Design and Analysis of Digital Watermarking, Information Embedding, and Data Hiding Systems

Part I: Concepts and System Overview

Author: Brian Chen
Ph. D. Dissertation, M. I. T.
Outline

- Introduction (Ch1)
- Mathematical modeling of digital watermarking (Ch2)
- Quantization index modulation (Ch3)
- Dither modulation (Ch5)
Information Embedding Systems

- Information embedding systems embed one signal, the so-called “embedded signal” or “watermark”, within another signal, a “host signal”.
  - The embedding must be robust to common degradations or intentional attacks.
  - As long as the host survives, the watermark should survive.
Information Embedding Applications

- Owner identification
- Copyright enforcement
- Broadcast monitoring
- Transaction monitoring
- Copy control
- Device control
- Data authentication
General Information Embedding Model

- An integer $m \in \{1, 2, \ldots, 2^{NRm}\}$ is embedded in the host signal vector $x \in R^N$ using some embedding function $s(x, m)$.
- $n$ is a perturbation vector corrupts the composite signal $s \in R^N$.
- The decoder extracts an estimate $\hat{m}$ of $m$ from the noisy channel output $y$. ($x$ is not always available)
Distortions Caused by Information Embedding and Channel Noises

- The degradation to the host signal must be acceptable and is usually measured by some distortion function between the host and the composite signal.

\[
D(s, x) = \frac{1}{N} \| s - x \|^2 \quad \text{Squared error distortion measure}
\]

\[
D(s, x) = \frac{1}{N} (s - x)^T W (s - x)
\]

- s is subject to a variety of signal corrupting manipulations such as lossy compression, addition of random noise, .... The degradations caused by the channel should not be too large, either.
Conflicting Goals

- Three conflicting goals of information embedding systems we want to achieve
  - Low distortion
  - High robustness
  - High rate

- The performance of a watermarking system is evaluated in terms of its achievable rate-distortion-robustness trade-offs.
**Equivalent Super-Channel Model**

- \( s(x,m) = e(x,m) + x \)
- The host signal \( x \) is viewed as a known state of this super-channel
- Information embedding problems are viewed as power-limited communication over a super-channel.

\[
\frac{1}{N} \left\| s - x \right\|^2 = \frac{1}{N} \left\| e \right\|^2
\]
Considering Channel Models

- Different viewpoints
  - Description of the degradations that can occurred to the composite signal
  - Description of the degradations against which we wish the embedder and decoder to be against

- Described probabilistically or deterministically

- Examples
  - Bounded perturbation channels: \[ \| y - s \| ^2 = \| n \| ^2 \leq \mathcal{N} \sigma_n^2 \]
    - Bounds the energy of the perturbation vector
  - Additive noise channels
    - The perturbation vector \( n \) is random and statistically independent of \( s \)
Classifying Embedding Methods

- Classifying by different types of signals
  - Not suitable for high-level discussions
- Classifying by different effects of host-signals due to system design
  - Host-interference non-rejecting methods
  - Host-interference rejecting methods

Only blind-host systems are of concern during this discussion
Host-interference Non-rejecting Methods

- The host signal is a source of interference in the system.
- The encoder do not sufficiently exploit knowledge of the host signal \( x \).
- Most currently important watermarking techniques are of this type.

\[
\begin{align*}
\mathbf{s}(x, m) &= x + \mathbf{w}(m) \\
\tilde{s} &= \mathbf{s}^T \nu = \tilde{x} + a(m) \\
\mathbf{w}(m) &= a(m) \nu \\
a(m) &= \tilde{s} - \tilde{x} \\
\mathbf{s} &= x + a(m) \nu \\
\mathbf{s} &= x + (\tilde{s} - \tilde{x}) \nu
\end{align*}
\]
Geometrical Interpretation

- $x$ and $n$ can be of any direction and constrained amplitude
Specific Weighted Embedding Functions

\[ \alpha(x) = \lambda x \]
\[ s = x + \lambda x v = x(1 + \lambda v) \]
\[ \log s = \log x + \log(1 + \lambda v) \]

- Additive in the log domain
- Thus, still host-interference non-rejecting in the log domain
- Host signal will still limit the performance
Host-interference Rejecting Example (I)

Host-interference Rejecting Example (II)

- The DCT coefficient is projected to a pseudo-random vector.
- Another projection vector can be of use to embed more than 1 bit.
The quantization step $T$ is decided by determining the frequency masking.

