Lecture 6 Scalar and Vector Quantization

Wen-Hsiao Peng, Ph.D

Multimedia Architecture and Processing Laboratory (MAPL)
Department of Computer Science, National Chiao Tung University

May 2013
Quantization

- Lossy compression method
  - Reduce distinct output values to a smaller set
  - Map an input value/vector to an approximated value/vector

- Approaches
  - Scalar quant. – quantize each sample separately
    - Uniform vs. Non-uniform
    - MSE vs. MAE vs. ....
  - Vector quant. – quantize a group of samples jointly
    - Uniform (lattice quantizer) vs. Non-uniform
    - MSE vs. MAE vs. ....

- Objectives
  1. Minimize distortion s.t. number of reconstruction levels
  2. Minimize distortion s.t. entropy rate constraint (or vice versa)
Example

Scalar quantization

- Quantization levels $L = 2^R$
- Reconstruction values $g_l, l = 1, 2, ..., L$
- Boundary values $b_l, l = 0, 1, ..., L$
- Quantizer mapping function

$$Q(f) = g_l \text{ if } f \in B_l = [b_{l-1}, b_l)$$
Quantizer Mapping Function

Non-uniform Quantizer

\[ Q(f) = g_l \text{ if } f \in B_l \]

Uniform Quantizer

\[ Q(f) = \left\lfloor \frac{f - f_{\text{min}}}{q} \right\rfloor \times q + \frac{q}{2} + f_{\text{min}} \]
Uniform vs. Non-uniform
Midrise vs. Midtread Quantizers

- Midrise uniform quantizer (even $L$)
  - "0" is a decision boundary
- Midtread uniform quantizer (odd $L$)
  - "0" is a reconstruction level
  - Better for small $L$ in video/image coding

Midrise vs. Midtread (for symmetric distributions)
Quantizer Distortion

\[ D_q = \sigma^2_{q(\text{granular})} + \sigma^2_{q(\text{overload})} \]

\( \sigma^2_{q(\text{granular})} \) our focus

0 for bounded inputs
Quantizer Distortion

Quantizer distortion $D_q$

\[ D_q = E(d(F, Q(F))) \]

\[ = \sum_{l \in \mathcal{L}} \left( \int_{f \in B_l} d(f, Q(f))p(f)df \right) \]

\[ = \sum_{l \in \mathcal{L}} P(B_l) \int_{B_l} d(f, Q(f))p(f|F \in B_l)df \]

\[ = D_{q,l} \]

Example: $d(F, Q(F)) = (F - Q(F))^2$

\[ D_q = E(d(F, Q(F))) = E \left( (F - Q(F))^2 \right) = \sigma_q^2 \]

where quant. error $F - Q(F)$ is usually of zero mean (as we shall show later)

\[ ^1\text{Depend on distortion measure } d(\cdot) \text{ and source distribution } p(f) \]
Scalar Quantizer
Uniform Scalar Quantizer

- Equal distances between adjacent boundaries and reconstruction values

\[ b_l - b_{l-1} = g_l - g_{l-1} = q \]

where

\[ b_l = f_{\text{min}} + l \times q \]
\[ g_l = \frac{b_l + b_{l-1}}{2} \]

\[ Q(f) = \left\lfloor \frac{f - f_{\text{min}}}{q} \right\rfloor \times q + \frac{q}{2} + f_{\text{min}} \]

Closed-form
Uniform Quantizer Optimized for Uniform Distribution

- Uniform distribution
  
  \[ p(f) = \begin{cases} 
  1/B & f \in (f_{\min}, f_{\max}) \\
  0 & \text{otherwise} 
  \end{cases} \]
  
  where \( B = f_{\max} - f_{\min} \)

- Distortion measure
  
  \[ d(F, Q(F)) = (F - Q(F))^2 \]

- Quantizer distortion
  
  \[ D_q = \sum_{l \in L} P(B_l)D_{q,l} = D_{q,l} = \frac{q^2}{12} = \sigma_f^2 2^{-2R} \]

  \[ SNR = 10 \log \frac{\sigma_f^2}{D_q} = (20 \log 2) R = 6.02 R \]

  where

  \[ D_{q,l} = \frac{q^2}{12}, q = \frac{B}{L}, L = 2^R, \text{signal variance } \sigma_f^2 = \frac{B^2}{12} \]
Minimum Mean Square Error Scalar Quantizer

- Distortion measure
  \[ d(\mathcal{F}, Q(\mathcal{F})) \equiv (\mathcal{F} - Q(\mathcal{F}))^2 \]

- Objective
  \[
  (b^*, g^*) = \operatorname{arg\,min}_{\{b, g\}} \left( \sum_{l \in L} \int_{b_{l-1}}^{b_l} (f - g_l)^2 p(f) df \right)
  \]
  \[ J(b, g) \]
  where
  \[
  b = (b_0, b_1, ..., b_L)^T, \quad g = (g_1, g_2, ..., g_L)^T
  \]

- Necessary conditions
  \[
  \nabla J(b^*, g^*) = 0 \Rightarrow \begin{align*}
  (1) \quad & \frac{\partial J(b^*, g^*)}{\partial b_l} = 0 \\
  (2) \quad & \frac{\partial J(b^*, g^*)}{\partial g_l} = 0
  \end{align*}
  \]
MMSE Scalar Quantizer

- **Solution**

  \[
  (1) \quad \frac{\partial J(b^*, g^*)}{\partial b_i} = (b_i^* - g_i^*)^2 p(b_i^*) - (b_i^* - g_{i+1}^*)^2 p(b_i^*) = 0
  \]

  \[
  (2) \quad \frac{\partial J(b^*, g^*)}{\partial g_i} = - \int_{b_{i-1}^*}^{b_i^*} 2(f - g_i^*) p(f) df = 0
  \]

- **Nearest-neighbor condition**

  \[
  (b_i^* - g_i^*)^2 = (b_i^* - g_{i+1}^*)^2 \quad (1)
  \]

  \[
  \Rightarrow \quad b_i^* = \frac{g_i^* + g_{i+1}^*}{2}
  \]

  \[
  \Rightarrow \quad B_i = \{ f : d(f, g_i) \leq d(f, g_{i'}), \forall i' \neq i \}
  \]

