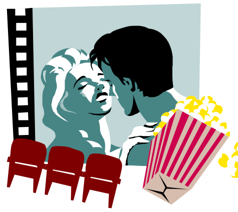
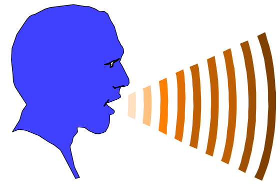





Video Processing & Communications

Foundation of Video Coding
Part II: Scalar and Vector Quantization

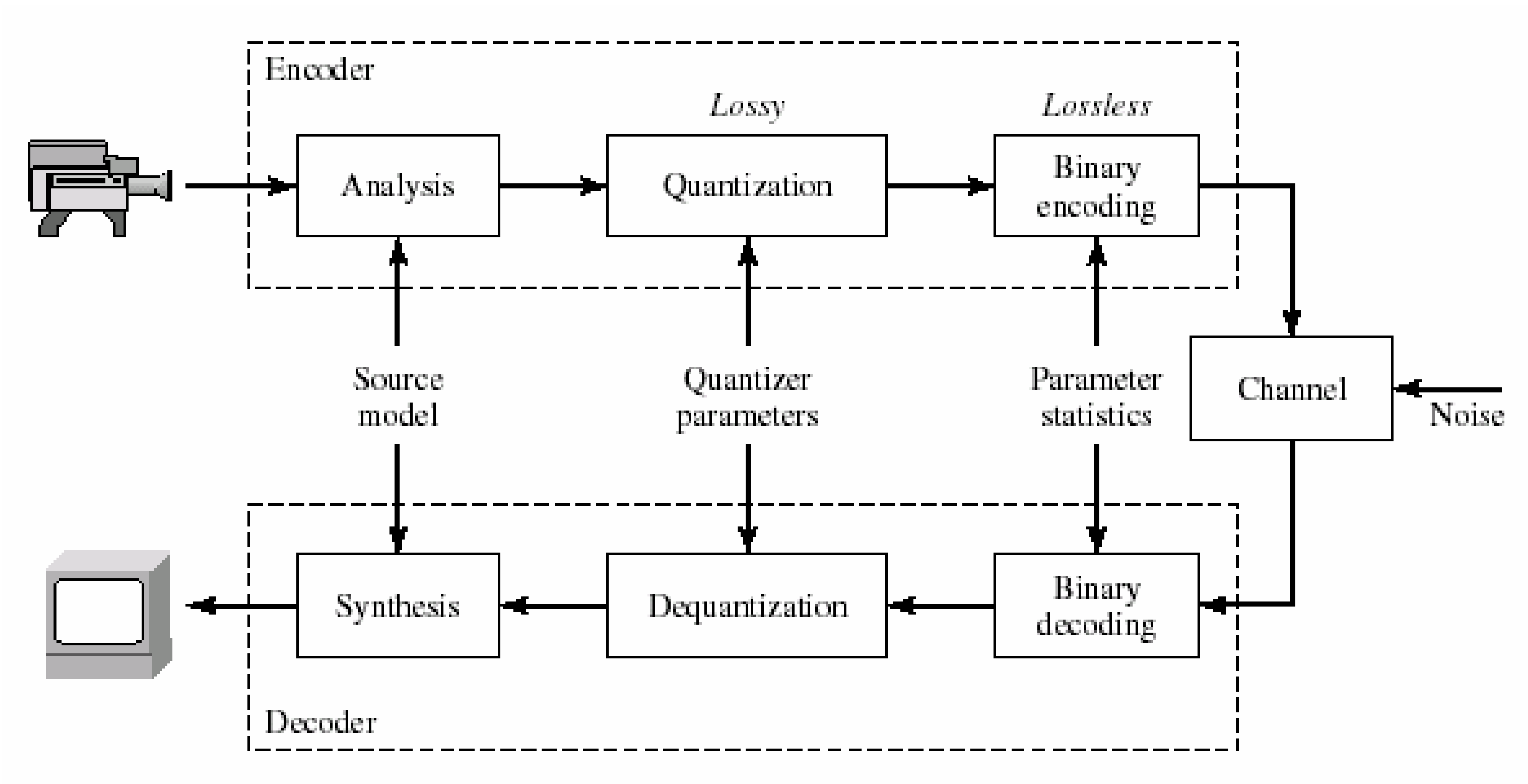
Yao Wang
Polytechnic University, Brooklyn, NY11201
http://eeweb.poly.edu/~yao

Based on: Y. Wang, J. Ostermann, and Y.-Q. Zhang, Video Processing and Communications, Prentice Hall, 2002.

