STATISTICAL MODELS FOR DIGITAL WATERMARKING

NG TEK MING

NATIONAL UNIVERSITY OF SINGAPORE

2007

STATISTICAL MODELS FOR DIGITAL WATERMARKING

NG TEK MING

(*B. Sc.* (*Hons.*), *M. Tech., M. Eng., NUS*)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY OF ENGINEERING

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2007

To

Ying Xin and Ying Tong

Acknowledgment

It has been a long and tiring journey. But this is my most fulfilling journey, with many fruitful moments, and with many sweet and exciting memories. Thanks to the following people for making this whole experience a pleasant, enjoyable and memorable one.

My supervisor, Prof. Hari Krishna Garg, is the most important person who has helped me to achieve my career goal. He is the person who has brought me into the world of academic and research. His talent in mathematics has always impressed me and has also inspired me in many ways. It has been a great pleasure to work with him and to learn from him. I am exremely grateful and indebted to him for all his help.

My family members have been very understanding and have given me the moral support during this period. They have sacrifice a lot for my education. It is their continuing love that keeps me going. I dedicate this thesis to my dearest Ying Xin and Ying Tong for the joy and happiness they bring into my life. Their laughters are the soothing music that help me to de-stress all my worries. When the going gets tough, a simple hug from them means a lot to me and has always

cheer me up. They are the most important people in my life and I am very proud of them.

Prof. Yeo Swee Ping, Prof. Daniel Chan Siu Hung and Prof. Loh Ai Poh are the people who have made it possible for me to work as a teaching assistant in the Electrical and Computer Engineering (ECE) department. This has not only given me the financial support to see me through my graduate studies but has given me the opportunity to teach which is what I enjoy most.

All my students have really made my years of teaching in ECE an extremely wonderful experience. The excellent teaching feedback from them, with kind and touching words, have been the source of encouragement and motivation in my career. Their comments and suggestions have helped me to improve my teaching over the years. How I wish I could turn back time to experience all these all over again.

Last but not least, I would also like to thank Mr. Eric Siow and all my friends from ECE for their support and help during my studies and work.

Contents

Summary

This thesis is directed towards the study of the likelihood ratio (LR) based detection method in detecting invisible watermarks in images.

LR detection method based on Bayes' decision theory has been considered for image watermarking in transform domain. The Neyman-Pearson criterion is used to derive a decision threshold to minimize the probability of missed detection subject to a given probability of false alarm. In order to achieve the optimum behavior of the LR detector, a probability distribution function (PDF) that models the distribution of the transform coefficients is required. This detection method first appeared in the literature for image watermarking in discrete Fourier transform (DFT) domain. Thresholding via Neyman-Pearson criterion is done by modeling magnitude of a set of DFT coefficients using a Weibull PDF. The same detection method has also been examined for image watermarking in the discrete wavelet transform (DWT) domain, where a set of DWT coefficients is modeled using a Gaussian PDF.

The Weibull and Gaussian distributions are special cases of the generalized gamma and generalized Gaussian distributions, respectively. These two general distributions also encompass many other well known and commonly used

distributions. This leads us to propose using the generalized gamma PDF and generalized Gaussian PDF to model transform coefficients of DFT and DWT, respectively, for LR detection. We consider a zero mean generalized Gaussian PDF as the mean of the DWT coefficients in a given subband is approximately zero. In addition, we also explore using a Laplacian PDF for LR detection in DWT domain. Decision rule and closed-form decision threshold are derived for all proposed models. New estimators are introduced for parameters of the generalized Gaussian and generalized gamma distributions. Our numerical experiments reveal that the proposed models can produce better LR detection.

Maximum a posteriori (MAP) detection is another statistical watermark detection method. It is simpler than LR detection in the sense that a decision threshold is not required. MAP detection has been considered for watermarking in discrete cosine transform (DCT) domain using a Laplacian PDF. We propose an MAP detector using a generalized Gaussian PDF in DWT domain, and show that it can result in improved detection.

An embedding scheme that is based on the additive embedding scheme is also included in our work. The proposed embedding scheme requires more computation but it can give better watermark robustness.

Abbreviations

CLT	:	Central Limit Theorem
CDF	:	Cumulative Distribution Function
DCT	:	Discrete Cosine Transform
DFT	:	Discrete Fourier Transform
DWT	:	Discrete Wavelet Transform
HVS	:	Human Visual System
MAP	:	Maximum A Posteriori
LR	:	Likelihood Ratio
MSE	:	Mean Square Error
PSNR	:	Peak Signal to Noise Ratio
PDF	:	Probability Distribution Function
RV	:	Random Variable
SI	:	Statistical Independence

List of Figures

List of Tables

Chapter 1

Introduction

This chapter gives an overview of our work. Section 1.1 describes briefly our areas of focus in digital watermarking. The objective of our research together with the contributions made are summarized in Section 1.2. A brief organization of the thesis is given in Section 1.3.

1.1 Digital Watermarking

A digital watermark is a mark placed on multimedia content for a variety of applications including copyright protection, copy protection, authentication, fingerprinting, broadcast monitoring, etc [11, 19, 40]. In recent years, digital watermarking has become a hot area of research due to the rapid development of multimedia networks and thus the need to prevent unauthorized duplication and distribution of multimedia content [4, 9, 10, 13, 18, 27]. In the literature, many digital watermarking algorithms have been developed and improved. Some are already being used in the multimedia industry.

In a digital watermarking system, the detection stage is a crucial stage. Good detection schemes enable the recovery of a watermark with low probability of false detection. Two types of false detection are possible during the detection process. A false alarm occurs if a watermark is reported to be present when it is not there. On the other hand, a missed detection occurs when an existence of a watermark is rejected even though one is present. The complexity of the detector, the type of watermark embedding method used, and the characteristics of the watermark channel are among other things that influence the accuracy of the detection process.

Traditionally, watermark detection algorithms are based on computing correlation between the watermarked media and the watermark itself. Correlation detection is usually preferred because of its simplicity. Another advantage is that the detection can be 'blind', i.e., the original media is not required in the detection process. However, correlation detection is known to be optimal only when the embedding process follows an additive scheme, and the media is drawn from Gaussian distributions [11].

More recent works on watermark detection are based on decision theory [2, 3, 5, 8, 14, 25]. For this type of detection, an accurate model for the probability distribution function (PDF) of the original media is required. Our main focus is in the work of Barni et al [3] where a likelihood ratio (LR) detection method based on Bayes' decision theory is proposed. In [3], an imperceptible watermark is inserted using a non-additive scheme to the discrete Fourier transform (DFT) of the

original image. This involves modeling the magnitude of a set of DFT coefficients using a Weibull PDF. A decision threshold is derived using the Neyman-Pearson criterion to minimize the missed detection probability subject to a given false alarm probability. The same detection method is also explored by Kwon et al [25] by considering the discrete wavelet transform (DWT) domain for watermark embedding. In [25], DWT coefficients are modeled using a Gaussian PDF.

Experimental results given in [3, 25] show that, in the context of robustness, the LR detector has a better performance than the correlation detector. Moreover, blind detection is also possible in LR detection by estimating the parameters of the PDF from the watermarked image [7].

1.2 Our Work

Our objective is to explore and to generalize the LR detection framework of Barni et al [3]. The research work reported here emphasizes on developing a wider range of PDF models for LR detection in transform domain watermarking. Although our numerical experiments are done for DWT and DFT domains, these models are also applicable in other transform domains, for example, the discrete cosine transform (DCT) domain. Also included in our work is an embedding scheme which is based on the additive scheme and a maximum a posteriori (MAP) detector which is quite similar to the LR detector but simpler.

1.2.1 Publications

Our work so far has resulted in few publications as listed below. The same list is also included in the bibliography at the end of the thesis. To avoid confusion, the same numbering is used.

Journal Paper

[32] T.M. Ng and H.K. Garg, "Maximum likelihood detection in DWT image watermarking using Laplacian modeling," *IEEE Signal Processing Letters*, Vol. 12, No. 4, pp. 285-288, Apr 2005.

[33] T.M. Ng and H.K. Garg, "Wavelet domain watermarking using maximum-likelihood detection," *Journal of Imaging Science and Technology*, Vol. 49, No. 3, pp. 303-308, May/June 2005.

[34] T.M. Ng and H.K. Garg, "A maximum a posteriori identification criterion for wavelet domain watermarking," *International Journal of Wireless and Mobile Computing: Special Issue on Mobile Systems and Applications*, 2005.

Conference Paper

[35] T.M. Ng and H.K. Garg, "Wavelet domain watermarking using maximum-likelihood detection," *Proc. SPIE Conf. on Security, Steganography, and Watermarking of Multimedia Contents VI*, Vol. 5306, San Jose, Jan 19-22, 2004.

[36] T.M. Ng and H.K. Garg, "A maximum a posteriori identification criterion for wavelet domain watermarking," *Proc. 24th IEEE Intl. Conf. on Distributed Computing Systems Workshop*, Tokyo, March 23-24, 2004.

[37] T.M. Ng and H.K. Garg, "An embedding scheme for bipolar watermark," *Proc. Intl. Conf. Sciences of Electronic, Technologies of Information and Telecommunications*, Tunisia, March 15-20, 2004.

[38] T.M. Ng and H.K. Garg, "Maximum likelihood detection in image watermarking using generalized gamma model," *Proc. 39th Asilomar Conference on Signals, Systems and Computers*, Monterey, pp. 1680-1684 Oct 28-Nov 2, 2005.

Note that [33] and [34] are extended versions of [35] and [36], respectively.

1.2.2 Contributions

Based on our publications, we briefly summarize the original work reported in this thesis. Further details are given in the following chapters.

i. **Energy Embedding Scheme**

The additive scheme is one of the simplest schemes to embed a watermark to an image. This is done by scaling and then adding elements of the watermark directly to the image pixels or transform coefficients of the image. For transform domain watermarking, elements of the watermark can be embedded to transform coefficients with highest magnitude. This is one

way to improve the robustness of the watermark. In [37], we introduce a new embedding scheme which is based on modifying the transform coefficients with highest 'energy'. Although the proposed scheme is mathematically more complex, it is shown that it can result in better watermark robustness.

ii. **General LR Detection Framework**

The LR detection method of Barni et al [3] is based on using a Weibull PDF to model the magnitude of the DFT coefficients of an image. It involves an approximation which is derived using Taylor's Theorem [23]. Their derivation is formulated in terms of the Weibull PDF. In [33, 35], we generalize this derivation as well as the whole LR detection framework to hold for any PDF model.

iii. **LR Detection Based on Laplacian Model**

Which PDF model to use depends on the transform domain under consideration. One guideline is to choose a PDF with shape that resembles closely the shape of the histogram of the transform coefficients. The Gaussian PDF is used by Kwon et al [25] to model DWT coefficients for LR detection. In [32], we consider the Laplacian PDF instead. A closed-form decision threshold is derived. It is shown that the Laplacian model can yield a better watermark detection as compared to the Gaussian model.

iv. **LR Detection Based on Generalized Gaussian Model**

In [25], watermark is inserted to the high resolution DWT subbands of the

image. The mean of the DWT coefficients in high resolution subbands is close to zero. This leads us to consider a zero mean generalized Gaussian PDF to model the transform coefficients [33, 35]. The Gaussian PDF and Laplacian PDF are special cases of the generalized Gaussian PDF. Our numerical experiments show that the zero mean generalized Gaussian model can produce better LR detection results. A closed-form decision threshold under the zero mean generalized Gaussian model is also derived.

v. **LR Detection Based on Generalized Gamma Model**

The generalized gamma distribution is another distribution that includes many common distributions as special cases. For example, the gamma, Weibull, and exponential distributions can be obtained from the generalized gamma distribution based on appropriate setting of parameters. Our work in [33, 35] has led us to consider generalizing the Weibull model of Barni et al [3] to a generalized gamma model. Besides deriving a closed-form decision threshold, we also introduce new estimators for the parameters of the distribution. The generalized gamma model is also shown to result in improved watermark detection.

vi. **MAP detector Based on Generalized Gaussian Model**

In another work of Barni et al [2], an MAP detector is proposed for DCT domain image watermarking using a Laplacian model. We introduce a similar MAP detector in [36] for DWT domain watermarking using a generalized Gaussian model. The watermark to be embedded in an image

is chosen from a predefined set of watermarks. In identifying the embedded watermark, the a posteriori probability corresponding to each watermark in the set is computed. The maximum of these a posteriori probabilities is the one belonging to the embedded watermark. Thus, in applications whereby the number of watermarks in the set is not too large, this can be a feasible method to identify the embedded watermark. Moreover, it eliminates the need for a decision threshold and therefore should result in a more accurate detection.

1.3 Outline of Thesis

This thesis is organized as follows. In Chapter 2, we give the background material required to peruse this thesis. A chapter each is then devoted to describe the work done for each of the contributions mentioned in the previous section. This begins with describing the fundamental watermark embedding schemes and watermark detection methods in Chapter 3. The energy embedding scheme is covered in this chapter. The general LR detection framework is derived in Chapter 4. In Chapter 5, we include the LR detection proposed by Kwon et al [25] which is based on a Gaussian model. This gives the insight into the derivations for the subsequent chapters. Chapter 6, 7 and 8 are devoted to describe LR detection based on Laplacian, generalized Gaussian and gamma models, respectively. Chapter 9 gives the MAP detector based on the generalized Gaussian model. Lastly, we summarize and conclude our work in Chapter 10. This includes mentioning a few

interesting areas in which further research can be conducted.

The general notations used in this thesis are as follows. Non-boldface letters are used to represent scalar quantities, sets and functions. Boldface letters are used for vectors and matrices. All vectors and matrices are real-valued and expressed in column form. The superscript T represents the transpose of vectors and matrices.

All figures and tables are placed at the end of the chapters.

Chapter 2

Background

This chapter briefly reviews the background material required for the understanding and development of the chapters that follow. Notations and terminologies introduced here will be used throughout the thesis.

2.1 Probability Theory

2.1.1 Random Variables and Their Characterization

A random variable (RV) X is a function that maps every outcome of a random experiment to a real value. A continuous RV can take uncountably many possible values while a discrete RV has only a finite or countable number of values. We assume that X is a continuous RV throughout. Associated with X is the probability distribution function (PDF) of X, denoted by $f_X(x)$. The probability that X will take values from a set R of real numbers may be obtained by integrating $f_X(x)$ over R. For example, if R is the interval $[a,\ b]$ then the

probability of the event $\{a \leq X \leq b\}$ is given as

$$P(\{a \leq X \leq b\}) = \int_a^b f_X(x)dx. \tag{2.1}$$

The cumulative distribution function (CDF) is defined as

$$F_X(x) = \int_{-\infty}^{x} f_X(u)du. \tag{2.2}$$

In other words, $F_X(x)$ is the probability that X takes values in $(-\infty,\ x]$.

The PDF and CDF of X give complete characterization of the behaviour of X. We are also interested in parameters associated with the PDF and CDF that provide us with partial but meaningful information about X. Two of the most widely used parameters are the mean and variance of X, defined as

$$\mu = E[X] = \int_{-\infty}^{+\infty} x f_X(x)dx \tag{2.3}$$

and

$$\sigma^2 = V[X] = \int_{-\infty}^{+\infty} (x-\mu)^2 f_X(x)dx, \tag{2.4}$$

respectively. The mean and variance are special cases of parameters called moments and central moments, respectively. The kth moment is defined as

$$\mu = E[X^k] = \int_{-\infty}^{+\infty} x^k f_X(x)dx, \tag{2.5}$$

where $k = 1$ yields the mean of X. The kth central moment is defined as

$$E[(X-\mu)^k] = \int_{-\infty}^{+\infty} (x-\mu)^k f_X(x)dx, \tag{2.6}$$

where $k = 2$ yields the variance of X.

2.1.2 Multidimensional Random Variables

In many situations, we encounter multiple RVs. For example, the values of pixels in an image can be considered as a collection of RVs. Multiple random variables are basically multi-dimensional functions. Let us just consider the case of two RVs, X_1 and X_2. Generalization to the multidimensional case is straightforward.

We can view X_1 and X_2 as a single two dimensional RV $\mathbf{X} = (X_1,\ X_2)$. A complete characterization of (X_1, X_2) is given by the joint PDF of (X_1, X_2) denoted as $f_{X_1,X_2}(x_1, x_2)$ or more compactly as $f_{\mathbf{X}}(\mathbf{x})$, where $\mathbf{x} = (x_1, x_2)$ is the realization of $\mathbf{X}$. The probability of the event $\{a \leq X_1 \leq b, c \leq X_2 \leq d\}$ is given as

$$P(a \leq X_1 \leq b, c \leq X_2 \leq d) = \int_a^b \int_c^d f_{X_1,X_2}(x_1, x_2) dx_1 dx_2. \tag{2.7}$$

The joint CDF, denoted as $F_{X_1,X_2}(x_1, x_2)$, is the probability

$$P(-\infty < X_1 < x_1, -\infty < X_2 < x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_{X_1,X_2}(x_1, x_2) dx_1 dx_2. \tag{2.8}$$

Of special interest is when X_1 and X_2 are statistically independent (SI). Then,

$$f_{X_1,X_2}(x_2, x_2) = f_{X_1}(x_1) f_{X_2}(x_2) \tag{2.9}$$

and

$$F_{X_1,X_2}(x_2, x_2) = F_{X_1}(x_1) F_{X_2}(x_2). \tag{2.10}$$

The conditional PDF of X_1 given that $X_2 = x_2$ is defined as

$$f_{X_1|X_2}(x_1|x_2) = \frac{f_{X_1,X_2}(x_1,\ x_2)}{f_{X_2}(x_2)} \tag{2.11}$$

for all values of x_2 such that $f_{X_2}(x_2) > 0$. With conditional PDF, we can obtain conditional probabilities of events associated with X_1 when the value of X_2 is given. For example,

$$P(a \leq X_1 \leq b | X_2 = x_2) = \int_a^b f_{X_1|X_2}(x_1|x_2) dx_1. \tag{2.12}$$

2.1.3 Sum of Random Variables

It is common to encounter a RV X defined as the linear sum of M RVs X_1, $X_2, \ldots, X_M$. In other words,

$$X = a_1 X_1 + a_2 X_2 + \ldots + a_M X_M, \tag{2.13}$$

where $a_1, a_2, \ldots, a_M$ are arbitrary scalars. The expectation operator is a linear operator, i.e.,

$$E[X] = a_1 E[X_1] + a_2 E[X_2] + \ldots + a_M E[X_M]. \tag{2.14}$$

If X_1, $X_2, \ldots, X_M$ are SI, then

$$V[X] = a_1^2 V[X_1] + a_2^2 V[X_2] + \ldots + a_M^2 V[X_M]. \tag{2.15}$$

The random variables X_i and X_j, $i \neq j$, are said to be uncorrelated if

$$E[X_i X_j] = E[X_i] E[X_j]. \tag{2.16}$$

Note that (2.15) also holds if X_1, $X_2, \ldots, X_M$ are uncorrelated.

2.1.4 Parameter Estimation

Often in practice, we are interested in characterizing a RV X associated with a large group of objects called population. This involves drawing statistical inferences about certain parameters of X. Instead of examining the entire population, which is usually impossible, we may work with a random sample $X_1, X_2, \ldots, X_N$ from the population. The size N of the sample is much smaller than the size of the population. Based on the sample, we create functions of $X_1, X_2, \ldots, X_N$ to estimate parameters of X. For example, the mean μ and variance σ^2 of the population can be estimated as

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} X_i \tag{2.17}$$

and

$$\hat{\sigma}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \hat{\mu})^2, \tag{2.18}$$

respectively. These are also what we call unbiased estimators, meaning $E[\hat{\mu}] = \mu$ and $E[\hat{\sigma}^2] = \sigma^2$.

Parameter estimation plays an important role in science and engineering. Finding an estimator that can measure the actual parameter in the best possible way is a crucial area for research.

2.1.5 Gaussian Distribution and Central Limit Theorem

We say that a RV X is a Gaussian RV, or simply X is Gaussian distributed, with mean μ and variance σ^2, if the PDF of X is given as

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]. \tag{2.19}$$

The Gaussian PDF is a bell-shaped curve symmetrical about μ. If $\mu = 0$ and $\sigma^2 = 1$ then X is known as the standard Gaussian RV. Related to the PDF of X is the complementary error function defined as

$$\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{+\infty} \exp(-u^2) du. \tag{2.20}$$

For $x > \mu$, the complementary error function is proportional to the area under the tail of the Gaussian PDF [43]. It may be helpful to express the complementary error function as

$$\text{erfc}(x) = 2[1 - F_U(\sqrt{2}x)], \tag{2.21}$$

where U is the standard Gaussian RV. This is easily obtained from (2.20) by a change of variable in the integral.

The Gaussian PDF is usually regarded as the most important PDF in probability theory. One of the reasons is because of the central limit theorem (CLT) as stated in the following.

Theorem 2.1 *(Central Limit Theorem) Let $X_1, X_2, \ldots, X_n$ be n SI RVs. Each $X_i, i = 1, 2, \ldots, n$, has an arbitrary PDF $f_{X_i}(x_i)$, mean μ_i and finite variance σ_i.*

Set

$$S_n = X_1 + X_2 + \ldots + X_n.$$

Then S_n approaches a Gaussian RV with mean $\sum_{i=1}^{n} \mu_i$ and variance $\sum_{i=1}^{n} \sigma_i^2$ as $n \to +\infty$.

In other words, if n is sufficiently large then we can approximate S_n as a Gaussian RV. How large is large can be quite subjective. It is mentioned in [12] that as a rule of thumb the value of n should be at least 30 for the application of CLT.

2.1.6 Transformation of Random Variables

Given a RV X with PDF $f_X(x)$, we can define a functional mapping $y = g(x)$. This gives rise to a RV $Y = g(X)$. If g is a monotonic function, then the PDF of Y is given as

$$f_Y(y) = f_X(x) \left| \frac{dx}{dy} \right|, \tag{2.22}$$

where $|\cdot|$ is the modulus (or absolute value) function. Note that as the left-hand side of (2.22) is a function of y, all quantities in the right-hand side must be expressed in terms of y. For example, consider $Y = aX + b$, where a and b are constants. This mapping is linear and monotonic. The PDF of Y expressed in terms of PDF of X is

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right). \tag{2.23}$$

Suppose now given that X is RV with uniform PDF in the interval $(0,\ 1)$, and the PDF of Y is also known. The unknown that we need to determine

is a monotonically increasing mapping g that transforms X to Y. Since g monotonically increasing, we can omit the modulus sign in (2.22) and write it as

$$f_Y(y)dy = f_X(x)dx. \tag{2.24}$$

The uniform PDF in the interval $(0,\ 1)$ is given as

$$f_X(x) = \begin{cases} 1 & x \in (0,\ 1) \\ 0 & \text{elsewhere} \end{cases}. \tag{2.25}$$

Integrating both sides of (2.24) yields

$$\int_{-\infty}^{y'} f_Y(y)dy = \int_{-\infty}^{x'} f_X(x)dx, \tag{2.26}$$

where $x' \in (0,\ 1)$ and $y' = g(x')$. Clearly, the right-hand side is equal to x' and the left-hand side is the CDF of Y evaluated at y'. That is

$$F_Y(y') = x' \tag{2.27}$$

or equivalently

$$y' = F_Y^{-1}(x'). \tag{2.28}$$

Thus, the required mapping g is the inverse of the CDF of Y. We note that (2.28) is useful in simulating a set of RVs from a known distribution. For example, we can generate a set of Gaussian RVs from uniform RVs.

2.2 Gamma Function

The gamma function is defined as

$$\Gamma(u) = \int_0^{+\infty} t^{u-1} \exp(-t)dt, \quad u > 0. \tag{2.29}$$

In particular, we have

$$\Gamma(1) = \int_0^{+\infty} \exp(-t)dt = 1 \tag{2.30}$$

and

$$\begin{aligned} \Gamma(1/2) &= \int_0^{+\infty} t^{-1/2} \exp(-t)dt \\ &= 2\int_0^{+\infty} \exp(-z^2)dz, \quad t = z^2 \\ &= 2\frac{\sqrt{\pi}}{2} \\ &= \sqrt{\pi}. \end{aligned} \tag{2.31}$$

The gamma function satisfy the recursive formula

$$\Gamma(u+1) = u\Gamma(u), \quad u > 0. \tag{2.32}$$

The Euler formula [22] is given as

$$\Gamma(u) = \frac{1}{u}\prod_{j=1}^{\infty}\left(1+\frac{1}{j}\right)^u \left(1+\frac{u}{j}\right)^{-1}. \tag{2.33}$$

Another infinite products expression is the Weierstrass formula,

$$\Gamma(u) = \left[u\exp(\zeta u)\prod_{k=1}^{+\infty}\left(1+\frac{u}{k}\right)\exp(-u/k)\right]^{-1}, \tag{2.34}$$

where $\zeta = \lim_{m\to+\infty}\{(1+1/2+1/3+\ldots+1/m) - \ln m\}$ is a positive constant called the Euler-Mascheroni constant.

2.3 Standard Image Processing Operations

Watermarked images transmitted over any communication channel may undergo some image processing operations. It is a challenge for any watermarking

algorithm to ensure that the embedded watermark survive these operations, and can be detected at the receiving end. We describe briefly some of these operations, and consider them in our numerical experiments later.

JPEG compression

An image is JPEG compressed to reduce the amount of data needed to represent it. Consequently, this reduces the space required to store the image or reduces the speed of transmission of the image. However, the greater the compression the more information from the image is lost and thus affecting the image quality. A quality factor is used to indicate the desired image quality after JPEG compression. It ranges from 0 to 100, where 0 indicates best compression and 100 indicates best image quality.

Low Pass Filtering

A low-pass filter passes on lower frequency components of an image, while attenuating or rejecting the higher frequency components. It is commonly used to reduce noise from an image. The image is blurred and smoothed from the effects of low-pass filtering.

Mathematically, low-pass filtering is implemented by performing a two-dimensional convolution between the image matrix and a filter kernel. For

example, a 3×3 filter kernel is given as

$$\begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}.$$

Convolution calculates a new intensity value for a pixel in the image based on the pixel's neighbours. Each neighbouring pixel contributes a percentage of its own to the calculation of the new pixel.

Median Filtering

A median filter uses a sorting of pixel intensity values to determine the pixel's filtered value. The input pixel is replaced by the median of the pixels contained around the pixel. For example, consider the following pixel values in a 3×3 window:

$$\begin{bmatrix} 3 & 6 & 3 \\ 3 & 12 & 4 \\ 5 & 3 & 3 \end{bmatrix}.$$

The pixel values are then sorted in increasing order as $\{3, 3, 3, 3, 3, 4, 5, 6, 12\}$. The median value is 3. Median filtering is effective in removing pixel values that are greatly different from the rest of the neighbourhood.

Gaussian Noise

Noise originates from the image formation process, transmission medium, recording process, etc., is usually modelled as an additive zero mean white

Gaussian noise process. In generating Gaussian noise using MATLAB [29], we need to specify the mean and variance of the noise.

Speckle Noise

Speckle noise is a granular noise that inherently exists in radar and ultrasound images. It is multiplicative in nature. In MATLAB [29], speckle noise is added to an image $\mathbf{I}$ using the equation $\mathbf{I}' = \mathbf{I} + \eta\mathbf{I}$, where η is uniformly distributed random noise with zero mean and variance v, and $\mathbf{I}'$ is the corrupted image.

Salt and Pepper Noise

Salt and pepper noise is caused by errors in the image transmission. In some cases, the corrupted pixels are set alternatively to zero or to the maximum value. It appears as black and white impulses on the image, giving the image a 'salt and pepper' like appearance. The noise is usually quantified by the percentage of pixels which are corrupted.

Cropping

In our numerical experiments, watermarked images are cropped to retain a rectangular portion at the centre. The missing portion is replaced by pixels with value zero so that the size of each cropped image remains the same.

Chapter 3

Watermark Insertion and Detection

A watermark can be embedded to the spatial or transform (frequency) domain of an image. In spatial domain, the pixels of the image are modified to blend in the watermark. Although spatial domain watermarking is considered easier to implement, it may not have the robustness to survive some of the common image processing operations mentioned in Chapter 2. Transform domain usually offers more protection against these operations by exploiting the characteristics of the human visual system (HVS) [20]. In transform domain watermarking, the transform coefficients of the image are modified instead to capture the watermark. Our focus throughout is on transform domain watermarking, namely the discrete wavelet transform (DWT) and discrete Fourier transform (DFT) domains.

3.1 Embedding Scheme

Let $\mathbf{x} = [x_1, x_2, \ldots, x_N]^T$ be the vector representing N transform coefficients selected to embed a watermark $\mathbf{w} = [w_1, w_2, \ldots, w_N]^T$. The corresponding transform coefficients of the watermarked image are represented as $\mathbf{y} = [y_1, y_2, \ldots, y_N]^T$. Watermark is usually inserted into the image transform coefficients using either the additive scheme

$$y_i = x_i + \alpha_i w_i \tag{3.1}$$

or non-additive scheme (also called the multiplicative scheme) as

$$y_i = x_i(1 + \alpha_i w_i), \tag{3.2}$$

for $i = 1, 2, \ldots, N$, where α_i is a positive scalar representing the embedding strength. The larger the embedding strength, the more robust is the watermark. However, this also means more distortion is being introduced into the image, thus the visual quality of the image may be affected. Therefore, it is important to tune the embedding strength to balance between robustness and imperceptibility of the watermark. The amount added to x_i is usually set to depend on its magnitude $|x_i|$. For example, the additive scheme can be formulated as

$$y_i = x_i + \alpha|x_i|w_i, \quad i = 1, 2, \ldots, N, \tag{3.3}$$

where α is a fixed constant.

