525.749 Image Compression, Packet Video and Video Processing

Dr. Nicholas Beser

E-mail: Nicholas.Beser@jhuapl.edu

11/09/09

Lecture 8 - Foundation of Video Coding

Part II: Scalar and Vector Quantization

http://webdev.apl.jhu.edu/~beser/525759/index.html

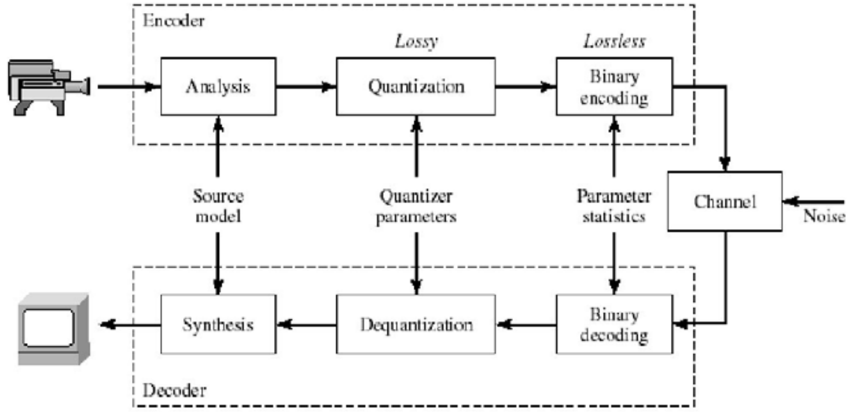
1

Outline

- Overview of video coding systems
- Scalar Quantization
- Vector Quantization
- Rate-distortion characterization of lossy coding
  - Operational rate distortion function
  - Rate distortion bound (lossy coding bound)

2

Components in a Coding System



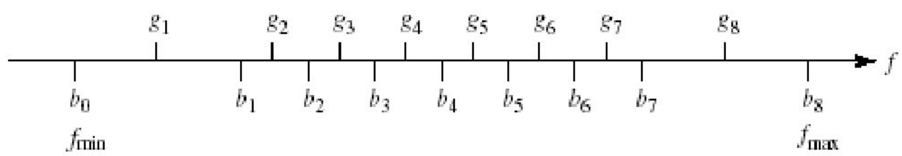
Lossy Coding

- **Original source is discrete**
  - **Lossless coding: bit rate >= entropy rate**
  - **One can further quantize source samples to reach a lower rate**
- **Original source is continuous**
  - **Lossless coding will require an infinite bit rate!**
  - **One must quantize source samples to reach a finite bit rate**
  - **Lossy coding rate is bounded by the mutual information between the original source and the quantized source that satisfy a distortion criterion**
- **Quantization methods**
  - **Scalar quantization**
  - **Vector quantization**

Scalar Quantization

- **General description**
- **Uniform quantization**
- **MMSE quantizer**
- **Lloyd algorithm**

