DIGITAL WATERMARKING AND INFORMATION EMBEDDING USING DITHER MODULATION

Brian Chen and Gregory W. Wornell
Department of Electrical Engineering and Computer Science,
and Research Laboratory of Electronics
Massachusetts Institute of Technology
Cambridge, MA

Abstract - A variety of related applications have emerged recently that require the design of systems for embedding one signal within another signal. We propose a new class of embedding methods called quantization index modulation (QIM) and develop an example of such a method called dither modulation in which the embedded information modulates the dither signal of a dithered quantizer. We also develop a framework within which one can analyze performance trade-offs among robustness, distortion, and embedding rate, and 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 that require the design of systems for embedding one signal, sometimes called an "embedded signal" or "watermark", within another signal, called a "host signal". 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. These applications include copyright notification and enforcement, authentication, and transmission of auxiliary information. These and other applications are described in [1], which also provides an overview of several proposed information-embedding algorithms.

Many algorithms belong to one of two classes: (1) additive techniques such as spread-spectrum in which a small pseudo-noise signal is added to the host signal and (2) quantize-and-replace strategies that replace a quantized host signal with another quantization value. A common example belonging to the second class is low-bit(s) modulation (LBM) in which the least significant bit(s) of the host signal are replaced by the embedded signal.

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.

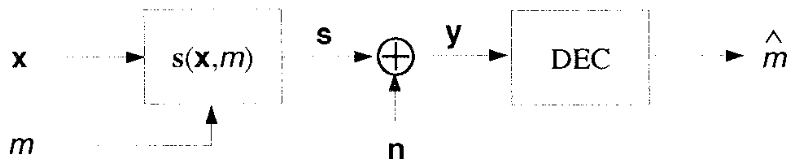

Figure 1: General information embedding problem model. An integer m is embedded in the host signal $\mathbf{x}$. A perturbation vector $\mathbf{n}$ corrupts the composite signal $\mathbf{s}$. The decoder extracts an estimate $\hat{m}$ of m from the channel output $\mathbf{y}$.

There has been relatively little performance analysis and much work remains to characterize the inherent trade-offs among the robustness of the embedding, the degradation to the host signal caused by the embedding, and the amount of data embedded. In this paper we introduce a framework for characterizing these trade-offs and develop a class of information-embedding systems, quantization index modulation (QIM) systems [2], that perform these trade-offs efficiently. We demonstrate that dither modulation, an example of a QIM system, offers significant advantages over previously proposed spread-spectrum and LBM techniques.

2. PROBLEM MODEL

Although a variety of information-embedding applications exist, many of these 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 signal could be a vector of pixel values or Discrete Cosine Transform (DCT) coefficients from an image, for example. We wish to embed at a rate of R bits per dimension (bits per host signal sample) so we can think of m as an integer chosen from the set $\{1, 2, \ldots, 2^{NR}\}$. An embedding function maps $\mathbf{x}$ and 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

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

The composite signal is passed 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 model the combined effects of these manipulations by the addition of a noise or perturbation vector $\mathbf{n} \in \Re^N$, which can be random or deterministic, signal independent or signal dependent. Thus, this channel model is completely general. However, we assume that the channel output $\mathbf{y}$ must still be a fair representation of the original signal so in this paper we often bound the energy of the perturbation vector,

$$||\mathbf{n}||^2 \leq N\sigma_n^2. \tag{2}$$

The decoder forms an estimate $\hat{m}$ of m based on $\mathbf{y}$. We can quantify the robustness of the system by the maximum allowable σ_n^2 such that we can still guarantee that $\hat{m} = m$. Alternatively, particularly if we wish to model $\mathbf{n}$ as random, we could characterize the reliability of the system by the probability of message error $\Pr[\hat{m} \neq m]$ or bit-error rate. The problem we face is to design an embedding function $s(\mathbf{x}, m)$ that achieves the best possible trade-off among the three parameters rate, distortion, and robustness (or reliability).

3. QUANTIZATION INDEX MODULATION

Specifying the performance requirements of an information-embedding system in terms of rate, distortion, and robustness leads quite naturally to the notion of quantization index modulation (QIM), as we develop in this section. In the last section, we consider the embedding function $s(\mathbf{x}, m)$ to be a function of two variables, the host signal and the embedded information. However, we can also view $s(\mathbf{x}, m)$ as a collection or ensemble of functions of $\mathbf{x}$, indexed by m. Henceforth, we denote the functions in this ensemble as $s(\mathbf{x}; m)$ to emphasize this view. The rate R determines the number of possible values for m, and hence, the number of functions in the ensemble. The distortion constraint suggests that each function in the ensemble is close to an identity function so that $s(\mathbf{x}; m) \approx \mathbf{x}$ for all m. That the system needs to be robust to noise suggests that the points in the range of one function in the ensemble should be "far away" in some sense from the points in the range of any other function. At the very least, the ranges should be non-intersecting. Otherwise, even in the absence of any noise, there will be some values of s from which one will not be able to uniquely determine m. This property, when considered with the near-identity property, suggests that the functions be discontinuous. Quantizers are just such a class of discontinuous, approximate-identity functions. QIM 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.