- **Centroid condition**

  \[
  g_i^* = \frac{1}{P(B_i)} \int_{b_{i-1}^*}^{b_i^*} f p(f) df = E(F | F \in B_i) \quad (2)
  \]

  \[
  \text{Conditional Mean}
  \]

- **Eqs. (1) & (2) generally DO NOT have a closed-form solution**
Properties of MMSE Scalar Quantizer

- **Symbols**
  - $\mathcal{F}$: quantizer input
  - $\mathcal{G}$: quantizer output
  - $Q = \mathcal{F} - \mathcal{G}$: quantization error

- $\mathcal{G}$ is an unbiased estimator of $\mathcal{F}$
  \[ E(\mathcal{G}) = E(\mathcal{F}), E(Q) = 0 \]

- $\mathcal{G}$ is orthogonal to (and uncorrelated with) $Q$, whereas $\mathcal{F}$ and $Q$ are correlated
  \[ E(\mathcal{G}Q) = 0, E(\mathcal{F}Q) \neq 0 \]

- Reduced signal variance
  \[ \sigma^2_{\mathcal{G}} = \sigma^2_{\mathcal{F}} - \sigma^2_Q \]

- Equalized error contribution
  \[ P(\mathcal{B}_l) D_{q,l} = \frac{D_q}{L}, \quad \forall l \in \mathcal{L} \]
MMSE Scalar Quantizers of Various Sources

- All sources are with zero mean and unit variance.²

### Quantizer Design \((y_j = g_j, x_j = b_j)\)

<table>
<thead>
<tr>
<th>R</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>pdf</td>
<td>(j)</td>
<td>(y_{j, \text{opt}})</td>
<td>(y_{j, \text{opt}})</td>
<td>(y_{j, \text{opt}})</td>
</tr>
<tr>
<td>U</td>
<td>1</td>
<td>0.000</td>
<td>0.866</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.000</td>
<td>0.866</td>
<td>0.129</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.000</td>
<td>0.866</td>
<td>1.083</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.000</td>
<td>0.866</td>
<td>1.516</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.000</td>
<td>0.866</td>
<td>1.939</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.000</td>
<td>0.866</td>
<td>2.361</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>0.000</td>
<td>0.866</td>
<td>2.784</td>
</tr>
<tr>
<td>L</td>
<td>1</td>
<td>0.000</td>
<td>0.798</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.000</td>
<td>0.798</td>
<td>0.501</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.000</td>
<td>0.798</td>
<td>1.050</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.000</td>
<td>0.798</td>
<td>1.748</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.000</td>
<td>0.798</td>
<td>2.446</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.000</td>
<td>0.798</td>
<td>3.144</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>0.000</td>
<td>0.798</td>
<td>3.841</td>
</tr>
<tr>
<td>(\Gamma)</td>
<td>1</td>
<td>0.000</td>
<td>0.707</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.000</td>
<td>0.707</td>
<td>0.533</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.000</td>
<td>0.707</td>
<td>1.253</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.000</td>
<td>0.707</td>
<td>2.073</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.000</td>
<td>0.707</td>
<td>2.893</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.000</td>
<td>0.707</td>
<td>3.725</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>0.000</td>
<td>0.707</td>
<td>4.557</td>
</tr>
<tr>
<td>(\Gamma)</td>
<td>1</td>
<td>0.000</td>
<td>0.577</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.000</td>
<td>0.577</td>
<td>0.327</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.000</td>
<td>0.577</td>
<td>1.047</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.000</td>
<td>0.577</td>
<td>2.073</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.000</td>
<td>0.577</td>
<td>3.089</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.000</td>
<td>0.577</td>
<td>4.105</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>0.000</td>
<td>0.577</td>
<td>5.128</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>R</th>
<th>pdf</th>
<th>SNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>U</td>
<td>6.02</td>
</tr>
<tr>
<td>2</td>
<td>U</td>
<td>12.04</td>
</tr>
<tr>
<td>3</td>
<td>U</td>
<td>18.06</td>
</tr>
<tr>
<td>4</td>
<td>U</td>
<td>24.08</td>
</tr>
<tr>
<td>5</td>
<td>U</td>
<td>30.10</td>
</tr>
<tr>
<td>6</td>
<td>U</td>
<td>36.12</td>
</tr>
<tr>
<td>7</td>
<td>U</td>
<td>42.14</td>
</tr>
<tr>
<td>1</td>
<td>G</td>
<td>4.40</td>
</tr>
<tr>
<td>2</td>
<td>G</td>
<td>9.30</td>
</tr>
<tr>
<td>3</td>
<td>G</td>
<td>14.62</td>
</tr>
<tr>
<td>4</td>
<td>G</td>
<td>20.22</td>
</tr>
<tr>
<td>5</td>
<td>G</td>
<td>26.01</td>
</tr>
<tr>
<td>6</td>
<td>G</td>
<td>31.89</td>
</tr>
<tr>
<td>7</td>
<td>G</td>
<td>37.81</td>
</tr>
<tr>
<td>1</td>
<td>L</td>
<td>3.01</td>
</tr>
<tr>
<td>2</td>
<td>L</td>
<td>7.54</td>
</tr>
<tr>
<td>3</td>
<td>L</td>
<td>12.64</td>
</tr>
<tr>
<td>4</td>
<td>L</td>
<td>18.13</td>
</tr>
<tr>
<td>5</td>
<td>L</td>
<td>23.87</td>
</tr>
<tr>
<td>6</td>
<td>L</td>
<td>29.74</td>
</tr>
<tr>
<td>7</td>
<td>L</td>
<td>35.69</td>
</tr>
<tr>
<td>1</td>
<td>(\Gamma)</td>
<td>1.76</td>
</tr>
<tr>
<td>2</td>
<td>(\Gamma)</td>
<td>6.35</td>
</tr>
<tr>
<td>3</td>
<td>(\Gamma)</td>
<td>11.52</td>
</tr>
<tr>
<td>4</td>
<td>(\Gamma)</td>
<td>17.07</td>
</tr>
<tr>
<td>5</td>
<td>(\Gamma)</td>
<td>22.85</td>
</tr>
<tr>
<td>6</td>
<td>(\Gamma)</td>
<td>28.73</td>
</tr>
<tr>
<td>7</td>
<td>(\Gamma)</td>
<td>34.67</td>
</tr>
</tbody>
</table>