SQ as Line Partition



Quantization Levels: L

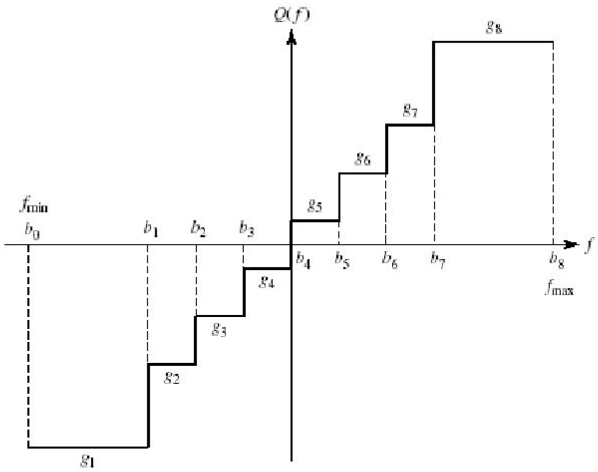
Boundary Values: b_i

Partition Regions: $B_l = [b_{l-1}, b_l)$

Reconstruction Values: g_l

Quantizer Mapping: $Q(f) = g_l, if \; f \in B_l$

Function Representation



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

Distortion Measure

- **General measure:**

$$D_q = E\{d_1(F, Q(F))\} = \int_{f \in B} d_1(f, Q(f))p(f)df$$

$$= \sum_{l \in L} P(B_l) D_{q,l}$$

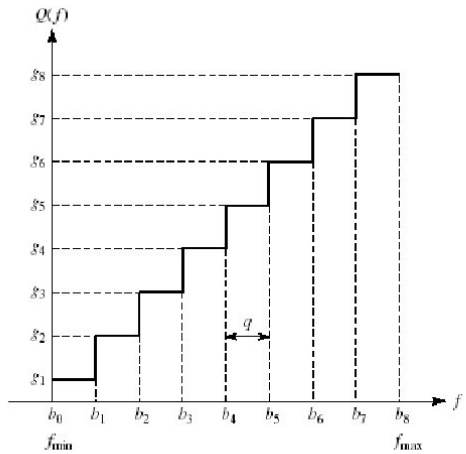
$$D_{q,l} = \int_{f \in B_l} d_1(f, g_l)p(f \mid f \in B_l)df.$$

- **Mean Square Error (MSE)**

$$\sigma_q^2 = E\left\{|F - Q(F)|^2\right\} = \sum_{l \in L} P(B_l) \int_{b_{l-1}}^{b_l} (f - g_l)^2 p(f \mid B_l)df.$$

IMAGE COMPRESSION, PACKET VIDEO AND VIDEO PROCESSING

Uniform Quantization



$$Q(f)=\left[\frac{f-f_{\min}}{q}\right]*q+\frac{q}{2}+f_{\min}.$$

Uniform Source:

$$p(f)=\begin{cases}1/B & f\in(f_{\min},f_{\max})\\ 0 & otherwise\end{cases}$$

$$\sigma_q^2=\frac{q^2}{12}=\sigma_f^2 2^{-2R}$$

$$SNR=10\log_{10}\frac{\sigma_f^2}{\sigma_q^2}$$

$$=(20\log_{10}2)R:=6.02R(dB)$$

9

IMAGE COMPRESSION, PACKET VIDEO AND VIDEO PROCESSING

Minimum MSE (MMSE) Quantizer

Determine b_l, g_l *to* minimize MSE

$$\sigma_q^2=E\left\{\left|F-Q(F)\right|^2\right\}=\sum_{l\in L}P(B_l)\int_{b_{l-1}}^{b_l}(f-g_l)^2\,p(f|B_l)df.$$

Setting $\dfrac{\partial\sigma_q^2}{b_l}=0,\dfrac{\partial\sigma_q^2}{g_l}=0$ yeilds:

$$b_l=\frac{g_l+g_{l+1}}{2},\ or\ B_l=\left\{f:d_1(f,g_l^{'})\le d_1(f,g_l^{'}),\forall l'\ne l\right\}$$ **(Nearest Neighbor Condition)**

$$g_l=E\left\{F\middle|F\in B_l\right\}=\int_{B_l}p(f|f\in B_l)df,$$ **(Centroid Condition)**

10

High Resolution Approximation

- **For a source with arbitrary pdf, when the rate is high so that the pdf within each partition region can be approximated as flat:**

$$\sigma_q^2 = \varepsilon^2 \sigma_f^2 2^{-2R}$$

$$\varepsilon^2 = \frac{1}{12}\left(\int_{-\infty}^{+\infty} \overline{p}(f)^{1/3} df \right)^3, \quad \overline{p}(f) = \sigma_f p(\sigma_f f)$$

