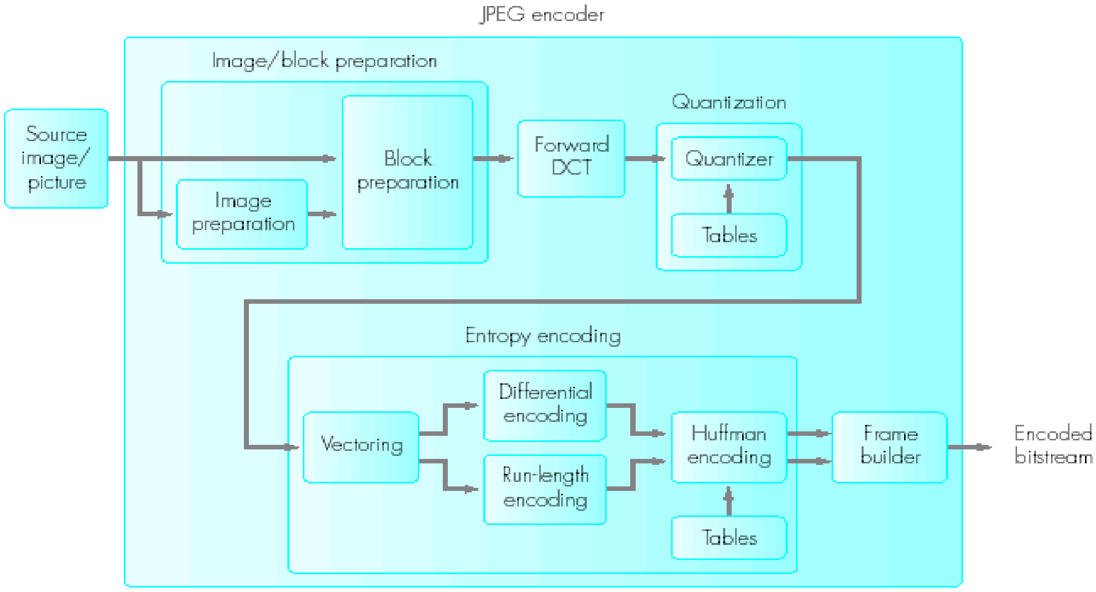
IAT 3

1.Explain JPEG encoder with a neat diagram.



Source image is made up of one or more 2-D matrices of values

2-D matrix is required to store the required set of 8-bit grey-level values that represent the image

For the colour image if a CLUT is used then a single matrix of values is required

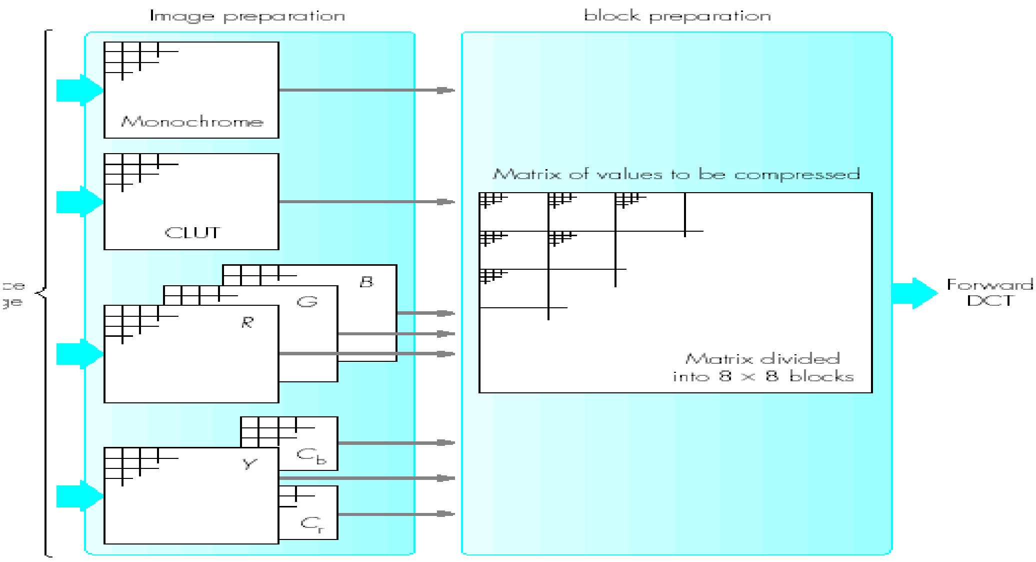
If the image is represented in R, G, B format then three matrices are required

If the Y, C_r, C_b format is used then the matrix size for the chrominance components is smaller than the Y matrix (Reduced representation)

Once the image format is selected then the values in each matrix are compressed separately using the DCT

In order to make the transformation more efficient a second step known as block preparation is carried out before DCT

In block preparation each global matrix is divided into a set of smaller 8X8 submatrices (block) which are fed sequentially to the DCT



