Audio Coding

- ***Fundamentals***
- ***Quantization***
- ***Waveform Coding***
- ***Subband Coding***

1. Fundamentals

- **Introduction**
- **Data Redundancy**
  - *Coding Redundancy*
  - *Spatial/Temporal Redundancy*
  - *Perceptual Redundancy*
- **Compression Models**
  - *A General Compression System Model*
  - *The Sourc/Channel e Encoder and Decoder*
- **Information Theory**
  - *Information*
  - *Entropy*
  - *Conditional Information & Entropy*
  - *Mutual Information*

1.1 Introduction

Compression

- *Reduce the amount of data required to represent a media*

Why Compression

- *Stereo Audio*
  - *16 bits for 96 dB*
  - *44.1 k sample rate*
  - *176.4 k bytes per second and 10Mbytes for a minute*
- *Video*
  - *525 x 360 x 30 x 3 = 17 MB/s or 136 Mb/s*
  - *1000 Mbytes for a minute*

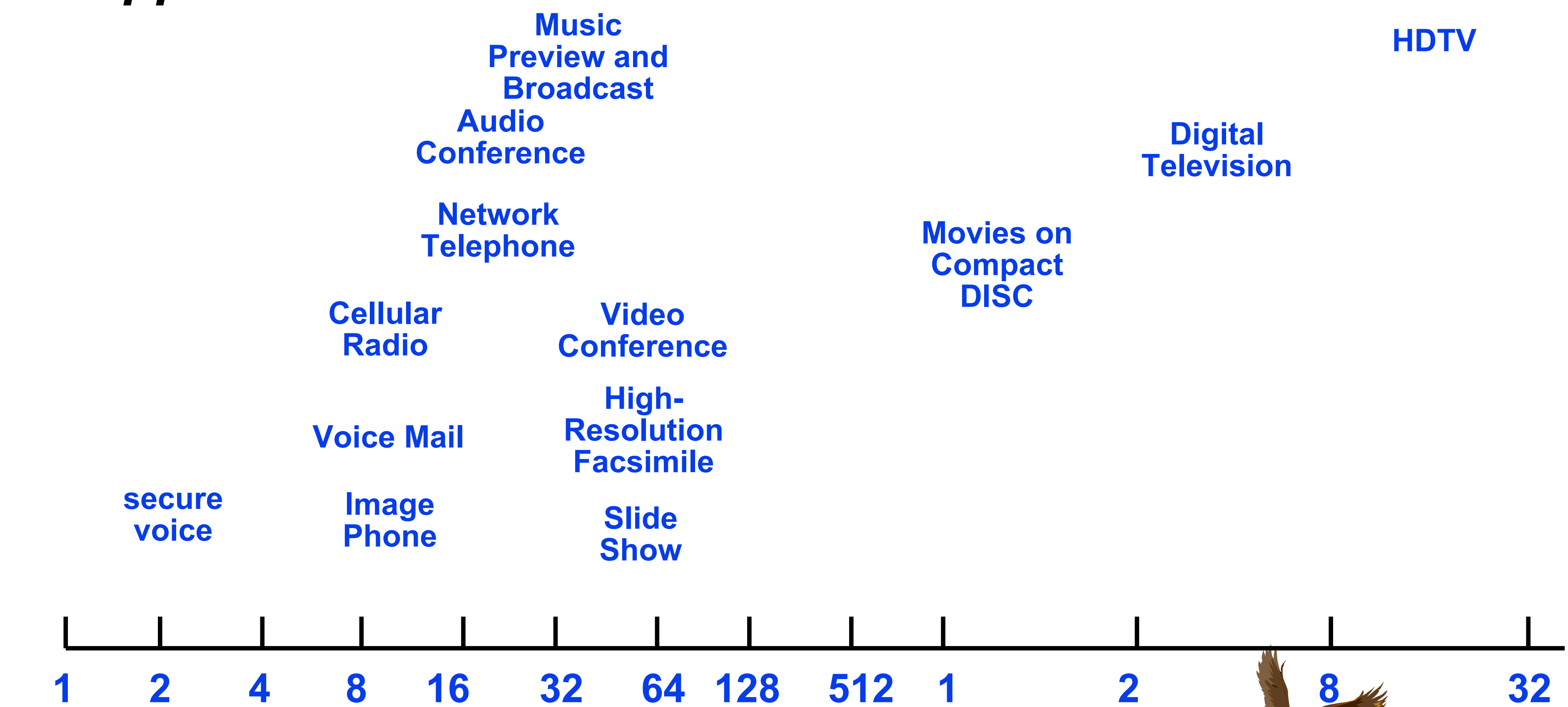
Compression is necessary for storage, communication, ...

1.1 Introduction (c.1)

- **Advantages of Digital over Analog Signals**
  - *Processing Flexibility and Facility*
  - *Ease of Precision Control*
  - *Higher Signal-to-Noise Resistance*
- **Techniques to Compress Data**
  - *Data Redundancies*
  - *Perceptual Effects*
  - *Applications Requirements*
- **Standards**
  - *Speed up the advance of related technology*
  - *Increase the compatibility*
  - *The landmarks of technical developments.*

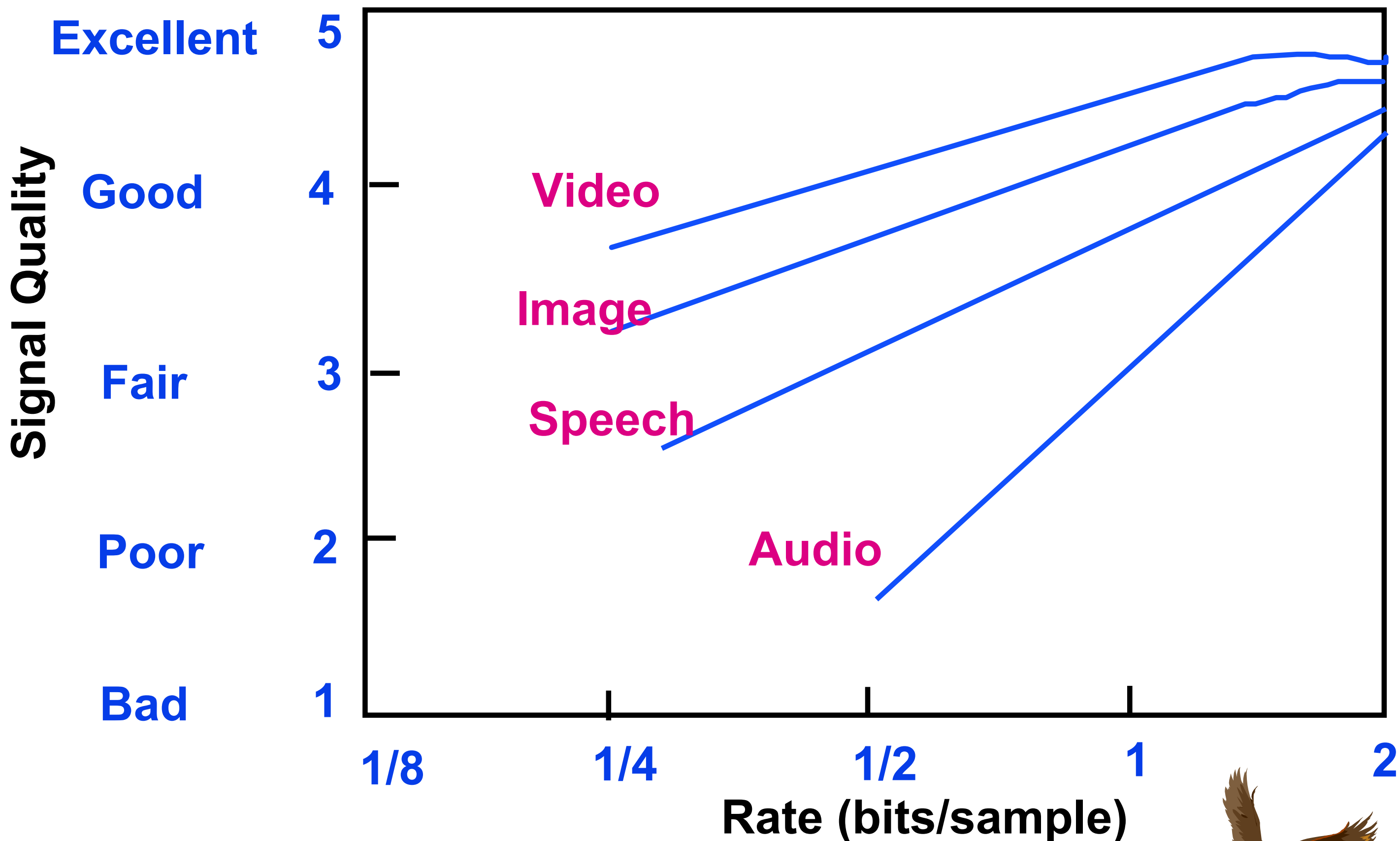
1.1 Introduction (c.2)

Applications



1.1 Introduction (c.3)

Current Technology



1.2 Redundancy

- Data v.s. Information
- Coding Redundancy
- Interdata Redundancy
- Perceptual Redundancy

1.2 Redundancy-- Data v.s. Information

Data Compression

- *Process of reducing the amount of data required to represent a given quantity of information.*

Data v.s. Information

- *Data are means by which information is conveyed.*

Data Redundancy

- *The part of data that contains no relevent information*
- *Not an abstract concept but a mathematically quantifiable entity*

1.2 Redundancies-- Data v.s. Information(c.1)

Example

- *If n_1 and n_2 denote the number of information carrying units for the same information*