$$\text{Uniform Source}: \varepsilon^2 = 1$$

$$\text{i.i.d Gaussian Source}: \varepsilon^2 = 2.71 (w/o\ VLC)$$

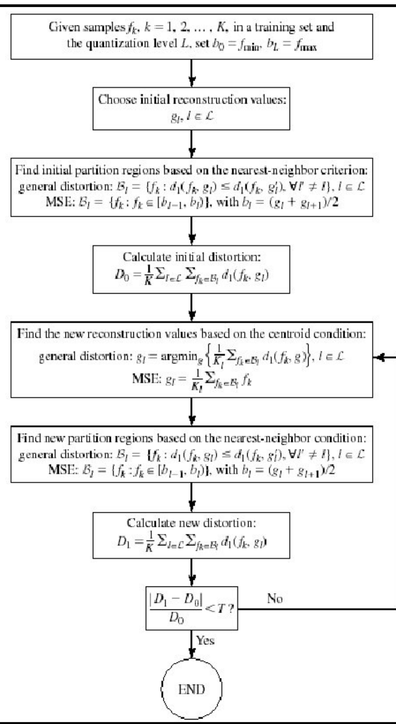
$$\text{Bound for Gaussian Source}: \varepsilon^2 = 1$$

11

IMAGE COMPRESSION, PACKET V

Lloyd Algorithm

- **Iterative algorithms for determining MMSE quantizer parameters**
- **Can be based on a pdf or training data**
- **Iterate between centroid condition and nearest neighbor condition**



12

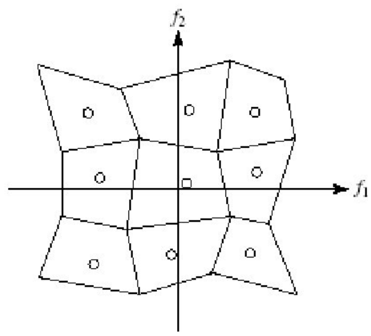
Vector Quantization

- General description
- Nearest neighbor quantizer
- MMSE quantizer
- Generalized Lloyd algorithm

Vector Quantization: General Description

- Motivation: quantize a group of samples (a vector) together, to exploit the correlation between these samples
- Each sample vector is replaced by one of representative vectors (or patterns) that often occur in the signal
- Applications:
  - Color quantization: Quantize all colors appearing in an image to N colors for display on a monitor that can only display N distinct colors at a time – Adaptive palette
  - Image quantization: Quantize every NxN block into one of the typical patterns (obtained through training). More efficient with larger block size, but block size are limited by complexity.

VQ as Space Partition



Original vector: $f \in R^N$

Quantization levels: L

Partition regions: B_l

Reconstruction vector (codeword): g_l

Quantizer mapping: $Q(f) = g_l, if \ f \in B_l$

Codebook: $C = \{g_l, l = 1,2,\ldots,L\}$

Bit rate: $R = \frac{1}{N}\log_2 L$

Every point in a region (B_l) is replaced by (quantized to) the point indicated by the circle (g_l)

IMAGE COMPRESSION, PACKET VIDEO AND VIDEO PROCESSING

Distortion Measure

General measure:

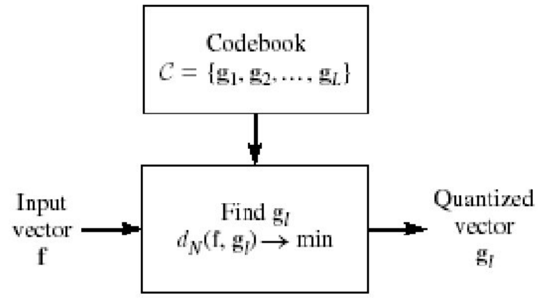
$$D_q = E\{d_N(F,Q(F))\} = \int_B p_N(f)d_N(f,Q(f))df$$

$$= \sum_{l=1}^{L} P(B_l)D_{q,l}$$

$$D_{q,l} = E\{d_N(F,Q(F))|F \in B_l\} = \int_{f \in B_l} P_N(f|f \in B_l)d_N(f,g_l)df,$$

$$\text{MSE}: \quad d_N(f,g) = \frac{1}{N}\sum_{n=1}^{N}(f_n - g_n)^2,$$

Nearest Neighbor (NN) Quantizer



$$B_l = \left\{ f \in R^N : d_N(f, g_l) \le d_N(f, g_{l'}), \forall l' \ne l \right\}$$

Challenge: How to determine the codebook?