Outline

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

Components in a Coding System



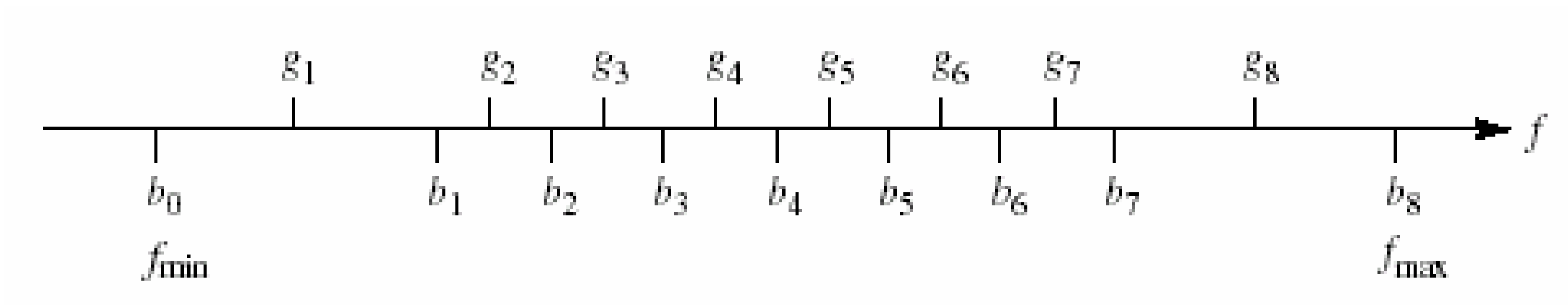
Lossy Coding

- 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

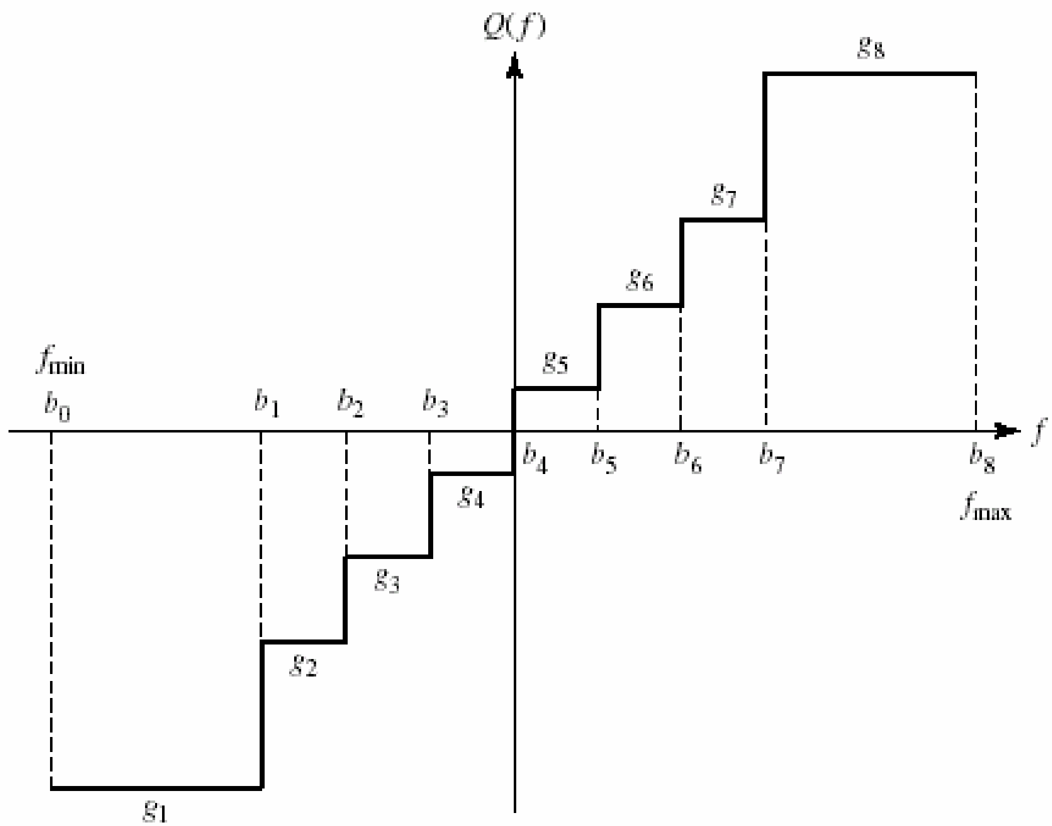
Boundary values: b_l

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

Reconstruction values: g_l

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

Function Representation



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

Distortion Measure

General measure:

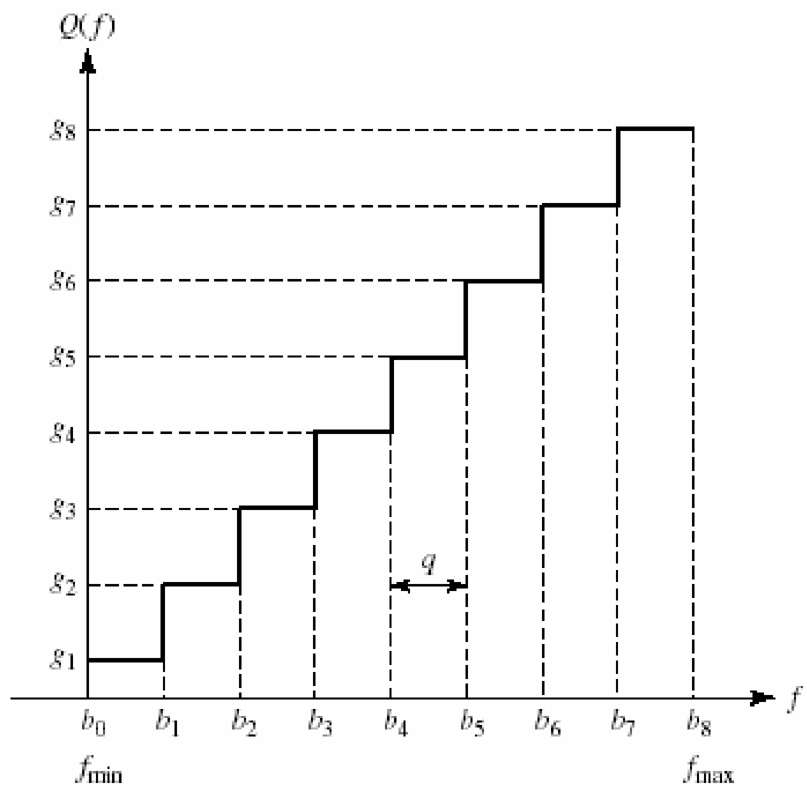
$$D_q = E\{d_1(\mathcal{F}, Q(\mathcal{F}))\} = \int_{f \in \mathcal{B}} d_1(f, Q(f))\, p(f)\, df$$
$$= \sum_{l \in \mathcal{L}} P(\mathcal{B}_l)\, D_{q,l}$$

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

Mean Square Error (MSE): $d_1(f,g) = (f-g)^2$

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

Uniform Quantization



$$Q(f) = \left\lfloor \frac{f - f_{\min}}{q} \right\rfloor * q + \frac{q}{2} + f_{\min},$$

Uniform source:

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

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

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

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

$$= 6.02 R \text{ (dB)}$$

Each additional bit provides 6dB gain!

Minimum MSE (MMSE) Quantizer

Determine b_l, g_l to minimize MSE

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

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

$b_l = \dfrac{g_l + g_{l+1}}{2},$ or $\mathcal{B}_l = \{f : d_1(f, g_l) \le d_1(f, g_{l'}), \forall l' \ne l\}.$ (Nearest Neighbor Condition)

$g_l = E\{\mathcal{F} \mid \mathcal{F} \in \mathcal{B}_l\} = \int_{\mathcal{B}_l} f\, p(f \mid f \in \mathcal{B}_l)\, df.$ (Centroid Condition)

- Special case: uniform source
  - MSE optimal quantizer = Uniform quantizer

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:

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

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

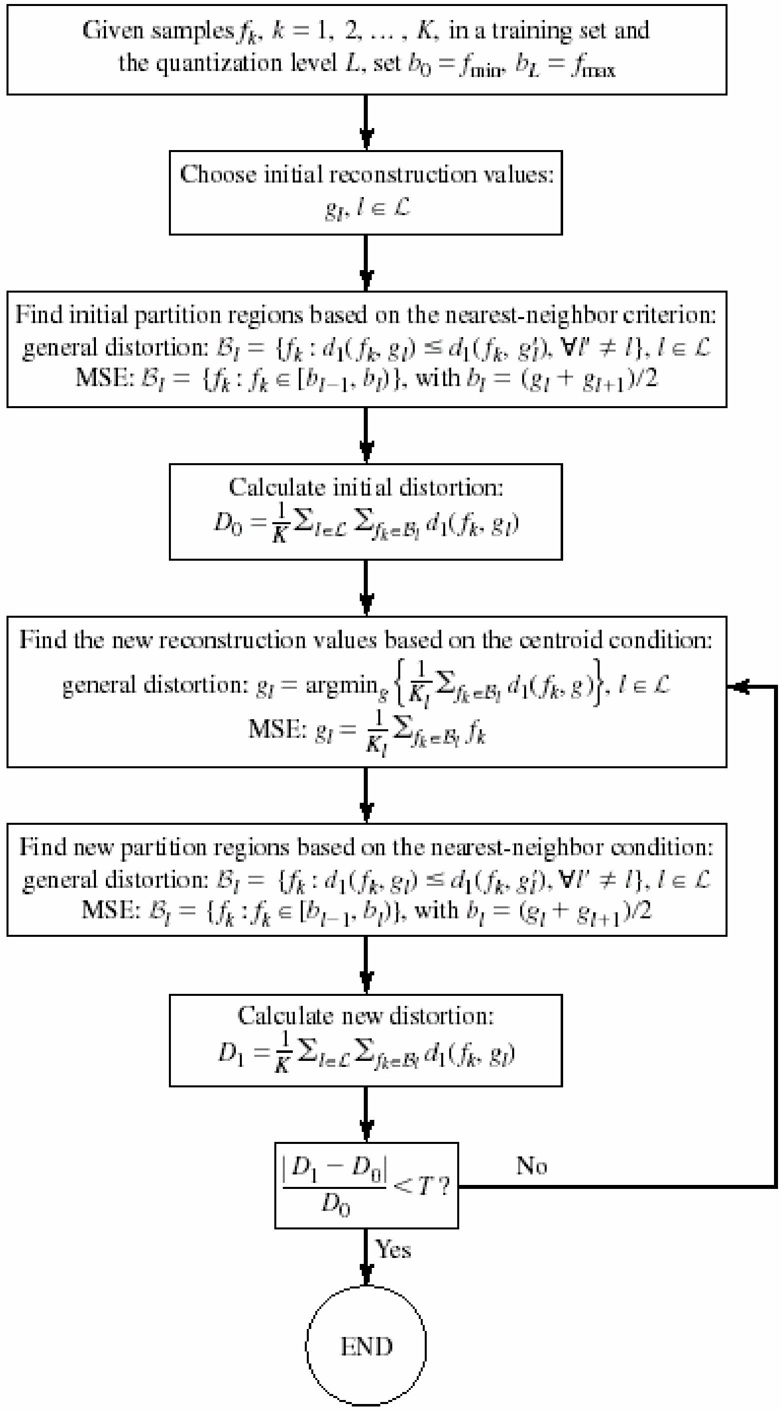
Uniform source: $\varepsilon^2 = 1$

i.i.d Gaussian source: $\varepsilon^2 = 2.71$ (w/o VLC)

Bound for Gaussian source: $\varepsilon^2 = 1$

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



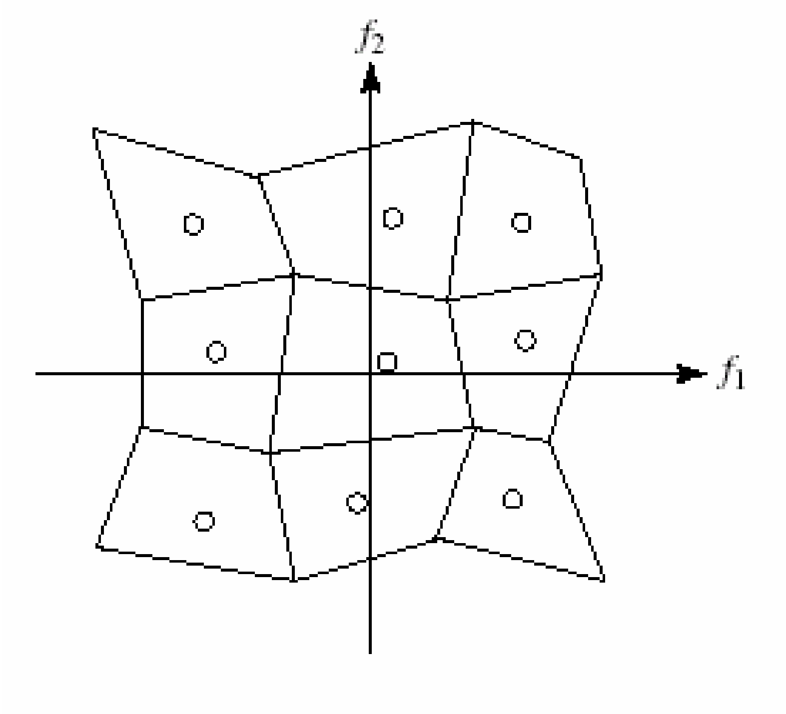
Vector Quantization

- 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 *L* colors for display on a monitor that can only display *L* distinct colors at a time – Adaptive palette
  - Image quantization: Quantize every *NxN* block into one of the *L* typical patterns (obtained through training). More efficient with larger block size, but block size are limited by complexity.

