Design and Statistical Analysis of a Hash-Aided Image Watermarking System

Jillian Cannons, *Student Member, IEEE,* and Pierre Moulin, *Fellow, IEEE*

***Abstract*—This paper develops a joint hashing/watermarking scheme in which a short hash of the host signal is available to a detector. Potential applications include content tracking on public networks and forensic identification. The host data into which the watermark is embedded are selected from a secret subset of the full-frame discrete cosine transform of an image, and the watermark is inserted through multiplicative embedding. The hash is a binary version of selected original image coefficients. We propose a maximum likelihood watermark detector based on a statistical image model. The availability of a hash as side information to the detector modifies the posterior distribution of the marked coefficients. We derive Chernoff bounds on the receiver operating characteristic performance of the detector. We show that host–signal interference can be rejected if the hash function is suitably designed. The relative difficulty of an eavesdropper's detection problem is also determined; the eavesdropper does not know the secret key used. Monte Carlo simulations are performed using photographic test images. Finally, various attacks on the watermarked image are introduced to study the robustness of the derived detectors. The joint hashing/watermarking scheme outperforms the traditional "hashless" watermarking technique.**

***Index Terms*—Authentication, Chernoff bounds, content-based retrieval, detection theory, eavesdropping, image hashing, image watermarking, likelihood ratio test.**

I. Introduction

THE DESIRE and ability to hide information without invoking suspicion have been present in society for thousands of years. Throughout generations, the techniques used to accomplish these covert goals have varied and, with the current prevalence of digital multimedia data, these methods continue to evolve. In particular, digital watermarking is a current technique which offers a means by which information can be inserted into digital data. To assist in the watermark embedding process, a key is often utilized. For example, the key could specify the location within the digital content at which the watermark is to be inserted. In many applications, invisible watermarking is employed, causing the watermark to be imperceptible in the host data. However, the watermark should also be resistant to attacks, creating a tradeoff between invisibility and robustness. The remaining element of the watermarking process is a watermark detector, which assesses whether or not an input object is watermarked. To maximize detection efficiency, it is beneficial to use the original content in making the decision. This framework is termed *private watermarking* and is rather expensive in terms of storage and computation. Consequently, *blind* or *public watermarking* is also possible, where no portion of the original data is present at the detector.

A common application of the watermarking process is within a content protection system. In such a setting, the owner of an original digital object desires the ability to manage its distribution. Increasingly often, the case is being made that, indeed, this responsibility should lie with the content owner or provider, since manufacturers of media players have no commercial interest in implementing security solutions as part of their devices [1]. To achieve the goal of the protection system, watermarks can be utilized to distinguish unauthorized and authorized instances of the content. A typical content protection system features a large number of audio–visual objects, users, and secret keys. Thus, subversion attempts beyond traditional signal processing attacks are a concern.

1) If multiple images are marked using the same key, traditional blind watermarking schemes display security weaknesses to attacks such as the *Holliman-Memon attack* [2]. To overcome attempts of this nature, different keys should be utilized for different images.
2) *Copy attacks*, as studied by Kutter *et al.* [3], in which a user illegally embeds a watermark derived from one image into a new image, can be attempted to compromise the integrity of the system.
3) If the attacker has access to a detector, repeated queries can be made, thereby permitting the development of a successive approximation strategy (*sensitivity analysis attack* [4]) to modify the content such that it is outside the acceptance region.
4) If preloaded, static keys are utilized, any breach of security in the system is inherently difficult to repair, leaving the system open to further attacks.

In order to combat these attacks, it is desirable to have image-dependent keys and a detector that either is operated by the content provider or that at least requires communication with the content provider. Such desirable features are present in a new generation of watermarking systems that allow content tracking and forensic tracking and identification [5], [6]. In a typical content-tracking application, the content provider watermarks each file sold with a digital signature establishing ownership. The

Manuscript received December 16, 2002; revised January 7, 2004. This work was supported in part by NSF Grants CDA-96-24396 and CCR-02-08809. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Bruno Carpentieri.

J. Cannons was with the Beckman Institute and Electrical and Computer Engineering Department, University of Illinois, Urbana, IL 61801 USA. She is now with the Electrical and Computer Engineering Department, University of California at San Diego, La Jolla, CA 92093 USA (e-mail: jcannons@ucsd.edu).

P. Moulin is with the Beckman Institute and Electrical and Computer Engineering Department, University of Illinois, Urbana, IL 61801 USA (e-mail: moulin@ifp.uiuc.edu).

Digital Object Identifier 10.1109/TIP.2004.834660

content provider also scans public networks to detect the possible presence of unauthorized copies.

Applications, such as content protection systems, motivate the development of reliable watermarking schemes which adhere to these guidelines. A preferable alternative to the expenses of private watermarking would be a process in which the detector need only store a short hash of the original image, where a hash is a function of the original data and a cryptographic key, producing a *secret, simplified representation of the content*. This solution will be referred to as *hash-aided watermarking* and may be thought of as an intermediary between private watermarking and public watermarking. In this context, the hash is very small, dramatically decreasing storage requirements relative to a private watermarking system. Perceptual image hashes have previously been used in applications such as authentication, registration, and retrieval. The presence of a hash effectively introduces image-dependent keys, as well as requires communication over a side channel between the content provider and the detector.

The concept of hash-aided watermarking has been discussed by Voyatzis and Pitas in 1999 [7] but has not received much attention in the literature, a recent exception being work by Roy and Chang [8] in the context of database watermarking. This void in the literature prompted the initiation of a quantitative performance study of hash-aided watermarking. A preliminary version of this work appeared in the first author's M.S. thesis [9]. This paper will demonstrate that, when combined with a statistically optimal detection test, the hash function may be designed to dramatically enhance detection performance and, in particular, offer host-signal rejection capabilities. The watermark embedding utilized in the development is multiplicative [10]; however, extensions to additive embedding are straightforward.

This paper first reviews the basic image hashing problem (Section II) and the statistical watermark detection problem (Section III). The new, hash-aided watermarking system is then formulated, and statistical signal detection theory is employed to derive an optimal watermark detector (Section IV). Performance bounds are derived to evaluate this detector (Section V). The difficulty of the watermark detection problem as seen from the perspective of an eavesdropper is also considered (Section VI). Attacks are briefly considered and are specialized to multiplicative white noise (Section VII). The problem of nuisance parameters at the detector is discussed (Section VIII). The watermarking system is evaluated using real-world test images (Section IX). From these analyses and experiments, conclusions are drawn regarding the effectiveness of the hash-aided watermarking framework (Section X).

Notation: Random variables are denoted by capital letters, and their individual values by lowercase letters. Boldface notation is used for vectors. The *probability density function* (pdf) of a random variable X is denoted by $p_X(x)$, and $P_X(\mathcal{A}) = \int_{\mathcal{A}} p_X(x)dx$ denotes the probability of a set $\mathcal{A}$ under this distribution. We denote by $1_{\{x\in\mathcal{A}\}}$ the indicator function of a set $\mathcal{A}$. The asymptotic equality relation $f(x) \sim g(x)$ as $x \to x_0$ means that $\lim_{x\to x_0}(f(x)/g(x)) = 1$. The notation for specific variables is summarized below.

$\mathbf{i_s}$	Original host image.
$\mathbf{i_x}$	Watermarked image.
$\mathbf{i_y}$	Received image.
$\mathbf{s} = \{s_i\}$	Original image coefficients.
$\mathbf{x} = \{x_i\}$	Watermarked image coefficients.
$\mathbf{w} = \{w_i\}$	Attacker's noise.
$\mathbf{y} = \{y_i\}$	Received image coefficients.
$\mathbf{h} = \{h_i\}$	Binary hash values.
h	Binary hash function.
$\tilde{s}_i, \tilde{x}_i, \tilde{w}_i, \tilde{y}_i$	Natural logarithm of s_i, x_i, w_i, y_i.
$\mathbf{k}$	Secret key known to embedder and detector.
$\mathbf{r}$	Pseudorandom sequence known to embedder and detector.
$\mathbf{m} = \{m_i\}$	Watermark vector.
$P_M(m)$	Empirical distribution of $\{m_i\}$.
ϵ	Watermark embedding strength.
$\mathcal{C}$	Secret subset of image coefficients used for watermarking.
$\mathcal{S}_h$	Subset of $\mathbb{R}$ associated with hash function h.
N_T	Total number of coefficients in public subset of image coefficients.
$N_{\mathcal{C}}$	Number of image coefficients in set $\mathcal{C}$.
N	Number of watermarked image coefficients.
ν	$N_{\mathcal{C}}/N_T$.
σ	Chernoff exponent.

II. Hashing

An image hash $h(\mathbf{i_s}, \mathbf{k})$ applied to an image $\mathbf{i_s}$ with a secret key $\mathbf{k}$ is a short binary string [11], [12]. To verify authenticity of a received image $\mathbf{i_y}$, one compares $h(\mathbf{i_y}, \mathbf{k})$ with the stored value $h^* = h(\mathbf{i_s}, \mathbf{k})$. The image hash function should have the following two properties.

1) *Robustness*: $h(\mathbf{i_y}, \mathbf{k}) \approx h(\mathbf{i_s}, \mathbf{k})$ if $\mathbf{i_y}$ is perceptually similar to $\mathbf{i_s}$.

2) *Resistance to Collisions*: The random strings $h(\mathbf{i_y}, \mathbf{k})$ and $h(\mathbf{i_s}, \mathbf{k})$ are nearly independent if $\mathbf{i_y}$ and $\mathbf{i_s}$ are unrelated.

Hence, the hash function should extract robust, randomized features of the image.

Authenticity may be assessed via the hypothesis test

$$\begin{aligned} H_0 &: \ \mathbf{i_y} \text{ is perceptually similar to } \mathbf{i_s} \\ H_1 &: \ \mathbf{i_y} \text{ is not perceptually similar to } \mathbf{i_s}. \end{aligned}$$

A typical decision rule is

$$\frac{1}{L} d_H \left(h(\mathbf{i_y}, \mathbf{k}), h^* \right) \underset{H_0}{\overset{H_1}{\gtrless}} \gamma \tag{2.1}$$

where d_H denotes Hamming distance, L is the length of the hash, and $\gamma \in [0, 1]$ is the threshold of the test; typical values are $L = 100$ and $\gamma = 0.75$.

III. Multiplicative Image Watermarking

This section first summarizes the standard multiplicative blind image watermarking problem [13]. A statistical image model is then described, and the corresponding likelihood ratio test for watermark detection is derived.

A. Basic Image Watermarking Problem

The basic image watermarking problem consists of an embedding process, an attack process, and a detection process. The N vector of host data into which the watermark is to be inserted $\mathbf{s} = \{s_i, 1 \leq i \leq N\}$ is derived from the host image $\mathbf{i_s}$. A common source of host data is image transform coefficients, and the watermark is inserted into a subset of these coefficients. The *discrete cosine transform* (DCT) is selected for use in this paper for simplicity, and the full-frame transform is employed to increase robustness against image resizing and other geometric attacks [14], [15]. A specific watermark vector $\mathbf{m} = \{m_i, 1 \leq i \leq N\}$ is to be embedded into the host vector. The multiplicative watermark embedding method generates each element of the watermarked data according to the formula [10]

$$x_i = s_i(1 + \epsilon m_i), \quad 1 \leq i \leq N \tag{3.1}$$

where ϵ is the strength of the watermark embedding. We assume without loss of generality that

$$\max_{1 \leq i \leq N} |m_i| = 1$$

then typically $\epsilon \approx 0.1$. Denote by $P_M(m)$ the *empirical distribution* of the $\{m_i\}$, namely

$$P_M(m) := \frac{1}{N} \sum_{i=1}^{N} 1_{\{m_i \leq m\}}, \quad -1 \leq m \leq 1. \tag{3.2}$$

For the commonly used binary symmetric distribution, we have $m_i \in \{\pm 1\}$ with equal frequency. Thus, $P_M(m) = 1/2$ for $-1 \leq m < 1$, and $P_M(m) = 1$ for $m = 1$. The watermarked data $\mathbf{x}$ are reinserted into the image to form a watermarked image $\mathbf{i_x}$.

With the watermark now embedded into the host data, the watermarked image is released into the public domain. The detector has no knowledge of the manipulations performed on the image, which could include simple image processing operations, or an attacker's attempts to remove the watermark.

When supplied with an image $\mathbf{i_y}$, the task of the detector is to determine whether or not the watermark vector $\mathbf{m}$ has been embedded. The detector produces a statistic indicating the degree of certainty that the given watermark is present. This statistic is then compared against a threshold to determine a yes or no result.

