or transmitted in compressed format. Instead of treating watermarking and compression separately, it is interesting and beneficial to look at joint design of watermarking and compression schemes. In contrast with a vast amount of research in digital watermarking, there are only a few research works in the domain of joint watermarking and compression (JWC). Some *ad hoc* JWC algorithms were proposed for applications in images, audio, and video, however, there is no unified design strategy until an joint compression and watermarking algorithm using fixed-rate scalar quantization (FRSQ), which is for the purpose of robustness in the presence of additive Gaussian attacks, was proposed by Wu and Yang [12]. In the following section, we will first briefly review the previous work of designing efficient embedding systems by quantization index modulation (QIM) developed in [10] and the JWC systems designed by using fix-rate scalar quantization proposed in [12]. Then, we propose an algorithm to design the JWC system using variable-rate scalar quantization (VRSQ) and it is shown that a potential gain can be obtained by using variable-rate scalar quantization to design JWC systems.

## 3.2 Review of Previous Work

Since the subject of watermarking and information embedding has been attracting a vast amount of attention, quite a lot information embedding schemes have been developed recently [3]. In [10], a coding strategy called quantization index modulation (QIM) proposed by Chen and Wornell is now considered as one of the most efficient embedding methods and it can achieve the embedding capacity for several scenarios when the statistics of the attack channel is known to the watermark encoder.

The basic idea of QIM can be explained by looking at the simple problem of embedding one bit in a real-valued sample. Here we have watermark $m \in \{0, 1\}$ (1-bit message), and covertext or host signal $s \in \mathcal{R}$ (1 sample). A scalar, uniform quantizer with step size is defined as $Q(s)$ with step size $\triangle$ is defined as $Q(s) = \Delta\lfloor\frac{s+\frac{\Delta}{2}}{\Delta}\rfloor$. We may use the function $Q(s)$ to generate two new *dithered* quantizers:

$$Q_i(s) = Q(s - d_i) + d_i, \quad i = 0, 1 \tag{3.1}$$

where $d_0 = -\frac{\Delta}{4}$ and $d_1 = \frac{\Delta}{4}$. The reproduction levels of quantizers $Q_0$ and $Q_1$ are shown as circles and crosses on the real line in Fig. 3.1

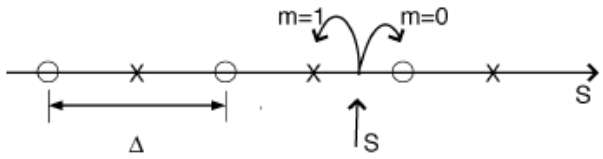


Figure 3.1: Embedding one bit into one sample using original QIM.

One can extend the above dither modulation approach to general quantizers $Q^m(s)$, $m \in \{0, 1\}$ where each $Q^m$ is a mapping from the real line $\mathcal{R}$ to a codebook $B^m = \{b_1^m, b_2^m, \ldots b_L^m\}$. Here all codebooks are assumed to be disjoint [1]. The output values, $b_j^m$, $1 \leq j \leq L$, are referred to as reconstruction points and $L$ is the size of the codebook $B^m$. At the receiver, upon receiving a distorted or corrupted watermarked signal $y$ , one has to form an estimate of the original watermark message so that

[1]The disjoint assumption makes the distinction between $m = 0$ and $m = 1$ easy and hence allows one to use a simple decoder such as the MD decoder; it can be well justified at high distortion-to-noise ratios (DNR). In general, however, if a sophisticated decoder such as the ML decoder which uses source statistics is applied, the codebooks should be allowed to overlap or not disjoint to get better performance at low DNRs.

the error probability $P\{\hat{m} \neq m\}$ is as small as possible. One simple approach is to apply a so-called MD decoder, which first chooses the reconstruction point closest to and then extracts the watermark accordingly, i.e.

$$\hat{m}(y) = arg \min_{m \in \{0,1\}} \|y - Q^m(y)\| \tag{3.2}$$

In [12], Wu and Yang proposed a joint watermarking and compression (JWC) strategy using fixed-rate scalar quantization to maximize robustness against additive white Gaussian (AWGN) attacks. In JWC, the quantization level $L$ is finite. Associated with the quantizer $Q^m$ is a partition of the real line $R$ into $L$ quantization cells $C_j^m$. The $j$th quantization cell

$$C_j^m = \{s \in R : Q^m(s) = b_j^m\} = [z_{j-1}^m, z_j^m) \tag{3.3}$$

is an interval corresponding to the input range of $b_j^m$, where $z_{j-1}^m$, and $z_j^m$ are defined as end points of the $C_j^m$ if $1 \leq j \leq L-1$, $z_0^m = -\infty$ and $z_L^m = +\infty$. By mapping $(m, s) \in \{0, 1\} \times R$ into $Q^m(s)$, the covertext signal is jointly watermarked and compressed. Thus, as a mapping from $\{0, 1\} \times R$ to $B^0 \cup B^1$ serves as a binary JWC encoding scheme using fixed-rate scalar quantization. To design a JWC system, an optimal decoding rule first needed to be found. By simulations [12], it has been shown that when distortion to noise ratio (DNR) is larger than 4.77dB, which is the minimum DNR required to achieve the embedding capacity of one bit per sample, the performance of the minimum distance (MD) decoder approaches that of the maximum likelihood (ML) decoder. That is to say, in the DNR region of practical interest, we can use the MD decoder instead of the ML decoder as the former has low implementation complexity. Based on MD decoding rule, the decoding bit error probability $P_e$ of the corresponding system is defined as follows

$$P_e = \frac{1}{2} \sum_{m \in \{0,1\}} \sum_{j=1}^{L} P(s \in C_j^m) P_{j,e}^m \tag{3.4}$$

where $P(s \in C_j^m)$ is the probability that $s$ lies in $C_j^m$. $P_{j,e}^m$ is the conditional decoding bit error probability given $m$ and given the fact that the covertext $s$ lies in the quantization cell $C_j^m$. In the case of an AWGN attack channel with a noise variance $\sigma_n^2$, the conditional bit error probability $P_{j,e}^m$ is given by

$$\begin{cases} P_{j,e}^0 = Q(\mid \frac{(b_L^0+b_L^1)-2b_j^0}{2\sigma_n} \mid) + \sum_{i=1}^{L-1} \mid Q(\mid \frac{(b_i^0+b_i^1)-2b_j^0}{2\sigma_n} \mid) - Q(\mid \frac{(b_i^1+b_{i+1}^0)-2b_j^0}{2\sigma_n} \mid) \mid \\ \\ P_{j,e}^1 = Q(\mid \frac{(b_1^0+b_1^1)-2b_j^1}{2\sigma_n} \mid) + \sum_{i=2}^{L} \mid Q(\mid \frac{(b_{i-1}^1+b_i^0)-2b_j^1}{2\sigma_n} \mid) - Q(\mid \frac{(b_i^0+b_i^1)-2b_j^1}{2\sigma_n} \mid) \mid \end{cases} \tag{3.5}$$

when $L$ is even and $Q(x) = (1/\sqrt{2\pi}) \int_x^\infty e^{-\frac{t^2}{2}} dt$. A similar formula can be obtained when $L$ is odd. Assume the squared error distortion measure is used. Since watermark messages $m$ are equally likely, the average embedding/quantization distortion can be expressed as

$$D(S,X) = \frac{1}{2} \sum_{m\in\{0,1\}} \sum_{j=1}^{L} \int_{z_{j-1}^m}^{z_j^m} (s - b_j^m)^2 p(s) ds \tag{3.6}$$

where $p(s)$ is the probability density function of the host signal. To design the optimal joint robust watermarking and compression system, is just to minimize the decoding error probability $P_e$ under the constraint of distortion no more than $D$, i.e. to solve the following constrained optimization problem:

$$\begin{cases} \text{Minimize} \quad P_e, \quad \text{subject to} \\ D(S,X) \le D \end{cases} \tag{3.7}$$

A Lagrangian method can be applied to solve the above problem, that is to convert it to the following unconstrained problem

$$W(B,Z,\lambda) = P_e(B) + \lambda D(B,Z) \tag{3.8}$$

where the codebook set $B = \{B^0, B^1\}$, the end point set $Z = \{z_1^0, z_2^0, \ldots, z_{L-1}^0, z_1^1, z_2^1, \ldots, z_{L-1}^1\}$ and $\lambda \ge 0$. The distortion function $D(B,Z)$ is defined as before in (3.6).

Note that if each point $z_j^m$ is force to have the relation as $z_j^m = \frac{1}{2}(b_j^m + b_{j=1}^m)$ for $1 \le j \le L$ to minimize the distortion, the bit error $P_e$ is a function of codebook set $B$. An alternating algorithm was developed based on Lloyd-Max algorithm [16] to solve the above unconstrained optimization problem and the convergence analysis of the algorithm was also stated therein.

## 3.3 Problem Formulation

Though in some applications fixed-rate scalar quantization is preferred with the advantage of low implementation complexity, low time delay and immunity to error propagation for transmission over noisy channel, more efficient compression could be achieved by applying variable-rate scalar quantization which uses entropy coding. This fact leads us to the following questions

- Is there any potential gain we can obtain if we design our JWC system using variable-rate scalar quantization?
- Based on the constraints on compression rate and encoding distortion, how to design the optimum JWC scheme to maximize the tradeoff between robustness and rate-distortion performance of the resulting systems?

So in the following section, we will develop a novel joint watermarking and compression system using variable-rate scalar quantization (VRSQ) to maximize the robustness against AWGN attacks. It is shown that potential gains of bit error probability versus DNR will be obtained.

Before formulating our optimization objective function, we first define the composite rate of the JWC system as the entropy of the stegotext $X^N$, i.e.

$$R = H(X^N) = H(f_N(S^N, m)) = \frac{1}{2}\left[H(Z^0) + H(Z^1)\right] + 1 \tag{3.9}$$

where $Z^0$ and $Z^1$ stand for the two end points of the partitions of the codebook $B = \{B^0, B^1\}$ and watermark $m$ is uniformly distributed. Normally, however, we use (3.9) so as not to tie our results to a particular entropy code, since there are a number of noiseless codes, e.g., arithmetic codes and Ziv-Lempel codes, that achieve average rates quite close to the codeword entropy. Easily to see that composite rate $R$ is the function of end point set $Z = \{Z^0, Z^1\}$, we can rewrite it as follows

$$R(Z) = 1 - \frac{1}{2}\sum_{m\in\{0,1\}}\sum_{i=1}^{L}\int_{z_{i-1}^m}^{z_i^m} p(s)ds \log \int_{z_{i-1}^m}^{z_i^m} p(s)ds \tag{3.10}$$

Now we formulate our objective is to solve the following constrained optimization problem:

$$\begin{cases} \text{Minimize } P_e, \quad \text{subject to} \\ D(S,X) \leq D \\ R(Z) \leq R_c \end{cases} \tag{3.11}$$

This constrained optimization problem was solved in the classic Lagrangian form,

$$J(B,Z,\lambda,\xi) = P_e(B,Z) + \lambda D(B,Z) + \xi R(Z) \tag{3.12}$$

with $\lambda \geq 0$ and $\xi \geq 0$. $B$ and $Z$ denote the codebook set and the end-point set respectively, however, the bit error probability is rewritten as

$$P_e(B,Z) = \frac{1}{2} \sum_{m\in\{0,1\}} \sum_{j=1}^{L} \int_{z_{j-1}^m}^{z_j^m} p(s) ds P_{j,e}^m. \tag{3.13}$$

And it is a function of codebook set $B$ and end point set $Z$. In order to make equations (3.13) and (3.9) holds all the time, the following two conditions need to be satisfied

$$\begin{cases} b_1^0 \leq b_1^1 \leq \ldots \leq b_j^0 \leq b_j^1 \leq \ldots b_L^0 \leq b_L^1. \\ z_{j-1}^m \leq b_j^m \leq z_j^m \quad \text{for} \quad 0 \leq j \leq L \end{cases} \tag{3.14}$$

which are the constraints given by the relationships between the elements of codebook set $B$ and the points in the end point set $Z$.

The minimization of the Lagrange function (3.12) also leads to the solution of the optimization problem in (3.11).

**Theorem 3.3.1.** *For any $\lambda \leq 0$ and $\xi \leq 0$ the codebook set $B^*(\lambda,\xi)$ and the end point set $Z^*(\lambda,\xi)$ which are the optimal solutions to the problem*

$$\min_B \min_Z J(B,Z,\lambda,\xi) \tag{3.15}$$

*subject to the conditions in (3.14) are also the optimal solutions to the constrained problem in (3.11) subject to the conditions in (3.14) when $D(B^*(\lambda,\xi), Z^*(\lambda,\xi)) = D$ and $R(Z^*(\lambda,\xi)) = R_c$.*

*Proof.* For the optimal solution $B^*(\lambda,\xi)$ and $Z^*(\lambda,\xi)$, we have

$$P_e(B^*,Z^*) + \lambda D(B^*,Z^*) + \xi R(Z^*) \leq P_e(B,Z) + \lambda D(B,Z) + \xi R(Z)$$

Equivalently, we have

$$P_e(B^*, Z^*) - P_e(B, Z) \leq \lambda(D(B, Z) - D(B^*, Z^*)) + \xi(R(Z) - R(Z^*)).$$

Since $D(B, Z) \leq D(B^*, Z^*) = D$ and $R(Z) \leq R(Z^*) = R_c$ and $\lambda \leq 0$, $\xi \leq 0$, we have

$$P_e(B^*, Z^*) \leq P_e(B, Z).$$

That is, $B^*$ and $Z^*$ are the optimal solutions to the rate and distortion constrained problem in (3.11). This complete the proof of the theorem. □

As we sweep $\lambda$ and $\xi$ over the range from zero to infinity, set of solutions $B^*(\lambda, \xi)$ and $Z^*(\lambda, \xi)$ and constraints $D(\lambda, \xi)$ and $R(\lambda, \xi)$ are obtained. We then find the optimal solutions $B^*(\lambda, \xi)$ and $Z^*(\lambda, \xi)$.

