MULTIMEDIA COMMUNICATION SYSTEM

RATE-DISTORTION THEORY
Rate distortion theory - Introduction

- Lossy compression: Decoded signal is only an approximation of original signal
- Rate distortion theory
  - calculates the minimum transmission bit-rate $R$ for a required picture quality
  - Information theoretical bounds for lossy compression

$\hat{X}$

Image Source $\xrightarrow{X}$ Coder $\xrightarrow{\hat{X}}$ Decoder $\xrightarrow{\text{Distortion } d}$

Bitrate at least $R$ for distortion $d \leq D$

- Results of rate distortion theory are obtained without consideration of a specific coding method
Rate distortion theory - Introduction

Goal: Lower the bit-rate $R$ by allowing some acceptable distortion $D$ of the signal.
Information Theoretical Bounds for Lossy Compression

- Rate distortion theory
  - R-D function for memoryless Gaussian sources with MSE distortion criterion
  - R-D function for Gaussian sources with memory
  - R-D function for images

Distortion

- Symbol (signal, image . . . ) \( X \) sent, \( \hat{X} \) received
- Single-letter distortion measure:
  \[
  \rho(x, \hat{x}) \geq 0 \\
  \rho(x, \hat{x}) = 0 \text{ for } x = \hat{x}
  \]
- Average distortion:
  \[
  d(X, \hat{X}) = E\{\rho(X, \hat{X})\} = \sum_x \sum_{\hat{x}} f_{x, \hat{x}}(x, \hat{x}) \rho(x, \hat{x})
  \]
- Distortion criterion:
  \[
  d(X, \hat{X}) \leq \hat{D}
  \]

Maximum permissible average distortion
Mutual Information

"Mutual information" is the average information that random variables $X$ and $Y$ convey about each other

- Reduction in uncertainty about $X$, if $Y$ is observed
- Reduction in uncertainty about $Y$, if $X$ is observed

$$ I(X;Y) = H(X) - H(X \mid Y) = H(Y) - H(Y \mid X) $$

$$ = \sum_x \sum_y f_{X,Y}(x, y) \log_2 \frac{f_{X,Y}(x, y)}{f_X(x) f_Y(y)} $$

Properties

$$ 0 \leq I(X;Y) = I(Y;X) $$
$$ I(X;Y) \leq H(X) $$
$$ I(X;Y) \leq H(Y) $$
Mutual Information: extension to continuous RV’s

- **Differential entropy**
  \[
  h(X) = -E\{\log_2 f_X(X)\} = -\int f_X(x) \log_2 f_X(x) \, dx
  \]

- **Differential conditional entropy**
  \[
  h(X|Y) = -E\{\log_2 f_{X|Y}(X,Y)\} = -\iint f_{X,Y}(x,y) \log_2 f_{X|Y}(x,y) \, dx \, dy
  \]

- **Mutual information**
  \[
  I(X;Y) = h(X) - h(X|Y) = h(Y) - h(Y|X)
  \]

- **Rate distortion function**
  \[
  R(D) = \inf_{f_{\hat{X}|X}:d(X,\hat{X})\leq D} \{I(X;\hat{X})\}
  \]
Mutual Information: extension to continuous RV's

Example 3.1 Let $X$ be Gaussian with mean 0 and variance $\sigma^2$. Then,

$$f_X(x) = \frac{1}{\sqrt{2\pi \sigma^2}} e^{-x^2/2\sigma^2}$$

and

$$h(X) = - \int f_X(x) \log_2 f_X(x) dx$$

$$= \int f_X(x) \left[ \frac{x^2}{2\sigma^2} + \ln \sqrt{2\pi \sigma^2} \right] dx$$

$$= \frac{\left[ \frac{1}{2} + \ln \sqrt{2\pi \sigma^2} \right]}{\ln 2}$$

$$= \frac{1}{2} \log_2 2\pi e\sigma^2$$

(3.5)

(3.6)

Theorem 3.3 Let $X$ and $Y$ be random variables each of variance $\sigma^2$ and let $X$ be Gaussian. Then $h(X) \geq h(Y)$. 
Continuous Random Processes

- **Continuous random process**
  - Series of random experiments at time instants \( t_n \), with \( n = 0, 1, 2, \ldots \), characterized by a series of continuous random variables \( X = \{X_n\} \)
  - Statistical properties of discrete-time random process \( X \): N-th order joint pdf
    \[
    f_{X_n}(x) = f_{X_n}(x_n, x_{n+1}, \ldots, x_{n+N-1})
    \]