²See Appendix for arbitrary PDF
MMSE Scalar Quantizer: High Rate Approximation

- **High rate approx.**: $p(f)$ is approximately flat in all intervals $B_l$

\[
D_q = \sum_{l \in \mathcal{L}} \int_{b_{l-1}}^{b_l} (f - g_l)^2 p(f) \, df
\]

\[
\approx \sum_{l \in \mathcal{L}} p(g_l) \int_{b_{l-1}}^{b_l} (f - g_l)^2 \, df
\]

\[
= \sum_{l \in \mathcal{L}} \frac{P(B_l)}{\Delta_l} \int_{b_{l-1}}^{b_l} (f - g_l)^2 \, df, \text{ where } \Delta_l = b_l - b_{l-1}
\]

- Centroid condition

\[
\frac{\partial D_q(b^*, g^*)}{\partial g_l} = 0 \Rightarrow \int_{b_{l-1}^*}^{b_l^*} (f - g_l^*) \, df = 0 \Rightarrow g_l^* = \frac{b_l^* + b_{l-1}^*}{2}
\]
MMSE Scalar Quantizer: High Rate Approximation

- Quantizer distortion

\[ D_q \approx \sum_{l \in \mathcal{L}} \frac{P(B_l)}{\Delta_l} \int_{b_{l-1}^*}^{b_l^*} (f - g_l^*)^2 df = \sum_{l \in \mathcal{L}} P(B_l) \frac{\Delta_l^2}{12} \]

where \( \Delta_l = b_l - b_{l-1} \)

- Equalized error contribution

\[ P(B_l) D_{q,l} = \frac{1}{12} P(B_l) \Delta_l^2 = \text{constant} \]

- Special case: uniform quantizer and uniform source

\[ P(B_l) = P(B_{l'}) , \Delta_l = \Delta_{l'} = \Delta, D_q = \frac{\Delta^2}{12} \]
MMSE Scalar Quantizer: High Rate Approximation

- Alternative expression (more useful)

\[
D_q \sim \sum_{l \in \mathcal{L}} P(B_l) \frac{\Delta_l^2}{12} \sim \sum_{l \in \mathcal{L}} \frac{p_{\mathcal{F}}(g_l^*) \Delta_l^3}{12} \sim \sigma_f^2 2^{-2R} \epsilon^2 \quad \text{(See Appendix)}
\]

where

\[
\epsilon^2 = \frac{1}{12} \left( \int_{f_{\min} / \sigma_f}^{f_{\max} / \sigma_f} \sqrt[3]{p_{\mathcal{F}}'}(f) df \right)^3
\]

\[p_{\mathcal{F}}'(f) = \sigma_f p_{\mathcal{F}}(\sigma_f f)\] is PDF of an unit variance source

- Uniform quantizer/source vs. non-uniform quantizer/source

Uniform : \[SNR \sim 10 \log \frac{\sigma_f^2}{D_q} = 10 \log 2^{2R} = 6.02R\]

Non-uniform : \[SNR \sim 10 \log \frac{\sigma_f^2}{D_q} = 6.02R - 20 \log \epsilon\]
MMSE Scalar Quantizer of Various Sources

- \( \Delta SNR = 6.02R \) (Uniform Source and Quantizer) — \( \max\{SNR\} \)
- Solid (PDF-opt. non-uniform quantizer) vs. Dashed (PDF-opt. uniform quantizer)

Uniform quant. becomes increasingly inefficient with increasing \( R \)
Non-uniform quant. attains a PDF-specific asymptote with increasing \( R \)
Companding: Compressing and Expanding

- A technique for analyzing quantizers at high rate
- Companding law $c(x)$ defines $B_l$
- Interval intensity

$$\frac{dc(g_l)}{dx} \approx \frac{2x_{\text{max}}}{L \Delta_l}$$

where

$c(x_{\text{max}}) = x_{\text{max}},$
$c(0) = 0,$
$c(x_{\text{min}}) = x_{\text{min}}$

$^a$Quantizers designed by $c(x)$ is no longer common
Companding: Compressing and Expanding

- Distortion measure
  \[ d(F, Q(F)) \equiv (F - Q(F))^2 \]

- Quantizer distortion (in terms of \( c(x) \))
  \[
  D_q \approx \sum_{l \in \mathcal{L}} P(B_l) \frac{\Delta_l^2}{12}
  \]
  \[
  \approx \frac{x_{\text{max}}^2}{3L^2} \sum_{l \in \mathcal{L}} p(g_l) \Delta_l \left( \frac{dc(g_l)}{dx} \right)^{-2}
  \]
  \[
  \approx \frac{x_{\text{max}}^2}{3L^2} \int_{x_{\text{min}}}^{x_{\text{max}}} p(x) \left( \frac{dc(x)}{dx} \right)^{-2} \, dx
  \]

- Constraint
  \[
  \int_{x_{\text{min}}}^{x_{\text{max}}} \frac{dc(x)}{dx} \, dx = c(x_{\text{max}}) - c(x_{\text{min}}) = 2x_{\text{max}}
  \]
MMSE Scalar Quantizer Design Using Companding

- **MMSE scalar quantizer**
  \[
  \min_{D_q} \frac{x_{\text{max}}^2}{3L^2} \int_{x_{\text{min}}}^{x_{\text{max}}} p(x) \left( \frac{dc(x)}{dx} \right)^{-2} \, dx \quad \text{s.t.} \quad \int_{x_{\text{min}}}^{x_{\text{max}}} \frac{dc(x)}{dx} \, dx = 2x_{\text{max}}
  \]

- **Lagrange multiplier** (find \( dc^*(x) / dx = g^* \))
  \[
  g^* = \arg \min \left( \int_{x_{\text{min}}}^{x_{\text{max}}} p(x) (g)^{-2} \, dx + \lambda \left( \int_{x_{\text{min}}}^{x_{\text{max}}} g \, dx - 2x_{\text{max}} \right) \right)
  \]

