Signal and Image Processing  (236327)

Tutorial 3

Quantization
Signal Digitization: The Quantization Role

• The signal **sampling** gives
  – A discrete signal of finite number of elements.
  – These elements are **real numbers**
    • hence, their **exact representation** would require an infinite amount of bits.

• **Quantization** is used for
  – representing **approximated-values** of the samples using a **finite number of bits**.
The Quantization Procedure

• The input is in the **continuous range** \([\phi_L, \phi_H]\).
• We would like to represent it using \(b\) bits that allow \(J = 2^b\) discrete representation levels \(\{r_i\}_{i=1}^{J}\).
• Correspondingly, the input range \([\phi_L, \phi_H]\) is partitioned by \(J + 1\) decision levels \(\{d_i\}_{i=0}^{J}\).
  
  – The mapping function:
    
    \[ Q(x) = r_i \quad for \quad x \in (d_{i-1}, d_i) \]

*Demonstration* \((b = 3)\):

![Diagram](image-url)
The Quantization Error for a Uniform Quantizer and a Uniform-Distributed Input

• Our input signal is \( \{ \varphi(t) | t \in [0,1] \} \), its values are assumed to be **uniformly distributed** in \( [\varphi_L, \varphi_H] \).

• Let us consider a **uniform quantizer of b bits**:
  - There are \( J = 2^b \) quantization-intervals of equal size \( \Delta_q = (\varphi_H - \varphi_L)/J \).
    - one of \( J, b \) or \( \Delta_q \) should be set, the other two are adjusted correspondingly.
  - Decision levels: \( d_i = \varphi_L + i\Delta_q \), \( i = 0,1,...,J \)
  - Representation levels: \( r_i = \varphi_L + (i + \frac{1}{2})\Delta_q \), \( i = 0,1,...,J - 1 \)

\[
\text{for } x \in [\varphi_L, \varphi_H]: \quad Q(x) = \varphi_L + \Delta_q \left( \left\lfloor \frac{x - \varphi_L}{\Delta_q} \right\rfloor + \frac{1}{2} \right)
\]
The Quantization Error
for a Uniform Quantizer and a Uniform-Distributed Input

\[
MSE_Q = \int_{\varphi_L}^{\varphi_H} \left( x - Q(x) \right)^2 p(x) \, dx = \sum_{i=1}^{J} \int_{d_{i-1}}^{d_i} \left( x - Q(x) \right)^2 p(x) \, dx = (*)
\]

setting
\[
d_i = \varphi_L + i \Delta \quad ; \quad d_{i-1} = \varphi_L + (i-1) \Delta \quad ; \quad Q(x) = d_{i-1} + 0.5 \Delta \; \text{for} \; x \in [d_{i-1}, d_i)
\]

and \( p(x) = \frac{1}{\varphi_H - \varphi_L} \) for \( x \in [\varphi_L, \varphi_H] \) gives

\[
(*) = \frac{1}{\varphi_H - \varphi_L} \sum_{i=1}^{J} \int_{\varphi_L + (i-1)\Delta}^{\varphi_L + i\Delta} \left( x - (\varphi_L + (i-1)\Delta + 0.5\Delta) \right)^2 \, dx = (**)
\]

integration-variable substitution: \( e_q \equiv x - (\varphi_L + (i-1)\Delta + 0.5\Delta) \) for each integral \( i = 1, \ldots, J \)

\[
(**) = \frac{1}{\varphi_H - \varphi_L} \sum_{i=1}^{J} \int_{-\Delta/2}^{\Delta/2} e_q^2 \, de_q = \frac{1}{\varphi_H - \varphi_L} J \int_{-\Delta/2}^{\Delta/2} e_q^2 \, de_q = \frac{1}{\varphi_H - \varphi_L} J \left( \frac{e_q^3}{3} \right)_{-\Delta/2}^{\Delta/2} = \frac{J}{\varphi_H - \varphi_L} \frac{\Delta}{12} = \frac{\Delta^2}{12}
\]
The Quantization Error
for a Uniform Quantizer and a Uniform-Distributed Input

Another point of view:
The error \( e_q \) is characterized as follows:

\( e_q \) is uniformly distributed in \([- \frac{\Delta q}{2}, \frac{\Delta q}{2}]\)

\[
p_{e_q}(x) = \begin{cases} 
  1/\Delta q & \text{for } -\frac{\Delta q}{2} \leq x \leq \frac{\Delta q}{2} \\
  0 & \text{otherwise}
\end{cases}
\]

Therefore,
the mean error: \( E\{e_q\} = 0 \) (due to symmetry around zero)

Let us calculate the MSE (the error variance, in this case):

\[
MSE_Q = \int_{-\Delta q/2}^{\Delta q/2} e_q^2 p(e_q) de_q = \frac{1}{\Delta q} \int_{-\Delta q/2}^{\Delta q/2} e_q^2 de_q = \frac{1}{\Delta q} \frac{1}{3} e_q^3 \bigg|_{-\Delta q/2}^{\Delta q/2} = \frac{\Delta q^2}{12}
\]
Optimal Scalar Quantization

