Source Coding (Data Compression)

Lossless Data Compression:

- Huffman Coding, Lempel-Ziv Algorithm (zip), Run-length Coding,

- Lossy Data Compression:

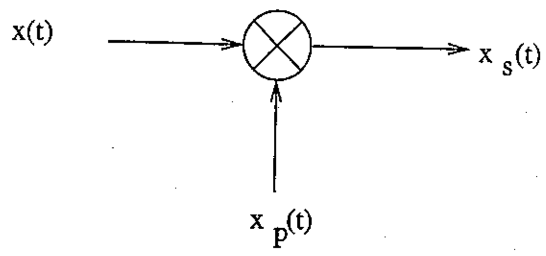
- PCM, DPCM, ADPCM, LPC, CELP, etc for Voice.

- JPEG for Image, H261 and MPEG for Video, etc.

Source coding in general consists of:

1- Sampling,
2- Quantization,
3- Coding (Labeling the quantization levels).

Sampling



Sampling Process

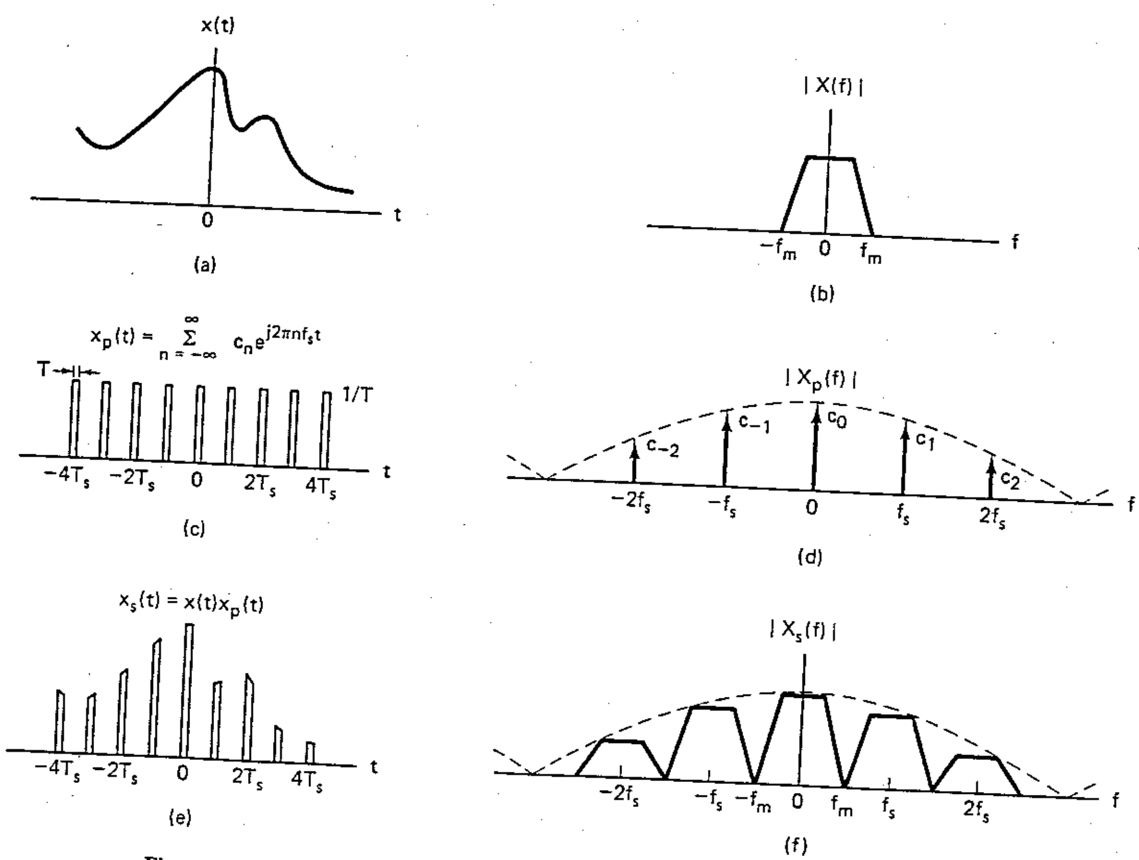
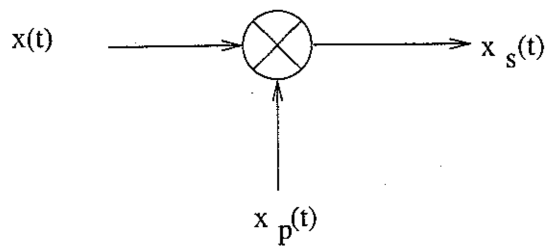


Figure Sampling theorem using the frequency shifting property of the Fourier transform.

Sampling



Sampling Process

$$x_s(t) = x(t)x_p(t)$$

where $x_p(t)$ is a periodic pulse and, as such, it has a Fourier Series representation:

$$x_p(t) = \sum_{n=-\infty}^{\infty} c_n e^{j2\pi n f_s t},$$

where $f_s = \frac{1}{T_s}$ is the sampling frequency and

$$c_n = \frac{1}{T_s} \frac{Sin(\frac{n\pi T}{T_s})}{\frac{n\pi T}{T_s}} = \frac{1}{T_s} Sinc(\frac{nT}{T_s}),$$

T is the pulse width.
Based on the above, we have

$$x_s(t) = x(t) \sum_{n=-\infty}^{\infty} c_n e^{j2\pi n f_s t}.$$

In the frequency domain:

$$X_s(f) = F\left\{x(t) \sum_{n=-\infty}^{\infty} c_n e^{j2\pi n f_s t}\right\} = \sum_{n=-\infty}^{\infty} c_n F\left\{x(t) e^{j2\pi n f_s t}\right\}$$

or

$$X_s(f) = \sum_{n=-\infty}^{\infty} c_n X(f - n f_s).$$

Sampling
Nyquist Theorem

In order to be able to recover the original signal from its samples, the sampling frequency should be at least twice the signal bandwidth:

$$f_s \geq 2W.$$

or equivalently, the time between two consecutive samples be less than $\frac{1}{2W}$.

$$T_s \leq \frac{1}{2W}.$$

Under sampling (sampling at a rate less than 2W) results in aliasing.

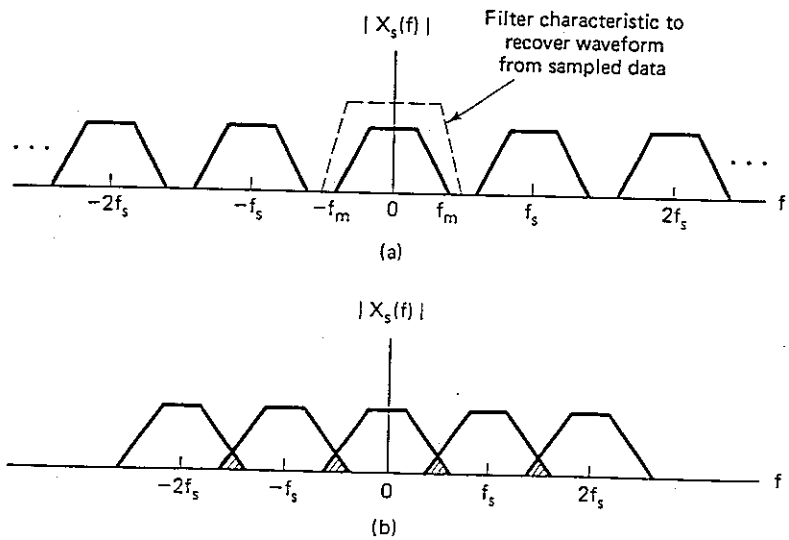


Figure. Spectra for various sampling rates. (a) Sampled spectrum ($f_s > 2f_m$). (b) Sampled spectrum ($f_s < 2f_m$).

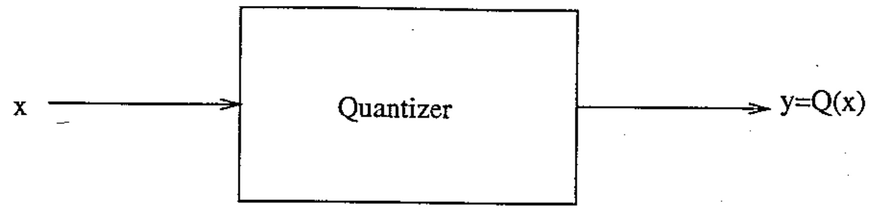
Example: For speech the bandwidth is from 0 to 3400 Hz. So, the minimum required sampling frequency is 6800 Hz. (6.8 kHz.). Usually, a sampling rate of 8 k samples/second is used. The, $T_s = 125\,\mu s$.