- **Optimal companding law** \( c^*(x) \)
  \[
  g^* = \frac{dc^*(x)}{dx} = \frac{x_{\text{max}}}{\int_0^{x_{\text{max}}} 3\sqrt{p(x)} \, dx} \sqrt[3]{p(x)}
  \]
  \[
  D_q^* \simeq \frac{2}{3L^2} \left( \int_0^{x_{\text{max}}} \sqrt[3]{p(x)} \, dx \right)^3 = \sigma_f^2 2^{-2R} \epsilon^2
  \]
Minimum Mean Absolute Error Quantizer

- Distortion measure
  \[ d(\mathcal{F}, Q(\mathcal{F})) \equiv |\mathcal{F} - Q(\mathcal{F})| \]

- Objective
  \[
  (b^*, g^*) = \arg\min_{\{b, g\}} \left( \sum_{l \in \mathcal{L}} \int_{b_{l-1}}^{b_l} |f - g_l| p(f) df \right)
  \]

  where
  \[ b = (b_0, b_1, ..., b_L)^T, g = (g_1, g_2, ..., g_L)^T \]

- Necessary conditions
  \[ \nabla J(b^*, g^*) = 0 \Rightarrow \]
  \[
  \begin{align*}
  (1) & \quad \frac{\partial J(b^*, g^*)}{\partial b_l} = 0 \\
  (2) & \quad \frac{\partial J(b^*, g^*)}{\partial g_l} = 0
  \end{align*}
  \]
Minimum Mean Absolute Error Quantizer

- **Solution**
  
  \[
  (1) \quad \frac{\partial J(b^*, g^*)}{\partial b_l} = |b_l^* - g_l^*| p(b_l^*) - |b_l^* - g_{l+1}^*| p(b_l^*) = 0
  \]
  
  \[
  (2) \quad \frac{\partial J(b^*, g^*)}{\partial g_l} = \int_{b_{l-1}^*}^{g_l^*} p(f) df - \int_{g_l^*}^{b_l^*} p(f) df = 0 \quad (\text{Appendix})
  \]

- **Nearest-neighbor condition (same as MMSE Quantizer)**

  \[
  |b_l^* - g_l^*| p(b_l^*) - |b_l^* - g_{l+1}^*| p(b_l^*) = 0 \Rightarrow b_l^* = \frac{g_l^* + g_{l+1}^*}{2}
  \]

- **Generalized centroid condition**

  \[
  \int_{b_{l-1}^*}^{g_l^*} p(f) df = \int_{g_l^*}^{b_l^*} p(f) df
  \]
Optimal Scalar Quantizer

- Nearest-neighbor condition (same as MMSE case and obvious)
  \[ B_i^* = \{ f : d(f, g_l) \leq d(f, g_{l'}), \forall l' \neq l \} \]

- Generalized centroid condition\(^3\)
  \[ g_l^* = \arg \min_{g_l} E(d(F, g) | F \in B_l) \]

Remarks

- Nearest-neighbor (distortion measure dependent) is NOT affected
- Centroid SHALL be adapted to distortion measure
- Source distribution MUST be known
- Reconstruction level is indexed by a fixed-length code

\(^3\)Solution depends on distortion measure
Lloyd-Max Algorithm

- Optimal quantizer design based on **training data**
  - Applicable when **source distribution is unknown**
  - Use **sample averages** to replace **expectations**

- Update reconstruction and boundary values iteratively
  1. **Initialization**
     - Choose initial reconstruction values
     - Calculate initial boundary values (nearest neighbor)
     - Calculate initial distortion
  2. **Iterations**
     - Find new reconstructions (centroid, sample mean)
     - Find new boundaries (nearest neighbor)
     - Calculate new distortion
     - Repeat till distortion converges
Lloyd-Max Algorithm

Given samples \( f_k, k = 1, 2, \ldots, K \), in a training set and the quantization level \( L \), set \( b_0 = f_{\min}, b_L = f_{\max} \).

1. Choose initial reconstruction values:
   \[ g_l, l \in L \]

2. Find initial partition regions based on the nearest-neighbor criterion:
   \[ B_i = \{ f_k : d_i(f_k, g_l) \leq \frac{1}{L} \sum_{l \in L} d_i(f_k, g_l) \}, \quad b_i = (g_l + g_{l+1})/2 \]
   MSE: \( B_i \)

3. Calculate initial distortion:
   \[ D_0 = \frac{1}{K} \sum_{k=1}^{K} \sum_{l \in L} d_i(f_k, g_l) \]

4. Find the new reconstruction values based on the centroid condition:
   \[ g_l = \arg \min_{g_l} \left\{ \frac{1}{K} \sum_{k=1}^{K} d_i(f_k, g_l) \right\}, \quad l \in L \]
   MSE: \( B_i \)

5. Find new partition regions based on the nearest-neighbor condition:
   \[ B_i = \{ f_k : d_i(f_k, g_l) \leq \frac{1}{L} \sum_{l \in L} d_i(f_k, g_l) \}, \quad b_i = (g_l + g_{l+1})/2 \]
   MSE: \( B_i \)

6. Calculate new distortion:
   \[ D_1 = \frac{1}{K} \sum_{k=1}^{K} \sum_{l \in L} d_i(f_k, g_l) \]

7. Update previous distortion:
   \[ D_0 = D_1 \]

8. Check convergence:
   \[ |D_1 - D_0| < T \]

9. Repeat steps 2-8 until convergence.

END
Entropy Constrained Optimal Scalar Quantizer

- Minimize quantizer output entropy \( s.t. \) quantizer distortion in MSE

\[
\min H_Q = - \sum_{l \in \mathcal{L}} P(B_l) \log_2 P(B_l)
\]

\[
\text{Quantizer output entropy}
\]

\[
s.t. \ D_q = \frac{x_{\text{max}}^2}{3L^2} \int_{x_{\text{min}}}^{x_{\text{max}}} p(x) \left( \frac{dc(x)}{dx} \right)^{-2} \, dx = D
\]

\[
\text{Quantizer distortion with } L \text{ Levels (high rate approx.)}
\]

- Parameters to be determined: \( L \) and \( \frac{dc(x)}{dx} \)
- Assumptions – (1) high rate, (2) the value of \( L \) is arbitrary
Entropy Constrained Optimal Scalar Quantizer