Let us quantize the continuous scalar \( t \in [L, H] \):
* \( J \) discrete representation levels: \( \{ r_1, r_2, ..., r_J \} \)
* Each \( r_i \) has a decision region \( D_i \subset [L, H] \)
* The mapping: \( Q(t) = r_i \) for \( t \in D_i \)

The MSE is

\[
MSE_Q = \int_L^H \left( t - Q(t) \right)^2 p(t) \, dt
\]

where \( p(t) \) is the distribution density of \( t \) over \([L, H]\).
Optimal Scalar Quantization

If we are given the regions \( D_i \), the error would be

\[
MSE_Q = \sum_{j=1}^{J} \int_{D_j} \left( t - r_j \right)^2 p(t) dt
\]

and the optimal \( r_i \) values are chosen according to

\[
\min_{r_i} \int_{D_i} \left( t - r_i \right)^2 p(t) dt \quad i = 1, \ldots, J
\]

\[
\Rightarrow \frac{\partial}{\partial r_i} \int_{D_i} \left( t - r_i \right)^2 p(t) dt = \int_{D_i} \frac{\partial}{\partial r_i} \left[ \left( t - r_i \right)^2 \right] p(t) dt = -2 \int_{D_i} (t - r_i) p(t) dt = 0
\]

which yields

\[
r_i = \frac{\int_{D_i} t \cdot p(t) dt}{\int_{D_i} p(t) dt}
\]

Given the decision regions,

the best representation levels are the "centers of mass" of the regions.
Optimal Scalar Quantization

Let us fix the representation levels \( \{r_i\}_{i=1}^{J} \) to \( J \) distinct values within \([L, H]\).
Assume that \( L \leq r_1 < r_2 < \ldots < r_J \leq H \).

Given a value \( t \in [L, H] \), we may ask to which level \( r_i \) should we assign \( t \), to minimize the MSE.

The squared-error for a given \( t \) can take only \( J \) values \((t - r_i)^2 \) \( i = 1, 2, \ldots, J \) and we achieve minimal error by assigning \( t \) to the closest representation level:

\[
D_i = [d_{i-1}, d_i) \quad \text{where} \quad d_i = \frac{r_i + r_{i+1}}{2} \quad \text{and} \quad d_0 = L, \quad d_J = H
\]

Given the representation levels we can find the decision regions, and given the decision regions we have a formula for the best representation levels.
Optimal Scalar Quantization

Determining the optimal quantizer requires both optimal representation levels \( \{r_i\}_{i=1}^J \) and decision regions \( \{d_i\}_{i=0}^J \), such that \( L = d_0 \leq r_1 < d_1 < r_2 < d_2 < \ldots < r_J \leq d_J = H \).

The error is given by \( MSE_Q = \sum_{i=1}^J \int_{d_{i-1}}^{d_i} (t - r_i)^2 p(t) \, dt \).

A necessary condition for optimality is given by \( \frac{\partial}{\partial r_i} MSE_Q = 0 \), \( \frac{\partial}{\partial d_i} MSE_Q = 0 \).

Recall that \( \frac{\partial}{\partial x} \int_y f(t) \, dt = f(x) \) and \( \frac{\partial}{\partial x} \int_x f(t) \, dt = -f(x) \).

developing the condition equations yields:

\[
\begin{align*}
\frac{\partial}{\partial r_i} MSE_Q &= -2 \int_{d_{i-1}}^{d_i} (t - r_i) p(t) \, dt = 0, \quad i = 1, \ldots, J \\
\frac{\partial}{\partial d_i} MSE_Q &= (d_i - r_i)^2 p(d_i) - (d_i - r_{i+1})^2 p(d_i) = 0, \quad i = 1, \ldots, J - 1
\end{align*}
\]

\[
\Rightarrow \begin{cases}
    \int_{d_{i-1}}^{d_i} t \cdot p(t) \, dt \\
    r_i = \frac{d_{i+1} - d_i}{d_{i+1} - d_i}, \quad i = 1, \ldots, J \\
    \int_{d_{i-1}}^{d_i} p(t) \, dt \\
    d_i = \frac{r_i + r_{i+1}}{2}, \quad i = 1, \ldots, J - 1
\end{cases}
\]
Optimal Scalar Quantization
Max-Lloyd Algorithm

The conditions for optimality have inter-dependency.
Hence, Lloyd and Max proposed the following iterative solution:

1. Initialization: set guess representation levels \( \{r_i\} \)
2. Compute the optimal decision levels for \( \{r_i\} \) and set to \( \{d_i\} \)
3. Compute the optimal representation levels for \( \{d_i\} \) and set to \( \{r_i\} \)
4. If stopping criteria has not met, return to (2).

Remark: the roles of the representation and decision levels can be substituted.