## 3.4 Algorithm Design

Since the objective optimization problem (3.12) with conditions (3.14) is a double-minimization problem, in principle, the following alternating minimization procedure can be used to solve it.

- Fix the codebook set $B$, find the optimal end point set $Z$ as follows

$$Z = \arg\min_{Z} \{P_e(B, Z) + \lambda D(B, Z) + \xi R(Z)\}.$$

- Fix the end point set $Z$, find the optimal codebook set $B$ as follows

$$B = \arg\min_{B} \{P_e(B, Z) + \lambda D(B, Z)\}.$$

However, it is difficult to find the minimization of the first step since the entropy function $R(Z)$ of the end-point set is there, which is virtually a concave function. So we adopt the typical method in generalized Lloyd-Max algorithm for vector quantization design [17]. It basically introduces another pmf $\Omega = \{\omega_i^m\}_{i=1}^L$ which refereed as the code-distributions. The optimal code-distributions are given as $\omega_i^m = \int_{z_{i-1}^m}^{z_i^m} p(s)ds$, which are just the probabilities of covertext $S$ falling into the partitions $C_j^m = [z_{i-1}^m, z_i^m)$. Using divergence inequality [2], we have the following fact which decouples the end-point set from the composite rate constraint:

**Fact 3.4.1.** *The entropy of a discrete random variable $X$ with pmf $p = \{p(i)\}$ can be written as*

$$H(X) = \min_{\omega} \sum_i p(i) \log \frac{1}{\omega(i)}$$

*where the minimum is over all sub-pmf's $\omega$, that is, all nonnegative $\omega = \{\omega(i)\}$ for which $\sum_i \omega(i) \leq 1$.*

So double-minimization problem (3.15) can be rewritten as

$$\min_{\{\omega_i^m\}} \min_{B,Z} J(B, Z, \lambda, \xi) = \min_{\{\omega_i\}} \min_{B,Z} \{P(B, Z) + \lambda D(B, Z) - \xi R'(Z)\}$$

$$= \min_{\{\omega_i^m\}} \min_{B,Z} \left\{ \frac{1}{2} \sum_{m\in\{0,1\}} \sum_{j=1}^{L} \int_{z_{i-1}^m}^{z_i^m} \left[ P_{j,e}^m + \lambda (s - b_j^m)^2 - \xi \log \omega_i^m \right] p(s) ds \right\}.$$

Here we omit the constant number 1 and this does not change the minimum. The proposed iterative algorithm for optimization problem (3.12) is summarized as follows.

---

**Algorithm: Joint watermarking and compression using variable-rate scalar quantization (JWC-VRSQ)**

---

1. Select an initial codebook set $B$ satisfying

   $$b_1^0 < b_1^1 < \ldots < b_j^0 < b_j^1 < \ldots < b_L^0 < b_L^1.$$

   The initial $Z$ is set as follows: $z_j^m = \frac{1}{2}(b_j^m + b_{j+1}^m)$ for $1 \leq j \leq L-1$. $z_0^m = -\infty$ and $z_L^m = \infty$. The initial code-distribution $\Omega$ is set as follows: $\omega_j^m = \int_{z_{j-1}^m}^{z_j^m} p(s) ds$. Compute $J(B, Z, \lambda, \xi)$ and denote it by $J^{(1)}$. Set $t = 1$, $B^{(1)} = B$, $\Omega^{(1)} = \Omega$ and $Z^{(1)} = Z$.

2. Fix end point set $Z^{(t)}$ and code-distribution $\Omega^{(t)}$. Update codebook set $B^{(t+1)}$ by

   $$\begin{aligned} B &= \arg\min_B J(B, Z, \lambda, \xi, \{\omega_i^m\}) \\ &= \arg\min_B \left\{ \frac{1}{2} \sum_{m\in\{0,1\}} \sum_{j=1}^{L} \int_{z_{i-1}^m}^{z_i^m} \left[ P_{j,e}^m + \lambda (s - b_j^m)^2 - \xi \log \omega_i^m \right] p(s) ds \right\} \end{aligned}$$

   subject to the conditions in (3.14).

3. Fix codebook set $B^{(t+1)}$ and probabilities for each partition $\omega_j^{m(t)}$. Update end-point set $Z^{(t+1)}$ by

$$z_j^m = \frac{1}{2}(b_j^m + b_{j+1}^m) + \frac{\xi(\log \omega_j^m - \log \omega_{j+1}^m)}{2\lambda(b_{j+1}^m - b_j^m)}.$$

for $1 \leq j \leq L-1$.

4. Fix codebook set $B^{(t+1)}$ and end point set $Z^{(t+1)}$, update code-distribution $\Omega^{(t+1)}$ by

$$\omega_j^m = \int_{z_{j-1}^m}^{z_j^m} p(s)ds. \tag{3.16}$$

Compute $J(B^{(t+1)}, Z^{(t+1)}, \lambda, \xi)$ and denote it by $J^{(t+1)}$.

5. If the minimum distance between distinct points in $B^{(t+1)}$ is less than $\epsilon_1$ or $J^{(t)} - J^{(t+1)} < \epsilon_2$ for some $t$ , where $\epsilon_1$ and $\epsilon_2$ are prescribed thresholds, stop; otherwise continue.

The core of the iterative JWC algorithm is Step 2 and Step 3, i.e. finding the optimal end point set $Z$ given codebook set $B$ and code-distribution $\{\omega_i\}$, and updating codebook set $B$ with code-distribution $\{\omega_i\}$ and end-point set $Z$. These two steps are addressed separately as follows and the convergence analysis of the algorithm is described thereafter.

### 3.4.1 Optimal End-point Set and Codebook Set Updating

Before updating codebook set $B$, we first rewrite the expression of bit error probability as

$$P_e(B, \{\omega_i^m\}) = \frac{1}{2} \sum_{m\in\{0,1\}} \sum_{j=1}^{L} \int_{\tilde{z}_{j-1}^m}^{\tilde{z}_j^m} p(s)ds P_{j,e}^m \tag{3.17}$$

where $\tilde{z}_j^m = \frac{1}{2}(b_j^m + b_{j+1}^m) + \frac{\xi(\log \omega_j^m - \log \omega_{j+1}^m)}{2\lambda(b_{j+1}^m - b_j^m)}$. So the objective function in Step 2 is rewritten as

$$\min_B \left\{ P_e(B, \{\omega_i^m\}) + \frac{1}{2} \sum_{m\in\{0,1\}} \sum_{j=1}^{L} \int_{z_{i-1}^m}^{z_i^m} \left[\lambda(s - b_j^m)^2 - \xi \log \omega_i^m\right] p(s)ds \right\} \tag{3.18}$$

When the end point set $Z$ and code distribution $\{\omega_i^m\}$ are fixed, we can update the reconstruction points by the feasible direction method in nonlinear programming [15] to minimize $J(B, Z, \lambda, \xi, \{\omega_i^m\})$ in Step 2. The feasible direction operation is an iterative mapping for the minimization of $J(B, Z, \lambda, \xi, \{\omega_i^m\})$. The $i$th iteration starts with the reconstruction points, which satisfies (3.14) and looks for a feasible direction of displacement such that a small step in that direction does not lead out of the constraint (3.14) and decreases strictly. We then move some distance in this direction, to obtain a new codebook set, which is better than the previous one in terms of the objective function $J(B, Z, \lambda, \xi)$; for instance, we may look for the minimum of in the direction $\mathbf{v}$, subject to not violating the constraint (3.14), i.e., the $i$th iteration generates improved reconstruction points by

$$b_j^{m(i+1)} = b_j^{m(i)} + \alpha^* \nu_j^m \tag{3.19}$$

where the optimum step size $\alpha^*$ is a solution of a (single variable) line search problem. The direction $\mathbf{v}$ can be generated by the following linear programming problem

$$\begin{cases} \text{Minimize } \eta \text{ subject to} \\ \nabla_{b_j^m} J(B, Z, \lambda, \xi)\mathbf{v} - \eta \le 0 \\ b_j^0 - b_j^1 + \nu_j^0 - \nu_j^1 - \eta \le 0 \quad 1 \le j \le L \\ b_j^1 - b_{j+1}^0 + \nu_j^1 - \nu_{j+1}^0 - \eta \le 0 \quad 1 \le j \le L-1 \\ b_j^m - b_{j+1}^m + \nu_j^m - \nu_{j+1}^m + \delta - \eta \le 0 \quad m \in \{0,1\} \\ \sum_{m\in\{0,1\}} \sum_{i=1}^{L} \mid \nu_i^m \mid \le 1 \end{cases} \tag{3.20}$$

Here we treat current end point set $Z$ as the function of codebook set $B$ and the updated code distribution $\{\omega_i^m\}$, i.e. $Z = Z(B, \{\omega_i^m\})$. $\delta = \sqrt{\frac{\xi}{\lambda} \log \frac{\omega_{j+1}^m}{\omega_j^m}}$ and $\nabla_{b_j^m} J(B, Z, \lambda, \xi)$ represents the gradient of $J$ with respect to $B$ only. That is, the direction $\mathbf{v}$ is an optimum solution of (3.20). (Note that (3.20) has to be solved at each iteration; at the $i$th iteration, $b_j^m$ is replaced by $b_j^{m(i)}$.) If $\eta < 0$ , then $\nabla_{b_j^m} J(B, Z, \lambda, \xi)\mathbf{v} < 0$ and hence $\mathbf{v}$ is a direction of descent. In view of [15], it can be shown that there exists a constant $\beta$ such that $b_j^m + \alpha\nu_j^m$, $m \in \{0, 1\}$, $1 \le j \le L$, satisfy the constraint (3.14) for any $0 \le \alpha \le \beta$. The optimum step size at the $i$th iteration is determined by the following formula

$$\alpha^* = \arg \min_{0\le\alpha\le\beta} J(b_j^{m(i)} + \alpha\upsilon_j^m, Z, \lambda, \xi). \tag{3.21}$$

With the above iterative mappings, the objective function decreases as long as $\eta < 0$.

To update the end point set in Step 3, we need to solve the following minimization problem

$$Z = \arg\min_{Z} \{P_e(B, Z) + \lambda D(B, Z) + \xi R'(Z, \{\omega_i^m\})\} \tag{3.22}$$

where

$$R'(Z, \{\omega_i^m\}) = -\frac{1}{2} \sum_{m\in\{0,1\}} \sum_{i=1}^{L} \int_{z_{i-1}^m}^{z_i^m} p(s)ds \log \omega_i^m.$$

Since we have treated the bit decoding error probability as the function of end point set $B$ and $\{\omega_i\}$, i.e. $P_e = P_e(B, Z(B, \{\omega_i\}))$ which now can be taken as a constant since $B$ and $\{\omega_i\}$ are known. Therefore, we can find the optimal solution of the above minimization by taking derivative of (3.22) with respect to $Z$. The minimum is obtained when

$$z_j^m = \frac{1}{2}(b_j^m + b_{j+1}^m) + \frac{\xi(\log \omega_j^m - \log \omega_{j+1}^m)}{2\lambda(b_{j+1}^m - b_j^m)} \tag{3.23}$$

for $1 \le j \le L-1$.

**Remark 3.4.1.** *In Step 3 of the above iterative algorithm, assuming end point set $Z$ is a function of codebook set $B$ is necessary. That is to guarantee that finding the exact decent direction of codebook set $B$ for one updating cycle, i.e. updating both $Z$ and $B$ once.*

### 3.4.2 Convergence Analysis

The convergence of the above algorithm is stated in the following theorem.

**Theorem 3.4.2.** *Fix $\lambda$ and $\xi$. Assume that the probability density function $p(s)$ of the covertext $S$ is continuous and has a finite support. Then the iterative minimization procedure described above with any initial codebook set satisfying (3.14) either terminates at a local optimum or the limit of any convergent subsequence of reconstruction points $b_j^{m(t)}$, $m \in \{0,1\}$, $1 \le j \le L$ and end points $z_j^{m(t)}$, $m \in \{0,1\}$, $1 \le j \le L-1$ is a local optimum.*

*Proof.* To prove this theorem, we will employ Zangwills convergence theorem (1969) [15]. The theorem states as follows: the convergence of the above iterative algorithm depends on the following three sufficient conditions.

1. The codebook set $B$ and the end point set $Z$ are contained in a bounded and closed domain.

2. There exists a continuous descent function.

3. The iterative mapping associated with the feasible direction operation for codebook set updating and the optimal partitions updating is closed (see [15] for the definition of closed mapping).

Under the assumption that $p(s)$ has a finite support, it is easy to see that $B$ and $Z$ are contained in a bounded and closed domain. Therefore, to apply Zangwills convergence theorem, it suffices to show that the point-to-set map defined by the alternative minimization procedure in Steps 2, 3 and 4 of the JWC-VRSQ algorithm is closed and there exists a continuous descent function relative to this map.

