Differential pulse code modulation

According to the Nyquist sampling criterion, a signal must be sampled at a sampling rate that is at least twice the highest frequency in the signal to be able to reconstruct it without aliasing. The samples of a signal that is sampled at that rate or close to generally have little correlation between each other (knowing a sample does not give much information about the next sample). However, when a signal is highly oversampled (sampled at several times the Nyquist rate, the signal does not change a lot between from one sample to another. Consider, for example, a sine function that is sampled at the Nyquist rate. Consecutive samples of this signal may alternate over the whole range of amplitudes from –1 and 1. However, when this signal is sampled at a rate that is 100 times the Nyquist rate (sampling period is 1/100 of the sampling period in the previous case), consecutive samples will change a little from each other. This fact can be used to improve the performance of quantizers significantly by quantizing a signal that is the difference between consecutive samples instead of quantizing the original signal. This will result in either requiring a quantizer with much less number of bits (less information to transmit) or a quantizer with the same number of bits but much smaller quantization intervals (less quantization noise and much higher SNR).

DPCM is a derivative of standard PCM and exploits the fact that , for most audio signal , the range of the difference in amplitude between successive samples of audio waveform is less than the range of the actual sample amplitude.

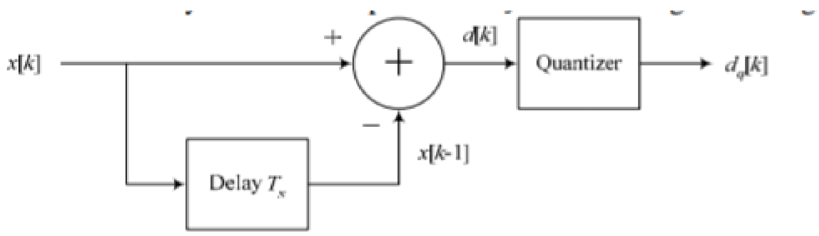
If the digitized difference signal is used to encode the waveform then fewer bits are required as compare to PCM signal with the same sampling rate.

Consider a signal $x(t)$ that is sampled to obtain the samples $x(kTs)$, where Ts is the sampling period and k is an integer representing the sample number. For simplicity, the samples can be written in the form $x[k]$, where the sample period Ts is implied. Assume that the signal $x(t)$ is sampled at a very high sampling rate. We can define $d[k]$ to be the difference between the present sample of a signal and the previous sample, or d[k]= x[k] –x[k -1]

Now this signal d[k] can be quantized instead of x[k] to give the quantized signal dq[k]. As mentioned above, for signals x(t) that are sampled at a rate much higher than the Nyquist rate, the range of values of d[k] will be less than the range of values of x[k].

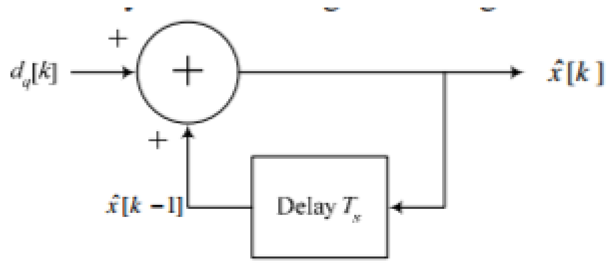
After the transmission of the quatized signal dq[k], theoretically we can reconstruct the original signal by doing an operation that is the inverse of the above operation. So, we can obtain an approximation of x[k] using

x^[k] = dq[k]+ x^[k-1]

