Multimedia Forensics and Security

Chang-Tsun Li
University of Warwick, UK

Information Science REFERENCE

INFORMATION SCIENCE REFERENCE

Hershey · New York

Acquisitions Editor:	Kristin Klinger
Managing Development Editor:	Kristin Roth
Assistant Managing Development Editor:	Jessica Thompson
Editorial Assistant:	Rebecca Beistline
Senior Managing Editor:	Jennifer Neidig
Managing Editor:	Jamie Snavely
Assistant Managing Editor:	Carole Coulson
Copy Editor:	Holly Powell
Typesetter:	Carole Coulson
Cover Design:	Lisa Tosheff
Printed at:	Yurchak Printing Inc.

Published in the United States of America by
Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue, Suite 200
Hershey PA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

and in the United Kingdom by
Information Science Reference (an imprint of IGI Global)
3 Henrietta Street
Covent Garden
London WC2E 8LU
Tel: 44 20 7240 0856
Fax: 44 20 7379 0609
Web site: http://www.eurospanbookstore.com

Library of Congress Cataloging-in-Publication Data

Multimedia forensics and security / Chang-Tsun Li, editor.

p. cm.

Includes bibliographical references and index.

Summary: "This book provides an in-depth treatment of advancements in the emerging field of multimedia forensics and security by tackling challenging issues such as digital watermarking for copyright protection, digital fingerprinting for transaction tracking, and digital camera source identification"--Provided by publisher.

ISBN 978-1-59904-869-7 (hardcover) -- ISBN 978-1-59904-870-3 (ebook)

1. Multimedia systems--Security measures. 2. Data encryption (Computer science) 3. Data protection. I. Li, Chang-Tsun.

QA76.575.M83187 2008

005.8--dc22

2008008467

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

All work contributed to this book set is original material. The views expressed in this book are those of the authors, but not necessarily of the publisher.

If a library purchased a print copy of this publication, please go to http://www.igi-global.com/agreement for information on activating the library's complimentary electronic access to this publication.

Chapter VII
Statistical Watermark Detection in the Transform Domain for Digital Images

Fouad Khelifi
The Institute of Electronics, Communications and Information Technology (ECIT), Queen's University Belfast, UK

Fatih Kurugollu
The Institute of Electronics, Communications and Information Technology (ECIT), Queen's University Belfast, UK

Ahmed Bouridane
The Institute of Electronics, Communications and Information Technology (ECIT), Queen's University Belfast, UK

ABSTRACT

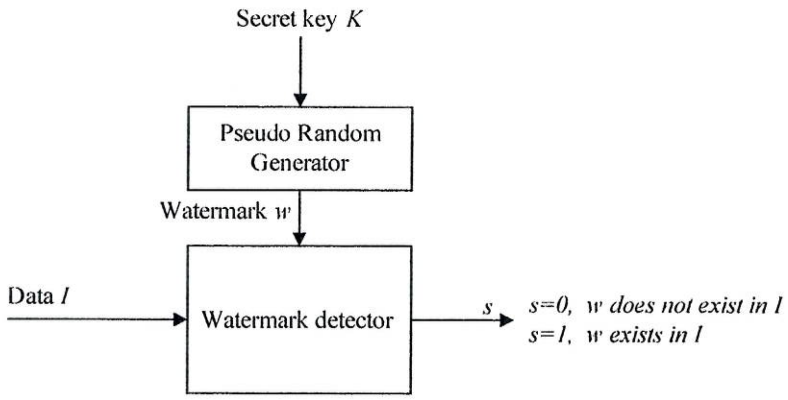
The problem of multiplicative watermark detection in digital images can be viewed as a binary decision where the observation is the possibility that watermarked samples can be thought of as a noisy environment in which a desirable signal, called watermark, may exist. In this chapter, we investigate the optimum watermark detection from the viewpoint of decision theory. Different transform domains are considered with generalized noise models. We study the effect of the watermark strength on both the detector performance and the imperceptibility of the host image. Also, the robustness issue is addressed while considering a number of commonly used attacks.

INTRODUCTION

Recently, we have seen an unprecedented advance in the use and distribution of digital multimedia data. However, illegal digital copying and forgery have become increasingly menacing as the duplication means are easy to use. This makes the protection of copyrighted original copies from illegal use and unrestricted broadcasting a very challenging task. These challenges and issues have involved the field of watermarking for the

Figure 1. Multi-bit watermark extraction system

Figure 2. One-bit watermark detection system



protection and security of digital data (Arnold, Schmucker, & Wolthusen, 2003).

Watermarking is the embedding of a hidden secondary data into the host data within an embedding distortion level. The robustness requires that the watermark information must be decoded/detected even if the watermarked data has undergone additional distortions. Existing systems are divided in two groups depending on the roles that watermarks play. In the first group, the watermarks are viewed as transmitted multi-bit information, where the decoder extracts the full version of the embedded message bit by bit. In such a case, the decoder already assumes that the input data is watermarked (Figure 1).

In the second group, known as one-bit watermarking, the watermarks serve as verification codes, as such a full decoding is not really necessary. It is used to decide whether or not a particular message or pattern is present in the host data (Figure 2).

In practice, the security of the entire one-bit watermarking systems is ensured by using a secret key, as commonly employed in communications, which is required to generate the watermark sequence. Only the legal owner of the watermarked data knows the key for the generation of that watermark and proves his/her ownership.

Depending on the embedding rule used in a watermarking system, the watermark is often additive or multiplicative (Langelaar, Setyawan, & Langendijk, 2000). To get better performances in terms of robustness and imperceptibility, both are used in the transform domain (Barni, Bartolini, Cappellini, & Piva, 1998; Cheng & Huang, 2001a). In fact, the energy compaction property exhibited in the transform domain suggests that the distortions introduced by a hidden data into a number of transform coefficients will be spread over all components in the spatial domain so as the change of the pixels values is less significant.

In additive watermarking, the watermark is simply added to a set of transformed coefficients with a scaling factor controlling the watermark strength. In the case of multiplicative watermarking, the watermark is inserted with respect to the transformed coefficients. Although additive watermarking is widely used in the literature for its simplicity, multiplicative watermarking exhibits a better exploitation of the characteristics of the human visual system (HVS) in digital images and, unlike additive watermarking, offers a data-dependent watermark casting.

This chapter provides an overview of the theory and practice of binary decision in one-bit multiplicative image watermarking in the transform domain. The likelihood formulation of the watermark detection problem is attractive for several reasons. First, it delivers an optimal solution in the binary decision in the sense that the probability of missed detection is minimized subject to a fixed false alarm probability. Second, the derivation of such a decision rule can easily be extended to multi-bit detection and thus can be used in the decoding of hidden data in multi-bit watermarking systems. Finally, it uses only few parameters to model the host data. These parameters are roughly unchanged even when the data holds the watermark and undergoes different attacks. Therefore, a fully blind detection can be used since the original data is not required. From a viewpoint of decision theory, the effect of multiplicative watermark strength on the visual quality of watermarked images and on the detection is investigated and different techniques proposed in the literature are reviewed in this chapter. The robustness issue is also addressed in this chapter while assessing a number of optimum detectors in the transform domain.

