Achievable Performance of Digital Watermarking Systems

Brian Chen Gregory W. Wornell

Department of Electrical Engineering and Computer Science,
and Research Laboratory of Electronics
Massachusetts Institute of Technology
Cambridge, MA, USA

E-mail: bchen@mit.edu gww@allegro.mit.edu

Abstract

A variety of digital watermarking applications have emerged recently that require the design of systems for embedding one signal (the "embedded signal" or "watermark") within another signal (the "host signal"). We develop a framework for analyzing achievable performance trade-offs of these systems among robustness, distortion, and embedding rate. We also describe a recently introduced class of embedding methods, quantization index modulation (QIM), in which an ensemble of quantizers is constructed and information is embedded by quantizing the host signal with a quantizer associated with the watermark. We introduce an implementation of such a method called spread-transform dither modulation where the embedded information modulates the dither signal of a dithered quantizer, which quantizes projections of the host signal onto a spreading vector. We show that QIM systems have considerable performance advantages over previously proposed spread-spectrum and low-bit modulation systems.

1. Introduction

A variety of related applications have emerged recently [8] that require the design of systems for embedding one signal, sometimes called an "embedded signal" or "watermark", within another signal, called a "host signal". These applications include copyright notification and enforcement, authentication, and transmission of auxiliary information. In each of the proposed applications, the embedding must be done such that the embedded signal causes no serious degradation to its host. At the same time, the host always carries the embedded signal, which can only be removed by causing significant damage to the host.

This work has been supported in part by ONR under Grant No. N00014-96-1-0930, by AFOSR under Grant No. F49620-96-1-0072, and by a NDSEG Fellowship.

Various information-embedding algorithms have been proposed [8] in this still emerging field. Some of the earliest proposed systems [1] employ a quantize-and-replace strategy: after first quantizing the host signal, these systems change the quantization value to embed information. A simple example of such a system is so-called low-bit(s) modulation (LBM), where the least significant bit(s) in the quantization of the host signal are replaced by a binary representation of the embedded signal. Recently, spread-spectrum based systems, which embed information by adding to the host signal a small pseudo-noise signal that is modulated by the embedded signal, have received considerable attention in the literature. (See the references in [8], for example.) However, as we demonstrate in this paper, spread-spectrum based systems offer relatively little robustness when the host signal is not known at the decoder. Intuitively, when the host signal is not known at the decoder, as is typical in many applications of interest, it is a source of interference. With a spread-spectrum system, the host signal is an additive interference that is often much larger, due to distortion constraints, than the pseudo-noise signal carrying the embedded information.

In this paper we introduce a framework for characterizing the inherent trade-offs among embedding rate, embedding-induced degradation, and robustness of information embedding methods and describe a recently introduced family of techniques called "quantization index modulation" (QIM) [3] that perform these trade-offs efficiently. We also explore a new, convenient realization of QIM, "spread-transform dither modulation", which offers significant advantages over previously proposed spread-spectrum and LBM techniques.

2. Problem model

Many information-embedding applications can be described by Fig. 1. We have some host signal vector $\mathbf{x} \in \Re^N$ in which we wish to embed some information m. This host

0-7695-0253-9/99 $10.00 © 1999 IEEE

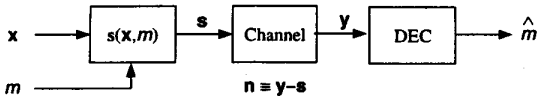

Figure 1. General information-embedding problem model. A message m is embedded in $\mathbf{x}$ using some embedding function $\mathbf{s}(\mathbf{x}, m)$. The channel corrupts $\mathbf{s}$. The decoder extracts an estimate $\hat{m}$ of m from $\mathbf{y}$.