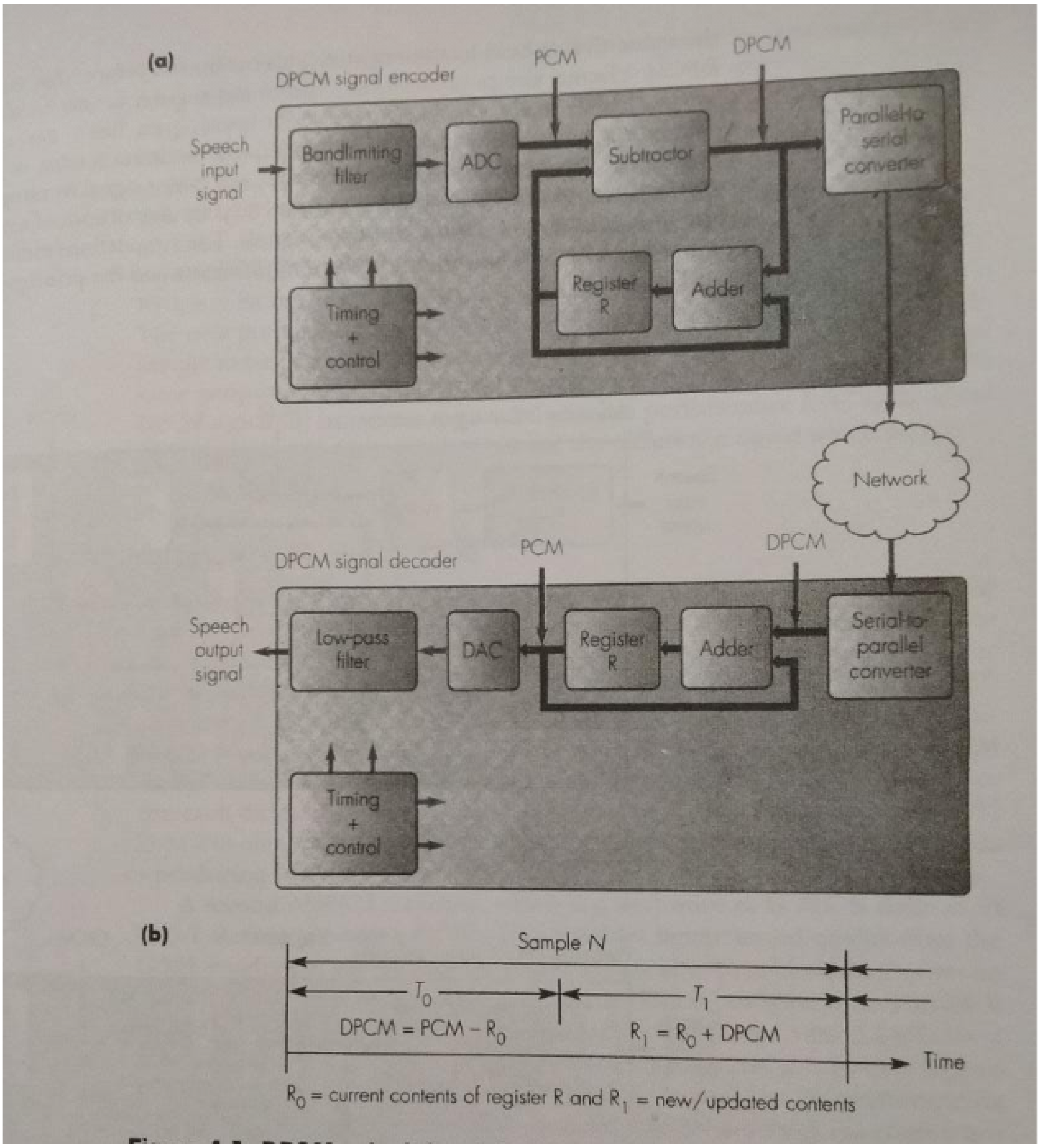


So, if dq[k] is close to d[k], it appears from the above equation that obtained x^[ ] k will be close to d[k]. However, this is generally not the case as will be shown later. The transmitter of the above system can be represented by the following block diagram

The receiver that will attempt to reconstruct the original signal after transmitting it through the channel can be represented by the following block diagram.



DPCM encoder and decoder



(a) Encoder/ decoder (B) encoder timing

- The previous digitized sample of the analog input signal is held in the register (temporary storage facility.
- The difference signal is computed by subtracting the current contents of the register from the new digitized sample output by the ADC
- The value in the register is then updated by adding to the current register contents the computed difference signal output by the subtractor prior to its transmission.