VQ as Space Partition



Original vector: $\mathbf{f} \in R^N$

Quantization levels: L

Partition regions: B_l

Reconstruction vector (codeword): $\mathbf{g}_l$

Quantizer mapping: $Q(\mathbf{f}) = \mathbf{g}_l, \text{ if } \mathbf{f} \in B_l$

Codebook: $C = \{\mathbf{g}_l, l = 1,2,\ldots,L\}$

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 ($\mathbf{g}_l$)

Distortion Measure

General measure:

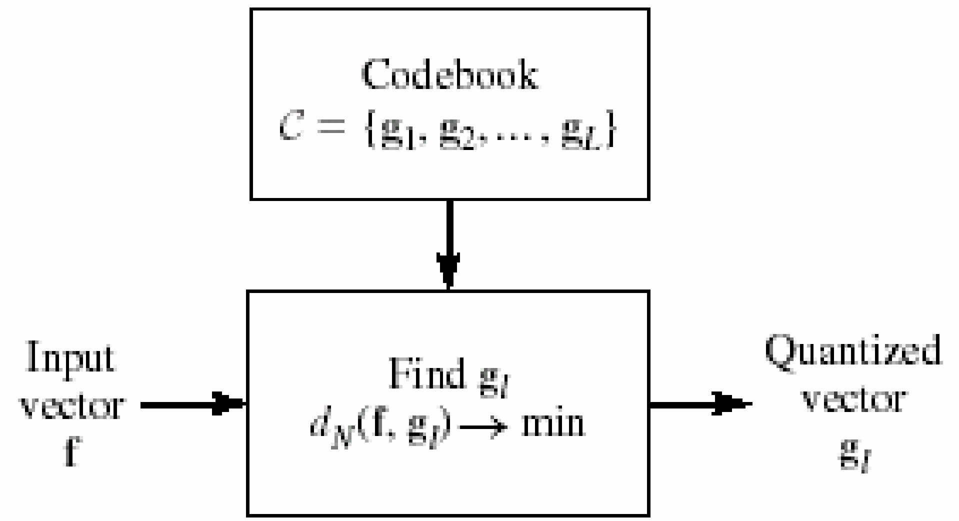
$$D_q = E\{d_N(\mathcal{F}, Q(\mathcal{F}))\} = \int_{\mathcal{B}} p_N(\mathbf{f}) d_N(\mathbf{f}, Q(\mathbf{f}))\, d\mathbf{f}$$
$$= \sum_{l=1}^{L} P(\mathcal{B}_l)\, D_{q,l}$$

$$D_{q,l} = E\{d_N(\mathcal{F}, Q(\mathcal{F})) \mid \mathcal{F} \in \mathcal{B}_l\} = \int_{\mathbf{f}\in\mathcal{B}_l} p_N(\mathbf{f} \mid \mathbf{f} \in \mathcal{B}_l) d_N(\mathbf{f}, \mathbf{g}_l)\, d\mathbf{f}.$$

MSE:

$$d_N(\mathbf{f}, \mathbf{g}) = \frac{1}{N} \sum_{n=1}^{N} (f_n - g_n)^2.$$

Nearest Neighbor (NN) Quantizer



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

Challenge: How to determine the codebook?

Complexity of NN VQ

- Complexity analysis:
  - Must compare the input vector with all the 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

$$\mathcal{B}_l = \{\mathbf{f} : d_N(\mathbf{f}, \mathbf{g}_l) \leq d_N(\mathbf{f}, \mathbf{g}_l'), \forall l' \neq l\}.$$