➔ **Relative Data Redundancy, R_d**

$$R_d = 1 - \frac{1}{C_r}$$

➔ **Compression ratio, C_r**

$$C_r = \frac{n_1}{n_2}$$

☒ *n2 >>n1 ==> large compression ratio and low relative redundancy.*

1.2 Redundancy-- Coding Redundancy

- **Redundancy Sources**
  - *The number of bits used to represent different symbols needs not be the same.*
- **Assume that the occurance probability of each symbol r_k is $p(r_k)$ and the number of bits used to represent r_k is $l(r_k)$**
  - *Average number of bits for a symbol is*

$$L_{avg} = \sum_k l(r_k)\, p(r_k)$$

- **Variable Length Coding**
  - *Assign fewer bits to the more probable symbols for compression.*

1.2 Redundancy-- Coding Redundancy(c.1)

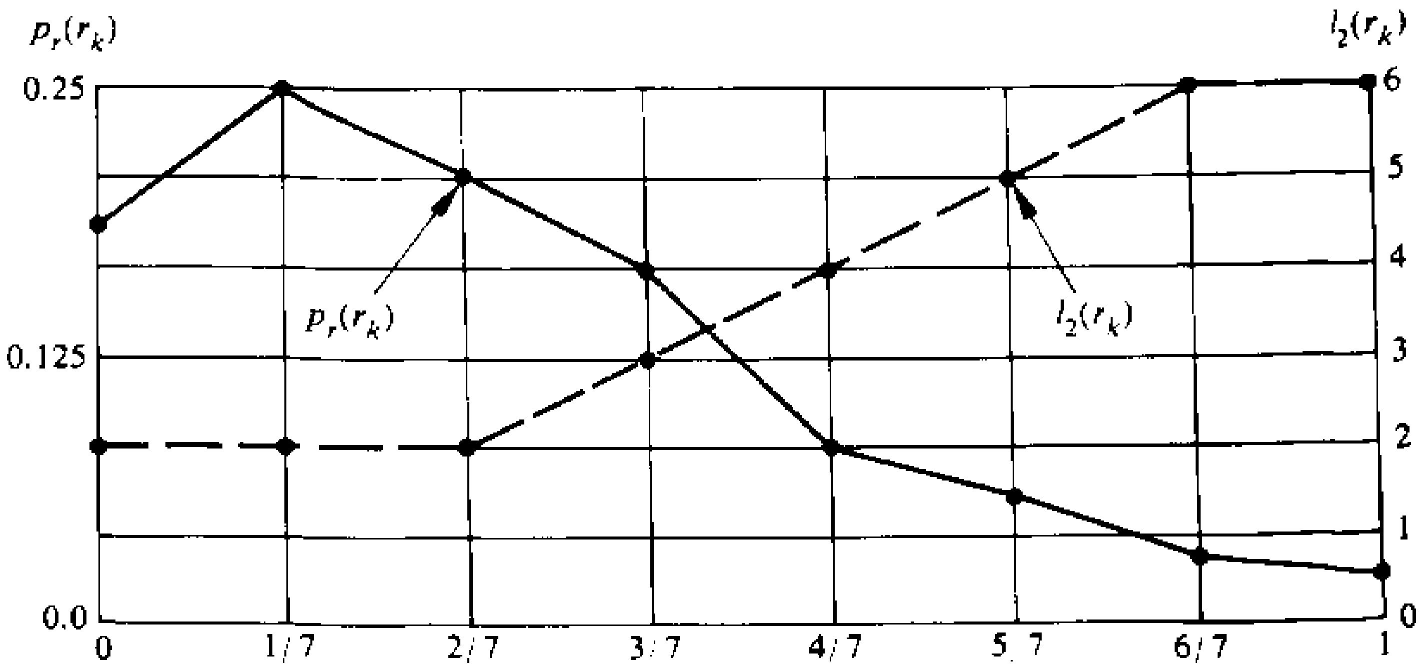
Variable-Length Coding Example

Table 6.1 Variable-Length Coding Example

r_k	$p_r(r_k)$	*Code 1*	$l_1(r_k)$	*Code 2*	$l_2(r_k)$
$r_0 = 0$	0.19	000	3	11	2
$r_1 = 1/7$	0.25	001	3	01	2
$r_2 = 2/7$	0.21	010	3	10	2
$r_3 = 3/7$	0.16	011	3	001	3
$r_4 = 4/7$	0.08	100	3	0001	4
$r_5 = 5/7$	0.06	101	3	00001	5
$r_6 = 6/7$	0.03	110	3	000001	6
$r_7 = 1$	0.02	111	3	000000	6

1.2 Redundancy-- Coding Redundancy(c.2)

- L_{avg} = 2.7 bits; C_r = 3/2.7=1.1; R_d = 1- (1/1.11) = 0.099
- Graphical representation the data compression



1.2 Redundancy-- Interdata Redundancy

- There is correlation between data
- The value of a data can be predicted from its neighbors
  - The information carried by individual data is relatively small.
- Other names
  - Interpixel Redundancy, Spatial Redundancy, Temporal Redundancy
- Ex.
  - Run-length coding

1.2 Redundancy-- Perceptual Redundancy

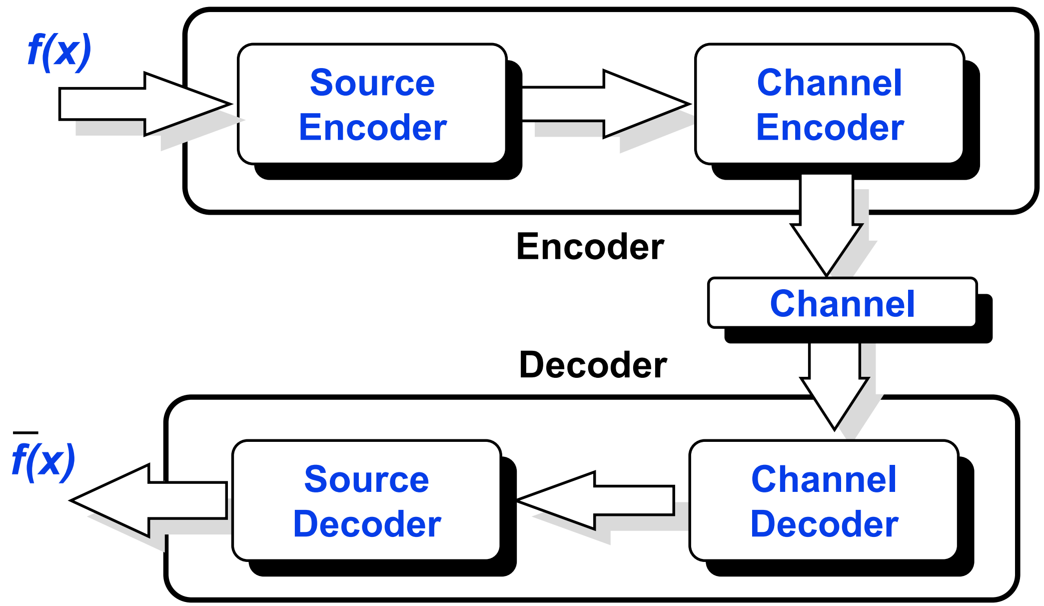
- **Certain information is not essential for normal perceptual processing**
- **Example:**
  - *Sharpe edges in an image.*
  - *Stronger sounds mask the weaker sounds.*
- **Other names**
  - *Psychovisual redundancy*
  - *Psychoacoustic redundancy*

1.3 Compression Models

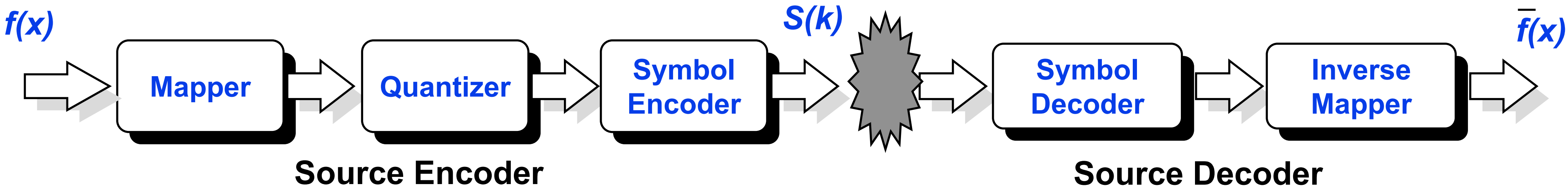
- A General Compression System Model
- The Source Encoder and Decoder
- The Channel Encoder and Decoder

1.3 Compression Models-- A General Compression System Model

- **Encoder**
  - *Create a set of symbol from input data*
- **Source Encoder**
  - *Removes input redundancies*
- **Channel Encoder**
  - *Increases the noise immunity of the source encoder output.*
- **Decoder**
  - *Reconstruct the input data*



1.3 Compression Models-- The Source Encoder and Decoder



- **Mapper**
  - *Transform the input data into a form designed to reduce interdata redundancies.*
- **Quantizer**
  - *Reduces the accuracy of the mapper output in accordance with some preestablished fidelity criterion.*
  - *Irreversible,reduce perceptual redundancy*
- **Symbol Encoder**
  - *Creates a fixed- or variable-length codeto represent the quantizer output and maps the output in accordance with the code.*
  - *Reduce coding redundancy*

1.3 Compression Models-- The Channel Encoder and Decoder

- Reduce the impact of channel noise by inserting a controlled form of redundancy.
- Example: (7, 4) Hamming Code
  - Encoding 4-bit word

$$h_1 = b_3 \oplus b_2 \oplus b_0 \qquad h_2 = b_3 \oplus b_1 \oplus b_0$$

$$h_4 = b_2 \oplus b_1 \oplus b_0 ; h_3 = b_3 ; h_5 = b_2 ; h_6 = b_1 ; h_7 = b_0$$