Let $\mathcal{A}$ denote the point-to-set map specified in Steps 2, 3 and 4 of the JWC-VRSQ algorithm. Starting with an initial codebook set $B^{(1)}$ and end point set $Z^{(1)}$ , the algorithm generates a sequence of codebook sets and end point sets $(B^{(t)}, Z^{(t)})$ for which $(B^{(t+1)}, Z^{(t+1)}) \in \mathcal{A}(B^{(t)}, Z^{(t)})$, i.e. $(B^{(t+1)}, Z^{(t+1)})$ is obtained from $(B^{(t)}, Z^{(t)})$ by one application of Step 2, 3 and 4 in the JWC-VRSQ algorithm. Since Step 2 includes two mini-steps, the point-to-set map $\mathcal{A}$ is actually a composition of five point-to-set maps: $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{A}_3$, $\mathcal{A}_4$ and $\mathcal{A}_5$. Here, $\mathcal{A}_1$ associates every end point set $Z$ with induced code distribution $\Omega$, given by (3.16), i.e.,

$$\mathcal{A}_1(B, Z) = \{(B, Z, \Omega) : \Omega \text{ is given by (3.16)}\}.$$

$\mathcal{A}_2$ associates every codebook set $B$, point set $Z$ with the direction $\mathbf{v}$, the optimal solution of (3.20), i.e.,

$$\mathcal{A}_2(B, Z, \Omega) = \{(B, Z, \Omega, \mathbf{v}) : \mathbf{v} \text{ is an optimal solution of (3.20)}\}.$$

The point-to-set map $\mathcal{A}_3$ associates $(B, Z, \Omega, \mathbf{v})$ with $(B + \alpha^*\mathbf{v}, Z, \mathbf{w})$, i.e.,

$$\mathcal{A}_3(B, Z, \Omega, \mathbf{v}) = \{(B + \alpha^*\mathbf{v}, Z, \Omega) : \alpha^* = \arg \min_{0 \le \alpha \le \beta} J(b_j^{m(i)} + \alpha v_j^m, Z, \lambda, \xi)\}.$$

$\mathcal{A}_4$ represents the map specified in Step 3, which maps $(B, Z, \Omega)$ into $B, \tilde{Z}$, where $\tilde{Z}$ is the end point set obtained from $B$ and $\Omega$ by (3.22). Finally, the last map $\mathcal{A}_5$ is minimizing the objective function $J(B, Z, \lambda, \xi)$ by mapping $(B, \tilde{Z}, \Omega)$ back into $(B, \tilde{Z})$ using (3.16) which is the same as $\mathcal{A}_1$. Since $\mathcal{A}_1$ is a continuous mapping, $\mathcal{A}_1$ is closed. To prove that $\mathcal{A}_2$ is closed, we directly apply lemma 5.3 in [15] to $J(B, Z, \lambda, \xi)$ as a function of both $B$ and $Z$. Note that in the corresponding linear programming problem involving the gradient of $J$ with respect to both $B$ and $Z$, there is no constraint on the direction with respect to $Z$. Therefore, the linear programming problem involving the gradient of $J$ with respect to both $B$ and $Z$ can be decomposed into two independent problems: one given by (3.20) and the other involving the direction with respect to $Z$ only. From this the closeness of $\mathcal{A}_2$ is proved. By using a similar argument to [ Theorem 3.1 and 3.3 ] in [15], one can also show that $\mathcal{A}_3$ is closed. Obviously, $\mathcal{A}_4$ is continuous and hence closed. Therefore, all $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{A}_3$, $\mathcal{A}_4$ and $\mathcal{A}_5$ are closed. Since $\mathcal{A}$ is a composition (or product) of $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{A}_3$, $\mathcal{A}_4$ and $\mathcal{A}_5$ in the indicated order, it shows that $\mathcal{A}$ is closed.

To show that there is a continuous descent function relative to $\mathcal{A}$, let us look at the objective function $J(B, Z, \lambda, \xi)$ itself, which is continuous with respect to $B$ and $Z$. As long as $\eta < 0$ in (3.20), the direction $\mathbf{v}$ is a descent direction for $J(B, Z, \lambda, \xi)$, and hence

$$J(B^{(t+1)}, Z^{(t+1)}, \lambda, \xi) < J(B^{(t)}, Z^{(t)}, \lambda, \xi).$$

On the other hand, if the optimum value $\eta$ of (3.20) is zero, then the present $B^{(t)}$ is a local optimum for the fixed $Z^{(t)}$. Subsequently, can not be updated by Step 4 of the JWC-VRSQ algorithm either. Thus, $(B^{(t)}, Z^{(t)})$ is a stationary point. This completes the proof of the theorem. □

**Remark 3.4.2.** *In the above, the source statistics is assumed to be known. If the source statistics are unknown, one can apply the proposed design algorithm to the training sets.*

## 3.5 Simulation and Comparison

Having described and analyzed algorithms for designing optimum binary JWC encoding schemes using variable-rate scalar quantization, in this section, we evaluate its performance by simulation and comparison with designing JWC systems us-

ing fixed-rate nonuniform scalar quantization in the presence of additive Gaussian attacks.

Consider i.i.d Gaussian covertexts with zero mean and unit variance. Assume that the squared error distortion is used, the minimum distance decoder is employed and the attack channel is an AWGN channel with variance $\sigma_n^2$. Compute and test the bit error probabilities for binary JWC schemes obtained from optimal fixed-rate scalar quantization (FRSQ) in [12] and variable-rate scalar quantization (VRSQ) described above respectively. We plot curves in terms of decoding bit error probability $P_e$ versus distortion noise ratio ($DNR$), where

$$DNR = 10 \log_{10} \frac{D(S,X)}{\sigma_n^2}$$

Fig. 3.2 plots the bit error probabilities versus DNR for the optimum binary JWC systems using VRSQ and FRSQ. To make the comparison fair, we assume that both of the two schemes have the same composite rate, which is $R_c = 4.15$ bits per sample and the encoding distortion constraint is $D = 0.019$. We can see that the optimal binary JWC systems using variable-rate scalar quantization achieve better performance than the optimal binary JWC systems using fixed-rate nonuniform scalar quantization. In particular, the optimum binary JWC systems using the variable-rate scalar quantization method provide about 0.3-dB DNR gain over those using fixed-rate nonuniform scalar quantization in a wide range.

In the simulation, 55 sample sequences of length $10^6$ were processed. The 94% confidence intervals for bit error probability were computed and found to be within 3% of the true value. The prescribed threshold values $\epsilon_1$ and $\epsilon_2$ were set to $10^{-12}$ and $10^{-18}$ respectively for the FRSQ algorithm. For the VRSQ method, $\epsilon_1$ and $\epsilon_2$ were set to $10^{-15}$ and $5 \times 10^{-15}$, respectively. Usually 1000 to 3000 iterations are needed to terminate both of the two algorithms. Although plenty of computing time is needed for running these two methods, the processes are offline. Once the quantization codebooks are determined, the watermark and compression process can be accomplished by the defined encoding rule.

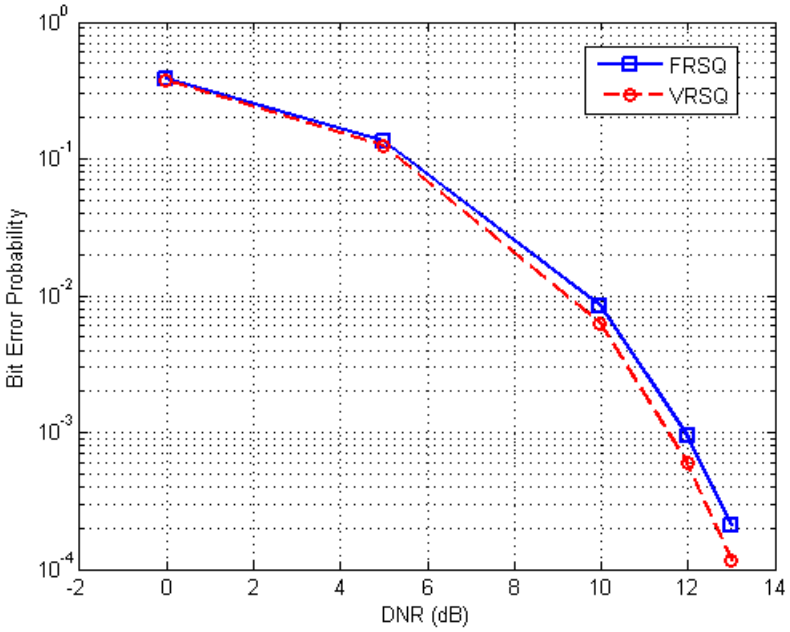


Figure 3.2: Decoding bit error probabilities comparison between VRSQ and FRSQ when composite rate is 4.15 with distortion constraint 0.019.

## 3.6 Chapter Summary

In this chapter, we have investigated the design of JWCs using variable-rate scalar quantization. The MD decoder is first selected as the decoding rule in our subsequent design. The binary JWC encoding scheme using variable-rate scalar quantization (VRSQ) are then presented. Simulation results show that optimum binary JWC systems using variable-rate scalar quantization are better than optimum binary JWC systems using fixed-rate scalar quantization (FRSQ) proposed in [12]. In comparison with the results of JWC systems using FRSQ, optimum binary JWC systems using VRSQ achieve about 0.3-dB DNR gain in the DNR region of practical interest.

# Chapter 4

# Joint JPEG Compression and Robust Watermarking

## 4.1 Introduction

Watermarks designed to survive legitimate and everyday usage of content are referred as *robust* watermarks. Examples of processes a watermark might need to survive include lossy compression, printing and scanning, format conversion, noise reduction and so on. In this chapter, we consider designing the joint compression and watermarking systems which have the robustness to a broad class of valumetric distortion attacks. In the real JWC applications, we have to design our watermark encoder to be compatible with the decoders in current multimedia compression standards, for instance, JPEG in image compression, MPEG-4 and H.264 in video compression. We propose two joint watermarking and compression schemes to embed the data or similar type of information invisibly into images with high payload. As JPEG is a widely used compression format [19] [18], in this chapter, we use JPEG compression as an example to investigate how to maintain or even improve the compression rate distortion performance of a JWC system after a watermark message is embedded. Specifically, given a watermark embedding rate, we develop a joint image compression and blind watermarking system to maximize the compression rate distortion performance while maintaining baseline JPEG decoder compatibility and satisfying the additional constraints imposed by watermarking.

In the following, we first review the previous work on JPEG optimization and a joint JPEG compression and watermarking algorithm proposed in the recent literature. Then, in Section 4.3, we develop a joint odd-even watermarking (OEW) and

JPEG compression algorithm to jointly optimize run-length coding, Huffman coding and quantization table selection which is subject to some constraint imposed by watermark embedding for the purpose of being robust to a class of standard JPEG recompression attacks and additive Gaussian noise attacks respectively. Iterative algorithms are then proposed to maximize the compression rate-distortion performance of the JPEG-compatible JWC systems under the robustness constraints. Then, to obtain the more robustness against other types of valumetric distortion attacks, in Section 4.5, we improve the OEW method to the zero-nonzero watermarking (ZNW) scheme which can survive a class of valumetric distortion attacks including recompression, additive Gaussian and amplitude scaling. Detailed experimental results and comparisons are given in Section 4.6.

## 4.2 Previous Work on JPEG Optimization and Joint JPEG compression and Watermarking

We now review the so called graph-based JPEG joint optimization [20] and a joint JPEG watermarking proposed based on it–DQW algorithm in [23].

### 4.2.1 Graph-based JPEG Joint Optimization

A JPEG encoder consists of three basic steps [19] [18]: The encoder first partitions an input image into $8 \times 8$ blocks and then processes these $8 \times 8$ image blocks one by one in raster scan order (baseline JPEG). Each of these $8 \times 8$ blocks is transformed from the pixel domain to the DCT domain by an $8 \times 8$ DCT. Then the resulting DCT coefficients are then uniformly quantized using an $8 \times 8$ quantization table, whose entries are the quantization step sizes for each frequency bin. After that, the DCT indices from the quantization are then entropy coded using run-length coding and Huffman coding. The JPEG syntax leaves the selection of the quantization step sizes and the Huffman codewords to the encoder provided the step sizes must be used to quantize all the blocks of an image. This framework offers significant opportunity to apply rate-distortion (R-D) optimization at the encoder where the quantization tables and the Huffman tables are two free parameters the encoder can optimize.

Inspired by the fixed-slope universal lossy data compression scheme considered in [24] [25], Yang and wang in [20] proposed a JPEG-compatible joint optimization

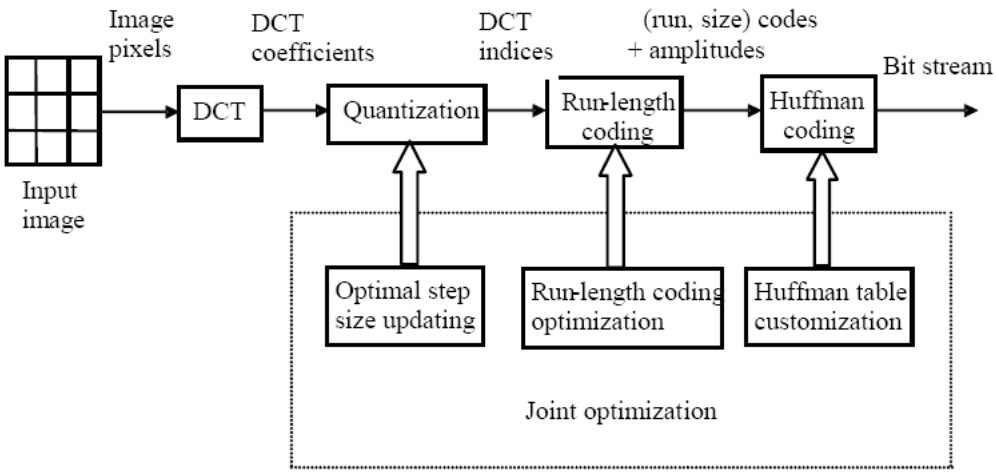


Figure 4.1: Block diagram of joint optimization of the run-length coding, Huffman coding, and quantization step sizes.

algorithm to maximize the compression performance over all possible sequences of run-size pairs $(R, S)$ followed by in category indices amplitudes $A$, all possible Huffman tables $H$, and all possible quantization tables $Q$ in the procedure of JPEG encoding as shown in Figure. 4.1. The free choice of these three parameters in the JPEG syntax provides ample opportunity for the optimization of the compression rate distortion performance. The authors also developed a neat graph-based run-length code iterative optimization algorithm that chooses the sequence $(R, S, A)$, Huffman table, and quantization table iteratively to solve the objective minimization function defined by (4.1).

$$\min_{(R,S,A),H,Q} J(\lambda) = d[I_0, (R, S, A)_Q] + \lambda r[(R, S), H] \tag{4.1}$$

where $d[I_0, (R, S, A)_Q]$ denotes the mean square error distortion between the original image $I_0$ and the reconstructed image determined by $(R, S, A)$ and $Q$ over all AC coefficients, $r[(R, S), H]$ denotes the compression rate for all AC coefficients resulting from the chosen $(R, S, A)$ and $H$, $\lambda$ is a fixed parameter that represents the tradeoff of rate for distortion, and $J(\lambda)$ is the Lagrangian encoding cost.