Figure 2 illustrates QIM information embedding for the $N = 2$ and $R = 1/2$ case. In this example, one bit is to be embedded so that $m \in \{1, 2\}$. The reconstruction points in $\Re^N$ of the two required quantizers are represented in Fig. 2 with $\times$'s and $\circ$'s. If $m = 1$, for example, $\mathbf{x}$ is quantized with the $\times$-quantizer, i.e., $\mathbf{s}$ is chosen to be the $\times$ closest to $\mathbf{x}$. If $m = 2$, $\mathbf{x}$ is quantized with the $\circ$-quantizer.

A few parameters of the ensemble conveniently characterize the performance of a QIM system. As noted above, the number of quantizers in the ensemble 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,

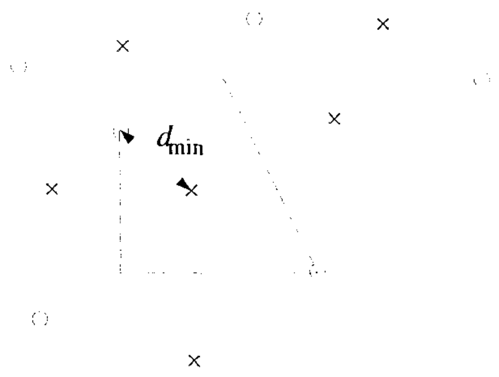

Figure 2: Quantization index modulation. The reconstruction points marked with ×'s ($m = 1$) and o's ($m = 2$) belong to two different quantizers. The minimum distance $d_{\min}$ measures the robustness to noise, and the sizes of the quantization cells, one of which is shown in the figure, determine the embedding-induced distortion.

where the minimum distance is defined as

$$d_{\min} \triangleq \min_{(i,j):i \neq j} \min_{(\mathbf{x}_i, \mathbf{x}_j)} \|\mathbf{s}(\mathbf{x}_i; i) - \mathbf{s}(\mathbf{x}_j; j)\|. \tag{3}$$

Intuitively, the minimum distance measures the size of noise vectors that can be tolerated by the system. For example, with bounded noise energy (2) a minimum distance decoder, which chooses the reconstruction point closest to the channel output, is guaranteed to not make an error as long as

$$\frac{d_{\min}^2}{4N\sigma_n^2} > 1. \tag{4}$$

Alternatively, for additive white Gaussian noise with variance σ_n^2, the error probability is $\sim Q(d_{\min}/(2\sigma_n))$ at high signal-to-noise ratio [3].

From the preceding discussion, we see that the non-zero minimum distance of QIM systems offers quantifiable robustness to noise. In contrast, spread-spectrum systems offer relatively little robustness to noise. 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 these systems is actually 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 Eq. (3). Thus, although these systems may be effective when the host signal is known at the decoder, in the often more typical case where the host signal is not known, they offer no guaranteed robustness to noise. Intuitively, when the host signal is not known at the decoder, it is a source of noise. With a spread-spectrum system, $\mathbf{x}$ is an additive noise that is often much larger than $\mathbf{w}$ due to the distortion constraint. The quantization that occurs with QIM, however, removes much of the noisiness introduced by $\mathbf{x}$ by reducing the number of possible values. We further quantify this robustness advantage in Sec. 4.

4. DITHER MODULATION

Dithered quantizers [4] are quantizer ensembles where the quantization cells and reconstruction points of every quantizer in the ensemble are shifted versions of some base quantizer $\mathbf{q}(\cdot)$. The shift is given by a dither vector $\mathbf{d}$, which in non-watermarking contexts is typically chosen pseudorandomly. In a dither modulation (DM) system, however, the dither vector is modulated by the embedded information. Specifically, we define a dither vector $\mathbf{d}(m)$ for each possible value of m. Thus, the embedding function is $\mathbf{s}(\mathbf{x}; m) = \mathbf{q}(\mathbf{x} + \mathbf{d}(m)) - \mathbf{d}(m)$.