In the next section, we review different state-of-the-art watermark detection techniques proposed in the literature. Then, we describe the problem of watermark detection from a decision theory viewpoint in the third section. The fourth section describes two different generalized distributions to model the behavior of the host data in the transform domain. In particular, the discrete wavelet transform (DWT) domain, the discrete cosine transform (DCT) domain, and the discrete Fourier transform (DFT) domain are considered. The watermark detectors derived in the fourth section will be assessed experimentally on a number of digital test images in the fifth section. Conclusions and further discussions are provided in the final two sections.

BACKGROUND

The straightforward approach for watermark detection is to compare the correlation between the candidate watermark $w^* = (w_1^*, \cdots, w_N^*)$ and image coefficients in which the actual watermark is inserted $y^* = (y_1^*, \cdots, y_N^*)$ with some threshold T as illustrated by Figure 3. This assumes that the actual watermark is strongly correlated with the watermarked samples and, thus, the correlation coefficient λ should be large to some extent (Cox, Kilian, Leighton, & Shamoon, 1997; Langelaar et al., 2000). However, such a correlation-based detection would be optimum only in the additive case under the assumption that the image coefficients follow a Gaussian distribution (Elmasry & Shi, 1999). This has been pointed out when formulating the watermark detection problem as a statistical hypothesis testing problem (Cheng & Huang, 2001a). In fact, the problem of watermark detection can be viewed as a binary decision where the observation is the possibly watermarked transformed coefficients, that is, this can be formulated as a problem of detecting a known signal in a noisy environment (Green & Swets, 1966) (Figure 4).

The statistical behavior of the transformed coefficients can be used to derive a decision rule that decides whether a candidate watermark presented to its input is actually embedded in the data (hypothesis H_1) or not (hypothesis H_0), thereby, conventional statistical signal detectors can be used. In additive watermarking, Hernandez, Amado, and Perez-Gonzalez (2000) used

Figure 3. Correlation-based watermark detector

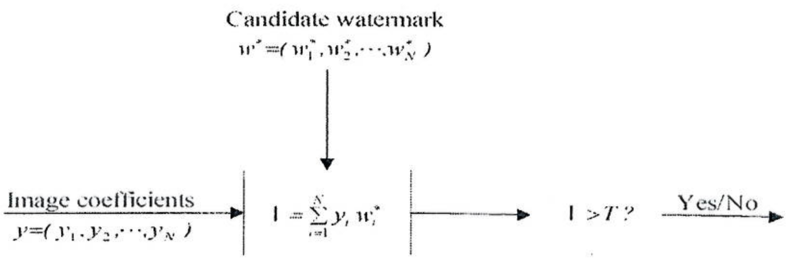
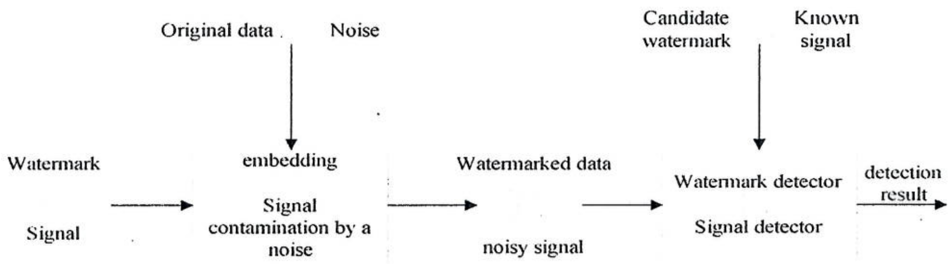


Figure 4. Block diagram illustrating the formulation of the problem of one-bit watermarking as a problem of signal detection in noisy observation.



the generalized Gaussian (GG) distribution to statistically model the 8×8 DCT coefficients. The resulting detector structure has been shown to outperform a correlation based counterpart. In Cheng and Huang (2001a) the GG model was also adopted for DWT coefficients. In the multiplicative case, a generalized correlator based on the GG distribution has been derived and applied in the DWT domain (Cheng & Huang, 2001b). Such a detector is locally optimum in the sense that it approaches the optimality for weak signals[1] (Miller & Thomas, 1972). In the DFT domain, the Weibull distribution has been employed, based on which an optimum detector has been derived (Barni, Bartolini, De Rosa, & Piva, 2001). Likewise, the same distribution has been used to derive a locally optimum detector structure (Cheng & Huang, 2002). Kwon, Lee, Kwon, Kwon, and Lee (2002) have derived an optimum multiplicative detector based on the normal distribution. However, although their proposed detector significantly outperforms the correlation-based detector for DWT-transformed images, such a distribution can not accurately model the host coefficients. In Ng and Grag (2005), the Laplacian model, which exhibits more accuracy, has been considered. Consequently, better performance in terms of detection has been achieved. It has also been used in Khelifi, Bouridane, and Kurugollu (2006) when investigating the problem of detector accuracy. Nevertheless, as many authors have pointed out, wavelet sub-bands can be modeled perfectly by a GG distribution (Calvagno, Ghirardi, Mian, & Rinaldo, 1997; Do & Vetterli, 2002). Thereby, the proposed detector in Ng and Grag (2004), using the GG model, has been proven to provide the best performance. In investigation on robust optimum detection of multiplicative watermarks, a class of locally optimum detectors has been developed based on the GG distribution for subband transformed domains (Cheng & Huang, 2003). In fact, in the presence of attacks, the multiplicative watermark detection problem can be formulated as a detection of a known signal in non-additive noise where the attacks are thought of as uncertainties in the observation model. Such a non-additive noise observation model has been extensively addressed in Blum (1994).

PROBLEM FORMULATION

The commonly used multiplicative embedding rule is (Barni et al., 2001; Cheng & Huang, 2003)

$$y_i = x_i\,(1 + \lambda\, w_i) \qquad (1)$$

where x = (x_1,..., x_n) is a sequence of data from the transformed original image, w = (w_1,...,w_N) is the watermark sequence, λ is a gain factor controlling the watermark strength, and y = (y_1,...,y_N) is the sequence of watermarked data. By relying on decision theory, the observation variables are the vector y of possibly marked coefficients. The candidate watermark vector at the input of the detector is denoted by $w^* -(w_1^*,\ldots,w_N^*)$. Let us define two regions (classes) W_0 and W_1 where $W_0 = \{w \mid w \neq w^*\}$ (hypothesis H_0), including w = 0 that corresponds to the case where no watermark is embedded, and $W_1 = \{w^*\}$ (hypothesis H_1). The likelihood ratio $\ell(\mathrm{y})$ is