B. Host Data Modeling

To derive a detector based upon statistical detection theory, a realistic, probabilistic representation of the host data must be determined. First, the DCT coefficients are assumed to be independent. Then, the zero-mean *power exponential* (PE) distribution (also called the *generalized Gaussian distribution*) is commonly used to model the distribution of the DCT coefficients of an image [16]. The PE distribution contains two parameters $\alpha > 0$ and $\beta > 0$, with α relating to the variance and β relating to the heaviness of the distribution tails. The distribution itself is given by

$$p_S(s) = \frac{\beta}{2\alpha\Gamma\left(\frac{1}{\beta}\right)} \exp\left\{-\left|\frac{s}{\alpha}\right|^{\beta}\right\}, \quad s \in \mathbb{R} \tag{3.3}$$

where Γ is the Gamma function. The parameters α and β are frequency dependent; the value of β usually ranges between 1.5 and 2.2 [15], [16]. In this paper, we assume that the selected DCT coefficients can be grouped into frequency regions with homogeneous statistics, e.g., the 16 regions described by Barni *et al.* [15], [16].

A slightly better (and also often used) model is the Weibull distribution

$$p_S(s) = \frac{\beta}{\alpha}\left(\frac{s}{\alpha}\right)^{\beta-1} \exp\left\{-\left(\frac{s}{\alpha}\right)^{\beta}\right\}, \quad s > 0 \tag{3.4}$$

where $\alpha > 0$ and $\beta > 0$.

Similar results may be derived under both models by noting that they are exponential families [17]

$$p_\theta(s) = H(s) \exp\{\theta T(s) + \psi(\theta)\} \tag{3.5}$$

where $T(s) = -|s|^{\beta}$, $\theta = \alpha^{-\beta}$, and

$$\begin{aligned} \psi(\theta) &= \frac{1}{\beta}\ln\theta, \quad H(s) = \frac{\beta}{2\Gamma\left(\frac{1}{\beta}\right)} \;:\; \text{PE} \\ \psi(\theta) &= \ln\theta, \quad\;\; H(s) = \beta s^{\beta-1} \quad :\; \text{Weibull.} \end{aligned}$$

C. Statistical Detection for Basic Image Watermarking

Statistical decision theory can be used to formulate a watermark detector based upon a *binary hypothesis test*, indicating the presence or absence of a particular watermark. The *likelihood ratio test* is an optimal binary hypothesis test [17], and, thus, is employed in this paper. Watermark detectors using this technique have previously been developed, specifically for use with the Weibull pdf [13].

The detection problem is formulated as a choice between two hypotheses H_0 (the image does not contain the specific watermark) and H_1 (the image contains the specific watermark). Each hypothesis has associated with it a distribution for the data, respectively, $p_{\mathbf{Y},0}(\mathbf{y})$ and $p_{\mathbf{Y},1}(\mathbf{y})$, where $\mathbf{y}$ are the possibly watermarked coefficients. The likelihood ratio test takes the form

$$L(\mathbf{y}) = \frac{p_{\mathbf{Y},1}(\mathbf{y})}{p_{\mathbf{Y},0}(\mathbf{y})} \underset{H_0}{\overset{H_1}{\gtrless}} e^{\gamma} \tag{3.6}$$

where e^{γ} is the threshold of the likelihood ratio test.

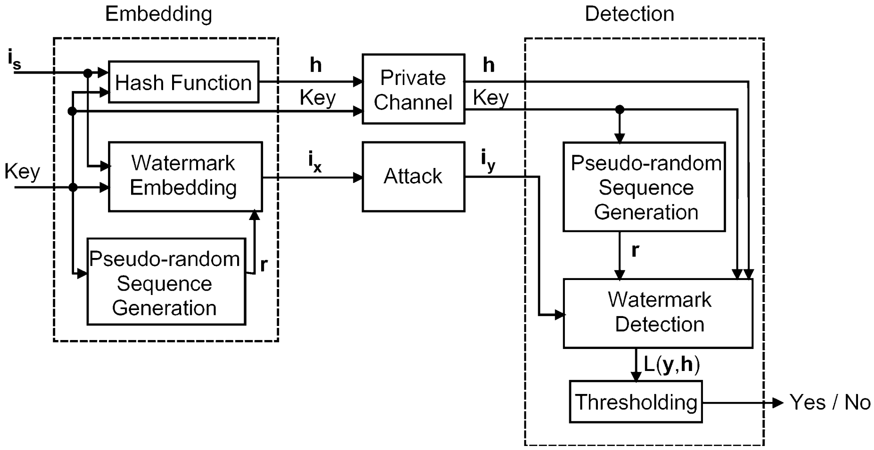


Fig. 1. General joint hashing/watermarking process.

The threshold value can be chosen based on a Neyman–Pearson hypothesis test [17].[1] The *probability of false alarm* P_{FA} (falsely detecting a watermark) is given by

$$\begin{aligned} P_{FA} &:= P[\text{choosing } H_1 \mid H_0 \text{ is true}] \\ &= P_0 [\ln L(\mathbf{y}) > \gamma] \\ &= \int_{\gamma}^{\infty} p_{L,0}(\ell) d\ell \end{aligned} \tag{3.7}$$

where $p_{L,0}(\ell)$ is the distribution of $\ln L(\mathbf{y})$ under H_0. Thus, if P_{FA} is specified, (3.7) can be solved for γ.

The likelihood ratio test detector is now specialized to the case when the PE or Weibull distribution is used to model the host coefficients. In this situation, the two hypotheses can be written as

$$\begin{aligned} H_0 &: \ y_i = s_i, & 1 \le i \le N \\ H_1 &: \ y_i = s_i(1 + \epsilon m_i), & 1 \le i \le N. \end{aligned}$$

Observe that $p_{Y_i,1}$ remains in the original family of distributions (PE or Weibull) for all values of $m_i \in [-1, 1]$. Using (3.5), the distributions under H_0 and H_1 are, respectively

$$p_{\mathbf{Y},0}(\mathbf{y}) = \prod_{i=1}^{N} H(y_i) \exp\left\{ - \left| \frac{y_i}{\alpha} \right|^{\beta} + \psi(\alpha^{-\beta}) \right\} \tag{3.8}$$

$$\begin{aligned} p_{\mathbf{Y},1}(\mathbf{y}) = \prod_{i=1}^{N} H(y_i) \exp\Bigg\{ &- \left| \frac{y_i}{\alpha(1 + \epsilon m_i)} \right|^{\beta} \\ &+ \psi\left(\alpha^{-\beta}(1 + \epsilon m_i)^{-\beta}\right) \Bigg\}. \end{aligned} \tag{3.9}$$

These distributions take a product form because the data $\{y_i\}$ are independent. The log-likelihood ratio is given by

$$\ln L(\mathbf{y}) = \sum_{i=1}^{N} \left(\psi\left((1+\epsilon m_i)^{-\beta}\right) + \left| \frac{y_i}{\alpha} \right|^{\beta} \left(1 - (1+\epsilon m_i)^{-\beta}\right) \right) \tag{3.10}$$

to be compared against the threshold γ. The detector that implements this test will be referred to as the *PE* or *Weibull detector*. For the typical case of small ϵ, a first-order Taylor series expansion of (3.10) around $\epsilon = 0$ yields

$$\ln L(\mathbf{y}) \sim \epsilon\beta \sum_{i=1}^{N} m_i \left(\left| \frac{y_i}{\alpha} \right|^{\beta} - \psi'(1) \right) \quad \text{as } \epsilon \to 0 \tag{3.11}$$

i.e., for both the PE and Weibull models, the test statistic is a correlation between the sequence $\{m_i\}$ and the transformed data $\{|y_i|^{\beta}\}$. We have $\psi'(1) = 1/\beta$ and $\psi'(1) = 1$ for PE and Weibull, respectively.

IV. Joint Image Hashing/Watermarking Problem

In this section, the basic watermarking problem is altered to include side information at the detector, resulting in a joint image hashing/watermarking scheme and a new likelihood ratio test. In Section III-A, the data available at the detector are the possibly watermarked image $\mathbf{i_y}$ and the watermark vector $\mathbf{m}$. This system can be modified to include side information $\mathbf{h}$ at the detector to increase system performance, as shown in Fig. 1. Note that $\mathbf{h} = 0$ yields the basic system of Section III-A, commonly referred to as *blind watermarking*. Furthermore, when $\mathbf{h} = \mathbf{i_s}$, the original, unmarked image is present at the detector, and the watermarking system is said to be *private*. Any intermediate case (such as in this paper, where $\mathbf{h} = h(\mathbf{i_s}, \mathbf{k})$ is a hash) can be thought of as *semi-blind watermarking*.

A. System Description

Our DCT-based joint hashing/watermarking system is diagrammed in Fig. 2. Its main features are described below.

1) Candidate Set: The DCT coefficients to be watermarked come from a publicly known region of size N_T, as depicted by the trapezoidal region in Fig. 3. To increase the security of the watermarking system, a secret subset of the DCT coefficients is defined, and forms a set $\mathcal{C}$ of $N_{\mathcal{C}}$ *candidate locations* for embedding. These locations are known by both the embedder and the detector. Let $\nu = N_{\mathcal{C}}/N_T$ denote the fraction of coefficients included in the candidate set. Note that in the majority of current

[1] While optimality in the Neyman–Pearson sense may require the use of a randomized likelihood ratio test, the family of nonrandomized tests (3.6) is optimal in all problems considered in this paper.

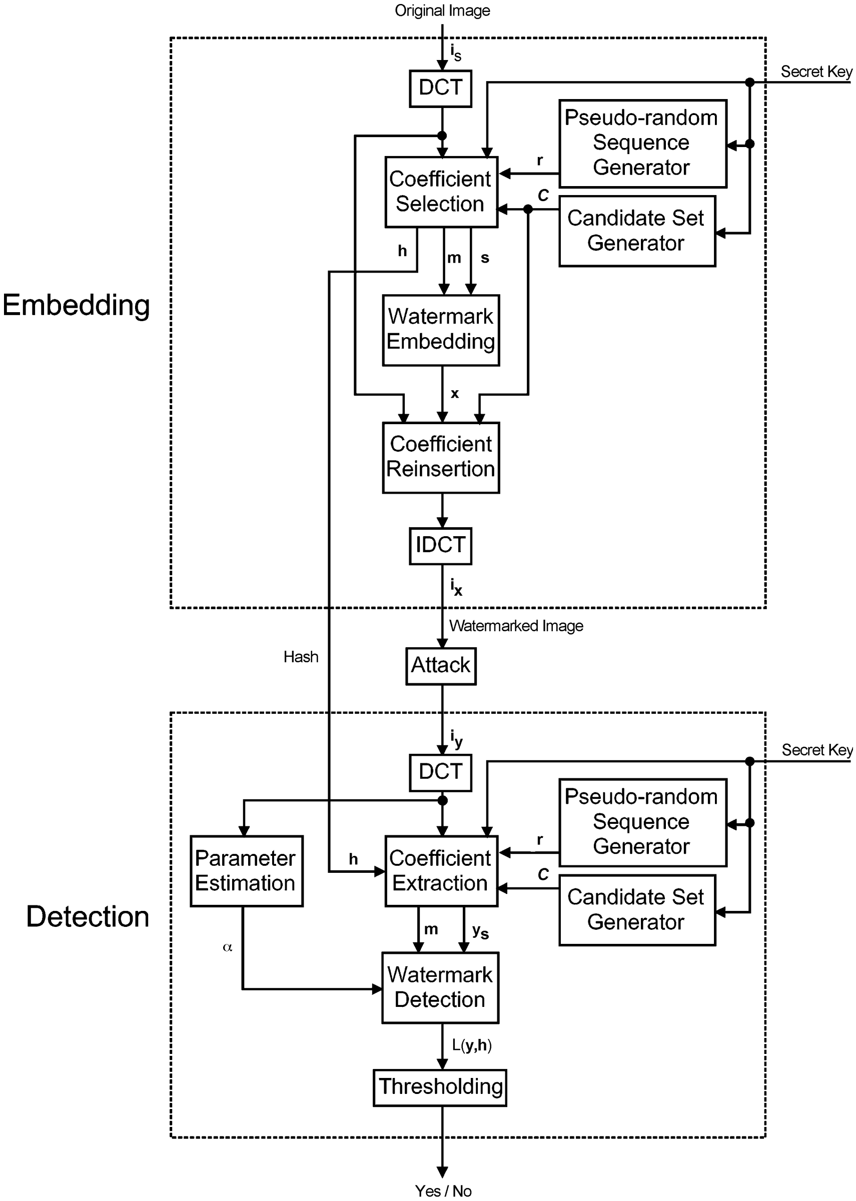


Fig. 2. Our DCT-based hashing/watermarking system.

watermarking schemes $\nu = 1$, indicating that no secret candidate set is employed.