A distortion measure is used to quantify visual degradation of the image due to the embedded watermark. One of the most widely used distortion

measures is the Peak Signal to Noise Ratio (PSNR) [24], [31]. For an $\mathcal{M} \times \mathcal{N}$ original (undistorted) image, let the pixel from the mth row and nth column be represented as $p_{m,n}$. The corresponding pixel in the watermarked image is represented as $\tilde{p}_{m,n}$. The PSNR is then given as

$$\text{PSNR} = \frac{\max_{m,n} p_{m,n}^2}{\frac{1}{\mathcal{M}\mathcal{N}} \sum_{j=1}^{\mathcal{M}} \sum_{k=1}^{\mathcal{N}} (p_{m,n} - \tilde{p}_{m,n})^2}. \tag{3.4}$$

It is usually measured in decibels (dB), i.e., PSNR(dB)=$10\log_{10}$(PSNR).

3.2 Detection Method

The likelihood ratio (LR) and maximum a posteriori (MAP) detection methods are formulated based on the multiplicative embedding scheme. These are discussed in the following chapters. Here, we look at detection methods for additive embedding scheme.

By denoting $|x| = [|x_1|, |x_2|, \ldots, |x_N|]^T$, we can write (3.3) as

$$\mathbf{y} = \mathbf{x} + \alpha|\mathbf{x}| \otimes \mathbf{w}, \tag{3.5}$$

where $\otimes$ denotes element by element multiplication of vectors. The watermarked image is usually subjected to common image processing operations or intended attacks to remove the watermark. Therefore, the watermarked image coefficients vector $\mathbf{y}$ may be distorted to $\tilde{\mathbf{y}} = [\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_N]^T$. If the distortion can be modeled as an additive noise $\mathbf{n} = [n_1, n_2, \ldots, n_N]^T$, then

$$\tilde{\mathbf{y}} = \mathbf{x} + \alpha|\mathbf{x}| \otimes \mathbf{w} + \mathbf{n}. \tag{3.6}$$

Watermark can be detected with or without the use of the original image. When the original image is used, a possibly distorted watermark $\tilde{\mathbf{w}}$ is first extracted from $\tilde{\mathbf{y}}$ by reversing the embedding scheme (3.3). The similarity between $\tilde{\mathbf{w}}$ and the embedded watermark is measured by

$$\text{sim}(\mathbf{w}, \tilde{\mathbf{w}}) = \frac{\sum_{i=1}^{N} w_i \tilde{w}_i}{\sqrt{\sum_{i=1}^{N} w_i \tilde{w}_i}}. \tag{3.7}$$

The watermark is declared to be present if $\text{sim}(\mathbf{w}, \tilde{\mathbf{w}})$ has a value greater than a predefined threshold.

As mentioned in Chapter 1, watermark detection without the use of the original image is termed blind detection. This may be more desirable and may have wider applications. Blind detection can be done by computing the correlation between $\tilde{\mathbf{y}}$ and $\mathbf{w}$, defined as

$$\text{cor}(\tilde{\mathbf{y}}, \mathbf{w}) = \frac{1}{N} \tilde{\mathbf{y}}^T \mathbf{w} = \frac{1}{N} \sum_{i=1}^{N} \tilde{y}_i w_i. \tag{3.8}$$

Substituting $\tilde{\mathbf{y}}$ from (3.6) in (3.8), we obtain

$$\begin{aligned} \text{cor}(\tilde{\mathbf{y}}, \mathbf{w}) &= \frac{1}{N} \left(\mathbf{x}^T \mathbf{w} + \alpha (|\mathbf{x}| \otimes \mathbf{w})^T \mathbf{w} + \mathbf{n}^T \mathbf{w} \right) \\ &= \frac{1}{N} \left(\sum_{i=1}^{N} x_i w_i + \alpha \sum_{i=1}^{N} |x_i| w_i^2 + \sum_{i=1}^{N} n_i w_i \right). \end{aligned} \tag{3.9}$$

The summations $\sum_{i=1}^{N} x_i w_i$ and $\sum_{i=1}^{N} n_i w_i$ are usually much smaller as compared to $\sum_{i=1}^{N} |x_i| w_i^2$. This is due to the uncorrelatedness between $\mathbf{w}$ and both $\mathbf{x}$ and $\mathbf{n}$. Therefore, we can approximate $\text{cor}(\tilde{\mathbf{y}}, \mathbf{w})$ as

$$\text{cor}(\tilde{\mathbf{y}}, \mathbf{w}) \approx \frac{\alpha}{N} \sum_{i=1}^{N} |x_i| w_i^2. \tag{3.10}$$

If this correlation value is greater than a predefined threshold, then the watermark is said to be present. Otherwise, it is said to be absent. The threshold has to be properly chosen to ensure a high level of accuracy in the detection process [1, 42].

3.3 Energy Embedding Scheme

In the embedding scheme (3.3), each magnitude $|x_i|$ is scaled by the same constant α. We can further introduce variable scaling here. Replacing α by a variable η_i yields

$$y_i = x_i + \eta_i |x_i| w_i, \quad i = 1, 2, \ldots, N. \tag{3.11}$$

Similar to (3.1), we can choose η_i according to the magnitude of x_i. Without loss of generality, we assume that $|x_1| \leq |x_2| \leq \ldots \leq |x_N|$, and consider setting η_i as

$$\eta_i = \beta \frac{|x_i|}{x_N}, \quad i = 1, 2, \ldots, N, \tag{3.12}$$

where β is another constant. Therefore, each scaling is again apportioned according to the magnitude of the coefficient. Thus, we can write (3.11) as

$$y_i = x_i + \frac{\beta}{|x_N|} |x_i|^2 w_i, \quad i = 1, 2, \ldots, N. \tag{3.13}$$

In (3.3), the constant α is set according to the magnitude of the transform coefficients. If we refer to $|x_i|^2$ as the energy of x_i, then in (3.13) the scalar β is set according to the energy of the transform coefficients. For convenience of discussion, we refer to (3.3) as the magnitude scheme and (3.13) as the energy

scheme. With the energy scheme, we assume that $\mathbf{y}$ is distorted to $\tilde{\mathbf{y}}$ under the same noise model as in (3.6). Then

$$\mathbf{y} = \mathbf{x} + \frac{\beta}{N|x_N|}\mathbf{x} \otimes \mathbf{x} \otimes \mathbf{w} + \mathbf{n}, \tag{3.14}$$

and by the same argument that leads to (3.10), we obtain.

$$\text{cor}(\tilde{\mathbf{y}}, \mathbf{w}) \approx \frac{\beta}{N|x_N|} \sum_{i=1}^{N} |x_i|^2 w_i^2. \tag{3.15}$$

In [37], we compare the magnitude and energy schemes for the embedding of bipolar watermarks. For a bipolar watermark, each of its coefficients is either 1 or -1, i.e., $w_i = 1$ or -1 for $i = 1, 2, \ldots, N$. Since $w_i^2 = 1$, we see that (3.10) and (3.15) reduce to

$$\text{cor}(\tilde{\mathbf{y}}, \mathbf{w}) \approx \frac{\alpha}{N} \sum_{i=1}^{N} |x_i| \tag{3.16}$$

and

$$\text{cor}(\tilde{\mathbf{y}}, \mathbf{w}) \approx \frac{\beta}{N|x_N|} \sum_{i=1}^{N} |x_i|^2, \tag{3.17}$$

respectively. It follows that to ensure a high correlation value for effective detection, we can choose $x_1, x_2, \ldots, x_N$ so that either the sum of their magnitudes is maximized under the magnitude scheme, or sum of their energies is maximized under the energy scheme. Theoretically, in the absence of noise, embedding using the energy scheme yields the same correlation detection capability as embedding using the energy scheme if

$$\frac{\beta}{N|x_N|} \sum_{i=1}^{N} |x_i|^2 = \frac{\alpha}{N} \sum_{i=1}^{N} |x_i|, \tag{3.18}$$

or equivalently

$$\beta = \frac{\alpha |x_N| \sum_{i=1}^{N} |x_i|}{\sum_{i=1}^{N} |x_i|^2}. \tag{3.19}$$

Substituting this β in (3.13), we obtain

$$y_i = x_i + \frac{\alpha |x_N| \sum_{i=1}^{N} |x_i|}{\sum_{i=1}^{N} |x_i|^2} |x_i|^2 w_i, \quad i = 1, 2, \ldots, N. \tag{3.20}$$

Our experiments in the next section compare the magnitude scheme with the energy scheme in (3.20).

3.4 Experimental Results

We consider watermark embedding in the discrete wavelet transform (DWT) domain. One of the many advantages of using DWT is that it is known to be better in handling aspects of the HVS as compared to the others transform domains [20], [26], [51]. This is desirable for making a watermark more robust to distortions.

The 512×512 grayscale images considered in our experiments are shown in Figure 3.1. A Daubechies filter [50] is used for DWT. Each of these original images is first transformed by DWT to obtain a multi-resolution decomposition. This separates the image into lower resolution subband (LL_1) and high resolution horizontal (HL_1), vertical (LH_1) and diagonal (HH_1) subbands. The process can be repeated to obtain a multiple level pyramid decomposition. A three-level decomposition is shown in Figure 3.2. The DWT coefficients $x_1, x_2, \ldots, x_N$ where

we embed the watermark are selected from the high resolution subbands where the human eye is less sensitive to noise [20]. For simplicity, we insert watermark by modifying only the coefficients of the HH_2 subband.

The values of α, as tabulated in Table 3.1, are selected so that the PSNR of each watermarked image is about 40 dB when the energy scheme is used. A PSNR of 40 dB and above is usually accepted as good image quality. When the magnitude scheme is used with the same values of α, the PSNR of the images are slightly higher. This is plotted in Figure 3.3. Therefore, as expected, the energy scheme degrades the images more than the amplitude scheme.

The robustness of the embedded watermark when using the energy scheme is examined as follows. A set of 1,000 bipolar watermarks is generated. We embed a watermark from this set to each of the images. Standard image processing operations are simulated and applied to the watermarked images to distort them. The correlation detector is then used to identify the embedded watermark for each of the distorted images. For example, Figure 3.4(a) shows the watermarked image 'Lena' using the energy scheme with $\alpha = 0.190$. The image is still of good quality with no visible distortion. Figure 3.4(b) shows the same image distorted by salt and pepper of density 0.4, i.e., approximately 40% of the pixels are affected. Even though the image is now visibly degraded, the correlation detector can still identify the embedded watermark as seen in the plot of correlation values in Figure 3.5. The highest peak at position 388 belongs to the embedded watermark.

In comparing the energy scheme with the magnitude scheme, the above

experiment is performed over 100 trials of different sets of 1,000 bipolar watermarks. The number of successful detections from all the trials are recorded for each of the images. Figure 3.6 shows the result for watermarked images distorted by salt and pepper noise of density 0.4. In Figure 3.7, the watermarked images are distorted by speckle noise with mean 0 and variance 0.6. In both cases correlation detector performs more effectively when the energy scheme is used. Our other experiments also reveal that the energy scheme yields better robustness when watermarked images are JPEG compressed. However, both schemes are comparable when watermarked images are exposed to Gaussian noise and low-pass filtering.

Figure 3.1: Test Images.

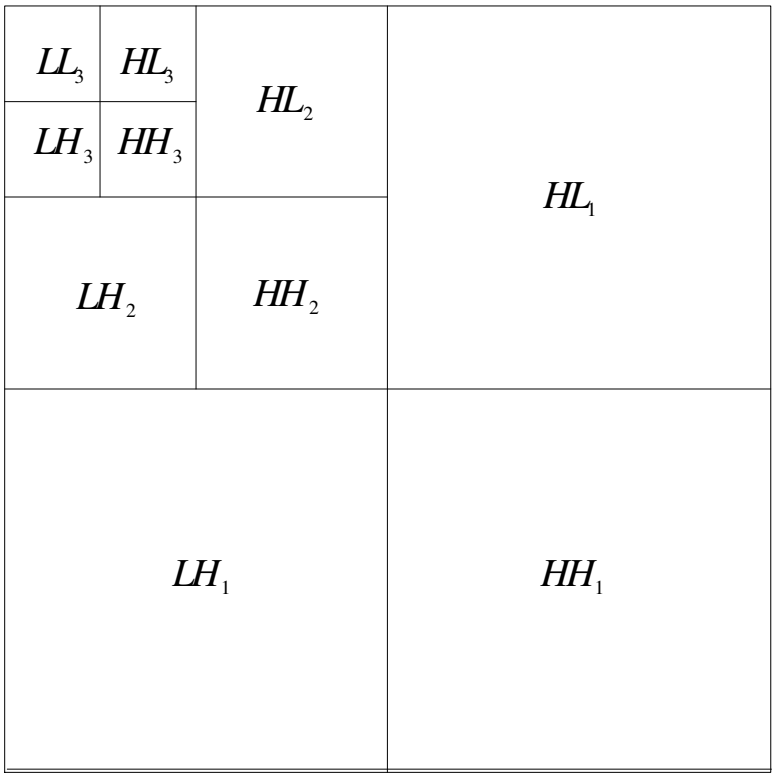

Figure 3.2: A DWT three-level pyramid decomposition of an image.

Table 3.1: Watermark embedding strength for images.

Image	α
Harbour	0.300
Lena	0.190
Fishing boat	0.325
Peppers	0.355
Barbara	0.255
Goldhill	0.405
Zelda	0.480
LAX	0.205

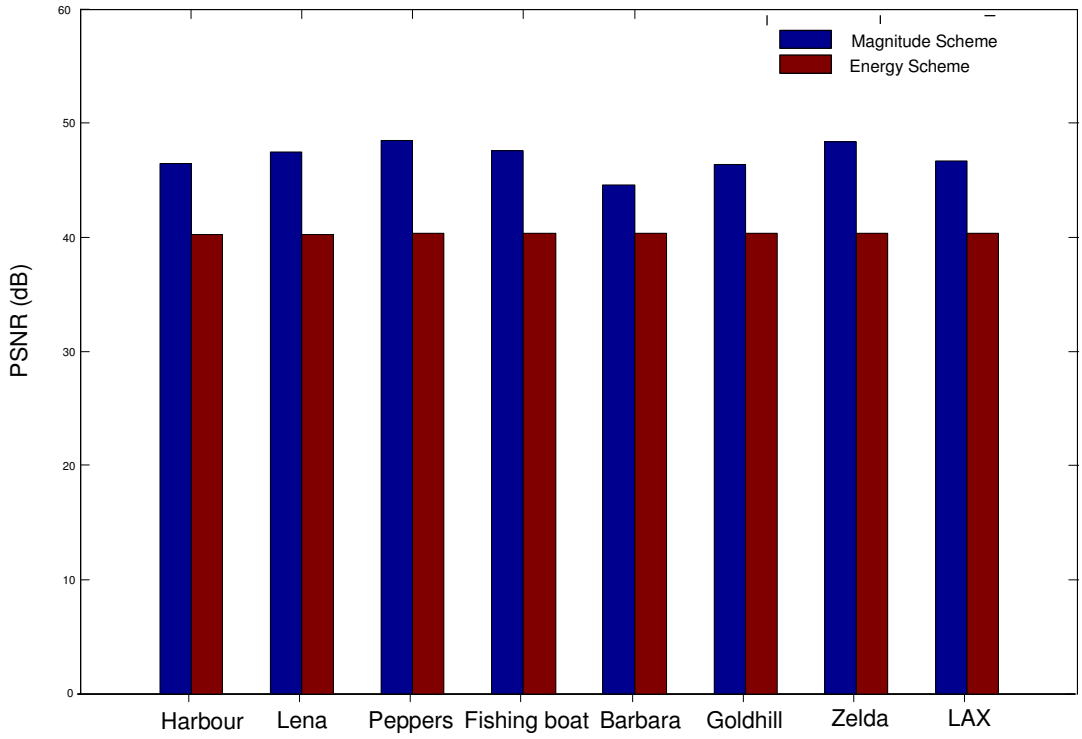

Figure 3.3: PSNR of watermarked images using magnitude and energy schemes.

(a)

(b)

Figure 3.4: (a) Image 'Lena' watermarked using energy scheme, (b) distorted version of (a) by salt and pepper noise.

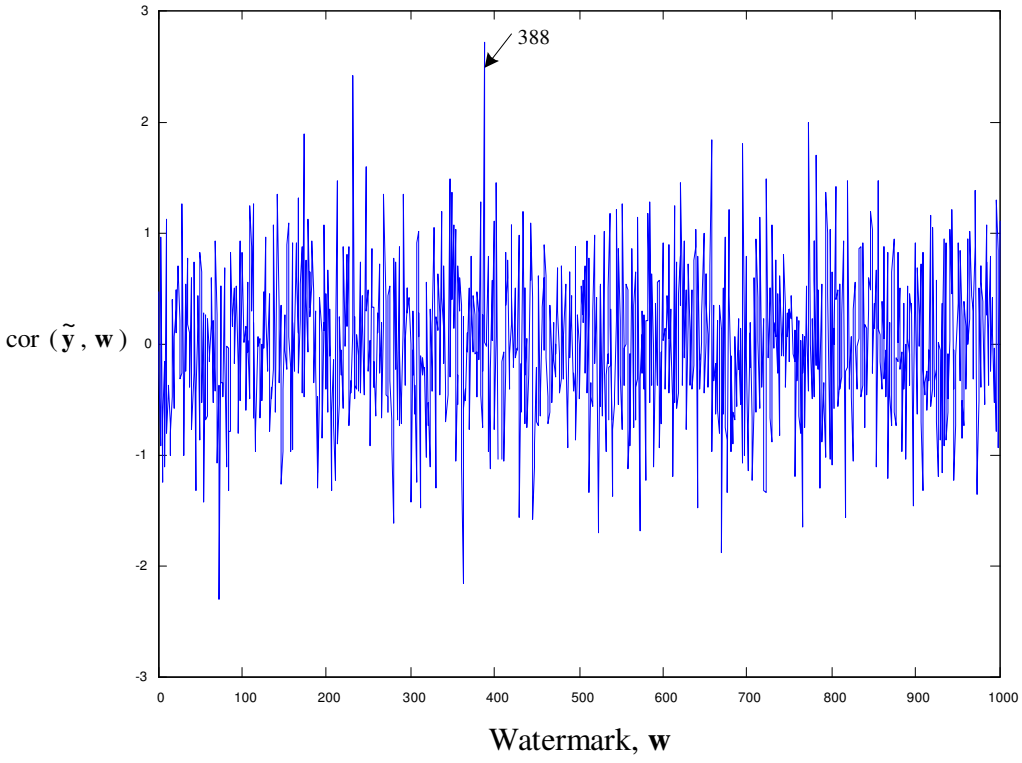

Figure 3.5: Correlation between $\tilde{y}$ and 1000 watermarks when energy scheme is used.

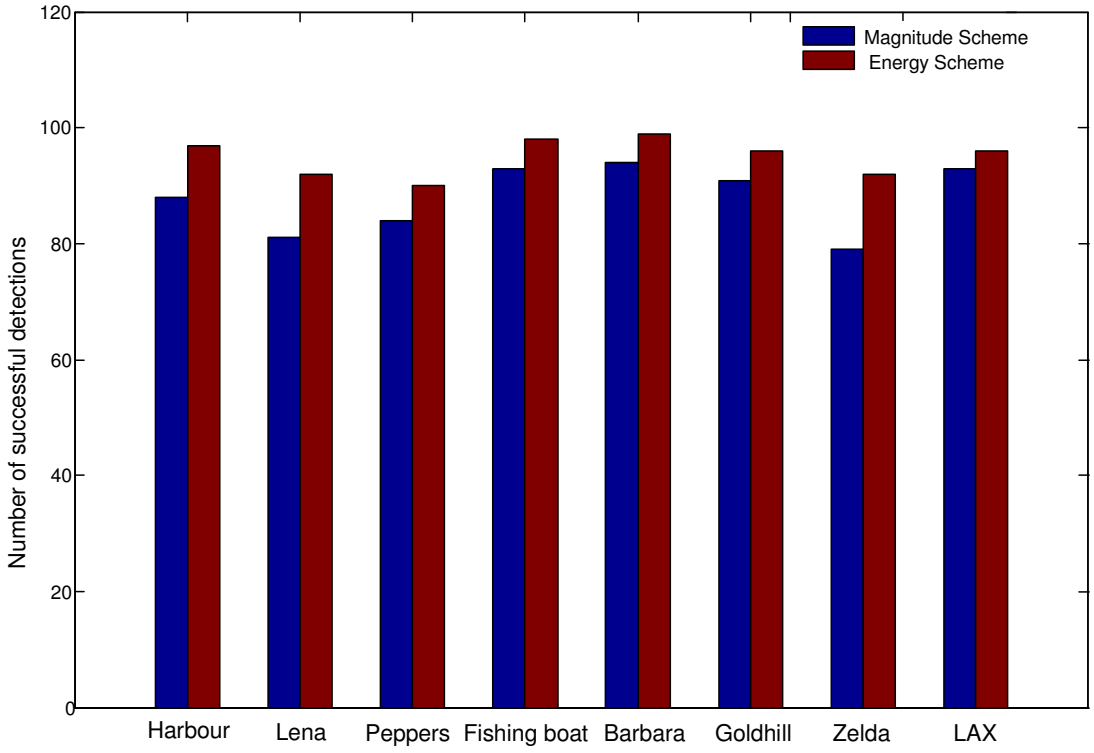

Figure 3.6: Robustness of watermark against salt and pepper noise.

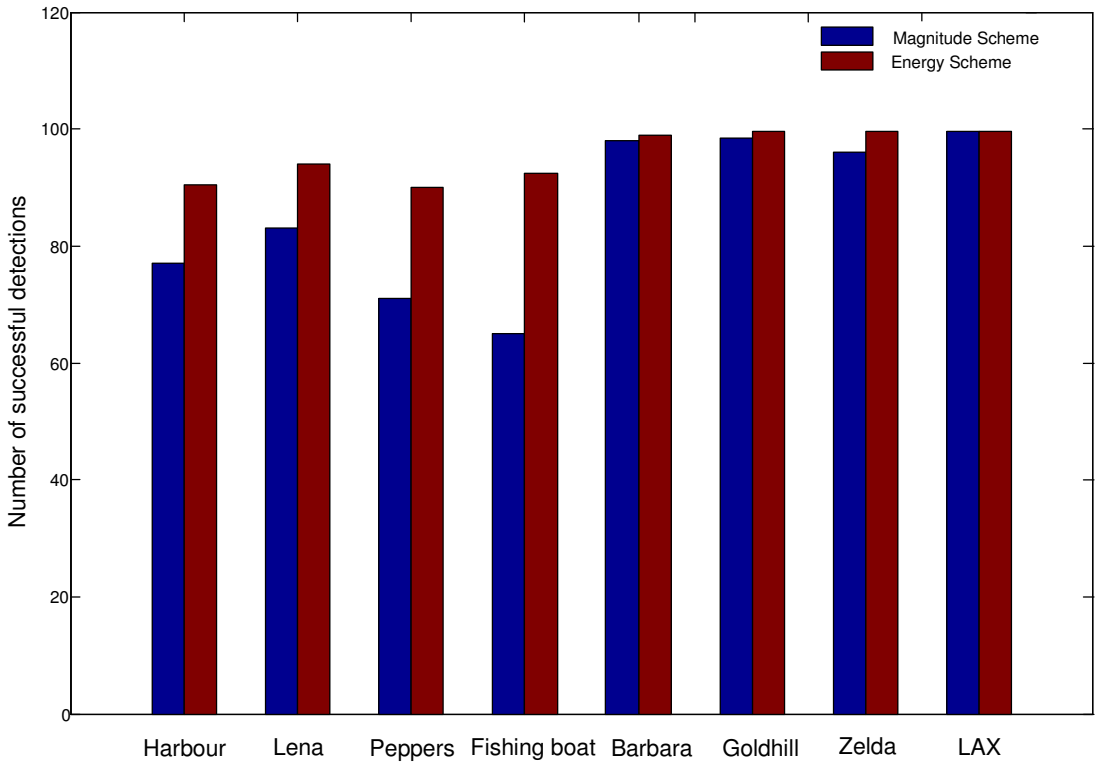

Figure 3.7: Robustness of watermark against speckle noise.

Chapter 4

LR Detection of Watermark

In this chapter we formulate the LR detection framework of Barni et al [3] in a general setting. Results presented here are valid for any probability distribution. Experimental procedure for Chapter 6, 7 and 8 is also given here.

4.1 LR Detection Framework

As in Chapter 3, we let $\mathbf{x} = [x_1,\ x_2, \ldots,\ x_N]^T$ be the vector representing N transform coefficients of an image selected to embed a watermark $\mathbf{w} = [w_1,\ w_2,\ \ldots,\ w_N]^T$, where $\mathbf{w}$ is chosen from a set M. The corresponding transform coefficients of the watermarked image is represented as $\mathbf{y} = [y_1,\ y_2,\ \ldots,\ y_N]^T$. We view x_i, w_i and y_i as realizations of the random variables X_i, W_i and Y_i, respectively, for $i = 1, 2, \ldots, N$. The PDFs of X_i, W_i and Y_i are denoted as $f_{X_i}(x_i)$, $f_{W_i}(w_i)$ and $f_{Y_i}(y_i)$, respectively. Embedding is done using the non-additive scheme,

$$y_i = x_i(1 + \alpha_i w_i), \tag{4.1}$$

for $i = 1, 2, \ldots, N$, where α_i is the embedding strength.

The components of the watermarks from the set M are assumed to be independent and uniformly distributed in $[-1, 1]$ so that

$$f_{\mathbf{W}}(\mathbf{w}) = \prod_{i=1}^{N} f_{W_i}(w_i) = \prod_{i=1}^{N} \frac{1}{2} = \frac{1}{2^N}. \tag{4.2}$$

The set M is thus the N-dimensional space of $[-1,\ 1]$, written as $[-1,\ 1]^N$. Specifically, if $\mathbf{w}^* = [w_1^*,\ w_2^*,\ \ldots,\ w_N^*]^T$ is the embedded watermark, we can write $M = M_0 \cup M_1$, where $M_0 = \{\mathbf{w} | \mathbf{w} \neq \mathbf{w}^*\}$ and $M_1 = \{\mathbf{w}^*\}$. Note that $\mathbf{w} = \mathbf{0} = [0,\ 0,\ \ldots,\ 0]^T$ that corresponds to a non-marked image is already included in M_0.

In LR detection, two hypotheses are established as follows:

$$H_0 \ : \ \mathbf{y} \text{ is not marked with } \mathbf{w}^*;$$

$$H_1 \ : \ \mathbf{y} \text{ is marked with } \mathbf{w}^*.$$

The hypothesis H_1 is accepted or equivalently the watermark $\mathbf{w}^*$ is detected if

$$l(\mathbf{y}) = \frac{f_{\mathbf{Y}}(\mathbf{y}|M_1)}{f_{\mathbf{Y}}(\mathbf{y}|M_0)} > \lambda, \tag{4.3}$$

where $f_{\mathbf{Y}}(\mathbf{y}|M_j), j = 0, 1$, are the conditional PDFs and λ is the LR decision threshold. The ratio $l(\mathbf{y})$ is called the likelihood ratio. The conditional PDF $f_{\mathbf{Y}}(\mathbf{y}|M_0)$ can be obtained by integrating $f_{\mathbf{Y}}(\mathbf{y}|\mathbf{w}) f_{\mathbf{W}}(\mathbf{w})$ over $[-1,\ 1]^N$, i.e.,

$$f_{\mathbf{Y}}(\mathbf{y}|M_0) = \int_{[-1,\ 1]^N} f_{\mathbf{Y}}(\mathbf{y}|\mathbf{w}) f_{\mathbf{W}}(\mathbf{w}) d\mathbf{w}. \tag{4.4}$$

Note that the Nth order integral in (4.4) should be taken over the set M_0. However, M_0 and $[-1\ 1]^N$ differ by a single point $\mathbf{w}^*$ which is of zero measure [45]. Thus, integrating over $[-1,\ 1]^N$ is the same as integrating over M_0.

The transform coefficients $X_1, X_2, \ldots, X_N$ are assumed to be independent [3, 47]. Under this assumption, the elements of $\mathbf{y}$ are conditionally independent and together with (4.2), we can express (4.4) as

$$f_{\mathbf{Y}}(\mathbf{y}|M_0) = \frac{1}{2^N} \prod_{i=1}^{N} \int_{-1}^{1} f_{Y_i}(y_i|w_i) dw_i. \tag{4.5}$$

The embedding strength α_i is set to be much smaller than 1 to make the watermark imperceptible. For small α_i, the integrals in (4.5) can be approximated using Taylor's Theorem.

Theorem 4.1 *(Taylor's Theorem) Let g be a real-valued function on $[a,\ a+\delta]$, $\delta > 0$, such that the $(n+1)$th derivative of g, $g^{(n+1)}(t)$, exists for every $t \in [a,\ a+\delta]$ and $g^{(n+1)}$ is continuous on $[a,\ a+\delta]$. Then, if $t \in [a,\ a+\delta]$, there exists a number ξ with $a \leq \xi \leq t$ such that*

$$\begin{aligned} g(t) &= g(a) + \frac{g^{(1)}(a)}{1!}(t-a) + \frac{g^{(2)}(a)}{2!}(t-a)^2 \\ &\quad + \cdots + \frac{g^{(n)}(a)}{n!}(t-a)^n + \frac{g^{(n+1)}(\xi)}{(n+1)!}(t-a)^{n+1}. \end{aligned} \tag{4.6}$$

The same result holds if $[a, a+\delta]$ is replaced by $[a-\delta, a]$.

The following lemma gives the approximation for (4.5).