Decoder

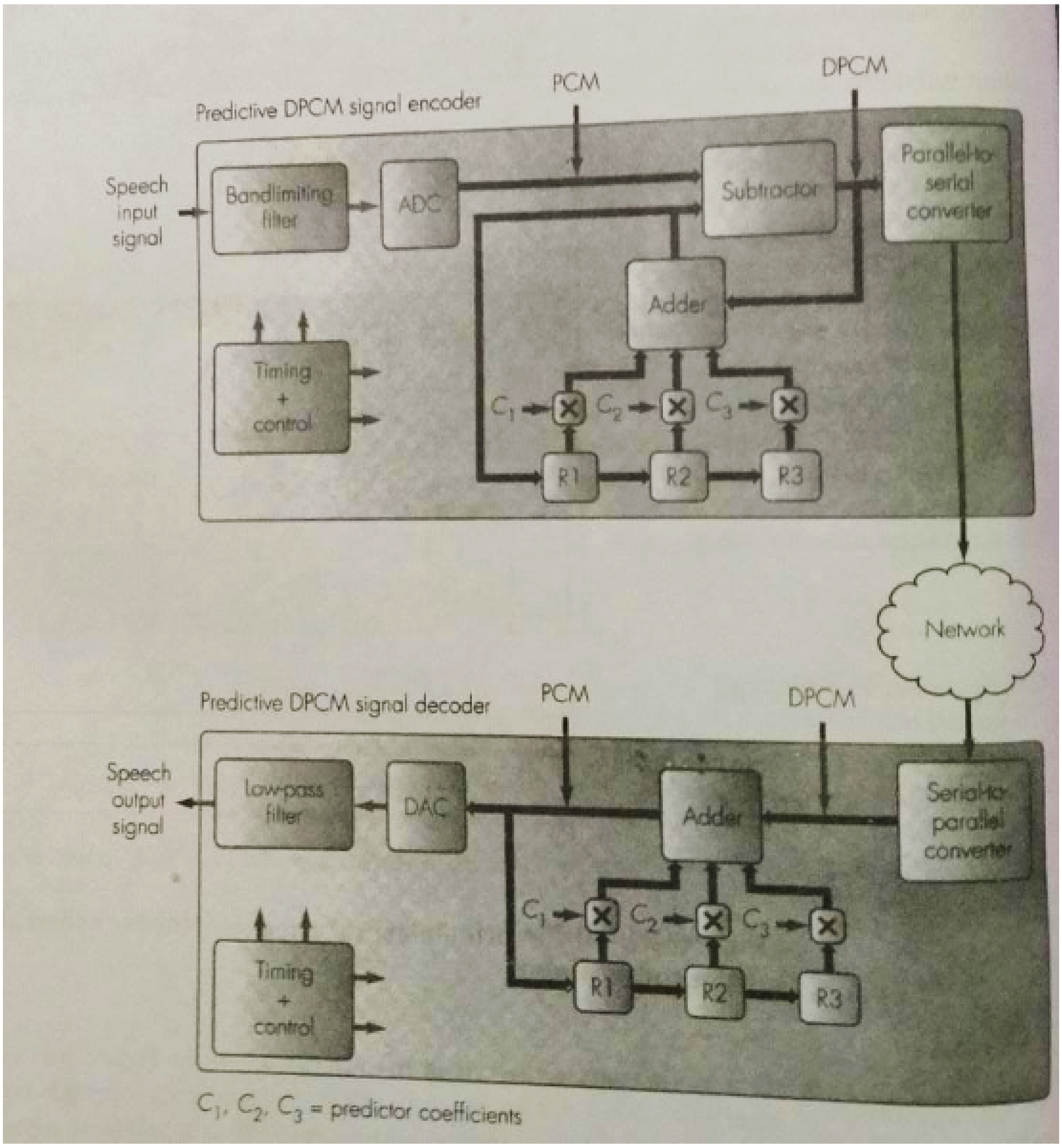
- It operates by simply adding the received difference signal (DPCM) to the previously computed signal held in the register(PCM)

The output of ADC is used directly . hence the accuracy of each computed difference signal (residual signal), is determined by the accuracy of the previous signal held in the register.

we are quantizing a difference signal and transmitting that difference over the channel, the reconstructed signal may suffer from one or two possible problems.

1. Accumulation of quantization noise: the above system suffer from the possible accumulation of the quantization noise. Unlike the quantization of a signal where quantization error in each sample of that signal is completely independent from the quantization error in other samples, the quantization error in this system may accumulate to the point that it will result in a reconstructed signal that is very different from the original signal.
2. Effect of transmission errors: in a regular PCM system, the effect of an error that happens in the transmitted signal is only limited to the sample in which the error occurs. In DPCM, an error that occurs in the transmitted signal will cause all the reconstructed samples at the receiver after that error occurs to have errors. Therefore, even if quantization error did not accumulate, an error caused by the channel will cause all successive samples to be wrong.