$$\ell(\mathrm{y}) = \frac{f_{\mathrm{y}}(\mathrm{y} \mid W_1)}{f_{\mathrm{y}}(\mathrm{y} \mid W_0)} \qquad (2)$$

where f_y (y | W) is the probability density function (pdf) of the vector y conditioned to W. In practice, y satisfies f_y (y | W_0) > 0 since the host samples have a non-zero occurrence probability. By relying on the fact that the components of y are independent of each other, we can write

$$f_y(y \mid W_0) = \prod_{i=1}^{N} f_y(y_i \mid W) \qquad (3)$$

Under the assumption that watermarks are uniformly distributed in [-1,1], W_0 consists of an infinite number of watermarks. Hence, according to the total probability theorem (Papoulis, 1991), $f_{yi}(y_i \mid W_0)$ can be written as

$$f_{yi}(y_i \mid W_0) = \int_{-1}^{1} f_{yi}(y_i \mid W) f_{wi}(w)\, dw \qquad (4)$$

where f_{wi} (w) is the pdf of w_i. Therefore

$$f_{yi}(y_i \mid W_0) = \tfrac{1}{2}\int_{-1}^{1} f_{yi}(y_i \mid W)\, dw \qquad (5)$$

Thus, combining equations (2), (3), and (5) yields

$$\ell(\mathrm{y}) = \frac{\prod_{i=1}^{N} f_{y_i}(y_i \mid w_i^*)}{\frac{1}{2^N}\prod_{i=1}^{N}\int_{-1}^{1} f_{y_i}(y_i \mid w)\, dw} \qquad (6)$$

By remembering the watermarking rule, the pdf f_y (y | w) of a watermarked coefficient y conditioned to a watermark value w is

$$f_y(y \mid w) = \frac{1}{1+\lambda w} f_x\left(\frac{y}{1+\lambda w}\right) \qquad (7)$$

where f_x (x) indicates the pdf of the original, non-watermarked coefficients.[2] Thus, equation (6) becomes

$$\ell(\mathrm{y}) = \frac{\prod_{i=1}^{N} \frac{1}{1+\lambda w_i^*} f_{x_i}\left(\frac{y_i}{1+\lambda w_i^*}\right)}{\frac{1}{2^N}\prod_{i=1}^{N}\int_{-1}^{1} \frac{1}{1+\lambda w} f_{x_i}\left(\frac{y_i}{1+\lambda w}\right) dw} \qquad (8)$$

Under the assumption λ « 1, the following approximation is widely adopted in the literature[3] (Barni & Bartolini, 2004; Barni et al., 2001)

$$\frac{1}{2}\int_{-1}^{1}\frac{1}{1+\lambda w}f_x\left(\frac{y}{1+\lambda w}\right)dw \approx f_x(y) \quad (9)$$

More specifically, the derivation of the likelihood ratio relies on the assumption below

$$f_y(y \mid w_0) \approx f_y(y \mid 0) \quad (10)$$

where by 0 the null watermark is meant. By evoking the approximation (9) and taking logarithm, the decision rule has the form (Barni & Bartolini, 2004)

$$\varphi(y) \begin{matrix} > T \boxtimes H_1 \\ < T \boxtimes H_0 \end{matrix} \quad (11)$$

where T is a properly chosen threshold, and

$$\phi(y) = \sum_{i=1}^{N}\left[ln\left(f_{x_i}\left(\frac{y_i}{1+\lambda w_i^*}\right)\right) - ln(f_{x_i}(y_i))\right]$$

$$\varphi(y) = \sum_{i=1}^{N} v_i \quad (12)$$

To obtain the threshold T, the Neyman-Pearson criterion is used in such a way that the missed detection probability is minimized, subject to a fixed false alarm probability (Ferguson, 1967).

$$P_{FA} = P(\phi(y) > T \mid W_0$$

$$= \int_T^{+\infty} f_\phi(\alpha \mid W_0)\, d\alpha \quad (13)$$

Note that the variable φ (y) is a sum of statistically independent terms. Thus, by invoking the central limit theorem (Papoulis, 1991), its pdf can be assumed to be a normal one. Therefore, the mean and variance of φ(y) are given by (Barni et al., 2001; Cheng & Huang, 2003)

$$\mu_\phi = \sum_{i=1}^{N}\mu_{v_i} \quad (14)$$

$$\sigma_\phi^2 = \sum_{i=1}^{N}\sigma_{v_i}^2 \quad (15)$$

where by μ_{v_i} and $\sigma_{v_i}^2$, it is meant the mean and the variance of v_i respectively. Under hypothesis $\prod_0$, in view of equation (10), we have

$$\mu_{v_i|y_i=x_i} = E\left[ln\left(f_{x_i}\left(\frac{x_i}{1+\lambda w_i^*}\right)\right) - ln(f_{x_i}(x_i))\right] \quad (16)$$

$$\sigma_{v_i|y_i=x_i}^2 = E\left[\left(ln\left(f_{x_i}\left(\frac{x_i}{1+\lambda w_i^*}\right)\right) - ln(f_{x_i}(x_i)) - \mu_{v_i}\right)^2\right] \quad (17)$$

Practically, the parameters of the pdf $f_{x_i}(\cdot)$ are not available (blind watermarking). Instead, the possibly marked coefficients are used. Finally, P_{FA} can be written as

$$P_{FA} = \int_T^{+\infty}\frac{1}{\sqrt{2\pi\sigma_\phi^2}}exp\left(-\left(\frac{\alpha-\mu_\phi}{2\sigma_\phi^2}\right)^2\right)d\alpha$$

$$= \frac{1}{2}erfc\left(\frac{T-\mu_\phi}{\sqrt{2\sigma_\phi^2}}\right) \quad (18)$$

where *erfc* is the complementary error function, given by

$$erfc(x) = \frac{2}{\pi}\int_x^{+\infty} e^{-t^2}\, dt \quad (19)$$

Hence

$$T = erfc^{-1}(2\,P_{FA})\sqrt{2\sigma_{\phi}^{2}} + \mu_{\phi} \qquad (20)$$

APPLICATION

Watermark detector based on the GG model. The GG model is broadly used in the literature to describe some probability distributions of digital data in the transform domain. It basically provides a good understanding of the statistical behavior of those signals that are impulsive and heavy-tailed such as the wavelet sub-bands (Calvagno et al., 1997; Do & Vetterli, 2002) and non dc coefficients in the DCT transform domain (Birney & Fischer, 1995). The probability density function of a zero-mean GG distribution is

$$f_X(x) = A \exp\left(-|\beta\, x|^{c}\right),\ c > 0 \qquad (21)$$

where $\beta = \frac{1}{\sigma}\sqrt{\left(\frac{\Gamma(3/c)}{\Gamma(1/c)}\right)}$, $A = \frac{\beta c}{2\Gamma(1/c)}$, σ is the standard deviation of the distribution, and $\Gamma(t) = \int_0^{+\infty} r^{t-1}\, e^{-r}\, dr$ is the Gamma function. It is worth mentioning that the GGD contains the Gaussian and the Laplacian distributions as special cases, with $c = 2$ and $c = 1$, respectively. As depicted in Figures 5 and 6, when $c \rightarrow 0$, it approaches the Dirac function. When $c \rightarrow \infty$, it tends to the uniform distribution. Practically, c can be computed by solving (Mallat, 1989; Sharifi & Leon-Garcia, 1995)

$$c = F^{-1}\left(\frac{E[|X|]}{\sigma}\right) \qquad (22)$$

where

$$F(z) = \frac{\Gamma(2/c)}{\Gamma(1/c)\Gamma(3/c)}$$

In Figure 7 we show the actual and estimated pdfs on a set of transformed coefficients in the DWT and DCT domains for 'Goldhill' 512×512 test image.