As a simple example, we consider the case where $\mathbf{q}(\cdot)$ is a uniform, scalar quantizer with step size Δ, R is between $1/N$ and 1, and the NR bits in m are used to binary amplitude modulate a length-R^{-1} pseudorandom sequence of $\pm\Delta/4$ with these NR sequences concatenated to form $\mathbf{d}(m)$. The reconstruction points of the quantizers in this case lie on hypercubic grids in $\Re^N$, the points of a given quantizer shifted by $\pm\Delta/2$ in each dimension relative to the points of any other quantizer over at least R^{-1} dimensions. Thus, the minimum distance (3) is $d_{\text{min}} = \sqrt{R^{-1}(\Delta/2)^2}$. If the quantization cells are sufficiently small such that $\mathbf{x}$ can be modeled as uniformly distributed within each cell, the expected squared-error distortion per sample (1) of a uniform, scalar quantizer is $\Delta^2/12$. Thus, with bounded noise energy and a minimum distance decoder, (4) can be used to compactly express the trade-off among distortion, robustness, and rate as

$$\boxed{\frac{3}{4}\frac{1}{NR}\frac{E[D(\mathbf{s},\mathbf{x})]}{\sigma_n^2} > 1.} \tag{5}$$

Thus, for example, at a fixed rate R to tolerate more noise energy σ_n^2 requires that we accept more expected distortion $E[D]$. Eq. (5) is convenient in relating design specifications to design parameters. For example, if the design specifications require an embedding rate of at least R and robustness to noise of at least σ_n^2 in energy per sample, then (5) gives the minimum embedding-induced distortion that must be introduced into the host signal, or equivalently the minimum quantization step size Δ, to achieve these specifications. Similar relationships to (5) for other QIM and DM systems [2], for example those employing error correction codes, can also be derived [5].

As mentioned in Sec. 3, spread-spectrum systems have $d_{\text{min}} = 0$, so no condition analogous to (5) exists under which error-free decoding is guaranteed. Furthermore, analysis in [5] establishes that LBM is 2.43 dB worse than DM in this case of bounded noise.

DM outperforms spread-spectrum not only in the bounded noise case, but also in the additive Gaussian noise case as well. A plot of bit-error rates, as measured by Monte Carlo simulations, is shown in Fig. 3. The host signal vectors are DCT coefficients of 8 × 8 blocks of an image. We see that at typical host SNRs of 30-40 dB, the host signal is too large an effective noise source for the spread-spectrum decoder to overcome.

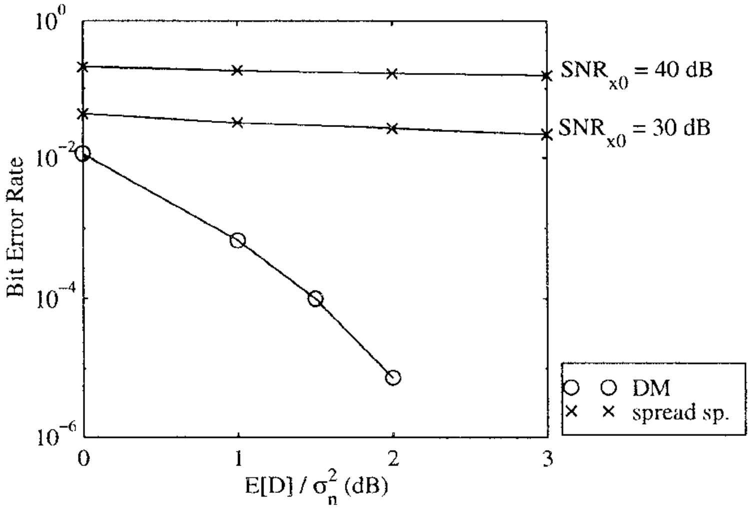

Figure 3: Dither modulation vs. spread spectrum on the additive Gaussian noise channel. $E[D]$ denotes expected distortion, σ_n^2 denotes noise variance, and SNR_{x0} is the ratio between $||\mathbf{x}||^2$ and $N\sigma_n^2$. $R = 1/64$ bits per pixel.

Finally, DM systems also outperform spread-spectrum and LBM systems when an adversary tries to remove the watermark by exploiting full knowledge of the embedding and decoding algorithms, including any keys, and any partial knowledge of **x** that might be available in the form of a probability density function, while working under an expected distortion constraint between **y** and **x**. Significantly, while DM systems can benefit from error-correction coding in this case, spread-spectrum and LBM systems do not [5]. Several other results on coded QIM systems are also developed in [5].

References

[1] M. D. Swanson, M. Kobayashi, and A. H. Tewfik, "Multimedia data-embedding and watermarking technologies," *Proceedings of the IEEE*, vol. 86, pp. 1064–1087, June 1998.

[2] 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 info.: MIT Technology Lic. Office.

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

[4] N. S. Jayant and P. Noll, *Digital Coding of Waveforms : Principles and Applications to Speech and Video.* Prentice-Hall, 1984.

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