The host signal determines the particular $x$ or $0$ that is chosen, but it does not inhibit the decoder’s ability to determine whether it is a $0$ or $x$. 

\[ \tilde{s} = q(\tilde{x}) + d(m) \]

\[ s = x + (\tilde{s} - \tilde{x})v \]
Generalized Least Bit Modulation (LBM)

- The quantization-and-perturbation process is equivalent to
  - Quantize $x$ with a quantizer of step size $\frac{\Delta}{2}$ whose reconstruction points are the union of $0$ and $x$.
  - Modulate the least significant bit in the bit label with the watermark bit to arrive at a composite signal bit label.
Characteristics of LBM

- The set of embedding intervals corresponding to a given value are the same to the set of embedding intervals to all other intervals.
- This is an unnecessary constraint for data embedding system. (Later, QIM will take advantage of this.)

\[
\{I_i(s_i) \mid s_i \in S_i\} = \{I_j(s_j) \mid s_j \in S_j\}, \quad \forall i, j \in \{1, \ldots, 2^{NR_m}\}
\]

\[
I_m(s_0) = \{x \mid s(x, m) = s_0\}
\]
Designing the Embedding Function (I)

- The embedding function can be viewed as an ensemble of functions of \( x \), indexed by \( m \).
- Since the embedding-introduced distortion should be small, each function in the ensemble must be close to an identity function.
- The system needs to be robust to perturbations. Thus 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.
  - E.g. non-intersecting

\[
\begin{align*}
S(x, m) & \Rightarrow S(x; m) \\
S(x; m) & \approx x, \forall m
\end{align*}
\]
Designing the Embedding Function (II)

- Approximate-identity property
  - the range of the functions should cover the space of possible $x$
- Non-intersecting property
  - discontinuous functions
- Quantization is such a discontinuous and approximate-identity function.
Quantization Index Modulation

- One bit is to be embedded, i.e.
- Two quantizers are required. The reconstruction points of each are represented with x and o.
  - If \( m = 1 \), \( x \) is quantized with the x-quantizer, i.e. \( s \) is chosen to be the x closest to \( x \).
  - If \( m = 2 \), \( x \) is quantized with the o-quantizer.
Convenience from an Engineering Perspective

- Information-embedding rate
  - The number of quantizers

- Embedding-introduced distortion
  - The size and shapes of the quantization cells

- Robustness of the embedding
  - The minimum distance between the sets of reconstruction points of different quantizers.

- The minimum distance decoder

\[
\begin{align*}
    d_{\min} &\equiv \min_{(i,j):i\neq j} \min_{(x_i,x_j)} \| s(x_i;i) - s(x_j;j) \| \\
    d_{\min} &\equiv \min_{(i,j):i\neq j} \| s(x;i) - s(x;j) \| \\
    \hat{m}(y) = \arg\min_{m} \min_{x} \| y - s(x;m) \| 
\end{align*}
\]
Robustness against Different Channels

- The bounded perturbation channel

\[ N\sigma_n^2 > \| n \|^2 > \left( \frac{d_{\text{min}}}{2} \right)^2 \Rightarrow \frac{d_{\text{min}}^2}{4N\sigma_n^2} < 1 \]

- The additive white Gaussian noise channel

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

\[ Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{t^2}{2}} \, dt \]
The embedding intervals for a pair of reconstruction points in generalized LBM are the same => The union of the two quantization cells.

In QIM, the quantization cells of the quantizers need not to be the same.
QIM v.s. Generalized LBM (II)

- The embedding caused distance are smaller in QIM than in generalized LBM though the minimum distance in the two cases are the same.
Dithering in Coarse Quantizers

- Techniques used to enhance quantization of coarse quantizers
Dither Modulation

- Quantization cells and reconstruction points of any quantizer in the ensemble are shifted versions of the quantization cells and reconstruction points of any other quantizer.
- Each embedded message is uniquely mapped to a different dither vector $d(m)$

$$s(x; m) = q(x + d(m)) - d(m)$$
Dither Modulation System (I)

- The NR_m 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=(k_u/k_c)/R_m

- We divided the host signal x into N/L non-overlapping blocks of length L and embed the i-th coded bit z_i in the i-th block.
Dither Modulation System (II)

- Two length $L$ dither sequence $d[k,0]$ and $d[k,1]$ and one length $L$ sequence of uniform, scalar quantizers with step sizes $\Delta_1, \ldots, \Delta_k$ are constructed by

\[
d[k,1] = \begin{cases} 
  d[k,0] + \frac{\Delta_k}{2}, & d[k,0] < 0 \\
  d[k,0] - \frac{\Delta_k}{2}, & d[k,1] \geq 0 
\end{cases}, \quad k = 1, \ldots, L
\]
Dither Modulation System (III)