Obviously, the GG model fits the empirical distributions perfectly. By replacing the pdf of the GG model in equation (12), the detector can be expressed as (Cheng & Huang, 2003; Ng & Grag, 2004)

Figure 5. pdf of the generalized Gaussian distribution for small values of c with $\sigma = 15$

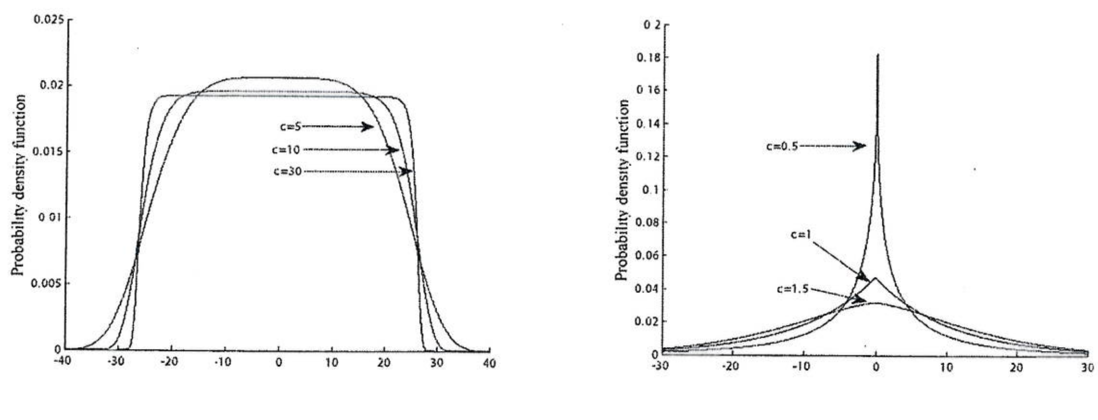


Figure 6. pdf of the generalized Gaussian distribution for large values of c with $\sigma = 15$

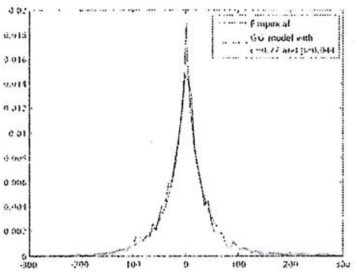
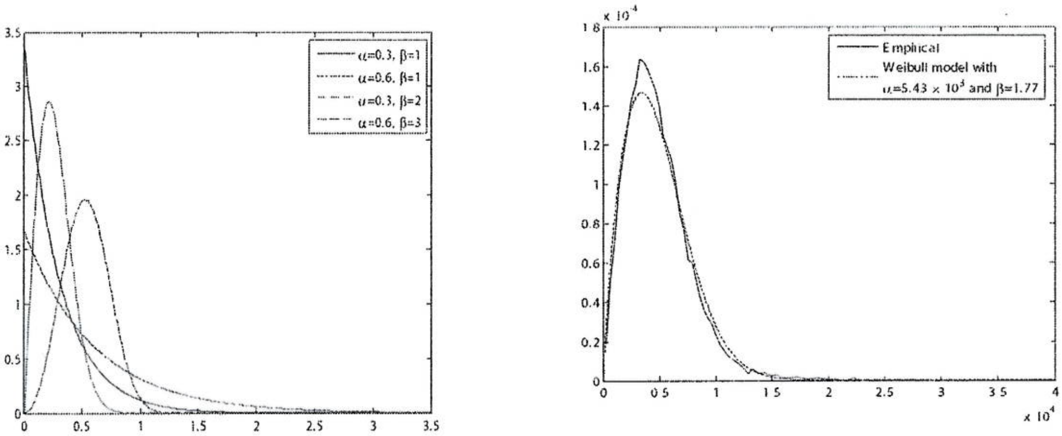


Figure 7. Fig. 7. Modeling of transformed coefficients for 'Goldhill' image. Left. DWT: third level sub-band HL3 coefficients using 9/7 biorthogonal filters. Right. DCT: a set of 12000 consecutive coefficients in a zigzag scan order of the whole DCT transformed image starting from sample 31000.



$$\phi(y) = \sum_{i=1}^{N} |\beta_i y_i|^{c_i} \left(1 - \frac{1}{(1+\lambda w_i^{*})^{c_i}}\right) \quad (23)$$

The threshold T can be determined from equation (20) where,

$$\mu_{\phi} = \sum_{i=1}^{N} \frac{1}{c_i}\left(1 - \frac{1}{(1+\lambda w_i^{*})^{c_i}}\right) \quad (24)$$

$$\sigma_{\phi}^{2} = \sum_{i=1}^{N} \frac{1}{c_i}\left(1 - \frac{1}{(1+\lambda w_i^{*})^{c_i}}\right)^{2} \quad (25)$$

Watermark Detector Based on the Weibull Model

The Weibull distribution offers a good flexibility to describe the statistical characteristics of the coefficients magnitude in the DFT domain (Barni et al., 2001). It is defined as

$$f_X(x) = \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} \exp\left(-\left(\frac{x}{\alpha}\right)^{\beta}\right) \quad (26)$$

for $x \geq 0$, where $\beta > 0$ is the shape parameter and $\alpha > 0$ is the scale parameter of the distribution. The n^{th} raw moment of a random variable x that follows the Weibull distribution is given by

$$m_n = \alpha^n \Gamma\left(1 + \frac{n}{\beta}\right) \quad (27)$$

Figure 8 shows a number of Weibull functions with different parameters. When $\beta = 1$, the Weibull distribution reduces to the exponential distribution. When β increases, it appears similar to a normal distribution. The same test image 'Goldhill' is also used to show the model function of a set of 10,000 consecutive DFT coefficient magnitudes selected in a zigzag order of the top right quadrant (Figure 9).