2) Hash Function: In our joint image hashing/watermarking system, a hash vector of length $N_{\mathcal{C}}$ is utilized, where each h_i takes the form of a 1-bit hash of a corresponding original image coefficient s_i, $i \in \mathcal{C}$. Specifically

$$h_i = 1_{\{s_i \in \mathcal{S}_h\}}, \quad 1 \leq i \leq N_{\mathcal{C}}$$

where $\mathcal{S}_h$ is a subset of $\mathbb{R}$, typically a union of intervals

$$\mathcal{S}_h = \bigcup_{j=1}^{j_{\max}} [a_j, b_j] \tag{4.1}$$

such that $a_1 < b_1 < a_2 < b_2 < \cdots < b_{j_{\max}}$. The N coefficients for which $h_i = 1$ form the host data set $\mathbf{s}$ and will be watermarked. Fig. 3 contains a conceptual representation of the different sets involved in the construction of $\mathbf{s}$.

One may think of $\{h_i\}$ as a bitplane in a particular binary decomposition of the original image coefficients $\{s_i\}$. The expected fraction of DCT coefficients in the set $\mathcal{S}_h$ is $P_S(\mathcal{S}_h) = \int_{\mathcal{S}_h} p_S(s)ds$ (in our original system design [9], $\mathcal{S}_h$ was a semi-infinite interval $\mathcal{S}_h = [\delta, \infty)$, where δ is a significance threshold). To increase the resistance of the scheme to eavesdroppers, the set $\mathcal{S}_h$ should depend on the secret key $\mathbf{k}$, in which case $P_S(\mathcal{S}_h)$ need also be averaged over $\mathbf{k}$. Referring to our conditions in Section II for a good image hashing system, we need:

1) sufficiently wide intervals $[a_j, b_j]$ so that similar images map to similar hash values;
2) sufficiently large a_1 to ensure that each feature is perceptually significant;
3) small $[P_S(\mathcal{S}_h)]^N$ to provide adequate resistance against collisions.

Example 4.1: The specific choice used in our experiments is as follows. Choose a significance threshold δ for the magnitude

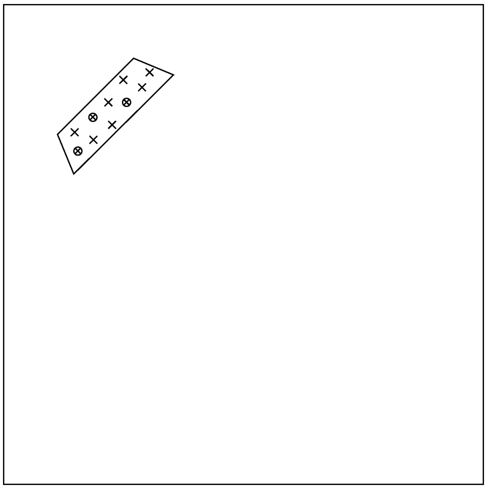


Fig. 3. Conceptual formation of a length N host vector $\mathbf{s}$. The host DCT coefficients are drawn from the trapezoidal region, of size N_T. The locations marked × represent DCT coefficients in the candidate set $\mathcal{C}$ (of size $N_{\mathcal{C}} = 10$ in this example). The locations marked $\otimes$ indicate elements of $\mathcal{C}$ whose magnitudes are in the set $\mathcal{S}_h$ $(N = 3)$.

of the DCT coefficients to be watermarked. Select $\Delta > 0$ and $j_{\max} \geq 1$. Select a_1 randomly in the interval $[\delta, \delta(1+\Delta)]$, according to a uniform distribution with a generating seed known to both the encoder and the detector. For all j, let $b_j = a_j(1+\Delta)$ and $a_{j+1} = b_j(1+\Delta)$. Typical parameter values are $\delta \approx \alpha$, $\Delta = 0.3$, and $j_{\max} = 3$. For the Weibull distribution with $\alpha = 0.1135$ and $\beta = 2$, this choice of parameters leads to $P_S(\mathcal{S}_h) \approx 0.3$ (averaging over $\mathbf{k}$).

3) Operating Methods: Two operational strategies are possible for the joint hashing/watermarking system.

1) The parameters ν and $N_{\mathcal{C}} = \nu N_T$ are fixed. The number $N = \#\{i \in \mathcal{C} : h_i = 1\}$ is, thus, image dependent and key dependent. These dependencies influence the construction of the watermark itself. One specification technique is for the embedder to be given a sequence $\mathbf{r}$ (where the r_i follow the distribution $P_M(m)$) of the same size $(N_{\mathcal{C}})$ as the candidate set, $\mathcal{C}$. Each element of this sequence is associated with a coefficient in $\mathcal{C}$. Those elements of $\mathbf{r}$ coinciding with the coefficients of $\mathbf{s}$ are denoted by $\mathbf{m}$. Thus, $\mathbf{m}$ is of length N and forms the actual watermark which will be embedded in the image. With the specification of the key used to generate the candidate set, the key used to generate the set $\mathcal{S}_h$, the hash vector $\mathbf{h}$, the sequence $\mathbf{r}$, and the possibly watermarked image $\mathbf{i_y}$, the detector is able to extract the appropriate N vector of coefficients $\mathbf{y}$, which is a sufficient statistic for the detection test.
2) The parameter N is fixed but $N_{\mathcal{C}}$ is image dependent and key dependent. The embedder implementation is as follows. A length N_T vector is obtained by a pseudorandom scrambling of the original N_T coefficients. The binary hash sequence $\{h_i\}$ is obtained by successively applying the function $1_{\{s \in \mathcal{S}_h\}}$ to the components of this scrambled vector, starting from the first component. This process is terminated when the Nth "1" in the sequence $\{h_i\}$ is obtained. The length of the hash vector at this point is $N_{\mathcal{C}}$. The watermark $\mathbf{m}$ is then embedded in the N coefficients for which $h_i = 1$. The watermark detector knows N, the key used for the scrambling algorithm, the key used to generate $\mathcal{S}_h$, and the hash vector. It can, thus, extract $\mathbf{y}$, the vector of N coefficients for which $h_i = 1$.

The first method is slightly more difficult to analyze statistically, hence the second method is used throughout this paper. Note that the second method fails in the case when fewer than N coefficients in the original size N_T dataset have magnitude in the set $\mathcal{S}_h$. Under our statistical assumptions on the coefficients, this probability is guaranteed to be very small if $N \ll N_T P_S(\mathcal{S}_h)$.

B. Statistical Detection for Joint Image Hashing/Watermarking

The appropriate likelihood ratio test must now be derived, incorporating the side information $\mathbf{h}$ in addition to the possibly watermarked coefficients $\mathbf{y}$. The pair $(\mathbf{y}, \mathbf{h})$ has a joint distribution $p_{\mathbf{Y},\mathbf{H},0}(\mathbf{y}, \mathbf{h})$ under H_0 (no watermark present) and a joint distribution $p_{\mathbf{Y},\mathbf{H},1}(\mathbf{y}, \mathbf{h})$ under H_1 (watermark present). The likelihood ratio is now written as

$$L(\mathbf{y}, \mathbf{h}) = \frac{p_{\mathbf{Y},\mathbf{H},1}(\mathbf{y}, \mathbf{h})}{p_{\mathbf{Y},\mathbf{H},0}(\mathbf{y}, \mathbf{h})} = \frac{p_{\mathbf{Y}|\mathbf{H},1}(\mathbf{y}|\mathbf{h})}{p_{\mathbf{Y}|\mathbf{H},0}(\mathbf{y}|\mathbf{h})} \overset{H_1}{\underset{H_0}{\gtrless}} e^{\gamma} \tag{4.2}$$

where e^{γ} is the threshold of the test.

C. Pulse-Modulated Distribution

Since the PE and Weibull distributions are well suited to modeling the DCT coefficients of an image, either one is a tempting choice for modeling the coefficients when designing the detector. However, in the joint hashing/watermarking scheme, these distributions are *not* representative of the coefficients selected for watermarking. Here, only the coefficients with magnitudes in the set $\mathcal{S}_h$ are watermarked, as specified by the hash. Thus, a new, *pulse-modulated* (PM) distribution is derived to correctly represent the actual posterior distribution of the chosen coefficients.

Denote by $p_S(s)$ the prior distribution of an unmarked coefficient s and consider the binary hash

$$h = 1_{\{s \in \mathcal{S}_h\}}.$$

After observing $h = 1$, the posterior distribution of s becomes

$$p_{PM}(s) := p(s|h=1) = \frac{1}{P_S(\mathcal{S}_h)} p_S(s) 1_{\{s \in \mathcal{S}_h\}}. \tag{4.3}$$

The PM distribution (4.3) is shown in Fig. 4 for the case when $p_S(s)$ is a PE distribution.

To evaluate the likelihood ratio (4.2), we derive

$$\begin{aligned} p_{\mathbf{Y}|\mathbf{H},0}(\mathbf{y}|\mathbf{h}) &= \prod_{i=1}^{N} p_{PM}(y_i) \\ &= \prod_{i=1}^{N} \frac{1}{P_S(\mathcal{S}_h)} p_S(y_i) 1_{\{y_i \in \mathcal{S}_h\}} \end{aligned} \tag{4.4}$$

and

$$\begin{aligned} &p_{\mathbf{Y}|\mathbf{H},1}(\mathbf{y}|\mathbf{h}) \\ &= \prod_{i=1}^{N} \frac{1}{1+\epsilon m_i} p_{PM}\left(\frac{y_i}{1+\epsilon m_i}\right) \\ &= \prod_{i=1}^{N} \frac{1}{P_S(\mathcal{S}_h)} \frac{1}{1+\epsilon m_i} p_S\left(\frac{y_i}{1+\epsilon m_i}\right) 1_{\left\{\frac{y_i}{1+\epsilon m_i} \in \mathcal{S}_h\right\}}. \end{aligned} \tag{4.5}$$

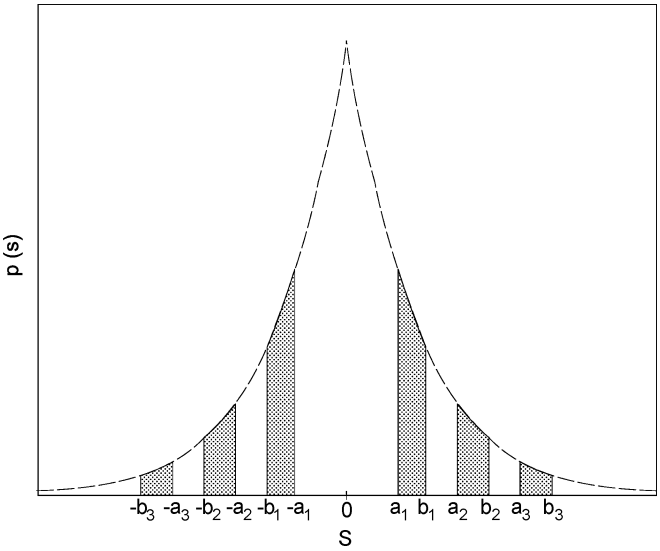


Fig. 4. Pulse-modulated PE distribution whose pulse support set $\mathcal{S}_h$ is a union of six intervals.

Hence, the log-likelihood ratio is given by

$$\ln L(\mathbf{y},\mathbf{h}) = \sum_{i=1}^{N} \ln L_i(y_i,h_i) \tag{4.6}$$

where

$$\ln L_i(y_i,h_i) = \begin{cases} \psi\big((1+\epsilon m_i)^{-\beta}\big) \\ \quad +\left|\frac{y_i}{\alpha}\right|^{\beta}\big(1-(1+\epsilon m_i)^{-\beta}\big), & \text{if } y_i \in \mathcal{S}_h \text{ and} \\ & \quad y_i \in (1+\epsilon m_i)\mathcal{S}_h \\ \infty, & \text{if } y_i \notin \mathcal{S}_h \text{ and} \\ & \quad y_i \in (1+\epsilon m_i)\mathcal{S}_h \\ -\infty, & \text{if } y_i \in \mathcal{S}_h \text{ and} \\ & \quad y_i \notin (1+\epsilon m_i)\mathcal{S}_h. \end{cases} \tag{4.7}$$

These results are quite intuitive. In the first case, when the set conditions $y_i \in \mathcal{S}_h$ and $y_i \in (1+\epsilon m_i)\mathcal{S}_h$ are satisfied, $\ln L_i(y_i,h_i)$ is the same as under the PE and Weibull models. When one of the set membership conditions is met but the other is not, one of the hypotheses is impossible, which is indicated by the presence of infinities in the statistic. Note that the probability of having the range conditions fail under both hypotheses is zero. Fig. 5 illustrates the three regions present in the decision statistic when $\mathcal{S}_h$ is a semi-infinite interval, $\mathcal{S}_h = [\delta,\infty)$. The likelihood ratio test based upon (4.6) will be referred to as the *PM detector*.