Example: For audio (music), the bandwidth extends up to 15 kHz. (or even 20 kHz.) and a sampling rate of 44 ksamples/sec. is used.

Quantization

Performance of Uniform Quantizer

Denote the input and output of the quantizer by x and y, respectively.



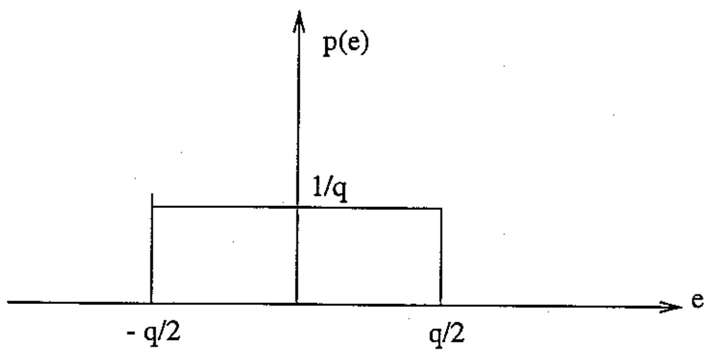
The quantization error (or distortion) e is defined as the difference between x and y,

$$e = y - x.$$

It is clear that,

$$-\frac{q}{2} \leq e \leq \frac{q}{2}.$$

For a uniform source, e has a uniform probability density function,



Quantization error variance (power) is:

$$\sigma_q^2 = \int_{-q/2}^{q/2} e^2 p(e) de = \int_{-q/2}^{q/2} e^2 \frac{1}{q} de = \frac{q^2}{12}.$$

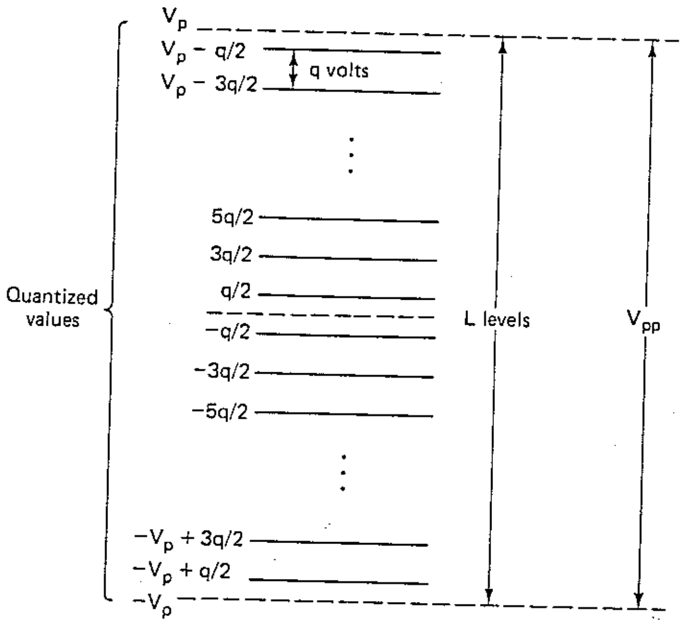
Quantization

Assume that the dynamic range of the signal to be quantized (the difference between the maximum and minimum value of the signal) is V_{pp}. Assume that we have an R bit quantizer. That is, we have to represent each sample of the signal with R. Then, we can have $L = 2^R$ levels. For example, an 8-bit Analog-to-Digital Converter (ADC) encodes each sample into one of 256 values.

Let's first consider a uniform quantizer.A uniform quantizer is only optimum if the source to be quantized is uniform (that is, its samples takes values in a given range with equal probability). In a uniform quantizer, the dynamic range of the source (V_{pp}) is divided into $L = 2^R$ equal segments and any source sample falling into a given segment is encoded (approximated) as the point in the middle of that segment.

The size of each segment (the step size) is,

$$q = \frac{V_{pp}}{L} = \frac{V_{pp}}{2^R}.$$



Quantization levels.

Quantization
Performance of Uniform Quantizer

Denoting the rms value of the signal as V_{rms}, the Signal-to-Quantization Noise Ratio is defined as:

$$(\frac{S}{N})_q = \frac{V_{rms}^2}{\sigma_q^2} = \frac{12V_{rms}^2}{q^2}.$$

Example: For a sinusoidal signal $Acos(2\pi f_c t)$, we have $V_{pp} = 2A$ and $V_{rms} = A/\sqrt{2}$. So, $q = \frac{2A}{L} = \frac{2A}{2^R}$and,

$$(\frac{S}{N})_q = \frac{12V_{rms}^2}{q^2} = \frac{3}{2} \times 2^{2R}.$$

or in dB:

$$SQNR = 10log(\frac{S}{N})_q = 6R + 1.7\,dB.$$

Each additional bit adds 6 dB to SQNR.

Example: For a random signal (one for which the dynamic range is not known exactly), the dynamic range is taken as a multiple of standard deviation. For example if the expected rms value of the signal is believed to be σ (the variance being σ^2), we may take the dynamic range as $\pm 4\sigma$ (assuming a Gaussian distribution, this ensures that more than 99% of the samples fall inside the dynamic range). In this case $V_{pp} = 8\sigma$ and $q = \frac{8\sigma}{2^R}$ and therefore,

$$(\frac{S}{N})_q = \frac{12\sigma^2}{q^2} = \frac{12\sigma^2}{\frac{64\sigma^2}{2^{2R}}} = \frac{3}{16}2^{2R}.$$

In decibels:

$$SQNR = 10log(\frac{S}{N})_q = 6R - 7.27\,dB.$$

Nonuniform Quantization

The problem with uniform quantizer is that it treats the source as if it were uniformly distributed, i.e., all levels had equal probability of occurring. But, in most cases of interest, e.g., speech, image, there are more low intensity components than high intensity samples.

In such cases, most of the levels are wasted (do not contribute to the improvement of SQNR). For this reason, non-uniform quantizer is used in such cases.

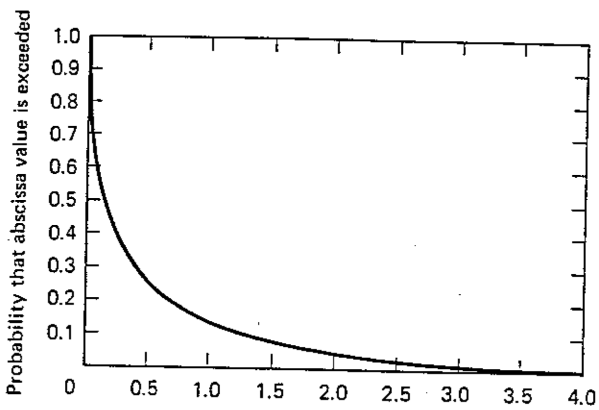
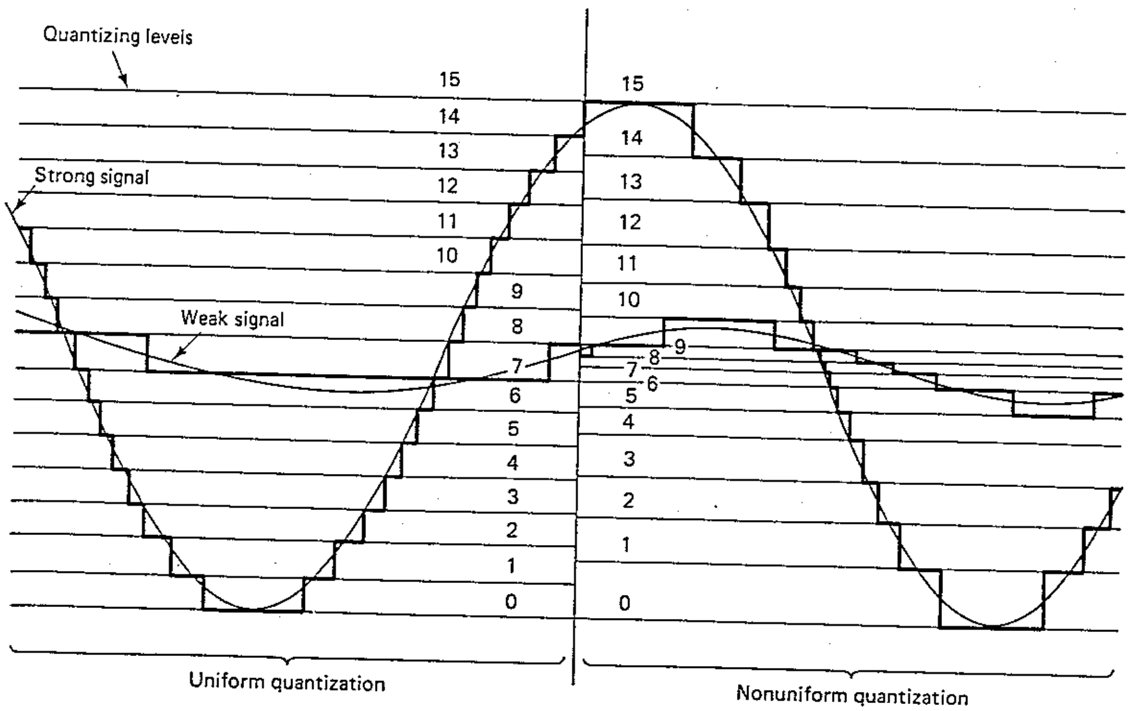


Figure Statistical distribution of single-talker speech signal magnitudes.