IAT 3

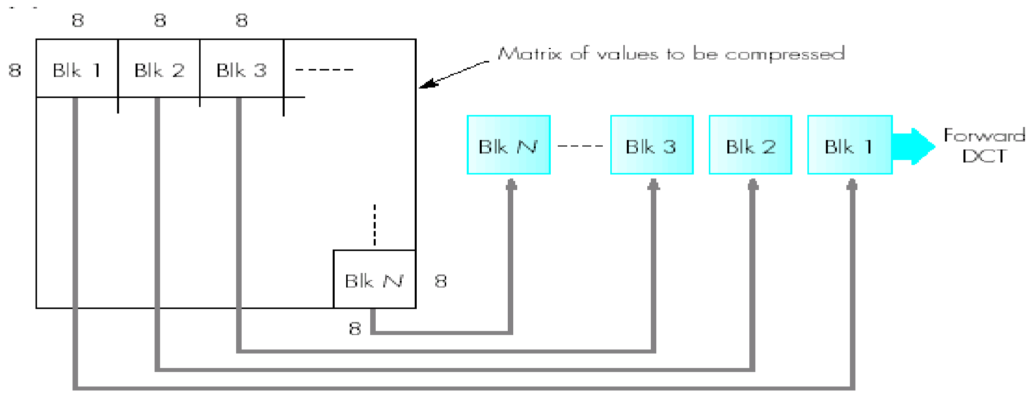
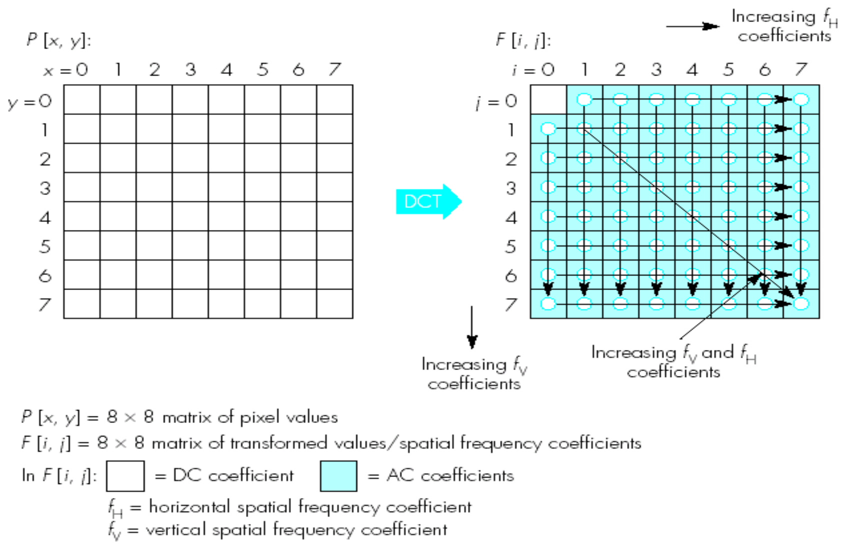
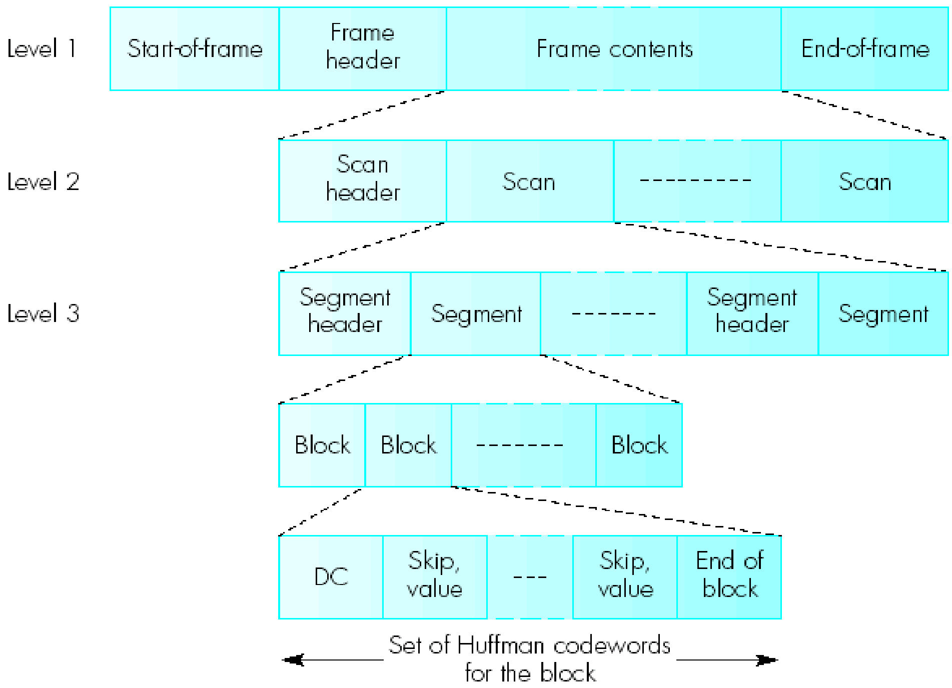


Image preparation



Using DCT there is very little loss of information during the DCT phase
The losses are due to the use of fixed point arithmetic
The main source of information loss occurs during the quantization and entropy encoding stages where the compression takes place
The human eye responds primarily to the DC coefficient and the lower frequency coefficients (The higher frequency coefficients below a certain threshold will not be detected by the human eye)
This property is exploited by dropping the spatial frequency coefficients in the transformed matrix (dropped coefficients cannot be retrieved during decoding)
In addition to classifying the spatial frequency components the quantization process aims to reduce the size of the DC and AC coefficients so that less bandwidth is required for their transmission (by using a divisor)
The sensitivity of the eye varies with spatial frequency and hence the amplitude threshold below which the eye will detect a particular frequency also varies
The threshold values vary for each of the 64 DCT coefficients and these are held in a 2-D matrix known as the quantization table with the threshold value to be used with a particular DCT coefficient in the corresponding position in the matrix
The choice of threshold value is a compromise between the level of compression that is required and the resulting amount of information loss that is acceptable
JPEG standard has two quantization tables for the luminance and the chrominance coefficients. However, customized tables are allowed and can be sent with the compressed image
Image compression

IAT 3



At the top level the complete frame-plus-header is encapsulated between a start-of-frame and an end-of-frame delimiter which allows the receiver to determine the start and end of all the information relating to a complete image

The frame header contains a number of fields

- the overall width and height of the image in pixels
- the number and type of components (CLUT, R/G/B, $Y/C_b/C_r$)
- the digitization format used (4:2:2, 4:2:0 etc.)