Each step applies an optimization, hence the MSE will decrease or not changed.
Convergence to a minima (however, not necessarily the global one) is guaranteed.
Max-Lloyd Quantizer

Exercise #1
The samples of a given signal have the following probability-density-function:

\[
p(x) = \begin{cases} 
2x, & 0 \leq x \leq 1 \\
0, & \text{else} 
\end{cases}
\]

Find the Max-Lloyd quantizer for two representation levels (J=2).
Max-Lloyd Quantizer

Analytic Solution:

input range borders: $d_0 = 0$, $d_2 = 1$

optimal decision level:

(1) $d_1 = \frac{r_1 + r_2}{2}$

optimal representation levels:

(2) $r_1 = \frac{\int_{d_0}^{d_1} x \ p(x) \ dx}{\int_{d_0}^{d_1} p(x) \ dx} = \frac{\int_{d_0}^{d_1} x \cdot 2 \ dx}{\int_{d_0}^{d_1} 2 \ dx} = \frac{2}{3} d_1$

(3) $r_2 = \frac{\int_{d_1}^{d_2} x \ p(x) \ dx}{\int_{d_1}^{d_2} p(x) \ dx} = \frac{\int_{d_1}^{d_2} x \cdot 2 \ dx}{\int_{d_1}^{d_2} 2 \ dx} = \frac{2}{3} \frac{1 - d_1^3}{1 - d_1^2}$

substituting (2) and (3) into (1) yields: $d_1 = \frac{2/3 d_1 + 2/3 \frac{1 - d_1^3}{1 - d_1^2}}{2} \Rightarrow d_1^3 - 2d_1 + 1 = 0 \Rightarrow d_1 = 0.618, -1.618$

setting the optimal $d_1$ into (2) and (3) gives the optimal representation levels: $r_1 = 0.412$ and $r_2 = 0.824$
Max-Lloyd Quantizer

Iterative Solution:

(1) Initial guess: representation levels of a uniform quantizer in the range $[0,1]$: $r_1^{(0)} = 0.25$ \quad $r_2^{(0)} = 0.75$

Iteration 1: (2) $d_1^{(1)} = \frac{r_1^{(0)} + r_2^{(0)}}{2} = \frac{1}{2}$

(3) $r_1^{(1)} = \frac{2}{3} d_1^{(1)} = \frac{1}{3}$ \quad $r_2^{(1)} = \frac{2}{3} \frac{1 - (d_1^{(1)})^3}{1 - (d_1^{(1)})^2} = 0.778$

Iteration 2: (2) $d_1^{(2)} = \frac{r_1^{(1)} + r_2^{(1)}}{2} = 0.556$

(3) $r_1^{(2)} = \frac{2}{3} d_1^{(2)} = 0.37$ \quad $r_2^{(2)} = \frac{2}{3} \frac{1 - (d_1^{(2)})^3}{1 - (d_1^{(2)})^2} = 0.799$

... 

Iteration 15: $d_1^{(15)} = 0.618$ \quad $r_1^{(15)} = 0.412$ \quad $r_2^{(15)} = 0.824$
Mean-Absolute-Error (MAE) Quantizer

The error is given by

\[ MAE_Q = \int_L^H |t - Q(t)|p(t)\,dt = \sum_{i=1}^J \int_{d_{i-1}}^{d_i} |t - r_i|p(t)\,dt \]

The optimization problem:

\[ \min_{\{d_i\}_{i=1}^{J-1},\{r_i\}_{i=1}^J} MAE_Q = \min_{\{d_i\}_{i=1}^{J-1},\{r_i\}_{i=1}^J} \sum_{i=1}^J \int_{d_{i-1}}^{d_i} |t - r_i|p(t)\,dt \]

Recall that

\[ \frac{\partial}{\partial x} \int_y^x f(t)\,dt = f(x) \quad \text{and} \quad \frac{\partial}{\partial x} \int_x^y f(t)\,dt = -f(x) \]

Hence,

\[ \frac{\partial}{\partial d_i} MAE_Q = |d_i - r_i|p(d_i) - |d_i - r_{i+1}|p(d_i) = 0 \quad \Rightarrow \quad d_i = \frac{r_i + r_{i+1}}{2} \quad i = 1, \ldots, J - 1 \]

\[ \frac{\partial}{\partial r_j} MAE_Q = \sum_{i=1}^J \int_{d_{i-1}}^{d_i} \frac{\partial}{\partial r_j} |t - r_i|p(t)\,dt = \int_{d_{j-1}}^{d_j} \text{sign}(t - r_j)p(t)\,dt = 0 \]

\[ \Rightarrow \quad \int_{d_{j-1}}^{d_j} p(t)\,dt - \int_{d_{j-1}}^{r_j} p(t)\,dt = 0 \quad \Rightarrow \quad \int_{r_j}^{d_j} p(t)\,dt = \int_{d_{j-1}}^{r_j} p(t)\,dt \quad j = 1, \ldots, J \]