  - Decoding

$$c_1 = h_1 \oplus h_3 \oplus h_5 \oplus h_7 \qquad c_2 = h_2 \oplus h_3 \oplus h_6 \oplus h_7$$

$$c_4 = h_4 \oplus h_5 \oplus h_6 \oplus h_7$$

1.4 Information Theory

- Information
- Entropy
- Conditional Information & Entropy
- Mutual Information

1.4 Information Theory (c.1)

Introduction

- What does Information Theory talk about ?
  - The field of information theory is concerned with the amount of uncertainty associated with the outcome of an experiment .
  - The amount of information we receive when the outcome is known depends upon how much uncertainty there was about its occurren

1.4 Information Theory (c.2)

Shannon formalism

- *A random event E that occurs with probability P(E) is said to contain*

$$I(E) = \log {}^{1}\!/_{P(E)} = -\log P(E)$$

  units of information

- *The information is a measure of uncertainty associated with event E -- the less likely is the event E , the more information we receive*
- *For example*
  P(E) = 1 => I(E) = 0 (no information is needed)
  P(E) = 1/2 => I(E) = 1 (one bit is needed)
  P(E) = 1/8 => I(E) = 3 (three bits are needed)

1.4 Information Theory (c.3)

Entropy

$$H(E) = E\{I(E)\} = \sum_{i=1}^{K} P(E) \bullet (-\log P(E))$$

- *The entropy is a measure of expected information across all outcomes of the random vector*
- *The higher entropy is, the more uncertainty it is and thus the more information associated with the sourse is needed*
- *For example, Huffman coding*

1.4 Information Theory (c.4)

Conditional Information

- ***The information received about X=x after we already know the outcome of Y=y***

$$I(X=x|Y=y) = -\log_2 P(X=x|Y=y)$$

Conditional Entropy

- ***The average of conditional information for I(x/y)***

$$H(X|Y) = \xi_{X,Y}\{I(X|Y)\}$$

$$= -\sum_X \sum_Y P(X=x, Y=y)\log_2 P(X=x|Y=y)$$

1.4 Information Theory (c.5)

Mutual Information

- *The shared information in two individual outcome*

$$M(X=x;Y=y)=I(X=x)-I(X=x|Y=y)$$

$$=\log_2\frac{P(X=x,Y=y)}{P(X=x)P(Y=y)}$$

Expected Mutual Information

- *The average mutual information*

$$\overline{M}(X;Y)=H(X)-H(X|Y)=H(Y)-H(Y|X)$$

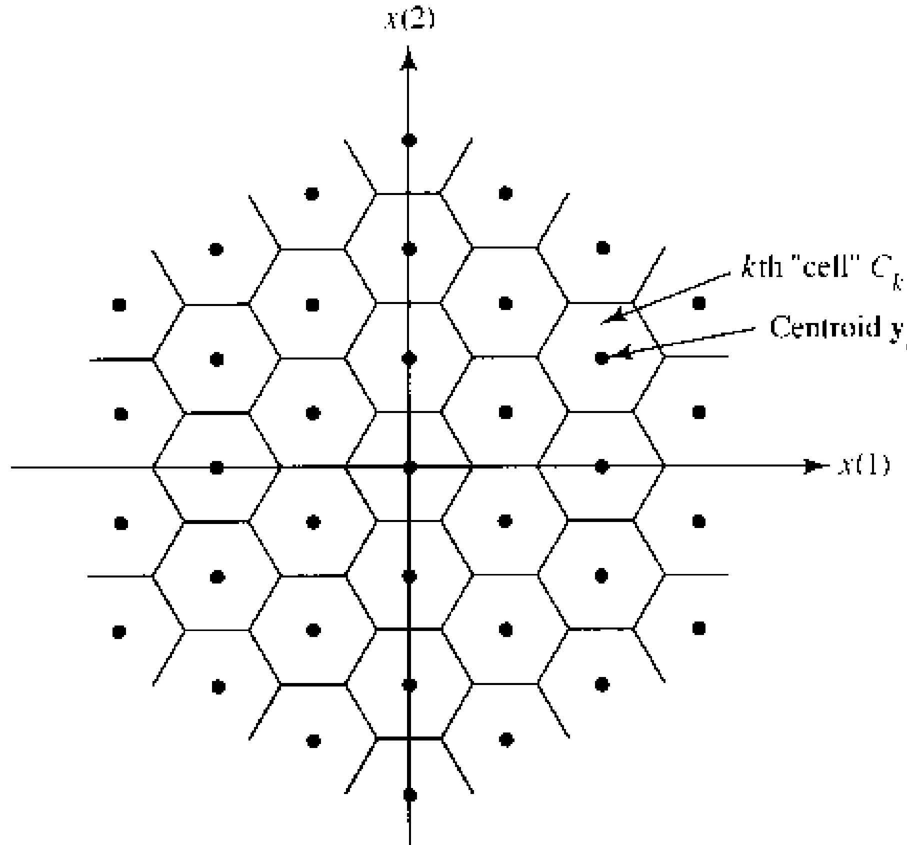
$$=\sum_X\sum_Y P(X=x,Y=y)\log_2\frac{P(X=x,Y=y)}{P(X=x)P(Y=y)}$$

1.5 Concluding Remarks

- **Data Redundancy**
  - Coding Redundancy
  - Spatial/Temporal Redundancy
  - Perceptual Redundancy
- **Compression Models**
  - A General Compression System Model
  - The Source Encoder and Decoder
  - The Channel Encoder and Decoder
- **Information Theory**
  - Information
  - Entropy
  - Conditional Information & Entropy
  - Mutual Information

2. Quantization

- Introduction
- Scalar Quantization
- Vector Quantization



2.1 Introduction

Concepts

- *Coding of Continuous Sources from a theoretical viewpoints.*
- *Quantization of the amplitude results in waveform distortion.*
- *The minimization of this distortion from the viewpoint of quantizer characteristics.*

Two Cases

Scalar quantization

- *the samples are processed "one " at a time*

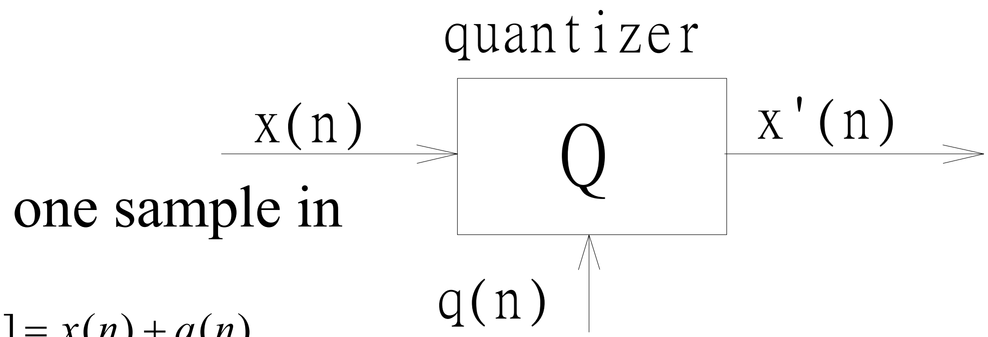
Vector quantization

- *a "block" of samples are quantized as a single entity*

2.2 Scalar Quantization

Quantization Error Optimization (Optimal Quantizer Design)

- ***Quantization Model***



$$\tilde{x}(n) = Q[x(n)] = x(n) + q(n)$$

$$q(n) = \tilde{x}(n) - x(n) = Q[x(n)] - x(n)$$

2.2 Scalar Quantization (c.1)

Optimum Design

- *Select x^(n) (output level) and x(n) (input level) for a particular optimization criterior.*

$$D = E\{h[q(n)]\} = \int_{-\infty}^{\infty} h(\zeta) f_q(\zeta) d\zeta$$

- *The optimizaion is to minimize*

$$D = \int_{-\infty}^{\infty} h[Q(\zeta) - \zeta] f_x(\zeta) d\zeta$$

- *Require the knowledge of the pdf together with the variance of the input signals.*

2.2 Scalar Quantization (c.2)

TABLE 7.7. Optimum Quantizers for Signals with a Gamma Distribution (Paez and Glisson, 1972).

Level	2		4		8		16		32	
	x_i	y_i	x_i	y_i	x_i	y_i	x_i	y_i	x_i	y_i
1	∞	0.577	1.205	0.302	0.504	0.149	0.229	0.072	0.101	0.033
2			∞	2.108	1.401	0.859	0.588	0.386	0.252	0.169
3					2.872	1.944	1.045	0.791	0.429	0.334
4					∞	3.799	1.623	1.300	0.630	0.523
5							2.372	1.945	0.857	0.737
6							3.407	2.798	1.111	0.976
7							5.050	4.015	1.397	1.245
8							∞	6.085	1.720	1.548
9									2.089	1.892
10									2.517	2.287
11									3.022	2.747
12									3.633	3.296
13									4.404	3.970
14									5.444	4.838
15									7.046	6.050
16									∞	8.043
D_{min}	0.6680		0.2326		0.0712		0.0196		0.0052	

2.3 Vector Quantization

Definition

$$\underline{x} = [\underline{x}(1)\ \underline{x}(2)\ \ldots\ldots\ldots\ \underline{x}(N)]$$

$$\underline{y} = [\underline{y}(1)\ \underline{y}(2)\ \ldots\ldots\ldots\ \underline{y}(N)]$$