At the next level a frame consists of a number of components each of which is known as a *scan*

The level two header contains fields that include:

- the identity of the components
- the number of bits used to digitize each component
- the quantization table of values that have been used to encode each component

Each *scan* comprises one or more *segments* each of which can contain a group of (8X8) *blocks* preceded by a header

This contains the set of Huffman codewords for each block

2.Explain LPC encoder and decoder with a help of neat diagram

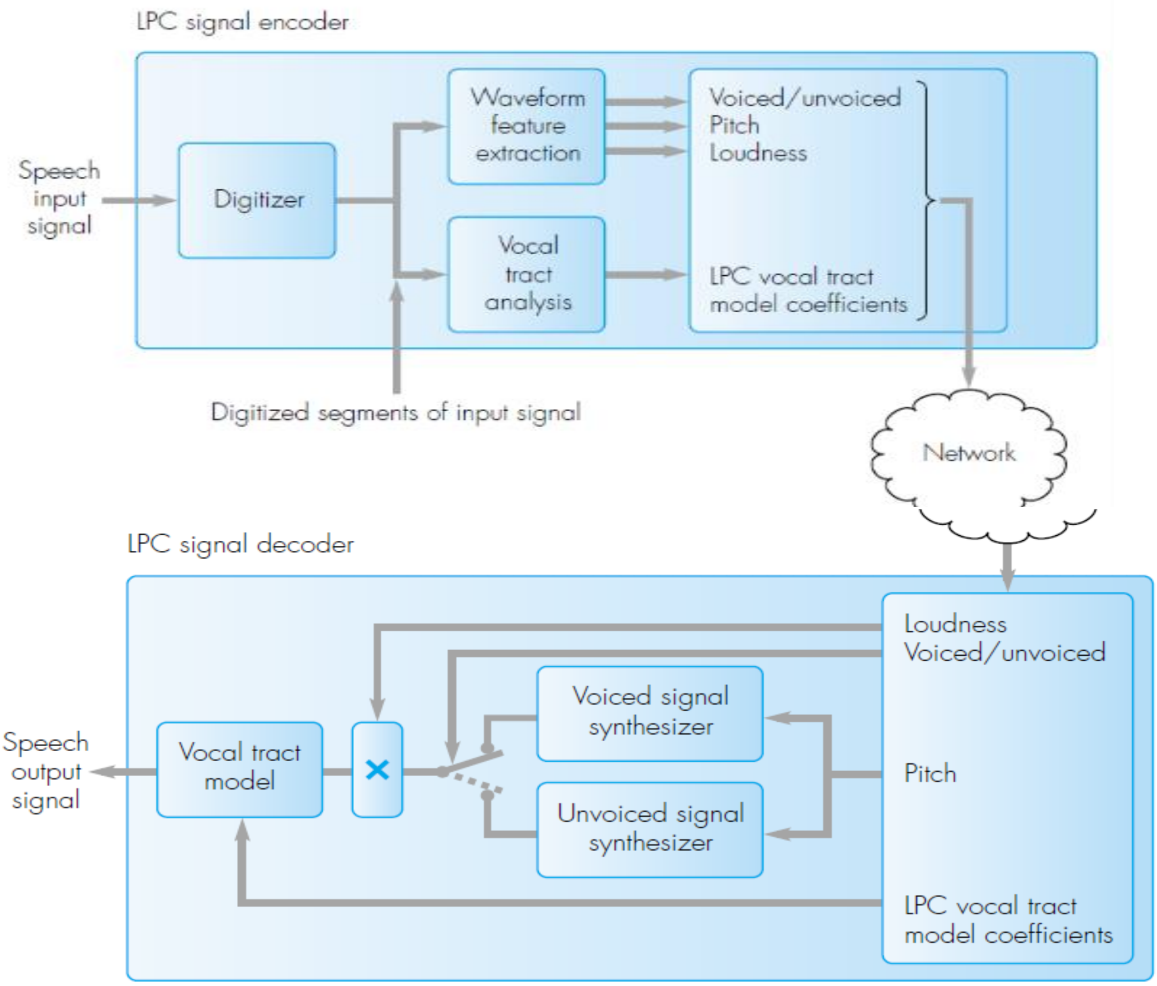
All algorithms must be sampling, digitization and quantization using DPCM / ADPCM
DSP crcuits help in analyzing the signal based on the required features (perceptual) and then quantized
Origin of sound is also important – vocal tract excitation parameters
Voiced sounds-generated through vocal chords
Unvoiced sounds – vocal chords are open
These are used with proper model of vocal tract to produce synthesized speech
After analyzing the audio waveform , These are then quantized and sent and the destination uses them,together with a sound synthesizer,to regenerate a sound that is perceptually comparable with the source audio signal.this is **LPC** technique.