D. Host-Signal Rejection

In conventional spread–spectrum blind watermarking systems, the host signal is a substantial source of noise which adversely affects detection performance. In the case of Gaussian hosts for instance, the probability-of-error exponents are inversely proportional to the variance of the host [18].

In our hash-aided system, the set $\mathcal{S}_h$ can be advantageously designed to reduce host-signal interference. Consider Fig. 6, which shows the PM distribution $p_{Y,0}(y) = p_{PM}(y)$ versus the scaled PM distribution $p_{Y,1}(y) = (1/(1+\epsilon m))p_{Y,0}(y/(1+\epsilon m))$ for a fixed value of m. Roughly speaking, the ability to discriminate between $p_{Y,0}$ and $p_{Y,1}$ is enhanced when the overlap of these two pdfs is reduced, which occurs when Δ is decreased. However, small Δ also results in a lack of robustness against attacks, so the choice of Δ is a tradeoff. The relative overlap of $p_{Y,0}$ and $p_{Y,1}$ depends weakly on the variance of the host signal's pdf, which could even be infinite. A more precise study of the probability of error appears in the next section.

The ability to reduce host–signal interference is reminiscent of the well-known *quantization index modulation* (QIM) technique, which has a similar property. The hash information tells the PM detector that the host signal belongs to a particular subset of signal space (a union of disjoint cells), and the detector tests for the presence of the watermark conditioned on that information. In contrast, the QIM embedding scheme preconditions the host signal so as to concentrate its distribution on a subset of signal space (typically a union of disjoint cells); the QIM detector must then infer whether the *received signal* belongs to a particular subset of signal space (also a union of disjoint cells). It is worth noting that the hash-aided watermarking scheme achieves host-signal rejection without preconditioning the host signal.

V. Chernoff Bounds

Binary hypothesis testing using a likelihood ratio test forms a conceptually simple statistical detector. However, detector performance analysis can often be quite complex. Thus, bounds on the *probability of detection* P_D (correct detection of a watermark) and the probability of false alarm P_{FA} are often sought. Here, we consider Chernoff bounds [17], [19], which are large-deviation bounds and are asymptotically tight in the exponent. The bound provided for P_{FA} is an upper bound, while the bound provided for P_D is a lower bound. These two bounds yield a useful lower bound on the *receiver operating characteristic* (ROC) curve (P_D versus P_{FA}, parameterized by the test threshold γ). This bound may be used to evaluate detector performance without the need for the large Monte Carlo simulations which are required when error probabilities are very low. Moreover, the Chernoff bounds may be used to appropriately design the hash function. In this section, Chernoff bounds are formulated for the general case of multiplicative watermarking, and then specialized to the likelihood ratio test detectors based on exponential and PM distributions. Finally, the bounds are evaluated numerically.

A. Chernoff Bounds for Multiplicative Watermarking

As noted earlier, when multiplicative watermarking is employed to insert a watermark into a set of host data, the distribution of each element y_i of the output under H_1 is a scaled version of the corresponding distribution under H_0

$$p_{Y_i,1}(y_i) = \frac{1}{1+\epsilon m_i} p_{Y,0}\left(\frac{y_i}{1+\epsilon m_i}\right). \tag{5.1}$$

This property allows Chernoff bounds to be constructed in a general sense, and later specialized to individual distributions.

The *Chernoff distance* between $p_{\mathbf{Y},0}$ and $p_{\mathbf{Y},1}$ is defined as [17]

$$D(\sigma, p_{\mathbf{Y},0}, p_{\mathbf{Y},1}) := -\ln \int p_{\mathbf{Y},0}^{1-\sigma}(\mathbf{y}) p_{\mathbf{Y},1}^{\sigma}(\mathbf{y}) d\mathbf{y} \geq 0 \tag{5.2}$$

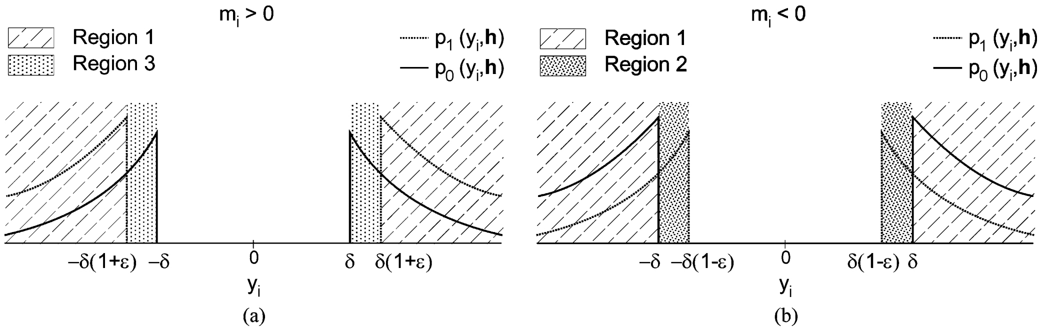


Fig. 5. Regions utilized in the likelihood ratio test statistic for the PM distribution with $\mathcal{S}_h = [\delta, \infty)$ (a) $m_i > 0$ and (b) $m_i < 0$.

for all $\sigma \in (0,1)$. The tightest bounds on P_{FA} and $P_{\text{miss}} = 1 - P_D$ are obtained when $\sigma = \sigma^*$ maximizes the function $\sigma\gamma + D(\sigma, p_{\mathbf{Y},0}, p_{\mathbf{Y},1})$. Due to the conditional independence of the $\{Y_i\}$, the Chernoff distance is additive over the components i

$$D(\sigma, p_{\mathbf{Y},0}, p_{\mathbf{Y},1}) = \sum_{i=1}^{N} D(\sigma, p_{Y,0}, p_{Y_i,1}). \tag{5.3}$$

Using (5.1)–(5.3), the Chernoff distance can be written in terms of $p_{Y,0}$

$$\begin{aligned} D(\sigma, p_{\mathbf{Y},0}, p_{\mathbf{Y},1}) &= \sum_{i=1}^{N} \mathcal{D}\left(p_{Y,0}; \sigma; \frac{1}{1+\epsilon m_i}\right) \\ &= N \int_{-1}^{1} \mathcal{D}\left(p_{Y,0}; \sigma; \frac{1}{1+\epsilon m}\right) \\ &\quad \times dP_M(m) \end{aligned} \tag{5.4}$$

where the nonnegative functional

$$\mathcal{D}(p_{Y,0}; \sigma, z) := -\ln \int p_{Y,0}^{1-\sigma}(y) z^{\sigma} p_{Y,0}^{\sigma}(yz), dy \tag{5.5}$$

is introduced to simplify the notation, and the distribution $P_M(m)$ is defined in (3.2). The value $\sigma = \sigma^*$ that maximizes $\sigma\gamma + D(\sigma, p_{\mathbf{Y},0}, p_{\mathbf{Y},1})$ is found numerically by solving a convex programming problem [17]. Note that $\mathcal{D}(p_{Y,0}; \sigma, 1/z) = \mathcal{D}(p_{Y,0}; 1-\sigma, z)$ and that $\mathcal{D}(p_{Y,0}; \sigma, z) + \mathcal{D}(p_{Y,0}; 1-\sigma, z) \leq 2\mathcal{D}(p_{Y,0}; 1/2, z)$ by concavity of the Chernoff distance with respect to σ. Hence, the optimal Chernoff exponent is $\sigma^* \sim 1/2$ as $\epsilon \to 0$ if the distribution P_M is symmetric around $m = 0$.

The Chernoff bounds on the various probabilities of error are [17], [19]

$$P_{FA} \leq \exp\{-\sigma\gamma - D(\sigma, p_{\mathbf{Y},0}, p_{\mathbf{Y},1})\} \tag{5.6}$$

$$\begin{aligned} 1 - P_D &= P_{\text{miss}} \\ &\leq \exp\{-(\sigma-1)\gamma - D(\sigma, p_{\mathbf{Y},0}, p_{\mathbf{Y},1})\} \end{aligned} \tag{5.7}$$

for all values of γ between $-D(p_{\mathbf{Y},0} \| p_{\mathbf{Y},1})$ and $D(p_{\mathbf{Y},1} \| p_{\mathbf{Y},0})$, where $D(p\|q)$ denotes Kullback–Leibler divergence from a distribution p to another distribution q. Assuming that the empirical distribution P_M in (3.2) converges to a limit, the Chernoff bounds are tight in the exponent as $N \to \infty$ for $\sigma = \sigma^*$ [17], [19]. The decision threshold γ is generally chosen to be proportional to $N: \gamma = N\bar{\gamma}$. Then, it follows from (5.4), (5.6), and (5.7) that P_{FA} and P_{miss} vanish exponentially with N, with respective rates

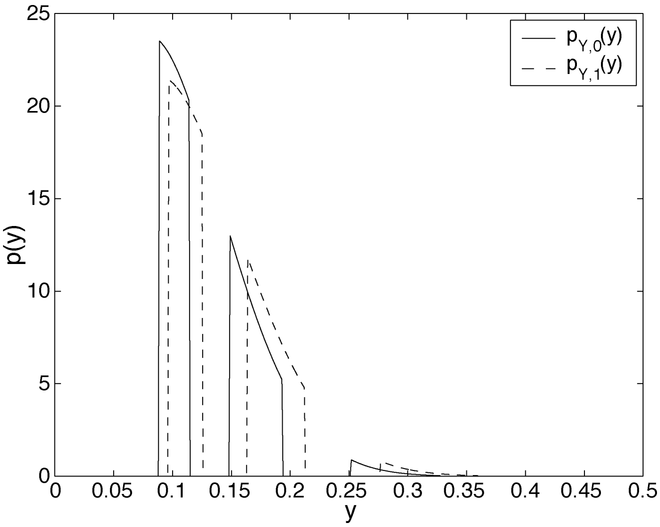


Fig. 6. Unmarked and positively marked PM-Weibull distributions.

$$\begin{aligned} E_F &:= -\lim_{N\to\infty} \frac{1}{N} \ln P_{FA} \\ &= \sigma^*\bar{\gamma} + \int_{-1}^{1} \mathcal{D}\left(p_{Y,0}; \sigma^*; \frac{1}{1+\epsilon m}\right) dP_M(m) \end{aligned} \tag{5.8}$$

$$\begin{aligned} E_M &:= -\lim_{N\to\infty} \frac{1}{N} \ln P_{\text{miss}} \\ &= (\sigma^* - 1)\bar{\gamma} + \int_{-1}^{1} \mathcal{D}\left(p_{Y,0}; \sigma^*; \frac{1}{1+\epsilon m}\right) \\ &\quad \times dP_M(m). \end{aligned} \tag{5.9}$$

The Chernoff bounds may now be specialized for a specific modeling distribution simply by substituting the desired $p_{Y,0}$ into the functional $\mathcal{D}$ defined in (5.5).

B. Exponential Distributions

For any exponential distribution of the form (3.5), direct evaluation of the Chernoff integral yields

$$D(\sigma, p_{\theta_0}, p_{\theta_1}) = \psi(\sigma\theta_0 + (1-\sigma)\theta_1) - \sigma\psi(\theta_0) - (1-\sigma)\psi(\theta_1). \tag{5.10}$$

The distributions $p_{Y,0}$ and $p_{Y_i,1}$ for our likelihood ratio test may be written as

$$p_{Y,0}(y_i) = p_{\theta_0}(y_i) \quad \text{and} \quad p_{Y_i,1}(y_i) = p_{\theta_1}(y_i)$$

with $\theta_0 = \alpha^{-\beta}$ and $\theta_1 = \theta_0(1+\epsilon m_i)^{-\beta}$. Define

$$\phi(\sigma, u) := \ln(\sigma + (1-\sigma)e^u) - (1-\sigma)u. \tag{5.11}$$

For the Weibull distribution, we have $\psi(\theta) = \ln\theta$, and so (5.10) becomes

$$D\left(\sigma, p_{\theta_0}, p_{\theta_1}\right) = \phi\left(\sigma, \ln\frac{\theta_1}{\theta_0}\right) \tag{5.12}$$

with $\ln(\theta_1/\theta_0) = -\beta\ln(1+\epsilon m_i)$. The Chernoff distance (5.4) becomes

$$D(\sigma, p_{\mathbf{Y},0}, p_{\mathbf{Y},1}) = N\int_{-1}^{1}\phi\left(\sigma, -\beta\ln(1+\epsilon m)\right)dP_M(m). \tag{5.13}$$

For the PE distribution, we have $\psi(\theta) = (1/\beta)\ln\theta$, and so (5.10) becomes

$$D\left(\sigma, p_{\theta_0}, p_{\theta_1}\right) = \frac{1}{\beta}\phi\left(\sigma, \ln\frac{\theta_1}{\theta_0}\right). \tag{5.14}$$