signal could be a vector of pixel values or Discrete Cosine Transform (DCT) coefficients from an image, for example. Alternatively, the host signal could be a vector of samples or transform coefficients, such as Discrete Fourier Transform (DFT) or linear prediction coding coefficients, from an audio or speech signal. We wish to embed at a rate of R_m bits per dimension (bits per host signal sample) so we can think of m as an integer, where

$$m \in \{1, 2, \ldots, 2^{NR_m}\}. \tag{1}$$

An embedding function maps the host signal $\mathbf{x}$ and embedded information m to a composite signal $\mathbf{s} \in \Re^N$ subject to some distortion constraint. For example, one might choose the squared-error distortion constraint

$$D(\mathbf{s}, \mathbf{x}) = \frac{1}{N}\|\mathbf{s} - \mathbf{x}\|^2 \leq D_{\max}. \tag{2}$$

The composite signal $\mathbf{s}$ passes through a channel, where it is subjected to various common signal processing manipulations such as lossy compression, addition of random noise, and resampling, as well as deliberate attempts to remove the embedded information. We let $\mathbf{y} \in \Re^N$ denote the output of the channel and define a perturbation vector to be the difference $\mathbf{n} \triangleq \mathbf{y} - \mathbf{s}$. The decoder forms an estimate $\hat{m}$ of the embedded information m based on the channel output $\mathbf{y}$. We would like the estimate to be reliable as long as the channel corruptions are not too severe. Thus, a measure of the robustness of our system is the severity of the channel corruptions that can be tolerated such that either we can guarantee that $\hat{m} = m$ or $\Pr[\hat{m} \neq m] < \epsilon$. Specific channels and the corresponding measures of corruption severity and robustness that are of interest in this paper are:

1. **bounded perturbation channels:** In this case, we consider the largest σ_n such that we can guarantee $\hat{m} = m$ whenever $\mathbf{n}$ satisfies

   $$\|\mathbf{y} - \mathbf{s}\|^2 = \|\mathbf{n}\|^2 \leq N\sigma_n^2. \tag{3}$$

   This channel model describes a maximum distortion constraint between the channel input and output and may be an appropriate model for the effect of a lossy compression algorithm, printing and scanning, and attempts by an active attacker to remove the embedded signal, for example.

2. **bounded host-distortion channels:** Some attackers may work with distortion constraint between the host signal, rather than the channel input, and the channel output since this distortion is the most direct measure of degradation to the host signal. For example, if an attacker has partial knowledge of the host signal, which may be in the form of a probability distribution, so that he or she can calculate this distortion, then it may be appropriate to bound the expected distortion $D_{\mathbf{y}} = E[D(\mathbf{y}, \mathbf{x})]$.

3. **JPEG channels:** The output of a JPEG channel is simply the JPEG-compressed version of the input, and the robustness measure is the worst-case tolerable JPEG quality factor.

4. **probabilistic channels:** In some contexts it is convenient to assume some probability distribution for $\mathbf{n}$. Although we point out in this paper how our analysis framework can be applied to these channels, we refer the reader to [4, 5, 2] for details.

We wish to have high rate, low distortion, and high robustness, but in general these three goals tend to conflict. Thus, the performance of an information embedding system can be measured in terms of its achievable trade-offs among these three parameters.

3. Quantization index modulation and dither modulation

Before examining the achievable performance trade-offs of several digital watermarking systems in various scenarios, we describe a recently introduced [4, 5] class of embedding systems called quantization index modulation (QIM) and a convenient realization of this class called dither modulation. We also highlight some properties of these systems that, as we shall see in later sections, lead to attractive performance advantages over spread spectrum and LBM techniques.

3.1. General QIM systems

We can view the embedding function $\mathbf{s}(\mathbf{x}, m)$ as an ensemble of functions of $\mathbf{x}$, indexed by m. We denote the functions in this ensemble as $\mathbf{s}(\mathbf{x}; m)$ to emphasize this view. In QIM systems [3], these functions are quantizers, which is convenient for at least two reasons. First, each individual quantizer is designed such that one can satisfy the distortion