Three feature which determine the perception of a signal by the ear are its:

IAT 3

- Pitch
- Period
- Loudness

The i/p waveform is first sampled and quantized at a defined rate
Segment- block of sampled signals are analyzed to define perceptual parameters of speech
The speech signal generated by the vocal tract model in the decoder is the present o/p signal of speech synthesizers and linear combination of previous set of model coefficients
The output of encoder is a set of frames ,each frame consists of fields for pitch and loudness



3,

IAT 3

A digitized video is to be compressed using the MPEG-1 standard. Assuming a frame sequence of:

IBBPBBPBBPBBI...

and average compression ratios of 10:1 (I), 20:1 (P) and 50:1 (B), derive the average bit rate that is generated by the encoder for both the NTSC and PAL digitization formats.

Answer:

Frame sequence = IBBPBBPBBPBBI...

Hence: 1/12 of frames are I-frames, 3/12 are P-frames, and 8/12 are B-frames.

and Average compression ratio $= (1 \times 0.1 + 3 \times 0.05 + 8 \times 0.02)/12$

$= 0.0342$ or 29.24:1

NTSC frame size:

Without compression $= 352 \times 240 \times 8 + 2\ (176 \times 120 \times 8)$

= 1.013760 Mbits per frame

With compression $= 1.01376 \times 1/29.24$

= 34.670 kbits per frame

Hence bit rate generated at 30 fps = 1.040 Mbps

4.Explain in detail the concept of streaming video across the internet.

IAT 3

6.2.4 Streaming Video across the Internet

Real-time transport of live video or stored video is the predominant part of real-time multimedia. On the other hand, video streaming refers to real-time transmission of stored video. There are two modes for transmission of stored video across the Internet: the download mode and the streaming mode. In the download mode, a user downloads the entire video file and then plays back the video file. However, full file transfer in the download mode usually suffers long and perhaps unacceptable transfer time. In the streaming mode, the video content need not be downloaded in full, but is being played out while parts of the content are being received and decoded. Due to its real-time nature, video streaming has bandwidth, delay and loss requirements. Designing mechanisms and protocols for streaming video pose

IAT 3

many challenges. Streaming video has six key areas: video compression, application-layer QoS control, continous media distribution services, streaming servers, media synchronization mechanisms and protocols for streaming media. Each of the six areas is a basic building block with which an architecture for streaming video can be built. The relations among these basic building blocks are illustrated in Figure 6.18 [6.64, 6.65]. Raw video and audio data are precompressed by video compression and audio compression algorithms and saved to storage devices. Upon the client's request, a streaming server retrieves compressed audio/video data from storage devices. Then, the application layer QoS control module adapts the audio-video bit streams according to the network states and QoS requirements. After the adaptation, the transport protocols packetize the compressed bit streams and send the audio-video packets to the Internet. Packets may be dropped as they experience excessive delays inside the Internet due to congestion. To improve the quality of audio-video transmission, continuous media distribution services are developed for the Internet. For packets that are successfully delivered to the receiver, they first pass through the transport layers and then are processed by the application layer before being decoded at the audio-video decoder. To achieve synchronization between video and audio presentations, media synchronization mechanisms are required. As it can be seen, these areas are closely related, and they are coherent constituents of the video streaming architecture.

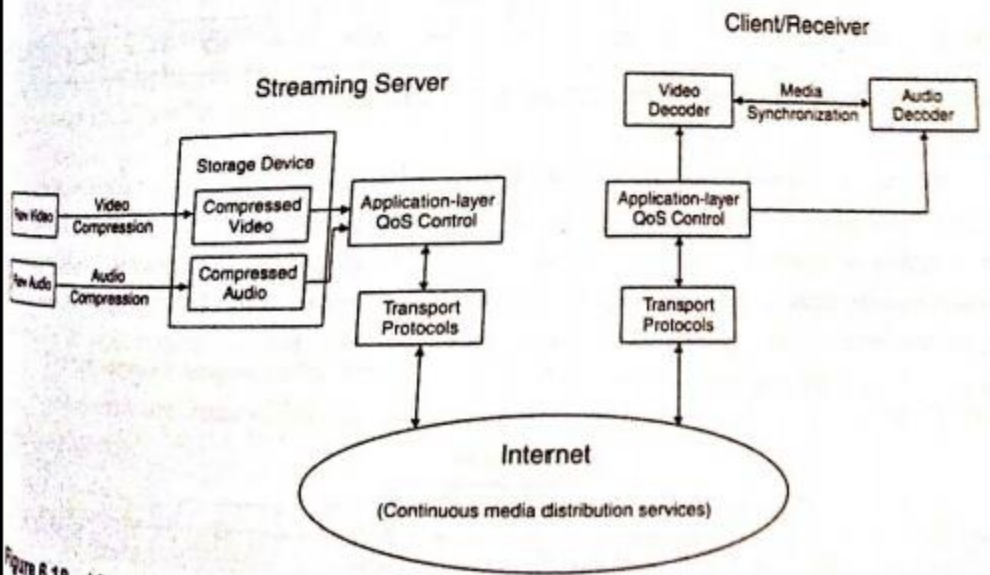


Figure 6.18 Video-streaming architecture [6.65]. ©2001 IEEE.

5.Explain the principal of video compression with the neat sketch.