- **Characterization of continuous random processes**
  - **Stationary**: Statistical properties are invariant to a shift in time
  - **Memoryless**: The random variables \( X_n \) are independent of each other
  - **Independent and identically distributed (iid)**: Stationary and memoryless
  - **Markov process**: Future outcomes do not depend on past outcomes, but only on the present outcome,
    \[
    f_{X_n}(x_n|x_{n-1}, \ldots) = f_{X_n}(x_n|x_{n-1})
    \]
    - Can be generated from zero-mean memoryless process \( Z = \{Z_n\} \),
      \[
      X_n = \rho(X_{n-1} - \mu_X) + \mu_X + Z_n, \text{ with } |\rho|<1
      \]
  - **Gauss-Markov process**: Markov process for which the random variables \( Z_n \) of the memoryless process \( Z \) has a Gaussian distribution
Definition:

\[ R(D) = \inf_{f_{\hat{X}|X}:d(X,\hat{X})\leq D} \{ I(X;\hat{X}) \} \]

- Shannon’s Source Coding Theorem (and converse):
  For a given maximum average distortion \( D \), the rate distortion function \( R(D) \) is the (achievable) lower bound for the transmission bit-rate.

- \( R(D) \) is continuous, monotonically decreasing for \( R>0 \) and convex

- Equivalently use distortion-rate function \( D(R) \)
Shannon Lower Bound

- It can be shown that \( h(X - \hat{X} \mid \hat{X}) = h(X \mid \hat{X}) \)

- Thus
  \[
  R(D) = \inf_{d \leq D} \{ h(X) - h(X \mid \hat{X}) \} = h(X) - \sup_{d \leq D} \{ h(X \mid \hat{X}) \} = h(X) - \sup_{d \leq D} \{ h(X - \hat{X} \mid \hat{X}) \}
  \]

- Ideally, the source coder would introduce errors \( X - \hat{X} \) that are statistically independent from the reconstructed signal \( \hat{X} \) (not always possible!).

- Shannon lower bound:
  \[
  R(D) \geq h(X) - \sup_{d \leq D} h(X - \hat{X}) = R_L(D)
  \]
Shannon Lower Bound (cont.)

- Mean squared error distortion measure:
  Gaussian PDF possesses largest entropy for given variance

\[
R(D) \geq h(X) - \sup_{d \leq D} h(X - \hat{X}) \\
= h(X) - \frac{1}{2} \log_2 2\pi e D
\]

- Equivalently

\[
D(R) \geq \frac{1}{2\pi e} 2^{2h(X)} 2^{-2R}
\]

- Distortion reduction by 6 dB requires 1 bit/sampel
Corollary 3.5  For an IID Gaussian process with variance $\sigma^2$, the MSE rate-distortion function is given by

\[
R(D) = \frac{1}{2} \log_2 2\pi e \sigma^2 - \frac{1}{2} \log_2 2\pi e D
\]

\[
= \frac{1}{2} \log_2 \frac{\sigma^2}{D}
\]

(3.9)

Note that if $D \geq \sigma^2$, $R(D) = 0$. In this case, fixing $\hat{X} = E[X]$ yields

\[E[(X - \hat{X})^2] = \sigma^2 \leq D.\]

Inverting the rate-distortion function, we get the distortion-rate function

\[D(R) = \sigma^2 2^{-2R}\]  

(3.10)
Shannon Lower bound (cont.)

Examples for Shannon lower bound for MSE distortion

- Shannon lower bound differs only in $h(S')$ for various distributions

- **Uniform pdf:**

\[
h(S) = \frac{1}{2} \log_2(12\sigma^2) \quad \rightarrow \quad D_L(R) = \frac{6}{\pi e} \cdot \sigma^2 \cdot 2^{-2R}
\]

- **Laplacian pdf:**

\[
h(S) = \frac{1}{2} \log_2(2e^2 \sigma^2) \quad \rightarrow \quad D_L(R) = \frac{e}{\pi} \cdot \sigma^2 \cdot 2^{-2R}
\]

- **Gaussian pdf:**

\[
h(S) = \frac{1}{2} \log_2(2\pi e \sigma^2) \quad \rightarrow \quad D_L(R) = \sigma^2 \cdot 2^{-2R}
\]
R(D) for a memoryless Gaussian source & MSE dist

- Gaussian source, variance $\sigma^2$
- Mean squared error
  $$d = E\{(X - \hat{X})^2\} \leq D$$
- Rule of thumb: 6 dB $\equiv$ 1 bit
- $R(D)$ for non-Gaussian sources with the same variance $\sigma^2$ is always below this Gaussian $R(D)$ curve.
  $$R_L(D) \leq R(D) \leq \frac{1}{2} \log_2 \frac{\sigma^2}{D}$$

$$R(D) = \frac{1}{2} \log_2 \left( \frac{\sigma^2}{D} \right)$$

$$SNR = 10 \log_{10} \left( \frac{\sigma^2}{D} \right) \text{ [dB]}$$
R(D) for Gaussian source with memory

- Stationary, band-limited, jointly Gaussian source with power spectral density $\Phi_{xx} (\omega)$