Uniform and nonuniform quantization of signals.

Nonuniform Quantization
Companding

Nonuniform quantizer can either be implemented directly by having a quantizer with variable step size or can be implemented using a uniform quantizer with a mapping that turns a non-uniform source into a uniform one by distorting it (COMpressing it). This compressed signal is then quantized using the uniform quantizer. The effect of compressing is then removed by performing the inverse of the compression operation (this is called exPANDing). The overall procedure is called companding.

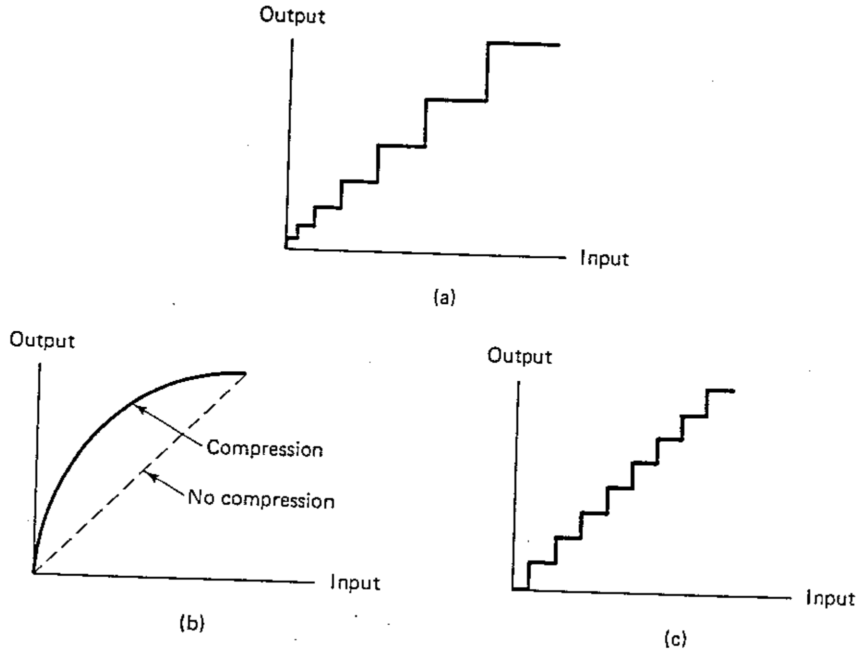


Figure (a) Nonuniform quantizer characteristic. (b) Compression characteristic. (c) Uniform quantizer characteristic.

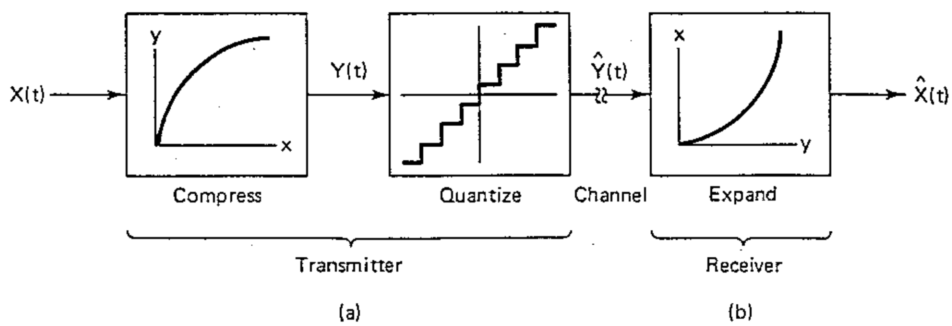


Figure Nonuniform quantizer as a sequence of compression, uniform quantization, and expansion.

Pulse Code Modulation (PCM)
Different PCM Standards

There are two standards for nonuniform quantizer companding, viz., $\mu -$ law for North America and A-law for Europe.

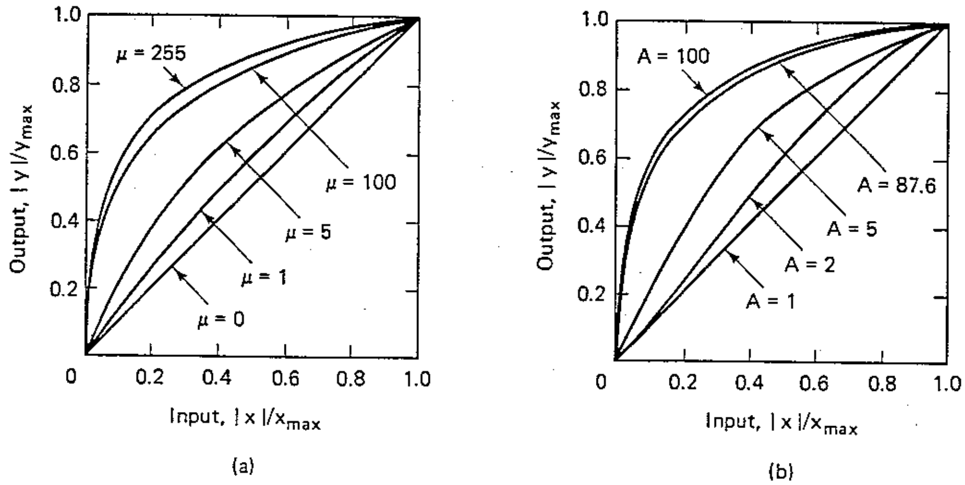


Figure Compression characteristics. (a) μ-law characteristic. (b) A-law characteristic.

μ-Law Compander. The μ-law compander is the Bell System (hence the North American standard) compression law. It is of the form

$$y = C(x) = y_{\max} \frac{\log_e [1 + \mu(|x| / x_{\max})]}{\log_e (1 + \mu)} \operatorname{sgn} x$$

The approximate behavior of this compressor in the regions corresponding to small and large values of the argument are

$$y = C(x) = \begin{cases} y_{\max} \dfrac{\mu(|x| / x_{\max})}{\log_e \mu} & \mu\left(\dfrac{|x|}{x_{\max}}\right) \ll 1 \\ y_{\max} \dfrac{\log_e [\mu(|x| / x_{\max})]}{\log_e \mu} & \mu\left(\dfrac{|x|}{x_{\max}}\right) \gg 1 \end{cases}$$

A-Law Compander. The A-law compander is the CCITT (hence the European) standard approximation to the logarithmic compression. The form of the compressor is

$$y = C(x) = \begin{cases} y_{\max} \dfrac{A(|x| / x_{\max})}{1 + \log_e A} \operatorname{sgn} x & 0 < \dfrac{|x|}{x_{\max}} < \dfrac{1}{A} \\ y_{\max} \dfrac{1 + \log_e [A(|x| / x_{\max})]}{1 + \log A} \operatorname{sgn} (x) & \dfrac{1}{A} < \dfrac{|x|}{x_{\max}} < 1 \end{cases}$$

The standard value of the parameter A is 87.56, and for this value, using an 8-bit conversion, the average SNR is 38.0 dB.

Piecewise Approximation of PCM Quantizer

In order to simplify the implement of the PCM codec, the full range is divided into 16 segments. Four bits specify the segment and the remaining 4 bits specify the location of the signal level in that segment.