These problems can be solved by using predictive DPCM signal encoder and decoder



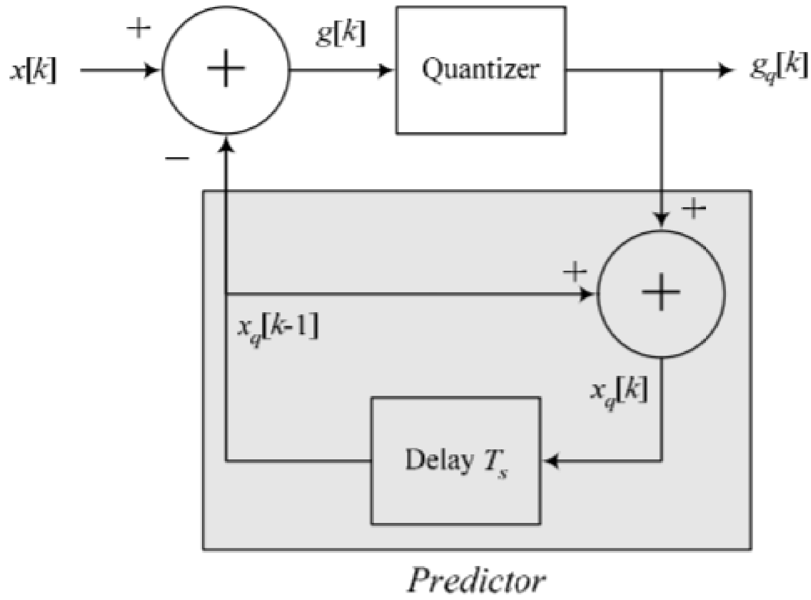
1. Eliminating the problem of accumulation of quantization noise: This problem can be solved by avoiding the quantization of the difference signal d[k] between x[k] and its previous sample x[k−1],or

   d[k]= x[k] −x[k -1]

   and quantizing instead a difference signal (we will call it g[k]) that is the difference between x[k] and the previous sample of its quantized form xq[k−1]. Therefore, g[k] is given by

$g[k] = x[k] - x_q[k-1]$

this will require applying the quantizer on the signal x[k] to obtain xq[k−1], which we are trying to avoid since the amplitude of x[k] is generally larger than the amplitude of a difference signal like d[k] or even g[k]. In fact, if both x[k] and g[k] are available, we can reconstruct the quantized form of x[k] using the following system.



In the above system, we can easily prove that the resulting signal xq[k] is the quantized form of x[k].

Now, the output of the quantizer is the quantized form of g[k] which can be represented by adding a quantization noise q[k] to the input of the quantizer. Therefore,

$g_q[k] = g[k] + q[k]$

Substituting for g[k] in gq[k] gives

$g_q[k] = g[k] = x[k] - x_q[k-1] + q[k]$

From the block diagram,

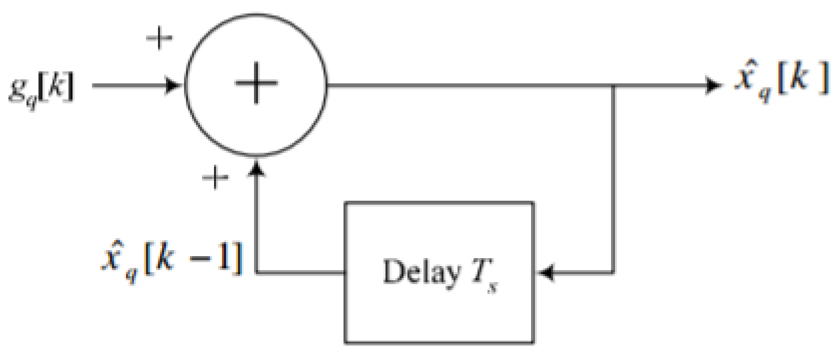
$X_q[k] = g_q[k] + x_q[k-1]$

=x[k]+q[k]

So, in fact, the function xq[k] is the quantized form of x[k] as seen by the last line of the above equation.