- Quantizer output entropy (in terms of $c(x)$)

$$H_Q = - \sum_{l \in \mathcal{L}} P(B_l) \log_2 P(B_l)$$

$$\leq - \sum_{l \in \mathcal{L}} p(g_l) \Delta_l \log_2 (p(g_l) \Delta_l)$$

$$\leq - \int_{x_{\text{min}}}^{x_{\text{max}}} p(x) \log_2 p(x) dx$$

$$+ \log_2 \frac{L}{2x_{\text{max}}} + \int_{x_{\text{min}}}^{x_{\text{max}}} p(x) \log_2 \frac{dc(x)}{dx} dx$$

$h(\mathcal{X}) = E(- \log_2 p(x))$ differential entropy

- Optimal quantizer is equivalent to finding $g^* = dc^*(x)/dx$ such that

$$\min \int_{x_{\text{min}}}^{x_{\text{max}}} p(x) \log_2 \frac{dc(x)}{dx} dx + \lambda \left( \int_{x_{\text{min}}}^{x_{\text{max}}} p(x) \left( \frac{dc(x)}{dx} \right)^{-2} dx - D' \right)$$

- Here we have assumed that $L$ takes on its best value
Uniform quantizer attains minimum output entropy regardless of PDF.

\[
\frac{dc^*(x)}{dx} = \sqrt{2\lambda \ln 2} = \text{constant}
\]

\[c^*(x) = x \text{ given } c(x_{\text{max}}) = x_{\text{max}}, c(0) = 0\]  \hspace{1cm} (4)

Optimal output entropy (from Eqs. (3) & (4))

\[H_Q^*(x) = h(X) - \log \Delta = h(X) - \frac{1}{2} \log_2 (12D_q)\]

Information-theoretical interpretation:

- \(h(X)\): average information amount at quantizer input
- \(\log_2 \Delta = h(Q)\): average information loss due to quantization (or average information conveyed by quantization error)
Comparison with Rate-Distortion Bound

- Memoryless Non-Gaussian Sources \((D(R)\) has NO closed-form\)

\[ \underbrace{L D(R)}_{\text{Lower Bound}} \leq D(R) \leq \underbrace{D(R)_G}_{\text{Upper Bound}} \]

where distortion \(D\) is in MSE and

\[
L D(R) = (2\pi e)^{-1} 2^{-2[R - h(X)]},
\]

\[
L R(D) = h(X) - \frac{1}{2} \log_2 2\pi eD
\]

- For a given distortion

\[ H_Q^*(x) - L R(D) = 0.255 \text{bits} \]

- Lower efficiency is caused by scalar quantization \(\Rightarrow\) vector quantization can be beneficial even for memoryless sources
Comparison with MMSE Scalar Quantizer

- For a given distortion, maximum bit rate reduction

\[
\max\{\Delta R\} = \frac{1}{2} \log_2 \frac{\sigma_f^2}{D_q} \epsilon^2 - H_Q^*(x) = \frac{1}{2} \log_2 \left(12\sigma_f^2 \epsilon^2\right) - h(\mathcal{X})
\]

- For a given rate, maximum SNR gain

\[
\max\{\Delta SNR\} = SNR_{EC} - SNR_{MMSE} = 6.02 \max\{\Delta R\} \text{ dB}
\]
Comparison with MMSE Scalar Quantizer

- $\max\{\Delta R\}$
  - Rate reduction with entropy constrained quantizer
  - A higher number of quantization levels

- $\Delta R_{pdf-opt}$
  - Rate reduction by entropy encoding MMSE quantizer outputs
  - A fixed number of quantization levels

<table>
<thead>
<tr>
<th>pdf</th>
<th>$h(X)$</th>
<th>$c^2$</th>
<th>$\Delta R_{pdf-opt}$ (bits/sample)</th>
<th>$\max{\Delta R}$ (bits/sample)</th>
<th>$\Delta SNR_{EC}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>$\frac{1}{2} \log_2 12 \sigma_x^2$</td>
<td>1.00</td>
<td>0.000</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>G</td>
<td>$\frac{1}{2} \log_2 2\pi e \sigma_x^2$</td>
<td>2.71</td>
<td>0.312</td>
<td>0.467</td>
<td>2.81</td>
</tr>
<tr>
<td>L</td>
<td>$\frac{1}{2} \log_2 2e^2 \sigma_x^2$</td>
<td>4.50</td>
<td>0.623</td>
<td>0.934</td>
<td>5.62</td>
</tr>
<tr>
<td>F</td>
<td>$\frac{1}{2} \log_2 \frac{4\pi}{3eC-1} \sigma_x^2$</td>
<td>5.68</td>
<td>1.138</td>
<td>1.707</td>
<td>10.28</td>
</tr>
</tbody>
</table>
Comparison with MMSE Scalar Quantizer

Gaussian Source
(MMSE Scalar Quantizer L=8)

- For a given $L$, $H_Q$ (entropy of quantizer output symbols) can be reduced at the cost of $\sigma_q^2$. How?
- For a given $\sigma_q^2$, the entropy can be reduced by increasing $L$; likewise, for a given output entropy $H_Q$, $\sigma_q^2$ can be reduced by increasing $L$. 
Summary

- $\Delta R_{pdf-opt}$
  - NOT the best value achievable by entropy coding
- $\max\{\Delta R\} = \left(\frac{3}{2}\right) \Delta R_{pdf-opt}$
  - Achieved with uniform quantizer and more quantization levels
  - Additional quantization levels are used for outer part of PDF
  - Fall short of R-D bound by 0.255bits (1.53dB) at high rates
  - Approach R-D bound at low rates
Vector Quantizer
Vector Quantization

- Why Vector Quantization?
  - Samples in a block are correlated
  - Some block patterns are more likely to occur than others

- Parameters
  - Quantization levels $L$
  - Reconstruction vectors (or Codewords) $g_l, l = 1, 2, \ldots, L$
  - Partition regions $B_l, l = 1, 2, \ldots, L$
  - Quantizer mapping function

$$g_l = Q(f) \text{ if } f \in B_l$$

where

$$f = [f_1, f_2, \ldots, f_N]^T, g_l = [g_l;1, g_l;2, \ldots, g_l;N]^T$$

- Bits/sample $R = \frac{1}{N} \lceil \log_2 L \rceil$
Nearest-Neighbor Vector Quantizer