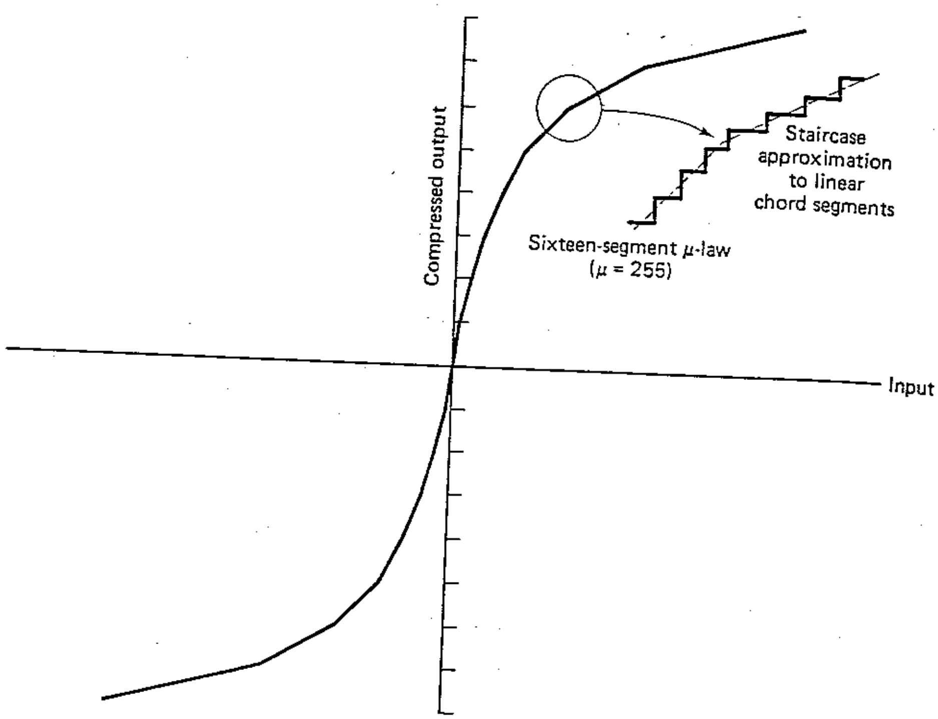
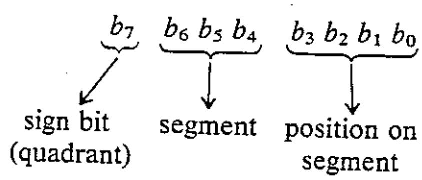
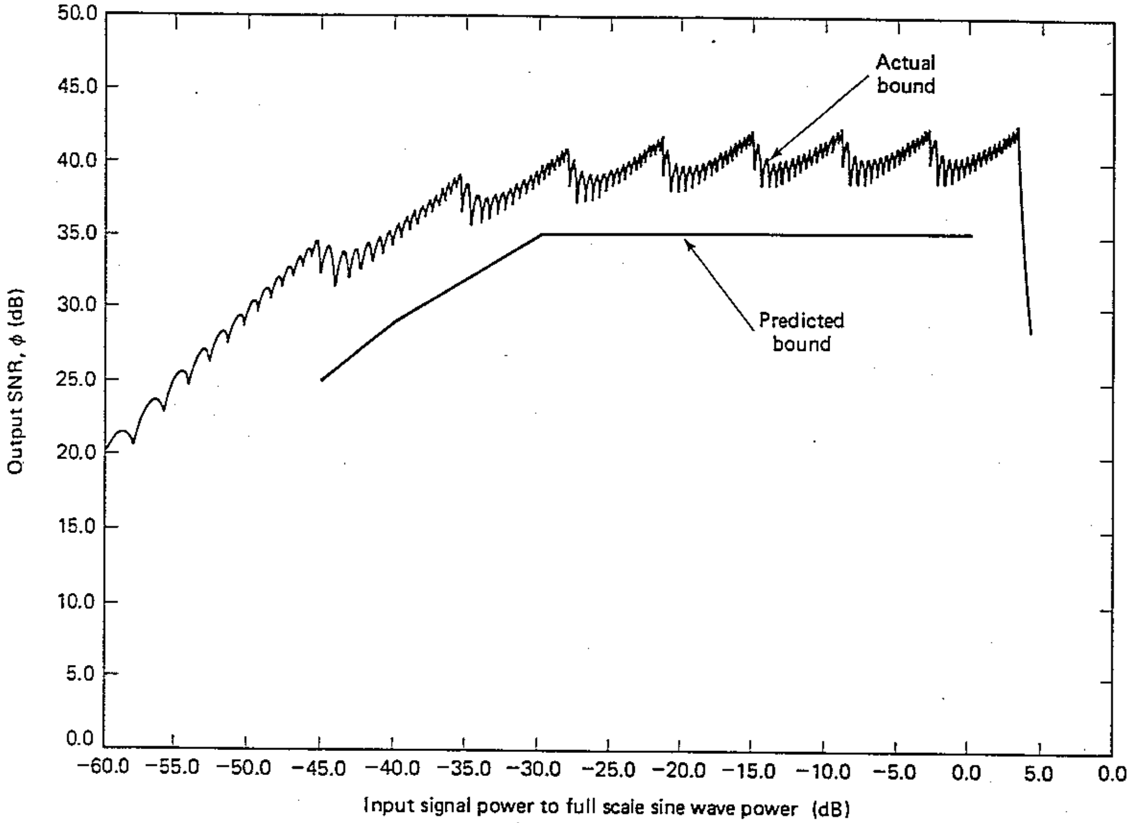


Figure Seven-bit compressed quantization with 16-segment approximation to μ-law.



Performance of PCM Quantizer



Predicted and measured SNR for a μ-law quantizer.

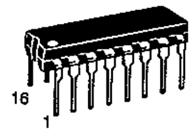
The 15-Segment Companding Characteristic ($\mu = 255$)

Linear Segment Number	Step-Size	Projections of Segment End Point onto the Horizontal Axis
0	2	± 31
$1a$, $1b$	4	± 95
$2a$, $2b$	8	± 223
$3a$, $3b$	16	± 479
$4a$, $4b$	32	± 991
$5a$, $5b$	64	± 2015
$6a$, $6b$	128	± 4063
$7a$, $7b$	256	± 8159

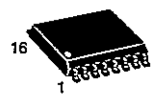
PCM Codec-Filter

MC145554
MC145557
MC145564
MC145567

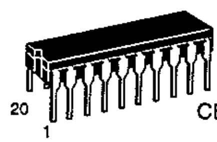
L SUFFIX
CERAMIC PACKAGE
CASE 620
MC145554/57



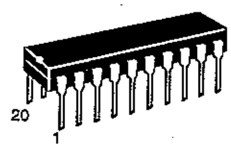
P SUFFIX
PLASTIC DIP
CASE 648
MC145554/57



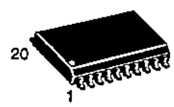
DW SUFFIX
SOG PACKAGE
CASE 751G
MC145554/57



L SUFFIX
CERAMIC PACKAGE
CASE 732
MC145564/67



P SUFFIX
PLASTIC DIP
CASE 738
MC145564/67



DW SUFFIX
SOG PACKAGE
CASE 751D
MC145564/67

The MC145554, MC145557, MC145564, and MC145567 are all per channel PCM Codec–Filters. These devices perform the voice digitization and reconstruction as well as the band limiting and smoothing required for PCM systems. They are designed to operate in both synchronous and asynchronous applications and contain an on–chip precision voltage reference. The MC145554 (Mu–Law) and MC145557 (A–Law) are general purpose devices that are offered in 16–pin packages. The MC145564 (Mu–Law) and MC145567 (A–Law), offered in 20–pin packages, add the capability of analog loopback and push–pull power amplifiers with adjustable gain.

These devices have an input operational amplifier whose output is the input to the encoder section. The encoder section immediately low–pass filters the analog signal with an active R–C filter to eliminate very–high–frequency noise from being modulated down to the pass band by the switched capacitor filter. From the active R–C filter, the analog signal is converted to a differential signal. From this point, all analog signal processing is done differentially. This allows processing of an analog signal that is twice the amplitude allowed by a single–ended design, which reduces the significance of noise to both the inverted and non–inverted signal paths. Another advantage of this differential design is that noise injected via the power supplies is a common–mode signal that is cancelled when the inverted and non–inverted signals are recombined. This dramatically improves the power supply rejection ratio.

After the differential converter, a differential switched capacitor filter band passes the analog signal from 200 Hz to 3400 Hz before the signal is digitized by the differential compressing A/D converter.

The decoder accepts PCM data and expands it using a differential D/A converter. The output of the D/A is low–pass filtered at 3400 Hz and sinX/X compensated by a differential switched capacitor filter. The signal is then filtered by an active R–C filter to eliminate the out–of–band energy of the switched capacitor filter.

These PCM Codec–Filters accept both long–frame and short–frame industry standard clock formats. They also maintain compatibility with Motorola's family of TSACs and MC3419/MC34120 SLIC products.