- Mean squared error distortion $d = E\{(X - \hat{X})^2\} \leq D$

- $R(D)$ function in parametric form

\[
D(\theta) = \frac{1}{2\pi} \int_\omega \min\{\theta, \Phi_{xx}(\omega)\} d\omega \\
R(\theta) = \frac{1}{2\pi} \int_\omega \max\left\{0, \frac{1}{2} \log \frac{\Phi_{xx}(\omega)}{\theta}\right\} d\omega
\]

- $R(D)$ for non-Gaussian sources with the same power spectral density is always lower
R(D) for Gaussian source with memory
R(D) for Gaussian source with memory

\[ D(R) = \gamma_X^2 \sigma^2 2^{-2R} \]  \hspace{1cm} (3.15)

where

\[ \gamma_X^2 = \frac{1}{\sigma^2} \exp \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln S_X(\omega) d\omega \right] \]

is known as the spectral flatness measure of \( \{X_n\} \). It is easily shown that \( \gamma_X^2 \leq 1 \) with equality if and only if \( S_X(\omega) = \sigma^2 \forall \omega \) (i.e., \( \{X_n\} \) is IID).
Rate distortion function for images

- Signal model: Gaussian source with auto-correlation function

\[ R_{ss} [n_x, n_y] = \exp\left(-\omega_0 \sqrt{n_x^2 + n_y^2}\right) \]

- Power spectral density (neglecting aliasing):

\[ \Phi_{ss} (\omega_x, \omega_y) = \frac{2\pi}{\omega_0^2} \left(1 + \frac{\omega_x^2 + \omega_y^2}{\omega_0^2}\right)^{-\frac{3}{2}} \]

\[ \omega_0 = -\ln(0.93) \]

Correlation between adjacent pixels
Rate distortion function for images (cont.)

- Mean squared error criterion: \( D = E\{ (S - \hat{S})^2 \} \)
- After numerical integration:
Rate distortion function for Gauss-Markov Process

Gauss-Markov process with correlation factor $\rho$ for MSE distortion

- Distortion rate function for $R \geq \log_2(1 + \rho)$ is given by
  \[
  D(R) = (1 - \rho^2) \cdot \sigma^2 \cdot 2^{-2R}
  \]

- Includes result for Gaussian iid sources ($\rho = 0$)
Summary: rate distortion theory

- Rate-distortion theory
  - minimum transmission bit-rate for given distortion
- Shannon Lower Bound
  - assumes statistical independence between distortion and reconstructed signal
- $R(D)$ for memoryless Gaussian source and MSE: 6 dB/bit
- $R(D)$ for Gaussian source with memory and MSE
  - Theoretical gain ~2.3 bits/sample by exploiting spatial redundancy in the video signal
  - Parametric formulation of rate distortion function
  - the most difficult to code among all sources with the same power spectral density (for MSE distortion)
PRACTICAL RATE-DISTORTION RELATIONSHIP IN VIDEO CODING
Rate control in video coding

- Rate control in video coding
  - select quantization parameters or coding modes at slice and macroblock levels to produce consistent quality image for a given bit budget

- A well-designed rate control strategy can
  - improve overall image quality for video transmission over a constant-bitrate channel

- Which distortion level & how many bits used for encoding?
  - Rate and Distortion are controlled by quantization
    - R-Q, D-Q models

- An example of rate control problem for encoding a frame

  Our goal is to select the quantization scales for each $B$ frame $(q_0, q_1, \cdots, q_{MB-1})$, so as to

  \[
  \text{minimize} \quad \sum_{i=0}^{MB-1} d_{Bi}(q_i) \quad \text{subject to} \quad \sum_{i=0}^{MB-1} r_{Bi}(q_i) \leq R_B
  \]
R-Q model

- R-Q model

\[
R = \alpha + \beta \log \frac{1}{Q}
\]

\[
R = \alpha + \frac{\beta}{Q^\gamma} \quad (0 < \gamma \leq 2)
\]
R-Q model

Fig. 3. (a) I-frame 30 of “tennis”; (b) I-frame 150 of “football”; (c) P-frame 48 of “tennis”; (d) P-frame 138 of “football.”
R-Q model

Fig. 2. I-frame R-Q curves of (a) frame 0–60 of “tennis” sequence; (b) frame 120–165 of “football” sequence.
R-Q model

- R-Q model (quadratic R-D model)

\[ B_i = a_1 \times Q_i^{-1} + a_2 \times Q_i^{-2} \]
Rate-Distortion Relationship

- rate-quantization (R-Q)
  \[ R_n(q) = \alpha_n/q + \beta_n \]
- distortion-quantization (D-Q)
  \[ D_n(q) = \delta_n q + \gamma_n \]
- parameters estimated at the decoder and sent to the encoder periodically