Lemma 4.1

$$f_{\mathbf{Y}}(\mathbf{y}|M_0) \approx f_{\mathbf{Y}}(\mathbf{y}|\mathbf{0}). \tag{4.7}$$

Proof: We need to show that $f_{\mathbf{Y}}(\mathbf{y}|M_0) \approx f_{\mathbf{Y}}(\mathbf{y}|\mathbf{0})$ when the embedding strength α_i, $i = 1, 2, \ldots, N$, is much smaller than 1. To begin with, each integral in (4.5) is expressed as

$$\begin{aligned}\int_{-1}^{1} f_{Y_i}(y_i|w_i)dw_i &= \int_{-1}^{1} \frac{1}{1+\alpha_i w_i} f_{X_i}\left(\frac{y_i}{1+\alpha_i w_i}\right) dw_i \\ &= \int_{y_i/(1+\alpha_i)}^{y_i/(1-\alpha_i)} \frac{y_i}{\alpha_i t^2} f_{X_i}(t)dt. \end{aligned} \tag{4.8}$$

The first equality is due to (2.22), and the second equality is obtained by substituting $t = y_i/(1+\alpha_i w_i)$. If α_i is much smaller than 1, then the integration is done over a very small interval. Using (4.6) with $g(t) = y_i f_{X_i}(t)/\alpha_i t^2$, we obtained

$$\frac{y_i}{\alpha_i t^2} f_{X_i}(t) \approx \frac{1}{\alpha_i y_i} f_{X_i}(y_i) + C_1(t - y_i), \tag{4.9}$$

where

$$C_1 = \frac{d}{dt}\left(\frac{y_i}{\alpha_i t^2} f_{X_i}(t)\right)\Big|_{t=y_i}. \tag{4.10}$$

Substituting (4.9) in (4.8) yields

$$\int_{-1}^{1} f_{Y_i}(y_i|w_i)dw_i \approx \int_{y_i/(1+\alpha_i)}^{y_i/(1-\alpha_i)} \frac{1}{\alpha_i y_i} f_{X_i}(y_i)dt + C_1 \int_{y_i/(1+\alpha_i)}^{y_i/(1-\alpha_i)} (t - y_i)dt. \tag{4.11}$$

Because α_i is much smaller than 1, the second integral on the right-hand side is approximately zero. The first integral on the right-hand side evaluates to

$$\int_{y_i/(1+\alpha_i)}^{y_i/(1-\alpha_i)} \frac{1}{\alpha_i y_i} f_{X_i}(y_i)dt = \frac{2f_{X_i}(y_i)}{1-\alpha_i^2} \approx 2f_{X_i}(y_i). \tag{4.12}$$

The approximation in (4.12) is again due to the small value of α_i as compared to 1. Thus,

$$\int_{-1}^{1} f_{Y_i}(y_i|w_i)dw_i \approx 2f_{X_i}(y_i) \tag{4.13}$$

and using this in (4.5), we obtain

$$f_{\mathbf{Y}}(\mathbf{y}|M_0) \approx \frac{1}{2^N}\prod_{i=1}^{N} 2f_{X_i}(y_i) = \prod_{i=1}^{N} f_{Y_i}(y_i|0) = f_{\mathbf{Y}}(\mathbf{y}|\mathbf{0}). \tag{4.14}$$

Q.E.D

Approximation (4.7) is valid for any PDF. It is first derived by Barni et al [3] for Weibull PDF. It is also used by Kwon et al [25] for Gaussian PDF.

With (4.7) and because of the conditional independence of $Y_1, Y_2, \ldots, Y_N$, the likelihood ratio $l(\mathbf{y})$ can be expressed in terms of the PDFs of $X_1, X_2, \ldots, X_N$ as

$$l(\mathbf{y}) = \frac{\prod_{i=1}^{N} f_{Y_i}(y_i|w_i^*)}{\prod_{i=1}^{N} f_{Y_i}(y_i|0)} = \frac{\prod_{i=1}^{N} \frac{1}{1+\alpha_i w_i^*} f_{X_i}\left(\frac{y_i}{1+\alpha_i w_i^*}\right)}{\prod_{i=1}^{N} f_{X_i}(y_i)}. \tag{4.15}$$

By taking the natural logarithm of the likelihood ratio, the LR decision rule given by (4.3) becomes

$$z(\mathbf{y}) = \sum_{i=1}^{N}\left[\ln f_{X_i}\left(\frac{y_i}{1+\alpha_i w_i^*}\right) - \ln f_{X_i}(y_i)\right] > \lambda', \tag{4.16}$$

where $\lambda' = \ln\lambda + \sum_{i=1}^{N}\ln(1+\alpha_i w_i^*)$ is the modified LR decision threshold.

4.2 Detection Under the Neyman-Pearson Criterion

When a watermarked image is distorted, the missed detection probability P_{MD} can be much larger than the false alarm probability P_{FA} [3]. To overcome this problem, the Neyman-Pearson criterion can be used to obtain the decision threshold λ' in such a way that the missed detection probability is minimized

subject to a specified false alarm probability, say P_{FA}^*. In view of (4.7), once P_{FA}^* has been fixed, λ' can be derived from

$$P_{FA}^* = P(z(\mathbf{y}) > \lambda' | M_0) = P(z(\mathbf{x}) > \lambda') = \int_{\lambda'}^{\infty} f_{z(\mathbf{X})}(z(\mathbf{x}))dz(\mathbf{x}), \tag{4.17}$$

where

$$z(\mathbf{x}) = z(\mathbf{y})\big|_{\mathbf{y}=\mathbf{x}} = \sum_{i=1}^{N} \left[\ln f_{X_i}\left(\frac{x_i}{1+\alpha_i w_i^*}\right) - \ln f_{X_i}(x_i) \right]. \tag{4.18}$$

By the central limit theorem (see Theorem 2.1), the PDF of $z(\mathbf{X})$ can be assumed to be Gaussian with mean

$$\mu_{z(\mathbf{X})} = E[z(\mathbf{X})] = \sum_{i=1}^{N} E\left(\ln f_{X_i}\left(\frac{x_i}{1+\alpha_i w_i^*}\right) - \ln f_{X_i}(x_i) \right) \tag{4.19}$$

and variance

$$\sigma_{z(\mathbf{X})}^2 = V[z(\mathbf{X})] = \sum_{i=1}^{N} V\left(\ln f_{X_i}\left(\frac{x_i}{1+\alpha_i w_i^*}\right) - \ln f_{X_i}(x_i) \right). \tag{4.20}$$

In this regard, (4.17) can be written as

$$\begin{aligned} P_{FA}^* &= \int_{\lambda'}^{\infty} \frac{1}{\sqrt{2\pi\sigma_{z(\mathbf{X})}^2}} \exp\left(-\frac{(z(\mathbf{x}) - \mu_{z(\mathbf{X})})^2}{2\sigma_{z(\mathbf{X})}^2} \right) dz(\mathbf{x}) \\ &= \frac{1}{2}\text{erfc}\left(\frac{\lambda' - \mu_{z(\mathbf{X})}}{\sqrt{2\sigma_{z(\mathbf{X})}^2}} \right), \end{aligned} \tag{4.21}$$

where $\text{erfc}(\cdot)$ is the complementary error function (see (2.20)). Hence, the decision threshold λ' is obtained as

$$\lambda' = \text{erfc}^{-1}(2P_{FA}^*)\sqrt{2\sigma_{z(\mathbf{X})}^2} + \mu_{z(\mathbf{X})}. \tag{4.22}$$

In view of (2.21), the quantity $\text{erfc}^{-1}(2P_{FA}^*)$ may be computed as

$$\text{erfc}^{-1}(2P_{FA}^*) = \frac{1}{\sqrt{2}} F_U^{-1}(1 - P_{FA}^*). \tag{4.23}$$

Some values of P_{FA}^* with corresponding $\text{erfc}^{-1}(2P_{FA}^*)$ are given in Table 4.1.

4.3 Experimental Procedure

In Chapters 6, 7 and 8, we use a standard procedure to test the performance of the LR detector under different PDF models. The steps are summarized as follows:

1. Set a value for P_{FA}^*
2. Select an image and obtain the transform coefficients vector, $\mathbf{x} = [x_1, x_2, \ldots, x_N]^T$.
3. Generate a set M' containing K watermarks. The components of each watermark are uniformly distributed in $[-1,\ 1]$.
4. Select a watermark $\mathbf{w}^* = [w_1^*, w_2^*, \ldots, w_N^*]^T$ from M'.
5. Choose α_i, $i = 1, 2, \ldots, N$, such that when $\mathbf{w}^*$ is embedded to $\mathbf{x}$ the PSNR of the watermarked image is about 45 dB.
6. Embed all the K watermarks from M' to the images to produce K watermarked images. Note that this also includes the embedding of $\mathbf{w}^*$.
7. Distort all the K watermarked images using a standard image processing operation, e.g., JPEG compression.
8. Estimate all the necessary parameters for $f_{X_i}(x_i), i = 1, 2, \ldots, N$.
9. Compute the decision threshold λ' using (4.22).

10. Compute $z(\tilde{\mathbf{y}})$ using (4.16), where $\tilde{\mathbf{y}}$ is the distorted version of $\mathbf{y}$, for all the K distorted watermarked images in Step 7, and then compare them with λ'.

11. If $z(\tilde{\mathbf{y}}) > \lambda'$ for $\mathbf{w}^*$ but not for any other watermarks in M', then the detection is said to be successful. Otherwise it is a failure.

For each image and for each standard image processing operation, Step 3 to Step 11 are repeated for 10,000 trials. The percentage of successful detections are recorded. For our experiments in Chapters 6, 7 and 8, we set $P^*_{FA} = 10^{-6}$ and $K = 100$.

Note that in Step 8 the parameters are estimated from the transfrom coefficients of the original image. The parameters can also be estimated from the transform coefficients of the watermarked image (possibly distorted). As the watermark is embedded in an imperceptible manner, estimation using the transform coefficients of the watermarked image should be close to that of the original image [7]. If estimation is done using the watermarked image, then the detection is 'blind'.

Table 4.1: P^*_{FA} versus $\text{erfc}^{-1}(2P^*_{FA})$.

P^*_{FA}	$\text{erfc}^{-1}(2P^*_{FA})$
10^{-2}	1.645
10^{-3}	2.185
10^{-4}	2.630
10^{-5}	3.016
10^{-6}	3.361
10^{-9}	4.241

Chapter 5

LR Detector Based on Gaussian Model

Besides being important in its role in CLT, Gaussian distribution is also popular because of its many special properties. In a variety of applications, data churned out from random experiments often yield histograms that resemble the bell shape. In this chapter, we look at LR detection in watermarking when the transform coefficients are modeled as Gaussian random variables. A detailed derivation of the LR decision rule and threshold is given.

5.1 LR Decision Rule

Let X_i be the transform coefficient of an image modeled as a Gaussian RV with mean μ_i and variance σ_i. As given by (2.19), the PDF of X_i is expressed as

$$f_{X_i}(x_i) = \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left(-\frac{(x_i-\mu_i)^2}{2\sigma_i^2}\right). \tag{5.1}$$

Figure 5.1 shows the plot of the standard Gaussian PDF.

Substituting (5.1) in (4.16), the LR decision rule becomes

$$\begin{aligned} z(\mathbf{y}) &= \sum_{i=1}^{N} \left\{ \ln \left[\frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{(y_i(1+\alpha_i w_i^*)^{-1} - \mu_i)^2}{2\sigma_i^2} \right) \right] \right. \\ &\qquad \left. - \ln \left[\frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{(y_i - \mu_i)^2}{2\sigma_i^2} \right) \right] \right\} \\ &= \sum_{i=1}^{N} \frac{1}{2\sigma_i^2} \left[(y_i - \mu_i)^2 - \left(\frac{y_i}{1+\alpha_i w_i^*} - \mu_i \right)^2 \right] > \lambda_g, \end{aligned} \tag{5.2}$$

where λ_g denotes the LR decision threshold under the Gaussian model. Derivation for λ_g requires more work as given in the next section.

5.2 LR Decision Threshold

Following (4.22), we express λ_g as

$$\lambda_g = \mathrm{erfc}^{-1}(2P_{FA}^*)\sqrt{2\sigma_{z(\mathbf{X})}^2} + \mu_{z(\mathbf{X})}, \tag{5.3}$$

where

$$z(\mathbf{X}) = \sum_{i=1}^{N} \frac{1}{2\sigma_i^2} \left[(X_i - \mu_i)^2 - \left(\frac{X_i}{1+\alpha_i w_i^*} - \mu_i \right)^2 \right]. \tag{5.4}$$

Thus, to derive the mean and variance of $z(\mathbf{X})$, we need to obtain the mean and variance of

$$g(X_i) = \frac{1}{2\sigma_i^2} \left\{ (X_i - \mu_i)^2 - \left(\frac{X_i}{1+\alpha_i w_i^*} - \mu_i \right)^2 \right\}. \tag{5.5}$$

5.2.1 Derivation for Mean of $z(\mathbf{X})$

Taking expectation on both sides of (5.5), we get

$$E[g(X_i)] = \frac{1}{2\sigma_i^2} \left\{ E[(X_i - \mu_i)^2] - \frac{1}{(1+\alpha_i w_i^*)^2} E[(X_i - \mu_i - \mu_i \alpha_i w_i^*)^2] \right\}$$

$$= \frac{1}{2\sigma_i^2}\left\{E[(X_i-\mu_i)^2] - \frac{1}{(1+\alpha_i w_i^*)^2}E[(X_i-\mu_i)^2] \right. \\ \left. -\frac{(\mu_i\alpha_i w_i^*)^2}{(1+\alpha_i w_i^*)^2} + \frac{2\mu_i\alpha_i w_i^*}{(1+\alpha_i w_i^*)^2}E[(X_i-\mu_i)]\right\}. \quad (5.6)$$

The central moments of X_i are given as [39]

$$E[(X_i-\mu_i)^n] = \begin{cases} 1\cdot 3\cdot 5\cdots(n-1)\sigma_i^n & n=2k \\ 0 & n=2k+1 \end{cases}. \quad (5.7)$$

Thus, $E[(X_i-\mu_i)^2] = \sigma_i^2$ and $E[(X_i-\mu_i)] = 0$. Substituting these in (5.6), we obtain

$$E[g(X_i)] = \frac{1}{2}\left[1 - \frac{1}{(1+\alpha_i w_i^*)^2}\right] - \frac{1}{2(1+\alpha_i w_i^*)^2}\left(\frac{\mu_i\alpha_i w_i^*}{\sigma_i}\right)^2. \quad (5.8)$$

Thus, the mean of $z(\mathbf{X})$ is given as

$$E[z(\mathbf{X})] = \sum_{i=1}^{N} E[g(X_i)] \\ = \sum_{i=1}^{N}\frac{1}{2}\left\{\left[1 - \frac{1}{(1+\alpha_i w_i^*)^2}\right] - \frac{1}{(1+\alpha_i w_i^*)^2}\left(\frac{\mu_i\alpha_i w_i^*}{\sigma_i}\right)^2\right\}. \quad (5.9)$$

5.2.2 Derivation for Variance of $z(\mathbf{X})$

In order to derive the variance of $z(\mathbf{X})$, we need to first derive the expectation of

$$g^2(X_i) = \frac{1}{4\sigma_i^4}\left\{(X_i-\mu_i)^4 - \frac{2}{(1+\alpha_i w_i^*)^2}[(X_i-\mu_i)^4 + (X_i-\mu_i)^2(\mu_i\alpha_i w_i^*)^2 \right. \\ -2(X_i-\mu_i)^3\mu_i\alpha_i w_i^*] + \frac{1}{(1+\alpha_i w_i)^4}[(X_i-\mu_i)^4 - 4(X_i-\mu_i)^3\mu_i\alpha_i w_i \\ \left. +6(X_i-\mu_i)^2(\mu_i\alpha_i w_i)^2 - 4(X_i-\mu_i)(\mu_i\alpha_i w_i)^3 + (\mu_i\alpha_i w_i)^4]\right\}. \quad (5.10)$$

By use of (5.7) again, we obtain

$$\begin{aligned} E[g^2(X_i)] &= \frac{1}{4\sigma_i^4}\left\{3\sigma_i^4 - \frac{2}{(1+\alpha_i w_i^*)^2}[3\sigma_i^4 + \sigma_i^2(\mu_i\alpha_i w_i^*)^2]\right. \\ &\quad \left. + \frac{1}{(1+\alpha_i w_i)^4}[3\sigma_i^4 + 6\sigma_i^2(\mu_i\alpha_i w_i)^2 + (\mu_i\alpha_i w_i)^4] \right\} \\ &= \frac{1}{4}\left\{3 - \frac{2}{(1+\alpha_i w_i^*)^2}\left[3 + \left(\frac{\mu_i\alpha_i w_i^*}{\sigma_i}\right)^2\right]\right. \\ &\quad \left. + \frac{1}{(1+\alpha_i w_i)^4}\left[3 + 6\left(\frac{\mu_i\alpha_i w_i}{\sigma_i}\right)^2 + \left(\frac{\mu_i\alpha_i w_i}{\sigma_i}\right)^4\right] \right\}. \end{aligned} \tag{5.11}$$

Combining (5.8) and (5.11), we obtain

$$\begin{aligned} V[g(X_i)] &= E[g^2(X_i)] - E^2[g(X_i)] \\ &= \frac{1}{4}\left\{3 - \frac{2}{(1+\alpha_i w_i^*)^2}\left[3 + \left(\frac{\mu_i\alpha_i w_i^*}{\sigma_i}\right)^2\right]\right. \\ &\quad \left. + \frac{1}{(1+\alpha_i w_i)^4}\left[3 + 6\left(\frac{\mu_i\alpha_i w_i}{\sigma_i}\right)^2 + \left(\frac{\mu_i\alpha_i w_i}{\sigma_i}\right)^4\right] \right\} \\ &\quad - \frac{1}{4}\left\{1 + \frac{1}{(1+\alpha_i w_i^*)^4} - \frac{2}{(1+\alpha_i w_i^*)^2} + \frac{1}{(1+\alpha_i w_i^*)^4}\left(\frac{\mu_i\alpha_i w_i}{\sigma_i}\right)^4\right. \\ &\quad \left. - \frac{2}{(1+\alpha_i w_i^*)^2}\left(\frac{\mu_i\alpha_i w_i^*}{\sigma_i}\right)^2 + \frac{2}{(1+\alpha_i w_i^*)^4}\left(\frac{\mu_i\alpha_i w_i^*}{\sigma_i}\right)^2 \right\}. \\ &= \frac{1}{2}\left[1 - \frac{1}{(1+\alpha_i w_i^*)^2}\right]^2 + \frac{1}{(1+\alpha_i w_i^*)^4}\left(\frac{\mu_i\alpha_i w_i^*}{\sigma_i}\right)^2. \end{aligned} \tag{5.12}$$

Thus, we obtain the variance of $z(\mathbf{X})$ as

$$\begin{aligned} V[z(\mathbf{X})] &= \sum_{i=1}^{N} V[g(X_i)] \\ &= \sum_{i=1}^{N}\left\{\frac{1}{2}\left[1 - \frac{1}{(1+\alpha_i w_i^*)^2}\right]^2 + \frac{1}{(1+\alpha_i w_i^*)^4}\left(\frac{\mu_i\alpha_i w_i^*}{\sigma_i}\right)^2\right\}. \end{aligned} \tag{5.13}$$

5.2.3 Closed-Form Expression for λ_g

Substituting (5.9) and (5.13) in (5.3) yields

$$\begin{aligned}\lambda_g &= \text{erfc}^{-1}(2P_{FA}^*)\left[2\sum_{i=1}^{N}\left\{\frac{1}{2}\left[1-\frac{1}{(1+\alpha_i w_i^*)^2}\right]^2\right.\right. \\ &\qquad \left.\left.+\frac{1}{(1+\alpha_i w_i^*)^4}\left(\frac{\mu_i\alpha_i w_i^*}{\sigma_i}\right)^2\right\}\right]^{1/2} \\ &\quad +\sum_{i=1}^{N}\frac{1}{2}\left\{\left[1-\frac{1}{(1+\alpha_i w_i^*)^2}\right]-\frac{1}{(1+\alpha_i w_i^*)^2}\left(\frac{\mu_i\alpha_i w_i^*}{\sigma_i}\right)^2\right\}.\end{aligned} \tag{5.14}$$

From the expression, for a fixed P_{FA}^*, we see that λ_g depends on μ_i, σ_i^2, α_i and w_i^* for $i = 1, 2, \ldots, N$.

5.3 Zero Mean Model

If the mean of the $X_i, i = 1, 2, \ldots, N$, is approximately zero, then (5.2) and (5.14) reduces to

$$z(\mathbf{y}) = \sum_{i=1}^{N}\frac{1}{2\sigma_i^2}\left[y_i^2-\left(\frac{y_i}{1+\alpha_i w_i^*}\right)^2\right] > \lambda_g, \tag{5.15}$$

and

$$\begin{aligned}\lambda_g &= \text{erfc}^{-1}(2P_{FA}^*)\left[2\sum_{i=1}^{N}\frac{1}{2}\left[1-\frac{1}{(1+\alpha_i w_i^*)^2}\right]^2\right]^{1/2} \\ &\quad +\sum_{i=1}^{N}\frac{1}{2}\left[1-\frac{1}{(1+\alpha_i w_i^*)^2}\right],\end{aligned} \tag{5.16}$$

respectively. This leads to simplification in the computation of the decision rule and threshold. However, now the decision threshold becomes independent of the

original images used for watermark embedding. The presence of μ_i and σ_i^2 in (5.14) may provide more flexibility in the tuning of λ_g to suit the different images for better watermark detection.

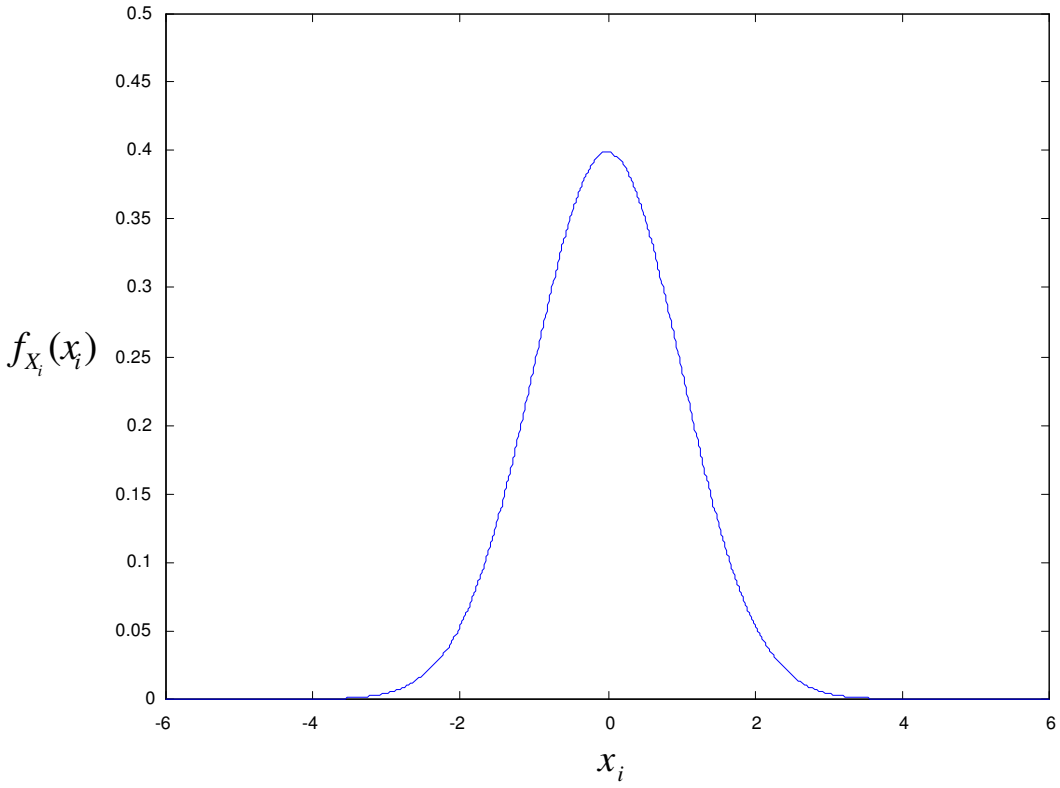

Figure 5.1: Gaussian PDF with $\mu_i = 0$ and $\sigma_i^2 = 1$

Chapter 6

LR Detector Based on Laplacian Model

Similar to Gaussian distribution, Laplacian distribution has a shape which is symmetrical about its mean. However, the peak of Laplacian distribution is sharp while the peak of Gaussian distribution is smooth. In this chapter, we model DWT coefficients using a Laplacian distribution. We show our work in deriving the LR decision rule and threshold under the Laplacian model. Results of numerical experiments that compare the Laplacian model and Gaussian model are given here.

6.1 LR Decision Rule

The Laplacian PDF is commonly used to model coefficients of DCT [6]. Here, we consider modeling the DWT coefficient X_i by a Laplacian RV. The PDF of X_i is

then expressed as

$$f_{X_i}(x_i) = \frac{b_i}{2}\exp(-b_i|x_i - \mu_i|), \quad -\infty < x_i < \infty, \tag{6.1}$$

with $b_i = \sqrt{2/\sigma_i^2}$, where σ_i^2 is the variance of X_i, and μ_i is the mean of X_i. Figure 6.1 shows then plot of Laplacian PDF with $\mu_i = 0$ and $\sigma_i^2 = 1$.

Substituting (6.1) in (4.16), we obtain the LR decision rule as

$$\begin{aligned} z(\mathbf{y}) &= \sum_{i=1}^{N}\left\{\ln\left[\frac{b_i}{2}\exp\left(-b_i\left|\frac{y_i}{1+\alpha_i w_i^*} - \mu_i\right|\right)\right] - \ln\left[\frac{b_i}{2}\exp(-b_i|y_i - \mu_i|)\right]\right\} \\ &= \sum_{i=1}^{N} b_i\left\{|y_i - \mu_i| - \frac{1}{|1+\alpha_i w_i^*|}\cdot|y_i - \mu_i - \mu_i\alpha_i w_i^*|\right\} > \lambda_l, \end{aligned} \tag{6.2}$$

where λ_l denotes the decision threshold under Laplacian model.

6.2 LR Decision Threshold

Compared to the Gaussian model in Chapter 5, it is much more complicated to derive a closed-form expression for LR decision threshold under the Laplacian model. As before, we need to first obtain the mean and variance of

$$z(\mathbf{X}) = \sum_{i=1}^{N} b_i\left\{|X_i - \mu_i| - \frac{1}{|1+\alpha_i w_i^*|}\cdot|X_i - \mu_i - \mu_i\alpha_i w_i^*|\right\}. \tag{6.3}$$

6.2.1 Derivation for Mean of $z(\mathbf{X})$

From (6.3), we see that to obtain the mean of $z(\mathbf{X})$, we need to derive the mean of

$$l(X_i) = b_i\left\{|X_i - \mu_i| - \frac{1}{|1+\alpha_i w_i^*|}\cdot|X_i - \mu_i - \mu_i\alpha_i w_i^*|\right\}. \tag{6.4}$$

Firstly,

$$\begin{aligned} E[|X_i - \mu_i|] &= \int_{-\infty}^{+\infty} |x_i - \mu_i| \cdot \frac{b_i}{2} \exp(-b_i|x_i - \mu_i|) dx_i \\ &= \frac{b_i}{2} \left\{ \int_{\mu_i}^{+\infty} (x_i - \mu_i) \exp(-b_i(x_i - \mu_i)) dx_i \right. \\ &\qquad \left. - \int_{-\infty}^{\mu_i} (x_i - \mu_i) \exp(b_i(x_i - \mu_i)) dx_i \right\}. \end{aligned} \tag{6.5}$$

Using integration by parts, we obtain

$$\begin{aligned} \int_{\mu_i}^{+\infty} (x_i - \mu_i) \exp(-b_i(x_i - \mu_i)) dx_i &= \frac{(x_i - \mu_i) \exp(-b_i(x_i - \mu_i))}{-b_i} \Big|_{\mu_i}^{+\infty} \\ &+ \int_{\mu_i}^{+\infty} \frac{\exp(-b_i(x_i - \mu_i))}{b_i} dx_i = \frac{1}{b_i^2} \end{aligned} \tag{6.6}$$

and

$$\begin{aligned} \int_{-\infty}^{\mu_i} (x_i - \mu_i) \exp(b_i(x_i - \mu_i)) dx_i &= \frac{(x_i - \mu_i) \exp(b_i(x_i - \mu_i))}{b_i} \Big|_{-\infty}^{\mu_i} \\ &- \int_{-\infty}^{\mu_i} \frac{\exp(b_i(x_i - \mu_i))}{b_i} dx_i = -\frac{1}{b_i^2}. \end{aligned} \tag{6.7}$$

Thus,

$$E[|X_i - \mu_i|] = \frac{b_i}{2} \left\{ \frac{1}{b_i^2} - \left(-\frac{1}{b_i^2} \right) \right\} = \frac{1}{b_i}. \tag{6.8}$$

Next, we derive $E[|X_i - \mu_i - \mu_i \alpha_i w_i^*|]$ by considering two cases.