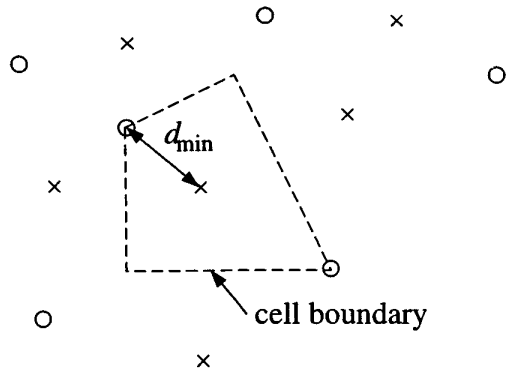

Figure 2. Quantization index modulation. The points marked with ×'s and ○'s belong to two different quantizers. If $m = 1$, the host signal is quantized to the nearest ×. If $m = 2$, the host signal is quantized to the nearest ○.

constraint. Second, the reconstruction points of each quantizer in the ensemble are "far away" in some sense from the reconstruction points of every other quantizer so that the system is robust to perturbations. Quantization index modulation refers to modulating an index or sequence of indices with the embedded information and quantizing the host signal with the associated quantizer or sequence of quantizers.

Fig. 2 illustrates this QIM information-embedding technique. In this example, one bit is to be embedded so that $m \in \{1, 2\}$. Thus, we require two quantizers, and their corresponding sets of reconstruction points in $\Re^N$ are represented in Fig. 2 with ×'s and ○'s. If $m = 1$, for example, the host signal is quantized with the ×-quantizer, i.e., $\mathbf{s}$ is chosen to be the × closest to $\mathbf{x}$. If $m = 2$, $\mathbf{x}$ is quantized with the ○-quantizer. Here, we see that the sets of reconstruction points of the two quantizers are "far away" from each other in the sense that there is some nonzero distance between every × point and every ○ point.

A few parameters of the ensemble conveniently characterize the performance of a QIM system. The number of quantizers in the ensemble equals the number of possible values for m, and hence, determines the information-embedding rate. The size and shape of the quantization cells determine the embedding-induced distortion. Finally, the minimum distance $d_{\min}$ between the sets of reconstruction points of different quantizers in the ensemble determines the robustness of the embedding, where the minimum distance is defined as

$$d_{\min} \triangleq \min_{(i,j):i\neq j} \min_{(\mathbf{x}_i,\mathbf{x}_j)} \|\mathbf{s}(\mathbf{x}_i; i) - \mathbf{s}(\mathbf{x}_j; j)\|. \quad (4)$$

Intuitively, the minimum distance measures the size of perturbation vectors that can be tolerated by the system. For example, in the case of the bounded perturbation channel, the energy bound (3) implies that a minimum distance decoder is guaranteed to not make an error as long as

$$\frac{d_{\min}^2}{4N\sigma_n^2} > 1. \quad (5)$$

In the case of an additive white Gaussian noise channel with a noise variance of σ_n^2, at high signal-to-noise ratio the minimum distance also characterizes the error probability of the minimum distance decoder [7],

$$\Pr[\hat{m} \neq m] \sim Q\left(\sqrt{\frac{d_{\min}^2}{4\sigma_n^2}}\right).$$

The minimum distance decoder to which we refer simply chooses the reconstruction point closest to the received vector, i.e.,

$$\hat{m}(\mathbf{y}) = \arg\min_m \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{s}(\mathbf{x}; m)\|. \quad (6)$$

If, which is often the case, the quantizers $\mathbf{s}(\mathbf{x}; m)$ map $\mathbf{x}$ to the nearest reconstruction point, then (6) can be rewritten as

$$\hat{m}(\mathbf{y}) = \arg\min_m \|\mathbf{y} - \mathbf{s}(\mathbf{y}; m)\|. \quad (7)$$

3.2. Spread-transform dither modulation

Dithered quantizers [6], are quantizer ensembles where the quantization cells and reconstruction points of any given quantizer in the ensemble are shifted versions of the quantization cells and reconstruction points of any other quantizer in the ensemble. To embed information, we simply modulate the amount of the shift, which is called the dither vector, by the embedded signal, i.e., each possible embedded signal maps uniquely onto a different dither vector $\mathbf{d}(m)$. The host signal is quantized with the resulting dithered quantizer to form the composite signal. Specifically, we start with some base quantizer $\mathbf{q}(\cdot)$, and the embedding function is

$$\mathbf{s}(\mathbf{x}; m) = \mathbf{q}(\mathbf{x} + \mathbf{d}(m)) - \mathbf{d}(m).$$

This type of information embedding is called dither modulation.