$\underline{x}(i)$, $\underline{y}(i)$, $1 \le i \le N$: real random variables

$\underline{x}$, $\underline{y}$: N- dimensional random vector

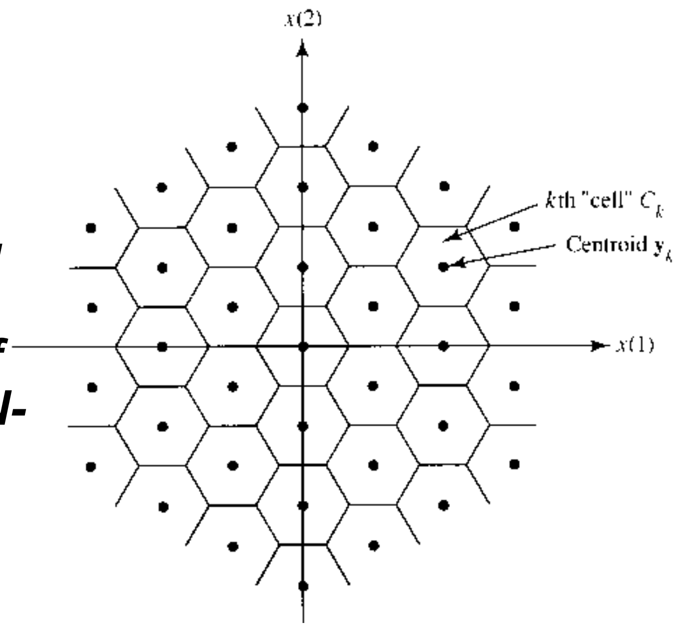
the vector $\underline{y}$ has a special distribution in that it may only take one of L (deterministic) vector values in R^N

2.3 Vector Quantization (c.1)

Vector quantization

$$\underline{y} = Q(\underline{x})$$

- *the vector quantization of x may be viewed as a pattern recognition problem involving the classification of the outcomes of the random variable x into a discrete number of categories or cell in N-space in a way that optimizes some fidelity criterion, such as mean square distortion.*



2.3 Vector Quantization (c.2)

VQ Distortion

$$D = \sum_{k=1}^{L} P(\underline{x} \in C_k) E\{d(\underline{x}, y_k) | \underline{x} \in C_k\}$$

$d(\underline{x}, y_k)$ are typically the distance measures in R^N, including l_1, l_2, l_∞ norm

VQ Optimization

- ***minimize the average distortion D.***

2.3 Vector Quantization (c.3)

Two conditions for optimality

- ***Nearest Neighbor Selection***

$$Q(x) = y_k \quad , \quad x \in C_k$$

$$iff \quad d(x, y_k) \le d(x, y_j) \quad \text{for } k \ne j, 1 \le j \le L$$

- ***minimize average distortion***

$$y_k = \arg\min_y D_k == \arg\min_y E\{d(x,y) | x \in C_k\}$$

$$= \arg\min_y \int \ldots \int_{x \in C_k} d(x,y) f_x(\xi_1 \ldots \xi_n) d\xi_1 \ldots d\xi_N$$

=> applied to partition the N-dimensional space into cell $\{C_k, 1 \le k \le L\}$ when the joint pdf $f_x(\bullet)$ is known.

3 Rate-Distortion Functions

- Introduction
- Rate-Distortion Function for a Gaussian Source
- Rate-Distortion Bounds
- Distortion Measure Methods

3.1 Introduction

Considering question

- *Given a source-user pair and a channel, under what conditions is it possible to design a communication system that reproduces the source output for the user with an average distortion that does not exceed some specified upper limit D?*
  - *The capacity (C) of a communication channel.*
  - *The rate distortion function (R(D))of a source-user pair.*

Rate-distortion function R(D)

- *A communication system can be designed that achieves fidelity D if and only if the capacity of the channel that connects the source to user exceeds R(D).*
- *The lower limit for data compression to achieve a certain fidelity subject to a predetermined distortion measure D.*

3.1 Introduction (cont.)

- ***Equations representations :***

Distortion D:

$$D = d(q) = \iint p(x)q(y|x)\rho(x,y)\,dx\,dy$$

Mutual information:

$$I(q) = \iint p(x)q(y|x)\log\frac{q(y|x)}{q(y)}\,dx\,dy$$

Rate distortion function R(D):

$$R(D) = \inf_{q \in Q_D} I(q), \qquad Q_d = \{q(y|x) : d(q) = D\}$$

$\rho(x,y)$: distortion measure for the source word

$\mathbf{x} = (x_1,\dots,x_n)$ reproduced as $\mathbf{y} = (\mathrm{y}_1,\dots,y_n)$

$$\rho_{\mathrm{n}}(\mathbf{x},\mathbf{y}) = n^{-1}\sum_{t=1}^{n}\rho(x_t,y_t)$$

The family $F_\rho = \{\rho_n, 1 \le n < \infty\}$ is called the single - letter fidelity criterion generated by ρ.

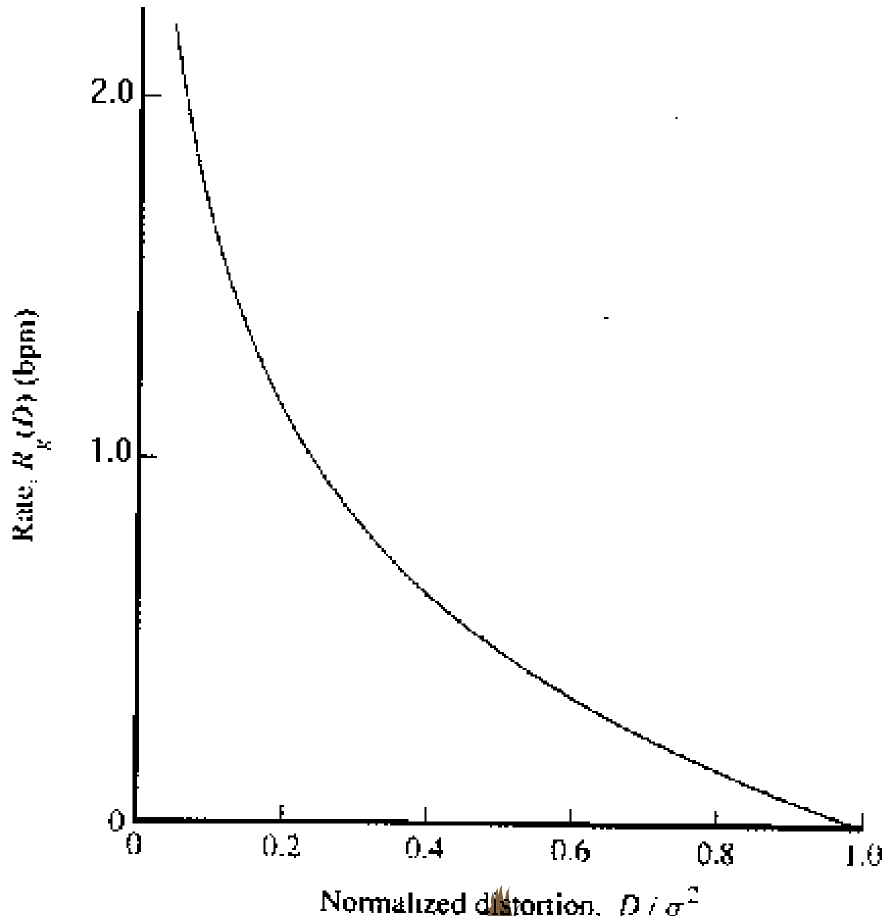
3.2 Rate-Distortion Bounds

- ***Introduction***
- ***Rate-Distortion Function for A Gaussian Source***
  - ***R(D) for a memoryless Gaussian source***
  - ***Source coding with a distortion measure***
- ***Rate-Distortion Bounds***
- ***Conclusions***

3.3 Rate-Distortion Function for A Gaussian Source

- **Rate-Distortion for a memoryless Gaussian source**
  - ***The minimum information rate (bpn) necessary to represent the output of a discrete-time, continuous-amplitude, memoryless stationary Gaussian source based on an MSE distortion measure per symbol.***
  - ***Equation***

$$R_g(D) = \begin{cases} \frac{1}{2}\log_2(\sigma_x^2 / D), & 0 \le D \le \sigma_x^2 \\ 0, & D \ge \sigma_x^2 \end{cases}$$



3.3 Rate-Distortion Function for A Gaussian Source (c.1)

- ***Source coding with a distortion measure (Shannon, 1959)***
  - ***There exists a coding scheme that maps the source output into codewords such that for any given distortion D, the minimum rate R(D) bpn is sufficient to reconstruct the source output with an average distortion that is arbitrarily close to D.***
  - ***Transform the R(D) to distortion-rate function D(R)***