C. PM Distributions

In this subsection, we find the Chernoff distance for PM distributions, which can be derived from the Chernoff distance associated with each individual pulse. Assume the minimum-pulse-width condition $a_j < b_j(1-\epsilon)$ for $1 \le j \le j_{\max}$. Let $\pi_j = P_S([a_j, b_j])$ denote the probability of pulse j ($1 \le j \le j_{\max}$) and

$$p_{0,j}(y) = \frac{p_{Y,0}(y)}{\pi_j}1_{\{y\in[a_j,b_j]\}}$$

denote the probability of y under H_0, conditioned on pulse j being used. For $1/(1+\epsilon) \le z \le 1/(1-\epsilon)$, the Chernoff distance (5.5) is given by

$$\begin{aligned}\mathcal{D}(p_{Y,0};\sigma,z) &= -\ln\sum_j\int_{\max\left(a_j,\frac{a_j}{z}\right)}^{\min\left(b_j,\frac{b_j}{z}\right)} p_{Y,0}^{1-\sigma}(y)z^\sigma p_{Y,0}^{\sigma}(yz)dy\\ &= -\ln\sum_j\pi_j\int_{\max\left(a_j,\frac{a_j}{z}\right)}^{\min\left(b_j,\frac{b_j}{z}\right)} p_{0,j}^{1-\sigma}(y)z^\sigma p_{0,j}^{\sigma}(yz)dy\\ &= -\ln\sum_j\pi_j\exp\left\{-\mathcal{D}(p_{0,j};\sigma,z)\right\}\end{aligned} \tag{5.15}$$

i.e., it is a geometrical average of the $\{\mathcal{D}(p_{0,j};\sigma,z), 1 \le j \le j_{\max}\}$. Given z, the ability to discriminate between the two distributions $p_{0,j}(y)$ and $zp_{0,j}(yz)$ depends strongly on the size of the support set of $p_{0,j}$. In fact, $\mathcal{D}(p_{0,j};\sigma,z)$ tends to infinity as $\min(b_j, b_j/z) \downarrow \max(a_j, a_j/z)$. Hence one can achieve host signal rejection by making the pulses more narrow, as illustrated by the example below.

Example 5.1: Consider a *piecewise-constant PM distribution* with $j_{\max}$ pulses. Let $b_j = (1+\Delta)a_j$, where $\Delta > \epsilon/(1-\epsilon)$ satisfies the minimum-pulse-width condition. Then, $p_{0,j}(y) = (1/a_j\Delta)1_{\{y\in[a_j,b_j]\}}$. For $1 \le z \le 1/(1-\epsilon) < 1+\Delta$, we have

$$\begin{aligned}\mathcal{D}(p_{0,j};\sigma,z) &= -\ln\int_{a_j}^{\frac{b_j}{z}}\frac{1}{a_j\Delta}z^\sigma dy\\ &= -\ln\frac{\left(\frac{b_j}{z}-a_j\right)z^\sigma}{a_j\Delta}\\ &= -\ln\left[\left(\frac{1+\Delta}{z}-1\right)\frac{z^\sigma}{\Delta}\right]\end{aligned} \tag{5.16}$$

which is independent of j. Hence, (5.15) becomes

$$\mathcal{D}(p_{Y,0};\sigma,z) = -\ln\left[\left(\frac{1+\Delta}{z}-1\right)\frac{z^\sigma}{\Delta}\right] \tag{5.17}$$

achieving host-signal rejection. For $1/(1+\epsilon) \le z \le 1$, we use $\mathcal{D}(p_{Y,0};\sigma,z) = \mathcal{D}(p_{Y,0};1-\sigma,1/z)$, where the right side is evaluated using (5.17).

Example 5.2: Consider a PM unit exponential distribution, where $b_j = a_j(1+\Delta)$ again. Thus, $p_{0,j}(y) = (e^{-y}/(e^{-a_j} - e^{-b_j}))1_{\{y\in[a_j,b_j]\}}$. For $1 \le z < 1+\Delta$, we have

$$\begin{aligned}&\mathcal{D}(p_{0,j};\sigma,z)\\ &\quad= -\ln\int_{a_j}^{\frac{b_j}{z}} z^\sigma\frac{e^{-(1-\sigma)y}e^{-\sigma yz}}{e^{-a_j}-e^{-b_j}}dy\\ &\quad= -\ln\int_{a_j}^{\frac{b_j}{z}}\frac{z^\sigma}{e^{-a_j}-e^{-b_j}}e^{-[1-\sigma(1-z)]y}dy\\ &\quad= -\ln\left[z^\sigma\frac{e^{-[1-\sigma(1-z)]a_j}-e^{-[1-\sigma(1-z)]\frac{b_j}{z}}}{\left(e^{-a_j}-e^{-b_j}\right)[1-\sigma(1-z)]}\right]\\ &\quad= -\ln\left[z^\sigma\frac{e^{\sigma(1-z)a_j}-e^{-\left[\frac{(1-\sigma(1-z))(1+\Delta)}{z}-1\right]a_j}}{\left(1-e^{-a_j\Delta}\right)[1-\sigma(1-z)]}\right].\end{aligned} \tag{5.18}$$

For $1/(1+\epsilon) \le z \le 1$, we use $\mathcal{D}(p_{0,j};\sigma,z) = \mathcal{D}(p_{0,j};1-\sigma,1/z)$, where the right side is evaluated using (5.18). Again, we obtain host-signal rejection.

Observe that if $a_j\Delta$ is small, the pdf $p_{0,j}(y)$ is approximately constant over its support, and the result coincides with that in Example 5.1.

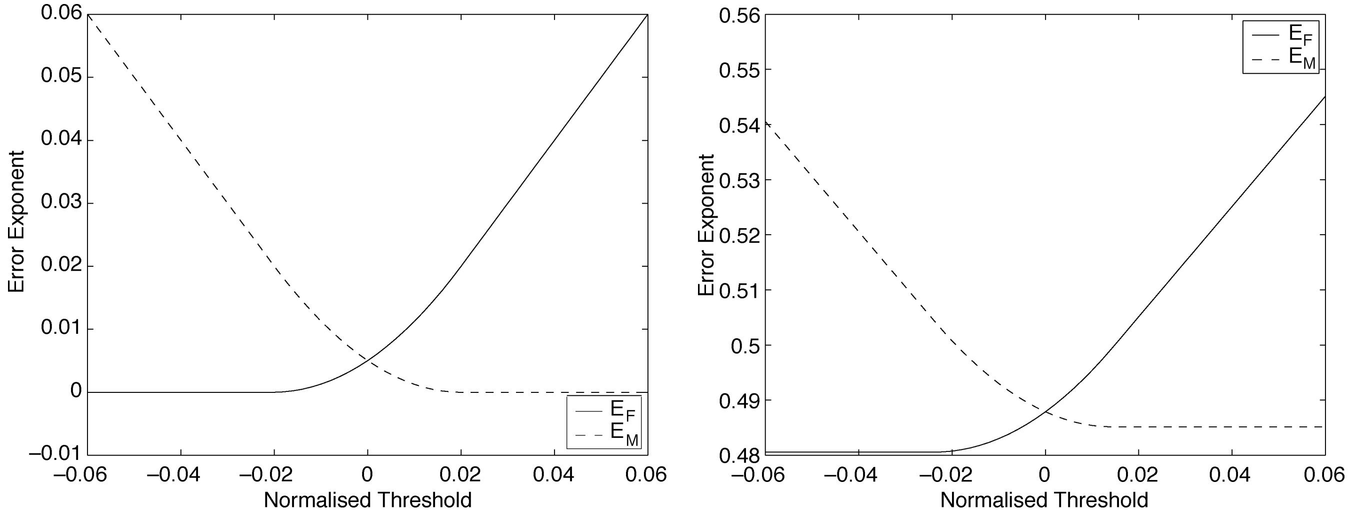
Example 5.3: Consider a PM-Weibull distribution, where $b_j = a_j(1+\Delta)$ again. The problem can be reduced to Example 2 using the transformation $\hat{s} = (s/\alpha)^\beta$, in which case the transformed variable $\hat{s}$ follows a unit exponential distribution. Then, we also define $\hat{a}_j = (a_j/\alpha)^\beta$, $\hat{b}_j = (b_j/\alpha)^\beta$, $\hat{z} = z^\beta$, and $\hat{\Delta} = (1+\Delta)^\beta - 1$. We obtain $\mathcal{D}(p_{0,j};\sigma,z)$ using (5.18) with $\hat{a}_j$, $\hat{z}$, and $\hat{\Delta}$ in place of a_j, z, and Δ.

D. Convexification of Chernoff Bounds

For each value of the exponent $\sigma \in (0,1)$, the Chernoff bounds (5.6) and (5.7), in which γ is viewed as a variable, define a lower bound $P_D = l_\sigma(P_{FA})$ on the ROC curve $P_D = f(P_{FA})$. This lower bound is a concave function. The upper envelope of these functions $P_D = l(P_{FA}) := \sup_{0<\sigma<1} l_\sigma(P_{FA})$ is a lower bound on the ROC curve as well. However, this envelope function is not necessarily concave. Denote by $c(P_{FA})$ the concave hull of the function $l(P_{FA})$. Since $l(P_{FA}) \le c(P_{FA}) \le f(P_{FA})$ (where the last inequality is due to the concavity of f), the convexified function c serves as an improved lower bound on the ROC when l is nonconcave. We found this technique to be useful for detection problems involving PM distributions, as detailed in the experiments in Section IX.

E. Evaluation of Chernoff Bounds

The Chernoff bounds for the Weibull and PM-Weibull distributions are now compared. For simplicity, a binary symmetric



(a)

(b)

Fig. 7. Error exponents E_F and E_M for (a) the Weibull detector and (b) the PM-Weibull detector. Here $\alpha = 0.1135$, $\beta = 2$, and $\mathcal{S}_h = [0.0881, 0.1145] \cup [0.1488, 0.1935] \cup [0.2515, 0.3270]$.

watermark distribution is employed, with $m_i \in \{-1, 1\}$. Typical values $\alpha = 0.1135$ and $\beta = 2$ are used (see Section IX). The set $\mathcal{S}_h$ that defines the hash function is given by $[0.0881, 0.1145] \cup [0.1488, 0.1935] \cup [0.2515, 0.3270]$, so $P_S(\mathcal{S}_h) = 0.3176$, while the embedding strength ϵ is selected to be 0.1. Figs. 7(a) and (b) show the error exponents (5.8) and (5.9) as a function of the normalized detection threshold $\bar{\gamma}$ for the Weibull and PM-Weibull detectors, respectively. Clearly, the error exponents are vastly different for the Weibull and PM-Weibull cases. Moreover, in the latter case, $E_F > 0$ (respectively, $E_M > 0$) as $\bar{\gamma} \to -\infty$ (respectively, $\bar{\gamma} \to \infty$) due to the presence of infinities in the log-likelihood ratio (4.7).

For both the Weibull and the PM-Weibull cases, the threshold $\gamma = 0$ maximizes the total-probability-of-error exponent $\min(E_F, E_M)$. The normalized Chernoff distances are computed for each coefficient model, using threshold $\gamma = 0$ and embedding strength $\epsilon = 0.1$. For the Weibull distribution, this normalized distance is approximately 0.05 while for the PM-Weibull distribution it is approximately 4.88, demonstrating the huge advantage of the PM-Weibull model over the Weibull model. That is, the same detection performance may be achieved using approximately 100 times fewer coefficients for watermarking.

VI. Eavesdropper's Detection Problem

Another problem worthy of study is the relative difficulty of the image watermark detection problem as seen by the detector compared to that seen by an eavesdropper. This eavesdropping situation can arise when watermarking techniques are employed in a distributed security system. For example, an eavesdropper may wish to detect watermarked images on the Internet. In another application, a security agency may insert a watermark to flag a suspicious image for further scrutiny by another agency. This mark should be difficult for the intended recipient of the image to detect. In both applications, the eavesdropper attempts to determine whether an object is watermarked without knowing the secret key and hash values.