The use of the Weibull distribution in equation (12) leads to the following watermark detector (Barni et al., 2001)

$$\phi(y) = \sum_{i=1}^{N} y_i^{\beta_i} \left(\frac{(1+\lambda w_i^*)^{\beta_i} - 1}{\alpha_i^{\beta_i}(1+\lambda w_i^*)^{\beta_i}}\right) \quad (28)$$

Equation (20) can be used to derive the threshold where,

Figure 8. Weibull functions for different values of parameters α and β

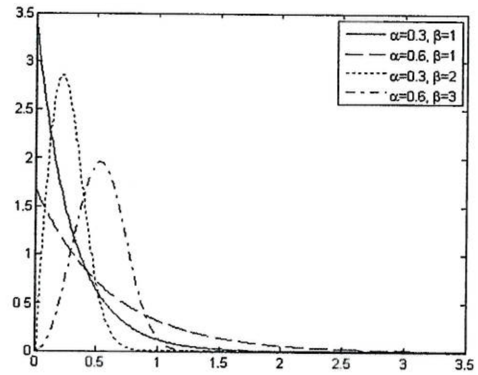
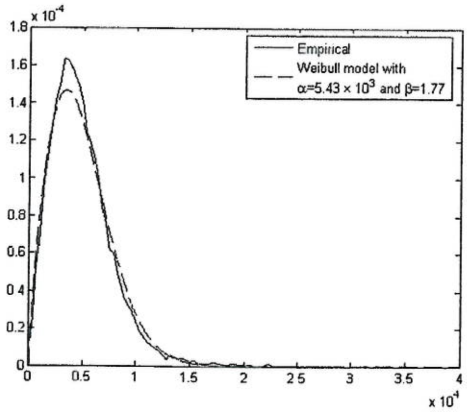


Figure 9. Modeling of DFT coefficients for 'Goldhill' image



$$\mu_{\phi} = \sum_{i=1}^{N}\left(\frac{(1+\lambda w_i^*)^{\beta_i}-1}{(1+\lambda w_i^*)^{\beta_i}}\right) \tag{29}$$

$$\sigma_{\phi}^2 = \sum_{i=1}^{N}\left(\frac{(1+\lambda w_i^*)^{\beta_i}-1}{(1+\lambda w_i^*)^{\beta_i}}\right)^2 \tag{30}$$

EXPERIMENTAL ANALYSIS

In this section, we evaluate the performance of the detectors discussed earlier on six 512×512 grayscale standard images with different information content. At the detector side, the original image is assumed to be unavailable and therefore the statistical model parameters used to perform a decision are directly computed from the watermarked and possibly corrupted images. Three issues are considered here. First, the imperceptibility of the watermark which is assessed visually and quantitatively by using the peak signal-to-noise ratio (PSNR) as a fidelity measure. Second, the detection performance with respect to the probability of false alarm and the probability of missed detection. We use the receiver operating characteristic (ROC) curve which is widely adopted in the literature for interpreting the performance of communications systems. Eventually, the robustness of the watermark is addressed while considering a number of commonly used attacks.

Imperceptibility

It is worth mentioning that the watermark sequence consists of a sequence of 12,288 random real numbers uniformly distributed in[4] [-1,1]. The watermark casting procedure in the DFT and DCT domains is similar to that used in Barni et al. (1998, 2001), respectively. In Cox et al. (1997), a set of the largest coefficients in the DCT domain were selected to hold the watermark sequence. However, this would require the original image to determine those components which hide the watermark since the watermarked data may be distorted. We have experimentally found that the approach used in Barni et al. (1998, 2001) to insert the watermark is more practical and brings comparable performance to that of Cox et al. (1997). This is justified by the fact that most of an image energy is packed into the low frequency components. Thus, the largest coefficients in the DCT and DFT domains are very likely to be found when scanning the image in a zigzag order starting from the first non-zero frequency coefficient. In the DWT domain, the watermark is embedded as suggested in Ng and Grag (2005). That is, all coefficients in the sub-bands of the third level, except the approximation sub-band, are used to insert the watermark sequence. In the DCT and DFT domains, the start point for embedding the watermark in a zigzag order has been selected in such a way the PSNRs obtained are close to those delivered in the DWT domain. In the experiments, the start point has been set to 1200 for the DFT and 4000 for the DCT. The PSNRs are plotted in Figures 10-12 for various test images.

In the evaluation of watermarking systems, two different types of imperceptibility assessment can be distinguished: *quality* and *fidelity* (Cox, Miller, & Bloom, 2002). Fidelity is a measure of the similarity between the watermarked data and the original one. The quality assessment, on the other hand, is concerned with obvious processing alterations without referring to the original data. In other words, only high quality of the watermarked data is targeted. In this chapter, fidelity is the perceptual measure of our concern.

From the first set of experiments, two relevant points arise when measuring the distortions by PSNRs. First, as pointed out in Jayant, Johnston, and Safranek (1993), although the automated fidelity test adopted here is often used in the literature, it does not reflect the true fidelity in some cases. Indeed, the visual imperceptibility also depends on the data content. The higher the frequency content exhibited in the image, the better the imperceptibility expected. Figure 13 shows an example of two test images with dif-

Figure 10. PSNR of watermarked images in the DWT domain

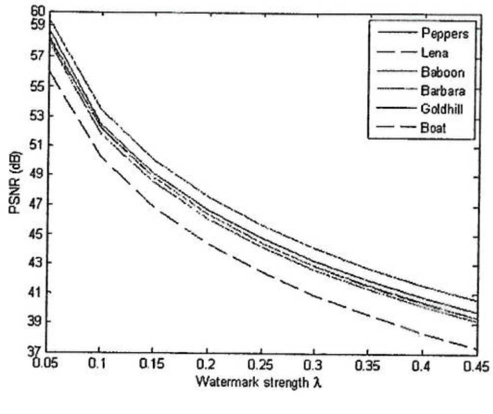


Figure 11. PSNR of watermarked images in the DFT domain

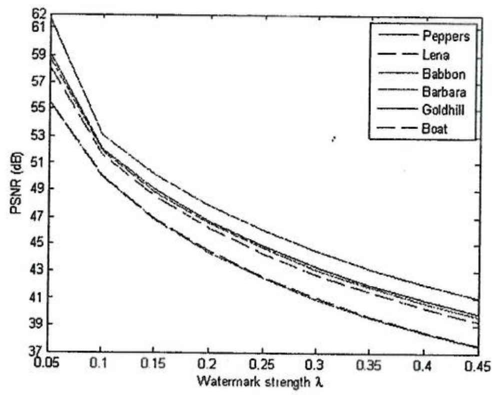


Figure 12. PSNR of watermarked images in the DCT domain

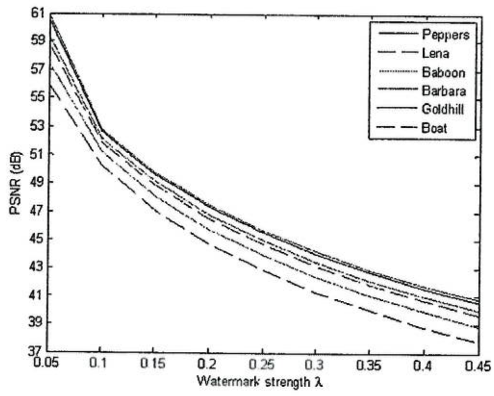
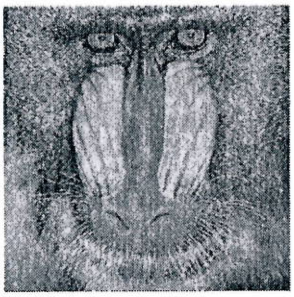
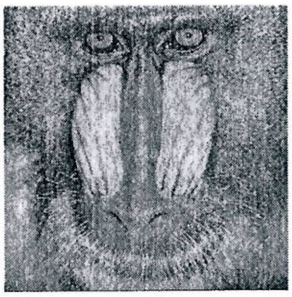


Figure 13. Illustration of a case in which PSNR cannot reflect the imperceptibility of the watermark. (a) Original Lena. (b) Watermarked Lena in the DFT domain PSNR=41.46 dB. (c) Original Baboon. (d) Watermarked Baboon in the DFT domain PSNR=39.62 dB.