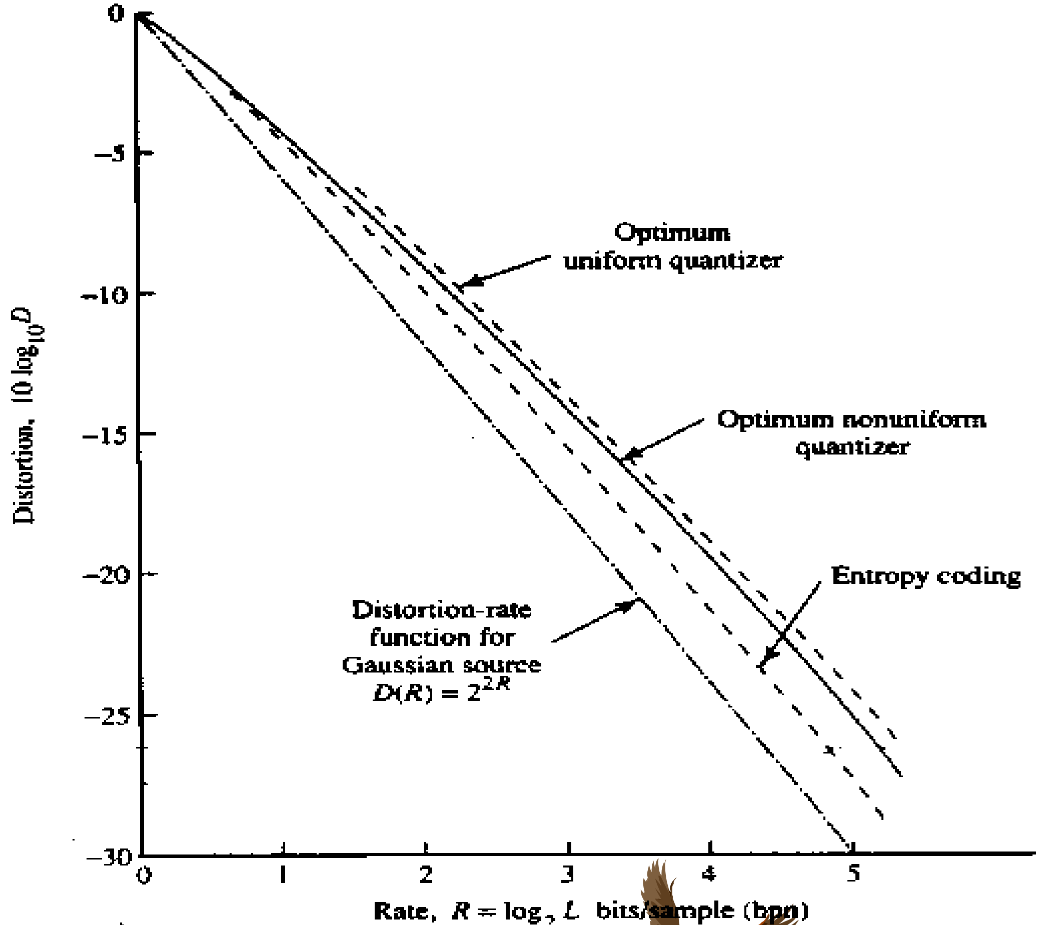
$$D_g(R) = 2^{-2R}\sigma_{\underline{x}}^2$$

Express in dB

$$10\log_{10} D_g(R) = -6R + 10\log_{10}\sigma_{\underline{x}}^2$$

3.3 Rate-Distortion Function for A Gaussian Source (c.2)

- Comparison between different quantizations



3.4 Rate-Distortion Bounds

Source:

- **_Memoryless, continuous-amplitude source with zero mean and finite variance $\sigma_{\underline{x}}^2$ with respect to the MSE distortion measure._**

Upper bound

- **_According to the theorem of Berger (1971), it implies that the Gaussian source requires the maximum rate among all other sources for a specified level of mean square distortion._**

$$R(D) \le \frac{1}{2}\log_2 \frac{\sigma_{\underline{x}}^2}{D} = R_g(D), \qquad 0 \le D \le \sigma_{\underline{x}}^2$$

$$D(R) \le D_g(R) = 2^{-2R}\,\sigma_{\underline{x}}^2$$

3.4 Rate-Distortion Bounds (c.1)

- **Lower bound**

$$R^*(D) = H(\underline{x}) - \frac{1}{2}\log_2 2\pi e D$$

$$D^*(R) = \frac{1}{2\pi e} 2^{-2|R - H(\underline{x})|}$$

$H(\underline{x})$: differential entropy

$$H(\underline{x}) \overset{def}{=} -\int_{-\infty}^{\infty} f_{\underline{x}(n)}(\xi) \log_2 f_{\underline{x}(n)}(\xi) d\xi$$

For Gaussian source:

$$f_{\underline{x}}(x) = \frac{1}{\sqrt{2\pi}\sigma_{\underline{x}}} e^{-x^2/2\sigma_{\underline{x}}^2}$$

$$H_g(\underline{x}) = \frac{1}{2}\log_2 2\pi e \sigma_{\underline{x}}^2$$

$$\Rightarrow R^*(D) = \frac{1}{2}\log_2 2\pi e \sigma_{\underline{x}}^2 - \frac{1}{2}\log_2 2\pi e D = \frac{1}{2}\log_2 \frac{\sigma_{\underline{x}}^2}{D}$$

3.4 Rate-Distortion Bounds (c.2)

- ***For Gaussian source, the rate-distortion, upper bound and lower bound are all identical to each other.***
- ***The bound of differential entropy***

$$10\log_{10} D^*(R) = -6R - 6[H_g(x) - H(x)]$$

$$10\log_{10} \frac{D_g(R)}{D^*(R)} = 6[H_g(x) - H(x)]$$

$$= 6[R_g(D) - R^*(D)]$$

$\Rightarrow$ The differential entropy is upper bounded by $H_g(x)$

3.4 Rate-Distortion Bounds (c.3)

Rate-distortion R(D) to channel capacity C

- For $C \geq R_g(D)$
  - The fidelity (D) can be achieved.
- For $R(D) \leq C < R_g(D)$
  - Achieve fidelity for stationary source
  - May not achieve fidelity for random source
- For $C < R(D)$
  - Can not be sure to achieve fidelity

3.4 Rate-Distortion Bounds (c.4)

TABLE 7.6. Differential Entropies and Rate-Distortion Comparisons of Four Common pdf's for Signal Models.

pdf	$f_{\underline{x}}(x)$	$H(\underline{x})$	$R_g(D) - R^*(D)$ (bpn)	$D_g(R) - D^*(R)$ (dB)
Gaussian	$\frac{1}{\sqrt{2\pi}\sigma_{\underline{x}}} e^{-x^2/2\sigma_{\underline{x}}^2}$	$\frac{1}{2}\log_2\left(2\pi e\sigma_{\underline{x}}^2\right)$	0	0
Uniform	$\frac{1}{2\sqrt{3}\sigma_{\underline{x}}},\ \lvert x\rvert \le \sqrt{3}\sigma_{\underline{x}}$	$\frac{1}{2}\log_2\left(12\sigma_{\underline{x}}^2\right)$	0.255	1.53
Laplacian	$\frac{1}{\sqrt{2}\sigma_{\underline{x}}} e^{-\sqrt{2}\lvert x\rvert/\sigma_{\underline{x}}}$	$\frac{1}{2}\log_2\left(2e^2\sigma_{\underline{x}}^2\right)$	0.104	0.62
Gamma	$\frac{\sqrt[4]{3}}{\sqrt{8\pi\sigma_{\underline{x}}\lvert x\rvert}} e^{-\sqrt{3}\lvert x\rvert/2\sigma_{\underline{x}}}$	$\frac{1}{2}\log_2\left(4\pi e^{0.423}\sigma_{\underline{x}}^2/3\right)$	0.709	4.25

3.5 Distortion Measure Methods

$$GG(m,\sigma,r) = k \exp\left\{-\left|c(x-m)\right|^{r}\right\}$$

$$k = \frac{rc}{2\Gamma(1/r)} \quad \text{and } c = \sqrt{\frac{\Gamma(3/r)}{\sigma^{2}\Gamma(1/r)}}$$

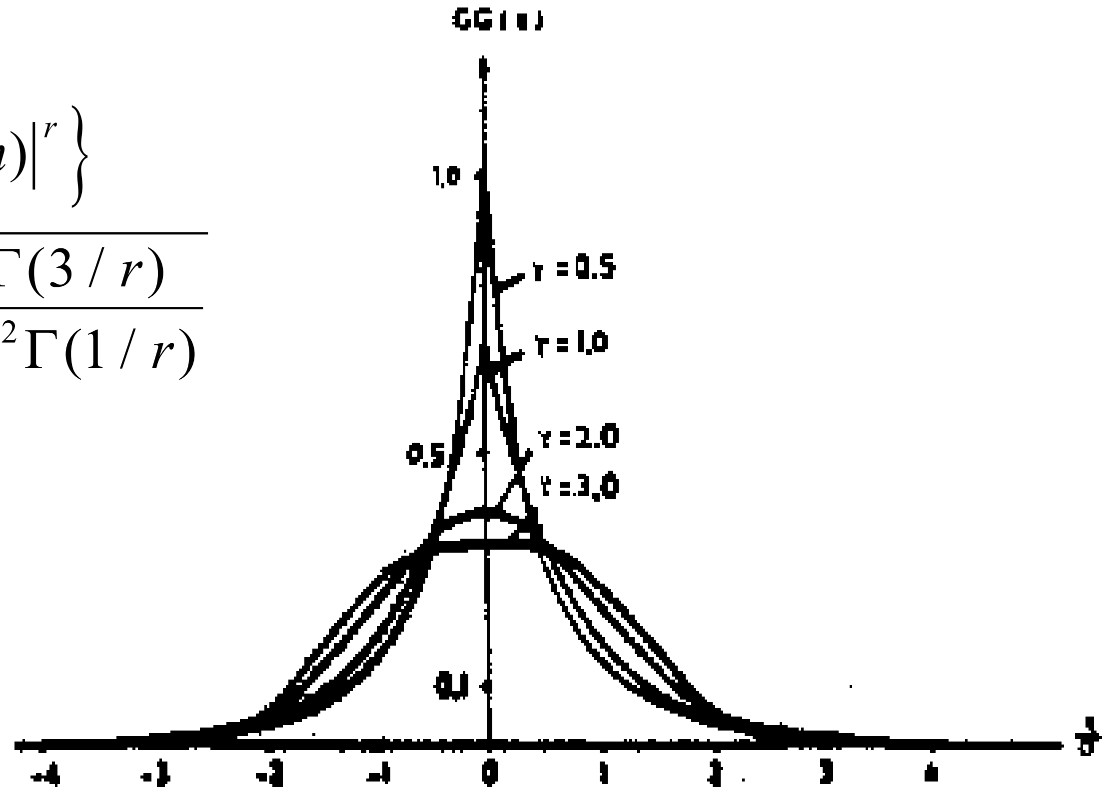
For different r:

$r = 1 \rightarrow$ Laplacian pdf

$r = 2 \rightarrow$ Gaussian pdf

$r = 0 \rightarrow$ constant pdf

$r = \infty \rightarrow$ uniform pdf



2. Generalized-Gaussian probability density function for differen parameter r.

4. Waveform Coding

- Introduction
- Pulse Code Modulation(PCM)
- Log-PCM
- Differential PCM
- Adaptive DPCM

4.1 Introduction

Two SCoding Categories

1 . Waveform coder

2 . Perceptual coder

Waveform Coding

- Methods for digitally representing the temporal or spectral characteristics of waveforms.