The iterative algorithm consists of two alternating steps, in which an optimal sequence $(R, S, A)$ is first determined given $Q$ and $H$, and then $Q$ and $H$ are updated when $(R, S, A)$ is fixed. The core of the iterative algorithm is a so called graph-based run-length coding (GBRLC) algorithm, which, given $Q$ and $H$, can

efficiently find an optimal sequence of $(R, S, A)$ to minimize the Lagrangian cost $J(\lambda)$. The optimal sequence $(R, S, A)$ is determined independently for each $8 \times 8$ image block as $J(\lambda)$ is block-wise additive. The graph utilized in the searching of the optimal sequence has 65 states ($0 \leq i \leq 64$). The first 64 states correspond to 64 DCT coefficient indices of an image block in zigzag order. Each state may have incoming connections from its previous 16 states, which correspond to the run $R$, in an $(R, S)$ pair. The last state is called *end* state. The *end* state may have incoming connections from all the other states, which correspond to the EOB (end-of-block) code, i.e, code $(0, 0)$. It may have incoming connections from all states $i$ ($i \leq 62$) where the indices are not equal to zeros. State 63 goes to state end without EOB code. For a given state $i$ ($i \leq 63$) and its predecessor $i$ ($0 \leq i \leq 15$), there are 10 parallel transitions between them which correspond to the size group $S$ in an $(R, S)$ pair. For each state $i$ where $i > 15$, there is one more transition form state $i-16$ to $i$ which corresponds to the pair $(15, 0)$, i.e., ZRL (zero run length) code. Associated with each transition $(r, s)$ is a cost defined as the incremental Lagrangian cost of going from state $i - r - 1$ to state $i$ when the ith DCT coefficient is quantized to size group $s$ (i.e., the coefficient index needs $s$ bits to represent its amplitude) and all the $r$ DCT coefficients appearing immediately before the $i$th DCT coefficient are quantized to zeros. Specifically, this incremental cost is equal to (4.2)

$$\sum_{j=i-r}^{i-1} C_j^2 + \mid C_j - q_i \cdot A_i \mid^2 + \lambda(-\log P(r, s) + s) \tag{4.2}$$

where $C_j$, $j = 1, 2, \ldots, 63$ is the $j$th DCT coefficient, $A_i$ is the chosen amplitude for the $i$th DCT index in size group $s$ that gives rise to the minimum distortion to $C_j$ among all allowed amplitudes within size group $s$, $q_i$ is the $i$th quantization step size and $P(r, s)$ is the probability of pair $(r, s)$, which determines the Huffman table $H$. Similarly, for the transition from state $i$ ($i \leq 62$) to the *end* state, its cost is defined as (4.3)

$$\sum_{j=i+1}^{63} C_j^2 + \lambda(-\log P(0, 0)) \tag{4.3}$$

With these definitions, every sequence of $(R, S)$ pairs of an $8 \times 8$ block corresponds to a path from state 0 to the end state with a Lagrangian cost. The authors then applied a fast dynamic programming algorithm to first find a minimum encoding cost for each state and then determine the optimal sequence $(R, S, A)$ for the whole graph which minimizes the Lagrangian cost.

### 4.2.2 Joint JPEG Compression and Differential Quantization Watermarking

Based on the aforementioned Graph-based JPEG optimization method, Yang and Wu developed a joint JPEG compression and differential quantization watermarking (DQW) algorithm [23] which embedded watermarks into images when images are compressed into JPEG format. The embedded watermark can be detected without the knowledge of the original image and the quantization step sizes in the process of joint embedding and compression mean while it can survive a class of standard JPEG recompression attacks.

The DQW strategy embeds binary watermarks into the JPEG compressed bit stream utilizing the difference of the DCT indices of corresponding positions between adjacent blocks. This procedure can be expressed as follows in (4.4)

$$\mid ID_{a,k} - ID_{b,k} \mid q_k(2m_{ab,k} - 1) \geq m_{ab,k}\Delta_{k,Q_{jpeg}} \tag{4.4}$$

where the watermark bit $m_{ab,k} = 1$ or 0, $ID_{a,k}$ and $ID_{b,k}$ denote the $k$the DCT coefficient indices in block $a$ and $b$ respectively, $q_k$ is the $k$th quantization step size in the quantization table of the proposed JWC alogrithm, and $\Delta_{k,Q_{jpeg}}$ is the $k$the quantization step size in the quantization table of the standard JPEG recompression attack with a quality factor equal to $Q_{jpeg}$. At decoder, the watermark is decoded using the decision rule as follows

$$\begin{cases} \hat{m} = 1, & \text{if } \mid \tilde{\theta}_{a,k} - \tilde{\theta}_{b,k} \mid \geq \delta \\ \hat{m} = 0, & \text{otherwise} \end{cases} \tag{4.5}$$

where $\tilde{\theta}_{a,k}$ and $\tilde{\theta}_{b,k}$ are the $k$th DCT coefficients in blocks $a$ and $b$ of the received and possibly attacked image. $\delta$ is set to $\frac{\Delta_{k,Q_{jpeg}}}{2}$ if the parameter $Q_{jpeg}$ is known at the watermark decoder; otherwise, it is set to 1. The watermark can be fully recovered without the knowledge of the original image and quantization step size if the quality factor of the standard JPEG recompression attack is not less than $Q_{jpeg}$.

By binding the graph-based JPEG optimization and the DQW embedding scheme together, a joint JPEG and DQW algorithm was then proposed. That is, given the watermark embedding rate, actually is to maximize the compression rate distortion performance while remaining faithful to the JPEG syntax and satisfying the additional constraint imposed by DQW embedding. It is indeed to solve the following minimization problem:

$$\begin{cases} \min_{(R,S,A),H,Q} J(\lambda) = d[I_0, (R,S,A)_Q] + \lambda r[(R,s),H] \\ \text{s.t.} \mid ID_{a,k} - ID_{b,k} \mid q_k(2m_{ab,k} - 1) \geq m_{ab,k}\Delta_{k,Q_{jpeg}} \end{cases} \tag{4.6}$$

A modified alternating algorithm was applied to find the local minimum of the cost function $J(\lambda)$ efficiently under watermarking embedding constraint in inequality (4.4). A trellis-and-tree based graph-based run-length coding (GBRLC) algorithm was also developed to embed multiple watermark bits per two blocks with relative low complexity.

## 4.3 Joint JPEG Compression and Robust Watermarking

Based on the GBRLC scheme of [20], we now develop two new joint watermarking and compression schemes to maximize the variability and flexibility a watermark encoder can enjoy when decoding compression syntaxes are given. Both of them can survive standard JPEG recompression attacks. It is shown that our proposed algorithms can achieve higher payload and better compression performance than the previous developed DQW and DEW algorithms.

### 4.3.1 Joint Compression and Odd-Even Watermarking

As described in the last section, the free choice of the three parameters in the JPEG syntax not only provides ample opportunity for the optimization of the compression rate distortion performance but also makes it possible to embed a watermark message into the JPEG compressed bit streams. In this section, we propose an odd-even watermarking (OEW) approach to embedding a watermark message into the compressed bit streams by modifying the quantized DCT coefficient indices in the process of JPEG compression, which can be fully recovered from the attacked images and the watermark decoder does not need to know the original image when decodes watermark messages but the quantization step sizes in the process of JWC are required.

In OEW, we embed binary watermarks into the DCT indices of each $8 \times 8$ DCT block by forcing the the quantized DCT coefficient indices to be odd or even according to the watermarks. This method can be viewed as a special case of the lookup-table (LUT) embedding [26] and quantization index modulation (QIM)

embedding [10]. In more details, we force the amplitude of DCT indices in the embedding positions to be even when a 0 is embedded or to be odd when a 1 is embedded. A more exact expression in mathematical form is that

$$A_i = 2k + m_i \quad k = 0, 1, 2, \ldots \tag{4.7}$$

where $A_i$ is the amplitude of the index of $i$th position ($1 \leq i \leq 64$) in zigzag order of each $8 \times 8$ DCT block, and $m_i = 0$ or $m_i = 1$ is the watermark embedded in $i$th position in zigzag order of this block. The OEW scheme also involves a constraint on the step size in the embedding positions in the quantization table, which is

$$q_i \geq \delta_{attack} \tag{4.8}$$

where $\delta_{attack}$ is the parameter corresponding to the attack channel and the proposed watermarking scheme can be implemented differently according to different classes of attacks. Therefore, in the following, we demonstrate how to implement this joint OEW and JPEG compression scheme for the robustness to standard JPEG recompression attacks and additive Gaussian noise attacks respectively.

### Recompression Attacks

Without loss of generality, we elaborate on the standard JPEG recompression attacks with different quality factors (QFs), that is, the watermarked images are compressed with a default quantization matrix scaled by various scaling factors (SF) to achieve different compression ratios [18]. SF increases with the decrease of QF. Mathematically, the relation is given by

$$SF = \begin{cases} \frac{50}{QF} & \text{if} \quad QF < 50 \\ 2 - \frac{QF}{50} & \text{if} \quad QF \geq 50 \end{cases} \tag{4.9}$$

where $QF$ is in the range of 0-100.

Let's denote the watermarked DCT coefficients in one $8 \times 8$ DCT block resulting from (4.7) as $\bar{\theta}_i$ for $1 \leq i \leq 64$. The property for the watermarked DCT coefficients in the presence of JPEG recompression attacks is shown as follows.

**Theorem 4.3.1.** *Let $\Delta_k$ be the kth quantization step size in the quantization table of the standard JPEG recompression attack and $\Delta_k \leq \Delta_{k,Q_{jpeg}}$. Define $\bar{\theta}_i \equiv I\bar{D}_i \cdot (\Delta_{k,Q_{jpeg}} + 1)$ and $I\tilde{D}_i \equiv$ Integer Round$(\frac{\tilde{\theta}_i}{\Delta_{k,Q_{jpeg}}+1})$ where $\tilde{\theta}_i$ is the DCT coefficient*

*of ith position in the corresponding DCT block after decoding and JPEG re-encoding attacks. Then, we have:*

$$\tilde{ID}_i = \bar{ID}_i. \tag{4.10}$$

*Proof.* Let $\tilde{\theta}_i = \bar{\theta}_i + r_i$. If the JWC image is fully decoded and the re-encoded in the JPEG recompression attacks, the round-off noise in the process of saving images should be considered. Therefore, we have $-\frac{\Delta_k+1}{2} < r_i < \frac{\Delta_k+1}{2}$, then

$$\tilde{\theta}_i - \frac{\Delta_k + 1}{2} \leq \bar{\theta} \leq \tilde{\theta}_i + \frac{\Delta_k + 1}{2}$$

Since

$$\tilde{\theta}_i - \frac{\Delta_{k,Q_{jpeg}} + 1}{2} \leq \tilde{ID}_i \cdot (\Delta_{k,Q_{jpeg}} + 1) \leq \tilde{\theta}_i + \frac{\Delta_{k,Q_{jpeg}} + 1}{2}$$

and $\Delta_k \leq \Delta_{k,Q_{jpeg}}$, we can see $\tilde{ID}_i = \bar{ID}_i$. This complete the proof of the theorem. □

The watermark bit $\hat{m}$ is then decoded by the following decision rule:

$$\hat{m} = m, \text{ if } |(\lfloor \frac{\tilde{\theta}_i}{\Delta_i} + 0.5 \rfloor)| = 2k + m. \tag{4.11}$$

where $\tilde{\theta}_i$ is the $i$th DCT coefficient in one block of the received image and $k$ is 0 or positive integer. $\Delta_i$ is the $i$th step size in the zigzag order of the quantization table. Here it is set to $\Delta_{i,Q_{jpeg}} + 1$ to guarantee zero error decoding. From theorem 4.3.1, the watermark can be fully recovered without the knowledge of the original image if the quality factor of the standard JPEG recompression attack is not less than $Q_{jpeg}$.

Given the watermark embedding rate, we next want to maximize the compression rate distortion performance while remaining faithful to the JPEG syntax and satisfying the additional constraints imposed by OEW scheme. That is, our problem is posed as a constrained optimization problem over all possible sequences of run-size pairs $(R, S)$ followed by in category indices amplitude $A$, all possible Huffman tables $H$ and all possible quantization tables $Q$

$$\begin{cases} \min_{(R,S,A),H,Q} r[(R,S),H] \text{ subjec to} \\ d[I_0, (R,S,A)_Q] \leq d_{budget} \\ A_i = 2k + m_i \quad k = 0, 1, 2, \ldots \\ q_i \geq \Delta_{i,Q_{jpeg}} + 1 \end{cases} \tag{4.12}$$

where $d[I_0, (R, S, A)_Q]$ denotes the distortion between the original image $I_0$ and reconstructed image determined by $(R, S, A)$ and $Q$ over all AC coefficients, and $r[(R, S), H]$ denotes the compression rate for all AC coefficients resulting from the chosen sequence $(R, S, A)$ and the Huffman table $H$ and $d_{budget}$ is the distortion constraint. With the help of the Lagrange multiplier, we may convert the distortion constrained problem into the following unconstrained problem

$$\min_{(R,S,A),H,Q} J(\lambda) = d[I_0, (R, S, A)_Q] + \lambda r[(R, S), H] \tag{4.13}$$

where $(R, S, A)$ and $Q$ are chosen in the set where the watermarking constraints are satisfied. Since a run-size probability distribution $P$ completely determines a Huffman table, we use $P$ to replace the Huffman table $H$ in the optimization process. The proposed iterative algorithm for optimization problem (4.13) is summarized as follows.