The MC145554/57/64/67 family of PCM Codec–Filters utilizes CMOS due to its reliable low–power performance and proven capability for complex analog/digital VLSI functions.

MC145554/57 (16–Pin Package)

- Fully Differential Analog Circuit Design for Lowest Noise
- Performance Specified for Extended Temperature Range of – 40 to + 85°C
- Transmit Band–Pass and Receive Low–Pass Filters On–Chip
- Active R–C Pre–Filtering and Post–Filtering
- Mu–Law Companding MC145554
- A–Law Companding MC145557
- On–Chip Precision Voltage Reference (2.5 V)
- Typical Power Dissipation of 40 mW, Power Down of 1.0 mW at ± 5 V

MC145564/67 (20–Pin Package) — All of the Features of the MC145554/57 Plus:

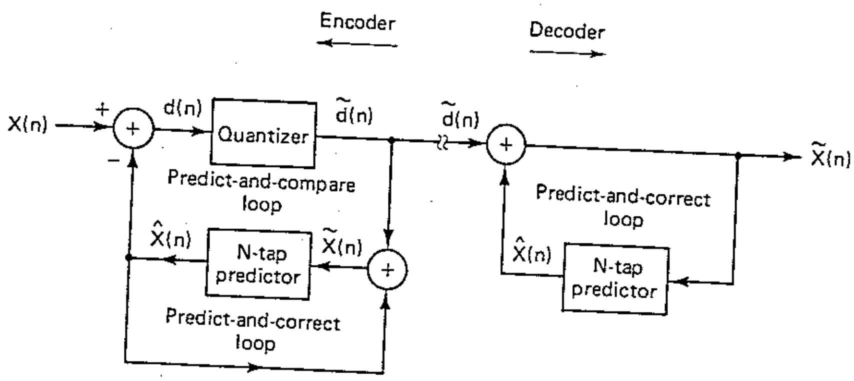
- Mu–Law Companding MC145564
- A–Law Companding MC145567
- Push–Pull Power Drivers with External Gain Adjust
- Analog Loopback

REV 1
9/95 (Replaces ADI1517)

MOTOROLA

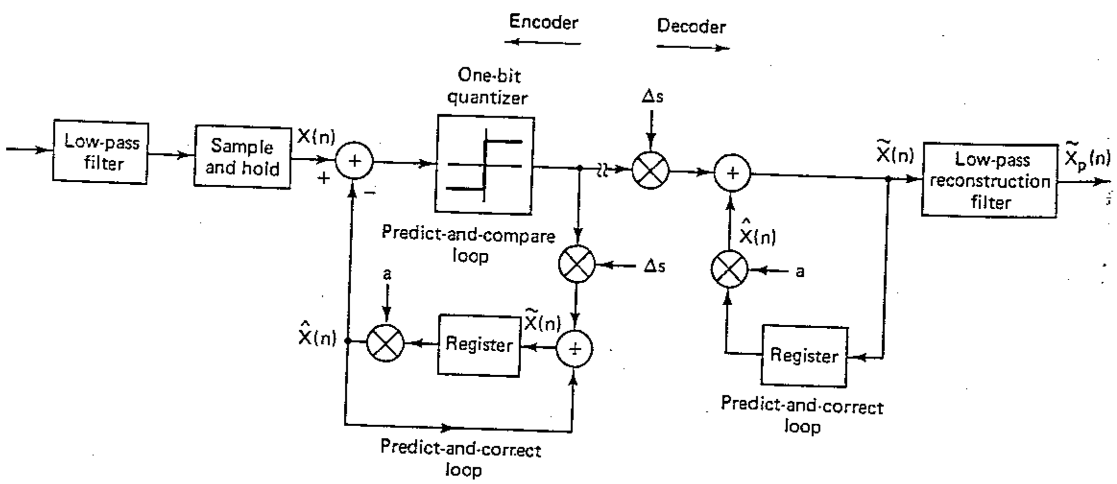
Differential Pulse Code Modulation (DPCM)

In order to take advantage of the correlation between the consecutive samples of the signal, one could, instead of quantizing the samples themselves, quantize the difference between the consecutive samples.



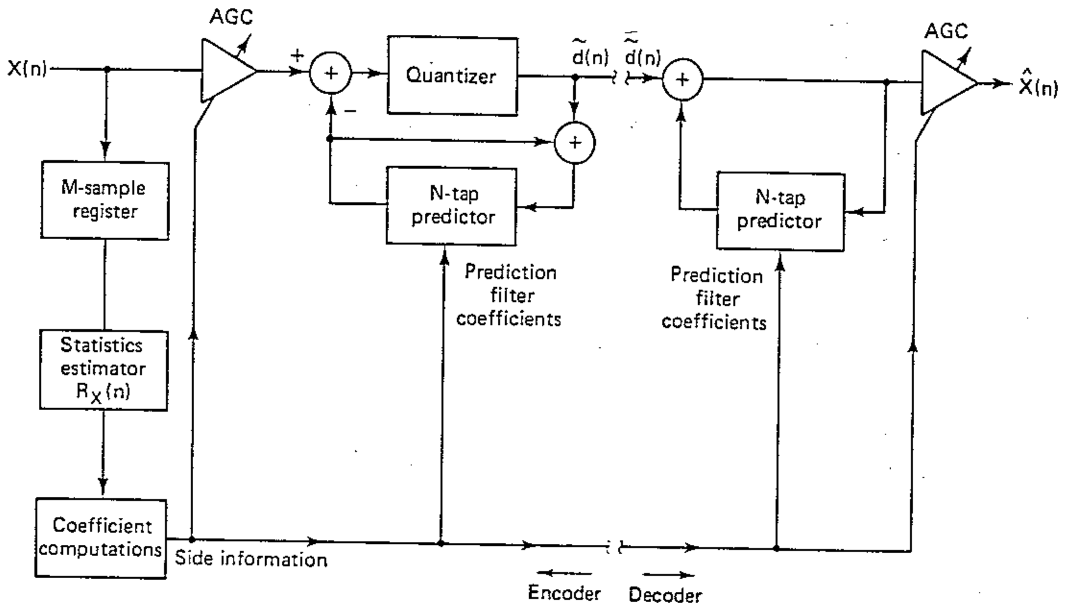
$$\tilde{X}(n) = \hat{X}(n) + \tilde{d}(n)$$

Delta Modulation: The implementation can be drastically simplified by using a single bit quantizer (just a comparator),

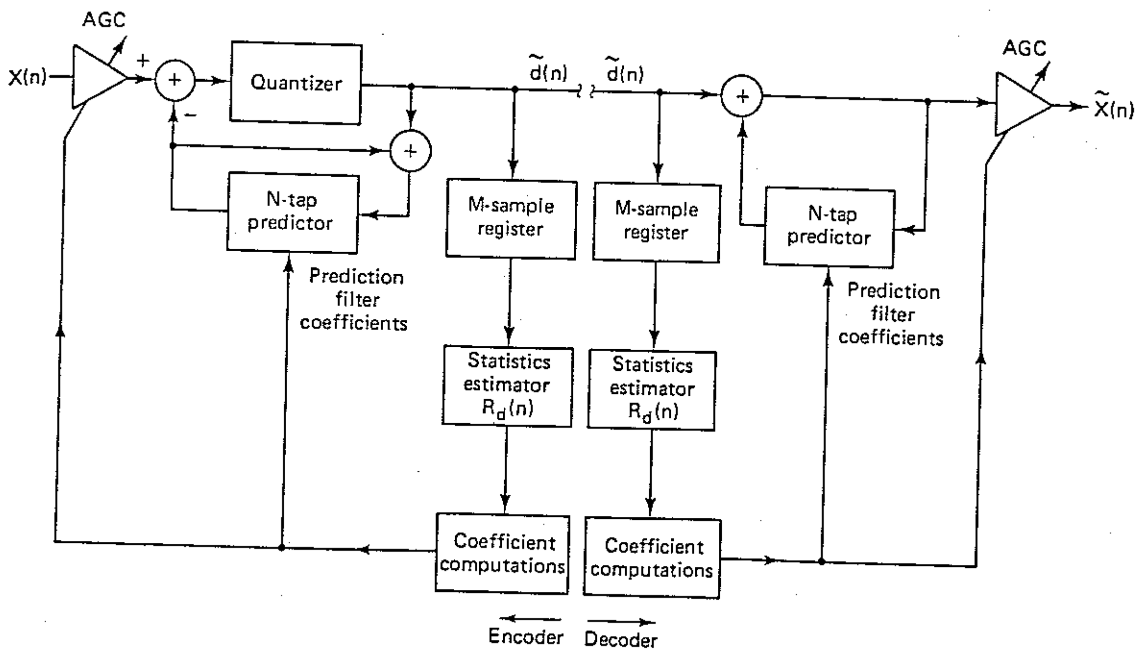


Adaptive Differential Pulse Code Modulation (ADPCM)

Extra improvement in performance or, equivalently, reduction in rate for the same performance could be achieved by making the prediction filter adaptive, i.e., changing the filter taps according to the source statistics.