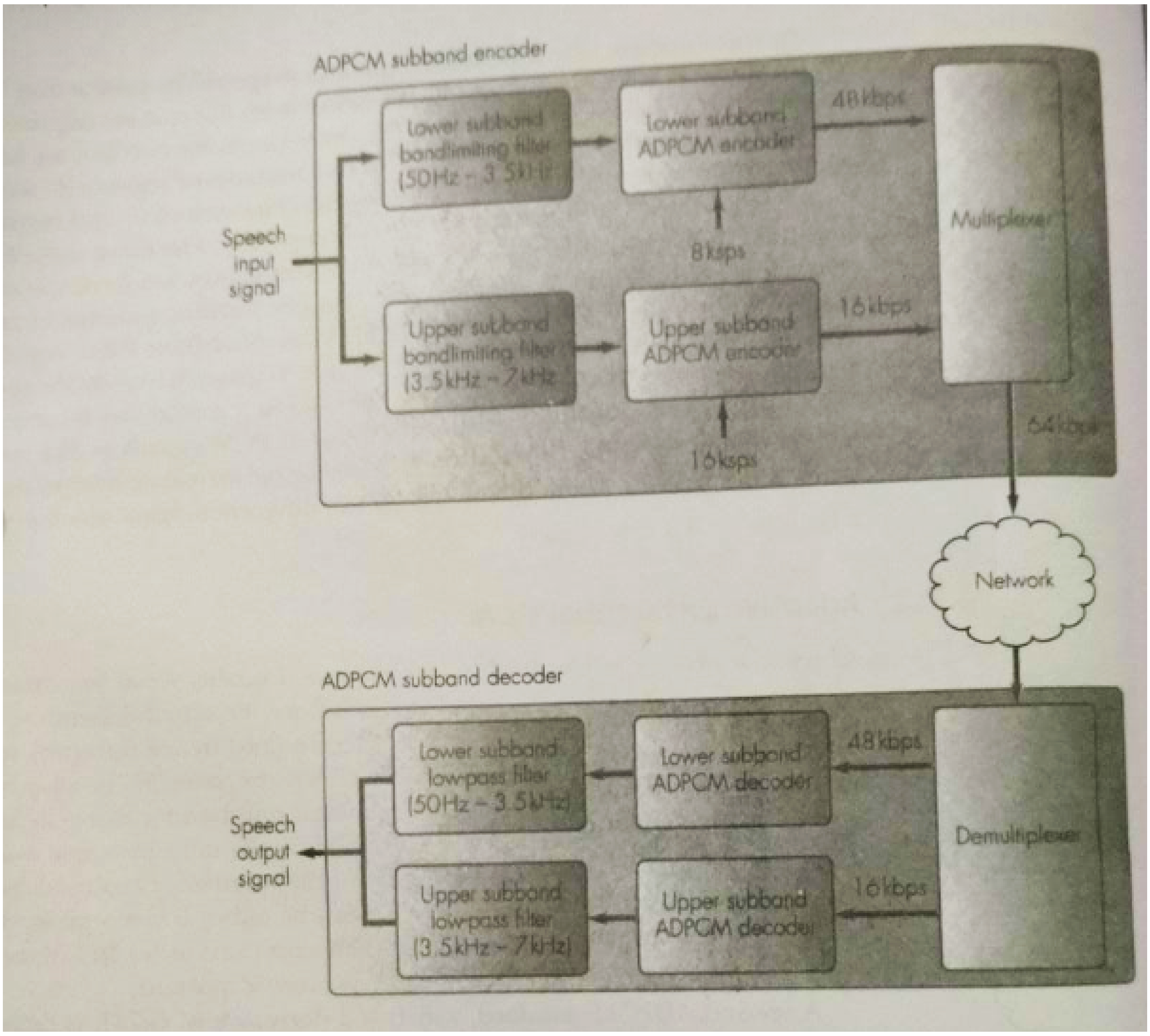
If we passed x[k] through the same quantizer in the block diagram above, we will get another function xq2[k] with samples that are generally different from xq[k].

At the receiver side of the DPCM system, we can use the “Predictor” since its input is the DPCM output gq[k] and its output is the desired signal xq[k]. Therefore the block diagram would be as follows.



2. **Reducing the effect of transmission errors**: as mentioned before, transmission errors result in errors in all the reconstructed samples of the input signal that come after the transmission error. The best method to combat this problem is to divide the data into sets of samples and resent the transmitter and receiver after the transmission of each set of samples. This way, a transmission error that occurs will affect only the samples of that part of the data. Once the system is reset, the effect of that error will stop.

Adaptive differential PCM



The principle of Adaptive differential PCM is using fewer bits to encode and hence transmit smaller difference value then for larger values. An international standard for this is defined in ITU-T recommendation G.721. as compare to DPCM an eight order predictor is used and the no. of bits used to quantize each difference value is varied.

A second ADPCM standard is defined in ITU-T recommendation G.722.

This provides better sound quality as compare to standard G.721 at the expense of added complexity. It uses an added technique known as sub band coding

The input speech bandwidth is extended to be from 50Hz through to 7KHz- as compare to 3.4KHz for a standard PCM .hence a wider bandwidth produces a higher fidelity speech signal.

To allow for higher signal bandwidth , prior to sampling , the audio input signal is first passed through two filters

- One filter passes signal frequencies in the range of 50Hz to 3.5kHz. it is called lower subband signal
- Other passes signal frequencies in the range of 3.5kHz to 7KHz. It is called upper subband signal.

Therefore the input signal is divided into two separate equal bandwidth signals.

Each is then sampled and encoded independently using ADPCM

- The use of two subband has the advantage that different bit rates can be used for each.
- The frequency components that are present in the lower subband signal have a higher perceptual importance than those in the upper subband.
- The operating bit rate can be 64,56,or 48 kbps.
- For example , with a bit rate of 64kbs, the lower subband is encoded at 48kbps and upper sideband at 16kbps.
- The two bit stream are then multiplexed together , to produce the transmitted signal(64kbps), in such a way that the decoder in the receiver is able to divide them back again into two separate strems for encoding.