Complexity of NN VQ

- Complexity analysis:
  - Must compare the input vector with every codewords
  - Each comparison takes N operations
  - Need L=2^{NR} comparisons
  - Total operation = N 2^{NR}
  - Total storage space = N 2^{NR}
  - Both computation and storage requirement increases exponentially with N!
- Example:
  - N=4x4 pixels, R=1 bpp: 16x2^16=2^20=1 Million operation/vector
  - Apply to video frames, 720x480 pels/frame, 30 fps: 2^20*(720x480/16)*30=6.8 E+11 operations/s !
  - When applied to image, block size is typically limited to <= 4x4
- Fast algorithms:
  - Structured codebook so that one can conduct binary tree search
  - Product VQ: can search subvectors separately

MMSE Vector Quantizer

- **Necessary conditions for MMSE**
  - **Nearest neighbor condition**

$$B_l = \{ f : d_N(f, g_l) \le d_N(f, g_l'), \forall l' \ne l \}.$$

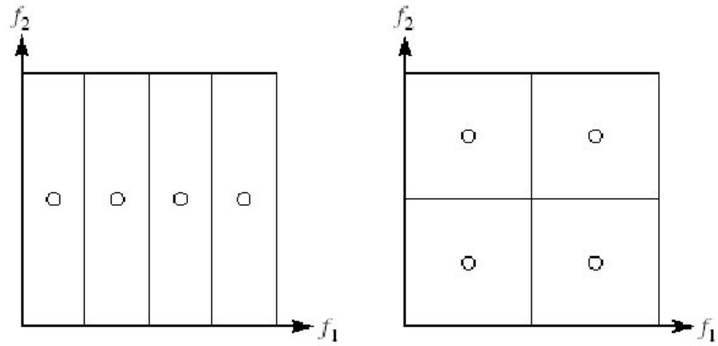
  - **Generalized centroid condition:**

$$g_l = \arg\min_g E\{ d_N(F, g) | F \in B_l \}.$$

  - **MSE as distortion:**

$$g_l = \int_{B_l} p(f \mid f \in B_l) f df = E\{ F | F \in B_l \}.$$

Caveats

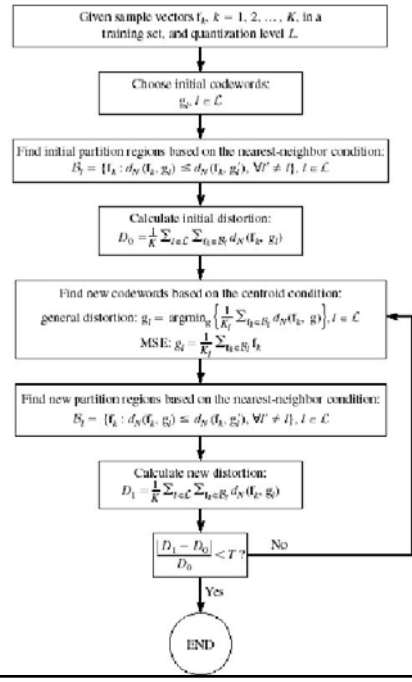


Both quantizer satisfy the NN and centroid condition, but the quantizer on the right is better!

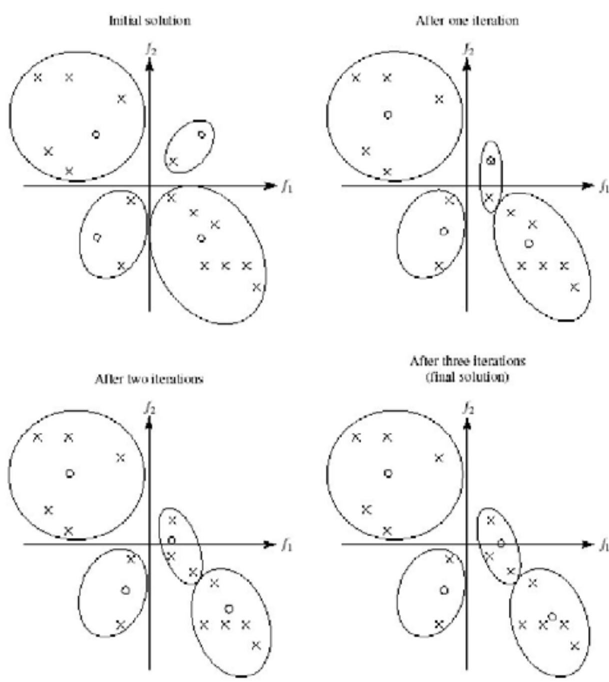
NN and centroid conditions are necessary but NOT sufficient for MSE optimality!