The digitization format defines the sampling rate that is used for the luminance ,Y ,and two chrominance,Cb and Cr
A technique used is based on combination of preceeding and succeeding frame .
Instead of video as set of compressed frames, difference between actual frame and predicted frame contents is sent- motion estimation and motion compensation .
Frame type
I-frame- Intracoded
I-frames are encoded without reference to any other frames
GOP:The number of frame between successive I-frames
P-frame:intercoded

encoding of a p-frame is relative to the contents of either a preceding I-frame or a preceding P-frame
The number of P-frames between I-frame is limited since any errors present in the first P-frame will be propagated to the next
B-frame:their contents are predicted using search regions in both past and future frames
PB-frame:this does not refer to a new frame type as such but rather the way two neighboring P- and B-frame are encoded as if they were a single frame
D-frame:only used in a specific type of application. It has been defined for use in movie/video-on-demand application

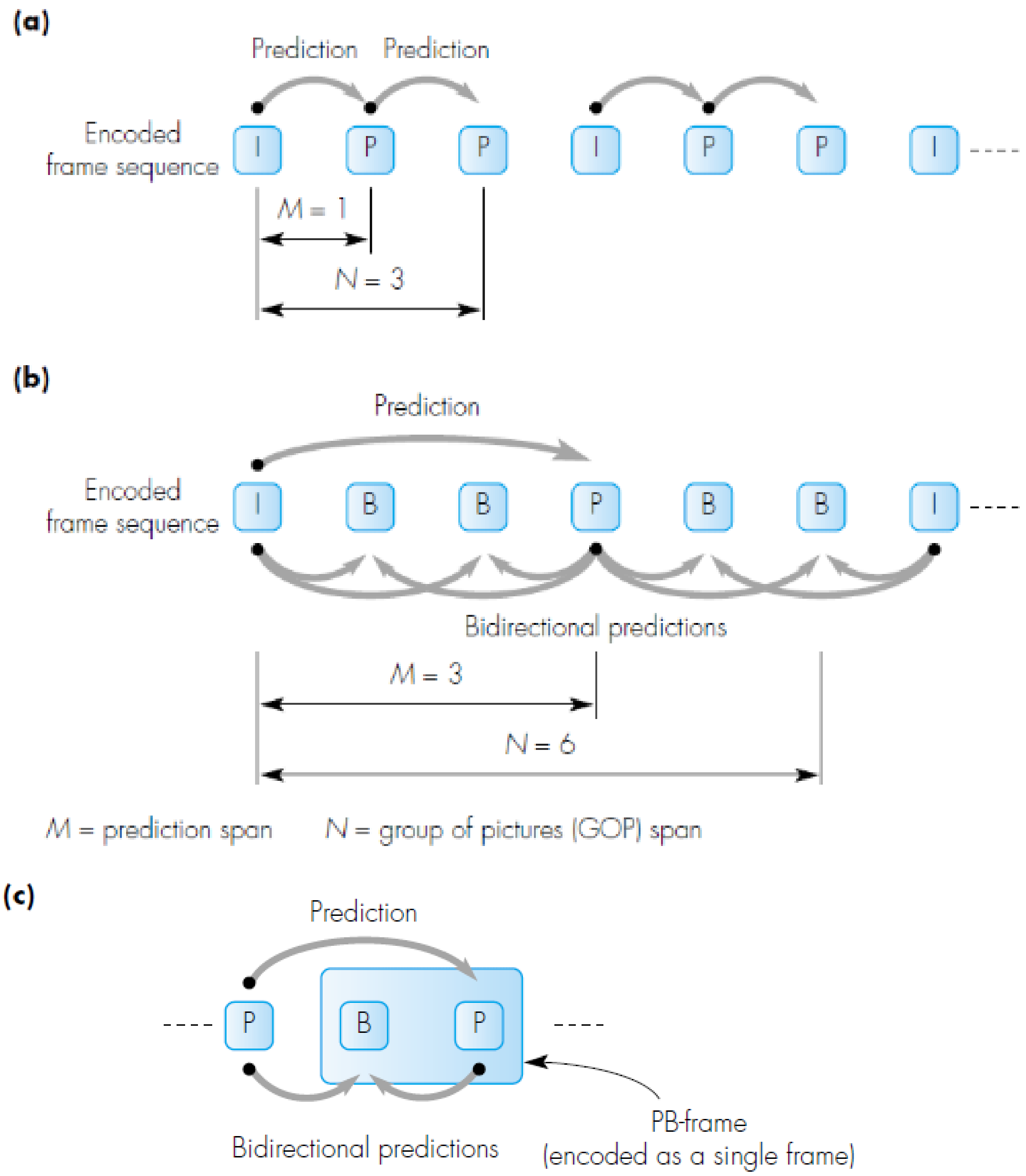
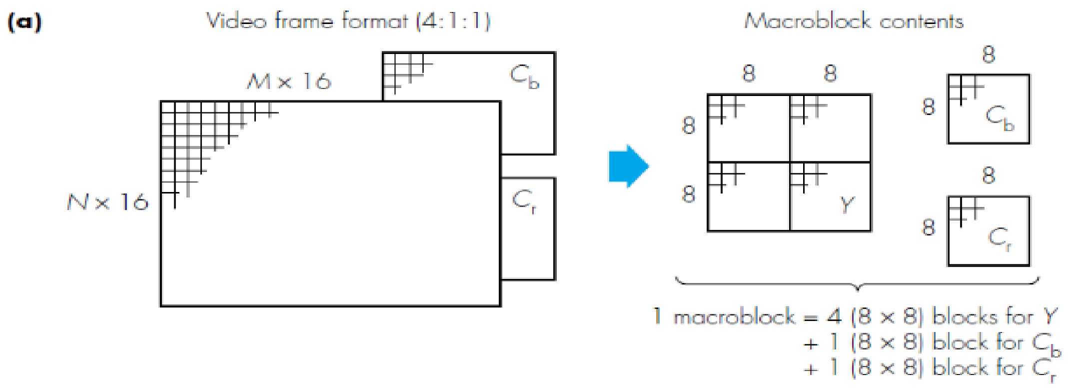