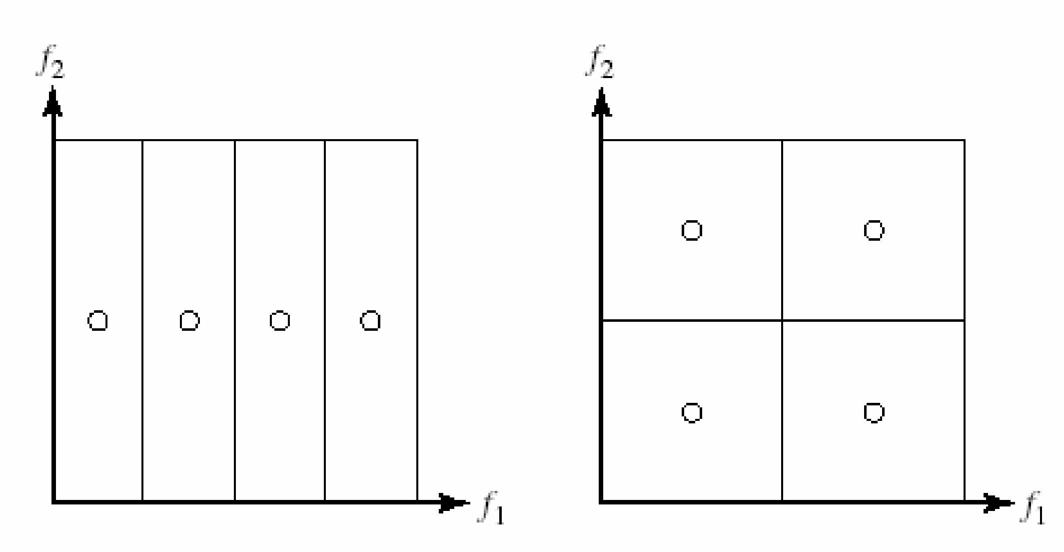
  - Generalized centroid condition:

$$\mathbf{g}_l = \operatorname{argmin}_{\mathbf{g}} E\{d_N(\mathcal{F}, \mathbf{g}) \mid \mathcal{F} \in \mathcal{B}_l\}.$$

  - MSE as distortion:

$$\mathbf{g}_l = \int_{\mathcal{B}_l} p(\mathbf{f} \mid \mathbf{f} \in \mathcal{B}_l)\mathbf{f}\, d\mathbf{f} = E\{\mathcal{F} \mid \mathcal{F} \in \mathcal{B}_l\}.$$