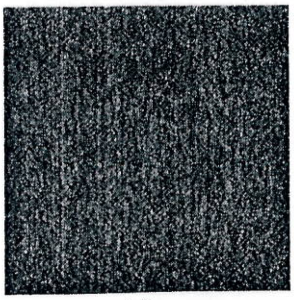

(a) *(b)*

(c) *(d)*

Figure 14. Visual effect of the watermark casting in different transform domains. (a) Watermarked Lena in the DWT domain PSNR=41.55. (b) Difference image with magnitudes multiplied by 20 in the DWT case. (c) Watermarked Lena in the DFT domain PSNR=41.46 dB. (d) Difference image with magnitudes multiplied by 20 in the DFT case.


(a) *(b)*

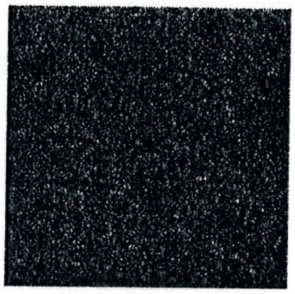



(c) *(d)*

continued on following page

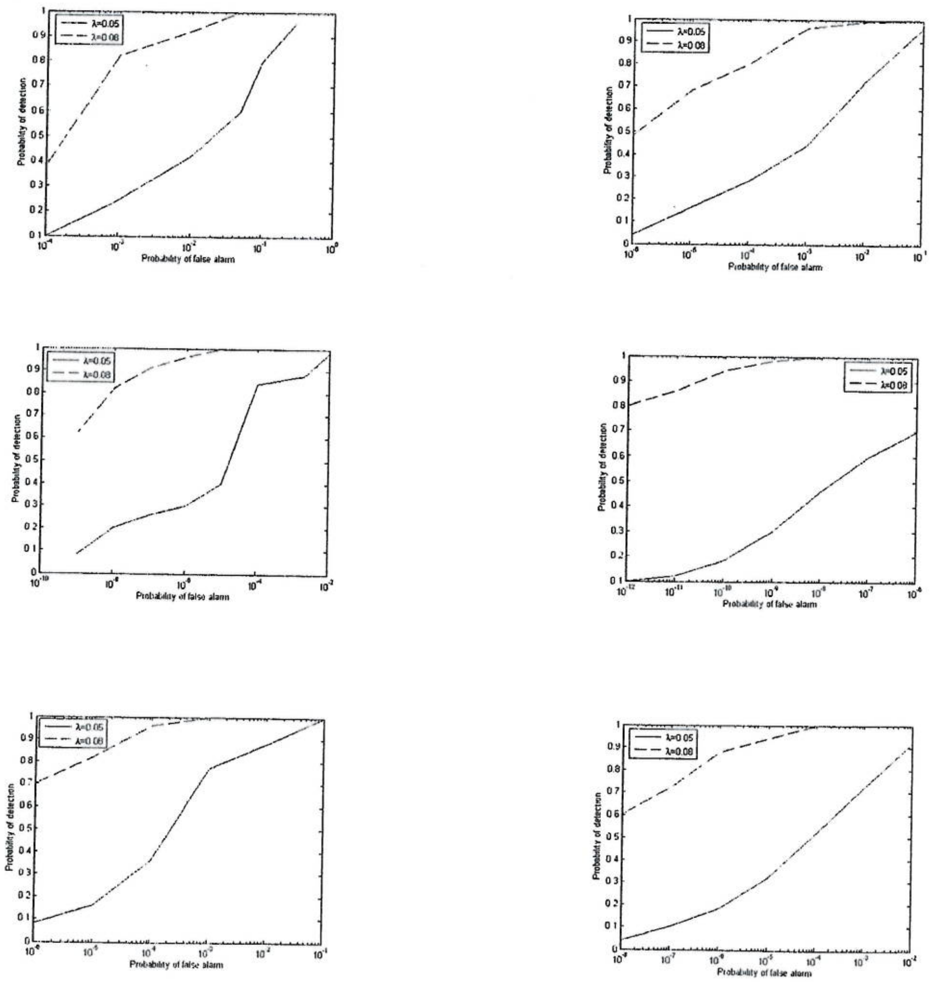
Figure 14. continued

(e)



(f)

Figure 15. ROC curves. Left side graphs: ROCs for Lena. Right side graphs: ROCs for Baboon. Top: watermarking in the DWT domain. Middle: watermarking in the DFT domain. Bottom: watermarking in the DCT domain.



ferent frequency content. Obviously, PSNR, as a perceptual model, suggests that the watermarked version of 'Lena' should be perceptually better than the watermarked 'Baboon' image. However, the watermarked 'Lena' shows more visible distortions when compared to the original image. The second point that should be mentioned is that a wavelet transform offers the best imperceptibility of the embedded watermarks when compared to the other transforms. This transform actually provides an image-dependent distortion that is mostly concentrated at edges and detail regions. It is well known that the HVS is less sensitive to such regions compared to uniform and non-textured regions. Figure 14 illustrates an example of the 'Lena' image holding the same watermark in different transform domains. From this figure, it can be seen that the watermarked images in the DFT and the DCT domains show more noticeable distortions when compared against the watermarked image in the DWT domain. The difference image clearly shows the image dependency of the watermarking in the DWT domain. The distortion, which is similarly introduced in the DFT and DCT domains, appears like a noise independent of the host image.

Detection Performance

In order to gauge the performance of the detectors, a number of test images were watermarked using λ=0.05 and λ=0.08 in different transform domains. ROC curves were used to assess the watermark detection. These represent the variation of the probability of correct detection against the probability of false alarm. A perfect detection would yield a point at coordinate (0,1) of the ROC space meaning that all actual watermarks were detected and no false alarms are found. Without loss of generality, the results on two sample images are plotted in Figure 15.