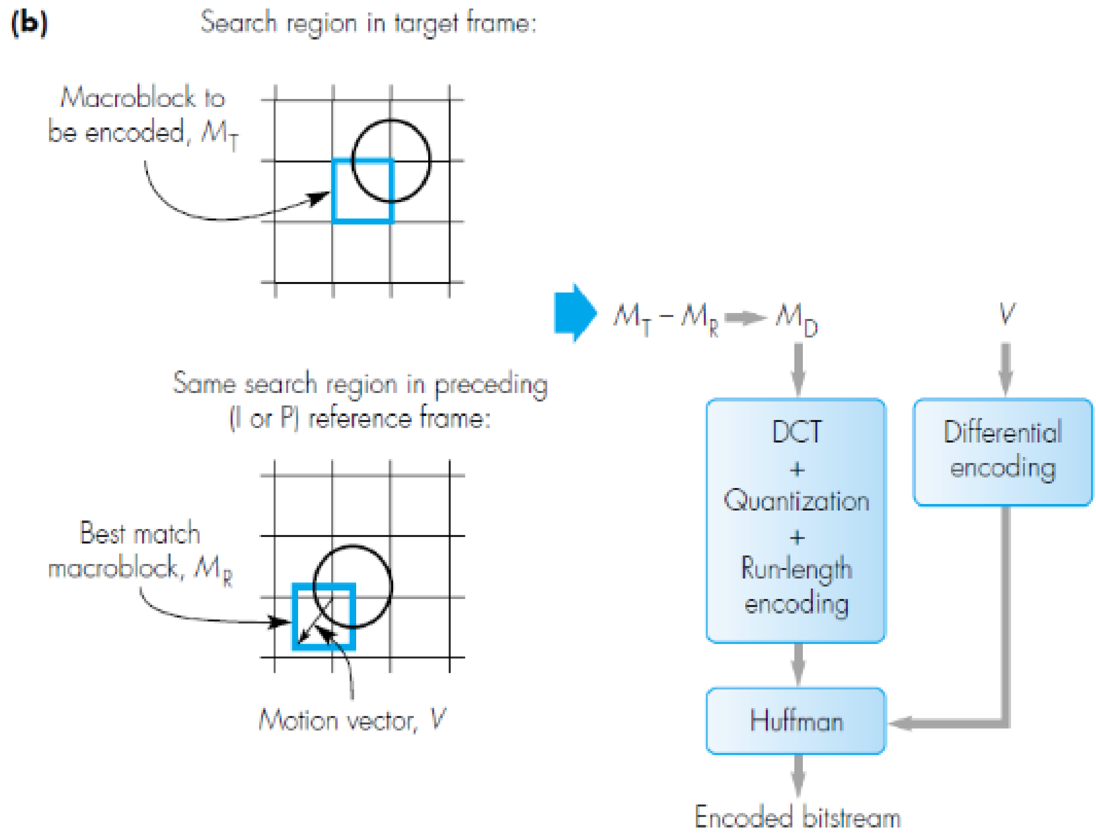
Figure 4.11 Example frame sequences with: (a) I- and P-frames only; (b) I-, P- and B-frames; (c) PB-frames.

MOTION ESTIMATION AND COMPENSATION

IAT 3

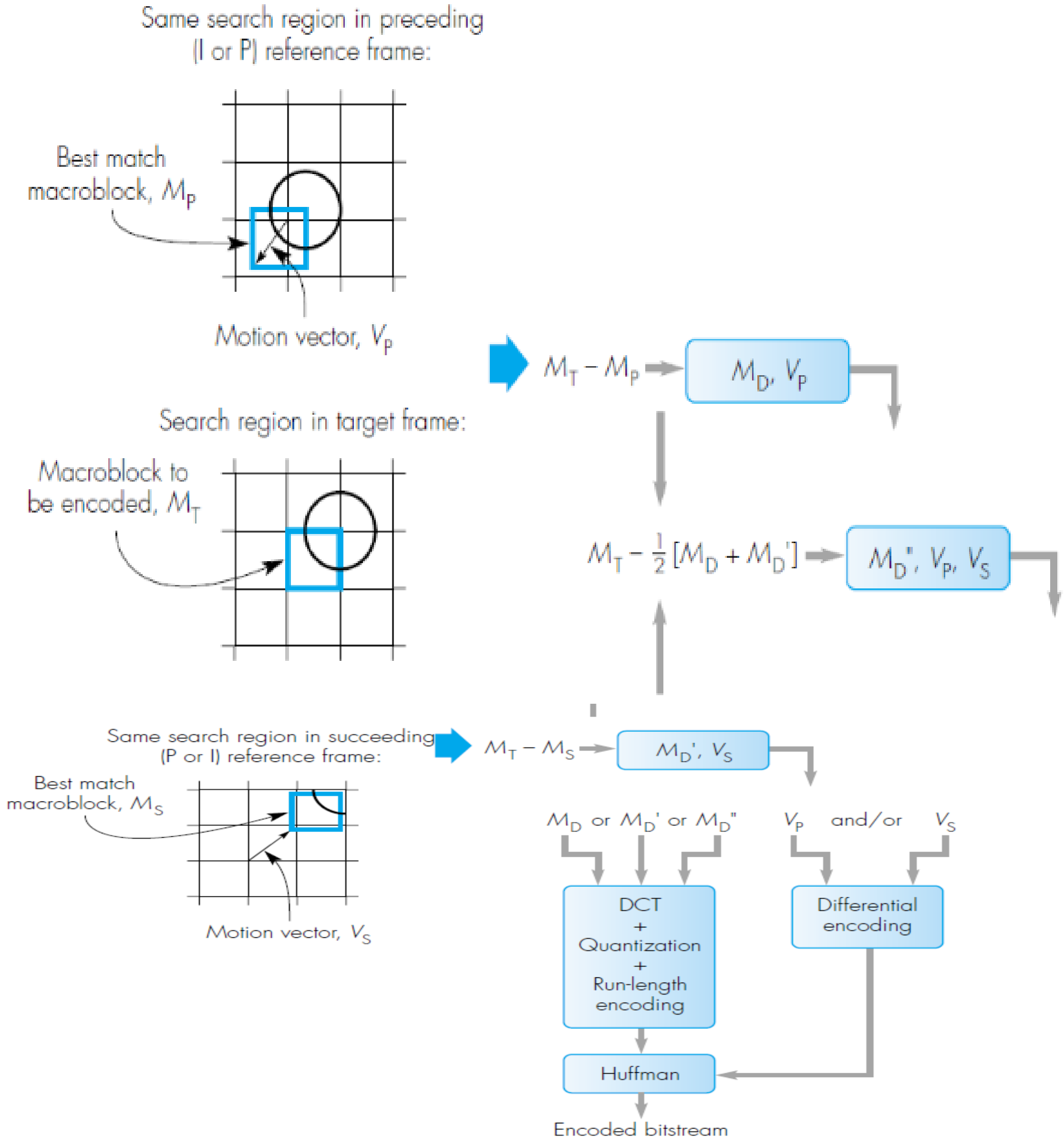
The encoded contents of both p and B frames are predicted by estimating any motion that has taken place between the present frame and the preceeding I or P frame and in B frames the succeeding P or I frames.
The digitized contents of Y matrix of each frame is divided into two dimensional matrix (16x16) pixels – macroblock
4:1:1 is considered and Cb Cr will be 8x8 pixels
Block size for DCT is also 8x8, A macroblock contains 4 DCT blocks for Y and one each for two chrominance signals
P frame encoding
For encoding p frame, the contents of each macroblock in the target frame is compared with the corresponding macroblock in the I or Pframe – reference frame.
If there is a close match then the address of the macroblock is coded else the search is continued for the nearby macroblock