A simple example of dither modulation that will be of interest in this paper is called binary spread-transform dither modulation (STDM) with uniform, scalar quantization of step size Δ. We assume that $1/N \leq R_m \leq 1$. One can convert a spread-spectrum system of the form $\mathbf{s}(\mathbf{x}; m) = \mathbf{x} + a(m)\mathbf{u}$ into a STDM system by replacing addition with quantization. Specifically, STDM involves the following steps:

- The NR_m information bits $\{b_1, b_2, \ldots, b_{NR_m}\}$ representing the embedded message m are error correction coded using a rate-k_u/k_c code to obtain a coded bit sequence $\{z_1, z_2, \ldots, z_{N/L}\}$, where

$$L = \frac{1}{R_m}(k_u/k_c).$$

- A spreading vector $\mathbf{u} \in \Re^L$ is chosen along with two sequences $d(\cdot, 0)$ and $d(\cdot, 1)$ of N/L dither values satisfying the constraint

$$d(i,1) = \begin{cases} d(i,0) + \Delta/2, & d(i,1) < 0 \\ d(i,0) - \Delta/2, & d(i,1) \geq 0 \end{cases}$$

for $1 \leq i \leq N/L$. For example, one could choose $d(i,0)$ pseudorandomly with a uniform distribution over $[-\Delta/2, \Delta/2]$.[1]

- The projection of the i-th length-L subvector of $\mathbf{x}$ onto $\mathbf{u}$ is quantized with a dithered quantizer using dither value $d(i, z_i)$. (In contrast, a spread spectrum system adds some value to this projection.)

An alternative form of dither modulation is presented in [5].

If the error correction code is a binary block code with a minimum Hamming distance of d_H, then the possible sequences of dither values will differ in at least d_H places, and the reconstruction points of the corresponding dithered quantizers are shifted by $\pm\Delta/2$ in each of these places. Thus, the minimum distance squared, as defined by (4), is

$$d^2_{\min} = d_H \left(\frac{\Delta}{2}\right)^2 = \gamma_c \frac{1}{LR_m}\left(\frac{\Delta}{2}\right)^2, \qquad (8)$$

where $\gamma_c = d_H(k_u/k_c)$. If the quantization cells are sufficiently small such that the source signal can be modeled as uniformly distributed within each cell, the expected squared-error distortion per dimension (2) of the uniform, scalar quantizers is

$$D_s = \frac{1}{L\Delta}\int_{-\Delta/2}^{\Delta/2} x^2\, dx = \frac{\Delta^2}{12L}. \qquad (9)$$

This information about the minimum distance and expected distortion can be combined to characterize the achievable performance of coded dither modulation, as is done in the next few sections.

[1] A uniform distribution for the dither sequence implies that the quantization error is statistically independent of the host signal and leads to fewer "false contours", both of which are generally desirable properties from a perceptual viewpoint [6].

4. Bounded perturbation channels

In this section we characterize the achievable performance of STDM, spread spectrum, and low-bit(s) modulation against bounded perturbation attacks.

Recall, the guaranteed error-free decoding condition for a minimum distance decoder (7) is given by (5). By substituting (8) and (9) into (5), one can precisely express the achievable trade-offs for STDM as

$$\boxed{\gamma_c \frac{3}{4}\frac{1}{NR_m}\frac{D_s}{\sigma_n^2} > 1.} \qquad (10)$$

(This expression also applies for other forms of dither modulation [5].) Thus, for example, at a fixed rate R_m to tolerate more perturbation energy σ_n^2 requires that we accept more expected distortion D_s. We also see that γ_c is the improvement or gain due to the error correction code. For example, an uncoded system has $\gamma_c = 1 = 0$ dB.