Obviously, the larger the watermark strength, the better detection performance is. In fact, from the standpoint of the signal detection theory, the watermark strength can be thought of as the signal amplitude. Higher values for signal amplitude make the hypotheses H_0 and H_1 more distinguishable. Also, it can easily be seen that the 'Baboon' image, which carries more details and texture information than 'Lena', does offers an improved detection. This is justified by the fact that most of the transform coefficients that hold the watermark for 'Baboon' are larger. Indeed, since the watermark casting depends on the coefficients magnitude in the multiplicative case (see equation (1)), the coefficients with larger magnitude ensure significant presence of the watermark.

From the experiments, the performance of the Weibull model-based detector in the DFT domain is significantly better than that obtained with the DCT detector which, in turn, outperforms the DWT detector for all the test images. Indeed, at the same probability of detection, the DFT-based detector delivers the smallest probability of false alarm in comparison with the DWT- and DCT-based detectors. The DFT domain offers an effective solution to embed the watermark in the sense that it permits attractive ability of detection. However, in practical situations, this should be viewed in connection with the robustness of watermarks when the host data is altered with intentional or unintentional manipulations. The following section is devoted to analysis of the watermark robustness in the transform domain.

Robustness Analysis

In the set of experiments carried out, six various well known images were used. Namely, 'Peppers', 'Lena', 'Baboon', 'Barbara', 'Goldhill', and 'Boat'. As mentioned earlier with regard to the imperceptibility, and in order to make comparison as fair as possible, λ is set to 0.18 for the DCT and DFT domains. For the DWT domain λ=0.2. The robustness of embedded watermarks is measured against standard image processing and geometric manipulations. Given six different watermarks, each one of them is embedded

Table 1. Robustness of embedded watermarks in the transform domain. QF: Quality Factor. v : variance. W: Window size. P: Number of pixels the image is shifted up and left. C: size of the cropped image (center part) compared to the original one. a: rotation angle (in degrees).

Attack		DWT	DFT	DCT
JPEG compression	QF=90	1	1	1
	QF=75	0	1	1
	QF=10	0	1	1
	QF=5	0	0	0
Gaussian white Noise addition	v=500	1	1	1
	v=1000	0	1	1
	v=2000	0	1	1
	v=4000	0	0	1
	v=8000	0	0	1
	v=1600	0	0	0
Median filtering	W=3×3	1	1	1
	W=5×5	0	1	1
	W=7×7	0	1	1
	W=9×9	0	0	0
Shifting	P=5	0	1	1
	P=10	0	1	1
	P=50	0	1	1
	P=150	0	1	1
	P=200	0	0	0
Cropping	C=95 %	0	1	0
	C=90 %	0	1	0
	C=80 %	0	1	0
	C=70 %	0	1	0
	C=60 %	0	0	0
Rotation	a=0.1	1	1	1
	a=0.2	0	1	1
	a=0.5	0	1	1
	a=0.6	0	0	0

into a test image. For each image, a set of 1000 watermarks is presented to the detector. Among these watermarks, there is only one watermark sequence that is actually embedded in that image. For all the watermarked images, the detection is said to be accurate (denoted by 1) if none of the fake watermarks is detected while the actual one is correctly detected. Otherwise, the detection will be inaccurate (denoted by 0). The theoretical probability of false alarm P_{FA}, which is used to determine the decision threshold in equation (20), has been fixed at 10^{-6}. The results are summarized in Table 1.

From the results, it can clearly be seen that the detector operating in the DWT domain performs poorly when compared with those implemented in the DCT and DFT domains. Unlike the DCT and DFT, which are invariant to small rotational and translational manipulations,[1] the DWT does not exhibit such desirable properties. Therefore, the embedded watermarks in the DWT domain cannot be recovered even for those attacks of slight visual impact on the watermarked images. It can also be observed that the DFT shows attractive ability to detect the embedded watermarks in small sub-parts of the watermarked images. This cannot be achieved in the DWT and DCT domains although significant portions from the watermarked images were maintained in some experiments (95% of the watermarked images). Finally, it should be mentioned that the DCT detector provides the best robustness to noise addition attacks. Overall, the DCT and DFT detectors perform closely well and show significant superiority over the DWT detector.

CONCLUSION

This chapter discussed the problem of multiplicative watermark detection in the transform domain. The optimum watermark detection is based on the information theory and can be formulated as a detection of a known signal in noisy observation. In one-bit watermarking, the requirements on the robustness and imperceptibility are essential. Robustness refers to the ability to detect the watermark after common signal processing operations. The challenges in the design of such watermarking systems are that the requirements are generally conflicting with each other. Indeed, a perfect imperceptibility of the watermark can be obtained while reducing the watermark strength. However, such improvements come at the price of the robustness.

The transform domain yields better performances in terms of imperceptibility and robustness. We have focused on the watermark detection in three different transform domains that are widely used in the literature: DWT, DFT, and DCT. The GG and the Weibull models have been used to describe the statistical behavior of the transform coefficients in order to derive the decision rule.

The watermark imperceptibility in the transform domain has extensively been discussed through several experiments. The difficulty in measuring the visual perceptibility of the watermark lies in the fact that the automated evaluation is deceiving in some particular cases. Indeed, such a perceptual assessment treats changes in all regions of an image data equally. However, human perception mechanisms are not uniform.

The experimental results have shown that the DWT offers better visual fidelity with an image dependent watermark casting effect when compared against the DFT and DCT. That is, the distortions introduced affect the high activity regions in which the HVS is less sensitive. In the second set of experiments, we have studied the ability of detection in the transform domain without any attack. ROC curves have been used to assess the detection performance. It has been observed that the DFT detector considerably outperforms the DCT detector which, in turn, provides a better detection than that obtained with a DWT detector. The robustness of the watermark has also been addressed through different common intentional and unintentional manipulations. With a comparable visual fidelity of the embedded watermarks, the DFT and DCT detectors, though performing very closely, have shown to provide significant improvements over a detection in a DWT domain.

FUTURE RESEARCH DIRECTIONS

Several research trends attempt to adapt the watermark embedding procedure to the sensitivity of the HVS. For this purpose, a perceptual model can be used to select the best suited coefficients to hold the watermark in the transform domain. However, one should pay attention to the effect of image processing manipulations on the perceptual masking at the detector side. Indeed, a vulnerable perceptual model would certainly mislead the detection of an embedded watermark. The watermark strength can also be selected in an adaptive way depending on the reliability of transformed coefficients to increase the robustness of the watermark. It would be interesting to consider translation/rotation invariant transforms such as Fourier-Mellin transform which may provide an efficient way to tackle the robustness issue against geometric attacks. The modeling of the signal distortions due to commonly used image processing manipulations such as low pass filtering and compression could be useful to develop a robust detector that considers such alterations as uncertainties in the statistical model of the host image. Finally, as shown in this chapter, further investigation is required to develop an effective assessment measure of the imperceptibility of watermarks since there is no metric available in the literature so far that satisfies the need perfectly.

REFERENCES

Arnold, M., Schmucker, M., & Wolthusen, S. D. (2003). *Techniques and applications of digital watermarking and content protection.* Artech House Computer Security.