- Quantized vector $\mathbf{g}_l$ determined by Nearest-Neighbor criterion
- Complexity increases exponentially with vector size $N$
  - Operations $NL$ (or $N2^{NR}$), storage $NL$ (or $N2^{NR}$)

\[ B_l = \{ \mathbf{f} \in \mathcal{R}^N : d_N(\mathbf{f}, \mathbf{g}_l) \leq d_N(\mathbf{f}, \mathbf{g}_{l'}) , \forall l' \neq l \} \]

\[ d_N(\mathbf{f}, \mathbf{g}_l) = \frac{1}{N} \sum_{n=1}^{N} (f_n - g_{l;n})^2 \]
Lattice Vector Quantizer

- **Realization of Uniform Vector Quantizer**
  - All partitions have same shape and size (good for uniform source)
  - Closed-form quantizer mapping function (low complexity)

- **Reconstruction vectors** $g_l$ are Lattice Points
  - Lattice $\Lambda$ (defined by a set of basis vectors $\{v_n\}$)

$$g_l = \sum_{n=1}^{N} m_l;n v_n = [V] m_l,$$

where

$$m_l \in \mathbb{Z}^N$$

$$[V] = [v_1, v_2, ..., v_N]$$

is Lattice Generating Matrix

- **Determination of quantized vector**
  1. $m = [V]^{-1} f$, where $m \in \mathbb{R}^N$ is a real index vector
  2. Evaluate distortions of $\hat{m} = [m]$ or $\lfloor m \rfloor$, where $\hat{m} \in \mathbb{Z}^N$
Quantizer Distortion

- Vector quantizer distortion

\[ D_q = E(d_N(\bar{\mathbf{F}}, Q(\bar{\mathbf{F}}))) \]
\[ = \int p_N(\mathbf{f})d_N(\mathbf{f}, Q(\mathbf{f}))d\mathbf{f} \]
\[ = \sum_{l=1}^{L} P(B_l) D_{q,l} \]

where

\[ D_{q,l} = \int_{B_l} p_N(\mathbf{f} | \bar{\mathbf{F}} \in B_l) d_N(\mathbf{f}, \mathbf{g}_l) d\mathbf{f} \]

- Example: lattice quantizer for uniform distribution

\[ D_q = \sum_{l=1}^{L} P(B_l) D_{q,l} = D_{q,l} \text{ and } D_{q,l} = \frac{1}{|\det[V]|} \int_{B_l} d_N(\mathbf{f}, \mathbf{0})d\mathbf{f} \]
Better space packing makes VQ outperform SQ even with i.i.d source.

- $d_N(f, g_I) = \max_{f \in B_I} d_N(f, g_I)$

- Rectangular lattice: 2 uniform scalar quantizers
Optimal Vector Quantizer

Objective

\[
(B_1^*, B_2^*, \ldots, B_L^*; g_1^*, g_2^*, \ldots, g_L^*) = \arg \min \ E(d(\hat{\mathbf{F}}, Q(\hat{\mathbf{F}))))
\]

\[
= \arg \min \left( \int d_N(f, Q(f)) p(f) df \right)
\]

\[
= \arg \min \left( \sum P(B_l) \int_{B_l} d_N(f, Q(f)) p(f|\hat{\mathbf{F}} \in B_l) df \right)
\]

Optimal Solution

1. Given \((g_1^*, g_2^*, \ldots, g_L^*)\), what would be optimal \((B_1, B_2, \ldots, B_L)\)?
2. Given \((B_1^*, B_2^*, \ldots, B_L^*)\), what would be optimal \((g_1, g_2, \ldots, g_L)\)?
Optimal Vector Quantizer

- Given \((g_1^*, g_2^*, ..., g_L^*)\), what would be optimal \((B_1, B_2, ..., B_L)\)?

\[
Q^*(f) = \arg\min_{Q(f) = \{g_1^*, g_2^*, ..., g_L^*\}} \left( \int d_N(f, Q(f)) p(f) \, df \right)
\]

\[
= \arg\min_{Q(f) = \{g_1^*, g_2^*, ..., g_L^*\}} d_N(f, Q(f))
\]

\[
B_i^* = \{ f : d_N(f, g_i^*) \leq d_N(f, g_{i'}^*), \forall i' \neq i \} \text{ (Nearest Neighbor)}
\]

- Given \((B_1^*, B_2^*, ..., B_L^*)\), what would be optimal \((g_1, g_2, ..., g_L)\)?

\[
g_i^* = \arg\min_g \int_{B_i^*} d_N(f, g) p(f) \, df \quad \overrightarrow{F} \in B_i^*
\]

\[
= \arg\min_g E \left( d_N(\overrightarrow{F}, g) | \overrightarrow{F} \in B_i^* \right) \text{ (Centroid)}
\]
Nearest-Neighbor and Centroid are **Necessary Conditions** (Candidates)
Lloyd-Max Algorithm

Given sample vectors \( \mathbf{f}_k, k = 1, 2, \ldots, K, \) in a training set, and quantization level \( L. \)

Choose initial codewords:

\[ g_l, l \in \mathcal{L} \]

Find initial partition regions based on the nearest-neighbor condition:

\[ B_l = \{ \mathbf{f}_k : d_N(\mathbf{f}_k, g_l) \leq d_N(\mathbf{f}_k, g_j), \forall j \neq l \}, l \in \mathcal{L} \]

Calculate initial distortion:

\[ D_0 = \frac{1}{K} \sum_{l \in \mathcal{L}} \sum_{\mathbf{f}_k \in B_l} d_N(\mathbf{f}_k, g_l) \]

Calculate new codewords based on the centroid condition:

\[ g_l = \arg \min_{g_l} \left\{ \frac{1}{K} \sum_{l \in \mathcal{L}} \sum_{\mathbf{f}_k \in B_l} d_N(\mathbf{f}_k, g_l) \right\} \]

MSE:

\[ g_l = \frac{1}{K} \sum_{l \in \mathcal{L}} \sum_{\mathbf{f}_k \in B_l} \mathbf{f}_k \]