Thus, the nonzero minimum distance of QIM systems offers quantifiable robustness to perturbations, even when the host signal is not known at the decoder. In contrast, spread-spectrum based systems offer relatively little robustness to perturbations if the host signal is not known at the decoder. These systems embed information by adding a pseudo-noise vector $\mathbf{w}(m)$ to the host signal, i.e., $\mathbf{s}(\mathbf{x}, m) = \mathbf{x} + \mathbf{w}(m)$.

The minimum distance of a spread-spectrum system is zero, which can be seen by setting $\mathbf{x}_j = \mathbf{x}_i + \mathbf{w}(i) - \mathbf{w}(j)$ during the minimization over $(\mathbf{x}_i, \mathbf{x}_j)$ in (4).[2] Thus, although these systems may be effective when the host signal is known at the decoder, in the more typical case when the host signal is not known, they offer no guaranteed robustness to perturbations, and hence, no expression analogous to (10) exists. As alluded to in Sec. 1, in a spread-spectrum system $\mathbf{x}$ is an additive interference that is often much larger than $\mathbf{w}$ due to the distortion constraint. The quantization that occurs with quantization index modulation, however, provides immunity against this host signal interference.

Although LBM systems also have nonzero minimum distance, the achievable performance trade-offs in this case are worse than those of dither modulation (10) by 2.43 dB [5].

5. Bounded host-distortion and in-the-clear attacks

As mentioned in Sec. 2, some attackers may exploit partial knowledge of the host signal. In these cases a bounded

[2] It may be tempting to define the minimum distance as $d_{ss} = \min_{i,j} \|\mathbf{w}(i) - \mathbf{w}(j)\|$. However, this distance determines the maximum tolerable $\|\mathbf{x} + \mathbf{n}\|$. In the typical cases where $\|\mathbf{x} + \mathbf{n}\| \gg \|\mathbf{n}\|$, a nonzero d_{ss} cannot be used to guarantee robustness against bounded perturbation attacks.

Table 1. Attacker's distortion penalties. The distortion penalty is the additional distortion that an attacker must incur to successfully remove a watermark.

Embedding System	Distortion Penalty $(D_{\mathbf{y}}/D_{\mathbf{s}})$
Quant. Index Mod.	$1+\frac{1}{4}\frac{d^2_{\min}/N}{D_{\mathbf{s}}} > 0$ dB
Spread Trans. Dith. Mod.	$1+\gamma_c\frac{3/4}{NR_m} > 0$ dB
Spread Spectrum	$-\infty$ dB
LBM	≤ 0 dB

host-distortion channel model, rather than a bounded perturbation channel model, may be appropriate.

In addition, these attackers may also exploit knowledge about the embedding and decoding processes. To limit the attackers' knowledge, some digital watermarking systems use keys, which allow only appropriate parties to embed and/or decode the embedded signal. However, in some scenarios it may be desirable to allow everyone to embed and decode watermarks without keys. For example, in a copyright notification system, everyone could embed the ASCII representation of a copyright notice such as, "Property of ..." in their copyrightable works. Such a system is analogous to the system currently used to place copyright notices in (hardcopies of) books, a system in which there is no need for a central authority to store or maintain separate keys or watermarks for each user. The widespread use of such a "no-key" system in which the watermark is "in the clear" requires only standardization of the decoder so that everyone will agree on the decoded watermark, and hence, the owner of the copyright.

In these scenarios, the ratio between $D_{\mathbf{y}}$ and $D_{\mathbf{s}}$ is the distortion penalty that an in-the-clear attacker must pay to remove the watermark and is a figure of merit measuring the trade-off between robustness and embedding-induced distortion at a given rate. Distortion penalties for QIM, spread-spectrum, and LBM systems are derived below and are shown in Table 1. We see that of the three systems considered, only QIM systems are robust enough such that the attacker must degrade the host signal quality to remove the watermark.