Generalized Lloyd Algorithm (LBG Algorithm)

- Start with initial codewords
- Iterate between finding best partition using NN condition, and updating codewords using centroid condition



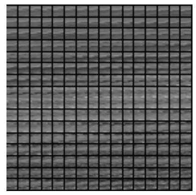
Example



LBG Example

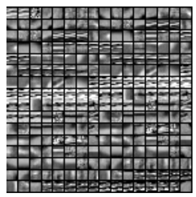
- **Training data created by concatenating 4 images**
- **Process chip size as 8x8**
- **Quantization set to 256**
- **Requires 61 iterations**
- **175 unique codes found**

Training Image



Source Image

Revised Training Image

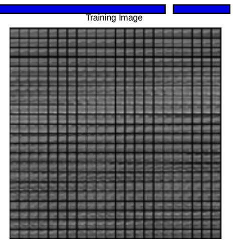


VQ Reconstruct Image

Large Code Book Example: LBG

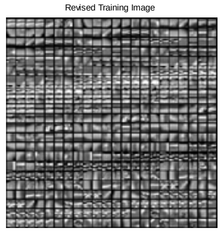
- **Training data created by concatenating 4 images**
- **Process chip size as 8x8**
- **Quantization set to 512**
- **Requires 59 iterations**
- **159 unique codes found**

Training Image

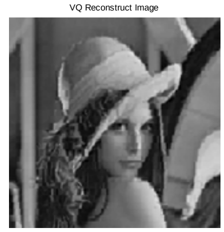


Source Image

Revised Training Image



VQ Reconstruct Image



Rate Distortion Bounds

- **Why is this important:**
  - **Rate distortion bounds implies that as you quantize the data, you introduce error**
  - **This quantization changes the data rate of the code**
  - **If you design a quantizer, you are trading code efficiency with distortion**
    - » **Optimal quantizers provide the most data with minimal error**
  - **The models for data will provide the probabilistic basis for determining the quality of the quantizer.**

IMAGE COMPRESSION, PACKET VIDEO AND VIDEO PROCESSING

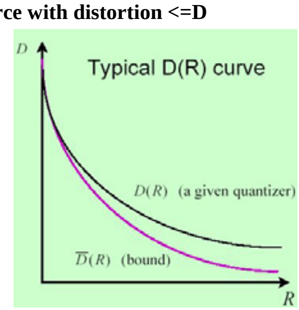
Rate-Distortion Characterization of Lossy Coding

- **Operational rate-distortion function of a quantizer:**
  - **Relates rate and distortion: R(D)**
  - **A vector quantizer reaches different point on its R(D) curve by using different number of codewords**
  - **Can also use distortion-rate function D(R)**
- **Rate distortion bound for a source**
  - **Minimum rate R needed to describe the source with distortion <=D**

$$\overline{R}(D) = \lim_{N\to\infty} \min_{q_N(g|f)\in Q_{n,N}} R_N(D; q_N(g|f))$$

$$Q_{D,N} = \{q_N(g|f) : E\{d_N(F,G)\} \le D\}$$

- **RD optimal quantizer:**
  - **Minimize D for given R or vice versa**



Lossy Coding Bound (Shannon Lossy Coding Theorem)

$$\bar{R}(D) = \lim_{N\to\infty} \min_{q_N(g|f)\in Q_{D,N}} R_N(D; q_N(g|f))$$

$$Q_{D,N} = \{q_N(g|f) : E\{d_N(F,G)\} \le D\}$$

$$\bar{R}(D) = \lim_{N\to\infty} \min_{q_N(g|f)\in Q_{D,N}} \frac{1}{N} I_N(F;G).$$

$I_N(F,G)$: **mutual information between *F* and *G*, information provided by *G* about *F***