Third standard is ITU-T recommendation G.726

- It also uses subband coding but with a speech bandwidth of 3.4Hz.
- The operating bit rate can be 40,32,24, or 16kbps.

Adaptive predictive coding

A higher level of compression can be obtained by making the predictor coefficents adaptive .this is the principle of Adaptive predictive coding

The optimum set of predictor coefficents continuously vary , since they are a function of the characteristics of the audio signal being digitized.

To exploit this property, the input speech signal is divided into fixed time segm.ents

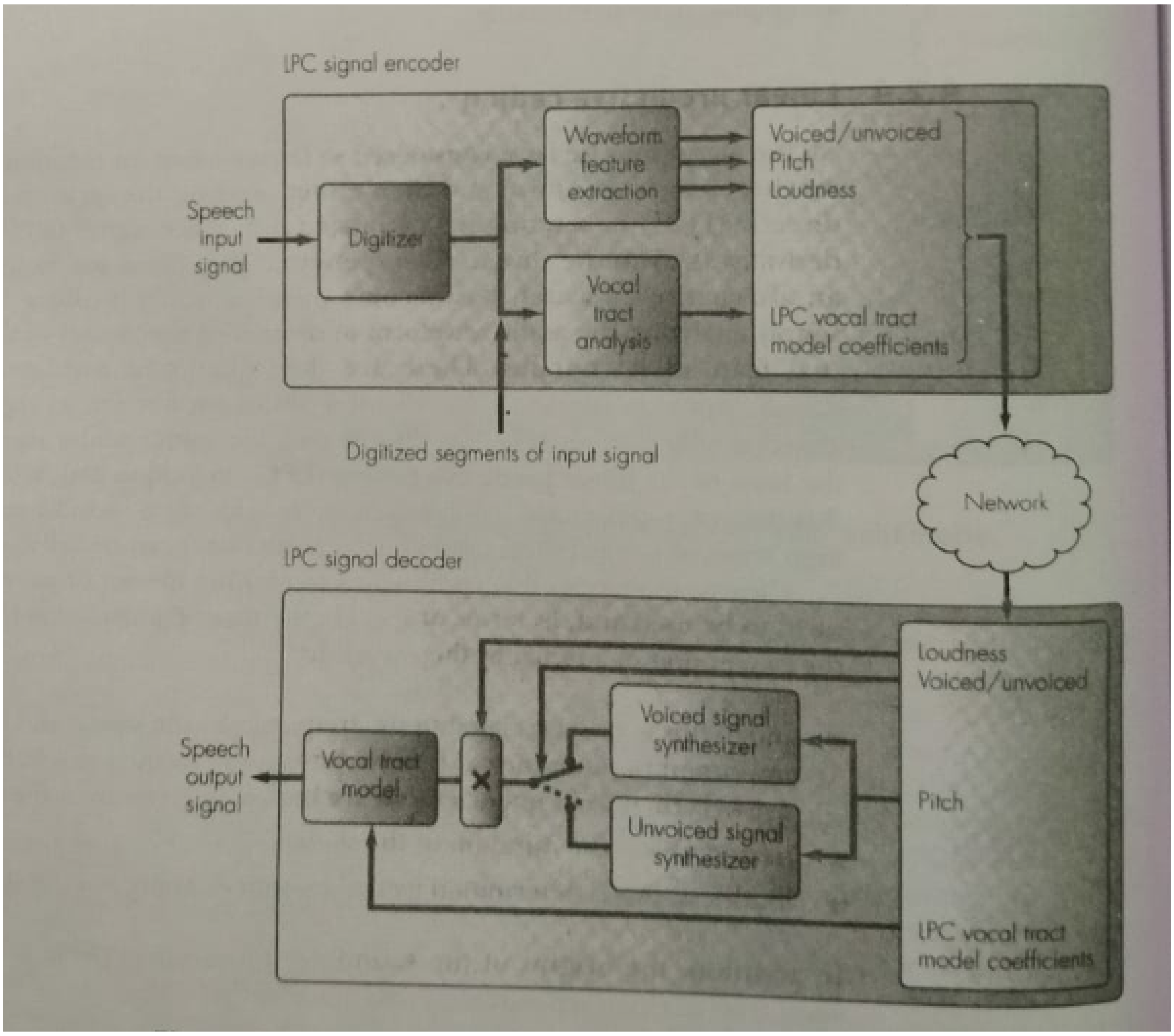
For each segment . the currently prevailing characteristics are determined .