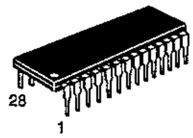
Forward adaptive prediction and quantization coding.



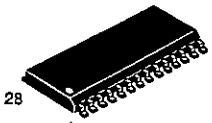
Backward adaptive prediction and quantization coding.

MC14LC5540

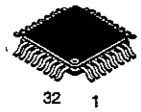
Technical Summary
ADPCM Codec


28
1

P SUFFIX
PLASTIC DIP
CASE 710


28
1

DW SUFFIX
SOG PACKAGE
CASE 751F


32 1

FU SUFFIX
TQFP
CASE 873A

ORDERING INFORMATION

MC14LC5540P	Plastic DIP
MC14LC5540DW	SOG Package
MC14LC5540FU	TQFP

This technical summary provides a brief description of the MC14LC5540 ADPCM Codec. A complete data book for the MC14LC5540 is available and can be ordered from your local Motorola sales office. The data book number is MC145540/D.

The MC14LC5540 ADPCM Codec is a single chip implementation of a PCM Codec–Filter and an ADPCM encoder/decoder, and therefore provides an efficient solution for applications requiring the digitization and compression of voiceband signals. This device is designed to operate over a wide voltage range, 2.7 to 5.25 V and, as such, is ideal for battery powered as well as ac powered applications. The MC14LC5540 ADPCM Codec also includes a serial control port and internal control and status registers that permit a microcomputer to exercise many built–in features.

The ADPCM Codec is designed to meet the 32 kbps ADPCM conformance requirements of CCITT Recommendation G.721–1988 and ANSI T1.301. It also meets ANSI T1.303 and CCITT Recommendation G.723–1988 for 24 kbps ADPCM operation, and the 16 kbps ADPCM standard, CCITT Recommendation G.726. This device also meets the PCM conformance specification of the CCITT G.714 Recommendation.

- Single 2.7 to 5.25 V Power Supply
- Typical 2.7 V Power Dissipation of 43 mW, Power–Down of 15 μW
- Differential Analog Circuit Design for Lowest Noise
- Complete Mu–Law and A–Law Companding PCM Codec–Filter
- ADPCM Transcoder for 64, 32, 24, and 16 kbps Data Rates
- Universal Programmable Dual Tone Generator
- Programmable Transmit Gain, Receive Gain, and Sidetone Gain
- Low Noise, High Gain, Three Terminal Input Operational Amplifier for Microphone Interface
- Push–Pull, 300 Ω Power Drivers with External Gain Adjust for Receiver Interface
- Push–Pull, 300 Ω Auxiliary Output Drivers for Ringer Interface
- Voltage Regulated Charge Pump to Power the Analog Circuitry in Low Voltage Applications
- Receive Noise Burst Detect Algorithm
- Order Complete Document as MC145540/D
- Device Supported by MC145537EVK ADPCM Codec Evaluation Kit

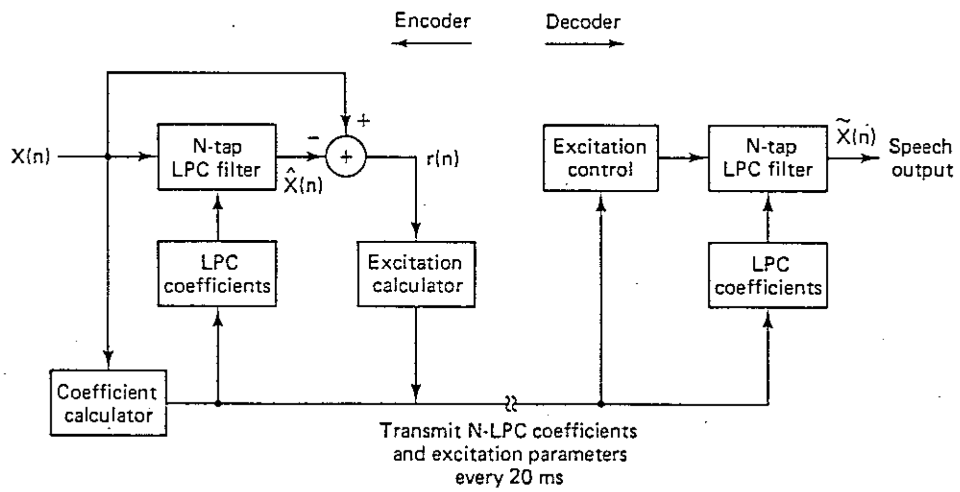
This document contains information on a product under development. Motorola reserves the right to change or discontinue this product without notice.

REV 2
6/97 TN97060200

Linear Predictive Coding (LPC)

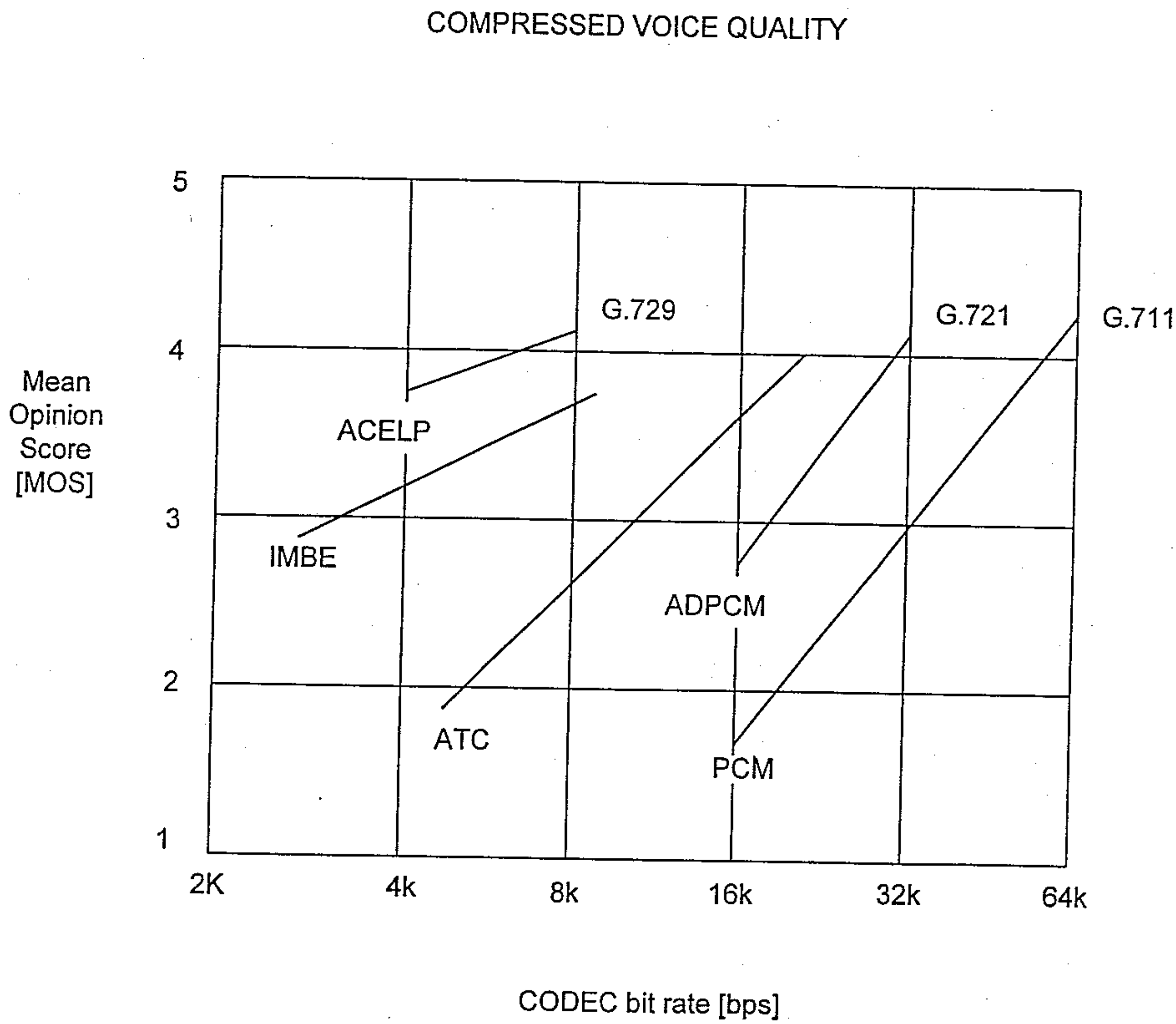
Taken from Digital Communications by B. Sklar

The adaptive predictors, described in Section 11.3.4, were designed to predict or form good estimates of an input speech signal. In the adaptive form, the prediction coefficients are recomputed as side information from periodic examination of the input data. Then the difference between the input and the prediction is transmitted to the receiver to resolve the prediction error. *Linear predictive coders* (LPCs) are the natural extension of N-tap predictive coders. When the filter coefficients are periodically computed with an optimal algorithm, the prediction is so good that there is (essentially) no prediction error information worth transmitting to the receiver. Rather than transmit these low-level prediction errors, the LPC system transmits the filter coefficients and the voiced/unvoiced excitation decision for the model. Thus the only data sent in LPC is the high-quality side information of the classic adaptive algorithm. An LPC model for voice synthesis is shown in Figure 11.34. The Texas Instruments Speak and Spell learning games use a 12-tap LPC speech synthesizer implemented by a single microchip.