If our watermarking algorithm is utilized, the eavesdropper knows neither the candidate region $\mathcal{C}$ nor the sequence $\{m_i\}$. When Operating Method 2 of Section IV-A is used, the eavesdropper knows N but not $\nu = (N_{\mathcal{C}}/N_T)$. The eavesdropper observes N_T coefficients $\{y_i, 1 \leq i \leq N_T\}$ (N of which are possibly marked) and evaluates two hypotheses: H_1 which states the data are marked and H_0 which states they are not. Under H_0, the distribution of the output $p_{\mathbf{Y},0}$ is simply given by a product of Weibull distributions. However, under H_1, since the eavesdropper does not know $\mathcal{C}$, she assumes a mixture distribution $p_{\mathbf{Y},1}^{\mathrm{eave}}$ of marked and unmarked coefficients. The mixture is obtained by averaging over all choices of $\mathcal{C}$ and $\mathbf{m}$. To formulate the distribution $p_{\mathbf{Y},1}^{\mathrm{eave}}$, a number of probability distributions are first defined

$$
\begin{aligned}
p_{Y,0}(y) &:= p_S(y) \quad \text{(unmarked distribution)}\\
p_Y^{\mathrm{UCPM}}(y) &:= \frac{1}{1 - P_S(\mathcal{S}_h)} p_S(y) 1_{\{y \notin \mathcal{S}_h\}}\\
&\quad \text{(unmarked complementary PM)}\\
p_Y^{\mathrm{UPM}}(y) &:= \frac{1}{P_S(\mathcal{S}_h)} p_S(y) 1_{\{y \in \mathcal{S}_h\}} \quad \text{(unmarked PM)}\\
p_{Y|M}^{\mathrm{MPM}}(y|m) &:= \frac{1}{1+\epsilon m} p_Y^{\mathrm{UPM}}\left(\frac{y}{1+\epsilon m}\right)\\
&\quad \text{(marked PM, conditioned on } M = m\text{)}\\
p_Y^{\mathrm{MPM}}(y) &:= \int_{-1}^{1} p_{Y|M}^{\mathrm{MPM}}(y|m) dP_M(m)\\
&\quad \text{(mixture marked PM)}\\
p_Y^{\mathcal{C}}(y) &:= (1 - P_S(\mathcal{S}_h)) p_Y^{\mathrm{UCPM}}(y) + P_S(\mathcal{S}_h) p_Y^{\mathrm{MPM}}(y)\\
&\quad \text{(mixture distribution for } y \text{ in } \mathcal{C}\text{).}
\end{aligned}
$$

Typically, the two mixture components of $p_Y^{\mathcal{C}}$ are well separated (they overlap only on the boundaries of the intervals forming $\mathcal{S}_h$). An *independent and identically distributed* (iid) distribution is assumed for $\{m_i\}$ in order to simplify the exposition (and maximize the difficulty of the eavesdropper's detection problem). With the notation above, the distribution of Y as seen by the eavesdropper under H_1 is given by

$$p_{Y,1}^{\text{eave}} := (1-\bar{\nu})p_{Y,0} + \bar{\nu} p_Y^{\mathcal{C}} \tag{6.1}$$

with $\bar{\nu} := \mathrm{E}[N_{\mathcal{C}}/N_T] = N/(N_T P_S(\mathcal{S}_h))$. The Chernoff distance between the distributions under H_0 and H_1 for the eavesdropper is given by

$$D\left(\sigma^*, p_{\mathbf{Y},0}, p_{\mathbf{Y},1}^{\text{eave}}\right) = N_T D\left(\sigma^*, p_{Y,0}, (1-\bar{\nu})p_{Y,0} + \bar{\nu} p_Y^{\mathcal{C}}\right). \tag{6.2}$$

As described in Section V, the situation is different for the actual detector, which has knowledge of $\mathcal{C}$ and the sequence $\{m_i\}$. Thus, for each coefficient, the detector knows which pair of distributions to consider for the two hypotheses. The Chernoff distance (5.4) as seen by the detector can be written as

$$D(\sigma^*, p_{\mathbf{Y},0}, p_{\mathbf{Y},1}) = N_T \bar{\nu} P_S(\mathcal{S}_h) \int_{-1}^{1} D\left(\sigma^*, p_{Y,0}, p_{Y|M=m}^{\text{MPM}}\right) dPM(m). \tag{6.3}$$

A comparison may now be made between the Chernoff distances (6.2) and (6.3) seen by the eavesdropper and the detector, to provide insight into the relative difficulty of the detection problems. With the above formulations and by the concavity of Chernoff distance [17], it is clear that

$$D(\sigma^*, p_{\mathbf{Y},0}, p_{\mathbf{Y},1}) > D\left(\sigma^*, p_{\mathbf{Y},0}, p_{\mathbf{Y},1}^{\text{eave}}\right)$$

i.e., the detection problem is more difficult for the eavesdropper than for the detector.

To quantify this effect, the Chernoff distances are evaluated over a range of possible $\bar{\nu}$ values using the same parameters as in Section V-E. For ease of comparison, the Chernoff distances are normalized with respect to N. Again using a binary symmetric watermark distribution, the resulting normalized Chernoff distance curves are given in Fig. 8. The eavesdropper observes a significantly smaller Chernoff distance than the informed detector (approximately an order of magnitude) and, hence, is much less able to detect the watermark.

VII. Attacks

This section considers the effect of attacks on system performance. To illustrate the idea, consider the following multiplicative noise model for attacks:

$$y_i = x_i w_i, \quad \forall i \tag{7.1}$$

where $\{W_i\}$ are iid distributed random variables following a distribution $p_W(w)$ and are independent of $\{X_i\}$. It is convenient to use a logarithmic transformation to map the problem into an additive model, $\tilde{y}_i = \tilde{x}_i + \tilde{w}_i$, where the tilde symbol indicates variables in the log domain ($\tilde{y} = \ln y$, etc.). Then, $p_{\tilde{Y}}(\tilde{y}) = e^{\tilde{y}} p_Y(e^{\tilde{y}})$ and $p_Y(y) = (1/y)p_{\tilde{Y}}(\ln y)$, etc. Also, define $\tilde{\mathcal{S}}_h = \{\tilde{y} : e^{\tilde{y}} \in \mathcal{S}_h\}$. The binary hypothesis test can be written as

$$\begin{aligned} H_0 &: \tilde{y}_i = \tilde{s}_i + \tilde{w}_i, \quad 1 \le i \le N \\ H_1 &: \tilde{y}_i = \tilde{s}_i + \tilde{w}_i + \ln(1+\epsilon m_i), \; 1 \le i \le N. \end{aligned} \tag{7.2}$$

Also, recall that Chernoff distance is invariant to invertible transformations of the data, so $D(\sigma, p_{\mathbf{Y},0}, p_{\mathbf{Y},1}) = D(\sigma, p_{\tilde{\mathbf{Y}},0}, p_{\tilde{\mathbf{Y}},1})$.

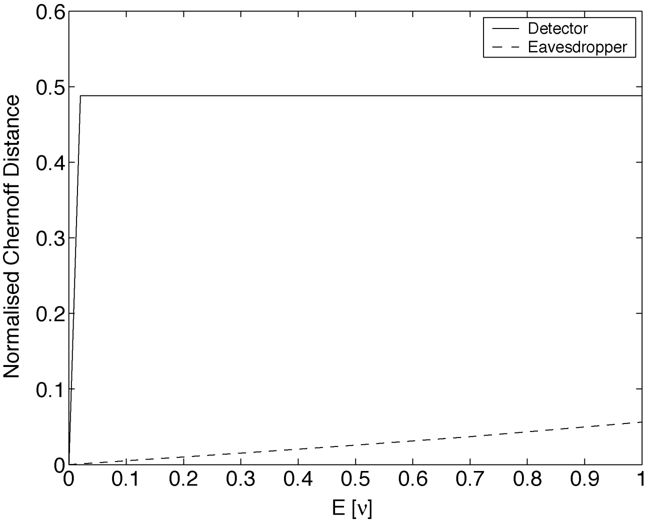


Fig. 8. Normalized Chernoff distances as seen by the detector and eavesdropper as a function of $\bar{\nu}$, the normalized size of the candidate set $\mathcal{C}$. The system parameters are the same as in Fig. 7.

The PM distributions assumed in the absence of attacks must be replaced with convolutions with $p_{\tilde{W}}$. In particular, from (4.3) and (7.2), we obtain $\tilde{p}_{PM}(\tilde{s}) = (1/P_{\tilde{S}}(\tilde{\mathcal{S}}_h))p_{\tilde{S}}(\tilde{s})1_{\{\tilde{s} \in \tilde{\mathcal{S}}_h\}}$ and

$$p_{\tilde{Y}_i,0}(\tilde{y}_i) = (\tilde{p}_{PM} \star p_{\tilde{W}})(\tilde{y}_i), \quad 1 \le i \le N \tag{7.3}$$

$$p_{\tilde{Y}_i,1}(\tilde{y}_i) = (\tilde{p}_{PM} \star p_{\tilde{W}})\left(\tilde{y}_i - \ln(1+\epsilon m_i)\right), \quad 1 \le i \le N. \tag{7.4}$$

The Chernoff distance is therefore given by (5.4), where $p_{Y,0}(y) = (1/y)p_{\tilde{Y},0}(\ln y)$. Greater blurring of the original, discontinuous p_{PM} due to p_W will yield a smaller Chernoff distance.

To illustrate this problem, it is useful to consider attacks of the form $p_{\tilde{W}}(\tilde{w}) = (1/\eta)q(\tilde{w}/\eta)$, where q is any normalized pdf, and η is a scale parameter; the attack becomes benign as $\eta \to 0$. Consider, for instance, the triangular pdf

$$q(\tilde{w}) = \max(0, 1 - |\tilde{w}|)$$

whose support set is [−1,1]. Fig. 9 gives the Chernoff distance in the presence of this attack as a function of η.

VIII. Nuisance Parameters

As well as the noise introduced by the attacker, there may be a small number of additional parameters that are unknown to the detector. For instance, the scale parameter α of the PE or Weibull distribution is generally not known to the detector, nor is a possible fixed scaling parameter c introduced by the attacker (in terms of the attack model in the previous section, we would have $y_i = cx_i$, where c is an unknown constant).

We propose using a noncoherent version of our likelihood ratio detector (4.2), treating unknown parameters such as α and c as nuisance parameters. The detector is a generalized likelihood ratio test

$$\tilde{L}(\mathbf{y},\mathbf{h}) = \frac{\sup_{\alpha,c} p_{\mathbf{Y}|\mathbf{H},1}(\mathbf{y}|\mathbf{h},\alpha,c)}{\sup_{\alpha,c} p_{\mathbf{Y}|\mathbf{H},0}(\mathbf{y}|\mathbf{h},\alpha,c)} \underset{H_0}{\overset{H_1}{\gtrless}} e^{\gamma} \tag{8.1}$$

where e^{γ} is the threshold of the test. This approach can be used if additional nuisance parameters are present, e.g., a parametric description of a point operation used by the attacker.

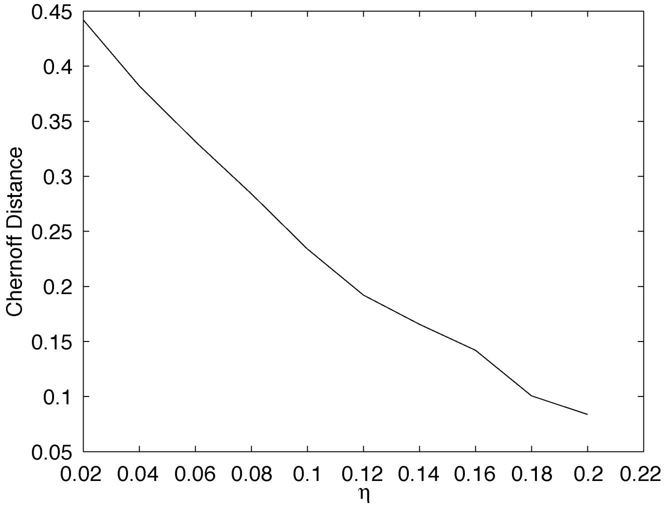


Fig. 9. Chernoff distance in the presence of a multiplicative white triangular noise attack, as a function of attack scale parameter η.

IX. Image Watermarking Experiments

The previous sections examined the design of optimal detectors and the derivation of performance bounds. Now, these detectors are applied to standard photographic test images to ascertain how well they perform on coefficients which do not necessarily follow the idealized distributions assumed in this paper.

A. Implementation Details

We selected the standard 512 × 512 *Lena* and *Baboon* images as test images and applied the full-frame DCT transform to them. For watermarking we considered only a trapezoidal region of coefficients in the low-mid frequencies: the union of Regions 6, 7, 8, 10, 11, and 12 (using the notation of [15], [16]), which has a total of 5286 elements. We chose $N = 10$ for both Lena and Baboon. The Weibull and PM-Weibull distributions are both considered for modeling the DCT coefficients at the detector. A fixed value of $\beta = 2$ is utilized (similar to [15], [16]); however, the detector must estimate the distribution parameter α. A maximum likelihood estimate is employed within each of Regions 6, 7, 8, 10, 11, and 12, and the resulting distribution parameters are given in Table I.

To determine the set $\mathcal{S}_h$, the method described in Example 4.1 is utilized. Since the α parameters must be estimated at the detector, the parameter δ defined in Example 4.1 is taken to be a coarsely quantized version of the average of the α parameters over the regions considered. The detector, knowing the candidate values for δ, can reliably retrieve the value of δ selected by the encoder. The resulting set $\mathcal{S}_h$ is given by $[0.0881, 0.1145] \cup [0.1488, 0.1935] \cup [0.2515, 0.3270]$.