Barni, M., & Bartolini, F. (2004). *Watermarking systems engineering.* Marcel Dekker Inc.

Barni, M., Bartolini, F., Cappellini, V., & Piva, A. (1998). A DCT-domain system for robust image watermarking. *Signal Processing, 66,* 357-372.

Barni, M., Bartolini, F., De Rosa, A., & Piva, A. (2001). A new decoder for the optimum recovery of non additive watermarks. *IEEE Transactions on Image Processing, 10,* 755-765.

Birney, K. A., & Fischer, T. R. (1995). On the modeling of DCT and subband image data for compression. *IEEE Transactions on Image Processing, 4,* 186-193.

Blum, R. S. (1994). Asymptotically robust detection of known signals in non-additive noise. *IEEE Transactions on Information Theory, 40,* 1612-1619.

Calvagno, G., Ghirardi, C., Mian, G. A., & Rinaldo, R. (1997). Modeling of subband image data for buffer control. *IEEE Transactions on Circuits and Systems for Video Technology, 7,* 402-408.

Cheng, Q., & Huang, T. S. (2001a). An additive approach to transform-domain information hiding and optimum detection structure. *IEEE Transactions on Multimedia, 3,* 273-284.

Cheng, Q., & Huang, T. S. (2001b). Optimum detection of multiplicative watermarks using locally optimum decision rule. In *Proceedings of the IEEE International Conference Multimedia and Expo,* Tokyo, Japan.

Cheng, Q., & Huang, T. S. (2002). Optimum detection and decoding of multiplicative watermarks in DFT domain. In *Proceedings of the IEEE International Conference on Acoustic, Speech, and Signal Processing,* Orlando, FL.

Cheng, Q., & Huang, T. S. (2003). Robust optimum detection of transform domain multiplicative watermarks. *IEEE Transactions on Signal Processing, 51,* 906-924.

Cox, I. J., Kilian, J., Leighton, F. T., & Shamoon, T. (1997). Secure spread spectrum watermarking for multimedia. *IEEE Transactions on Image Processing, 6,* 1673-1687.

Cox, I. J., Miller, M. L., & Bloom, J. A. (2002). *Digital watermarking.* Academic Press.

Do, M. N., & Vetterli, M. (2002). Wavelet-based texture retrieval using generalized Gaussian density and Kullbackleibler distance. *IEEE Transactions on Image Processing, 11,* 146-158.

Elmasry, G. F., & Shi, Y. Q. (1999). Maximum likelihood sequence decoding of digital image watermarks. In *Proceedings of SPIE Security and Watermarking of Multimedia Contents* (pp. 425-436). San Jose, CA.

Ferguson, T. (1967). *Mathematical statistics: A decision theoretical approach.* Academic Press.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Hernandez, J. R., Amado, M., & Perez-Gonzalez, F. (2000). DCT-domain watermarking techniques for still images: Detector performance analysis and a new structure. *IEEE Transactions on Image Processing, 9,* 55-68.

Jayant, N., Johnston, J., & Safranek, R. (1993). Signal compression based on models of human perception. *Proceedings of the IEEE, 81,* 1385-1422.

Khelifi, F., Bouridane, A., & Kurugollu, F. (2006). On the optimum multiplicative watermark detection in the transform domain. In *Proceedings of the IEEE International Conference on Image Processing.*

Kwon, S.-G., Lee, S.-H., Kwon, K.-K., Kwon, K.-R., & Lee, K. (2002). Watermark detection algorithm using statistical decision theory. In *Proceedings of the IEEE International Conference on Multimedia and Expo,* Lausanne, Switzerland.

Langelaar, G. C., Setyawan, I., & Langendijk, R. L. (2000). Watermarking digital image and video data. *IEEE Signal Processing Magazine, 17,* 20-46.

Mallat, S. G. (1989). A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 11,* 674-693.

Miller, J. H., & Thomas, J. B. (1972). Detectors for discrete-time signals in non-Gaussian noise. *IEEE Information Theory, 18,* 241-250.

Ng, T. M., & Grag, H. K. (2004). Wavelet domain watermarking using maximum likelihood detection. In *Proceedings of SPIE Conference on Security, Steganography, Watermarking Multimedia Contents, 5306,* San Jose, CA.

Ng, T. M., & Grag, H. K. (2005). Maximum likelihood detection in DWT domain image watermarking using Laplacian modeling. *IEEE Signal Processing Letters, 12,* 285-288.

Papoulis, A. (1991). *Probability, random variables, and stochastic processes.* McGraw-Hill.

Sharifi, K., & Leon-Garcia, A. (1995). Estimation of shape parameters for generalized Gaussian distributions in subband decomposition of video. *IEEE Transactions on Circuits and Systems for Video Technology, 5,* 52-56.

ADDITIONAL READING

Avciba, I., Sankur, B., & Sayood, K. (2002). Statistical evaluation of image quality measures. *Journal of Electronic Imaging, 11*(2), 206-223.

Barni, M., Bartolini, F., De Rosa, A., & Piva, A. (2003). Optimum decoding and detection of multiplicative watermarks. *IEEE Transactions on Signal Processing, 52*(4), 1118-1123.

Briassouli, A., & Strintzis, M. (2004). Optimal watermark detection under quantization in the transform domain. *IEEE Transactions on Circuits and Systems for Video Technology, 14*(12), 1308-1319.

Briassouli, M. A., & Strintzis, M. (2004). Locally optimum nonlinearities for DCT watermark detection. *IEEE Transactions on Image Processing, 13*(12), 604-1617.

Huang, X., & Zhang, B. (2007). Statistically robust detection of multiplicative spread-spectrum watermarks. *IEEE Transactions on Information Forensics and Security, 2*(1), 1-13.

Karybali, I. G., & Berberidis, K. (2006). Efficient spatial image watermarking via new perceptual masking and blind detection schemes. *IEEE Transactions on Information Forensics and Security, 1*(2), 256-274.

Khelifi, F., Bouridane, F., Kurugollu, F., & Thompson, A. I. (2005). An improved wavelet-based image watermarking technique. In *Proceedings of IEEE International Conference on Advanced Video and Signal Based Surveillance,* Italy.

Lu, C.-S. (2004). *Multimedia security: Steganography and digital watermarking techniques for protection of intellectual property.* Hershey, PA: Idea Group.

Noorkami, M., & Mersereau, R. M. (2007). A framework for robust watermarking of H.264-encoded video with controllable detection performance. *IEEE Transactions on Information Forensics and Security, 2*(1), 14-23.

ENDNOTES

1. The DCT is actually the real part of the DFT. Approximately, it shares the same invariance properties with the DFT to small rotations and translations.
2. To make the formulation of the detection problem as general as possible, the pdf is not specified.
3. Taylor's series can be used up to the first order to approximate the expression in the integral about the point $\lambda \approx 0$.