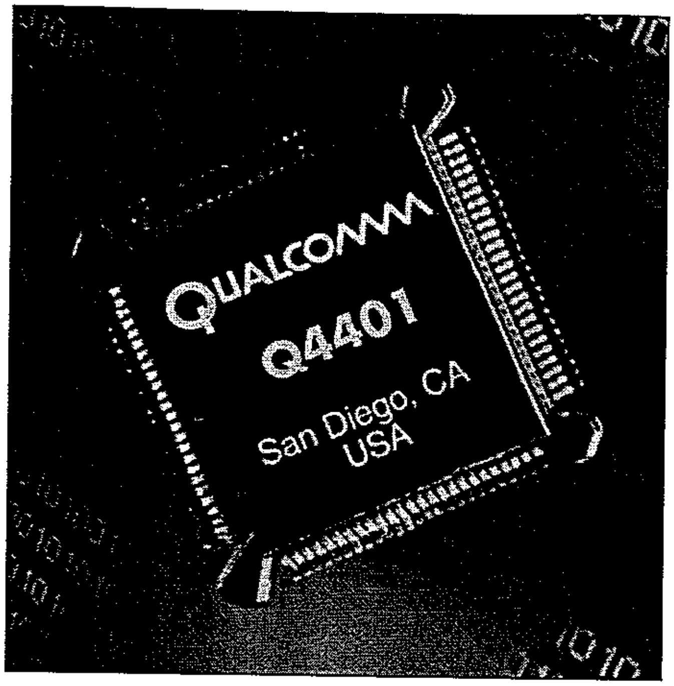
Linear predictive coefficient (LPC) speech modeling.

Comparison of Different Speech CODECs



Q4401

VARIABLE RATE VOCODER



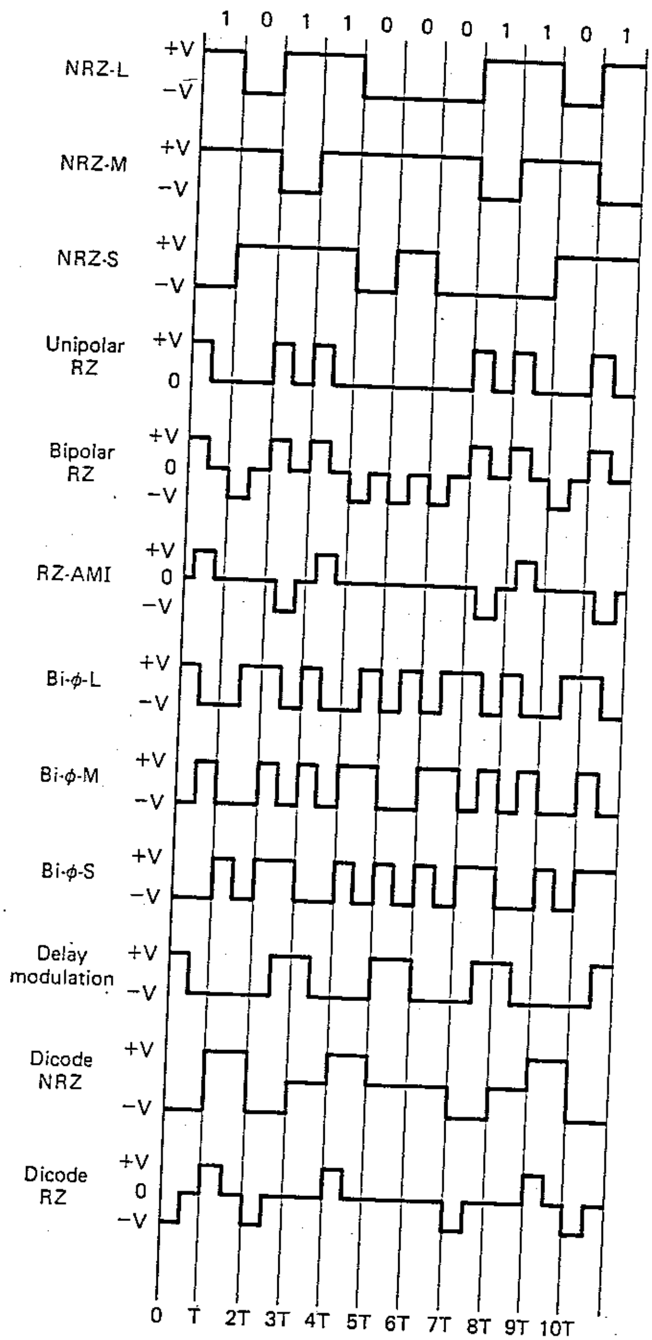
GENERAL DESCRIPTION

The QUALCOMM Q4401 Variable Rate Vocoder is a full-duplex speech Encoder and Decoder that produces near toll-quality speech at compressed data rates of under 9.6 kilobits per second (kbps). The Q4401 provides a single-chip solution to the speech compression requirements for digital telephone, wireless communications, voice storage, and speech synthesis systems. The Q4401 uses the proprietary QUALCOMM Codebook Excited Linear Predictive (QCELP) speech coding algorithm to achieve high speech quality at low data rates.

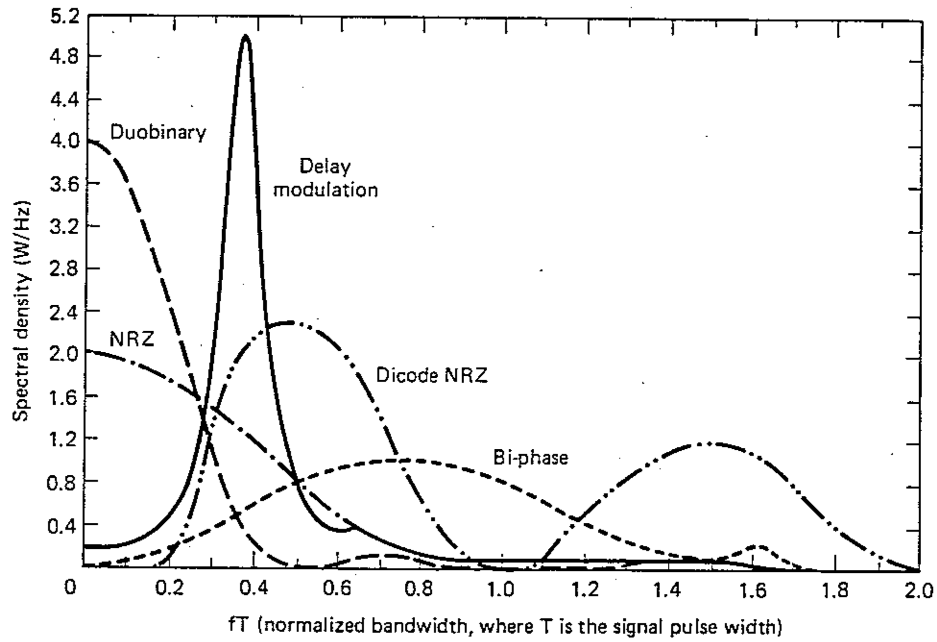
The Q4401 can encode speech at fixed or variable data rates. In Fixed Rate Mode, the Q4401 can code speech at rates of 4 kbps, 4.8 kbps, 8 kbps or 9.6 kbps. In Variable Rate Mode, the Q4401 automatically adjusts the data rate from 800 bps to 8 kbps (Normal Variable Rate Mode) or from 800 bps to 9.6 kbps (Enhanced Variable Rate Mode) every 20 milliseconds (ms). When in Variable Rate Mode, the Q4401 codes speech at under 7 kbps in continuous speech applications and at under 3.5 kbps in typical two-way telephone conversations, without degrading the speech quality.