Vocoders

- Parametric Coders, the parameters characterize the short-term spectrum of a sound.
- These parameters specify a mathematical model of human speech production suited to a particular sound.

4.2 Pulse Code Modulation

The Quantized Waveform $\hat{s}(n)$

$s(n) = \hat{s}(n) + q(n)$

Applying uniform quantizer

- The quantization noise can be modeled by a stationary random process $\underline{q}$ in which each of the random variables $\underline{q}(n)$ has the uniform pdf.

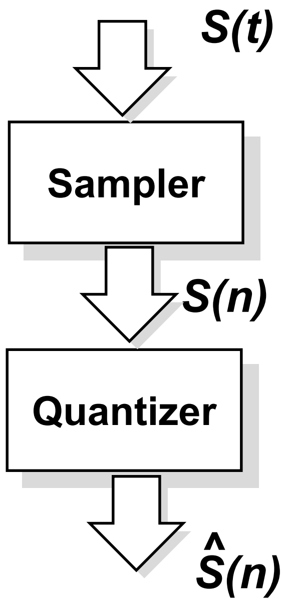
$$f_{\underline{q(n)}}(\xi) = \frac{1}{\Delta}, \quad -\frac{\Delta}{2} \le \zeta \le \frac{\Delta}{2}$$

The step size is 2^{-R}. The mean square value is

$$\xi\{\underline{q}^2(n)\} = \frac{\Delta^2}{12} = \frac{2^{-2R}}{12}$$

Measured in decibels

$$10\log_{10}\frac{\Delta^2}{12} = 10\log_{10}\frac{2^{-2R}}{12} = -6R - 10.79 \quad dB$$



4.3 Log PCM

Concepts

- *Small-signal amplitudes occur more frequently than large-signal amplitudes in speech signals*
- *Human hearing exhibits a logarithmic sensitivity*

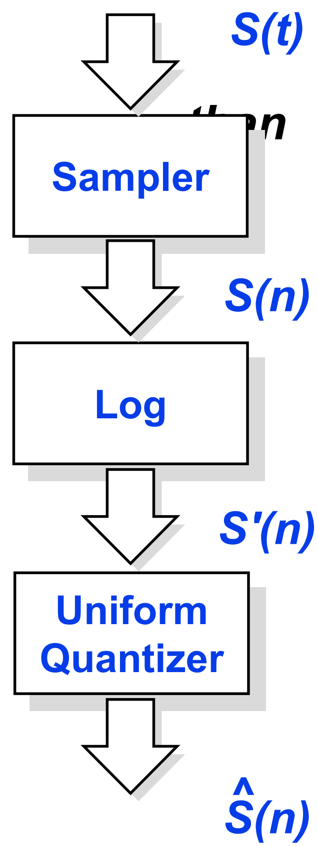
Two Nonuniform quantizer

- *u-law (a standard in the United States and Canada)*

$$|y| = \frac{\log(1+\mu|s|)}{\log(1+\mu)}$$

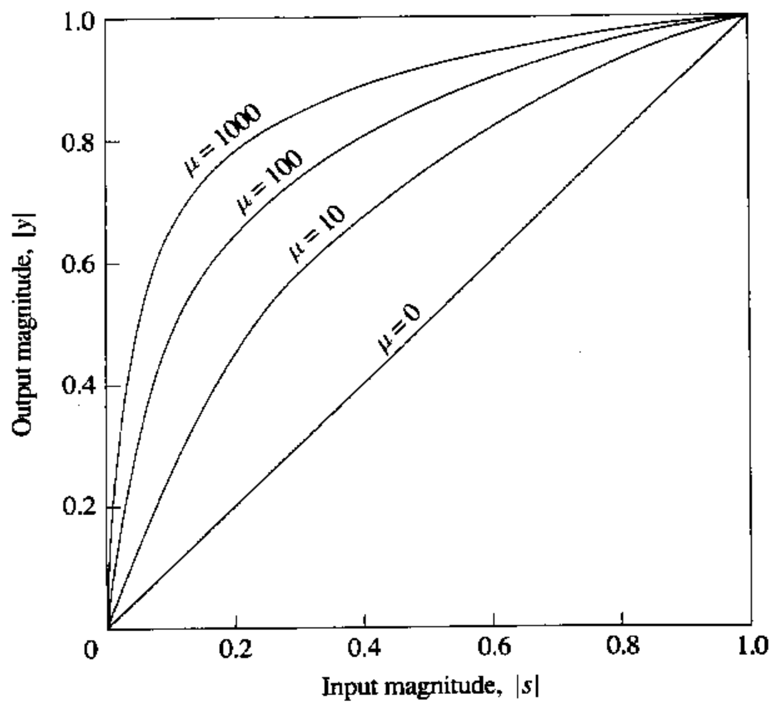
- *A-law (European standard)*

$$|y| = \frac{\log A|s|}{1+\log A}$$

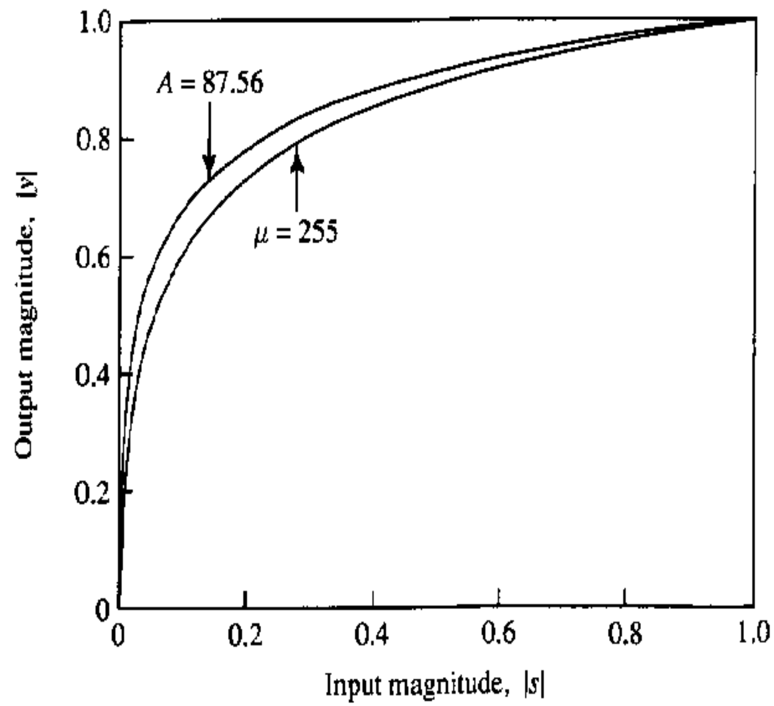


4.3 Log PCM(c.1)

- ***Input-Output Magnitude Characteristic of u-Law***



- ***Two Compression Functions***



4.4 Differential PCM (DPCM)

Concepts

- *In PCM , each sample is coded independently of all the other samples.*
- *The average changes in amplitude between samples are very small.*
  ==> Temporal Redundancy.