Referring to our description of Operating Method #2 in Section IV-A, $\nu = (N_{\mathcal{C}}/N_T)$ is image dependent and key dependent. We obtained $\nu \approx 0.0075$ on average for Lena (yielding $N_{\mathcal{C}} \approx 40$) and $\nu \approx 0.0105$ for Baboon (yielding $N_{\mathcal{C}} \approx 56$). In other words, typically 0.75% and 1.05% of the transform coefficients were used for watermarking Lena and Baboon, respectively.

For each detector considered, a Monte Carlo simulation was used to determine the detection and false alarm probabilities P_D and P_{FA}. The simulations were performed over a range of thresholds, with 10^5 runs for each threshold. A binary symmetric watermark distribution was utilized, and a new watermark and key were generated for each trial.

TABLE I
Estimated α Parameters for the DCT Coefficients ($\beta = 2$)

Region	Lena	Baboon
6	0.1388	0.1475
7	0.1255	0.1520
8	0.1136	0.1533
10	0.1043	0.2018
11	0.1065	0.1849
12	0.0920	0.1863

B. Simulation Results

To present the detection results for the image data, two types of figures are considered for each detector-image pair. The first shows the distribution of the test statistic and the second plots P_D and P_{FA} for an embedding strength of $\epsilon = 0.1$.

The distributions of the test statistics are shown for the Weibull and PM-Weibull distributions in Fig. 10 for Lena. For the PM-Weibull distribution, the infinite values which may be present in the statistics are represented pictorially using $\pm 10^8$. For the Weibull distribution, the distributions of the statistics under H_0 and H_1 are not well separated, indicating the difficulty in choosing between the two hypotheses. On the other hand, the distributions resulting from the use of the PM-Weibull distribution are extremely well separated, with the majority of the trials yielding $|\ln L(\mathbf{y}, \mathbf{h})| = \infty$. The results for the Baboon image were quite similar, suggesting that detector performance should not be highly dependent on the choice of image.

The P_D and P_{FA} curves based on the Monte Carlo simulations are shown in Fig. 11 along with the corresponding convexified Chernoff bounds (refer to Section V-D). Performance is similar for the Lena and Baboon images for both the Weibull detector and the PM-Weibull detector. For both images, the PM-Weibull detector produces detection probabilities within the range [0.99,1] for false alarm values in the range [0,0.01]. The performance using the conventional Weibull detector is dramatically worse.

Fig. 11 also provides the ROC for synthetic data generated from the Weibull distribution assumed for Lena and Baboon. The slight discrepancy between the ROC for the synthetic data and for the image data is due to the imperfections of the Weibull model for images.

C. Image Watermarking Experiments With Attacks

This section considers the detection of watermarks in the presence of attacks. Since the development of the detectors of Section IV-B did not include a model of an attack, these detectors are no longer optimal. Therefore, the robustness of these detectors in the presence of attacks is now studied. Two attack methods are considered: *multiplicative white triangular*

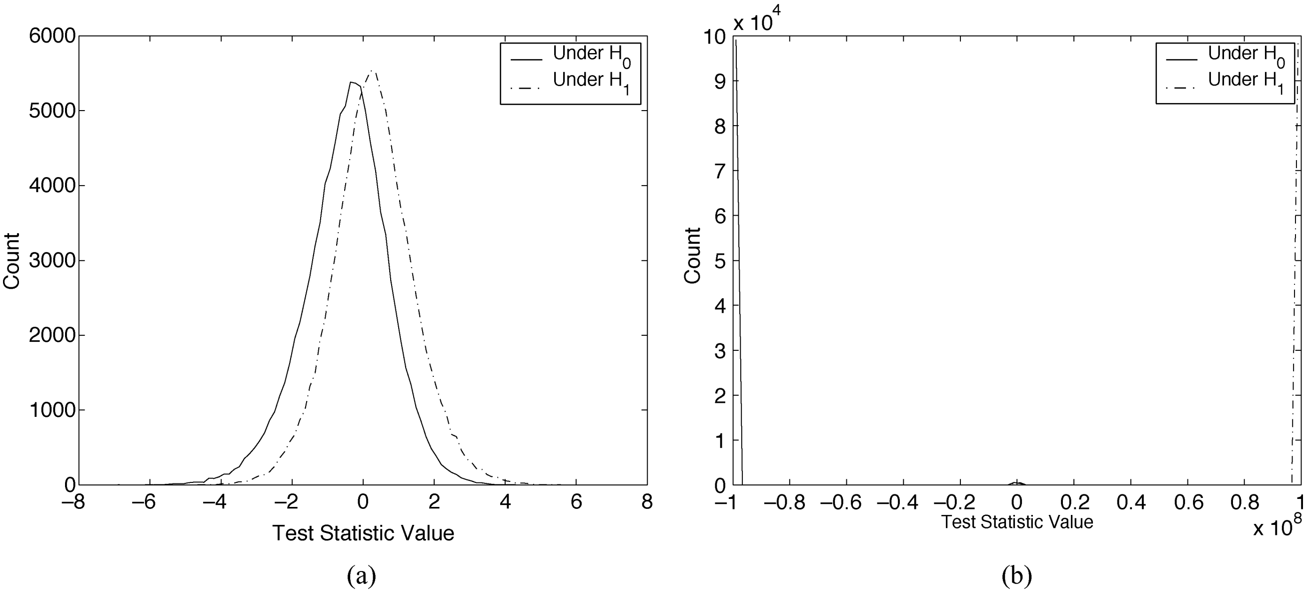


Fig. 10. Distribution of the test statistic for the Lena image using the (a) Weibull and (b) PM-Weibull distributions.

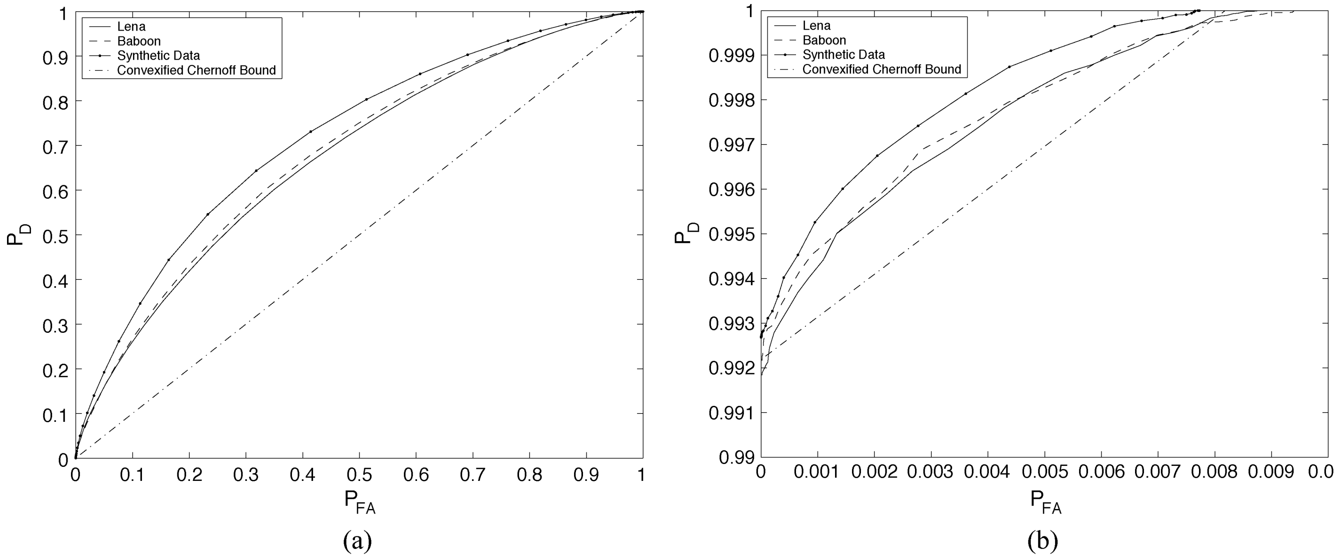
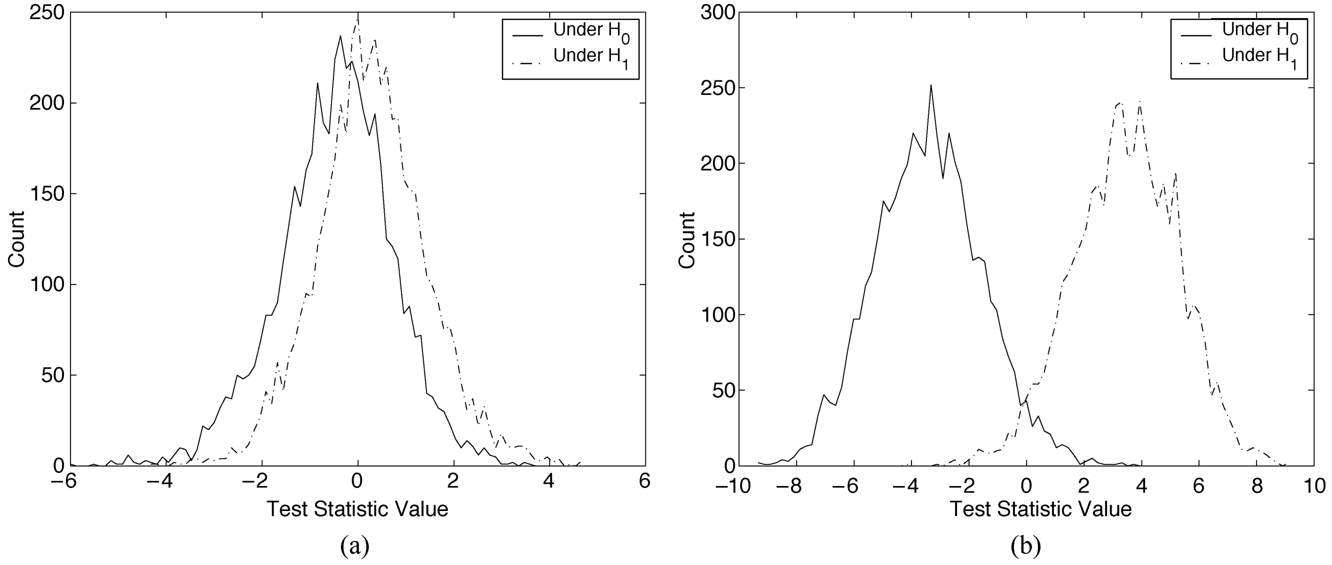


Fig. 11. ROC curves for the Lena, Baboon, and synthetic data, and the convexified Chernoff bounds, for the (a) Weibull and (b) PM-Weibull distribution. Here, $N = 10$, $\beta = 2$, and $\epsilon = 0.1$.

noise (MWTN) (see Section VII), and *JPEG compression*. The MWTN attack is limited to Regions 6, 7, 8, 10, 11, and 12, since the sequence is known to be embedded in a subset of these coefficients; however, the JPEG attack is applied to the entire image. For the MWTN attack, a more robust version of the PM-Weibull detector is used, namely the infinities in the test statistic are replaced with finite constant values, which are selected as ±1 in the experiments.

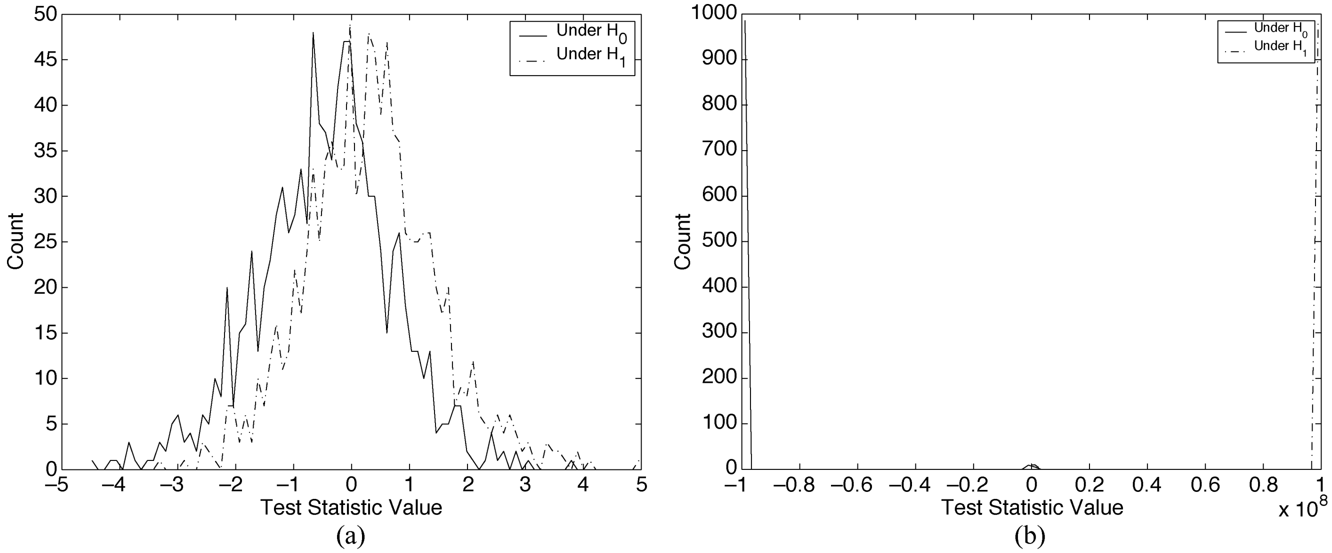
To evaluate the performance of the detectors based on the PM-Weibull and Weibull distributions, Monte Carlo simulations are performed with the Lena image using 5000 and 1000 trials for each of the attack types, respectively. The *mean squared error* (MSE) introduced through watermarking is denoted by D_1, while that introduced by the attack is denoted by D_2. The MSE for both attacks is large: $D_2 = 10D_1$. To illustrate detector performance, plots of the test statistic distributions and ROC curves are included.