Caveats ☹

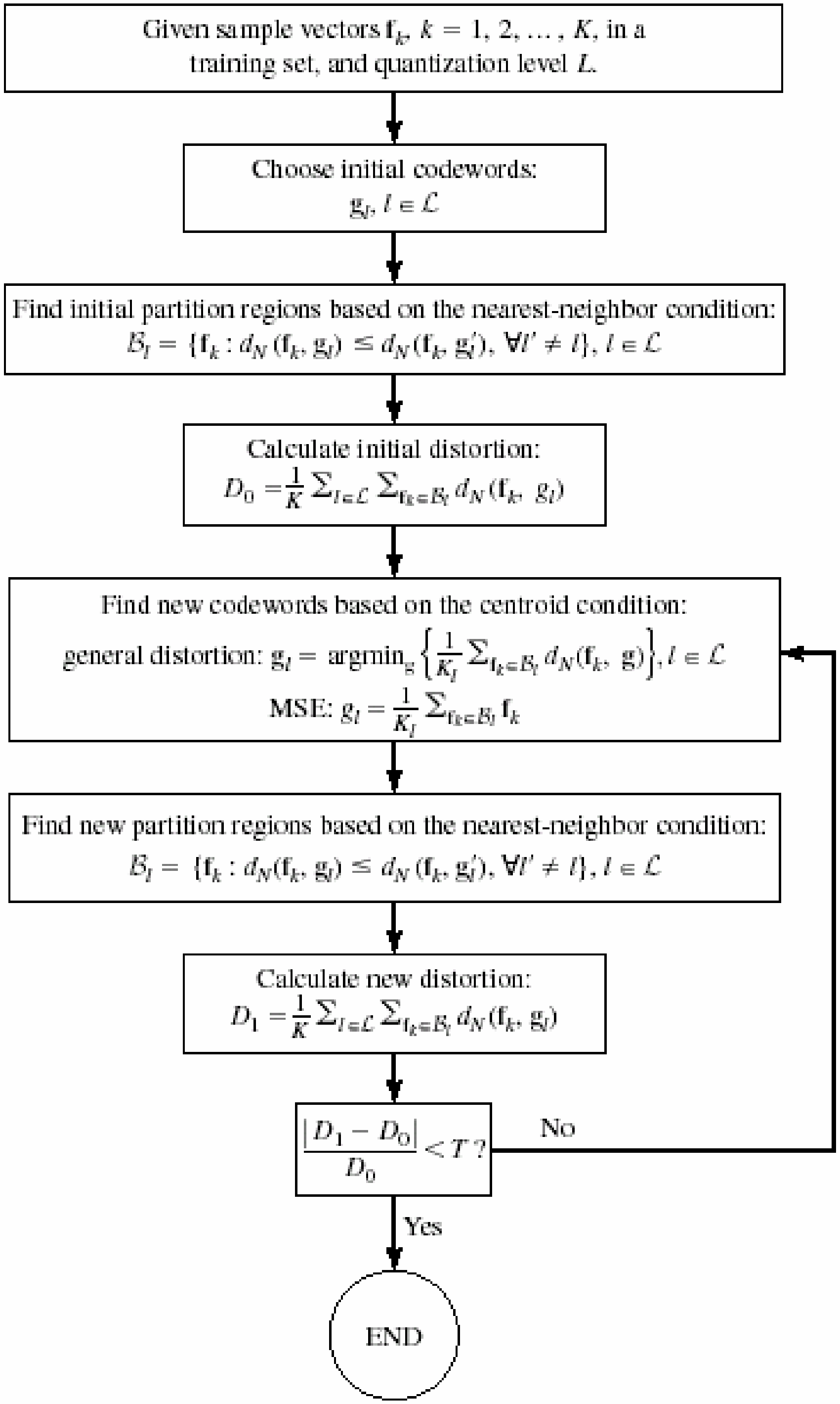


Both quantizers satisfy the NN and centroid condition, but the quantizer on the right is better!

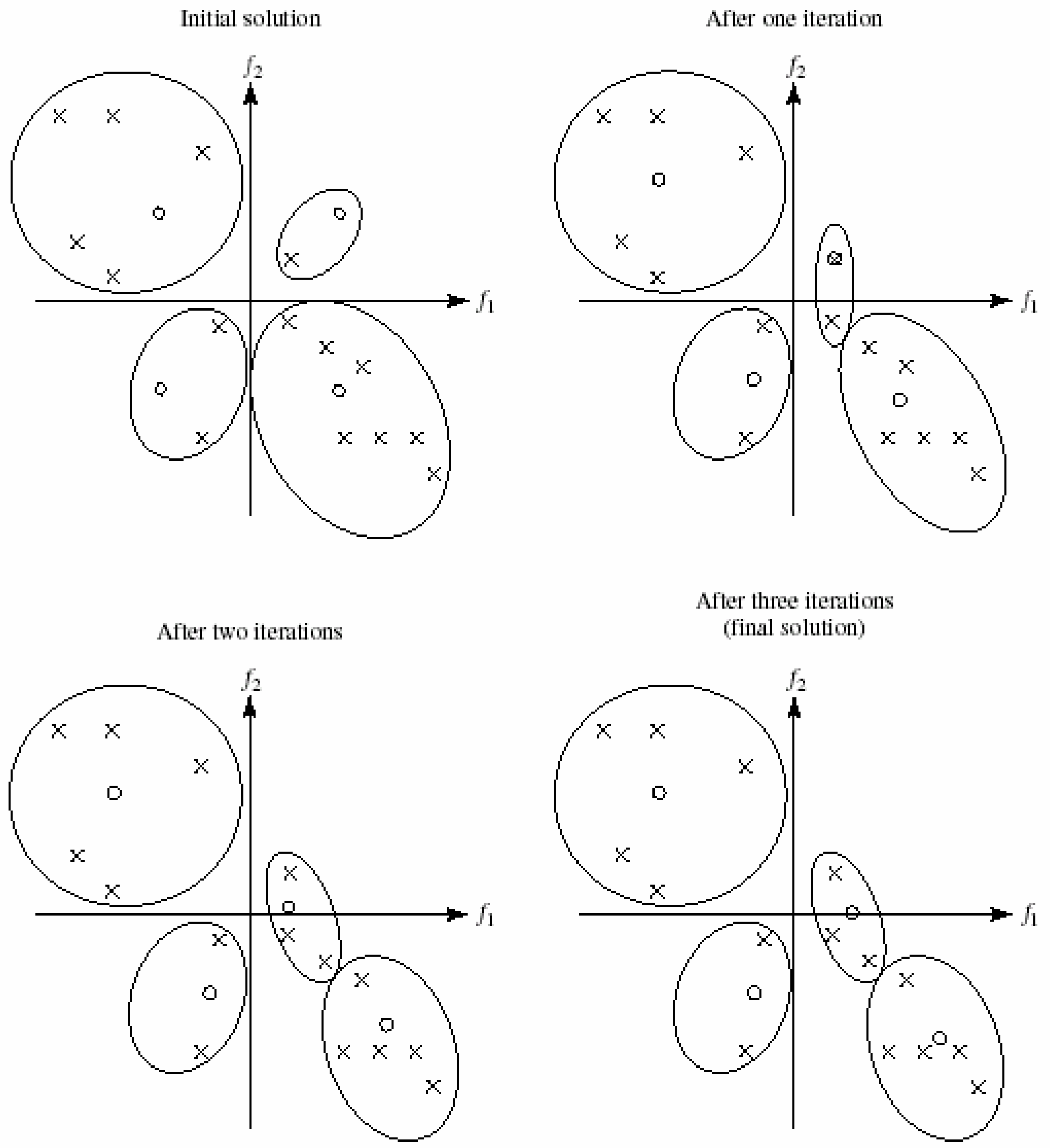
NN and centroid conditions are necessary but NOT sufficient for MSE optimality!