Find new partition regions based on the nearest-neighbor condition:

\[ B_l = \{ \mathbf{f}_k : d_N(\mathbf{f}_k, g_l) \leq d_N(\mathbf{f}_k, g_j), \forall j \neq l \}, l \in \mathcal{L} \]

Calculate new distortion:

\[ D_1 = \frac{1}{K} \sum_{l \in \mathcal{L}} \sum_{\mathbf{f}_k \in B_l} d_N(\mathbf{f}_k, g_l) \]

Update previous distortion:

\[ D_0 = D_1 \]

If \( D_1 - D_0 < T \) then:

END

No

Yes
Lloyd-Max Algorithm
Entropy Constrained Vector Quantizer

Objective

\[
\min D_q \text{ s.t. } \sum_{l \in \mathcal{L}} P(\mathcal{B}_l) \log_2 P(\mathcal{B}_l) = RN
\]

Solution (Necessary Conditions)

Lagrangian

\[
(\mathcal{B}_1^*, \mathcal{B}_2^*, \ldots, \mathcal{B}_L^*, \mathbf{g}_1^*, \mathbf{g}_2^*, \ldots, \mathbf{g}_L^*)
\]

\[
= \arg \min \left( D_q + \lambda^* \left( \sum_{l \in \mathcal{L}} P(\mathcal{B}_l) \log_2 P(\mathcal{B}_l) - RN \right) \right)
\]

\[
J(\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_L; \mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_L) \quad ^4
\]

\[
\lambda^* \text{ must be chosen such that }
\]

\[
\sum_{l \in \mathcal{L}} P(\mathcal{B}_l^*) \log_2 P(\mathcal{B}_l^*) = RN
\]

\[^4\text{Lloyd-Max Algorithm is still applicable by replacing } D_q \text{ with } J(\cdot) \]
Gaussian-Markov Source + 8-D VQ

http://www.stanford.edu/class/ee368b/Handouts/06-Quantization.pdf
Memoryless Laplacian + 8-D VQ

http://www.stanford.edu/class/ee368b/Handouts/06-Quantization.pdf
Appendix
Properties of MMSE Quantizer

- $\mathcal{G}$ is an unbiased estimator of $\mathcal{F}$.

$$E(\mathcal{G}) = \sum_l P(\mathcal{B}_l) g_l$$
$$= \sum_l P(\mathcal{B}_l) \left( \int_{\mathcal{B}_l} f \ p(f | \mathcal{F} \in \mathcal{B}_l) \, df \right)$$
$$= \sum_l \int_{\mathcal{B}_l} f \ p(f) \, df$$
$$= \int_{\mathcal{B}} f \ p(f) \, df = E(\mathcal{F})$$

- $\mathcal{Q} = \mathcal{F} - \mathcal{G}$ is of zero mean.
Properties of MMSE Quantizer

- \( \mathcal{G} \) is orthogonal to (and uncorrelated with) \( \mathcal{Q} \).

\[
E(\mathcal{GQ}) = E(\mathcal{G}(\mathcal{F} - \mathcal{G})) \\
= E_\mathcal{G}E_{\mathcal{F}|\mathcal{G}}((\mathcal{G}(\mathcal{F} - \mathcal{G}))) \\
= E_\mathcal{G}(\mathcal{G}E_{\mathcal{F}|\mathcal{G}}(\mathcal{F}) - \mathcal{G}^2) \\
= E_\mathcal{G}(\mathcal{G}^2 - \mathcal{G}^2) \\
= 0
\]

- Note that \( E(\mathcal{FQ}) = E(\mathcal{F}(\mathcal{F} - \mathcal{G})) = E(\mathcal{F}^2) - E(\mathcal{FG}) = E(\mathcal{F}^2) - E(\mathcal{G}^2) \neq 0 \)
Properties of MMSE Quantizer

- Reduced signal variance

\[ \sigma_G^2 = \sigma_F^2 - \sigma_Q^2 \]

- By definition

\[
\begin{align*}
\sigma_Q^2 &= E((F - G)^2) \\
&= E\left(\left((F - \mu_F) - (G - \mu_G)\right)^2\right) \\
&= \sigma_F^2 + \sigma_G^2 - 2E((F - \mu_F)(G - \mu_G)) \\
&= \sigma_F^2 + \sigma_G^2 - 2\left(E(FG) - \mu_F \mu_G\right) \\
&= \sigma_F^2 + \sigma_G^2 - 2\left(E(G^2) - \mu_G^2\right) \\
&= \sigma_F^2 - \sigma_G^2
\end{align*}
\]

where \( \mu_F = \mu_G = E(G) \) and \( E(FG) = E(G^2) \)
MMSE Scalar Quantizer of Various Sources

Given a source $F' = (F - \mu_F) / \sigma_F$ of zero mean and unit variance

$$P(F' \leq f') = P(F \leq \sigma_F f' + \mu_F)$$

$$C_{F'}(f') = C_F(\sigma_F f' + \mu_F)$$

$$\Rightarrow \frac{d}{df'} C_{F'}(f') = \frac{d}{df'} C_F(\sigma_F f' + \mu_F)$$

$$\Rightarrow p_{F'}(f') = \sigma_F p_F(\sigma_F f' + \mu_F) \text{ or } p_F(f) = \frac{1}{\sigma_F} p_{F'}\left(\frac{f - \mu_F}{\sigma_F}\right)$$

where

$p_{F'}(f')$ and $p_F(f)$ are PDF of $F'$ and $F$, respectively
MMSE Scalar Quantizer of Various Sources

**Optimal Scalar Quantizer for $\mathcal{F}'$**

1. $\tilde{b}_l = \frac{\tilde{g}_l + \tilde{g}_{l+1}}{2}$
2. $\tilde{g}_l = E(\mathcal{F}'|\mathcal{F}' \in \tilde{\mathcal{B}}_l) = \frac{\int_{\tilde{b}_{l-1}}^{\tilde{b}_l} f' p_{\mathcal{F}'}(f') df'}{\int_{\tilde{b}_{l-1}}^{\tilde{b}_l} p_{\mathcal{F}'}(f') df'}$