IAT 3

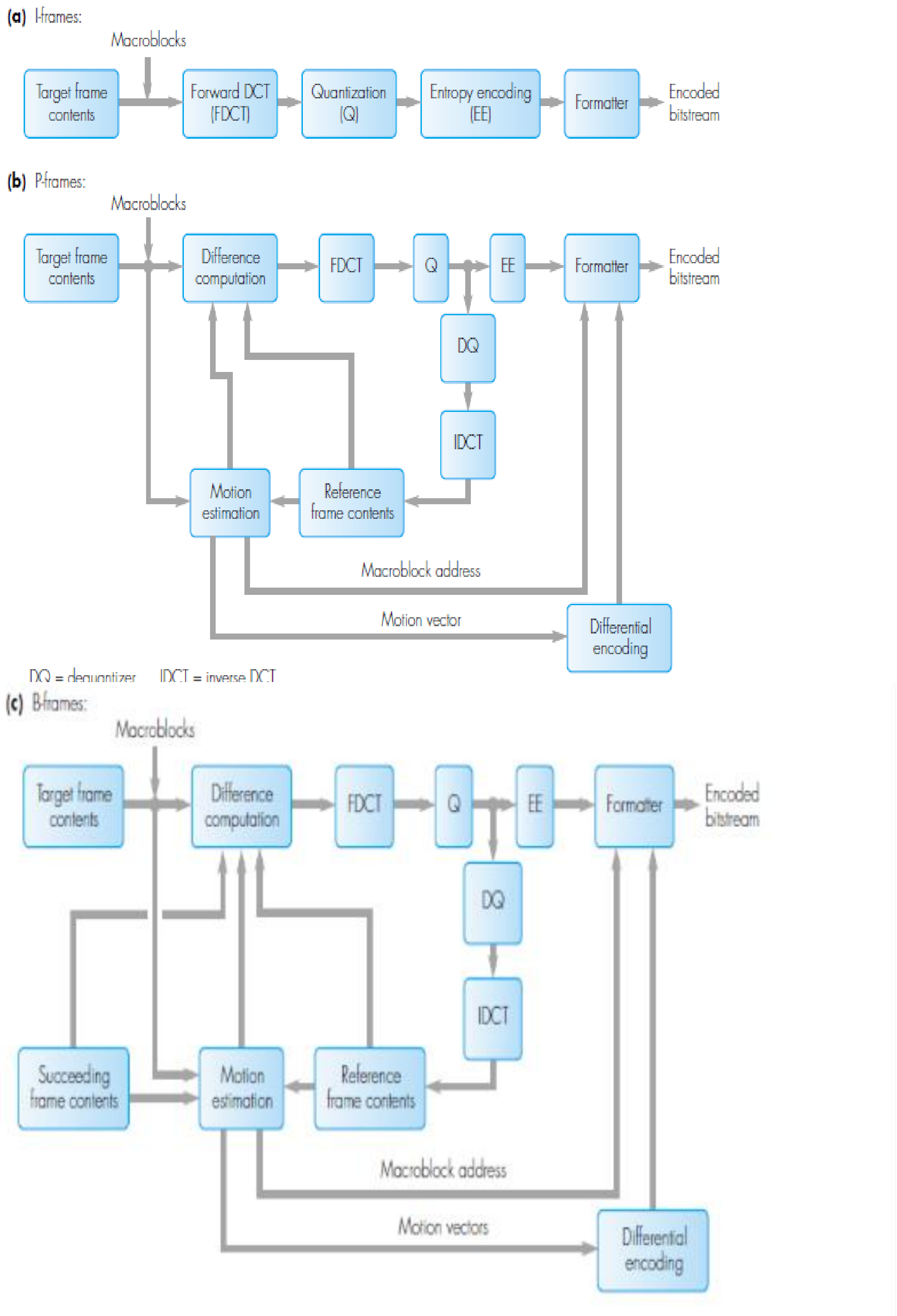
If a close match is found then two parameters are encoded. Motion vector, the(x,y) offset of the macroblock being encoded, and the location of the block of pixels in the ref frame and prediction error
Offset can be on macroblock or pixel boundary
Mv is known as the single pixel resolution.
prediction error- three matrices for Y cb cr each containing the difference values between target MB and the set of pixels in the search area which produced the close match.
MVs are encoded using Differential encoding and resulting codeword are huffmann coded.
B FRAME ENCODING
Motion estimation is with preceeding Ior Pframe and the immediate succeeding I or P frame
Mv and the difference matrices are computed using first preceeding frame as ref and then succeeding frame as ref.
Third mv and the difference is calculated using target and the mean of the two predicted values.
The set with least difference matrices is selected and encoded similar to p frame.
Mv resolution is termed as half pixel resolution