Approach

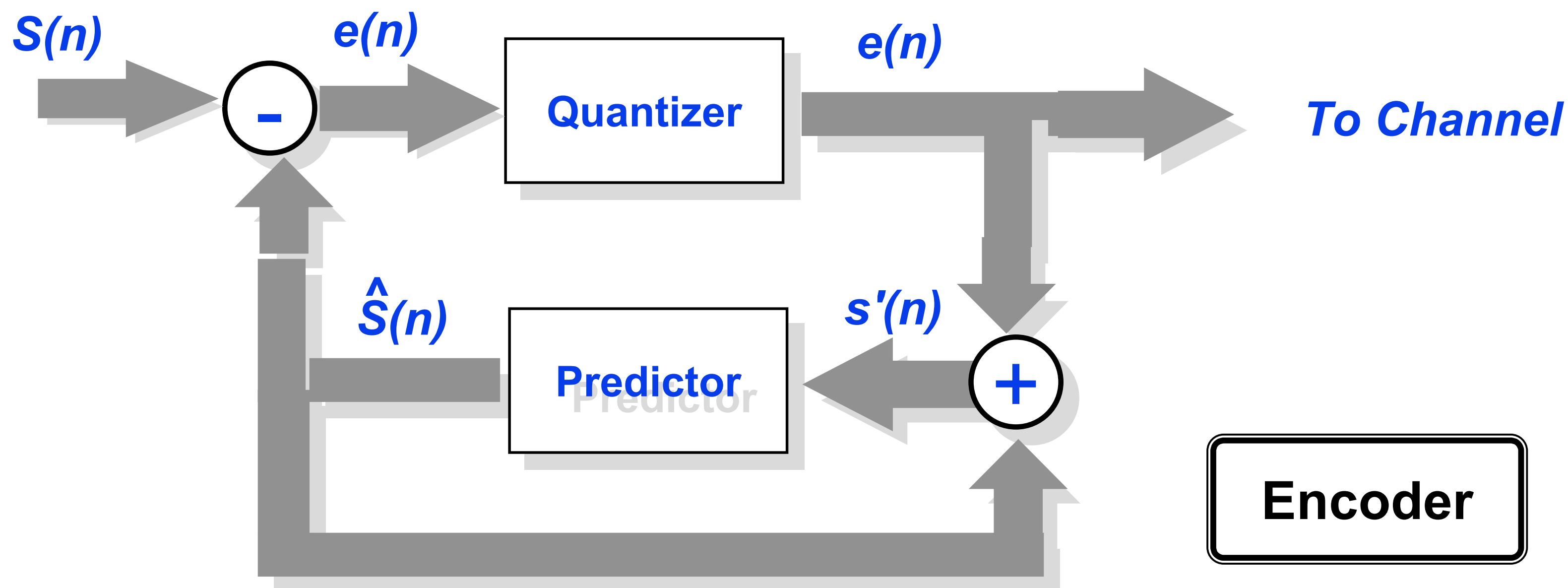
- *Encode the differenced sequence*

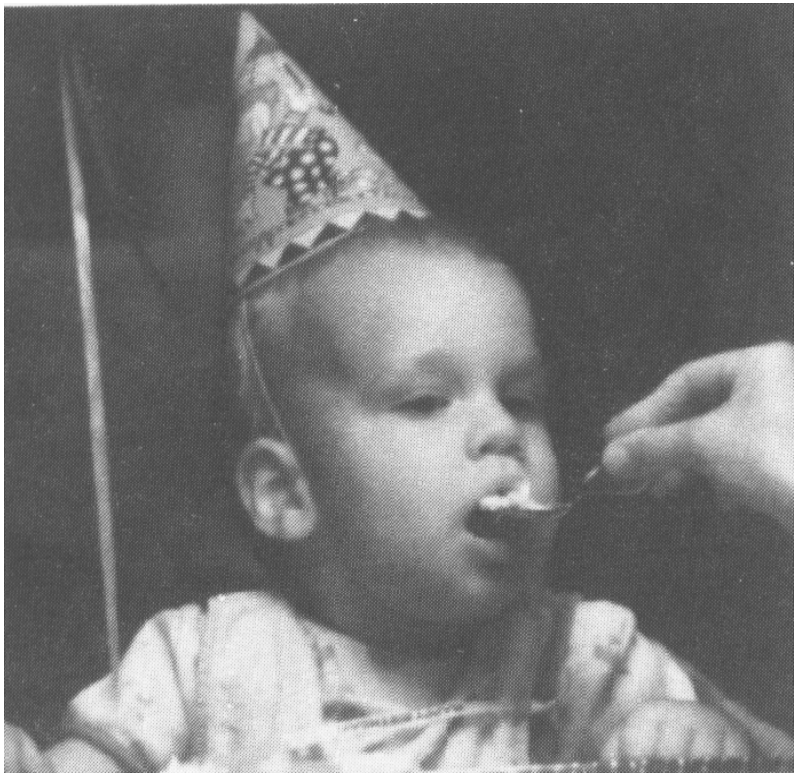
  ex. $e(n) = s(n) - s(n-1)$

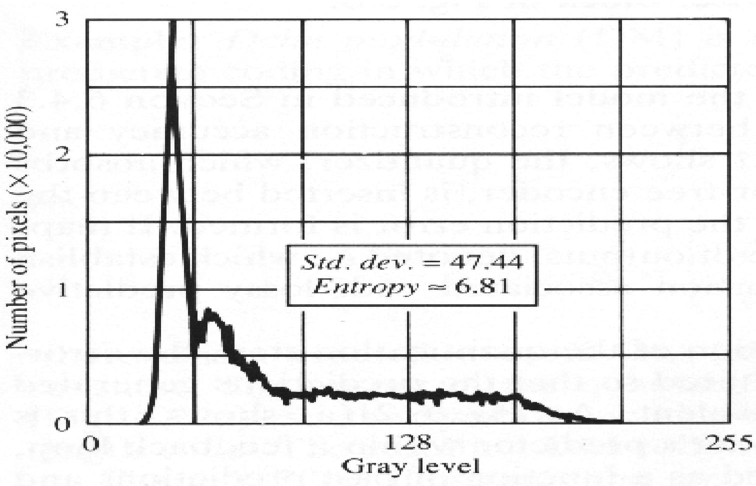
  ex. Typical predictor $$A(z) = \sum_{i=1}^{p} a_i z^{-i}$$
- *Fewer bits are required to represent the differences*
- *PCM & DPCM encoders are designed on the basis that the source output is stationary*
- *DPCM performs better than PCM at and below 32 kbits/s*

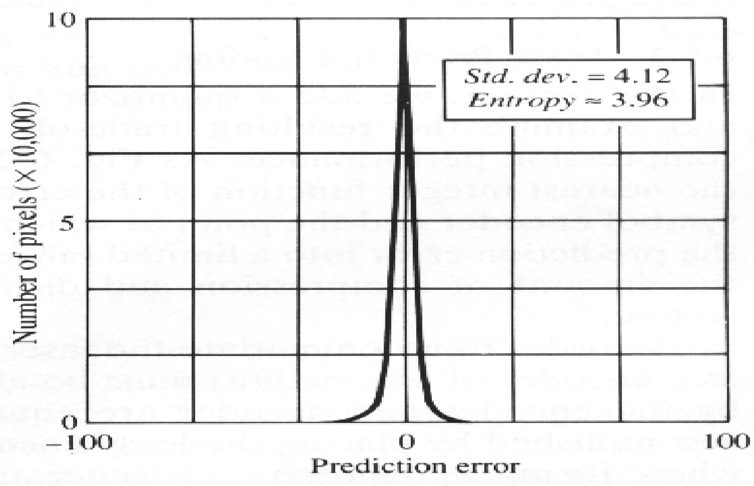
4.4 Differential PCM (c.1)

Block Diagram of a DPCM






3
2
1
0
Number of pixels (×10,000)
Std. dev. = 47.44
Entropy ≈ 6.81
0
128
255
Gray level


10
5
0
Number of pixels (×10,000)
Std. dev. = 4.12
Entropy ≈ 3.96
−100
0
100
Prediction error

4.5 Adaptive DPCM

Considerations

- *Speech signals are quasi-stationary in natual*

Concepts

- *Adapt to the slowly time-variant statistics of the speech signal*
- *Adaptive quantizer is used*
- *Feedforward and feedback adaptive quantizer.*

Example

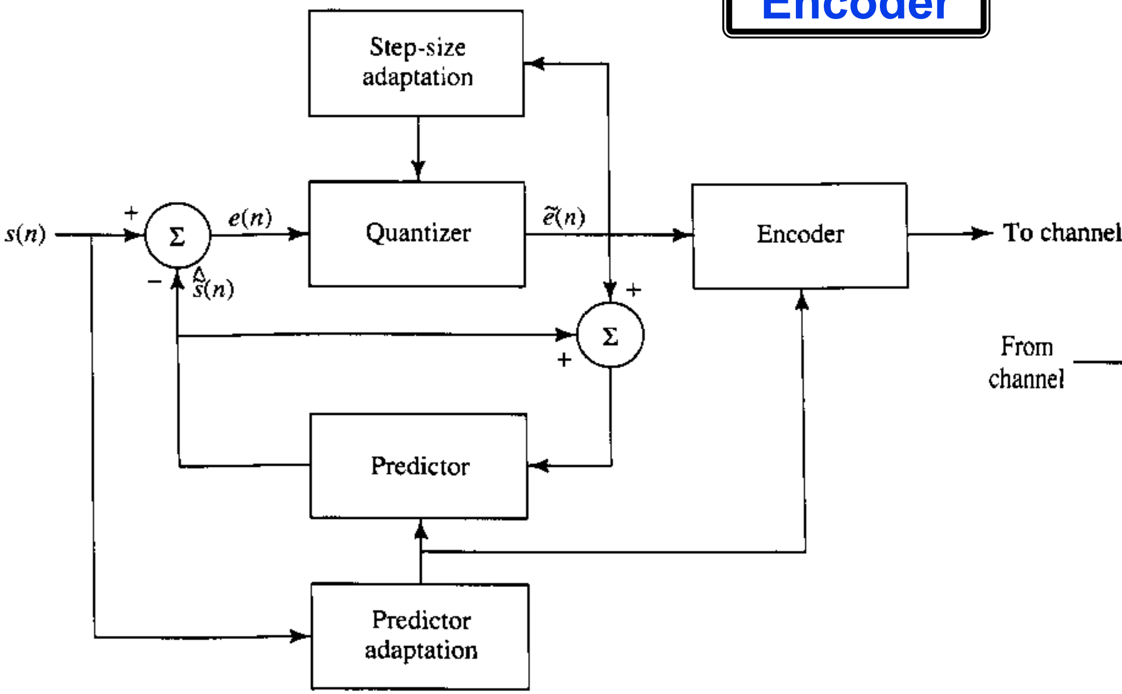
- *looks at only one previosly quantized sample and either expands or compresses the quantizer intervals.*