The Q4401 is a masked ROM version of a digital signal processor (DSP) device. Digitized speech is transferred to and from the Q4401 via a digital serial interface that connects to a 64 kbps µ-law or A-law speech codec. Compressed speech packets are transferred to and from the Q4401 via an 8-bit parallel data bus interface that connects to standard microprocessor buses. The Q4401 is also controlled via this processor interface.

Various PCM Waveforms



Spectral Densities ofVarious PCM Waveforms

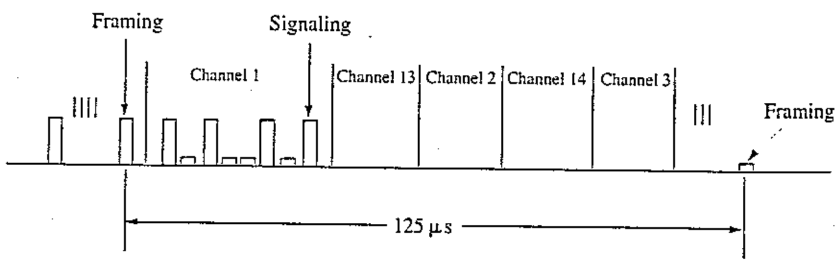


Digital Hierarchy

DS0: Single Voice Channel

Sampling speech at a rate of 8000 samples/sec. and quantizing each sample with an 8 bit PCM -> 64 kbps.

DS1 or T1: Equivalent of 24 PCM (DS0) channels.



There are 8x24+1 bits in a 125 usec frame -> 1.544 Mbps. One bit out of every 193 bit is spent for synchronization. So, the actual rate is 1.536 Mbps. Furthermore, every sixth frame, the least significant bit of each channel is used for signaling. That is each sample is actually encoded with 7 5/6 bits instead of 8 bits.

E1 (European Standard): Equivalent of 32 PCM channels: 31 information channels+1 signaling channels. The total rate is 64 kbpsx32 = 2.048 Mbps.

DS2 or T2: Equivalent of four T1 or 96 PCM channels -> 6.312Mbps.

DS3 or T3: Equivalent of 28 T1 or 672 PCM channels -> 44.736 Mbps.

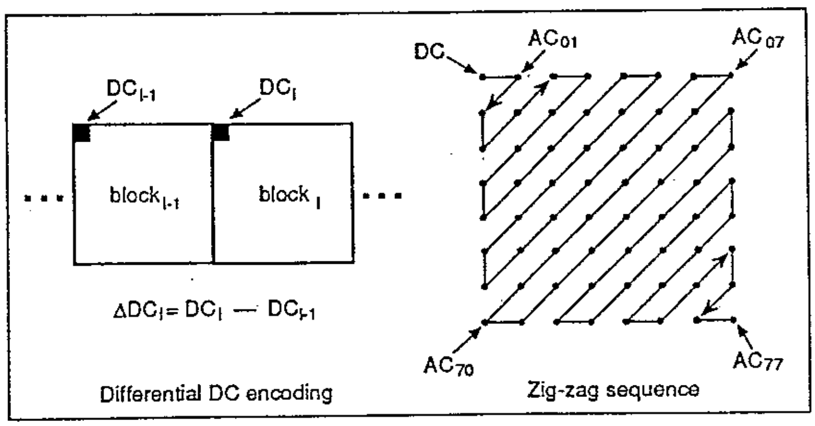
DS4 or T4: Equivalent of 4032 PCM channels -> 274.176 Mbps.

JPEG Image Compression Algorithm

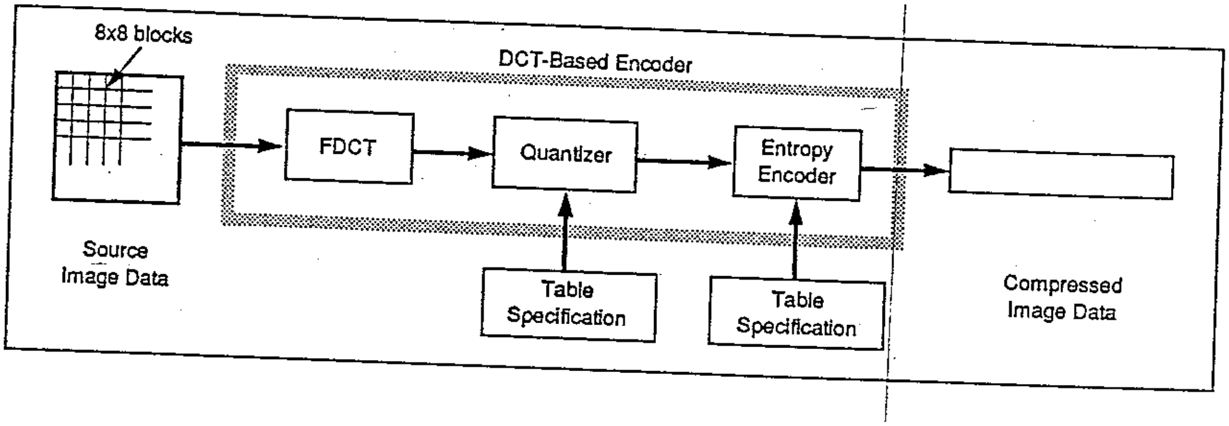
JPEG Developed by Joint Photographic Expert Group is the result of collaboration between CCITT and ISO.

TJPEG is a Transform Coding Technique consisting of the following steps:

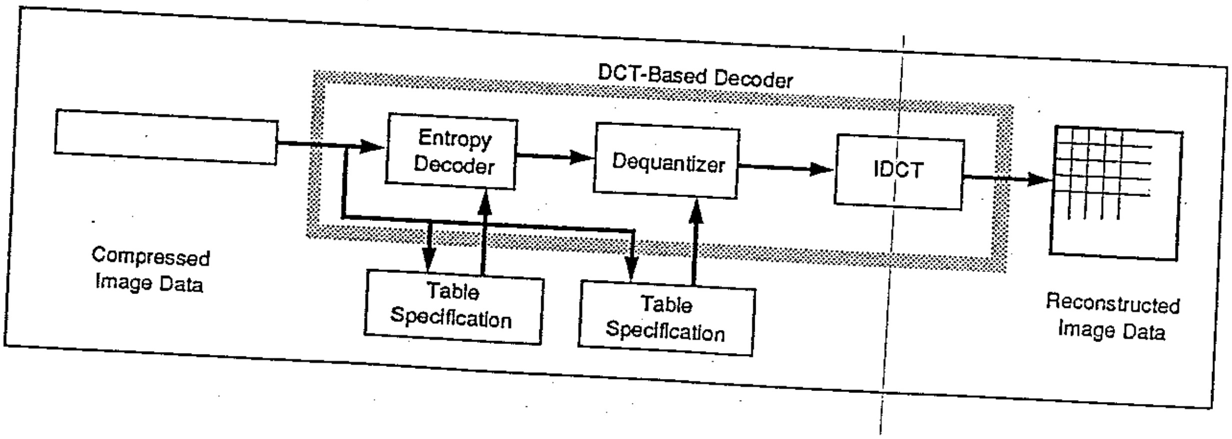
1- Segmenting the image into 8x8 blocks,
2- Taking Discrete Cosine Transform (DCT) of each block resulting in 63 AC and 1 DC coefficient,
3- Quantizing the DCT coefficients using step sizes that depend on the required performance, to utilize the inter-block correlation, the DC (mean) coefficients are quantized using DPCM,
4- Compressing the output of quantizer using entropy coding (Hufmann Coding). That is assigning smaller number of bits to more likely (usually small magnitude) values.



JPEG Encoder



JPEG Decoder



MPEG Video Compression

In Video Compression in addition to removing spatial (intra-frame) redundancy, one tries to also remove temporal (inter-frame) redundancy using motion compensated interpolation.

In MPEG, temporal redundancy removal is achieved using three types of pictures:

I or Intra-pictures: these are quite high resolution pictures (low compression) and are used for random access and reference for predicted pictures,

P or Predicted pictures: these are encoded with reference to a previous picture (either Intra- or Predicted) and are used as reference for future Predicted pictures.

B or Bidirectional predicted pictures: these are most compressed images and require both a past and a future picture for prediction. The B pictures are never used as reference for other pictures.