Case I: $\mu_i \alpha_i w_i^* > 0$

$$\begin{aligned} E[|X_i - \mu_i - \mu_i \alpha_i w_i^*|] &= \int_{-\infty}^{+\infty} |X_i - \mu_i - \mu_i \alpha_i w_i^*| \cdot \frac{b_i}{2} \exp(-b_i|x_i - \mu_i|) dx_i \\ &= \frac{b_i}{2} \left\{ - \int_{-\infty}^{\mu_i} (x_i - \mu_i - \mu_i \alpha_i w_i^*) \exp(b_i(x_i - \mu_i)) dx_i \right. \\ &\quad - \int_{\mu_i}^{\mu_i + \mu_i \alpha_i w_i^*} (x_i - \mu_i - \mu_i \alpha_i w_i^*) \exp(-b_i(x_i - \mu_i)) dx_i \\ &\quad \left. + \int_{\mu_i + \mu_i \alpha_i w_i^*}^{+\infty} (x_i - \mu_i - \mu_i \alpha_i w_i^*) \exp(-b_i(x_i - \mu_i)) dx_i \right\} \\ &= \frac{b_i}{2} \{-A - B + C\}, \end{aligned} \tag{6.9}$$

where A, B, and C are the three integrals evaluated as

$$\begin{aligned}
A &= \int_{-\infty}^{\mu_i} (x_i - \mu_i - \mu_i\alpha_i w_i^*) \exp(b_i(x_i - \mu_i))dx_i \\
&= \int_{-\infty}^{\mu_i} (x_i - \mu_i) \exp(b_i(x_i - \mu_i))dx_i - \int_{-\infty}^{\mu_i} \mu_i\alpha_i w_i^* \exp(b_i(x_i - \mu_i))dx_i \\
&= \frac{(x_i - \mu_i) \exp(b_i(x_i - \mu_i))}{b_i}\Big|_{-\infty}^{\mu_i} - \left(\mu_i\alpha_i w_i^* + \frac{1}{b_i}\right) \int_{-\infty}^{\mu_i} \exp(b_i(x_i - \mu_i))dx_i \\
&= -\frac{\mu_i\alpha_i w_i^*}{b_i} - \frac{1}{b_i^2}, \qquad (6.10)
\end{aligned}$$

$$\begin{aligned}
B &= \int_{\mu_i}^{\mu_i+\mu_i\alpha_i w_i^*} (x_i - \mu_i - \mu_i\alpha_i w_i^*) \exp(-b_i(x_i - \mu_i))dx_i \\
&= \int_{\mu_i}^{\mu_i+\mu_i\alpha_i w_i^*} (x_i - \mu_i) \exp(-b_i(x_i - \mu_i)) \\
&\qquad - \int_{\mu_i}^{\mu_i+\mu_i\alpha_i w_i^*} \mu_i\alpha_i w_i^* \exp(-b_i(x_i - \mu_i))dx_i \\
&= \frac{(x_i - \mu_i) \exp(-b_i(x_i - \mu_i))}{-b_i}\Big|_{\mu_i}^{\mu_i+\mu_i\alpha_i w_i^*} \\
&\qquad - \left(\mu_i\alpha_i w_i^* - \frac{1}{b_i}\right) \int_{\mu_i}^{\mu_i+\mu_i\alpha_i w_i^*} \exp(-b_i(x_i - \mu_i))dx_i \\
&= -\frac{\mu_i\alpha_i w_i^* \exp(-b_i\mu_i\alpha_i w_i^*)}{b_i} - \left(\mu_i\alpha_i w_i^* - \frac{1}{b_i}\right) \left[-\frac{\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i} + \frac{1}{b_i}\right] \\
&= -\frac{\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i^2} - \frac{\mu_i\alpha_i w_i^*}{b_i} + \frac{1}{b_i^2}, \qquad (6.11)
\end{aligned}$$

and

$$\begin{aligned}
C &= \int_{\mu_i+\mu_i\alpha_i w_i^*}^{+\infty} (x_i - \mu_i - \mu_i\alpha_i w_i^*) \exp(-b_i(x_i - \mu_i))dx_i \\
&= \int_{\mu_i+\mu_i\alpha_i w_i^*}^{-\infty} (x_i - \mu_i) \exp(-b_i(x_i - \mu_i))dx_i \\
&\qquad - \int_{\mu_i+\mu_i\alpha_i w_i^*}^{-\infty} \mu_i\alpha_i w_i^* \exp(-b_i(x_i - \mu_i))dx_i \\
&= \frac{(x_i - \mu_i) \exp(-b_i(x_i - \mu_i))}{-b_i}\Big|_{\mu_i+\mu_i\alpha_i w_i^*}^{-\infty} \\
&\qquad + \left(\frac{1}{b_i} - \mu_i\alpha_i w_i^*\right) \int_{\mu_i+\mu_i\alpha_i w_i^*}^{-\infty} \exp(-b_i(x_i - \mu_i))dx_i
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\mu_i\alpha_i w_i^* \exp(-b_i\mu_i\alpha_i w_i^*)}{b_i} + \frac{\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i^2} - \frac{\mu_i\alpha_i w_i^* \exp(-b_i\mu_i\alpha_i w_i^*)}{b_i} \\
&= \frac{\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i^2}. \qquad (6.12)
\end{aligned}
$$

Substituting (6.10)-(6.12) in (6.9), we obtain

$$
\begin{aligned}
E[|X_i - \mu_i - \mu_i\alpha_i w_i^*|] &= \frac{b_i}{2}\left\{\frac{\mu_i\alpha_i w_i^*}{b_i} + \frac{1}{b_i^2} + \frac{\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i^2}\right. \\
&\quad \left. + \frac{\mu_i\alpha_i w_i^*}{b_i} - \frac{1}{b_i^2} + \frac{\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i^2}\right\} \\
&= \frac{\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i} + \mu_i\alpha_i w_i^*. \qquad (6.13)
\end{aligned}
$$

Case II: $\mu_i\alpha_i w_i^* < 0$

$$
\begin{aligned}
E[|X_i - \mu_i - \mu_i\alpha_i w_i^*|] &= \int_{-\infty}^{+\infty} |X_i - \mu_i - \mu_i\alpha_i w_i^*| \cdot \frac{b_i}{2}\exp(-b_i|x_i - \mu_i|)dx_i \\
&= \frac{b_i}{2}\left\{-\int_{-\infty}^{\mu_i+\mu_i\alpha_i w_i^*} (x_i - \mu_i - \mu_i\alpha_i w_i^*)\exp(b_i(x_i - \mu_i))dx_i\right. \\
&\quad + \int_{\mu_i+\mu_i\alpha_i w_i^*}^{\mu_i} (x_i - \mu_i - \mu_i\alpha_i w_i^*)\exp(b_i(x_i - \mu_i))dx_i \\
&\quad \left. + \int_{\mu_i}^{+\infty} (x_i - \mu_i - \mu_i\alpha_i w_i^*)\exp(-b_i(x_i - \mu_i))dx_i\right\} \\
&= \frac{b_i}{2}\{-D + E + F\}, \qquad (6.14)
\end{aligned}
$$

where D, E, and F are the three integrals evaluated as

$$
\begin{aligned}
D &= \int_{-\infty}^{\mu_i+\mu_i\alpha_i w_i^*} (x_i - \mu_i - \mu_i\alpha_i w_i^*)\exp(b_i(x_i - \mu_i))dx_i \\
&= \int_{-\infty}^{\mu_i+\mu_i\alpha_i w_i^*} (x_i - \mu_i)\exp(b_i(x_i - \mu_i))dx_i \\
&\qquad - \int_{-\infty}^{\mu_i+\mu_i\alpha_i w_i^*} \mu_i\alpha_i w_i^* \exp(b_i(x_i - \mu_i))dx_i \\
&= \frac{(x_i - \mu_i)\exp(b_i(x_i - \mu_i))}{b_i}\Bigg|_{-\infty}^{\mu_i+\mu_i\alpha_i w_i^*} \\
&\qquad - \left(\mu_i\alpha_i w_i^* + \frac{1}{b_i}\right)\int_{-\infty}^{\mu_i+\mu_i\alpha_i w_i^*} \exp(b_i(x_i - \mu_i))dx_i
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\mu_i\alpha_i w_i^* \exp(b_i\mu_i\alpha_i w_i^*)}{b_i} - \left(\mu_i\alpha_i w_i^* + \frac{1}{b_i}\right)\frac{\exp(b_i\mu_i\alpha_i w_i^*)}{b_i} \\
&= -\frac{\exp(b_i\mu_i\alpha_i w_i^*)}{b_i^2}, \qquad (6.15)
\end{aligned}
$$

$$
\begin{aligned}
E &= \int_{\mu_i+\mu_i\alpha_i w_i^*}^{\mu_i} (x_i - \mu_i - \mu_i\alpha_i w_i^*)\exp(b_i(x_i-\mu_i))dx_i \\
&= \int_{\mu_i+\mu_i\alpha_i w_i^*}^{\mu_i} (x_i-\mu_i)\exp(b_i(x_i-\mu_i)) \\
&\qquad - \int_{\mu_i+\mu_i\alpha_i w_i^*}^{\mu_i} \mu_i\alpha_i w_i^* \exp(b_i(x_i-\mu_i))dx_i \\
&= \frac{(x_i-\mu_i)\exp(b_i(x_i-\mu_i))}{b_i}\Big|_{\mu_i+\mu_i\alpha_i w_i^*}^{\mu_i} \\
&\qquad - \left(\mu_i\alpha_i w_i^* + \frac{1}{b_i}\right)\int_{\mu_i+\mu_i\alpha_i w_i^*}^{\mu_i} \exp(b_i(x_i-\mu_i))dx_i \\
&= -\frac{\mu_i\alpha_i\ w_i^*\exp(b_i\mu_i\alpha_i w_i^*)}{b_i} - \left(\mu_i\alpha_i w_i^* + \frac{1}{b_i}\right)\left[\frac{1}{b_i} - \frac{\exp(b_i\mu_i\alpha_i w_i^*)}{b_i}\right] \\
&= -\frac{\mu_i\alpha_i\ w_i^*\exp(b_i\mu_i\alpha_i w_i^*)}{b_i} - \frac{\mu_i\alpha_i\ w_i^*}{b_i} + \frac{\mu_i\alpha_i\ w_i^*\exp(b_i\mu_i\alpha_i w_i^*)}{b_i} \\
&\qquad - \frac{1}{b_i^2} + \frac{\exp(b_i\mu_i\alpha_i w_i^*)}{b_i^2} \\
&= -\frac{\mu_i\alpha_i\ w_i^*}{b_i} - \frac{1}{b_i^2} + \frac{\exp(b_i\mu_i\alpha_i w_i^*)}{b_i^2}, \qquad (6.16)
\end{aligned}
$$

and

$$
\begin{aligned}
F &= \int_{\mu_i}^{+\infty} (x_i - \mu_i - \mu_i\alpha_i w_i^*)\exp(-b_i(x_i-\mu_i))dx_i \\
&= \int_{\mu_i}^{-\infty} (x_i-\mu_i)\exp(-b_i(x_i-\mu_i))dx_i - \int_{\mu_i}^{-\infty} \mu_i\alpha_i w_i^*\exp(-b_i(x_i-\mu_i))dx_i \\
&= \frac{(x_i-\mu_i)\exp(-b_i(x_i-\mu_i))}{-b_i}\Big|_{\mu_i}^{-\infty} \\
&\qquad + \left(\frac{1}{b_i} - \mu_i\alpha_i w_i^*\right)\int_{\mu_i}^{-\infty} \exp(-b_i(x_i-\mu_i))dx_i \\
&= \left(\frac{1}{b_i} - \mu_i\alpha_i w_i^*\right)\frac{1}{b_i}. \qquad (6.17)
\end{aligned}
$$

Substituting (6.15)-(6.17) in (6.14), we obtain

$$\begin{aligned} E[|X_i - \mu_i - \mu_i\alpha_i w_i^*|] &= \frac{b_i}{2}\left\{\frac{\exp(b_i\mu_i\alpha_i w_i^*)}{b_i^2} - \frac{\mu_i\alpha_i\ w_i^*}{b_i}\right. \\ &\qquad \left. - \frac{1}{b_i^2} + \frac{\exp(b_i\mu_i\alpha_i w_i^*)}{b_i^2} + \frac{1}{b_i^2} - \frac{\mu_i\alpha_i w_i^*}{b_i}\right\} \\ &= \frac{\exp(b_i\mu_i\alpha_i w_i^*)}{b_i} - \mu_i\alpha_i w_i^*. \end{aligned} \tag{6.18}$$

Combining the results for (6.13) and (6.18), we obtain

$$E[|X_i - \mu_i - \mu_i\alpha_i w_i^*|] = \frac{\exp(-b_i|\mu_i\alpha_i w_i^*|)}{b_i} + |\mu_i\alpha_i w_i^*|. \tag{6.19}$$

With (6.8) and (6.19), the mean of $l(X_i)$ is given as

$$\begin{aligned} E[l(X_i)] &= b_i\left\{\frac{1}{b_i} - \frac{1}{|1+\alpha_i w_i^*|}\left[\frac{\exp(-b_i|\mu_i\alpha_i w_i^*|)}{b_i} + |\mu_i\alpha_i w_i^*|\right]\right\}. \\ &= 1 - \frac{1}{|1+\alpha_i w_i^*|}\cdot\left[\exp(-b_i|\mu_i\alpha_i w_i^*|) + b_i|\mu_i\alpha_i w_i^*|\right]. \end{aligned} \tag{6.20}$$

and hence

$$\begin{aligned} E[z(\mathbf{X})] &= \sum_{i=1}^{N} E[l(X_i)] \\ &= \sum_{i=1}^{N}\left\{1 - \frac{1}{|1+\alpha_i w_i^*|}\cdot\left[\exp(-b_i|\mu_i\alpha_i w_i^*|) + b_i|\mu_i\alpha_i w_i^*|\right]\right\}. \end{aligned} \tag{6.21}$$

6.2.2 Derivation for Variance of $z(\mathbf{X})$

In order to obtain variance of $z(\mathbf{X})$, we need to derive the mean of

$$\begin{aligned} l^2(X_i) &= b_i^2\left\{|X_i - \mu_i|^2 + \frac{1}{|1+\alpha_i w_i^*|^2}\cdot|X_i - \mu_i - \mu_i\alpha_i w_i^*|^2\right. \\ &\qquad \left. - \frac{2|X_i - \mu_i||X_i - \mu_i - \mu_i\alpha_i w_i^*|}{|1+\alpha_i w_i^*|}\right\}. \end{aligned} \tag{6.22}$$

Note that $E[|X_i - \mu_i|^2] = E[(X_i - \mu_i)^2] = \sigma_i^2 = 2/b_i^2$. Since $E[X_i - \mu_i] = 0$,

$$\begin{aligned} E[|X_i - \mu_i - \mu_i\alpha_i w_i^*|^2 &= E[(X_i - \mu_i - \mu_i\alpha_i w_i^*)^2] \\ &= E[(X_i - \mu_i)^2] + (\mu_i\alpha_i w_i^*)^2 - 2\mu_i\alpha_i w_i^* E[(X_i - \mu_i)] \\ &= \frac{2}{b_i^2} + (\mu_i\alpha_i w_i^*)^2. \end{aligned} \tag{6.23}$$

It is now left to derive $E[|X_i - \mu_i||X_i - \mu_i - \mu_i\alpha_i w_i^*|]$. As before, we need to consider two separate cases.

Case I: $\mu\alpha_i w_i^* > 0$

$$\begin{aligned} &E[|X_i - \mu_i||X_i - \mu_i - \mu_i\alpha_i w_i^*|] \\ &= \int_{-\infty}^{+\infty} |X_i - \mu_i||X_i - \mu_i - \mu_i\alpha_i w_i^*| \cdot \frac{b_i}{2}\exp(-b_i|x_i - \mu_i|)dx_i \\ &= \frac{b_i}{2}\Bigg\{ \int_{-\infty}^{\mu_i} (x_i - \mu_i)(x_i - \mu_i - \mu_i\alpha_i w_i^*)\exp(b_i(x_i - \mu_i))dx_i \\ &\quad - \int_{\mu_i}^{\mu_i + \mu_i\alpha_i w_i^*} (x_i - \mu_i)(x_i - \mu_i - \mu_i\alpha_i w_i^*)\exp(-b_i(x_i - \mu_i))dx_i \\ &\quad + \int_{\mu_i + \mu_i\alpha_i w_i^*}^{+\infty} (x_i - \mu_i)(x_i - \mu_i - \mu_i\alpha_i w_i^*)\exp(-b_i(x_i - \mu_i))dx_i \Bigg\} \\ &= \frac{b_i}{2}\{A' - B' + C'\}, \end{aligned} \tag{6.24}$$

where A', B' and C' are evaluated as

$$\begin{aligned} A' &= \int_{-\infty}^{\mu_i} (x_i - \mu_i)(x_i - \mu_i - \mu_i\alpha_i w_i^*)\exp(b_i(x_i - \mu_i))dx_i \\ &= \int_{-\infty}^{\mu_i} (x_i - \mu_i)^2\exp(b_i(x_i - \mu_i))dx_i \\ &\qquad - \int_{-\infty}^{\mu_i} \mu_i\alpha_i w_i^*(x_i - \mu_i)\exp(b_i(x_i - \mu_i))dx_i \\ &= \frac{(x_i - \mu_i)^2\exp(b_i(x_i - \mu_i))}{b_i}\Bigg|_{-\infty}^{\mu_i} \\ &\qquad - \left(\frac{2}{b_i} + \mu_i\alpha_i w_i^*\right)\int_{-\infty}^{\mu_i} (x_i - \mu_i)\exp(b_i(x_i - \mu_i))dx_i \end{aligned}$$

$$
\begin{aligned}
&= -\left(\frac{2}{b_i}+\mu_i\alpha_i w_i^*\right)\left\{\frac{(x_i-\mu_i)\exp(b_i(x_i-\mu_i))}{b_i}\Big|_{-\infty}^{\mu_i}\right.\\
&\qquad\qquad\left.-\int_{-\infty}^{\mu_i}\frac{\exp(b_i(x_i-\mu_i))}{b_i}dx_i\right\}\\
&= -\left(\frac{2}{b_i}+\mu_i\alpha_i w_i^*\right)\left\{\frac{\exp(b_i(x_i-\mu_i))}{-b_i^2}\Big|_{-\infty}^{\mu_i}\right\}\\
&= \frac{2}{b_i^3}+\frac{\mu_i\alpha_i w_i^*}{b_i^2}, \qquad (6.25)
\end{aligned}
$$

$$
\begin{aligned}
B' &= \int_{\mu_i}^{\mu_i+\mu_i\alpha_i w_i^*}(x_i-\mu_i)(x_i-\mu_i-\mu_i\alpha_i w_i^*)\exp(-b_i(x_i-\mu_i))dx_i\\
&= \int_{\mu_i}^{\mu_i+\mu_i\alpha_i w_i^*}(x_i-\mu_i)^2\exp(-b_i(x_i-\mu_i))dx_i\\
&\qquad -\int_{\mu_i}^{\mu_i+\mu_i\alpha_i w_i^*}\mu_i\alpha_i w_i^*(x_i-\mu_i)\exp(-b_i(x_i-\mu_i))dx_i\\
&= \frac{(x_i-\mu_i)^2\exp(-b_i(x_i-\mu_i))}{-b_i}\Big|_{\mu_i}^{\mu_i+\mu_i\alpha_i w_i^*}\\
&\qquad +\left(\frac{2}{b_i}-\mu_i\alpha_i w_i^*\right)\int_{\mu_i}^{\mu_i+\mu_i\alpha_i w_i^*}(x_i-\mu_i)\exp(-b_i(x_i-\mu_i))dx_i\\
&= -\frac{(\mu_i\alpha_i w_i^*)^2\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i}\\
&\qquad +\left(\frac{2}{b_i}-\mu_i\alpha_i w_i^*\right)\left\{\frac{(x_i-\mu_i)\exp(-b_i(x_i-\mu_i))}{-b_i}\Big|_{\mu_i}^{\mu_i+\mu_i\alpha_i w_i^*}\right.\\
&\qquad\qquad\left.-\int_{\mu_i}^{\mu_i+\mu_i\alpha_i w_i^*}\frac{\exp(-b_i(x_i-\mu_i))}{-b_i}dx_i\right\}\\
&= -\frac{(\mu_i\alpha_i w_i^*)^2\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i}\\
&\qquad +\left(\frac{2}{b_i}-\mu_i\alpha_i w_i^*\right)\left\{-\frac{\mu_i\alpha_i w_i^*\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i}-\frac{\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i^2}+\frac{1}{b_i^2}\right\}\\
&= -\frac{\mu_i\alpha_i w_i^*\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i^2}+\frac{2}{b_i^3}-\frac{2\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i^3}-\frac{\mu_i\alpha_i w_i^*}{b_i^2}, \qquad (6.26)
\end{aligned}
$$

and

$$
\begin{aligned}
C' &= \int_{\mu_i+\mu_i\alpha_i w_i^*}^{+\infty}(x_i-\mu_i)(x_i-\mu_i-\mu_i\alpha_i w_i^*)\exp(-b_i(x_i-\mu_i))dx_i\\
&= \int_{\mu_i+\mu_i\alpha_i w_i^*}^{+\infty}(x_i-\mu_i)^2\exp(-b_i(x_i-\mu_i))dx_i
\end{aligned}
$$

$$
\begin{aligned}
& - \int_{\mu_i+\mu_i\alpha_i w_i^*}^{+\infty} \mu_i\alpha_i w_i^*(x_i-\mu_i)\exp(-b_i(x_i-\mu_i))dx_i \\
= & \frac{(x_i-\mu_i)^2\exp(-b_i(x_i-\mu_i))}{-b_i}\Big|_{\mu_i+\mu_i\alpha_i w_i^*}^{+\infty} \\
& + \left(\frac{2}{b_i}-\mu_i\alpha_i w_i^*\right)\int_{\mu_i+\mu_i\alpha_i w_i^*}^{+\infty}(x_i-\mu_i)\exp(-b_i(x_i-\mu_i))dx_i \\
= & \frac{(\mu_i\alpha_i w_i^*)^2\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i} \\
& + \left(\frac{2}{b_i}-\mu_i\alpha_i w_i^*\right)\left\{\frac{(x_i-\mu_i)\exp(-b_i(x_i-\mu_i))}{-b_i}\Big|_{\mu_i+\mu_i\alpha_i w_i^*}^{+\infty}\right. \\
& \left. - \int_{\mu_i+\mu_i\alpha_i w_i^*}^{+\infty}\frac{\exp(-b_i(x_i-\mu_i))}{-b_i}dx_i\right\} \\
= & \frac{(\mu_i\alpha_i w_i^*)^2\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i} \\
& + \left(\frac{2}{b_i}-\mu_i\alpha_i w_i^*\right)\left\{\frac{\mu_i\alpha_i w_i^*\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i}+\frac{\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i^2}\right\} \\
= & \frac{\mu_i\alpha_i w_i^*\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i^2}+\frac{2\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i^3}. \qquad (6.27)
\end{aligned}
$$

Substituting (6.25)-(6.27) in (6.24) yields

$$
\begin{aligned}
E[|X_i-\mu_i||X_i-\mu_i-\mu_i\alpha_i w_i^*|] = & \frac{b_i}{2}\left\{\frac{2}{b_i^3}+\frac{\mu_i\alpha_i w_i^*}{b_i^2}+\frac{\mu_i\alpha_i w_i^*\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i^2}\right. \\
& -\frac{2}{b_i^3}+\frac{2\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i^3}+\frac{\mu_i\alpha_i w_i^*}{b_i^2} \\
& +\frac{\mu_i\alpha_i w_i^*\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i^2} \\
& \left.+\frac{2\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i^3}\right\} \\
= & \frac{\mu_i\alpha_i w_i^*}{b_i}+\frac{\mu_i\alpha_i w_i^*\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i} \\
& +\frac{2\exp(-b_i\mu_i\alpha_i w_i^*)}{b_i^2}. \qquad (6.28)
\end{aligned}
$$

Case II: $\mu\alpha_i w_i^* < 0$

$$E[|X_i-\mu_i||X_i-\mu_i-\mu_i\alpha_i w_i^*|]$$

$$
\begin{aligned}
&= \int_{-\infty}^{+\infty} |X_i - \mu_i||X_i - \mu_i - \mu_i\alpha_i w_i^*| \cdot \frac{b_i}{2}\exp(-b_i|x_i - \mu_i|)dx_i \\
&= \frac{b_i}{2}\left\{ \int_{-\infty}^{\mu_i+\mu_i\alpha_i w_i^*} (x_i - \mu_i)(x_i - \mu_i - \mu_i\alpha_i w_i^*)\exp(b_i(x_i - \mu_i))dx_i \right. \\
&\quad - \int_{\mu_i+\mu_i\alpha_i w_i^*}^{\mu_i} (x_i - \mu_i)(x_i - \mu_i - \mu_i\alpha_i w_i^*)\exp(b_i(x_i - \mu_i))dx_i \\
&\quad \left. + \int_{\mu_i}^{+\infty} (x_i - \mu_i)(x_i - \mu_i - \mu_i\alpha_i w_i^*)\exp(-b_i(x_i - \mu_i))dx_i \right\} \\
&= \frac{b_i}{2}\{D' - E' + F'\}, \qquad (6.29)
\end{aligned}
$$

where D', E' and F' are evaluated as

$$
\begin{aligned}
D' &= \int_{-\infty}^{\mu_i+\mu_i\alpha_i w_i^*} (x_i - \mu_i)(x_i - \mu_i - \mu_i\alpha_i w_i^*)\exp(b_i(x_i - \mu_i))dx_i \\
&= \int_{-\infty}^{\mu_i+\mu_i\alpha_i w_i^*} (x_i - \mu_i)^2\exp(b_i(x_i - \mu_i))dx_i \\
&\qquad - \int_{-\infty}^{\mu_i+\mu_i\alpha_i w_i^*} \mu_i\alpha_i w_i^*(x_i - \mu_i)\exp(b_i(x_i - \mu_i))dx_i \\
&= \frac{(x_i - \mu_i)^2\exp(b_i(x_i - \mu_i))}{b_i}\Big|_{-\infty}^{\mu_i+\mu_i\alpha_i w_i^*} \\
&\qquad - \left(\frac{2}{b_i} + \mu_i\alpha_i w_i^*\right)\int_{-\infty}^{\mu_i+\mu_i\alpha_i w_i^*} (x_i - \mu_i)\exp(b_i(x_i - \mu_i))dx_i \\
&= \frac{(\mu_i\alpha_i w_i^*)^2\exp(b_i\mu_i\alpha_i w_i^*)}{b_i} \\
&\qquad - \left(\frac{2}{b_i} + \mu_i\alpha_i w_i^*\right)\left\{ \frac{(x_i - \mu_i)\exp(b_i(x_i - \mu_i))}{b_i}\Big|_{\infty}^{\mu_i+\mu_i\alpha_i w_i^*} \right. \\
&\qquad\qquad \left. - \int_{\infty}^{\mu_i+\mu_i\alpha_i w_i^*} \frac{\exp(b_i(x_i - \mu_i))}{b_i}dx_i \right\} \\
&= \frac{(\mu_i\alpha_i w_i^*)^2\exp(b_i\mu_i\alpha_i w_i^*)}{b_i} \\
&\qquad - \left(\frac{2}{b_i} + \mu_i\alpha_i w_i^*\right)\left\{ \frac{\mu_i\alpha_i w_i^*\exp(b_i\mu_i\alpha_i w_i^*)}{b_i} - \frac{\exp(b_i\mu_i\alpha_i w_i^*)}{b_i^2} \right\} \\
&= -\frac{\mu_i\alpha_i w_i^*\exp(b_i\mu_i\alpha_i w_i^*)}{b_i^2} + \frac{2\exp(b_i\mu_i\alpha_i w_i^*)}{b_i^3}, \qquad (6.30)
\end{aligned}
$$

$$
E' = \int_{\mu_i+\mu_i\alpha_i w_i^*}^{\mu_i} (x_i - \mu_i)(x_i - \mu_i - \mu_i\alpha_i w_i^*)\exp(b_i(x_i - \mu_i))dx_i
$$

$$
\begin{aligned}
&= \int_{\mu_i+\mu_i\alpha_i w_i^*}^{\mu_i} (x_i-\mu_i)^2 \exp(b_i(x_i-\mu_i))dx_i \\
&\qquad - \int_{\mu_i+\mu_i\alpha_i w_i^*}^{\mu_i} \mu_i\alpha_i w_i^*(x_i-\mu_i)\exp(b_i(x_i-\mu_i))dx_i \\
&= \frac{(x_i-\mu_i)^2\exp(b_i(x_i-\mu_i))}{b_i}\bigg|_{\mu_i+\mu_i\alpha_i w_i^*}^{\mu_i} \\
&\qquad - \left(\frac{2}{b_i}+\mu_i\alpha_i w_i^*\right)\int_{\mu_i+\mu_i\alpha_i w_i^*}^{\mu_i}(x_i-\mu_i)\exp(b_i(x_i-\mu_i))dx_i \\
&= -\frac{(\mu_i\alpha_i w_i^*)^2\exp(b_i\mu_i\alpha_i w_i^*)}{b_i} \\
&\qquad - \left(\frac{2}{b_i}+\mu_i\alpha_i w_i^*\right)\left\{\frac{(x_i-\mu_i)\exp(b_i(x_i-\mu_i))}{b_i}\bigg|_{\mu_i+\mu_i\alpha_i w_i^*}^{\mu_i}\right. \\
&\qquad \left. - \int_{\mu_i+\mu_i\alpha_i w_i^*}^{\mu_i}\frac{\exp(b_i(x_i-\mu_i))}{b_i}dx_i\right\} \\
&= -\frac{(\mu_i\alpha_i w_i^*)^2\exp(b_i\mu_i\alpha_i w_i^*)}{b_i} \\
&\qquad - \left(\frac{2}{b_i}+\mu_i\alpha_i w_i^*\right)\left\{-\frac{\mu_i\alpha_i w_i^*\exp(b_i\mu_i\alpha_i w_i^*)}{b_i}-\frac{1}{b_i^2}+\frac{\exp(b_i\mu_i\alpha_i w_i^*)}{b_i^2}\right\} \\
&= \frac{\mu_i\alpha_i w_i^*\exp(b_i\mu_i\alpha_i w_i^*)}{b_i^2}+\frac{2}{b_i^3}-\frac{2\exp(b_i\mu_i\alpha_i w_i^*)}{b_i^3}+\frac{\mu_i\alpha_i w_i^*}{b_i^2}, \qquad (6.31)
\end{aligned}
$$

and

$$
\begin{aligned}
F' &= \int_{\mu_i}^{+\infty}(x_i-\mu_i)(x_i-\mu_i-\mu_i\alpha_i w_i^*)\exp(-b_i(x_i-\mu_i))dx_i \\
&= \int_{\mu_i}^{+\infty}(x_i-\mu_i)^2\exp(-b_i(x_i-\mu_i))dx_i \\
&\qquad - \int_{\mu_i}^{+\infty}\mu_i\alpha_i w_i^*(x_i-\mu_i)\exp(-b_i(x_i-\mu_i))dx_i \\
&= \frac{(x_i-\mu_i)^2\exp(-b_i(x_i-\mu_i))}{-b_i}\bigg|_{\mu_i}^{+\infty} \\
&\qquad + \left(\frac{2}{b_i}-\mu_i\alpha_i w_i^*\right)\int_{\mu_i}^{+\infty}(x_i-\mu_i)\exp(-b_i(x_i-\mu_i))dx_i \\
&= \left(\frac{2}{b_i}-\mu_i\alpha_i w_i^*\right)\left\{\frac{(x_i-\mu_i)\exp(-b_i(x_i-\mu_i))}{-b_i}\bigg|_{\mu_i}^{+\infty}\right. \\
&\qquad \left. - \int_{\mu_i}^{+\infty}\frac{\exp(-b_i(x_i-\mu_i))}{-b_i}dx_i\right\}
\end{aligned}
$$

$$
\begin{aligned}
&= \left(\frac{2}{b_i} - \mu_i\alpha_i w_i^*\right)\left\{\frac{1}{b_i^2}\right\} \\
&= \frac{2}{b_i^3} - \frac{\mu_i\alpha_i w_i^*}{b_i^2}. \qquad (6.32)
\end{aligned}
$$