**Algorithm 1: Joint compression and OEW under recompression attacks**

1. Initialize a run-size distribution $P_0$ from the given image $I_0$ and an initial quantization table $Q_0$. Set $t = 0$, and specify a tolerance $\epsilon$ as the convergence criterion. Fix a value of $\lambda$.

2. Fix $P_t$ and $Q_t$ for any $t \geq 0$. Find an optimal sequence $(R_t, S_t, A_t)$ that achieves the following minimum

   $$\min_{(R,S,A)} J(\lambda) = d[I_0, (R, S, A)_{Q_t}] + \lambda r[(R, S), P_t]$$

   meanwhile satisfying (4.7). Denote $d[I_0, (R, S, A)_{Q_t}] + \lambda r[(R, S), P_t]$ by $J^t(\lambda)$. For $t > 0$, if $J^{t-1}(\lambda) - J^t(\lambda) \leq \epsilon$, stop the iterative algorithm and output $(R_t, S_t, A_t)$ and $Q_t$; otherwise, go to the next step to continue the iteration.

3. Fix $(R_t, S_t, A_t)$. Update $Q_t$ and $P_t$ into $Q_{t+1}$ and $P_{t+1}$ respectively so that $Q_{t+1}$ and $P_{t+1}$ together achieve the following minimum

   $$\min_{Q,P} J(\lambda) = d[I_0, (R_t, S_t, A_t)_Q] + \lambda r[(R_t, S_t), P]$$

   while satisfying the constraint that $q_i > \Delta_{i,Q_{jpeg}}$. Here $q_i$ is the $i$th quantization step of the quantization table $Q$ in zigzag order where a watermark is embedded.

4. Go to Step 2 for further iterations with $t = t + 1$.

Based on the joint optimization approach for JPEG compression proposed in [20], we develop the following algorithm to fulfill the two steps respectively.

*A. Joint OEW and GBRLC optimization*

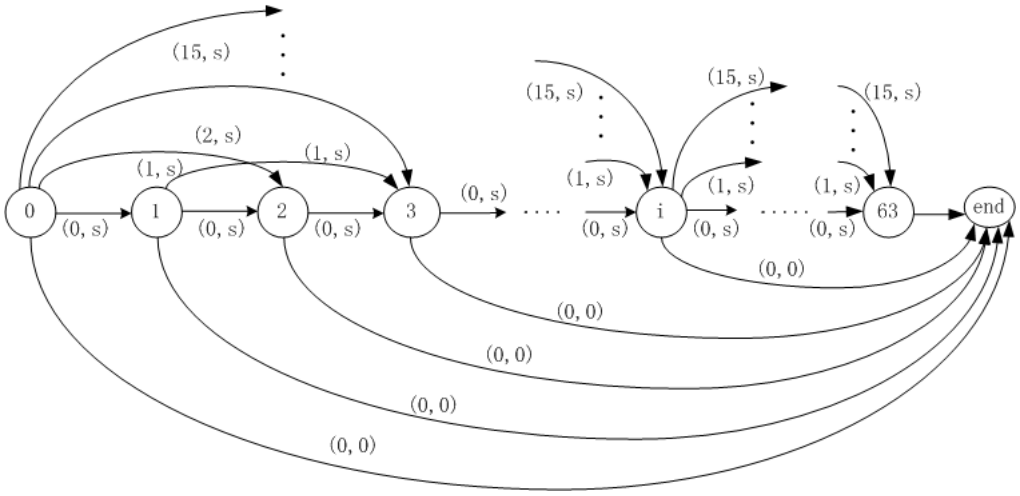


Figure 4.2: Graphic representation of sequences of run-size pairs of an $8 \times 8$ block, where $s$ takes values from 0 to 10 in $(15, s)$ and values from 1 to 10 in other cases.

Fix $H$ and $Q$. We can see that $J(\lambda)$ is block-wise additive when no watermark is embedded. Therefore, a graph-based run-length coding (GBRLC) algorithm developed in [20] and also described in Section 4.2 can be used to find the optimal sequence $(R, S, A)$ for each DCT block independently. We call it a joint graph-based run-length coding and watermark embedding (GBRLCWE) procedure. Basically, we define a directed graph with 65 nodes (or states). As shown in Fig. 4.2, the first 64 states, numbered as $i = 0, 1, \ldots, 63$, correspond to the 64 DCT indices of an $8 \times 8$ image block in zigzag order. Each state $i$ ($i \leq 63$) may have incoming connections from its previous 16 states $j$ ($j < i$) , which correspond to the run, $R$, in an $(R, S)$ pair (in JPEG syntax, $R$ takes value from 0 to 15). For a given state $i$ ($i \leq 63$) and its predecessor $i - r - 1$ ($0 \leq r \leq 15$), there are 10 parallel transitions between them which correspond to the size group, $S$, in an $(R, S)$ pair. For simplicity, we only draw one transition in the graph shown in Figure 4.2; the complete graph needs the expansion of $S$. For each state $i$ where $i > 15$, there is one more transition from state $i-16$ to $i$ which corresponds to the pair $(15, 0)$, i.e., ZRL (zero run length) code.

However, when state $i$ is watermarked using odd-even embedding method, some void $(R, S)$ pairs which contradict with embedding constraints should not appear.

In more details, first, we divide size group $S$ into three subsets, i.e. $S = \{S^0, S^1, S^\emptyset\}$ where $S^0$, $S^1$ and $S^\emptyset$ represent for the size group for watermark 0, watermark 1 and without watermark embedded according to the odd-even embedding constraint; Second, when watermark 1 is embedded into state $i$, the run-length pair $(R, S)$ and the EOB code which is the code $(0, 0)$ after the $j$th $(j \leq i)$ coefficient can not go though state $i$. Therefore, for this case, we treat state $i$ as the stopping state, i.e. the state before state $i$ can not reach the state after state $i$ by whatever run-length pair $(R, S)$ or EOB code, and the state after state $i$ treat it as the starting state 0. After making these changes, we assign to each transition $(r, s)$ from state $i - r - 1$ to state $i$ a cost which is defined as the incremental Lagrangian cost of going from state $i - r - 1$ to state $i$ when the $i$th DCT coefficient is soft-quantized to size group $s$ and all the $r$ DCT coefficients appearing immediately before the $i$th DCT coefficient are soft-quantized to zero. Specifically, this incremental cost is equal to

$$\sum_{j=i-r}^{i-1} C_j^2 + \mid C_j - q_i \cdot A_i \mid^2 + \lambda(-\log P(r, s) + s) \tag{4.14}$$

where the $(r, s)$ pair belongs to the $(R, S)$ set described above for each state $i$.

A more elaborate step-by-step description of the algorithm follows. As an initialization, the algorithm pre-calculates $\lambda \cdot (-\log P(r, s) + s)$ for each run-size pair $(r, s)$ which is valid based on the given run-size distribution $P$. The minimum cost to state 0 (DC coefficient) is initialized as 0 since it does not effect the run-length coding optimization. The algorithm starts with state 1 (the first AC coefficient). The cost associated with each path is calculated using (4.14), where the first term in (4.14) is pre-calculated, and $A_i$ is determined as follows. For simplicity, we only consider positive indices here; negative indices are processed similarly by symmetry. Suppose $A_i'$ is the output of the hard-decision quantizer with step size $q_i$ in response to the input $C_i$ , and it falls into the size group $s' \in S$. If $s = s'$ , $A_i$ is chosen as $A_i'$ since it results in the minimum distortion for $C_i$ in this size group. If $s < s'$, $A_i$ is chosen as the largest amplitude in size group $s \in S$ since this largest amplitude results in the minimum distortion in size group $s$. Similarly, $s > s'$ , $A_i$ is chosen as the smallest amplitude in size group $s$. After the ten incremental costs have been calculated out, we can find the minimum cost to state 1 from state 0 by adding the least incremental cost from state 0 to state 1 to the minimum cost to state 0. Record this minimum cost as well as the run-size pair $(r, s)$ and $A_i$ which results in this minimum cost at state 1. The procedure continues to the next coefficient and so on until the minimum cost to the last coefficient at state 63 is sorted out. By backtracking from the end state with the help of the stored pairs $(r, s)$ and

amplitudes $A_i$ in each state, one can find the optimal path from state 0 to the end state among all the possible paths.

This procedure is a full dynamic programming method, and always gives us the optimal solution. To further reduce its computational complexity, we do not need to compare the incremental costs among the 10 or 11 parallel transitions from one state to another state. Instead, it may be sufficient for us to compare only the incremental costs among the transitions associated with size group $s-1$, $s$ and $s+1$, where $s$ is the size group corresponding to the output of the given hard-decision quantizer satisfying the watermark embedding constraints.

*B. Optimal quantization table updating*

Fix $(R, S, A)$. Huffman table is updated according to empirical run-size distribution and we only need to minimize the distortion part in the Lagrangian cost as the compression rate does not depend on the quantization table $Q$ once $(R, S, ID)$ is given. That is, we need to find the minimum of $d[I_0; (R, S, A)_Q]$ among all possible $Q$ subject to the constraint that $q_i > \Delta_{i,Q_{jpeg}}$ for the $i$th position with watermark embedded. By applying the minimum mean square error criteria with respect to the distortion function, we can obtain the optimum quantization step size

$$\hat{q}_i = \frac{\sum_{j=1}^{NumBlk}(C_{i,j} \cdot A_{i,j})}{\sum_{j=1}^{NumBlk} A_{i,j}^2} \quad i = 0, 1, \ldots 63. \tag{4.15}$$

where $NumBlk$ is the number of $8 \times 8$ blocks in an image. If there is a watermark bit embedded at the $i$th position of a DCT block pair in zigzag order, the quantization step size is determined as follows

$$\hat{q}_i = \max \left\{ (\Delta_{i,Q_{jpeg}} + 1), \lfloor \frac{\sum_{j=1}^{NumBlk}(C_{i,j} \cdot A_{i,j})}{\sum_{j=1}^{Num_Blk} A_{i,j}^2} + 0.5 \rfloor \right\} \tag{4.16}$$

where $\Delta_{i,Q_{jpeg}}$ is the step size of $i$th position in the quantization table of JPEG recompression with quality factor $Q_{jpeg}$.

**AWGN attacks**

In the following, we consider designing a joint JPEG compression and watermarking system which is robust to a class of AWGN attacks. We still utilize odd-even embedding method and it can be shown that an efficient tradeoff among robustness, embedding rate and rate distortion performance can be obtained.

Assume that the additive noise involved by attackers have Gaussian distribution with zero mean and variance $\sigma_n^2$. We first compute the decoding bit error probability $P_{e,i}$ of $i$th position in zigzag order in each $8 \times 8$ DCT block. That is,

$$\begin{aligned} P_{e,i} &= \frac{1}{2} \sum_{m \in \{0,1\}} \sum_{j=-\infty}^{\infty} P(s \in C_{i,j}^m) P_{e,i,j}^m \\ &= \frac{1}{2} \sum_{m \in \{0,1\}} \sum_{j=-\infty}^{\infty} P(s \in C_{i,j}^m) \sum_{i=-\infty}^{\infty} \mid Q(\mid \frac{(2i+0.5)\Delta_i}{\sigma_n} \mid) - Q(\mid \frac{(2i+1.5)\Delta_i}{\sigma_n} \mid) \mid \\ &= \sum_{i=-\infty}^{\infty} \mid Q(\mid \frac{(2i+0.5)\Delta_i}{\sigma_n} \mid) - Q(\mid \frac{(2i+1.5)\Delta_i}{\sigma_n} \mid) \mid \end{aligned} \tag{4.17}$$

where $C_{i,j}^m$ denotes the interval where the DCT transformed signal $s$ is quantized as $j\Delta_i$ and $P_{e,i,j}^m$ is the conditional decoding bit error probability when the watermark $m$ is given and the signal $s$ is quantized as $j\Delta_i$.

It is not hard to prove that the bit error probability of the $i$th position $P_{e,i}$ is a decreasing function for all $\frac{\Delta_i}{\sigma_n} > 0$. We show this property of $P_{e,i}$ in numerical way in Fig. 4.3.

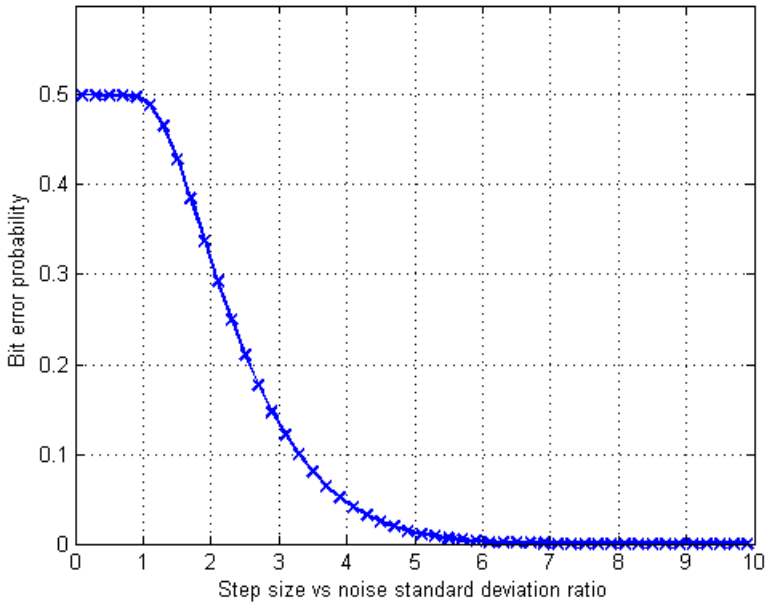


Figure 4.3: Bit error probability $P_{e,i}$ versus $\frac{\Delta_i}{\sigma_n}$.