Generalized Lloyd Algorithm (LBG Algorithm)

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



Example


Initial solution
After one iteration
After two iterations
After three iterations
(final solution)
f_1
f_2

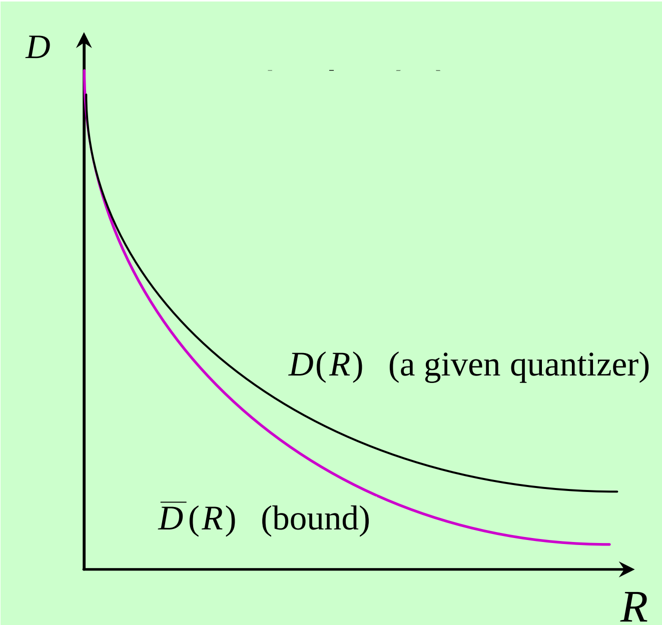
Rate-Distortion Characterization of Lossy Coding

- Operational rate-distortion function of a quantizer:
  - Relates rate and distortion: R(D)
  - A vector quantizer reaches a different point on its R(D) curve by using a 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

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

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

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



Lossy Coding Bound (Shannon Lossy Coding Theorem)

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

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

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

$I_N(\mathbf{F},\mathbf{G})$*:* mutual information between *F* and *G*, information provided by *G* about *F*
$Q_{D,N}$: all coding schemes (or mappings q(**g**|**f**)) that satisfy distortion criterion $d_N(f,\mathbf{g}) <= D$

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

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

$h(\mathbf{F})$*:* differential entropy of source ***F***
$R_G(D)$*:* RD bound for Gaussian source with the same variance
i.i.d. Gaussian source requires highest bit rate!

RD Bound for Gaussian Source

- i.i.d. 1-D Gaussian: $\bar{D}(R) = \sigma^2 2^{-2R}.$
- i.i.d. N-D Gaussian with independent components:

$$\bar{D}(R) = \left(\prod_n \sigma_n^2\right)^{1/N} 2^{-2R}.$$

- N-D Gaussian with covariance matrix **C**:

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

- Gaussian source with power spectrum (FT of correlation function) $S(e^{j\omega})$

$$\bar{R}(D) = \frac{1}{4\pi}\int_{-\pi}^{\pi} \log_2 \frac{S(e^{j\omega})}{D}\, d\omega.$$

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 discussed 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
- Written assignment
  - Prob. 8.8,8.11,8.14
- Computer assignment
  - Option 1: Write a program to perform vector quantization on a gray scale image using 4x4 pixels as a vector. You should design your codebook using all the blocks in the image as training data, using the generalized Lloyd algorithm. Then quantize the image using your codebook. You can choose the codebook size, say, L=128 or 256. If your program can work with any specified codebook size L, then you can observe the quality of quantized images with different L.
  - Option 2: Write a program to perform color quantization on a color RGB image. Your vector dimension is now 3, containing R,G,B values. The training data are the colors of all the pixels. You should design a color palette (i.e. codebook) of size L, using generalized Lloyd algorithm, and then replace the color of each pixel by one of the color in the palette. You can choose a fixed L or let L be a user-selectable variable. In the later case, observe the quality of quantized images with different L.