The distributions of the test statistics for both modeling distributions are given in Figs. 12 and 13 for MWTN and JPEG compression, respectively. Little change in the Weibull statistic distributions is present, for either attack type, from the corresponding distributions when no attack is present, as seen in Fig. 10. Thus, the performance of the Weibull detector is not significantly affected by the MWTN or JPEG compression attacks. For the PM-Weibull case, the clipping of infinities in the MWTN case significantly alters the distribution of the test statistics; however, a strong separation under H_0 and H_1 is still present. Conversely, for the JPEG attack, the majority of the distribution mass is again located at the infinity points. The small influence of the JPEG attack is expected, since the attack is ap-



(a) (b)

Fig. 12. Distribution of the test statistic for Lena under a MWTN attack for the (a) Weibull and (b) PM-Weibull distributions.



(a) (b)

Fig. 13. Distribution of the test statistic for Lena under a JPEG compression attack for the (a) Weibull and (b) PM-Weibull distributions.

plied to the entire image, causing the marked coefficients to be altered less for a fixed D_2.

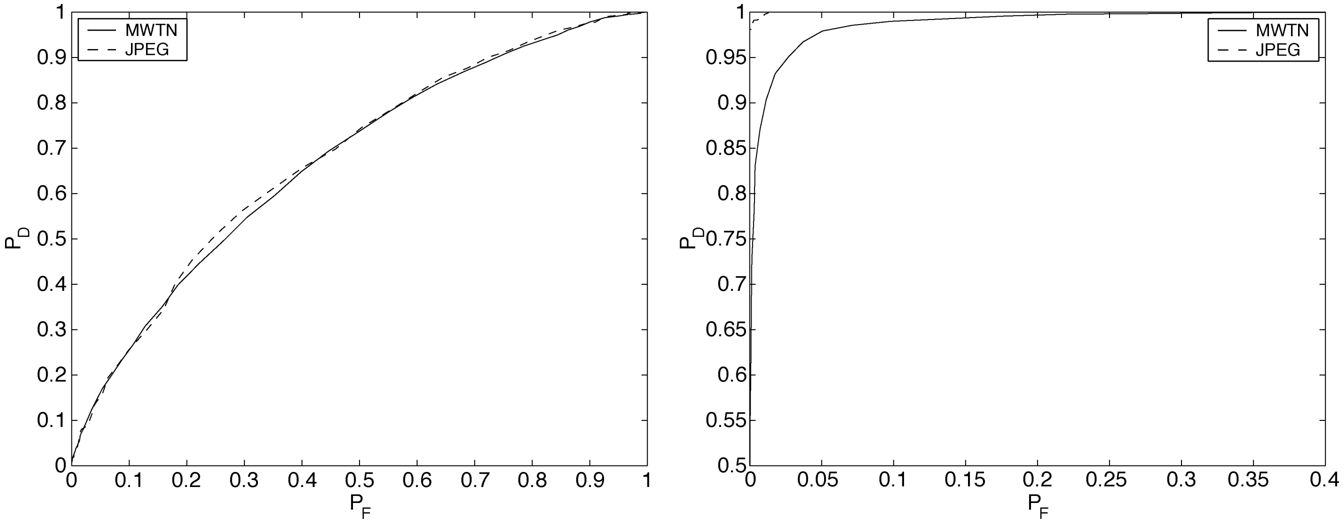
To further quantify the performance of these detectors, ROC curves are plotted in Fig. 14. The curves for the Weibull detector are virtually identical to those for the unattacked case, given in Fig. 11. On the other hand, a slight decrease in performance is present for the PM-Weibull case (particularly for the MWTN) when compared with the data resulting from no attack, Fig. 11. This decrease is consistent with the change in the statistic distribution previously observed. However, the performance of the PM-Weibull detector remains notably superior to that of the Weibull detector.

X. Conclusion

Statistical modeling and signal detection theory provide a structured framework in which optimal watermarking systems can be developed and studied. The encouraging results obtained clearly reveal the potential of joint image hashing/watermarking as a viable means of information protection.

In the setup considered in this paper, the detector has access to side information about the original image in the form of an image hash (1 bit of information for each original DCT coefficient at secret locations), creating a joint hashing/watermarking system. The inclusion of the side information permits the development of detectors which offer extremely high performance, even for very short watermarks $(\mathrm{length} \approx 10)$. A pulse-modulated (PM) distribution for modeling the selected coefficients was developed as a consequence of the particular hash selected, and the Neyman–Pearson detector was formulated.

Chernoff bounds were derived to analyze the performance of this detector. Evaluation of the bounds revealed tremendous increases in detection performance over hash-free systems. The Chernoff distance as seen by an eavesdropper attempting to detect the watermark was evaluated and found to be low; this result



(a)

(b)

Fig. 14. ROC curves for the MWTN and JPEG compression attacks for the (a) Weibull and (b) PM-Weibull distributions $(D_2 = 10D_1)$.

quantifies the increase in difficulty for the eavesdropper's detection problem.

Monte Carlo simulations were employed to analyze the performance of the detectors using data gathered from real-world test images. The PM-Weibull detector displayed vastly superior performance over the Weibull detector, offering decreases of over 95% in false alarm probabilities, for the same detection probability.

Experiments were also conducted to study the effect of multiplicative white triangular noise and JPEG compression. The Weibull detector is nearly unaffected by the introduction of an attack, while the PM-Weibull detector is slightly hindered. However, the PM-Weibull detector still significantly outperforms the Weibull detector. It is likely that further improvements can be obtained by taking attacks into account in the design of the watermark detector; this enhancement is a topic of future research.

ACKNOWLEDGMENT

The authors would like to thank Prof. T. Kalker for suggesting forensic watermarking applications.

REFERENCES

[1] P. Kocher, J. Jaffe, B. Jun, C. Laren, and N. Lawson, "Self-protecting digital content," Cryptography Research, Inc., 2003.

[2] M. Holliman, N. Memon, and M. Yeung, "On the need for image dependent keys in watermarking," in *Proc. 2nd Workshop on Multimedia*, Newark, NJ, 1999.

[3] M. Kutter, S. Voloshinovskij, and A. Herrigel, "The watermark copy attack," *Proc. SPIE*, vol. 3657, pp. 226–239, 1999.

[4] I. J. Cox and J.-P. M. G. Linnartz, "Public watermarks and resistance to tampering," in *Proc. IEEE Int. Conf. Image Processing*, Santa Barbara, CA, 1997.

[5] "Watermarking patents to activated content," Digimarc Press Release. [Online]. Available: http://www.audiowatermarking.co.uk/patents.htm (2002, July).

[6] M. van der Veen, A. Lemma, and T. Kalker, "Watermarking and fingerprinting for electronic music delivery," *Proc. SPIE*, vol. 5306, Jan. 2004.

[7] G. Voyatzis and I. Pitas, "The use of watermarks in the protection of digital multimedia products," *Proc. IEEE*, vol. 87, pp. 1197–1207, July 1999.

[8] S. Roy and E.-C. Chien, "Watermarking with knowledge of image database," in *Proc. IEEE Int. Conf. Image Processing*, Barcelona, Spain, 2003.

[9] J. Cannons, "Optimal detection of multiplicative watermarks using image hashes," M.S. thesis, Univ. Illinois at Urbana-Champaign, Urbana, IL, 2002.

[10] I. Cox, J. Kilian, T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. Image Processing*, vol. 6, pp. 1673–1687, Dec. 1997.

[11] R. Venkatesan, M. J. S.-M. Koon, and P. Moulin, "Robust image hashing," in *Proc. IEEE Int. Conf. Image Processing*, vol. 3, Vancouver, BC, Canada, 2000, pp. 664–666.

[12] J. Fridrich and M. Goljan, "Robust hash functions for digital watermarking," in *Proc. Int. Conf. Information Technology: Coding and Computing*, Las Vegas, NV, 2000, pp. 178–183.

[13] 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.

[14] R. Wolfgang, C. Podilchuk, and E. Delp, "Perceptual watermarks for digital images and video," *Proc. IEEE*, vol. 87, pp. 1108–1126, July 1999.

[15] M. Barni, F. Bartolini, A. De Rosa, and A. Piva, "Capacity of full frame DCT image watermarks," *IEEE Trans. Image Processing*, vol. 9, pp. 1450–1455, Aug. 2000.

[16] M. Barni, F. Bartolini, A. Piva, and F. Rigacci, "Statistical modeling of full frame DCT coefficients," in *Proc. 9th Eur. Signal Processing Conf. (EUSIPCO)*, Rhodes, Greece, 1998, pp. 1513–1516.

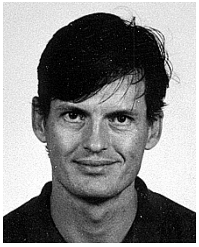
[17] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York: Springer-Verlag, 1994.

[18] P. Moulin and A. Ivanović, "The zero-rate spread-spectrum watermarking game," *IEEE Trans. Signal Processing*, vol. 51, pp. 1098–1117, Apr. 2003.

[19] H. L. VanTrees, *Detection, Estimation and Modulation Theory*. New York: Wiley, 1968.

Jillian Cannons (S'99) received the B.Sc. degree in computer engineering (with distinction) from the University of Manitoba, Winnipeg, MB, Canada, in 2000 and the M.S. degree in electrical engineering from the University of Illinois at Urbana-Champaign, in 2002.

She is currently pursuing the Ph.D. degree at the University of California at San Diego, La Jolla, where she is a member of the Information Coding Laboratory. Her current research interests include the areas of information theory, data compression, and source coding.



Pierre Moulin (F'03) received the Ingénieur civil électricien degree from the Faculté Polytechnique de Mons, Belgium, in 1984 and the M.Sc. and D.Sc. degrees in electrical engineering from Washington University, St. Louis, MO, in 1986 and 1990, respectively.

He was a Researcher at the Faculté Polytechnique de Mons from 1984 to 1985 and at the Ecole Royale Militaire, Brussels, Belgium, from 1986 to 1987. He was a Research Scientist at Bell Communications Research, Morristown, NJ, from 1990 until 1995. In 1996, he joined the University of Illinois at Urbana-Champaign, Urbana, where he is currently Professor in the Department of Electrical and Computer Engineering, Research Professor at the Beckman Institute and the Coordinated Science Laboratory, and affiliate faculty member in the Department of Statistics. His fields of professional interest are image and video processing, compression, statistical signal processing and modeling, nonparametric function estimation, information theory, information hiding, and the application of multiresolution signal analysis, optimization theory, and fast algorithms to these areas.

Dr. Moulin has served as an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY and the IEEE TRANSACTIONS ON IMAGE PROCESSING, Co-Chair of the 1999 IEEE Information Theory Workshop on Detection, Estimation, Classification, and Imaging, Chair of the 2002 NSF Workshop on Signal Authentication, and Guest Associate Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY 2000 Special Issue on Iinformation—Theoretic Imaging and the IEEE TRANSACTIONS ON SIGNAL PROCESSING 2003 Special Issue on Data Hiding. From 1998 to 2003, he was a member of the IEEE IMDSP Technical Committee. Currently, he is Area Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING and Guest Editor of the upcoming IEEE TRANSACTIONS ON SIGNAL PROCESSING Supplement Series on Secure Media. He received a 1997 Career Award from the National Science Foundation and a IEEE Signal Processing Society 1997 Senior Best Paper award. He is also coauthor (with Juan Liu) of a paper that received an IEEE Signal Processing Society 2002 Young Author Best Paper award. He was 2003 Beckman Associate of the UIUC's Center for Advanced Study.