5.1. Quantization index modulation

We first consider the robustness of quantization index modulation. We assume that all reconstruction points $\mathbf{s}$ lie at the centroids of their respective quantization cells. Therefore, for any $\mathbf{y}$ that is a distance $\|\mathbf{n}\|$ away from $\mathbf{s}$,

$$D_{\mathbf{y}} = D_{\mathbf{s}} + \frac{\|\mathbf{n}\|^2}{N},$$

which is shown formally in [5]. For a successful attack, $\|\mathbf{n}\| \geq d_{\min}/2$ so our figure of merit for a quantization index modulation system is

$$\frac{D_{\mathbf{y}}}{D_{\mathbf{s}}} \geq 1 + \frac{1}{4}\frac{d^2_{\min}/N}{D_{\mathbf{s}}}. \tag{11}$$

Thus, for any QIM system of nonzero $d_{\min}$, the attacker's distortion penalty is always greater than 1 (0 dB), indicating that to remove the watermark, the attacker must degrade the host signal quality beyond the initial distortion caused by the embedding of the watermark.

In the special case of binary STDM with uniform, scalar quantization, Eq. (8) gives $d^2_{\min}$ and Eq. (9) gives the distortion $D_{\mathbf{s}}$. Thus, the attacker's distortion penalty (11) that must be paid to defeat the watermark in this case is

$$\boxed{\frac{D_{\mathbf{y}}}{D_{\mathbf{s}}} \geq 1 + \gamma_c\frac{3/4}{NR_m},}$$

the same distortion penalty as for other forms of dither modulation [5]. We see that the distortion penalty increases with the power γ_c of the error correction code.

5.2. Spread-spectrum modulation

The embedding function of a spread-spectrum system is $\mathbf{s} = \mathbf{x} + \mathbf{w}(m)$ so the resulting distortion is $D_{\mathbf{s}} = \|\mathbf{w}\|^2/N > 0$. An in-the-clear attacker can decode the message m and subtract the corresponding pseudo-noise vector $\mathbf{w}(m)$ from $\mathbf{s}$, completely removing the watermark and obtaining the original host signal in the process, $\mathbf{y} = \mathbf{s} - \mathbf{w}(m) = \mathbf{x}$. Hence, the resulting distortion penalty is

$$\frac{D_{\mathbf{y}}}{D_{\mathbf{s}}} = \frac{0}{D_{\mathbf{s}}} = -\infty \text{ dB}.$$

5.3. Low-bit(s) modulation

The embedding function of a LBM system never alters the most significant bits of the host signal, so one possible attack is to simply remodulate the least significant bits of $\mathbf{s}$ with some message $m' \neq m$. Then, both $\mathbf{s}$ and $\mathbf{y}$ are low-bit(s) modulated versions of $\mathbf{x}$, so their distortions must be equal, particularly if the distortions are averaged over all possible choices of m and m'. Thus, the attacker's distortion penalty in this case is

$$\frac{D_{\mathbf{y}}}{D_{\mathbf{s}}} = 1 = 0 \text{ dB},$$

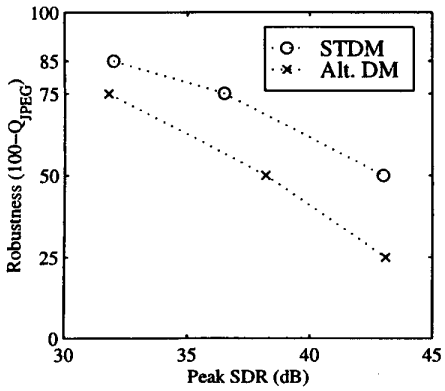

Figure 3. Achievable robustness-distortion trade-offs of dither modulation on the JPEG channel. $R_m = 1/320$. The bit-error rate is less than 5×10^{-6}.

This result applies regardless of whether error correction coding is used. Also, although the distortion penalty for this particular attack is 0 dB, this attack is not necessarily the best that an attacker could choose, so 0 dB is only an upper bound on the distortion penalty.