Substituting (6.30)-(6.32) in (6.29) yields

$$
\begin{aligned}
E[|X_i-\mu_i||X_i-\mu_i-\mu_i\alpha_i w_i^*|] &= \frac{b_i}{2}\left\{-\frac{\mu_i\alpha_i w_i^*\exp(b_i\mu_i\alpha_i w_i^*)}{b_i^2}\right. \\
&\quad +\frac{2\exp(b_i\mu_i\alpha_i w_i^*)}{b_i^3} - \frac{\mu_i\alpha_i w_i^*\exp(b_i\mu_i\alpha_i w_i^*)}{b_i^2} \\
&\quad -\frac{2}{b_i^3} + \frac{2\exp(b_i\mu_i\alpha_i w_i^*)}{b_i^3} - \frac{\mu_i\alpha_i w_i^*}{b_i^2} \\
&\quad \left.+\frac{2}{b_i^3} - \frac{\mu_i\alpha_i w_i^*}{b_i^2}\right\} \\
&= -\frac{\mu_i\alpha_i w_i^*}{b_i} - \frac{\mu_i\alpha_i w_i^*\exp(b_i\mu_i\alpha_i w_i^*)}{b_i} \\
&\quad +\frac{2\exp(b_i\mu_i\alpha_i w_i^*)}{b_i^2}. \qquad (6.33)
\end{aligned}
$$

Combining the results of (6.28) and (6.33), we obtain

$$
\begin{aligned}
E[|X_i-\mu_i||X_i-\mu_i-\mu_i\alpha_i w_i^*|] &= \frac{|\mu_i\alpha_i w_i^*|}{b_i} + \frac{|\mu_i\alpha_i w_i^*|\exp(-b_i|\mu_i\alpha_i w_i^*|)}{b_i} \\
&\quad + \frac{2\exp(-b_i|\mu_i\alpha_i w_i^*|)}{b_i^2}. \qquad (6.34)
\end{aligned}
$$

Thus,

$$
\begin{aligned}
E[l^2(X_i)] &= b_i^2\left\{E[|X_i-\mu_i|]^2 + \frac{1}{|1+\alpha_i w_i^*|^2}E[|X_i-\mu_i-\mu_i\alpha_i w_i^*|^2]\right. \\
&\quad \left.-\frac{2E[|X_i-\mu_i||X_i-\mu_i-\mu_i\alpha_i w_i^*|]}{|1+\alpha_i w_i^*|}\right\} \\
&= b_i^2\left\{\frac{2}{b_i^2} + \frac{1}{|1+\alpha_i w_i^*|^2}\left\{\frac{2}{b_i^2} + (\mu_i\alpha_i w_i^*)^2\right\}\right. \\
&\quad -\frac{2}{|1+\alpha_i w_i^*|}\left\{\frac{|\mu_i\alpha_i w_i^*|}{b_i} + \frac{|\mu_i\alpha_i w_i^*|\exp(-b_i|\mu_i\alpha_i w_i^*|)}{b_i}\right. \\
&\quad \left.\left.+\frac{2\exp(-b_i|\mu_i\alpha_i w_i^*|)}{b_i^2}\right\}\right\} \\
&= 2 + \frac{2}{|1+\alpha_i w_i^*|^2} + \frac{(b_i\mu_i\alpha_i w_i^*)^2}{|1+\alpha_i w_i^*|^2} - \frac{2b_i|\mu_i\alpha_i w_i^*|}{|1+\alpha_i w_i^*|}
\end{aligned}
$$

$$- \frac{2b_i|\mu_i\alpha_i w_i^*|\exp(-b_i|\mu_i\alpha_i w_i^*|)}{|1+\alpha_i w_i^*|} - \frac{4\exp(-b_i|\mu_i\alpha_i w_i^*|)}{|1+\alpha_i w_i^*|}. \tag{6.35}$$

Squaring both sides of (6.20) yields

$$\begin{aligned} E^2[l(X_i)] &= 1 + \frac{(b_i\mu_i\alpha_i w_i^*)^2}{|1+\alpha_i w_i^*|^2} + \frac{\exp(-2b_i|\mu_i\alpha_i w_i^*|)}{|1+\alpha_i w_i^*|^2} \\ &\quad + \frac{2b_i|\mu_i\alpha_i w_i^*|\exp(-b_i|\mu_i\alpha_i w_i^*|)}{|1+\alpha_i w_i^*|^2} - \frac{2b_i|\mu_i\alpha_i w_i^*|}{|1+\alpha_i w_i^*|} \\ &\quad - \frac{2\exp(-b_i|\mu_i\alpha_i w_i^*|)}{|1+\alpha_i w_i^*|}. \end{aligned} \tag{6.36}$$

Hence,

$$\begin{aligned} V[z(\mathbf{X})] &= \sum_{i=1}^{N} V[l(X_i)] \\ &= \sum_{i=1}^{N} \left\{E[l^2(X_i))] - E^2[l(X_i)]\right\} \\ &= \sum_{i=1}^{N} \left\{ \left\{ 2 + \frac{2}{|1+\alpha_i w_i^*|^2} + \frac{(b_i\mu_i\alpha_i w_i^*)^2}{|1+\alpha_i w_i^*|^2} - \frac{2b_i|\mu_i\alpha_i w_i^*|}{|1+\alpha_i w_i^*|} \right.\right. \\ &\quad \left. - \frac{2b_i|\mu_i\alpha_i w_i^*|\exp(-b_i|\mu_i\alpha_i w_i^*|)}{|1+\alpha_i w_i^*|} - \frac{4\exp(-b_i|\mu_i\alpha_i w_i^*|)}{|1+\alpha_i w_i^*|} \right\} \\ &\quad - \left\{ 1 + \frac{(b_i\mu_i\alpha_i w_i^*)^2}{|1+\alpha_i w_i^*|^2} + \frac{\exp(-2b_i|\mu_i\alpha_i w_i^*|)}{|1+\alpha_i w_i^*|^2} \right. \\ &\quad + \frac{2b_i|\mu_i\alpha_i w_i^*|\exp(-b_i|\mu_i\alpha_i w_i^*|)}{|1+\alpha_i w_i^*|^2} - \frac{2b_i|\mu_i\alpha_i w_i^*|}{|1+\alpha_i w_i^*|} \\ &\quad \left.\left. - \frac{2\exp(-b_i|\mu_i\alpha_i w_i^*|)}{|1+\alpha_i w_i^*|} \right\} \right\} \\ &= \sum_{i=1}^{N} \left\{ 1 + \frac{1}{|1+\alpha_i w_i^*|^2}\left\{2 - \exp(-2b_i|\mu_i\alpha_i w_i^*|)\right\} \right. \\ &\quad - \frac{2\exp(-b_i|\mu_i\alpha_i w_i^*|)}{|1+\alpha_i w_i^*|} \\ &\quad \left. -2b_i|\mu_i\alpha_i w_i^*|\exp(-b_i|\mu_i\alpha_i w_i^*|)\left\{ \frac{1}{|1+\alpha_i w_i^*|^2} + \frac{1}{|1+\alpha_i w_i^*|} \right\} \right\}. \end{aligned} \tag{6.37}$$

6.2.3 Closed-Form Expression for λ_l

With (6.21) and (6.37), we obtain the expression for λ_l as

$$\begin{aligned}
\lambda_l &= \mathrm{erfc}^{-1}(2P_{FA}^*)\sqrt{2\sigma_{z(\mathbf{X})}^2} + \mu_{z(\mathbf{X})} \\
&= \mathrm{erfc}^{-1}(2P_{FA}^*)\left[2\sum_{i=1}^{N}\left\{1 + \frac{1}{|1+\alpha_i w_i^*|^2}\left\{2 - \exp(-2b_i|\mu_i\alpha_i w_i^*|)\right\}\right.\right. \\
&\qquad - \frac{2\exp(-b_i|\mu_i\alpha_i w_i^*|)}{|1+\alpha_i w_i^*|} \\
&\qquad \left.\left. -2b_i|\mu_i\alpha_i w_i^*|\exp(-b_i|\mu_i\alpha_i w_i^*|)\left\{\frac{1}{|1+\alpha_i w_i^*|^2} + \frac{1}{|1+\alpha_i w_i^*|}\right\}\right\}\right]^{1/2} \\
&\qquad + \sum_{i=1}^{N}\left\{1 - \frac{1}{|1+\alpha_i w_i^*|}\cdot\left[\exp(-b_i|\mu_i\alpha_i w_i^*|) + b_i|\mu_i\alpha_i w_i^*|\right]\right\}. \qquad (6.38)
\end{aligned}$$

Note that $b_i = \sqrt{2/\sigma_i^2}$. Thus, for a fixed P_{FA}^*, the decision threshold λ_l depends on μ_i, σ_i^2, α_i and w_i^* for $i = 1, 2, \ldots, N$.

6.3 Zero Mean Model

The decision rule and threshold of the zero mean model are given as

$$z(\mathbf{y}) = \sum_{i=1}^{N} b_i|y_i|\left\{1 - \frac{1}{|1+\alpha_i w_i^*|}\right\} > \lambda_l \qquad (6.39)$$

and

$$\lambda_l = \mathrm{erfc}^{-1}(2P_{FA}^*)\left[2\sum_{i=1}^{N}\left\{1 - \frac{1}{|1+\alpha_i w_i^*|}\right\}^2\right]^{1/2} + \sum_{i=1}^{N}\left\{1 - \frac{1}{|1+\alpha_i w_i^*|}\right\}, \qquad (6.40)$$

respectively. As in the Gaussian model, the decision threshold becomes independent of the original images.

6.4 Experimental Results

The Laplacian model is compared with the Gaussian model using 512×512 grayscales images as shown in Chapter 3 (see Figure 3.1). The zero mean Laplacian and Gaussian models are also included in the comparison. Also, as in Chapter 3, each image is transformed by DWT using a Daubechies filter to obtain a three-level pyramid decomposition (see Figure 3.2). For simplicity, watermark embedding is done in all coefficients belonging to the high-resolution subbands LH_3, HL_3 and HH_3. Each subband has $4,096$ identically distributed coefficients, and therefore $N = 12,288$.

Blind detection is used, i.e., the original image is not required in the detection process. This is done by estimating μ_i and σ_i^2 from the possibly distorted watermarked image. Let $\mathcal{B}$ be one of the subbands LH_3, HL_3 and HH_3. All coefficients in B are assumed to be identically distributed, i.e., they have identical PDF. If $x_i \in \mathcal{B}$, then μ_i and σ_i^2 are estimated from the unbiased estimators

$$\hat{\mu}_i = \frac{1}{N_{\mathcal{B}}} \sum_{y \in \mathcal{B}} y \tag{6.41}$$

and

$$\hat{\sigma}_i = \frac{1}{N_{\mathcal{B}} - 1} \sum_{y \in \mathcal{B}} (y - \hat{\mu}_i)^2, \tag{6.42}$$

respectively, where $N_{\mathcal{B}} = 4,096$ and y is the corresponding DWT coefficient in $\mathcal{B}$ of the watermarked image.

A constant embedding strength α is used for all the coefficients in LH_3, HL_3 and HH_3. Each value of α, as tabulated in Table 6.1, is chosen so that the PSNR

of the corresponding watermarked image is about 45 dB.

The robustness of the watermark is tested under different standard image processing operations using the procedure given in Section 4.3. Table 6.2 shows the results for watermarked images compressed by JPEG with a 50% quality factor. In Table 6.3, watermarked images are blurred using a 4×4 spatial filter. In Tables 6.4, 6.5 and 6.6, watermarked images are corrupted by Gaussian noise of zero mean and variance equals to 0.5, speckle noise of variance 0.1, and salt and pepper noise covering 30% of the pixels, respectively. Lastly, Table 6.7 shows results due to cropping. Each watermarked image is cropped to retain only 400×400 pixels at the center, where the missing portion is replaced by zero pixels so that the size of each image remains at 512×512. Overall, the results reveal that Laplacian model generally yields a better watermark detection than the Gaussian model.

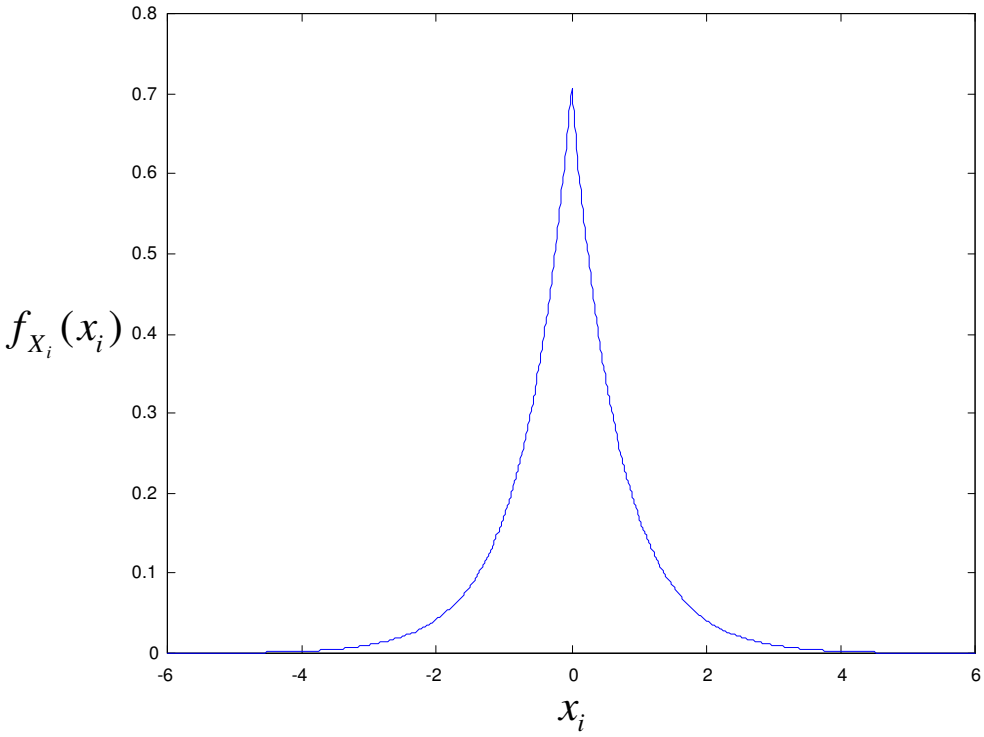

Figure 6.1: Laplacian PDF with $\mu_i = 0$ and $\sigma_i^2 = 1$.

Table 6.1: Watermark embedding strength for images.

Image	α
Harbour	0.205
Lena	0.185
Fishing boat	0.175
Peppers	0.165
Barbara	0.185
Goldhill	0.205
Zelda	0.215
LAX	0.190

Table 6.2: Percentage of successful detections under JPEG compression.

Image	Laplacian	Laplacian (zero mean)	Gaussian	Gaussian (zero mean)
Harbour	99.88	99.87	99.62	99.60
Lena	99.78	99.74	98.30	98.43
Fishing boat	99.82	99.80	99.23	99.22
Peppers	99.86	99.86	99.01	99.09
Barbara	99.93	99.95	99.13	99.10
Goldhill	99.90	99.93	98.33	98.30
Zelda	99.78	99.75	98.43	98.45
LAX	99.81	99.80	99.14	99.12

Table 6.3: Percentage of successful detections under low pass filtering.

Image	Laplacian	Laplacian (zero mean)	Gaussian	Gaussian (zero mean)
Harbour	99.87	99.88	99.39	99.40
Lena	99.45	99.49	99.03	99.03
Fishing boat	99.70	99.74	99.66	99.60
Peppers	99.56	99.52	99.67	99.64
Barbara	99.49	99.43	99.21	98.20
Goldhill	99.81	99.80	99.48	99.48
Zelda	99.38	99.38	99.07	99.03
LAX	99.65	99.64	99.55	99.55

Table 6.4: Percentage of successful detections under Gaussian noise.

Image	Laplacian	Laplacian (zero mean)	Gaussian	Gaussian (zero mean)
Harbour	99.77	99.75	99.42	99.35
Lena	99.56	99.58	99.10	99.14
Fishing boat	99.39	99.39	99.11	99.10
Peppers	99.76	99.77	99.12	99.14
Barbara	99.42	99.40	99.01	99.01
Goldhill	99.39	99.42	99.20	99.20
Zelda	99.19	99.17	99.32	99.33
LAX	99.27	99.30	99.01	99.00

Table 6.5: Percentage of successful detections under speckle noise.

Image	Laplacian	Laplacian (zero mean)	Gaussian	Gaussian (zero mean)
Harbour	99.71	99.68	99.35	99.33
Lena	99.49	99.50	99.06	99.08
Fishing boat	99.52	99.49	99.17	99.19
Peppers	99.69	99.69	99.22	99.22
Barbara	99.87	99.84	99.21	99.25
Goldhill	99.81	99.85	99.69	99.66
Zelda	99.72	99.77	99.36	99.36
LAX	99.68	99.65	99.75	99.71

Table 6.6: Percentage of successful detections under salt and pepper noise.

Image	Laplacian	Laplacian (zero mean)	Gaussian	Gaussian (zero mean)
Harbour	99.79	99.75	99.28	99.19
Lena	99.87	99.82	99.73	99.68
Fishing boat	99.67	99.63	99.60	99.67
Peppers	99.61	99.56	99.40	99.35
Barbara	99.80	99.85	99.66	99.59
Goldhill	99.68	99.68	99.31	99.31
Zelda	99.85	99.80	99.59	99.50
LAX	99.77	99.76	99.08	99.16

Table 6.7: Percentage of successful detections under cropping.

Image	Laplacian	Laplacian (zero mean)	Gaussian	Gaussian (zero mean)
Harbour	99.78	99.75	99.23	99.23
Lena	99.82	99.81	99.50	99.54
Fishing boat	99.87	99.86	99.23	99.19
Peppers	99.35	99.35	99.31	99.25
Barbara	99.56	99.53	99.60	99.60
Goldhill	99.81	99.79	99.34	99.30
Zelda	99.56	99.55	99.31	99.31
LAX	99.34	99.34	99.22	99.20

Chapter 7

LR Detector Based on Generalized Gaussian Model

The generalized Gaussian distribution is a general distribution that encompasses many important distributions. The Gaussian and Laplacian distributions are two of its special cases. In this chapter, we show our work in deriving the LR decision rule and threshold under a zero mean generalized Gaussian distribution. We compare this general model with the Gaussian and Laplacian models.

7.1 LR Decision Rule

Let the transform coefficients X_i be modeled by a zero mean generalized Gaussian RV. The PDF of X_i is then given as

$$f_{X_i}(x_i) = a_i \exp(-b_i^{\gamma_i}|x_i|^{\gamma_i}), \quad -\infty < x_i < +\infty, \tag{7.1}$$

where σ_i and $\gamma_i > 0$ are the variance and shape parameter of the distribution, respectively. The positive constant a_i and b_i are given as

$$a_i = \frac{b_i \gamma_i}{2\Gamma(1/\gamma_i)} \tag{7.2}$$

and

$$b_i = \frac{1}{\sigma_i}\sqrt{\frac{\Gamma(3/\gamma_i)}{\Gamma(1/\gamma_i)}}, \tag{7.3}$$

respectively, where Γ is the gamma function as given by (2.29). Note that $\gamma_i = 1$ yields the Laplacian PDF and $\gamma_i = 2$ yields the Gaussian PDF. Figure 7.1 shows the shape of the generalized Gaussian PDF with $\mu_i = 0$ and $\sigma_i^2 = 1$.

Substituting (7.1) in (4.16), we obtain the LR decision rule as

$$\begin{aligned} z(\mathbf{y}) &= \sum_{i=1}^{N} \left\{ \ln\left[a_i \exp\left(-b_i^{\gamma_i} \left| \frac{y_i}{1+\alpha_i w_i^*} \right|^{\gamma_i} \right) \right] - \ln\left[a_i \exp\left(-b_i^{\gamma_i} |y_i|^{\gamma_i} \right) \right] \right\} \\ &= \sum_{i=1}^{N} |y_i|^{\gamma_i} b_i^{\gamma_i} \left[1 - \frac{1}{|1+\alpha_i w_i^*|^{\gamma_i}} \right] > \lambda_{gg}, \end{aligned} \tag{7.4}$$

where λ_{gg} is the LR decision threshold under generalized Gaussian model. Further, with the substitution of (7.3), we can write (7.4) as

$$z(\mathbf{y}) = \sum_{i=1}^{N} |y_i|^{\gamma_i} \left[\frac{1}{\sigma_i}\sqrt{\frac{\Gamma(3/\gamma_i)}{\Gamma(1/\gamma_i)}} \right]^{\gamma_i} \left[1 - \frac{1}{|1+\alpha_i w_i^*|^{\gamma_i}} \right] > \lambda_{gg}. \tag{7.5}$$

7.2 LR Decision Threshold

In view of (4.22), in order to derive λ_{gg}, we need to first obtain the mean and variance of

$$z(\mathbf{X}) = \sum_{i=1}^{N} |X_i|^{\gamma_i} \left[\frac{1}{\sigma_i}\sqrt{\frac{\Gamma(3/\gamma_i)}{\Gamma(1/\gamma_i)}} \right]^{\gamma_i} \left[1 - \frac{1}{|1+\alpha_i w_i^*|^{\gamma_i}} \right]. \tag{7.6}$$

7.2.1 Derivation for Mean of $z(\mathbf{X})$

Since the term $\left[\frac{1}{\sigma_i}\sqrt{\frac{\Gamma(3/\gamma_i)}{\Gamma(1/\gamma_i)}}\right]^{\gamma_i}\left[1-\frac{1}{|1+\alpha_i w_i^*|^{\gamma_i}}\right]$ in (7.6) is constant with respect to X_i, we only need to obtain the mean and variance of $|X_i|^{\gamma_i}$. More generally, we give an expression for $E[|X_i|^n]$, where n is a real constant, in the following lemma.

Lemma 7.1 *If X_i is a zero mean generalized Gaussian random variable and n is a real constant, then*

$$E[|X_i|^n] = \sigma_i^n \frac{\Gamma((n+1)/\gamma_i)}{\Gamma^{1-n/2}(1/\gamma_i)\Gamma^{n/2}(3/\gamma_i)}. \tag{7.7}$$

Proof: By straightforward integration,

$$\begin{aligned} E[|X_i|^n] &= \int_{-\infty}^{+\infty} |x_i|^n a_i \exp(-b_i^{\gamma_i}|x_i|^{\gamma_i})dx_i \\ &= 2a_i \int_0^{+\infty} x_i^n \exp(-(b_i x_i)^{\gamma_i})dx_i \\ &= \frac{2a_i}{\gamma_i b_i^{n+1}} \int_0^{+\infty} t^{\frac{n+1}{\gamma_i}-1}\exp(-t)dt, \quad t=(b_i x_i)^{\gamma_i} \\ &= \frac{2a_i}{\gamma_i b_i^{n+1}}\Gamma\left(\frac{n+1}{\gamma_i}\right). \end{aligned} \tag{7.8}$$

Substituting a_i and b_i as given in (7.2) and (7.3), respectively, we obtain (7.7).

Q.E.D

In particular, when n is a positive even integer, the even moments of X_i is given by (7.7). When $n=\gamma_i$, it follows from the property of gamma function that

$$\Gamma((n+1)/\gamma_i) = \Gamma(1+1/\gamma_i) = \frac{1}{\gamma_i}\Gamma(1/\gamma_i). \tag{7.9}$$

Substituting (7.9) in (7.7) yields

$$E[|X_i|^{\gamma_i}] = \frac{\sigma_i^{\gamma_i}}{\gamma_i}\left[\frac{\Gamma(1/\gamma_i)}{\Gamma(3/\gamma_i)}\right]^{\gamma_i/2}. \tag{7.10}$$

With this, we obtain the mean of $z(\mathbf{X})$ as

$$\begin{aligned}\mu_{z(\mathbf{X})} &= \sum_{i=1}^{N} E[|X_i|^{\gamma_i}]\left[\frac{1}{\sigma_i}\sqrt{\frac{\Gamma(3/\gamma_i)}{\Gamma(1/\gamma_i)}}\right]^{\gamma_i}\left[1-\frac{1}{|1+\alpha_i w_i^*|^{\gamma_i}}\right] \\ &= \sum_{i=1}^{N}\frac{1}{\gamma_i}\left[1-\frac{1}{|1+\alpha_i w_i^*|^{\gamma_i}}\right]. \end{aligned} \tag{7.11}$$

7.2.2 Derivation for Variance of $z(\mathbf{X})$

When $n = 2\gamma_i$, using the property of gamma function again, we obtain

$$\begin{aligned}\Gamma((n+1)/\gamma_i) &= \Gamma(2+1/\gamma_i) \\ &= \left(1+\frac{1}{\gamma_i}\right)\Gamma(1+1/\gamma_i) \\ &= \left(1+\frac{1}{\gamma_i}\right)\frac{1}{\gamma_i}\Gamma(1/\gamma_i). \end{aligned} \tag{7.12}$$

Substituting (7.12) in (7.7) yields

$$E[|X_i|^{2\gamma_i}] = \left(1+\frac{1}{\gamma_i}\right)\frac{\sigma_i^{2\gamma_i}}{\gamma_i}\left[\frac{\Gamma(1/\gamma_i)}{\Gamma(3/\gamma_i)}\right]^{\gamma_i}. \tag{7.13}$$

With (7.10) and (7.13), the variance of $|X_i|^{\gamma_i}$ is given as

$$\begin{aligned}V[|X_i|^{\gamma_i}] &= E[|X_i|^{2\gamma_i}] - (E[|X_i|^{\gamma_i}])^2 \\ &= \left(1+\frac{1}{\gamma_i}\right)\frac{\sigma_i^{2\gamma_i}}{\gamma_i}\left[\frac{\Gamma(1/\gamma_i)}{\Gamma(3/\gamma_i)}\right]^{\gamma_i} - \frac{\sigma_i^{2\gamma_i}}{\gamma_i^2}\left[\frac{\Gamma(1/\gamma_i)}{\Gamma(3/\gamma_i)}\right]^{\gamma_i} \\ &= \frac{\sigma_i^{2\gamma_i}}{\gamma_i}\left[\frac{\Gamma(1/\gamma_i)}{\Gamma(3/\gamma_i)}\right]^{\gamma_i}. \end{aligned} \tag{7.14}$$

Finally, we obtain the variance of $z(\mathbf{X})$ as

$$\begin{aligned}\sigma^2_{z(\mathbf{X})} &= \sum_{i=1}^{N} V[|X_i|^{\gamma_i}] \left[\frac{1}{\sigma_i}\sqrt{\frac{\Gamma(3/\gamma_i)}{\Gamma(1/\gamma_i)}}\right]^{2\gamma_i} \left[1 - \frac{1}{|1+\alpha_i w_i^*|^{\gamma_i}}\right]^2 \\ &= \sum_{i=1}^{N} \frac{1}{\gamma_i}\left[1 - \frac{1}{|1+\alpha_i w_i^*|^{\gamma_i}}\right]^2. \end{aligned} \tag{7.15}$$

7.2.3 Closed-Form Expression for λ_{gg}

With (7.11) and (7.15), the decision threshold λ_{gg} is given as

$$\begin{aligned}\lambda_{gg} &= \text{erfc}^{-1}(2P^*_{FA})\sqrt{2\sigma^2_{z(\mathbf{X})}} + \mu_{z(\mathbf{X})} \\ &= \text{erfc}^{-1}(2P^*_{FA}) \left[2\sum_{i=1}^{N} \frac{1}{\gamma_i}\left[1 - \frac{1}{|1+\alpha_i w_i^*|^{\gamma_i}}\right]^2\right]^{1/2} \\ &\qquad + \sum_{i=1}^{N} \frac{1}{\gamma_i}\left[1 - \frac{1}{|1+\alpha_i w_i^*|^{\gamma_i}}\right]. \end{aligned} \tag{7.16}$$

For a fixed P^*_{FA}, λ_{gg} depends on γ_i, α_i and w_i^*. Note that, for $i = 1, 2, \ldots, N$, if $\gamma_i = 1$ then (7.16) reduces to the decision threshold of the zero mean Laplacian model in (6.40), and if $\gamma_i = 2$ then it reduces to the decision threshold of the zero mean Gaussian model in (5.16).

7.3 Parameter Estimation

For blind detection, in view of (7.5) and (7.16), the variance σ_i^2 and shape parameter γ_i are estimated from transform coefficients of the watermarked image. As in Chapter 6, all our experiments for generalized Gaussian model are conducted in DWT domain.

Let $\mathcal{B}$ be the DWT subband containing x_i and having $N_{\mathcal{B}}$ identically distributed coefficients. As in (6.42), the unbiased estimator of σ_i^2 is given as

$$\hat{\sigma}_i^2 = \frac{1}{N_{\mathcal{B}} - 1} \sum_{y \in \mathcal{B}} y^2, \tag{7.17}$$

where y is the corresponding DWT coefficient of the watermarked image (possibly distorted) in $\mathcal{B}$.