$$\Delta_{n+1} = \Delta_n M_i(|\hat{x}|)$$

4.5 Adaptive DPCM(c.1)

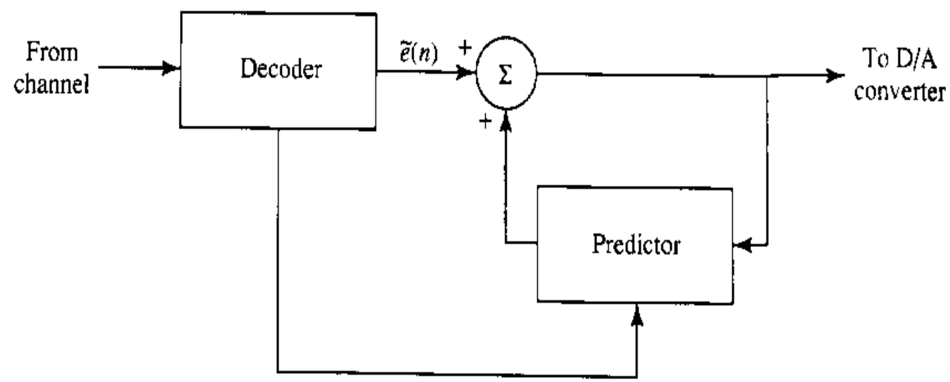
Block Diagram

Encoder



(a) Encoder

Decoder



(b) Decoder

4.5 Adaptive DPCM(c.2)

Adaptive Step Sizes

$$\Delta_{n+1} = \Delta_n M_i(|\hat{x}|)$$

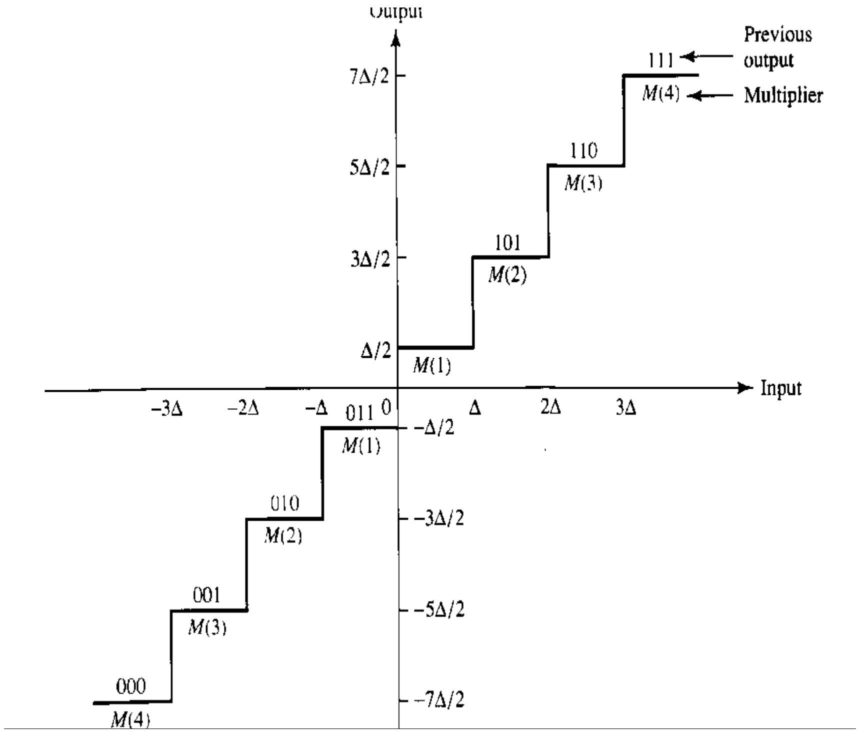


TABLE 7.8. Multiplication Factors for Adaptive Step Size Adjustment (Jayant, 1974).

	PCM			DPCM		
	2	3	4	2	3	4
$M(1)$	0.60	0.85	0.80	0.80	0.90	0.90
$M(2)$	2.20	1.00	0.80	1.60	0.90	0.90
$M(3)$		1.00	0.80		1.25	0.90
$M(4)$		1.50	0.80		1.70	0.90
$M(5)$			1.20			1.20
$M(6)$			1.60			1.60
$M(7)$			2.00			2.00
$M(8)$			2.40			2.40

4.5 Adaptive DPCM(c.3)

CCITT G.721 standard (1988)

- *Adaptive quantizer*
  - *quantize e(n) into 4 bits words.*
- *Adaptive predictor*
  - *Pole-zero predictor with 2 poles, 6 zeros.*
  - *Coefficients are estimated using a gradient algorithm and the stability is checked by testing two roots of A(z).*
- *The performance of the coder in terms of MOS is above 4.*
- *The G.721 ADPCM algorithm was modified to accomodate 24 and 40 kbits/s in G.723.*
- *The performance of ADPCM degrades quickly for rates below 24 kbits/s.*

4.6 Summary

- Introduction
- Pulse Code Modulation(PCM)
- Log-PCM
- Differential PCM
- Adaptive DPCM

5. Subband Coding

Concepts

- *Exploits the redundancy of the signal in the frequency domain.*
- *Quadrature-Mirror Filter for subband coding.*
- *The opportunitylies in both the short-time power spectrum and the the perceptual properties of the human ear.*

Standards

- *AT&T voice store-and-forward standard.*
  - *16 or 24 kbits/s*
  - *Five-band nonuniform tree-structured QMF band in conjunction with ADPCM coders*
  - *The frequecy range for each band are 0-0.5, 0.5-1, 1-2, 2-3, 3-4 kHz.*
  - *{4/4/2/2/0} for 16 kbits and {5/5/4/3/0} for 24 kbits.*
  - *The one-way delay is less than 18 ms.*

5. Subband Coding (c.1)

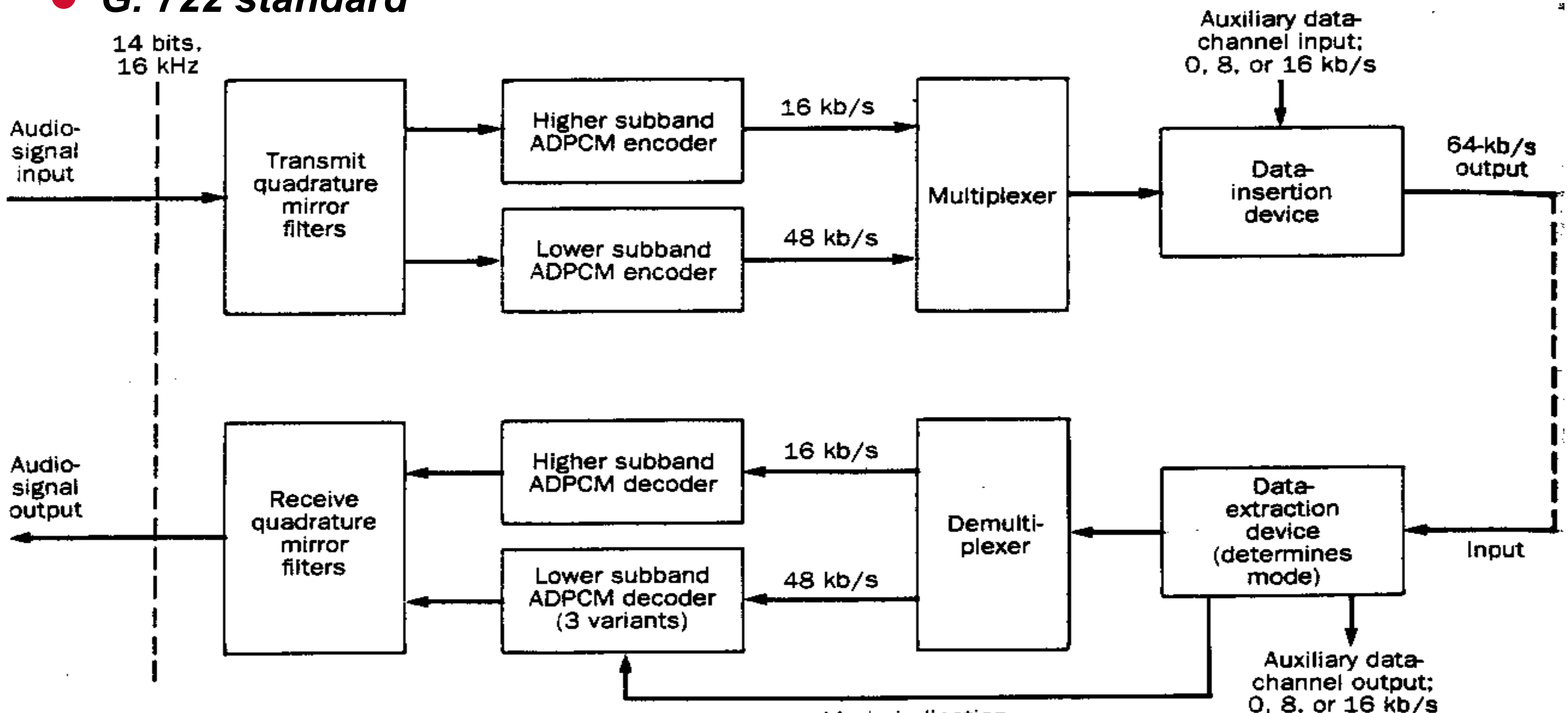
CCITT G.722

- G. 722 algorithm at 64 kb/s have an equivalent SNR gain of 13 db over the G.721.
- Low-frequency parts permit operation at 6, 5, or 4 bits (64, 56, and 48 kb/s) per sample with graceful degradation of quality.
- Two-band subband coder with ADPCM coding of each subband.
- The low- and high-frequency subbands are quantized using 6 and 2 bits per sample, respectively.
- The filter banks produce a communication delay of about 3 ms.
- The MOS at 64 kbits/s is greater than 4 for music signals.

5. Subband Coding (c.2)

Two-Band Subband Coder for 64-kb/s coding of 7-kHz Audio

- *G. 722 standard*



6. Transform Coding

Concepts

- The transform components of a unitary transform are quantized at the transmitter and decoded and inver-transformed at receiver.

Unitary transforms

- Karhunen-Loeve Transform
  - Optimal in the sense that the transform components are maximally decorrelated for any given signal.
  - Data dependent.
- The Discrete Cosine Transform
  - Near optimal.
- The Fast Fourier Transform
  - Approaches that of DCT for very large block.