6. JPEG channels

The robustness of digital watermarking algorithms to common lossy compression algorithms such as JPEG is of considerable interest. A natural measure of robustness is the worst tolerable JPEG quality factor for a given bit-error rate at a given distortion level and rate. However, it is difficult to obtain expressions analogous to (10) in closed form. Fortunately, one can find individual achievable operating points by numerical simulation, as we demonstrate in this section for both STDM and the alternative form of dither modulation described in [5].

Achievable distortion-robustness trade-offs at an embedding rate of $R_m = 1/320$ bits per grayscale pixel are shown in Fig. 3 at various JPEG quality factors (Q_{JPEG}). The peak SDR is defined as the ratio between the square of the maximum possible pixel value and the average embedding-induced distortion per pixel. The host and composite signals, both 512-by-512 images, are shown in Fig. 4. The actual embedding is performed in the DCT domain using 8-by-8 blocks ($f_1, f_2 \in \{0, 1/16, \ldots, 7/16\}$) and low frequencies ($f_1^2 + f_2^2 \leq 1/4$), with 1 bit embedded across 5 DCT blocks. STDM is better than the alternative form of dither modulation [5] by about 5 dB at $100 - Q_{JPEG}$ of 50 and 75.

Although no bit errors occurred during the simulations

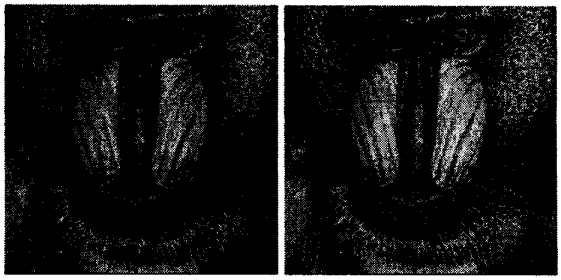

Figure 4. Host (left) and composite (right) image. After 25%-quality JPEG compression of the composite image, all bits were extracted without error. $R_m = 1/320$. Peak SDR of composite image is 36.5 dB.

used to generate Fig. 3, we estimate the bit-error rate to be at most 5×10^{-6}. At an embedding rate of 1/320, one can only embed 819 bits in the host signal image, which is not enough to measure bit-error rates this low. However, one can estimate an upper bound on the bit-error rate by measuring the bit-error rate ϵ at an embedding rate five times higher ($R = 1/64$) and calculating the coded bit-error probability of a rate-1/5 repetition code when the uncoded error probability is ϵ assuming independent errors, which can approximately be obtained by embedding the repeated bits in spatially separated places in the image.

References

[1] J. M. Barton. Method and apparatus for embedding authentication information within digital data. United States Patent #5,646,997. Issued July 8, 1997.

[2] B. Chen and G. W. Wornell. Dither modulation and quantization index modulation: New methods for digital watermarking and information embedding. Preprint.

[3] B. Chen and G. W. Wornell. System, method, and product for information embedding using an ensemble of non-intersecting embedding generators. U.S. patent pending. Licensing information: MIT Technology Licensing Office.

[4] B. Chen and G. W. Wornell. Digital watermarking and information embedding using dither modulation. In *Proc. of MMSP-98*, pp. 273–278, Redondo Beach, CA, Dec. 1998.

[5] B. Chen and G. W. Wornell. Dither modulation: A new approach to digital watermarking and information embedding. In *Proc. of SPIE: Security and Watermarking of Multimedia Contents*, vol. 3657, San Jose, CA, Jan. 1999.

[6] N. S. Jayant and P. Noll. *Digital Coding of Waveforms*. Prentice-Hall, 1984.

[7] E. A. Lee and D. G. Messerschmitt. *Digital Communication*. Kluwer Academic Publishers, 2nd edition, 1994.

[8] M. D. Swanson, M. Kobayashi, and A. H. Tewfik. Multimedia data-embedding and watermarking technologies. *Proc. of the IEEE*, 86(6):1064–1087, June 1998.