- The i-th block of x is quantized with the dithered quantizer using the dither sequence $d[k, z_i]$

- The corresponding minimum distance decoder

$$\hat{z}_i = \arg\min_{l \in \{0, 1\}} \sum_{k=(i-1)L+1}^{iL} (y[k] - s_y[k; l])^2, i = 1, ..., N / L$$
Dither Modulation System (IV)
Minimum Distance

- The square of the minimum distance over all $N$ dimensions

$$d_{\text{min}}^2 = d_H \sum_{k=1}^{l} \left( \frac{\Delta_k}{2} \right)^2$$

Hamming distance of the error correction code

$$= \left( d_H \frac{k_u}{k_c} \right) \frac{1}{4LR_m} \sum_k \Delta_k^2$$

Gain of the error correction code

$$= \gamma_c \frac{1}{4LR_m} \sum_k \Delta_k^2$$
Distortion-normalized Squared Minimum Distance

- The overall average expected distortion

\[ D_s = \frac{1}{12} L \sum_k \Delta_k^2 \]

- The distortion-normalized squared minimum distance

\[ d_{\text{norm}}^2 = \frac{d_{\text{min}}^2}{D_s} = \frac{3\gamma_c}{R_m} \]
Spread-Transform Dither Modulation (I)

- The reconstruction points of two quantizers for embedding one bit in a block of two samples where the quantization step sizes are the same for both samples

\[ d_{\text{min}} = \Delta / \sqrt{2} \]
Apply a unitary transform first so that the first transform coefficient is the component of the host signal in the direction of \( v \). The second transform coefficient is the component orthogonal to \( v \).
Spread-Transform Dither Modulation (III)

- Since in both cases, the minimum distance and the average square error distortion are the same, the robustness against bounded perturbations is the same in both cases.
- However, since the decision region of the later case contains the decision region of the first one, for additive noise channels, the probability of a correct decision is higher in the later case.
Spread-Transform Dither Modulation (IV)

\[ s(x, m) = (\tilde{x} + a(m))v + (x - \tilde{x}v) \]
STDM v.s. AMSS

**STDM**

\[ \tilde{s} = q(\tilde{x} + d(m)) - d(m) \]

\[ \Delta = \sqrt{12LD_s} \]

\[ \min_{(\tilde{x}_1, \tilde{x}_2)} \left| s(\tilde{x}_1, 1) - s(\tilde{x}_2, 2) \right|^2 = \frac{\Delta^2}{4} = 3LD_s \]

\[ \tilde{y} = \tilde{s} + n \]

\[ SNR_{STDM} = \frac{3LD_S}{P(\tilde{n})} \]

**AMSS**

\[ \tilde{s} = \tilde{x} + a(m) \]

\[ a(m) = \pm \sqrt{LD_s} \]

\[ |a(1) - a(2)|^2 = 4LD_s \]

\[ \tilde{y} = a(m) + \tilde{x} + \tilde{n} \]

\[ SNR_{AMSS} = \frac{4LD_s}{P(\tilde{x} + \tilde{n})} \]

\[ \frac{SNR_{STDM}}{SNR_{AMSS}} = \frac{3}{4} \frac{P(\tilde{x} + \tilde{n})}{P(\tilde{n})} \]
STDM v.s. LBM

\[ s = q(x) + d(m) \]

\[ \frac{1}{N} E \left[ \| s - x \|^2 \right] = \frac{1}{N} E \left[ \| q(x) - x + d(m) \|^2 \right] \]

\[ = \frac{1}{N} E \left[ \| q(x) - x \|^2 \right] + \frac{1}{N} E \left[ \| d(m) \|^2 \right] \]

\[ = \left( \frac{1}{12L} + \frac{1}{16L} \right) \sum_k \Delta_k^2 \]

\[ d_{\text{min}}^2 = \gamma_c \frac{1}{4LR_m} \sum_k \Delta_k^2 \]

\[ d_{\text{norm}}^2 = \frac{12\gamma_c}{7R_m} \]
To be continued…

Part I: Concepts and System Overview
- Introduction (Ch1)
- Mathematical modeling of digital watermarking (Ch2)
- Quantization index modulation (Ch3)
- Dither modulation (Ch5)

Part II: Theoretical Basis
- Typicality (Ch3)
- Joint typicality (Ch8)
- Differential entropy (Ch9)
- Gaussian channel (Ch10)

- Communication with side information
- Writing on dirty paper

Part III: Information Theoretic Analysis
- Information theoretic respective (Ch4)
- Gaussian channels (Ch6)
- Intentional Attacks (Ch7)
- Conclusions (Ch8)