**Optimal Scalar Quantizer for $\mathcal{F}$**

1. $b_l = \frac{g_l + g_{l+1}}{2} = \sigma_\mathcal{F} \tilde{b}_l + \mu_\mathcal{F}$
2. $g_l = E(\mathcal{F}|\mathcal{F} \in \mathcal{B}_l) = \frac{\int_{b_{l-1}}^{b_l} f p_{\mathcal{F}}(f) df}{\int_{b_{l-1}}^{b_l} p_{\mathcal{F}}(f) df} = \sigma_\mathcal{F} \tilde{g}_l + \mu_\mathcal{F}$
MMSE Scalar Quantizer of Various Sources

- Let
  \[ f' = \frac{f - \mu_F}{\sigma_F} \]

- Assume
  \[ \tilde{b}_l = \frac{b_l - \mu_F}{\sigma_F} \]

- Then optimal \( g_l \) for \( F \) can be obtained as follows:
  \[
  g_l = E(F|F \in B_l) = \frac{\int_{b_{l-1}}^{b_l} f \ p_F(f) \, df}{\int_{b_{l-1}}^{b_l} p_F(f) \, df} = \frac{\int_{b_{l-1}}^{b_l} f \ \frac{1}{\sigma_F} p_F'(\frac{f - \mu_F}{\sigma_F}) \, df}{\int_{b_{l-1}}^{b_l} \frac{1}{\sigma_F} p_F'(\frac{f - \mu_F}{\sigma_F}) \, df}
  \]
  \[
  = \frac{\int_{b_{l-1}}^{b_l} (\sigma_F f' + \mu_F) \ p_F'(f') \, df'}{\int_{b_{l-1}}^{b_l} p_F'(f') \, df'} = \sigma_F \tilde{g}_l + \mu_F \quad (5)
  \]

- From Eq. (5), the assumption \( \tilde{b}_l = (b_l - \mu_F) / \sigma_F \) is justified
  \[ b_l = (g_l + g_{l+1}) / 2 = \sigma_F \tilde{b}_l + \mu_F \]
MMSE Scalar Quantizer: High Rate Approximation

- **Quantizer Distortion**

\[
D_q \approx \sum_{l \in \mathcal{L}} \frac{P(B_l) \Delta_l^2}{12} = \sum_{l \in \mathcal{L}} \frac{p(g_l^*) \Delta_l^3}{12} = \sum_{l \in \mathcal{L}} \frac{\alpha_l^3}{12}
\]

where

\[
\alpha_l = \sqrt[3]{p(g_l^*) \Delta_l}
\]

- **Observe that**

\[
\sum_{l \in \mathcal{L}} \alpha_l = \sum_{l \in \mathcal{L}} \sqrt[3]{p(g_l^*) \Delta_l} \approx \int_{f_{\min}}^{f_{\max}} \sqrt[3]{p(f)} df = \text{constant}
\]

- **MMSE Scalar Quantizer**

\[
\min \sum_{l \in \mathcal{L}} \frac{\alpha_l^3}{12} \quad \text{s.t.} \quad \sum_{l \in \mathcal{L}} \alpha_l = C
\]

\[\Rightarrow \alpha_l^* = \text{constant} = \frac{1}{L} \int_{f_{\min}}^{f_{\max}} \sqrt[3]{p(f)} df\]
MMSE Scalar Quantizer High Rate Approximation

Given

\[ \alpha_i^* = \frac{1}{L} \int_{f_{\min}}^{f_{\max}} \sqrt[3]{p(f)} df \]

Quantizer distortion

\[ D_q \overset{!}{=} \sum_{l \in L} \frac{(\alpha_i^*)^3}{12} = L \frac{(\alpha_i^*)^3}{12} \]

\[ = \frac{1}{L^2} \frac{1}{12} \left( \int_{f_{\min}}^{f_{\max}} \sqrt[3]{p(f)} df \right)^3 \]

\[ = 2^{-2R} \frac{1}{12} \left( \int_{f_{\min}}^{f_{\max}} \sqrt[3]{p(f)} df \right)^3 \]

where \( L = 2^R \)
MMSE Scalar Quantizer High Rate Approximation

- Given $F' = F / \sigma_f$ is an unit variance source
  
  $$P(F' \leq f') = P(F \leq \sigma_f f')$$

  $$C_{F'}(f') = C_F(\sigma_f f')$$

  $$\Rightarrow \frac{d}{df'} C_{F'}(f') = \frac{d}{df'} C_F(\sigma_f f')$$

  $$\Rightarrow p_{F'}(f') = \sigma_f p_F(\sigma_f f') \text{ or } p_F(f) = \frac{1}{\sigma_f} p_{F'}\left(\frac{f}{\sigma_f}\right)$$

- Quantizer distortion

  $$D_q \approx 2^{-2R} \left( \int_{f_{min}}^{f_{max}} \frac{3}{\sqrt[3]{p_F(f)}} df \right)^3 = 2^{-2R} \left( \int_{f_{min}}^{f_{max}} \frac{1}{\sigma_f} p_{F'}\left(\frac{f}{\sigma_f}\right) df \right)^3$$

  $$= \sigma_f^2 2^{-2R} \epsilon^2$$

  where $\epsilon^2 = \frac{1}{12} \left( \int_{f_{min}}^{f_{max}} \frac{3}{\sqrt[3]{p_{F'}(f)}} df \right)^3$
MMAE Scalar Quantizer

- Observe that

\[
\int_{b_{l-1}}^{b_l} |f - g_l| p(f) df
= \int_{b_{l-1}}^{g_l} (g_l - f) p(f) df + \int_{g_l}^{b_l} (f - g_l) p(f) df
= g_l \int_{b_{l-1}}^{g_l} p(f) df - \int_{b_{l-1}}^{g_l} f p(f) df + \int_{g_l}^{b_l} f p(f) df - g_l \int_{g_l}^{b_l} p(f) df
\]

- Then

\[
\frac{\partial}{\partial g_l} J(b^*, g^*) = \frac{\partial}{\partial g_l} \left( \int_{b_{l-1}}^{b_l} |f - g_l| p(f) df \right)
= \int_{b_{l-1}}^{g_l} p(f) df - \int_{g_l}^{b_l} p(f) df
\]
References

1. Y. Wang, et. al - Video Processing and Communications
2. Jayant, N. S., et. al - Digital Coding of Waveforms