By taking $n = 1$ in (7.7), and then squaring both sides, we note that the mean and shape parameter of X_i can be related as [28]

$$\frac{E^2[|X_i|]}{\sigma_i^2} = \frac{\Gamma^2(2/\gamma_i)}{\Gamma(1/\gamma_i)\Gamma(3/\gamma_i)} = s(\gamma_i). \tag{7.18}$$

An estimator for γ_i can then be obtained as

$$\hat{\gamma_i} = s^{-1}\left(\frac{1}{\hat{\sigma}_i^2}\left[\frac{1}{N_{\mathcal{B}}}\sum_{y \in \mathcal{B}} |y|\right]\right). \tag{7.19}$$

where $\frac{1}{N_{\mathcal{B}}}\sum_{y \in \mathcal{B}} |y|$ is used to estimate $E[|X|]$. One way to solve (7.19) is to approximate the inverse of s using any of the well-known function interpolation method [21]. The knowledge of the range of s is important to achieve the desire accuracy in the interpolation process. Lemma 7.2 gives the exact range of s.

Lemma 7.2 *The function s as defined in (7.18) is a strictly increasing function with*

$$\lim_{\gamma_i \to +\infty} s(\gamma_i) = 3/4 \tag{7.20}$$

and

$$\lim_{\gamma_i \to 0^+} s(\gamma_i) = 0. \tag{7.21}$$

Proof: By use of the Euler formula given by (2.33), we can express $s(\gamma_i)$ as

$$\begin{aligned} s(\gamma_i) &= \frac{\left[\frac{\gamma_i}{2}\prod_{j=1}^{\infty}\left(1+\frac{1}{j}\right)^{\frac{2}{\gamma_i}}\left(1+\frac{2}{\gamma_i j}\right)^{-1}\right]^2}{\left[\gamma_i\prod_{j=1}^{\infty}\left(1+\frac{1}{j}\right)^{\frac{1}{\gamma_i}}\left(1+\frac{1}{\gamma_i j}\right)^{-1}\right]\left[\frac{\gamma_i}{3}\prod_{j=1}^{\infty}\left(1+\frac{1}{j}\right)^{\frac{3}{\gamma_i}}\left(1+\frac{3}{\gamma_i j}\right)^{-1}\right]} \\ &= \frac{3}{4}\prod_{j=1}^{\infty}\frac{\left(1+\frac{2}{\gamma_i j}\right)^{-2}}{\left(1+\frac{1}{\gamma_i j}\right)^{-1}\left(1+\frac{3}{\gamma_i j}\right)^{-1}}. \end{aligned} \tag{7.22}$$

If γ_i tends to $+\infty$, then the infinite product tends to 1. Thus, $s(\gamma_i)$ tends to 3/4. On the other hand, with straightforward rearrangement and expansion, we can express (7.22) as

$$s(\gamma_i) = \frac{3}{4}\prod_{j=1}^{\infty}\frac{(\gamma_i j)^2+4\gamma_i j+3}{(\gamma_i j)^2+4\gamma_i j+4}. \tag{7.23}$$

Now if γ_i tends to 0 from the right, then each term in the infinite product tends to 3/4. Thus, the infinite product tends to 0. Finally, we note that, for $\gamma_i \geq 0$,

$$\begin{aligned} \frac{d}{d\gamma_i}\left[\frac{(\gamma_i j)^2+4\gamma_i j+3}{(\gamma_i j)^2+4\gamma_i j+4}\right] &= \frac{2\gamma_i j^2+4j}{((\gamma_i j)^2+4\gamma_i j+4)^2} \\ &= \frac{2j}{(\gamma_i j+2)^3} > 0. \end{aligned} \tag{7.24}$$

This implies that s is strictly increasing, i.e., it increases from 0 to 3/4 when γ_i increases from 0 to $+\infty$.

Q.E.D

Figure 7.2 shows the plot of $s(\gamma_i)$ versus γ_i. In [46], the reciprocal function $r(\gamma_i) = 1/s(\gamma_i)$ is referred to as the generalized Gaussian ratio function, and γ_i is estimated as

$$\hat{\gamma}_i = r^{-1}\left(\frac{\hat{\sigma}_i^2}{\left[\frac{1}{N_{\mathcal{B}}}\sum_{y\in\mathcal{B}}|y|\right]^2}\right) \tag{7.25}$$

instead. However, since the range of $r(\gamma_i)$ is infinite, it is more practical to estimate γ_i using (7.19).

Similar types of estimator can be obtained by considering higher absolute moments of X_i. A rearrangement of (7.7) yields

$$\frac{\sigma_i^n}{E[|X_i|^n]} = \frac{\Gamma^{1-n/2}(1/\gamma_i)\Gamma^{n/2}(3/\gamma_i)}{\Gamma((n+1)/\gamma_i)}. \tag{7.26}$$

This leads us to define

$$\phi_n(\gamma_i) = \frac{\Gamma^{1-n/2}(1/\gamma_i)\Gamma^{n/2}(3/\gamma_i)}{\Gamma((n+1)/\gamma_i)}. \tag{7.27}$$

It follows that $\phi_2(\gamma_i) = 1$ and $\phi_i^2(\gamma_i) = 1/s(\gamma_i)$. The range of ϕ_n is given by Lemma 7.3.

Lemma 7.3 *The function ϕ_n as defined in (7.27) is a strictly decreasing function for $n = 1$ and strictly increasing function for $n > 2$ with*

$$\lim_{\gamma_i \to +\infty} \phi_n(\gamma_i) = \frac{n+1}{3^{n/2}} \tag{7.28}$$

and

$$\lim_{\gamma_i \to 0^+} \phi_n(\gamma_i) = \begin{cases} +\infty & n = 1 \\ 1 & n = 2 \\ 0 & n > 2 \end{cases}. \tag{7.29}$$

Proof: By use of the Weierstrass formula (2.34), we can express $\phi_n(\gamma_i)$ as

$$\begin{aligned} \phi_n(\gamma_i) &= \frac{\left[\frac{1}{\gamma_i} e^{\frac{\zeta}{\gamma_i}} \prod_{k=1}^{+\infty} \left(1 + \frac{1}{\gamma_i k}\right) e^{-\frac{1}{\gamma_i k}}\right]^{n/2-1} \left[\frac{3}{\gamma_i} e^{\frac{3\zeta}{\gamma_i}} \prod_{k=1}^{+\infty} \left(1 + \frac{3}{\gamma_i k}\right) e^{-\frac{3}{\gamma_i k}}\right]^{-n/2}}{\left[\frac{n+1}{\gamma_i} e^{\frac{(n+1)\zeta}{\gamma_i}} \prod_{k=1}^{+\infty} \left(1 + \frac{n+1}{\gamma_i k}\right) e^{-\frac{n+1}{\gamma_i k}}\right]^{-1}} \\ &= \frac{n+1}{3^{n/2}} \prod_{k=1}^{+\infty} \frac{\left(1 + \frac{1}{\gamma_i k}\right)^{n/2-1} \left(1 + \frac{3}{\gamma_i k}\right)^{-n/2}}{\left(1 + \frac{n+1}{\gamma_i k}\right)^{-1}}. \end{aligned} \tag{7.30}$$

When $\gamma_i \to +\infty$, the infinite products tend to one resulting in $\phi_n(\gamma_i) \to (n+1)/3^{n/2}$. On the other hand, with some rearrangements, we can express (7.30) as

$$\phi_n(\gamma_i) = \frac{n+1}{3^{n/2}} \prod_{k=1}^{+\infty} \left(\frac{\gamma_i k + 1}{\gamma_i k + 3} \right)^{n/2} \frac{(\gamma_i k + n + 1)}{(\gamma_i k + 1)}. \tag{7.31}$$

With this, when $\gamma_i \to 0^+$, each term in the infinite products tend to $(n+1)/3^{n/2}$. Thus, (7.29) follows readily. Moreover,

$$\begin{aligned}
&\frac{d}{d\gamma_i} \left(\frac{\gamma_i k + 1}{\gamma_i k + 3} \right)^{n/2} \frac{\gamma_i k + n + 1}{\gamma_i k + 1} \\
&= \frac{(\gamma_i k + 3)^{n/2}(\gamma_i k + 1)[(\gamma_i k + 1)^{n/2} k + (\gamma_i k + n + 1) \cdot \frac{kn}{2}(\gamma_i k + 1)^{n/2-1}]}{(\gamma_i k + 3)^n (\gamma_i k + 1)^2} \\
&- \frac{(\gamma_i k + 1)^{n/2}(\gamma_i k + n + 1)[(\gamma_i k + 3)^{n/2} k + (\gamma_i k + 1)\frac{kn}{2}(\gamma_i k + 3)^{n/2-1}]}{(\gamma_i k + 3)^n (\gamma_i k + 1)^2} \\
&= \frac{k(\gamma_i k + 3)^{n/2-1}(\gamma_i k + 1)^{n/2}}{(\gamma_i k + 3)^n (\gamma_i k + 1)^2} \\
&\qquad \times \Big\{ [(\gamma_i k + 3)(\gamma_i k + 1) + \frac{n}{2}(\gamma_i k + n + 1)(\gamma_i k + 3)] \\
&\qquad\quad - [(\gamma_i k + 3)(\gamma_i k + n + 1) + \frac{n}{2}(\gamma_i k + n + 1)(\gamma_i k + 1)] \Big\} \\
&= \frac{k(\gamma_i k + 3)^{n/2-1}(\gamma_i k + 1)^{n/2}}{(\gamma_i k + 3)^n (\gamma_i k + 1)^2} \\
&\qquad \times \Big\{ -n(\gamma_i k + 3) + \frac{n}{2}(\gamma_i k + n + 1)[(\gamma_i k + 3) - (\gamma_i k + 1)] \Big\} \\
&= \frac{k(\gamma_i k + 3)^{n/2-1}(\gamma_i k + 1)^{n/2}}{(\gamma_i k + 3)^n (\gamma_i k + 1)^2} \{n(n-2)\} \\
&= \frac{kn(\gamma_i k + 1)^{n/2-2}(n-2)}{(\gamma_i k + 3)^{1+n/2}}.
\end{aligned} \tag{7.32}$$

This means, for $\gamma_i \geq 0$, each term in the infinite products is strictly decreasing when $n = 1$ and strictly increasing when $n > 2$. Thus, the same goes for $\phi_n(\gamma_i)$.

Q.E.D

Figure 7.3 shows the plot of $\phi_n(\gamma_i)$ versus γ_i. For $n > 2$, since the range of $\phi_n(\gamma_i)$ is finite, we can also consider

$$\hat{\gamma}_i = \phi_n^{-1}\left(\hat{\sigma}_i^n \left[\frac{1}{N_{\mathcal{B}}}\sum_{y\in\mathcal{B}}|y|^n\right]^{-1}\right) \tag{7.33}$$

as an estimator for γ_i, where $\frac{1}{N_{\mathcal{B}}}\sum_{y\in\mathcal{B}}|y|^n$ is used to estimate $E[|X_i|^n]$.

7.4 Non-Zero Mean Model

In this section, we give extension to the non-zero mean generalized Gaussian model. This is more of a theoretical interest as our numerical experiments reveal that there is no improvement over the zero mean model, which is somewhat expected as the coefficients from the DWT subbands have approximately zero mean.

The generalized Gaussian PDF with non-zero mean μ_i is expressed as

$$f_{X_i}(x_i) = a_i \exp(-b_i^{\gamma_i}|x_i - \mu_i|^{\gamma_i}), \tag{7.34}$$

where the positive constants a_i and b_i are as defined in (7.2) and (7.3), respectively. Similar to the way (7.8) is derived, it can be shown that

$$E[|X_i - \mu_i|^n] = \frac{2a_i}{\gamma_i b_i^{n+1}}\Gamma\left(\frac{n+1}{\gamma_i}\right), \tag{7.35}$$

and with substitution of a_i and b_i yields

$$E[|X_i - \mu_i|^n] = \sigma_i^n \frac{\Gamma((n+1)/\gamma_i)}{\Gamma^{1-n/2}(1/\gamma_i)\Gamma^{n/2}(3/\gamma_i)}. \tag{7.36}$$

The LR decision rule under the non-zero mean model is

$$z(\mathbf{y}) = \sum_{i=1}^{N} b^{\gamma_i} \left[|y_i - \mu_i|^{\gamma_i} - \left| \frac{y_i}{1+\alpha_i w_i^*} - \mu_i \right|^{\gamma_i} \right] > \lambda_{gg}. \tag{7.37}$$

In order to obtain the LR decision threshold λ_{gg}, we need to find the mean and variance of $b^{\gamma_i} \left[|y_i - \mu_i|^{\gamma_i} - \left| \frac{y_i}{1+\alpha_i w_i^*} - \mu_i \right|^{\gamma_i} \right]$. Equivalently, we need to find the mean and variance for both $|X_i - \mu_i|^{\gamma_i}$ and $\left| \frac{X_i}{1+\alpha_i w_i^*} - \mu_i \right|^{\gamma_i}$. First, setting $n = \gamma_i$ in (7.35) and by use of (2.32), we obtain

$$E[|X_i - \mu_i|^{\gamma_i}] = \frac{2a_i}{\gamma_i^2 b_i^{\gamma_i+1}} \Gamma\left(\frac{1}{\gamma_i}\right) \tag{7.38}$$

and

$$\begin{aligned} V[|X_i - \mu_i|^{\gamma_i}] &= E[|X_i - \mu_i|^{2\gamma_i}] - E^2[|X_i - \mu_i|^{\gamma_i}] \\ &= \frac{2a_i}{\gamma_i^2 b_i^{2\gamma_i+1}} \left(1 + \frac{1}{\gamma_i}\right) \Gamma\left(\frac{1}{\gamma_i}\right) - \frac{4a_i^2}{\gamma_i^4 b_i^{2\gamma_i+2}} \Gamma^2\left(\frac{1}{\gamma_i}\right). \end{aligned} \tag{7.39}$$

Next, we derive the mean and variance of $\left| \frac{X_i}{1+\alpha_i w_i^*} - \mu_i \right|^{\gamma_i}$. Note that

$$E[(X_i - \mu_i)^{2k}] = E[|X_i - \mu_i|^{2k}] \tag{7.40}$$

for $k = 0, 1, 2, \ldots$. Thus, (7.35) can be used to obtain the even central moments of X_i. On the other hand, the odd central moments of X_i are zero, i.e.,

$$E[(X_i - \mu_i)^{2k+1}] = 0 \tag{7.41}$$

for $k = 0, 1, 2, \ldots$.

When γ_i is an even integer, by use of (7.35), (7.40) and (7.41), we obtain

$$E\left[\left| \frac{X_i}{1+\alpha_i w_i^*} - \mu_i \right|^{\gamma_i} \right]$$

$$
\begin{aligned}
&= \frac{1}{(1+\alpha_i w_i^*)^{\gamma_i}} E[(X_i - \mu_i - \mu_i\alpha_i w_i^*)^{\gamma_i}] \\
&= \frac{1}{(1+\alpha_i w_i^*)^{\gamma_i}} E\left[\sum_{k=0}^{\gamma_i} \frac{\gamma_i!}{k!(\gamma_i-k)!}(X_i-\mu_i)^k(-\mu_i\alpha_i w_i^*)^{\gamma_i-k}\right] \\
&= \frac{1}{(1+\alpha_i w_i^*)^{\gamma_i}} \sum_{k=0}^{\gamma_i} \frac{\gamma_i!}{k!(\gamma_i-k)!} E[(X_i-\mu_i)^k](-\mu_i\alpha_i w_i^*)^{\gamma_i-k} \\
&= \frac{2a_i(\gamma_i-1)!}{(1+\alpha_i w_i^*)^{\gamma_i}} \sum_{k=0,2,4,\ldots,\gamma_i} \frac{(-\mu_i\alpha_i w_i^*)^{\gamma_i-k}}{k!(\gamma_i-k)!b_i^{k+1}} \Gamma\left(\frac{k+1}{\gamma_i}\right).
\end{aligned}
\tag{7.42}
$$

With (7.42), we also obtain

$$
\begin{aligned}
&V\left[\left|\frac{X_i}{1+\alpha_i w_i^*} - \mu_i\right|^{\gamma_i}\right] \\
&\quad = E\left[\left|\frac{X_i}{1+\alpha_i w_i^*} - \mu_i\right|^{2\gamma_i}\right] - E^2\left[\left|\frac{X_i}{1+\alpha_i w_i^*} - \mu_i\right|^{\gamma_i}\right] \\
&\quad = \frac{2a_i(2\gamma_i-1)!}{(1+\alpha_i w_i^*)^{2\gamma_i}} \sum_{k=0,2,4,\ldots,2\gamma_i} \frac{(-\mu_i\alpha_i w_i^*)^{2\gamma_i-k}}{k!(2\gamma_i-k)!b_i^{k+1}} \Gamma\left(\frac{k+1}{2\gamma_i}\right) \\
&\qquad - \left[\frac{2a_i(\gamma_i-1)!}{(1+\alpha_i w_i^*)^{\gamma_i}} \sum_{k=0,2,4,\ldots,\gamma_i} \frac{(-\mu_i\alpha_i w_i^*)^{\gamma_i-k}}{k!(\gamma_i-k)!b_i^{k+1}} \Gamma\left(\frac{k+1}{\gamma_i}\right)\right]^2.
\end{aligned}
\tag{7.43}
$$

Note that in evaluating (7.38), (7.39), (7.42) and (7.43) we need to estimate μ_i, σ_i^2 and γ_i from the transform coefficients of the watermarked image. The estimators for μ_i and σ_i^2 are

$$
\hat{\mu}_i = \frac{1}{N_{\mathcal{B}}} \sum_{y\in\mathcal{B}} y \tag{7.44}
$$

and

$$
\hat{\sigma}_i^2 = \frac{1}{N_{\mathcal{B}}-1} \sum_{y\in\mathcal{B}} (y-\hat{\mu}_i)^2, \tag{7.45}
$$

respectively. Similar to (7.18), with non-zero mean, we set

$$\frac{E^2[|X_i - \mu_i|]}{\sigma_i^2} = \frac{\Gamma^2(2/\gamma_i)}{\Gamma(1/\gamma_i)\Gamma(3/\gamma_i)} = s(\gamma_i), \tag{7.46}$$

and use

$$\hat{\gamma_i} = s^{-1}\left(\frac{1}{\hat{\sigma}_i^2}\left[\frac{1}{N_{\mathcal{B}}}\sum_{y\in\mathcal{B}}|y-\hat{\mu}_i|\right]\right) \tag{7.47}$$

as an estimator for γ_i.

For $\gamma_i = 1$, we have the Laplacian model as derived in Chapter 6. Other than this, there is no close-form solution for the mean and variance of $\left|\frac{X_i}{1+\alpha_i w_i^*} - \mu_i\right|^{\gamma_i}$ when γ_i is not an even integer. However, we can estimate the mean and variance of $\left|\frac{X_i}{1+\alpha_i w_i^*} - \mu_i\right|^{\gamma_i}$ in two ways. One way is to first approximate γ_i to the nearest even integer and then use (7.42) and (7.43). Another way is to estimate from the transform coefficients of the image.

7.5 Experimental Results

The zero mean generalized Gaussian model is compared with the Laplacian and Gaussian models using the same experiment setting as in Section 6.4. That is, watermark embedding is done in DWT domain to the coefficients of the high-resolution subbands LH_3, HL_3 and HH_3. Blind detection is used with the variance and shape parameter estimated using (7.17) and (7.19), respectively. The embedding strengths for the different images are similar to that in Table 6.1. Tables 7.1-7.6 show results for watermark robustness under the same standard

image processing operations used in Chapter 6. Results from Chapter 6 for the non zero mean Laplacian and Gaussian models are included for comparison. Overall, the generalized Gaussian model yields a better detection result. The estimators of the shape parameter considered here may also find applications in other areas of science and engineering where the generalized Gaussian distribution is used.

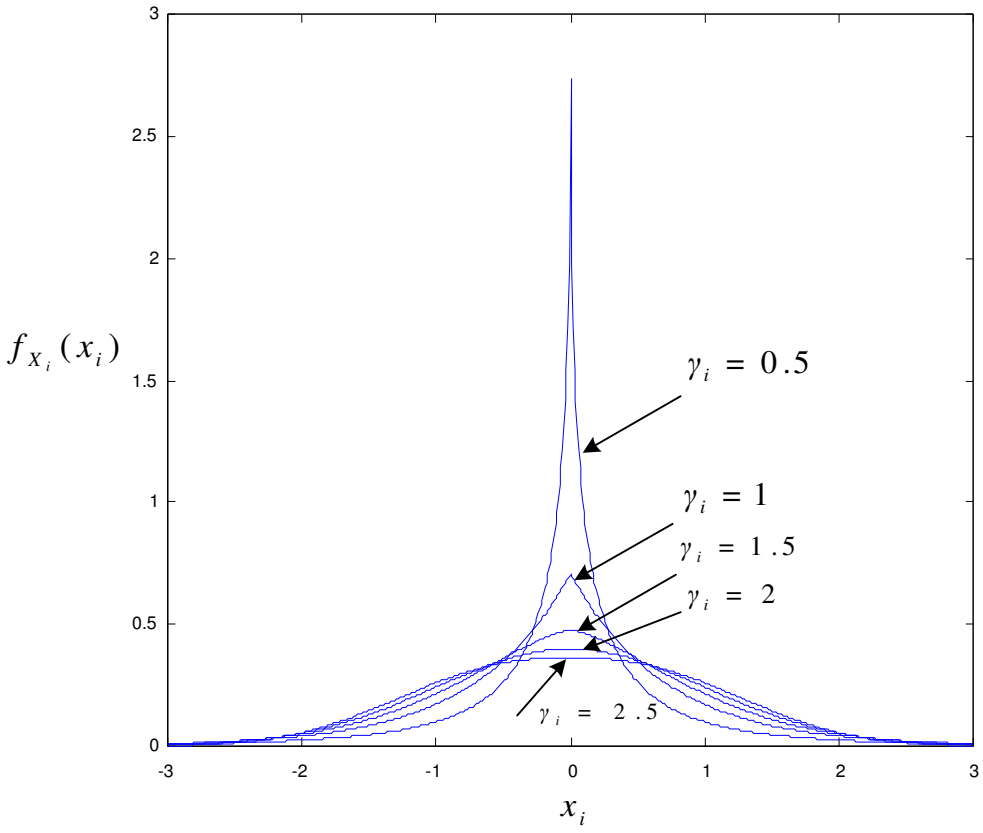

Figure 7.1: Generalized Gaussian PDF with $\mu_i = 0$ and $\sigma_i^2 = 1$.

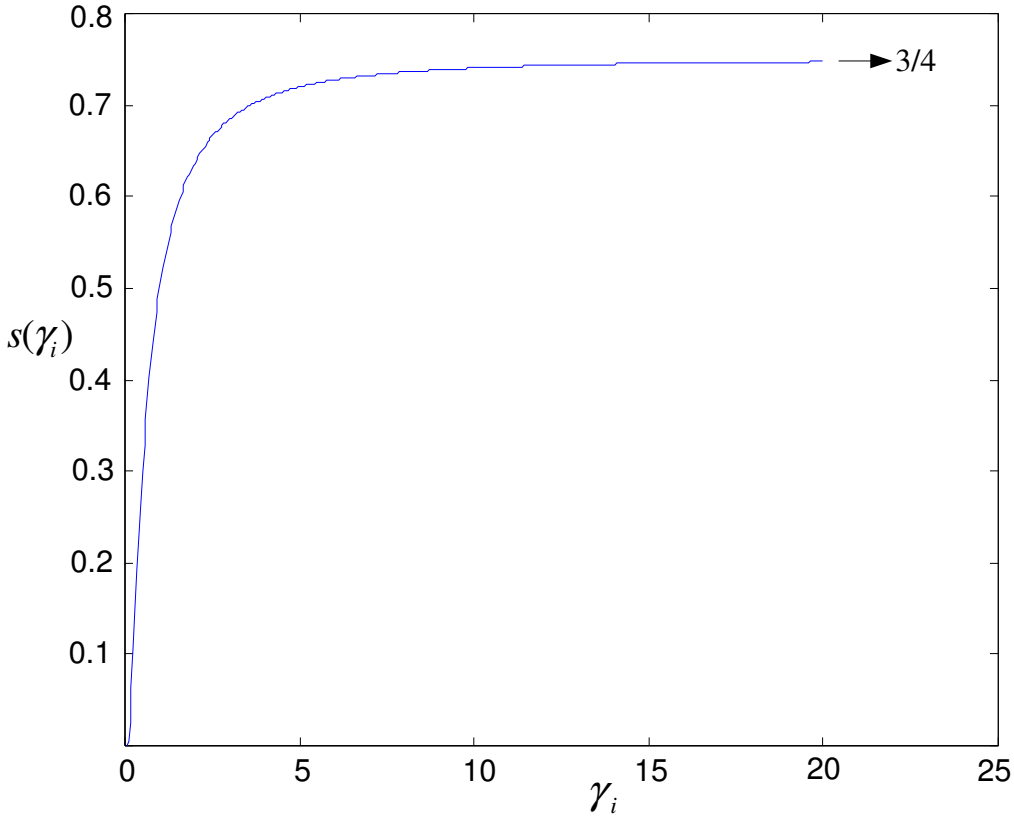

Figure 7.2: Plot of $s(\gamma_i)$ versus γ_i.

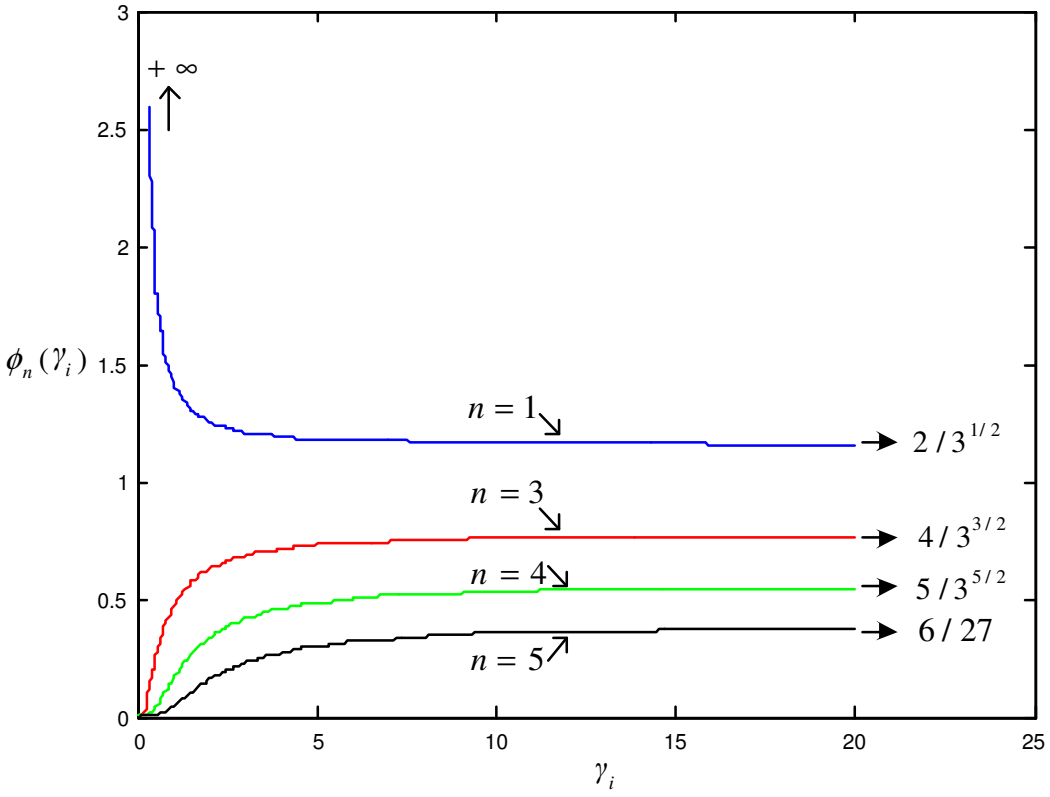

Figure 7.3: Plot of $\phi_n(\gamma_i)$ versus γ_i for $n = 1, 3, 4$, and 5.

Table 7.1: Percentage of successful detections under JPEG compression.

Image	Generalized Gaussian	Laplacian	Gaussian
Harbour	99.80	99.88	99.62
Lena	99.83	99.78	98.30
Fishing boat	99.89	99.82	99.23
Peppers	99.88	99.86	99.01
Barbara	99.97	99.93	99.13
Goldhill	99.93	99.90	98.33
Zelda	99.81	99.78	98.43
LAX	99.89	99.81	99.14

Table 7.2: Percentage of successful detections under low pass filtering.

Image	Generalized Gaussian	Laplacian	Gaussian
Harbour	99.92	99.87	99.39
Lena	99.67	99.45	99.03
Fishing boat	99.89	99.70	99.66
Peppers	99.50	99.56	99.67
Barbara	99.81	99.49	99.21
Goldhill	99.86	99.81	99.48
Zelda	99.46	99.38	99.07
LAX	99.78	99.65	99.55

Table 7.3: Percentage of successful detections under Gaussian noise.

Image	Generalized Gaussian	Laplacian	Gaussian
Harbour	99.80	99.77	99.42
Lena	99.67	99.56	99.10
Fishing boat	99.50	99.39	99.11
Peppers	99.83	99.76	99.12
Barbara	99.77	99.42	99.01
Goldhill	99.49	99.39	99.20
Zelda	99.38	99.19	99.32
LAX	99.31	99.27	99.01

Table 7.4: Percentage of successful detections under speckle noise.

Image	Generalized Gaussian	Laplacian	Gaussian
Harbour	99.85	99.71	99.35
Lena	99.62	99.49	99.06
Fishing boat	99.63	99.52	99.17
Peppers	99.80	99.69	99.22
Barbara	99.90	99.87	99.21
Goldhill	99.77	99.81	99.69
Zelda	99.88	99.72	99.36
LAX	99.74	99.68	99.75

Table 7.5: Percentage of successful detections under salt and pepper noise.