Based on the expression of bit error probability of the $i$th position, we set up

our Language objective function as follows.

$$\min_{(R,S,A),H,Q} J(\lambda,\xi) = \frac{1}{Numbits} \sum_{i=1}^{Numbits} P_{e,i} + \xi(d[I_0,(R,S,A)_Q] + \lambda r[(R,S),H]) \tag{4.18}$$

where $Numbits$ denotes the number of bits to be embedded per DCT block. Therefore, we have the following alternating algorithm:

---

**Algorithm 2: Joint compression and OEW under AWGN attacks**

---

1. Initialize a run-size distribution $P_0$ from the given image $I_0$ and an initial quantization table $Q_0$. Set $t = 0$, and specify a tolerance $\epsilon$ as the convergence criterion. Fix a value of $\lambda$.

2. Fix $P_t$ and $Q_t$ for any $t \geq 0$. Find an optimal sequence $(R_t, S_t, A_t)$ that achieves the following minimum

   $$\min_{(R,S,A)} J(\lambda,\xi) = d[I_0,(R,S,A)_{Q_t}] + \lambda r[(R,S),P_t]$$

   while satisfying (4.7) for all the watermark embedding position $i$ in each DCT block.

3. Fix $(R_t, S_t, A_t)$. Update $Q_t$ and $P_t$ into $Q_{t+1}$ and $P_{t+1}$ respectively so that $Q_{t+1}$ and $P_{t+1}$ together achieve the following minimum

   $$\min_{Q,P} \left\{ J(\lambda,\xi) = \frac{1}{Numbits} \sum_{i=1}^{Numbits} P_{e,i} + \xi(d[I_0,(R,S,A)_Q] + \lambda r[(R,S),H]) \right\}$$

   Denote $d[I_0,(R,S,A)_{Q_t}] + \lambda r[(R,S),P_t]$ by $J^t(\lambda)$.

4. For $t > 0$, if $J^{t-1}(\lambda) - J^t(\lambda) \leq \epsilon$, stop the iterative algorithm; otherwise, continue.

Here, Step 3 is different from the previous recompression attack case. Based on the property of decoding bit error probability in $i$the position $P_{e,i}$, we find the optimal solution of step size $\hat{q}_i$ as follows:

1. Update $i$th step size in the quantization table as

   $$\hat{q}_i = \frac{\sum_{j=1}^{NumBlk}(C_{i,j} \cdot A_{i,j})}{\sum_{j=1}^{NumBlk} A_{i,j}^2} \quad i = 0, 1, \ldots 63. \tag{4.19}$$

2. If there is a watermark bit embedded into state $i$, keeping increasing $\hat{q}_i$ by 1, until $J^{(t)}(\lambda, \xi)$ begin to increase in Step 3.

Since the bit error probability is a decreasing function of step size $q_i$ for fixed additive noise variance $\sigma_n^2$, by doing above steps it is guaranteed for us to obtain the optimal quantization table $Q$ in Step 3.

### 4.3.2 Joint Compression and Zero-Nonzero Watermarking

Having described the OEW algorithm for designing a joint image compression and watermarking with baseline JPEG decoder compatibility and robustness to JPEG recompression attacks and AWGN attacks, we denote that the quality fact set up in the process of JWC embedding is required at the watermark decoder which means that this scheme is semi-blind. However sometimes it is undesirable when the image is sent through existing large-scale, heterogeneous networks. On the other hand, watermarks are designed to survive legitimate and everyday usage of content. A class of very common distortion attack for everyday usage is so called *valumetric distortions* which change the values of individual pixels of images [1]. In the following, we design a zero-nonzero watermarking (ZNW) scheme for full-blind watermarking with baseline JPEG decoder compatibility which can survive a class of valumetric distortion attacks including additive Gaussian noise, recompression and amplitude scaling. Compared with the previous designed OEW scheme, the ZNW scheme sacrifices some compression performance and embedding rate but obtains more robustness against other types of valumetric distortion attacks.

Similarly to OEW, in ZNW, we embed watermark bit into the DCT indices of every $8 \times 8$ DCT block to force the amplitudes of the DCT indices to be zero when watermark 0 is embedded or nonzero when watermark 1 is embedded. We express it in a mathematical form is that

$$\begin{cases} A_i = 0 \text{ when } m = 0; \\ A_i \neq 0 \text{ when } m = 1. \end{cases} \tag{4.20}$$

To maintain certain robustness, we introduce another constraint on the $i$th step size in zigzag order in the quantization table where a watermark is embedded, i.e.

$$q_i \geq \delta_{attack} \tag{4.21}$$

where $\delta_{attack}$ is the parameter corresponding to the attacks.

At the watermark decoder, the knowledge of original image and the quantization step in the process of JWC are not required. The watemark decoder decodes the watermark bit $\hat{m}$ as follows,

$$\begin{cases} \hat{m} = 0 \text{ when } A_i = 0; \\ \hat{m} = 1 \text{ when } A_i \neq 0. \end{cases} \tag{4.22}$$

Having the embedding strategy, we now need to optimize the rate-distortion function for compression under the constraints given by watermark embedding, i.e.

$$\begin{cases} \min_{(R,S,A),H,Q} J(\lambda) = d[I_0, (R,S,A)_Q] + \lambda r[(R,S), H] \\ \text{subject to (4.20) and (4.21)} \end{cases}$$

An alternating algorithm can also be used for solving the minimization problem:

- Fix $P$ and $Q$. Find an optimal sequence $(R, S, A)$ to minimize $J(\lambda)$ while satisfying (4.20) for all the positions where watermarks are embedded.
- Fix $(R, S, A)$. Update $H$ and $Q$ respectively to achieve the minimum of $J(\lambda)$ while satisfying (4.21)

The detailed embedding procedure is similar to OEW scheme: when implementing the first step of the above algorithm, void run-length pairs $(R, S)$ and void EOB codes for $i$th DCT index where a watermark is embedded should be omitted during the joint ZNW and GBRLC optimization procedure. The size group $S$ in an run-length pair is divided into three subsets which are $S^{\emptyset}$, $S^0$ and $S^1$. They represent that state $i$ has no watermark embedded, watermark 0 embedded and watermark 1 embedded and include different elements according to the zero-nonzero embedding scheme.

When standard JPEG recompression attacks are considered, the optimal $i$th step size in the quantization table $\hat{q}_i$ with watermark embedding is given by

$$\hat{q}_i = \max\{\Delta_{i,Q_{jpeg}}, \lfloor \frac{\sum_{j=1}^{NumBlk}(C_{i,j} \cdot A_{i,j})}{\sum_{j=1}^{Num_Blk} A_{i,j}^2} + 0.5 \rfloor\}$$

where $\Delta_{i,Q_{jpeg}}$ is the step size of $i$th position in the quantization table of JPEG recompression with quality factor $Q_{jpeg}$. It can be seen that by using above ZNW scheme, the watermarks embedded into compressed images can be fully recovered in the presence of standard JPEG recompression attacks with a quality factor not less than $Q_{jpeg}$.

## 4.4 Experiment Results

Having described and analyzed two joint JPEG compression and robust watermarking algorithms against JPEG recompression attacks, in this section, we first evaluate the performance of the proposed algorithms and make a comparison with both the DQW algorithm and the DEW algorithm proposed in the recent literature. Comparative studies of the embedding performance are shown by rate-distortion (R-D) curves, with the distortion being measured by peak signal-to-noise ratio ($PSNR$) which is defined as

$$PSNR = 10 \log_{10} \frac{255^2}{MSE}$$

In practice, watermarked content will be subjected to a variety of distortions before reaching the watermark decoder. Therefore, the watermark encoder is impossible to have the knowledge of attackers in advance. In the this section, we assume that the watermark encoder and decoder are all designed for JPEG-recompression attacks and the robustness to other types of valumetric distortion attacks are also stated therein. Experimental results show that the proposed ZNW scheme against recompression attacks also achieves good robustness to other types of valumetric distortion attacks including additive Gaussian noise and valumetric scaling.

### 4.4.1 DCT Block and Coefficient Positions Selection for Watermark Embedding

Before applying the proposed JWC algorithms to jointly watermarking and compress an image, positions of $8 \times 8$ DCT blocks in the image and DCT coefficients within these blocks need to be determined for watermarking embedding. One possible way is to randomly select the positions based on a secret key. This requires the transmission of the secret key to the watermark decoder, which is sometimes undesirable. In this work, we follow the similar method proposed by Wu in [14], i.e. selecting the positions based on empirical data training to obtain a good tradeoff between perceptual quality and compression RD performance. The position information is then published so that it is the same for all the images. In particular, we select the DCT coefficients with positions from 9 to 20 of an $8 \times 8$ DCT block in zigzag order for watermark embedding. Fig. 4.4 shows the R-D performance of the $512 \times 512$ Lena image after applying the proposed OEW JWC algorithm robust for recompression attack when 1 bit of watermark information per DCT block is

embedded at different DCT positions in the zigzag order. The quality factor for the standard JPEG attacks, $Q_{jpeg}$ is set to 50. Form Fig. 4.4 we can see that compression performance degrades gradually as the embedding position is larger than 21. Similar results were obtained for ZNW JWC, other images and other standard JPEG recompression attacks.

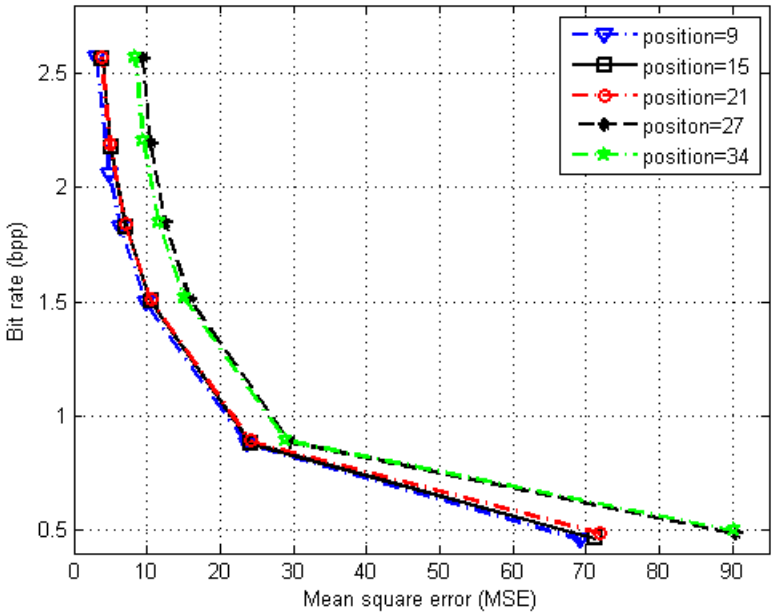


Figure 4.4: Influence of the embedding position on the compression R-D performance.

### 4.4.2 Robust Experiments and Comparisons

In our experiment, watermark embedding rate is set to 1 bit per DCT block in a $512 \times 512$ raw image, i.e. 4096 bits per image if all the DCT blocks all used for embedding. The convergence threshold $\epsilon$ for the two proposed JWC algorithms is set to 0.01 and the resulting average number of iterations is around 10 which can be completed within 12 seconds for a PC with an AMD Turion (tm) $64 \times 2$ $TL$-58+1.90GHz and 2GB memory and a float DCT transform algorithm was implemented.

Figure 4.5 and Figure 4.6 show the PSNR performance of the proposed OEW algorithm, the ZNW algorithm and the DQW JWC algorithm developed in [23] for the $512 \times 512$ image Lena and Barbara respectively. In the figures, OEW stands for

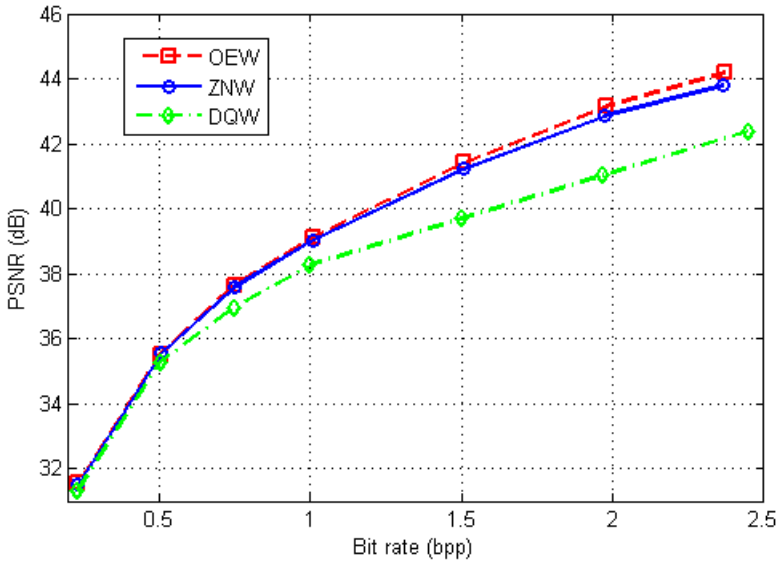


Figure 4.5: Comparison of compression performance for Lena $512 \times 512$.