The optimum set of coefficents are then computed .

These coefficents are used to predict more accurately the previous signal

This type of compression can reduce the bandwidth requirement to 8kbps, and obtain an acceptable perceived quality

Linear predictive coding



LPC methods are the most widely used in speech coding, speech synthesis, speech recognition, speaker recognition and verification and for speech storage – LPC methods provide extremely accurate estimates of speech parameters, and does it extremely efficiently.

PCM method is based on sending the quantized samples directly and DPCM method is based on sending the quantized difference signal and its derivative

but LPC method involves analysing the audio waveform to determine a selection of the perceptual feature it contains.

These are then quantized and sent .

The destination uses them , together with the sound synthesizer. To regenerate a sound that is perceptually comparable source audio signal .

Therefore it is necessary to identify the perceptual feature to be used .the three feature which determine the perception of a signal by the ear are:-

- Pitch
- Period
- Loudness

Pitch this is closely related to the frequency of the signal , it is important because the ear is more sensitive to the freq. in the range of 2-5kHz than to freq. that are higher or lower than these.

Period this is the duration of the signal

Loudness this is determined by the amount of energy in the signal.

Vocal tract excitation parameters

- Vocal sounds- these are generated through the vocal chords

  Eg. Sound relating to the letter m,v.

- Unvoiced sounds- the vocal chords are open

  Eg. Sound relating to the letters f, s.

When the perceptual feature have been obtained from the source wave form , it is possible to use them with a suitable model of vocal tract, to generate a synthesized version of original speech signal.

LPC encoder and decoder

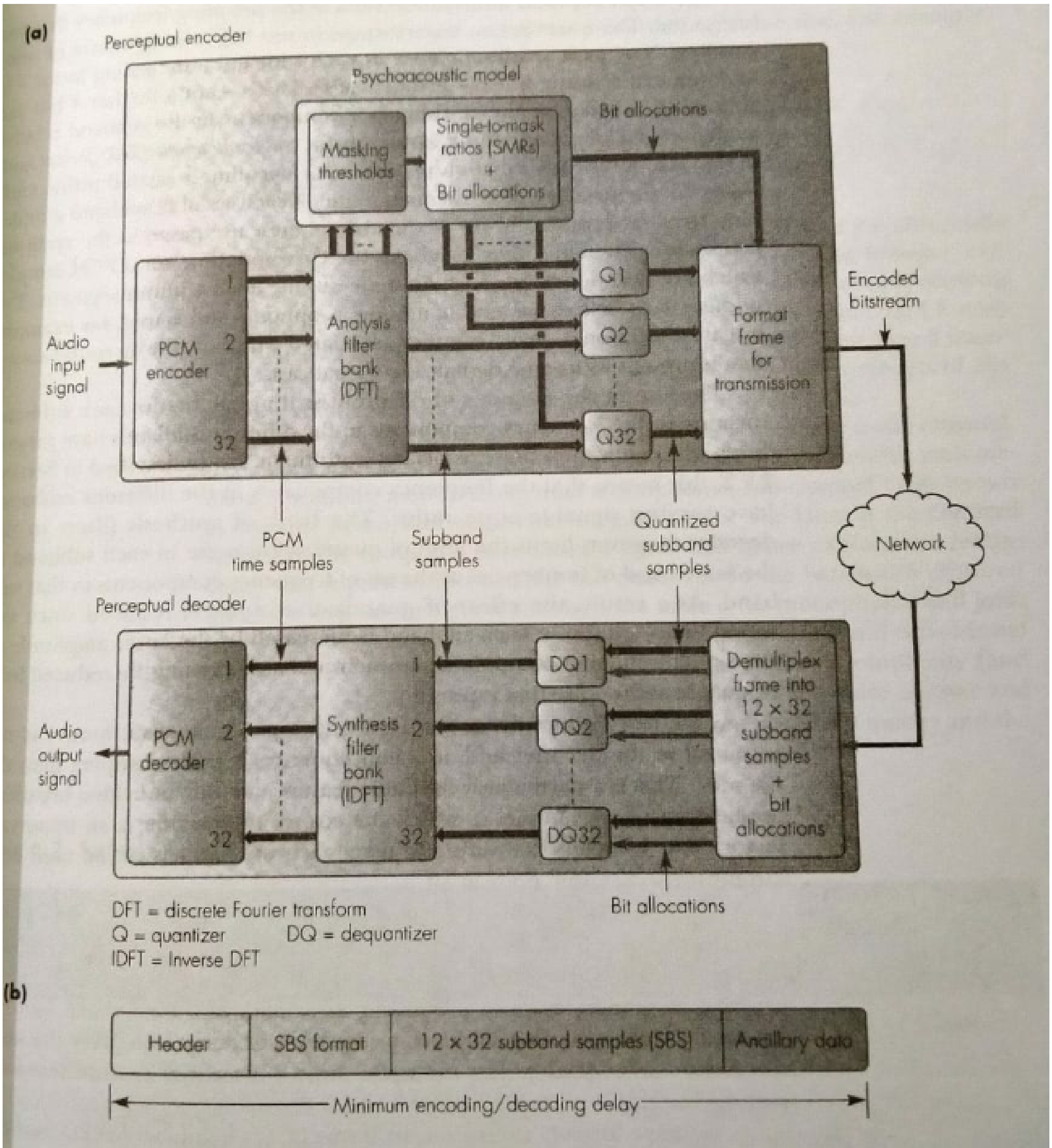
- The input speech waveform is first sampled and quantized at a defined rate
- A block of digitized samples (segment) is then analysed to determine the various perceptual parameters of the speech
- The speech signal generated by the vocal tract model in the decoder is function of the present output of the speech synthesizer plus a linear function of the previous set of model coefficents
- Hence the vocal tract model used is adaptive .
- The output of the encoder is a string of frames , one for each segment .
- Each frame contains a field for pitch , loudness and a notification of whether the signal is voiced or unvoiced,