Mutual Information measures the mutual dependence of two variables

$$I_N(F;G) = \sum_{f\in F}\sum_{g\in G} p(f,g)\log\frac{p(f,g)}{p(f)p(g)}$$

Lossy Coding Bound (Shannon Lossy Coding Theorem)

$Q_{D,N}$*: all coding schemes (or mappings q(g|f)) that satisfy distortion criterion *$d_N(f,g)$<=D

$$\bar{R}_L(D) \le \bar{R}(D) \le \bar{R}_G(D),$$

$$\bar{R}_L(D) = \bar{h}(F) - \frac{1}{2}\log_2 2\pi e D = \frac{1}{2}\log_2\frac{Q(F)}{D},$$

$R_G(D)$: **RD bound for Gaussian source with the same variance**

$R_L(D)$: is the Shannon lower bound

$\bar{h}(F)$: **entropy of source *F***

Q(F): **is the entropy power of *F***

For a Gaussian source, *$R_L(D)=R_G(D)$*

Bit Rate Limits

- **Among all sources with the same variance, The Gaussian source requires the highest bit rate to satisfy the same distortion criteria.**
- **Optimal coding scheme that achieves the RD bound is the one in which the quantization error sequence is i.i.d. Gaussian source with variance D, with a differential entropy**

$$\frac{1}{2}\log_2 2\pi e D$$

- **The lower bound is the difference between the differential entropy of the original source and that of the quantization error.**

RD Bound for Gaussian Source

- **i.i.d. 1-D Gaussian:**

$$\overline{D}(R) = \sigma^2 2^{-2R}.$$

$$\overline{R}(D) = \begin{cases} \frac{1}{2}\log_2 \frac{\sigma^2}{D} & 0 \le D < \sigma^2 \\ 0 & D \ge \sigma^2 \end{cases}$$

- **i.i.d. N-D Gaussian with independent components**

$$\overline{R}(D) = \frac{1}{2}\log_2 \frac{\left(\prod_n \sigma_n^2\right)^{1/N}}{D}, \quad \overline{D}(R) = \left(\prod_n \sigma_n^2\right)^{1/N} 2^{-2R}.$$

$$\textit{where} \quad D \le \min\{\sigma_n^2\}$$

i.i.d. Vector Gaussian Source with Correlated Components

- **If [C] represents the covariance matrix of the N components of each vector sample and let:**

$$\lambda_n, n = 1,2,\ldots N$$

- **Represent the eigenvalues of [C], RD bound for this source is given by**

$$\begin{cases} R(\alpha) = \frac{1}{N}\sum_n \max\left\{0, \frac{1}{2}\log_2 \frac{\lambda_n}{\alpha}\right\} \\ D(\alpha) = \frac{1}{N}\sum_n \min\{\alpha, \lambda_n\} \end{cases}$$

$$\textit{where} \quad \alpha \in (0, \max\{\lambda_n\})$$

For Small Distortion

- **When D is sufficiently small that the preceding reduces to:**

$$\bar{R}(D) = \frac{1}{2}\log_2 \frac{\left(\prod_n \lambda_n\right)^{1/N}}{D} = \frac{1}{2}\log_2 \frac{|\det[C]|^{1/N}}{D},$$

$$\bar{D}(R) = \left(\prod_n \lambda_n\right)^{1/N} 2^{-2R} = |\det[C]|^{1/N} 2^{-2R}.$$

- **Obtained by transforming the original vector using a transform matrix consisting of the eigenvectors of |C| and then applying the RD Bound to the transformed vector which will now have independent components.**

Summary

- **Coding system:**
  - original data -> model parameters -> quantization-> binary encoding
- **Quantization:**
  - Scalar quantization:
    - » Uniform quantizer
    - » MMSE quantizer (Nearest neighbor and centroid condition)
  - Vector quantization
    - » Nearest neighbor quantizer
    - » MMSE quantizer
    - » Generalized Lloyd alogorithm
  - Uniform quantizer
    - » Can be realized by lattice quantizer (not discuss here)
- **Rate distortion characterization of lossy coding**
  - Bound on lossy coding
  - Operational RD function of practical quantizers

Homework

- **Reading assignment:**
  - Sec. 8.5-8.7, 8.3.2,8.3.3