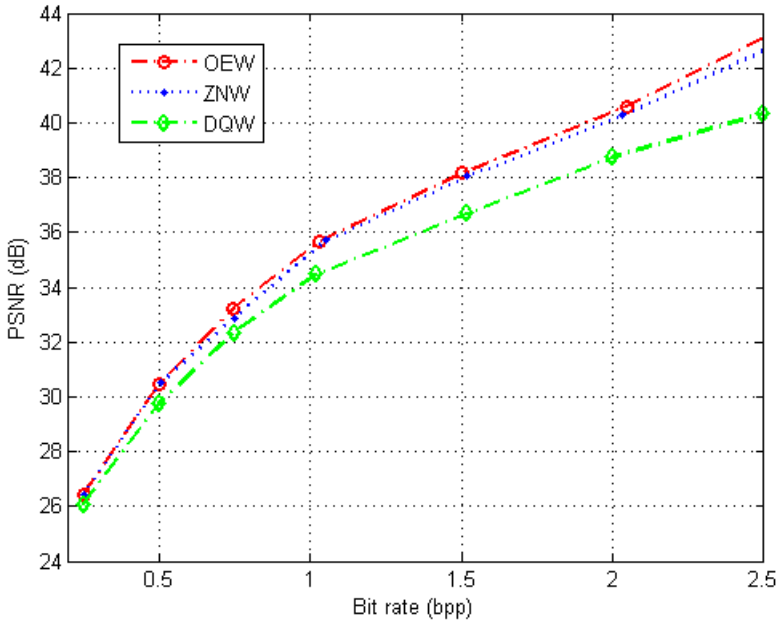


Figure 4.6: Comparison of compression performance for Barbara $512 \times 512$.

the odd-even JWC scheme, ZNW stands for the zero-nonzero JWC algorithm and DQW stands for the differential quantization JWC scheme. In our experiment, $\frac{1}{4}$ DCT blocks in the image were used for embedding watermarks in the OEW and ZNW algorithms with embedding rate is 1 bit per DCT block. However, for the benchmark DQW algorithm, half of all the DCT blocks were used for embedding with embedding rate was 1 bit per two blocks. In both of the OEW and ZNW algorithms, we embedded one bit watermarks at the 10th DCT coefficient position in the zigzag order while at the 13th DCT coefficient position in the zigzag order per block pair for the DQW algorithm. The quality factor for the standard JPEG recompression attacks, $Q_{jpeg}$ was set to 50 and totally 1024 watermark bits were embedded into each image. From the figures, we can see both of the proposed two can achieve better performance than the DQW algorithm. In particular, when the compression rate is equal to 1 bpp, the proposed JWC encoders can achieve 0.76 dB PSNR gain over the DQW JWC encoder for the Lena image and 1.22 dB PSNR gain for the Barbara image. Similar results were obtained for other test images and other standard JPEG recompression attacks.

Fig. 4.7 plots the comparison performance of the proposed two JWC schemes with the DQW scheme for the $512 \times 512$ Lena image at different watermark embedding rates. Half of all DCT blocks or $\frac{3}{4}$ of all DCT blocks in the image were selected for watermark embedding which result in different embedding rate at 2048 bits per image and 3072 bits per image respectively. For the DQW algorithm, we embedded 2 bits and 3 bits of information per two DCT blocks respectively and half of all blocks were used for watermark embedding. The quality factor for the standard JPEG recompression attack, $Q_{jpeg}$, was set to 50. It can be seen from the figure that as the watermark embedding rate increases, the compression rate distortion performance decreases as expected. The higher the embedding rate is, the more PSNR dB gain for proposed JWC schemes over the DQW algorithm can be obtained. In particular, the proposed OEW and ZNW algorithms achieve about 1.76 dB and 1.48 dB PSNR gain over the DQW algorithm respectively with embedding rate 3072 bits per image and compression rate 1 bpp. Another phenomena is that when embedding increases from 2048 bits per image to 3072 bits per image, the RD performance of the proposed two algorithms decreases quite small compared with the DQW algorithm which means that our proposed JWC algorithms are more suitable for high embedding rate case than the previous DQW scheme.

Fig. 4.8 illustrates the compression performance of the proposed OEW and ZNW schemes for $512 \times 512$ image Lena in the present of different standard JPEG compression attacks with quality factor equal to 25 and 50 when 1024 bits of infor-

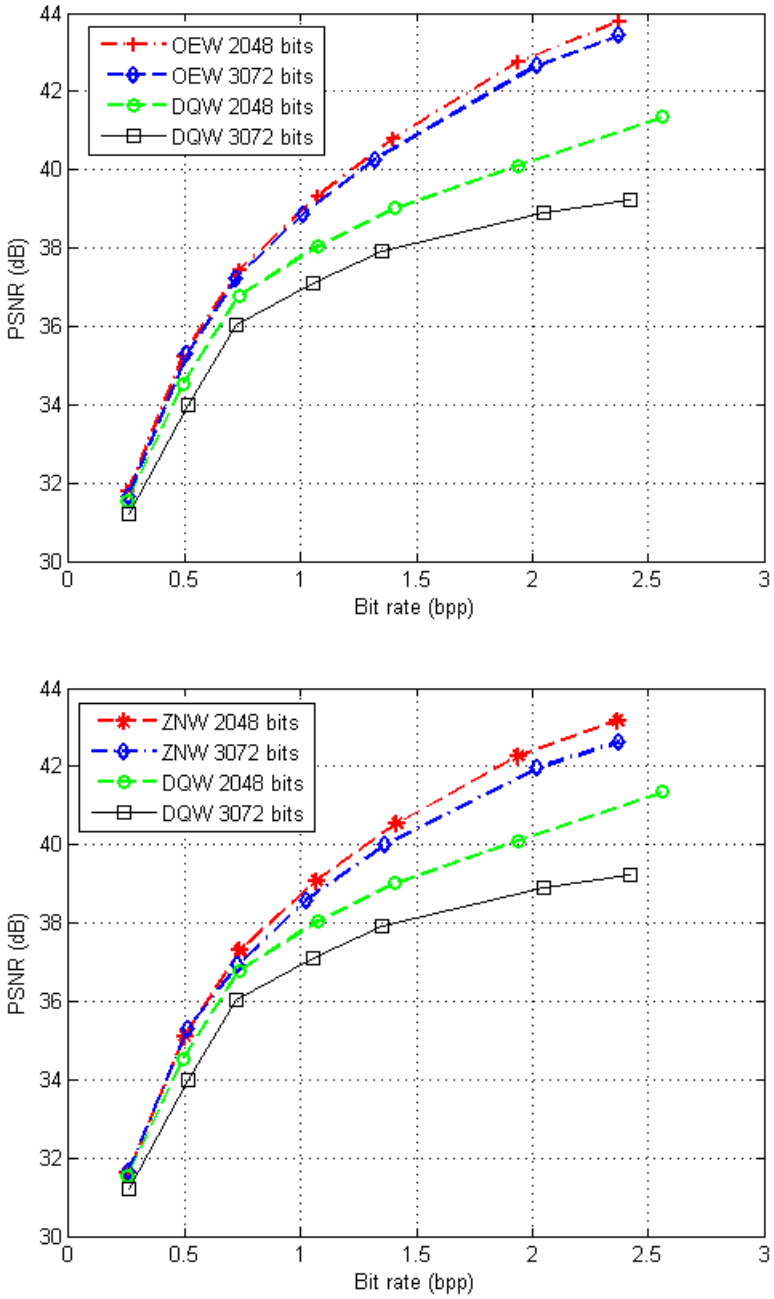


Figure 4.7: Comparison performance between OEW, ZNW and DQW algorithms at different embedding rates for $512 \times 512$ Lena.

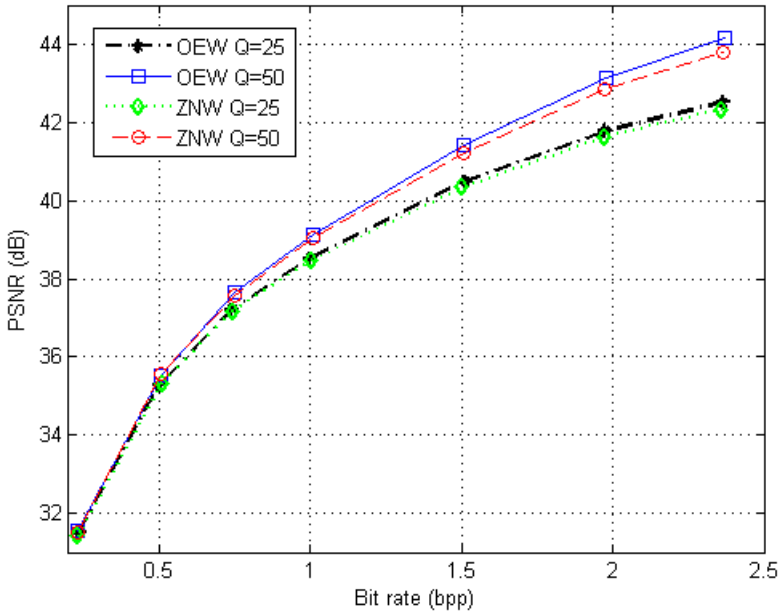


Figure 4.8: Comparison performance of the OEW scheme and the ZNW scheme under different attacks for $512 \times 512$ Lena.

mation were embedded into $\frac{1}{4}$ of all DCT blocks in the image. Form Fig. 4.8, we can see that the semi-blind JWC algorithm, OEW, get less gain over the full-blind JWC algorithm, ZNW, when quality factor $Q_{jpeg}$ decreases, i.e. the worse attack channels. The reason for this is that the ZNW JWC scheme which is modified based on the OEW scheme, sacrifices the RD performance to earn more robustness. When more robustness is needed, the advantage of the ZNW scheme comes out.

Fig. 4.9 compares the compression performance of the proposed OEW and ZNW algorithm with that of the DQW algorithm and the DEW algorithm for the $512 \times 512$ Lena image. In this figure, DQW stands for the differential quantization JWC proposed in [23], STD-JPEG stands for standard JPEG compression and DEW stands for the joint watermarking and compression scheme proposed in [21]. In [21], watermarks were embeded in the JPEG/MPEG streams by selectively removing high frequency DCT coefficients in certain image regions. However, the compression performance was not optimized. In all of the OEW, ZNW, DQW and DEW algorithms, 64 bits of watermark information were embedded and the quality factor of the standard JPEG recompression attack is set to 25. At such low embedding rate, the RD curves of the OEW and ZNW emerge together, we plot

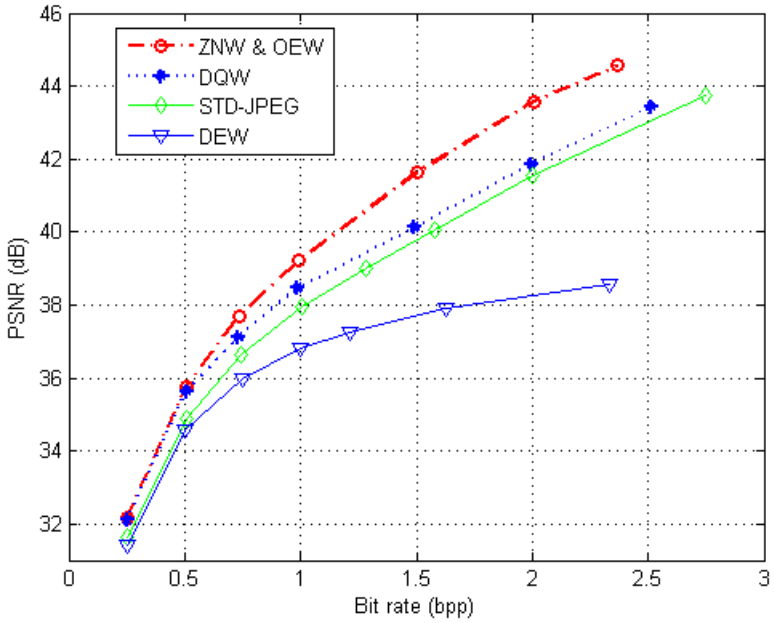


Figure 4.9: Comparison performance of between the proposed OEW and ZNW scheme, the DQW scheme and the DEW scheme.

one of them to stand both of them. The RD curve of the standard JPEG compression we plot here is to be the benchmark for measuring the RD performance of the proposed JWC schemes. It can be seen that the proposed OEW and ZNW scheme can achieve better RD performance than both of the DQW algorithm and the DEW algorithm. In particular, at compression rate of 1 bit per pixel, the OEW and ZNW algorithms achieve about 0.76 dB and 2.46 dB PSNR gain over the DQW algorithm and the DEW algorithm respectively. Similar results were obtained for other test images.

### Robustness to Gaussian Noise and Valumetric Scaling

In the following, we report robustness results for addition of Gaussian noise and valumetric scaling attacks. Since the watermark encoder do not have the knowledge of attack channels in advance, we still use the JWC encoder/decoder designed for the JPEG recompression attacks. We compare the results of the proposed OEW and ZNW algorithms with the informed coding and embedding algorithm (ICIE) proposed in [28].

The results of robustness with respect to Gaussian noise are summarized in

Fig. 4.10. Normally distributed noise with zero mean and standard deviation $\sigma_n$ was added to each of the watermarked images. The experiment was repeated for different standard deviations, $\sigma_n$, and the decoding bit error probability has been computed. The parameter of quality factor $Q_{jpeg}$ in process of JWC is set to 25 and it is assumed to be known at the watermark decoder. From the figure, we can see that ZNW algorithm achieves better decoding bit error probability performance than the OEW algorithm. Both of these two algorithms achieve better performance than the ICIE algorithm in [28]. In particular, when the decoding bit error probability is equal to 0.2, the standard deviation of the Gaussian noise is about 5.5, 10.5 and 32.5 for the ICIE, OEW and ZNW algorithms respectively.