MPEG(motion picture expert group) audio coders

Motion picture expert group was formed by the ISO to formulate a set of standards relating to a range of multimedia applications that involve the use of video with sound.

The coder associated with the audio compression are known a MPEG audio coders and a number of these uses perceptual coding.

MPEG encoder and decoder



ure 4.8 MPEG perceptual coder schematic: (a) encoder/decoder implementation ematic (b) example frame format.

- The time varying audio input signal is first sampled and quantizedusing PCM , then the sampling rate and the number of bits per sample is being determined
- The bandwidth that is available for transmission is divide into a number of freq. subbands using a bank of analysis filter(critical

band filter)

- critical band filter map each set of 32 PCM samples into equivalent set of 32 freq. samples , one per subband. Each freq. subband is of equal width, each is known as subband samples
- Eg. Assume 32 subband and a sampling rate of 32ksps, i.e maximum signal freq. of 16kHz , therefore each subband has a bandwidth of 500Hz.
- In a basic encoder, the time duration of each sampled segment of the audion input signal is equal to the time to accumulate 12 sucessive sets of 32 PCM anf hence subband samples . that is time duration equal to (12 * 32) PCM samples.
- critical band filter also determines the maximum amplitude of the 12 subband samples in each subband . each is known as Scaling factor fo the subband . these are passed together with the set of frequency samples in each subband to the quantizer block and to the psychoacoustic model
- The processing associated with both the freq. and temporal masking carried out by the psychoacoustic model which is performed concurrently witn the filtering and analysing operation
- DFT (discrete fourier transform) is used to transform 12 sets of 32 PCM samples into an equivalent set of frq. Components.
- Then using the masking properties and hearing threshpld of subband , the psychoacoustic model determines the various masking effect of this set of signals, the output of this model is known as signal to mask ratio.
- This ratio indicates those freq, components whose amplitude below the related audible threshold
- The set of scaling factors are used to determine the quantization accuracy
- All the frequency components in a sampled segment are encoded and these are carried in a frame the format of which shown in fig.

Header = it contains information related to sampling freq.

Subband sample format-the peak amplitude level in each subband is first quantized using 6 bits – giving 1 to 64 levels and further 4 bits are then used to quantize the 12 freq. components

- Decoder

Dequantizer is used to determine the magnitude of each 32 subband samples , and then passed these to the synthesis filter bank.. the bank then produces the corresponding set of PCM samples which are decoded to produce the time varying analog output segment.

Axcillary data field –it is used to carry the additional coded samples

- Decoder is less complex than the encoder because psychoacoustic model

Is not required in decoder.

Layer	Application	Compressed bit rate	Quality	Example input-to-output delay
1	Digital audio cassette	32 – 448 kbps	Hi-fi quality at 192 kbps per channel	20ms
2	Digital audio and digital video broadcasting	32 – 192 kbps	Near CD-quality at 128 kbps per channel	40ms
3	CD-quality audio over low bit rate channels	64 kbps	CD-quality at 64 kbps per channel	60ms

MPEG layer 1,2,3 perceptual encoder