Image	Generalized Gaussian	Laplacian	Gaussian
Harbour	99.85	99.79	99.28
Lena	99.90	99.87	99.73
Fishing boat	99.73	99.67	99.60
Peppers	99.81	99.61	99.40
Barbara	99.87	99.80	99.66
Goldhill	99.82	99.68	99.31
Zelda	99.90	99.85	99.59
LAX	99.66	99.77	99.08

Table 7.6: Percentage of successful detections under cropping.

Image	Generalized Gaussian	Laplacian	Gaussian
Harbour	99.90	99.78	99.23
Lena	99.93	99.82	99.50
Fishing boat	99.86	99.87	99.23
Peppers	99.66	99.35	99.31
Barbara	99.79	99.56	99.60
Goldhill	99.83	99.81	99.34
Zelda	99.70	99.56	99.31
LAX	99.59	99.34	99.22

Chapter 8

LR Detector Based on Generalized Gamma Model

The generalized gamma distribution is another general distribution. The Weibull, gamma, exponential distributions are some of its special cases. Generalized gamma distribution is used to characterize positive random variables. In this chapter, we discuss our work in extending the Weibull model of Barni et al [3] to generalized gamma model for LR detection of watermarks. This includes giving the LR decision rule and deriving a closed-form LR decision threshold. Numerical simulations are used to draw comparison with the Weibull model.

8.1 LR Detection Rule

If the transform coefficient X_i is modeled by a generalized gamma PDF [48], then its PDF is expressed as

$$f_{X_i}(x_i) = \frac{p_i}{a_i^{p_i\nu_i}} \frac{x_i^{p_i\nu_i - 1}}{\Gamma(\nu_i)} \exp\left[-\left(\frac{x_i}{a_i} \right)^{p_i} \right], \quad x_i > 0, \tag{8.1}$$

with positive shape p_i, scale a_i, and power ν_i parameters. The exponential ($p_i = \nu_i = 1$), Weibull ($\nu_i = 1$) and gamma ($p_i = 1$) are included as special cases. Figure 8.1 shows the shape of the generalized gamma PDF.

Substituting (8.1) in (4.16), we obtain the LR decision rule as

$$\begin{aligned} z(\mathbf{y}) &= \sum_{i=1}^{N} \left\{ \ln \left[\frac{p_i}{a_i^{p_i\nu_i}\Gamma(\nu_i)} \left(\frac{y_i}{1+\alpha_i w_i^*} \right)^{p_i\nu_i - 1} \exp\left[- \left(\frac{y_i}{(1+\alpha_i w_i^*)a_i} \right)^{p_i} \right] \right] \right. \\ &\qquad \left. - \ln \left[\frac{p_i}{a_i^{p_i\nu_i}\Gamma(\nu_i)} y_i^{p_i\nu_i - 1} \exp\left[- \left(\frac{y_i}{a_i} \right)^{p_i} \right] \right] \right\} \\ &= \sum_{i=1}^{N} y_i^{p_i} \frac{(1+\alpha_i w_i^*)^{p_i} - 1}{a_i^{p_i}(1+\alpha_i w_i^*)^{p_i}} > \lambda_{ga}, \end{aligned} \tag{8.2}$$

where λ_{ga} is the LR decision threshold under generalized gamma model.

8.2 LR Decision Threshold

In view of (4.22), the mean and variance of

$$z(\mathbf{X}) = \sum_{i=1}^{N} X_i^{p_i} \frac{(1+\alpha_i w_i^*)^{p_i} - 1}{a_i^{p_i}(1+\alpha_i w_i^*)^{p_i}} \tag{8.3}$$

are required to obtain λ_{ga}. Equivalently, we need to find $E[X_i^{p_i}]$ and $V[X_i^{p_i}]$. For a real constant c,

$$\begin{aligned} E[X_i^c] &= \int_0^{+\infty} x_i^c \frac{p_i}{a_i^{p_i\nu_i}} \frac{x_i^{p_i\nu_i - 1}}{\Gamma(\nu_i)} \exp\left[- \left(\frac{x_i}{a_i} \right)^{p_i} \right] dx_i \\ &= \frac{a_i^c}{\Gamma(\nu_i)} \int_0^{+\infty} t^{\nu_i + c/p_i - 1} \exp(-t) dt, \quad t = \left(\frac{x_i}{a} \right)^{p_i} \\ &= a_i^c \frac{\Gamma(\nu_i + c p_i^{-1})}{\Gamma(\nu_i)}. \end{aligned} \tag{8.4}$$

With (8.4), it is straightforward to obtain

$$\mu_{z(\mathbf{X})} = \sum_{i=1}^{N} E[X_i^{p_i}] \frac{(1+\alpha_i w_i^*)^{p_i} - 1}{a_i^{p_i}(1+\alpha_i w_i^*)^{p_i}}$$

$$
\begin{aligned}
&= \sum_{i=1}^{N} \left\{ a_i^{p_i} \frac{\Gamma(\nu_i+1)}{\Gamma(\nu_i)} \right\} \frac{(1+\alpha_i w_i^*)^{p_i}-1}{a_i^{p_i}(1+\alpha_i w_i^*)^{p_i}} \\
&= \sum_{i=1}^{N} \frac{\Gamma(\nu_i+1)[(1+\alpha_i w_i^*)^{p_i}-1]}{\Gamma(\nu_i)(1+\alpha_i w_i^*)^{p_i}} \qquad (8.5)
\end{aligned}
$$

and

$$
\begin{aligned}
\sigma^2_{z(\mathbf{X})} &= \sum_{i=1}^{N} V[X_i^{p_i}] \frac{[(1+\alpha_i w_i^*)^{p_i}-1]^2}{a_i^{2p_i}(1+\alpha_i w_i^*)^{2p_i}} \\
&= \sum_{i=1}^{N} \left\{ E[X_i^{2p_i}] - E^2[X_i^{p_i}] \right\} \frac{[(1+\alpha_i w_i^*)^{p_i}-1]^2}{a_i^{2p_i}(1+\alpha_i w_i^*)^{2p_i}} \\
&= \sum_{i=1}^{N} \left\{ a_i^{2p_i} \frac{\Gamma(\nu_i+2)}{\Gamma(\nu_i)} - a_i^{2p_i} \frac{\Gamma^2(\nu_i+1)}{\Gamma^2(\nu_i)} \right\} \frac{[(1+\alpha_i w_i^*)^{p_i}-1]^2}{a_i^{2p_i}(1+\alpha_i w_i^*)^{2p_i}} \\
&= \sum_{i=1}^{N} \frac{[\Gamma(\nu_i+2)\Gamma(\nu_i) - \Gamma^2(\nu_i+1)][(1+\alpha_i w_i^*)^{p_i}-1]^2}{\Gamma^2(\nu_i)(1+\alpha_i w_i^*)^{2p_i}}. \qquad (8.6)
\end{aligned}
$$

Thus, the closed-form of λ_{ga} is given as

$$
\begin{aligned}
\lambda_{ga} &= \mathrm{erfc}^{-1}(2P_{FA}^*) \left[2 \sum_{i=1}^{N} \frac{[\Gamma(\nu_i+2)\Gamma(\nu_i) - \Gamma^2(\nu_i+1)][(1+\alpha_i w_i^*)^{p_i}-1]^2}{\Gamma^2(\nu_i)(1+\alpha_i w_i^*)^{2p_i}} \right]^{1/2} \\
&\quad + \sum_{i=1}^{N} \frac{\Gamma(\nu_i+1)[(1+\alpha_i w_i^*)^{p_i}-1]}{\Gamma(\nu_i)(1+\alpha_i w_i^*)^{p_i}}. \qquad (8.7)
\end{aligned}
$$

For a fixed P_{FA}^*, λ_{ga} depends on ν_i, α_i and w_i^*.

8.3 Weibull Model

When $\nu_i = 1$, then X_i is a Weibull RV with PDF

$$
f_{X_i}(x_i) = \frac{p_i}{a_i^{p_i}} x_i^{p_i-1} \exp\left[-\left(\frac{x_i}{a_i} \right)^{p_i} \right], \quad x_i > 0. \qquad (8.8)
$$

The decision rule in (8.2) remains the same as it is independent of ν_i. However, the decision threshold λ_{ga} becomes

$$\lambda_{ga} = \text{erfc}^{-1}(2P^*_{FA}) \left[2\sum_{i=1}^{N} \frac{[(1+\alpha_i w_i^*)^{p_i} - 1]^2}{(1+\alpha_i w_i^*)^{2p_i}} \right]^{1/2} + \sum_{i=1}^{N} \frac{[(1+\alpha_i w_i^*)^{p_i} - 1]}{(1+\alpha_i w_i^*)^{p_i}}, \tag{8.9}$$

which is also independent of ν_i. It is simpler to compute (8.7) than (8.9). However, the present of ν_i in (8.7) provides greater flexibility in adjusting λ_{ga} according to different images for better watermark detection result.

8.4 Parameter Estimation

As in [3], watermark is embedded in the magnitude of a set of DFT coefficients of an image. Embedding is done to different regions identified in the magnitude of DFT spectrum. Let $\mathcal{B}$ be the region containing x_i and having $N_{\mathcal{B}}$ coefficients. All coefficients in $\mathcal{B}$ are assumed to be identically distributed.

For blind detection, all parameters are estimated from the magnitude of the DFT coefficients of the watermarked image. The unbiased estimators of μ_i and σ_i^2 are

$$\hat{\mu}_i = \frac{1}{N_{\mathcal{B}}} \sum_{y \in \mathcal{B}} y \tag{8.10}$$

and

$$\hat{\sigma}_i^2 = \frac{1}{N_{\mathcal{B}} - 1} \sum_{y \in \mathcal{B}} (y - \hat{\mu}_i)^2, \tag{8.11}$$

respectively, where y is the corresponding magnitude of the DFT coefficient of the watermarked image in $\mathcal{B}$. Estimating all three parameters, a_i, p_i and ν_i, is rather difficult [15, 49]. Here, we consider fixing either ν_i or p_i and then estimate the other two parameters.

Using (8.4), we obtain the mean and variance of X_i as

$$\mu_i = E[X_i] = a_i \frac{\Gamma(\nu_i + p_i^{-1})}{\Gamma(\nu_i)} \tag{8.12}$$

and

$$\begin{aligned} \sigma_i^2 &= V[X_i] \\ &= E[X_i^2] - E^2[X_i] \\ &= a_i^2 \frac{\Gamma(\nu_i + 2p_i^{-1})}{\Gamma(\nu_i)} - a_i^2 \frac{\Gamma^2(\nu_i + p_i^{-1})}{\Gamma^2(\nu_i)}, \end{aligned} \tag{8.13}$$

respectively. By manipulating (8.12) and (8.13), the expresssion

$$\frac{\Gamma^2(\nu_i + p_i^{-1})}{\Gamma(\nu_i + 2p_i^{-1})\Gamma(\nu_i)} = \frac{\mu_i^2}{\sigma_i^2 + \mu_i^2} \tag{8.14}$$

follows readily.

If ν_i is fixed, say $\nu_i = \nu_0$, then the left-hand side of (8.14) is solely a function of p_i. We define

$$\varphi(p_i) = \frac{\Gamma^2(\nu_0 + p_i^{-1})}{\Gamma(\nu_0 + 2p_i^{-1})\Gamma(\nu_0)} \tag{8.15}$$

and propose estimating p_i and a_i as

$$\hat{p}_i = \varphi^{-1}\left(\frac{\hat{\mu}_i^2}{\hat{\sigma}_i^2 + \hat{\mu}_i^2}\right) \tag{8.16}$$

and

$$\hat{a}_i = \hat{\mu}_i \frac{\Gamma(\nu_0)}{\Gamma(\nu_0 + \hat{p}_i^{-1})}, \tag{8.17}$$

respectively.

Similarly, if p_i is fixed, say $p_i = p_0$, then the left-hand side of (8.14) is solely a function of ν_i. We then define

$$\psi(\nu_i) = \frac{\Gamma^2(\nu_i + p_0^{-1})}{\Gamma(\nu_i + 2p_0^{-1})\Gamma(\nu_i)}, \tag{8.18}$$

and propose estimating ν_i and a_i as

$$\hat{\nu}_i = \psi^{-1}\left(\frac{\hat{\mu}_i^2}{\hat{\sigma}_i^2 + \hat{\mu}_i^2}\right) \tag{8.19}$$

and

$$\hat{a}_i = \hat{\mu}_i \frac{\Gamma(\hat{\nu}_i)}{\Gamma(\hat{\nu}_i + p_0^{-1})}, \tag{8.20}$$

respectively. As in the estimation of the generalized Gaussian shape parameter in Chapter 6, we solve (8.16) and (8.19) by approximating the inverse functions φ^{-1} and ψ^{-1} using function interpolation methods [21]. Both ϕ and ψ are constructed in such a way that their range are finite as shown in Lemma 8.1 and Lemma 8.2.

Lemma 8.1 *The function φ as defined in (8.15) is a strictly increasing function with*

$$\lim_{p_i \to +\infty} \varphi(p_i) = 1 \tag{8.21}$$

and

$$\lim_{p_i \to 0^+} \varphi(p_i) = 0. \tag{8.22}$$

Proof: Clearly, (8.21) follows from definition of φ in (8.15). Using the Weierstrass formula (2.34), we can express $\varphi(p_i)$ as

$$\begin{aligned}
\varphi(p_i) &= \frac{\left[\frac{(p_i\nu_0+1)}{p_i}e^{\frac{\zeta(p_i\nu_0+1)}{p_i}}\prod_{k=1}^{+\infty}\left(1+\frac{p_i\nu_0+1}{p_ik}\right)e^{-\frac{(p_i\nu_0+1)}{p_ik}}\right]^{-2}}{\left[\frac{(p_i\nu_0+2)}{p_i}e^{\frac{\zeta(p_i\nu_0+2)}{p_i}}\prod_{k=1}^{+\infty}\left(1+\frac{p_i\nu_0+2}{p_ik}\right)e^{-\frac{(p_i\nu_0+2)}{p_ik}}\right]^{-1}} \\
&\qquad \times \frac{1}{\left[\nu_0 e^{\zeta\nu_0}\prod_{k=1}^{+\infty}\left(1+\frac{\nu_0}{k}\right)e^{-\frac{\nu_0}{k}}\right]^{-1}} \\
&= \frac{p_i\nu_0(p_i\nu_0+2)}{(p_i\nu_0+1)^2}\prod_{k=1}^{+\infty}\frac{p_i(p_ik+p_i\nu_0+2)(k+\nu_0)}{(p_ik+p_i\nu_0+1)^2} \\
&= \prod_{k=0}^{+\infty}\frac{p_i(p_ik+p_i\nu_0+2)(k+\nu_0)}{(p_ik+p_i\nu_0+1)^2}. \qquad (8.23)
\end{aligned}$$

Letting $p_i \to 0^+$ yields (8.22). Moreover,

$$\begin{aligned}
\frac{d}{dp_i}\frac{p_i(p_ik+p_i\nu_0+2)}{(p_ik+p_i\nu_0+1)^2} &= \frac{(p_ik+p_i\nu_0+1)^2[2(p_ik+p_i\nu_0+1)]}{(p_ik+p_i\nu_0+1)^4} \\
&\qquad - \frac{p_i(p_ik+p_i\nu_0+2)[2(k+\nu_0)(p_ik+p_i\nu_0+1)]}{(p_ik+p_i\nu_0+1)^4} \\
&= \frac{2(p_ik+p_i\nu_0+1)}{(p_ik+p_i\nu_0+1)^4} > 0, \qquad (8.24)
\end{aligned}$$

for $p_i > 0$. This means each term in the infinite products in (8.23) is strictly increasing. Thus, φ is also strictly increasing, and consequently has finite range.

Q.E.D

Lemma 8.2 *The function ψ as defined in (8.18) is a strictly increasing function with*

$$\lim_{\nu_i \to +\infty} \psi(\nu_i) = 1 \qquad (8.25)$$

and

$$\lim_{\nu_i \to 0^+} \psi(\nu_i) = 0. \tag{8.26}$$

Proof: Since $\lim_{\nu_i \to 0^+} \Gamma(\nu_i) = +\infty$, (8.26) follows for the definition of ψ. Next, note that φ and ψ are similar in form but differ in the fixed variable. Thus, in view of (8.23), we can express $\psi(\nu_i)$ as

$$\psi(\nu_i) \;=\; \prod_{k=0}^{+\infty} \frac{p_0(p_0k + p_0\nu_i + 2)(k + \nu_i)}{(p_0k + p_0\nu_i + 1)^2}. \tag{8.27}$$

With some rearrangements, we obtain

$$\psi(\nu_i) \;=\; \prod_{k=0}^{+\infty} \frac{p_0(p_0k\nu_i^{-1} + p_0 + 2\nu_i^{-1})(k\nu_i^{-1} + 1)}{(p_0k\nu_i^{-1} + p_0 + \nu_i^{-1})^2}, \tag{8.28}$$

and then letting $\nu_i \to +\infty$ yields (8.25). We conclude that ψ is strictly increasing since in (8.27),

$$\frac{d}{d\nu_i} \frac{p_0(p_0k + p_0\nu_i + 2)(k + \nu_i)}{(p_0k + p_0\nu_i + 1)^2} = \frac{2p_0}{(p_0k + p_0\nu_i + 1)^3} > 0, \tag{8.29}$$

for $\nu_i > 0$. This again implies that the range of ψ is finite.

Q.E.D

Figure 8.2 and 8.3 show the plot of φ and ψ, respectively.

8.5 Experimental Results

Using the same set of 512×512 test images given in Figure 3.1, we explore three special cases of the generalized gamma model, (i) $\nu_i = 1$ (ii) $\nu_i = 1.1$, and (iii)

$p_i = 1$. In (i), we have the Weibull model [3]. In [3], a_i and p_i are estimated using a maximum likelihood scheme which requires the use of iterative method like the Newton-Raphson method. In this regard, our estimators are considered simpler. In (iii), we have the gamma model. All these models are selected because the plot of their PDFs closely resembles the shape of the histograms derived from the subsets of $x_1, x_2, \ldots, x_N$.

In [3], the DFT magnitude spectrum is divided into 16 regions. Here, for simplicity of testing, we consider only two regions, as shown in Figure 8.4 as two squares at the upper half of the DFT matrix. Each region has 2,500 identically distributed coefficients. The watermark is embedded into these coefficients and then duplicated to the corresponding coefficients in the two regions at the lower half of the DFT matrix. This is done to preserve the symmetry property of the DFT magnitude spectrum. A constant embedding strength $\alpha = 0.3$ is used for all the coefficients.

The robustness of the watermark is tested under the same standard image processing operations as in Chapters 6 and 7, using the procedure given in Section 4.3. The results are shown in Tables 8.1-8.6. Generally, our experiments show that the generalized gamma model gives rise to better watermark detection. Our choice of $\nu_i = 1.1$ is based on trial on error. Further research is required to decide the optimum choice of ν_i or p_i so that the remaining two parameters can be estimated.

The estimators proposed here are also found to be better than the moment

estimators in [16, 17], where comparisons are made in terms of mean square error. Experiments are done by simulating generalized gamma RVs using (2.28), and then estimating the parameters.

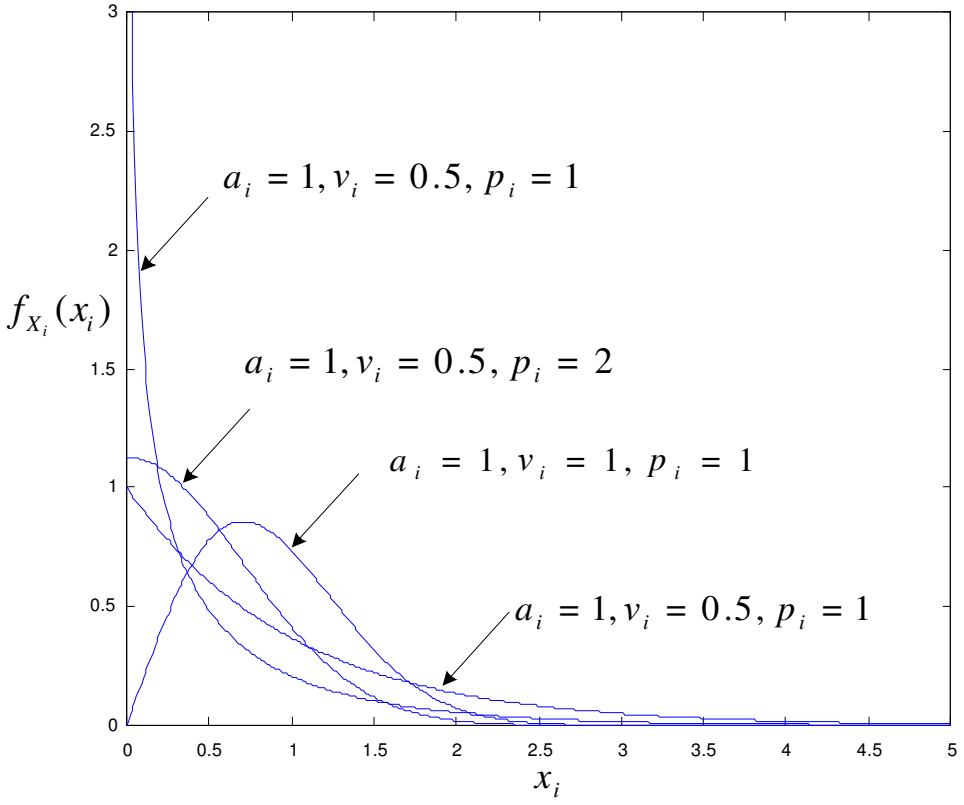

Figure 8.1: Generalized Gamma PDF.

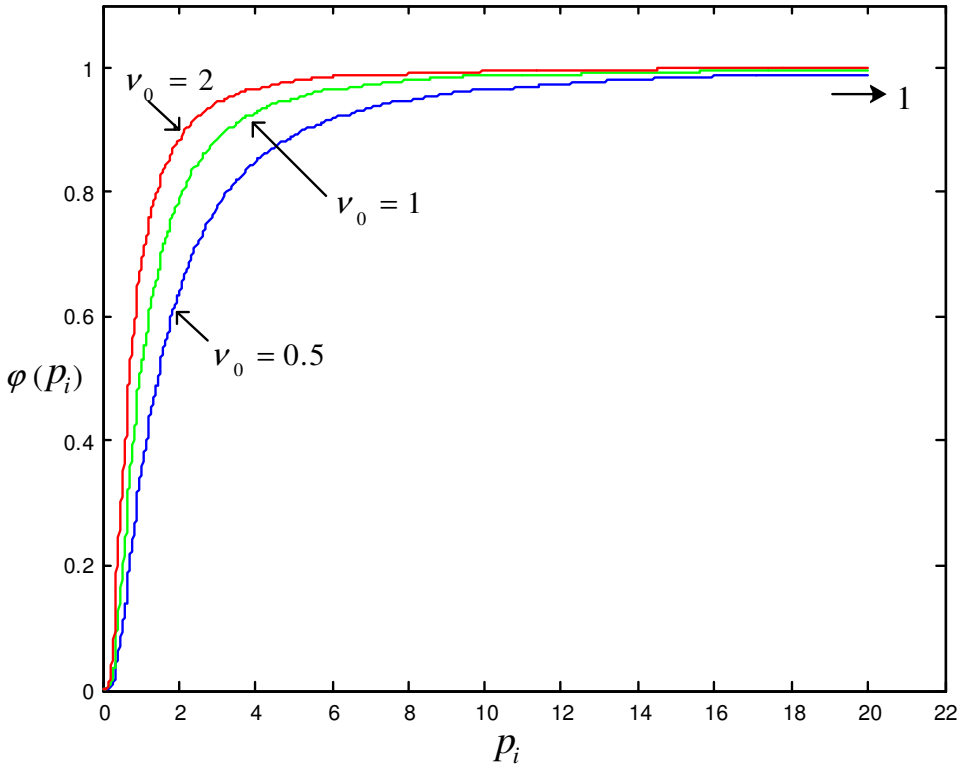

Figure 8.2: Plot of $\varphi(p_i)$ versus p_i for $\nu_0 = 0.5, 1$ and 2.

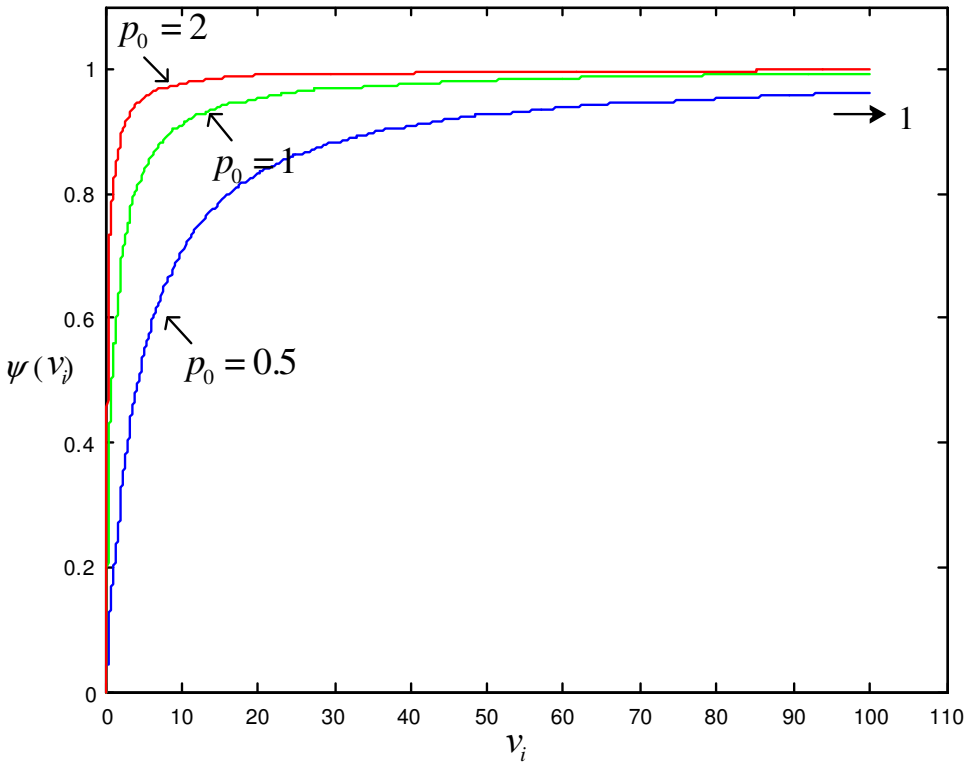

Figure 8.3: Plot of $\psi(\nu_i)$ versus ν_i for $p_0 = 0.5, 1$ and 2.

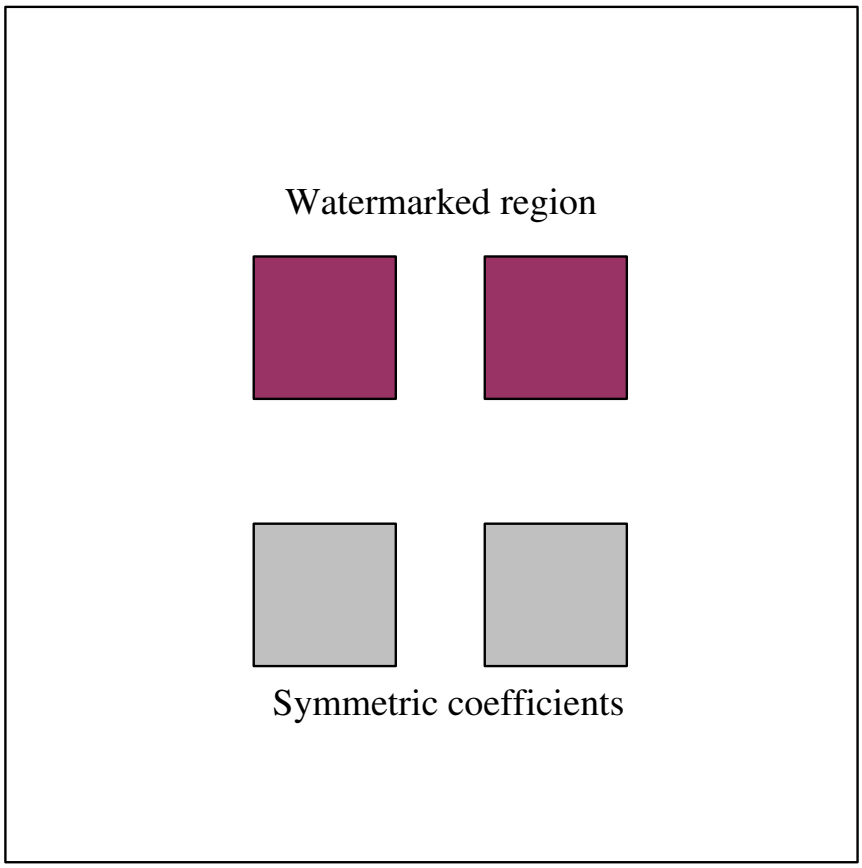

Figure 8.4: Watermark region in DFT(magnitude) matrix.

Table 8.1: Percentage of successful detections under JPEG compression.

Image	Weibull	Generalized Gamma ($\nu_i = 1.1$)	Gamma
Harbour	99.63	99.78	99.59
Lena	92.44	97.36	91.47
Fishing boat	93.79	98.33	89.79
Peppers	93.69	98.37	90.72
Barbara	95.34	96.99	94.59
Goldhill	99.84	99.86	99.78
Zelda	99.21	99.08	99.14
LAX	97.34	98.69	97.19

Table 8.2: Percentage of successful detections under low pass filtering.

Image	Weibull	Generalized Gamma ($\nu_i = 1.1$)	Gamma
Harbour	99.62	99.78	99.57
Lena	99.01	99.32	98.89
Fishing boat	99.01	99.10	99.08
Peppers	99.83	99.91	99.74
Barbara	99.32	99.63	99.19
Goldhill	99.36	99.41	99.28
Zelda	98.76	99.11	98.99
LAX	99.20	99.39	99.13

Table 8.3: Percentage of successful detections under Gaussian noise.