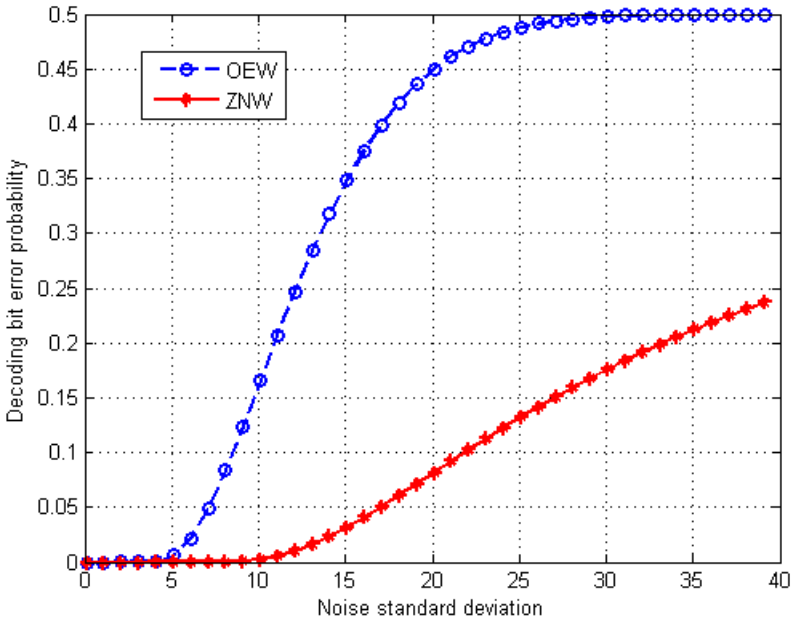


Figure 4.10: Robustness versus Gaussian noise.

Another simple, but important distortion is changing amplitude. That is $\mathbf{x}_n = \nu\mathbf{x}$, where $\mathbf{x}$ is the stegotext or watermarked image and $\nu$ is a scaling factor. This corresponds to a change of brightness and contrast for images and video. This attack is of particular interest for us, and is indeed the main weakness of the watermarking schemes of QIM proposed in [10]. Since the odd-even embedding scheme is a special case of QIM, this type of attacks effect much on the proposed OEW JWC algorithm. The results of valuemetric scaling attacks are reported in Fig. 4.11. From the figure, we can see that the ZNW algorithm achieves much better decoding error probability performance than both of the OEW algorithm and

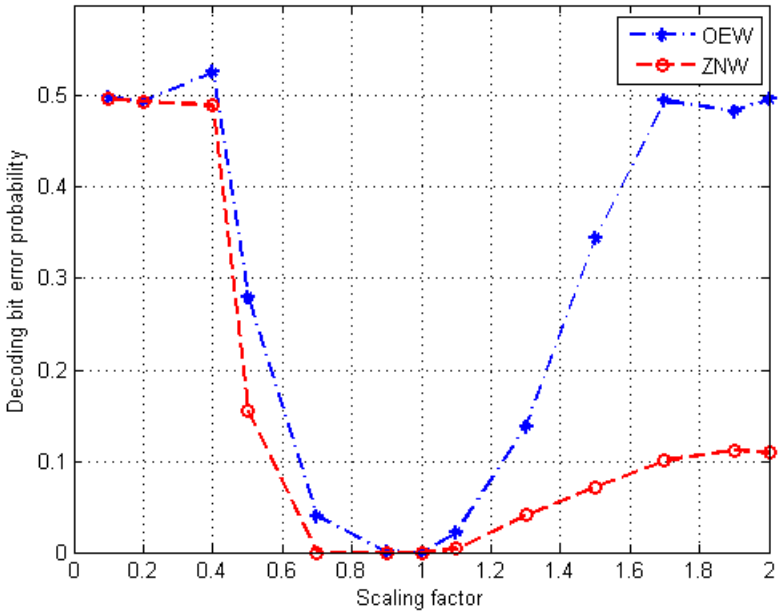


Figure 4.11: Robustness versus valumetric scaling up and down.

ICIE algorithm (see [28] for the performance). In particular, when the decoding error probability is less than 0.2, the scaling factor is in the range of 0.5-2 and 0.1-1.2 for the ZNW and ICIE algorithms respectively. That means our ZNW JWC scheme can achieve better performance than ICIE scheme for the scaling intensities up attacks. We need to mention that in the case of the scaling factor is less than 0.5 the perceptual quality of the images degrades heavily. This has been showed in Fig. 4.12 by $512 \times 512$ Lena image. In Fig. 4.12, the Lena image, with 4096 bits of watermarks embedded at compression rate 0.757 bpp using ZNW JWC scheme, was attacked by scaling with scaling factor 0.5 and AWGN attacks with standard deviation $\sigma_n = 20$ respectively. We can see that the perceptual quality of the Lena image degrades a lot after attacks.

Figure 4.12: Watermarked Lena image without attacks, attacked by scaling with scaling factor 0.5 and attacked by Gaussian noise with standard deviation $\sigma_n = 20$.

## 4.5 Chapter Summary

In this chapter, we have presented an integrated approach to JPEG compression and robust watermarking–a watermark is embedded in the process of compression which is faithful to the JPEG syntax. We proposed two algorithms, i.e. OEW and ZNW to embed watermarks into compressed bitstreams. The OEW algorithm is semi-blind, i.e. the watermark decoder should have the JPEG quality factor $Q_{jpeg}$ used in the process of joint compression and embedding while the ZNW scheme is full-blind where the watermark can be decoded only based on the received signal. We developed joint graph-based run-length coding and watermarking procedures to embed watermarks while optimizing the compression performance. It has been shown that the proposed JWC algorithms achieves better performance than the DQW and DEW algorithms proposed in the recent literature. The proposed ZNW algorithm designed for recompression attacks also has quite good robustness against other types of valumetric distortion attacks including additive Gaussian noise and valumetric scaling in everyday usage.

# Chapter 5

# Conclusions and Future Research

## 5.1 Conclusions

In digital watermarking, a watermark is embedded into a covertext resulting in a watermarked signal which is robust to certain distortion caused by either standard data processing in a friendly environment or malicious attacks in an unfriendly environment. As watermarked signals are highly desired to be compressed in real-world applications, in this thesis, we present the design, analysis and application of joint compression and watermarking systems. To broaden the application scope of digital watermarking, we have not only developed the JWC scheme using variable-rate scalar quantization but also proposed two new joint image compression and blind watermarking algorithms to jointly watermark and compress an image with baseline JPEG decoder compatible. The main results of this thesis are elaborated as follows.

- *JWC using variable-rate scalar quantization*: Using variable-rate scalar quantization for watermarking and compression, in Chapter 3, we have investigated how to design JWC systems to maximize the robustness in the presence of additive Gaussian noise attacks under constraints on both compression distortion and composite rate. We measured the decoding bit error probability by MD decoding rule, under consideration of low computation complexity. In comparison with the previous designed JWC systems using fixed-rate scalar quantization, the optimal JWC systems using variable-rate scalar quantization have been demonstrated to achieve better performance in the DNR region of practical interest.

- *Joint image compression and blind watermarking with baseline JPEG decoder compatible*: In Chapter 4, we have investigated the application of JWC utilizing variable-rate quantization where an efficient tradeoff among compression performance, embedding rate and robustness is desired to be obtained. We have developed two joint compression and watermarking schemes, i.e. OEW and ZNW, which are to maximize the embedding or compression performance while maintaining baseline JPEG decoders compatible. The watermark embedded by using these two schemes can survive a class of standard JPEG recompression attacks with high payload. To maximize the compression performance, two alternating algorithms have been developed to jointly optimize run-length coding, Huffman coding and quantization table selection subject to the additional watermark embedding constraints. The proposed two algorithm are demonstrated to have better embedding/compression performance than the DQW and DEW algorithm proposed in the recent literature. In particular, the ZNW algorithm designed for recompression attacks has also been shown to be robust to other types of valumetric distortion attacks including additive Gaussian noise and valumetric scaling.

## 5.2 Directions for Future Research

As a technique to protect copyright for digital content, digital watermarking has been recently one of the most active research fields in both the academic world and the industry. However, still a lot of theoretic and practical problems on optimal watermarking system design are unsolved. Among them, the following two questions related to our research are of particular interest:

1. As we know that watermark embedding can be viewed as a game between encoder/decoder and attacker. When joint compression and watermarking is considered, i.e. there is a constraint on the composite rate, what is the equilibrium or saddlepoints of this game?

2. How to extend the work of joint JPEG-compatible image compression and blind watermarking in this thesis to the design of JWC algorithms for audio and video signals? In particular, it will be more interesting if these JWC algorithms are backward compatible with industrial standards like MP3 and H.264 to achieve efficient tradeoffs among the embedding rate, compression rate, distortion and robustness.

# References


[1] I. J. Cox, M. L. Miller, and J. A. Bloom. *Digital Watermarking*. New York: Morgan Kaufmann, 2001.

[2] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. New York: John Wiley and Sons, 2006.

[3] P. Moulin and R. Koetter. Data-Hiding Codes. *Proceedings of IEEE*, vol. 93, No. 12, pp. 2083-2127, Dec. 2005.

[4] S. I. Gel'fand and M. S. Pinsker. Coding for channel with random parameters. *Probl. Control Inf. Theory*, vol. 9, no. 1, pp. 19-31, 1980.

[5] P. Moulin and J. A. O'Sullivan. Information-Theoretic Analysis of Information Hiding. *IEEE Transactions Information Theory*, vol. 49, pp. 563-593, March 2003.

[6] A. S. Cohen and Amos Lapidoth. The GaussianWatermarking Game. *IEEE Transactions Information Theory*, vol. 48, pp. 1639-1667, June 2002.

[7] D. Karakos and A. Papamarcou. A Relationship Between Quantization and Watermarking Rates in the Presence of Additive Gaussian Attacks. *IEEE Transactions Information Theory*, vol. 49, pp. 1970-1982, August 2003.

[8] A. Somekh-Baruch and N. Merhav. On the capacity game of public watermarking systems. *IEEE Trans. Inf. Theory*, vol. 50, no. 3, pp. 511-524, Mar. 2004.

[9] A. Maor and N. Merhav. On Joint Information Embedding and Lossy Compression in the Presence of a Memoryless Attack. *IEEE Transactions Information Theory*, vol. 51, no. 9, pp. 3166-3175, September 2005.

[10] B. Chen and G. W. Wornell. Quantization Index Modulation: A Class of Provably Good Methods for Digital Watermarking and Information Embedding. *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1423-1443, May 2001.

[11] E. Martinian, G. W. Wornell and B. Chen. Authentication with Distortion Criteria. *IEEE Trans. Inform. Theory*, vol. 51, no. 7, pp. 2523-2542, July 2005.

[12] Guixing Wu and En-Hui Yang. Joint Watermarking and Compression Using Scalar Quantization for Maximizing Robustness in the Presence of Additive Gaussian Attacks. *IEEE Trans. on Signal Processing*, Vol. 53, no. 2, pp. 834-844 Feb. 2005.

[13] Ludovic Guillemot and Jean-Marie Moureaux. Indexing Lattice Vectors in a Joint Watermarking and Compression Scheme. *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP*, 2007.

[14] Guixing Wu. *On the design and analysis of quantization based-digital watermarking system*. PhD thesis, Unversity of Waterloo, 2005.

[15] M. Minoux, *Mathematical Programming: Theory and Algorithms*. New York: Wiley, 1986.

[16] A. Gersho and R. Gray. *Vector Quantization and Signal Compression*. Boston, MA: Kluwer, 1992.

[17] P. A. Chou, T. Lookabaugh, R. M. Gray. Entropy-constrained vector quantization. *IEEE Trans. Signal Processing*, vol. 37, no. 1, pp. 31-42, Jan 1989

[18] ISO/IEC 10918-1 and ITU-T Recommendation T.81. *Information technology-digital compression and coding of continuous-tone still images: Requirements and guidelines*, 1994.

[19] G. K. Wallace. The JPEG still picture compression standard. *IEEE Trans. Consumer Electronics*, vol. 38, no. 1, pp. 18-34, Feb. 1992

[20] E.-H. Yang and L. Wang. Joint optimization of run-length coding, Huffman coding and quantization table with complete baseline JPEG decoder compatibility. *IEEE trans. on Image Processing*, Vol. 18, No.1, pp.63-74, Jan. 2009

[21] G.-C. Langelaar and R. L. Lagedijk. Optimal differential energy watermarking of DCT encoded images and video, *IEEE Trans. Image Processing*. Vol. 10, No.1 , pp. 148-158, Jan. 2001.

[22] Ching-Yung Lin and Shih-Fu Chang. A Robust Image Authentication Method Distinguishing JPEG Compression from Malicious Manipulation. *IEEE Trans. Circuits Syst. Video Technol.*, Vol. 11, No.2 , pp. 153-168, Feb. 2001.

[23] E.-H. Yang and G. Wu. Joint Compression and Blind Watermarking: A Case Study in the JPEG-Compatible Scenario. in *Proc. of 43rd annual Allerton Conference on Communication, Control and Computing*, 2005.

[24] E. h. Yang, Z. Zhang, and T. Berger. Fixed-slope universal lossy data compression. *IEEE Transactions on Information Theory*, vol. 43, pp. 1465-1476, Sep. 1997.

[25] E. h. Yang and Z. Zhang. Variable-rate trellis source encoding. *IEEE Transactions on Information Theory*, vol. 45, pp. 586-608, Mar. 1999.

[26] M. Wu and B. Liu. Data hiding in image and video: part I-fundamental issues and solutions. *IEEE Trans. Image Processing*, vol. 12, pp. 685-695, June 2003.

[27] Ingemar J. Cox, Joe Kilian, F. Thomson Leighton, and Talal Shamoon. Secure Spread Spectrum Watermarking for Multimedia. *IEEE Trans. Image Processing*, vol. 6, No.12 pp. 1673-1687, Dec. 1997.

[28] Matt L. Miller, Gwenael J. Doerr, and Ingemar J. Cox. Applying Informed Coding and Embedding to Design a Robust High-Capacity Watermark. *IEEE Trans. Image Processing*, vol.13, no.6, pp. 792-807, June 2004.

[29] F. Hartung and B. Girod. Watermarking of uncompressed and compressed video. *IEEE Trans. Singnal Processing*, vol. 66, no. 3, pp. 283-301, May 1998.

[30] L. Xie and G. R. Arce. A class of authentication digital watermarks for secure multimedia communication. *IEEE Trans. Image Processing*, vol. 10, no. 11, pp. 1754-1764, Nov. 2001.