Image	Weibull	Generalized Gamma ($\nu_i = 1.1$)	Gamma
Harbour	99.88	99.90	99.72
Lena	99.71	99.83	99.62
Fishing boat	97.63	98.37	98.01
Peppers	99.61	99.78	99.43
Barbara	99.94	99.56	99.27
Goldhill	99.78	99.89	99.38
Zelda	99.47	99.46	99.13
LAX	99.67	99.73	99.49

Table 8.4: Percentage of successful detections under speckle noise.

Image	Weibull	Generalized Gamma ($\nu_i = 1.1$)	Gamma
Harbour	99.93	99.99	99.77
Lena	99.95	99.98	99.85
Fishing boat	99.78	99.95	99.96
Peppers	99.85	99.90	99.89
Barbara	99.97	99.95	99.88
Goldhill	99.86	99.97	99.36
Zelda	99.67	99.79	99.69
LAX	99.79	99.81	99.69

Table 8.5: Percentage of successful detections under salt and pepper noise.

Image	Weibull	Generalized Gamma ($\nu_i = 1.1$)	Gamma
Harbour	99.12	99.18	99.10
Lena	99.73	99.93	99.68
Fishing boat	99.67	99.59	99.61
Peppers	99.69	99.72	99.65
Barbara	99.77	99.89	99.85
Goldhill	99.88	99.92	99.79
Zelda	99.94	99.96	99.82
LAX	99.69	99.68	99.63

Table 8.6: Percentage of successful detections under cropping.

Image	Weibull	Generalized Gamma ($\nu_i = 1.1$)	Gamma
Harbour	99.97	99.98	99.69
Lena	100.00	100.00	100.00
Fishing boat	99.93	99.12	99.89
Peppers	100.00	100.00	100.00
Barbara	99.81	99.07	99.19
Goldhill	99.67	99.79	99.62
Zelda	99.83	99.86	99.79
LAX	99.88	99.91	99.80

Chapter 9

MAP Detection of Watermark

MAP detection is considered simpler than LR detection as it does not require a decision threshold. In this chapter, we consider an MAP detector formulated under the generalized Gaussian distribution in the DWT domain.

9.1 MAP Detector

We use a slightly different set of notations from the previous chapters to describe the MAP detector.

Let $W = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_K\}$ be a set of K available watermarks. Each watermark is represented as a vector of N elements, $\mathbf{w}_k = [w_{k,1}, w_{k,2}, \ldots, w_{k,N}]^T$, for $k = 1, 2, \ldots, K$. A watermark from W, say $\mathbf{w}_j$, is selected and embedded to $\mathbf{x} = [x_1, x_2, \ldots, x_N]^T$, the DWT coefficients vector of the original image. The corresponding DWT coefficients vector of the watermarked image is denoted as $\mathbf{y} = [y_1, y_2, \ldots, y_N]^T$. The multiplicative embedding scheme is used with MAP

detector [2],

$$y_i = x_i + \alpha x_i w_{j,i}, \quad i = 1, 2, \ldots, N, \tag{9.1}$$

where α is a fixed embedding strength. The value of α is set to be much smaller than 1.

The decision rule of the MAP detector is based on comparing the a posteriori probabilities

$$P(\mathbf{w}_k|\mathbf{y}), \quad k = 1, 2, \ldots, K. \tag{9.2}$$

The watermark that corresponds to the maximum of this set of a posteriori probabilities is said to be the embedded watermark. It can be shown that this rule maximizes the probability of a correct decision and, hence the probability of error is minimized [43].

Using Bayes' rule, the a posteriori probabilities can be written as

$$P(\mathbf{w}_k|\mathbf{y}) = \frac{P(\mathbf{y}|\mathbf{w}_k)P(\mathbf{w}_k)}{P(\mathbf{y})}, \quad k = 1, 2, \ldots, K, \tag{9.3}$$

where $P(\mathbf{y}|\mathbf{w}_k)$ is the conditional probability of $\mathbf{y}$ given $\mathbf{w}_k$, and $P(\mathbf{w}_k)$ is the a priori probability of the kth watermark being embedded. Since the denominator $P(\mathbf{y})$ can be expanded as

$$P(\mathbf{y}) = \sum_{k=1}^{K} P(\mathbf{y}|\mathbf{w}_k)P(\mathbf{w}_k), \tag{9.4}$$

it thus follows from (9.3) that the computation of the a posteriori probabilities require the knowledge of $P(\mathbf{y}|\mathbf{w}_k)$ and $P(\mathbf{w}_k)$ for $k = 1, 2, \ldots, K$.

It is usually assumed that $P(\mathbf{w}_k) = 1/K$ for $k = 1, 2, \ldots, K$, i.e., the K watermarks are equally probable. Moreover, the denominator $P(\mathbf{y})$ does not depend on $\mathbf{w}_k$. As a result, the decision rule reduces to finding the watermark that corresponds to

$$\max_{k=1,2,\ldots,K} P(\mathbf{y}|\mathbf{w}_k). \tag{9.5}$$

Further, we also assume that the coefficients of DWT subbands are independent [2, 47]. Under this assumption, the elements of $\mathbf{y}$ are conditionally independent, and we can express (9.5) as

$$\max_{k=1,2,\ldots,K} \prod_{i=1}^{N} P(y_i|w_{k,i}). \tag{9.6}$$

We view x_i and y_i as the realizations of the random variables X_i and Y_i. If $f_{X_i}(x_i)$ is the PDF of X_i for $i = 1, 2, \ldots, N$, then (9.6) can be written as

$$\max_{k=1,2,\ldots,K} \prod_{i=1}^{N} \frac{1}{1+\alpha w_{k,i}} f_{X_i}\left(\frac{y_i}{1+\alpha w_{k,i}}\right). \tag{9.7}$$

The natural logarithm of (9.7) is usually used,

$$\max_{k=1,2,\ldots,K} \sum_{i=1}^{N} \left\{ \ln\left(f_{X_i}\left(\frac{y_i}{1+\alpha w_{k,i}}\right) - \ln(1+\alpha w_{k,i}) \right) \right\}. \tag{9.8}$$

Here, we note that the choice of the watermark elements must be such that

$$1 + \alpha w_{k,i} > 0, \tag{9.9}$$

so that $y_i/(1+\alpha w_{k,i})$ and the natural logarithm are well defined. We refer to the maximum value in (9.8) as the MAP peak.

In applications where K is not too large, it is feasible to identify the embedded watermark from (9.8). Moreover, it eliminates the need for a decision threshold, and therefore should result in a more accurate detection.

9.2 Generalized Gaussian Model

We propose modeling X_i as a generalized Gaussian RV. Then, as given in (7.34), the PDF of X_i is

$$f_{X_i}(x_i) = a_i \exp(-b_i^{\gamma_i}|x_i - \mu_i|^{\gamma_i}), \tag{9.10}$$

where μ_i, σ_i and γ_i are the mean, variance and shape parameter of the distribution, respectively. The positive constants a_i and b_i are given as

$$a_i = \frac{b_i \gamma_i}{2\Gamma(1/\gamma_i)} \tag{9.11}$$

and

$$b_i = \frac{1}{\sigma_i}\sqrt{\frac{\Gamma(3/\gamma_i)}{\Gamma(1/\gamma_i)}}, \tag{9.12}$$

respectively. Substituting (9.10) in (9.8), we obtain

$$\max_{k=1,2,\ldots,K} \left\{ \sum_{i=1}^{N} \ln a_i - b_i^{\gamma_i} \left| \frac{y_i}{1+\alpha w_{k,i}} - \mu_i \right|^{\gamma_i} - \ln(1+\alpha w_{k,i}) \right\}. \tag{9.13}$$

For blind detection, in computing (9.13), we need to estimate μ_i, σ_i^2 and γ_i from the watermarked image. We can use estimators similar to those discussed in Section 7.4. That is,

$$\hat{\mu}_i = \frac{1}{N_{\mathcal{B}}} \sum_{y \in \mathcal{B}} y, \tag{9.14}$$

$$\hat{\sigma}_i^2 = \frac{1}{N_{\mathcal{B}} - 1} \sum_{y \in \mathcal{B}} (y - \hat{\mu}_i)^2, \tag{9.15}$$

and

$$\hat{\gamma}_i = s^{-1} \left(\frac{1}{\hat{\sigma}_i^2} \left[\frac{1}{N_{\mathcal{B}}} \sum_{y \in \mathcal{B}} |y - \hat{\mu}_i| \right] \right), \tag{9.16}$$

where $\mathcal{B}$ is the DWT subband containing X_i with $N_{\mathcal{B}}$ coefficients.

9.3 Correlation Detector

If correlation detection is used, then the watermark $\mathbf{w}_j$ is embedded as [2]

$$y_i = x_i + \alpha |x_i| w_{j,i}, \quad i = 1, 2, \ldots, N. \tag{9.17}$$

The correlation between $\mathbf{y}$ and the watermarks in W is defined as

$$\text{cor}(\mathbf{y}, \mathbf{w}) = \frac{1}{N} \sum_{i=1}^{N} y_i w_{k,i}, \quad k = 1, 2, \ldots, K. \tag{9.18}$$

In the absent of distortion, the embedded watermark is the one that corresponds to

$$\max_{k=1,2,\ldots,K} \text{cor}(\mathbf{y}, \mathbf{w}_k). \tag{9.19}$$

9.4 Experiment Results

For the experiments here, the same 512×512 grayscale images in Figure 3.1 are used. Similarly a Daubechies filter is used to obtain the DWT of these images. To ensure that (9.9) is satisfied, the set W contains watermarks with uniformly

distributed components in $[-1,\ 1]$. We set $K = 1000$, i.e., W contains 1,000 watermarks. One of the watermarks from W is selected and embedded in the DWT coefficients of the images. This is done using both (9.1) and (9.17) to produce two set of watermarked images. Embedding is done in all the coefficients in LH_3, HL_3 and HH_3 subbands. The embedding strengths for the images are as given in Table 6.1.

Standard image processing operations are simulated and applied to the watermarked images to distort them. The proposed detector, MAP detector under Laplacian modeling [2], and correlation detector are then used to identify the embedded watermark for each of the distorted images. For example, in Figure 9.1, we see the plot of the MAP detector response for watermarked image 'Barbara' that is low pass filtered using a 4×4 spatial filter. The MAP peak at position 388 belongs to the embedded watermark.

In comparing the performance of the detectors, each of the experiment is performed over 10,000 trials. For all the detectors, the percentage of successful detections from all these trials are recorded for all images. Table 9.1 shows results for watermarked images compressed by JPEG with a 5% quality factor. In Table 9.2, the watermarked images are low pass filtered using a 4×4 spatial filter. In Table 9.3, the watermarked images are first corrupted by salt and pepper noise that covers 30% of the pixels, and then smoothened by a 3×3 median filter. Lastly, in Table 9.4, each watermarked image is cropped to retain only 400×400 pixels at the center. The missing portion is replaced by zero pixels so that the

size of each image remains at 512×512.

Our experiment results reveal that generally the proposed detector is more effective when (i) the watermarked images are low pass filtered, and (ii) when they are corrupted by salt and pepper noise and then median filtered. The detectors have about the same performance when the watermarked images are JPEG compressed or cropped. In [34], watermark robustness under cropping can be improved by extending the embedding to LH_2, HL_2 and HH_2 subbands.

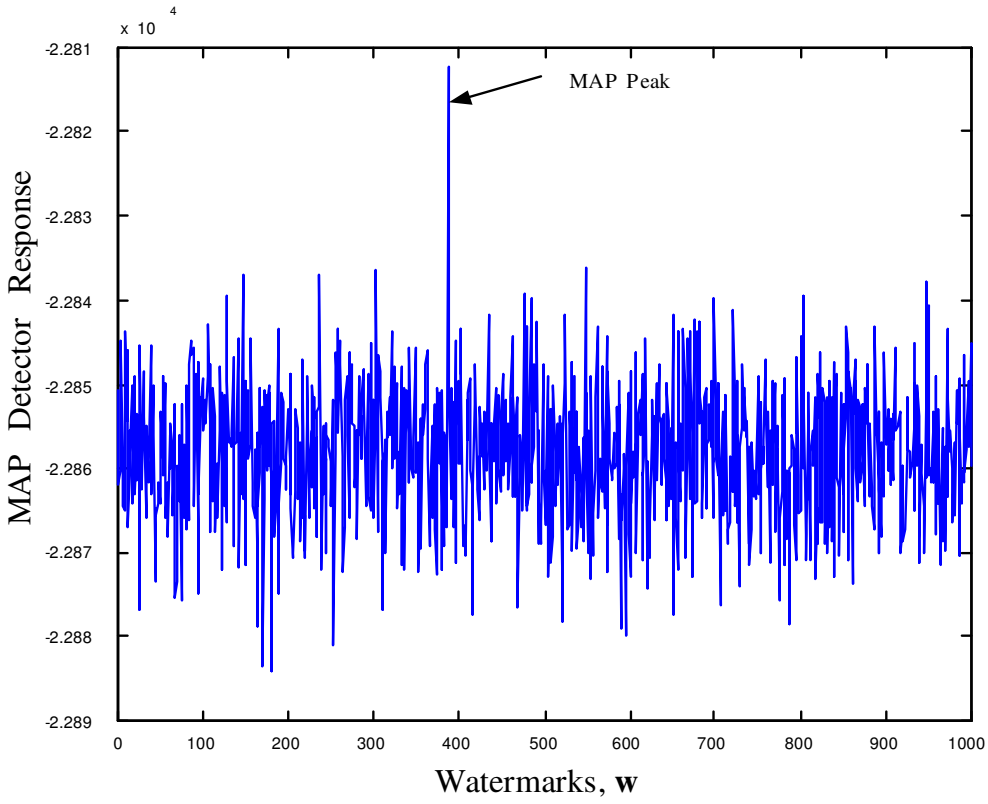

Figure 9.1: Response of MAP detector to 1,000 watermarks for watermarked image 'Barbara' after low pass filtering.

Table 9.1: Percentage of successful detections under JPEG compression.

Image	MAP (Gen. Gaussian)	MAP (Laplacian)	Correlation
Harbour	100.00	100.00	99.86
Lena	99.10	98.25	100.00
Fishing boat	99.83	99.87	99.96
Peppers	100.00	99.01	100.00
Barbara	99.64	99.65	99.50
Goldhill	100.00	100.00	100.00
Zelda	100.00	100.00	100.00
LAX	99.95	99.90	99.76

Table 9.2: Percentage of successful detections under low pass filtering.

Image	MAP (Gen. Gaussian)	MAP (Laplacian)	Correlation
Harbour	100.00	99.98	99.81
Lena	91.56	90.99	77.08
Fishing boat	99.93	99.91	99.78
Peppers	100.00	100.00	96.89
Barbara	95.61	94.73	92.29
Goldhill	100.00	99.87	99.25
Zelda	100.00	100.00	100.00
LAX	100.00	99.94	99.37

Table 9.3: Percentage of successful detections under salt and pepper noise, and followed by median filtering.

Image	MAP (Gen. Gaussian)	MAP (Laplacian)	Correlation
Harbour	99.89	99.81	97.67
Lena	99.80	99.84	99.73
Fishing boat	100.00	99.98	98.91
Peppers	100.00	100.00	99.96
Barbara	99.36	99.26	99.18
Goldhill	100.00	100.00	100.00
Zelda	100.00	99.98	100.00
LAX	99.98	99.96	99.27

Table 9.4: Percentage of successful detections under cropping.

Image	MAP (Gen. Gaussian)	MAP (Laplacian)	Correlation
Harbour	99.82	99.77	99.69
Lena	95.93	95.89	80.78
Fishing boat	96.46	99.95	99.92
Peppers	99.13	99.09	99.36
Barbara	88.09	89.64	100.00
Goldhill	100.00	100.00	100.00
Zelda	100.00	100.00	100.00
LAX	100.00	99.89	90.75

Chapter 10

Epilogue

10.1 Conclusion

Due to the rapid development of multimedia network systems, digital media can be accessed, processed, and stored with ease. The incredible growth of wireless technologies has also made it possible to meet the demand for the availability of multimedia content anyplace and anytime. However, this also leads to the problem of unauthorized duplication and distribution of digital media. Thus, there is an increasing need for mechanisms to protect the security and intellectual property rights of multimedia data over the wired and wireless channels. Digital watermarking has become a popular and effective solution to meet this demand.

In this thesis, we have focused on studying LR detection in image watermarking, where thresholding is done via Neyman-Pearson criterion. Specifically, we have extended the LR detection framework of Barni et al [3] to cover a wider range of probability distribution models. Our original contributions

to this work are summarized as follows:

i. We have shown in Lemma 4.1 that the approximation $f_{\mathbf{Y}}(\mathbf{y}|M_0) \approx f_{\mathbf{Y}}(\mathbf{y}|\mathbf{0})$ holds for any PDF model. With this, we have given in Chapter 4 a general setting to the LR detection framework of Barni et al [3].

ii. For a given PDF model, it is straightforward to obtain the LR decision rule. However, the LR decision threshold usually requires more work to derive. In Chapter 5, we have given the derivation for the closed-form expression of the LR decision threshold under the Gaussian model. The Gaussian model's LR decision threshold reported in [25] is found to be incorrect.

iii. As compared to the Gaussian model, the LR decision threshold under Laplacian model is much more complicated to derive. This is due to the presence of the absolute value sign in the Laplacian PDF expression. In Chapter 6, we have given a complete derivation for the closed-form expression of the LR decision threshold under the Laplacian model. Our experimental results show that the Laplacian model can yield a better watermark detection result than the Gaussian model in DWT domain.

iv. The mean of the DWT coefficients in the high resolution subbands is approximately zero. This leads us to consider using a zero mean generalized Gaussian PDF for LR detection in DWT domain. In Chapter 7, we have given a complete derivation for the closed-form expression of the LR decision threshold under the generalized Gaussian model. We have shown in Lemma

7.2 that the function s defined in (7.18) is of finite range. This facilitates the estimation of the shape parameter via function interpolation. Other estimators based on higher absolute moments of the generalized Gaussian RV are also given. Our experimental results show that the generalized Gaussian model can perform better than the Gaussian and Laplacian models.

v. In Chapter 8, we have given the derivation for the closed-form expression of the LR decision threshold under the generalized gamma model. This can be seen as an extension to the Weibull model of Barni et al [3]. For the work here, new estimators for the parameters of the generalized gamma PDF have been proposed. These estimators are also useful in areas like reliability analysis [15, 41] where the generalized gamma PDF is widely used. Our experimental results show that the LR detector under the generalized Gaussian gamma model can perform better than that for the Weibull model.

Other related contributions include:

i. In Chapter 9, we have formulated the MAP detector under the generalized Gaussian model for watermark detection in DWT. The MAP detector is considered simpler than the LR detector as a decision threshold is not required. We have shown that the generalized Gaussian model yields better detection than the Laplacian model of [2].

ii. We have introduced the energy embedding scheme in Chapter 3 based on modifying the additive scheme. The energy embedding scheme requires more difficult to compute but it is shown that it can make the watermark more robust.

10.2 Suggestions for Further Research

A few interesting areas in which progress can be made are as follows:

1. To perform a detailed performance evaluation for the various LR detection models. This includes examining and comparing the models under a wider range of image processing operations and distortions.

2. To perform a theoretical error analysis of the LR detection method. Since LR detection of Barni et al [3] is based on some approximations, it would be interesting and challenging to study the errors produced by the different PDF models.

3. To explore the application of information theory in statistical watermarking. Majority of the publications in watermarking have focused on novel ways to embed information in media and then to detect it. However, most of these publications lack the mathematical theory describing the fundamental limits of any information-hiding system. Information theoretic watermarking is aimed to provide a theoretical basis for a generic version of the information-hiding problem [30].

4. To find methods to improve robustness of statistical watermarking schemes under geometric attacks [44]. Many watermarks for images and video content are sensitive to geometric distortions. For example, simple rotation, scaling, translation, etc., of an image can prevent detection of a watermark.

5. To explore areas of reversible watermarking using statistical schemes. In a number of domains, such as military, legal and medical imaging, although some embedding distortion is admissible, permanent loss of signal fidelity is undesirable. This highlights the need for reversible (lossless) data embedding and recovery techniques.

Bibliography

[1] M. Barni, F. Bartolini, V. Cappellini, and A. Piva, "A DCT-domain system for robust image watermarking," *Signal Processing*, 66, pp. 357-372, 1998.

[2] M. Barni, F. Bartolini, V. Cappellini, A. Piva and F. Rigacci, "A M.A.P identification criterion for DCT-based watermarking," *9th European Signal Processing Conf.*, EUSIPCO' 98, Sept 8-11, 1998.

[3] M. Barni, F. Bartolini, A. De Rosa and A. Piva, "A new decoder for the optimum recovery of nonadditive watermarks," *IEEE Trans. Image Processing*, Vol. 10, No. 5, pp. 755-766, 2001.

[4] F. M. Boland, J. J. K. O'Ruanaidh, and C. Dautzenberg, "Watermarking digital images for copyright protection," *Fifth Int. Conf. on Image Processing and its Appl.*, pp. 326-330, 1995

[5] T.P.-C. Chen and T. Chen, "A framework for optimal blind watermark detection," *Proc. ACM Multimedia 2001*, Workshop on Multimedia and Security: New Challenges, Ottawa, Canada, Oct. 2001.

[6] Q. Cheng and T.S. Huang, "A DCT-domain blind watermarking system using optimum detection on Laplacian model," *Proc. Int. Conf. on Image Processing*, Vol 1, Vancouver, pp. 454-457, 2000.

[7] Q. Cheng and T.S. Huang, "Blind digital watermarking for images and videos and performance analysis," *IEEE Int. Conf. Multimedia and Expo*, Vol. 1, pp. 389-392, 2000.

[8] Q. Cheng and T.S. Huang, "Robust optimum detection of transform domain multiplicative watermarks," *IEEE Trans. Signal Processing*, Vol. 51, No. 4, pp. 906-924, 2003.

[9] I.J. Cox, J. Killian, T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. Image Process.*, 6, 12, pp. 1673-1687, 1997.

[10] I.J. Cox, M.L. Miller, and J.A. Bloom, "Watermarking applications and their properties," *Proc. Int. Conf. Information Technology: Coding and Computing - ITCC 2000*, pp. 6-10, 2000.

[11] I.J. Cox, M.L. Miller and J.A. Bloom, *Digital Watermarking*, Morgan Kaufmann Publishers, San Francisco, 2002.

[12] J.L. Devore, *Probability and Statistics for Engineering and Sciences*, Brooks/Cole, USA, 2004.

[13] F. Hartung and M. Kutter, "Multimedia watermarking techniques," *Proc. IEEE*, vol. 87, no. 7, pp. 1079-1107, July 1999.

[14] J.R. Hernández and F. Pérez-González, "Statistical analysis of watermarking schemes for copyright protection of images," *Proc. IEEE*, Vol. 87, No. 7, pp. 1142-1166, 1999.

[15] H. Hirose, "Parameter estimation for the 3-parameter gamma distribution using the continuation method," *IEEE Trans. Reliability*, 47, pp. 188-196, 1998.

[16] T.-Y. Hwang and P.-H. Huang, "On new moment estimation of parameters of the generalized gamma distribution using it's characterization," *Ann. Inst. Statist. Math.* Vol. 54, No. 4, pp. 840-847, 2002.

[17] T.-Y. Hwang and P.-H. Huang, "On new moment estimation of parameters of the generalized gamma distribution using it's characterization," *4th Int. Conf. Mathematical Methods in Reliability: Methodology and Practice*, Sante Fe, New Mexico, June 21-25, 2004.

[18] T. Kalker and J. Haitsma. "Efficient detection of a spatial spread-spectrum watermark in MPEG video streams," *Proc. Int. Conf. Image Processing,* Vancouver, BC, Canada, pp. 434-437, Sept 2000.

[19] S. Katzenbeisser and F.A.P. Petitcolas (Editors), *Information Hiding: Techniques for Steganography and Digital Watermarking*, Artech House, Boston, 2000.

[20] Y.-S. Kim, O.-H. Kwon and R.-H. Park, "Wavelet based watermarking method for digital images using the human visual system," *Electronic Letters*, Vol. 35, No.6, 1999.

[21] D. Kincaid and W. Cheney, *Numerical Analysis*, 2nd Edition, Brooks/Cole Pub. Co., 1996.

[22] S.G. Krantz, *Handbook of Complex Variables*, Birkhäuser, Boston, pp. 156. 1999.

[23] E. Kreyszig, *Advanced Engineering Mathematics*, 8th Edition, John Wiley & Sons, New York, 1999.

[24] M. Kutter and F.A.P. Petitcolas, "A fair benchmark for image watermarking systems," *Security and Watermarking of Multimedia Contents*, SPIE-3657, pp. 226-239, 1999.

[25] S.G. Kwon, S.-H. Lee, K.-K. Kwon, K.-R. Kwon and K.-I. Lee, "Watermark detection algorithm using statistical decision theory," *Proc IEEE Int. Conf. on Multimedia and Expo*, Vol. 1, pp. 561-564, 2002.

[26] G.C. Langelaar, I. Setyawan and R.L. Lagendijk, "Watermarking digital image video data: a state-of-art overview," *IEEE Signal Processing Magazine*, pp. 20-46, Sept. 2000.

[27] M. Maes, T. Kalker, J.-P.M.G. Linnartz, J. Talstra, G.F.G. Depovere and J. Haitsma, "Digital watermarking for DVD video copy protection," *IEEE Signal Processing Mag.*, vol. 17, pp. 47-57, Sept. 2000.

[28] S.G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 11, No. 7, pp. 674-693, July 1989.

[29] The MathWorks, Inc., *MATLAB: Image Processing Toolbox*, 1993.

[30] P. Moulin and J.A. O'Sullivan, "Information-Theoretic Analysis of Watermarking," *Proc. ICASSP'00*, Istanbul, Turkey, June 2000.

[31] N. Nikolaidis, V. Solachidis, A. Tefas, and I. Pitas, "Watermarking detection: benchmarking perspectives," *Proc IEEE Conf. Multimedia and Expo*, Vol. 2, pp. 493-496, Aug 26-29, 2002.

[32] T.M. Ng and H.K. Garg, "Maximum likelihood detection in DWT domain image watermarking using Laplacian modeling," *IEEE Signal Processing Letters*, Vol. 12, No. 4, pp. 285-288, Apr 2005.

[33] T.M. Ng and H.K. Garg, "Wavelet domain watermarking using maximum-likelihood detection," *Journal of Imaging Science and Technology*, Vol. 49, No. 3, pp. 303-308, May/June 2005.

[34] T.M. Ng and H.K. Garg, "A maximum a posteriori identification criterion for wavelet domain watermarking," *International Journal of Wireless and Mobile Computing: Special Issue on Mobile Systems and Applications*, 2005.

[35] T.M. Ng and H.K. Garg, "Wavelet domain watermarking using maximum-likelihood detection," *Proc. SPIE Conf. on Security, Steganography, and Watermarking of Multimedia Contents VI*, Vol. 5306, San Jose, Jan 19-22, 2004.

[36] T.M. Ng and H.K. Garg, "A maximum a posteriori identification criterion for wavelet domain watermarking," *Proc. 24th IEEE Intl. Conf. on Distributed Computing Systems Workshop*, Tokyo, March 23-24, 2004.

[37] T.M. Ng and H.K. Garg, "An embedding scheme for bipolar watermark," *Proc. Intl. Conf. Sciences of Electronic, Technologies of Information and Telecommunications*, Tunisia, March 15-20, 2004.

[38] T.M. Ng and H.K. Garg, "Maximum likelihood detection in image watermarking using generalized gamma model," *Proc. 39th Asilomar Conference on Signals, Systems and Computers*, Monterey, pp. 1680-1684, Oct 28-Nov 2, 2005.

[39] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, Inc., New York, pp. 110, 1991.

[40] F. A. P. Petitcolas, R. Anderson, and M. G. Kuhn, "Information hiding - a survey," *Proc. IEEE*, vol. 87, no. 7, pp. 1062-1077, July 1999.

[41] T. Pham and J. Almhana, "The generalized gamma distribution: Its hazard rate and stress-strength model," *IEEE Trans. on Reliability*, Vol. 44, No. 3, pp. 392-397, Sept 1995.

[42] A. Piva, M. Barni, F. Bartolini and V. Cappellini, "Threshold selection for correlation-based watermark detection," *Proc. COST254 Workshop*, L'Aquila, Italy, Apr. 1998.

[43] J.G. Proakis, *Digital Communications*, 3rd Edn, McGraw-Hill, Inc., New York, 1995.

[44] J.J.K. Ó Ruanaidh and T. Pun, "Rotation, scale and translation invariant spread spectrum digital image watermarking," *Signal Processing*, 66, 3, pp. 303-317, May 1998.

[45] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, Inc., New York, 1986.

[46] K. Sharifi and A. Leon-Garcia, "Estimation of shape parameter for generalized Gaussian distributions in subband decompositions of video," *IEEE Trans. Circuits and Systems for Video Tech.,* Vol. 5, No.1, pp. 52-56, 1999.

[47] E.P. Simoncelli and R.W. Buccigrossi, "Embedded wavelet image compression based on a joint probability model," *4th IEEE Int. Conf. Image Processing*, Santa Barbara, CA, Oct 26-29, 1997.

[48] E.W. Stacy, "A generalization of the gamma distribution," The Annals of Mathematical Statistics, Vol. 33, No. 3, pp. 1187-1192, 1962.

[49] E.W. Stacy and G.A. Mihram, "Parameter estimation for a generalized gamma distribution," *Technometrics*, 7, pp. 349-358, 1965.

[50] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*, Prentice Hall, New Jersey, 1995.

[51] X.-G. Xia, C.G. Boncelet and G.R. Acre, "Wavelet transform based watermark for digital images," *Optics Express*, Vol. 3, No. 12, pp. 497-511, 1998.