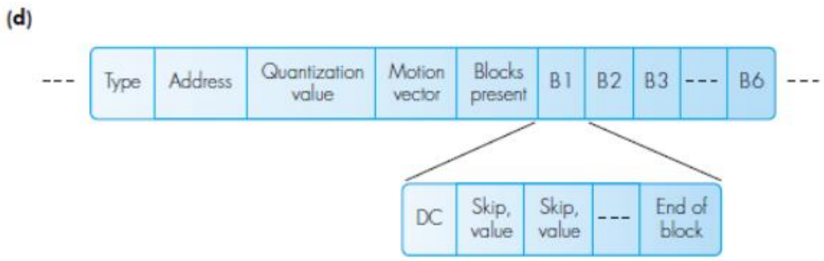
For four block of Y, two for chrominance, each macroblock requires 8x8 pixel blocks to be encoded.

IAT 3

For p frames the encoding of each macroblock is dependent on the output of motion estimation
For I frame three steps are : Forward DCT, Quantization and entropy coding
which depends on contents of macroblock encoded and the contents of macroblock in the search area of ref frame which produces closest match.
Three possibilities are:
If the two contents are same, only address in the macroblock in ref frame is encoded
If the two contents are very close, both the mv, the difference matrices are encoded.
If no match is found ,target macroblok encoded similar to I frame



IAT 3

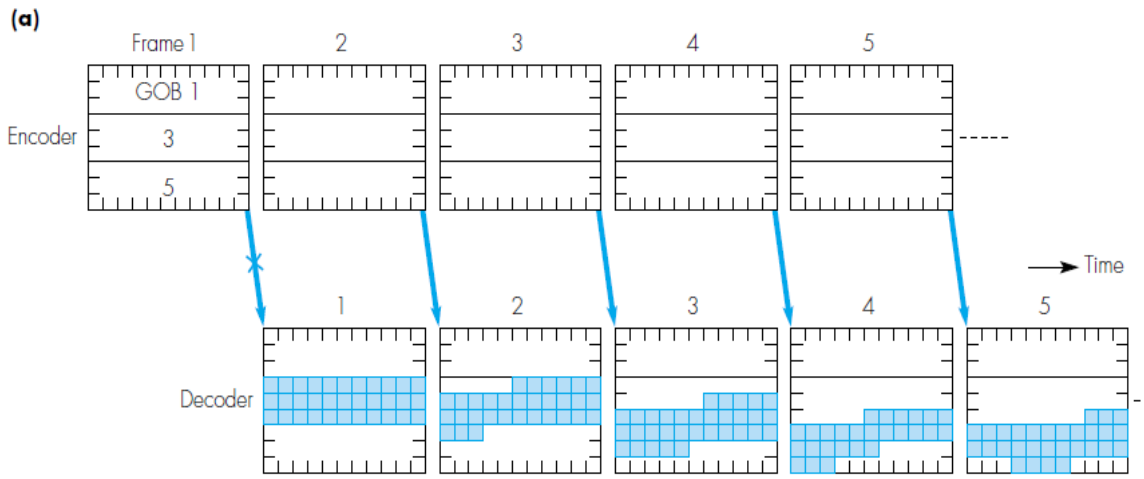


Motion estimation contains the search logic – uses computed difference values, considering
Target frame and ref frame contents, decompresses by dequantizes and IDCT block
After the target frame is completely compressed the difference values are used to update the ref frame contents for next frame encoding
The type of encoding for each macroblok is identified by formatter
Typical format is as shown
Type field indicates the type of frame being encoded- I/P / B
Address identifies the location of maroblock in the frame
Quantiztion value is the threshold value, to quantize all DCT coeffecients
Motion vector is the encoded vector
Blocks – the six 8x8 block that make up the macroblock
Decoding of received bitstream is simpler as it does not require estimation
At the receipt of the bitstream, each new frame is assembled a macroblock at a time
Decoding of I frame is same as JPEG
To decode p frame, the decoder keeps the copy of the preceeding I or P frame in a buffer and uses it along with encoded information of each macroblock to build the Y, Cb, Cr matrices for new frame in second buffer.
With uncoded macroblocks, the macroblocks address is used to locate the macroblocks in the previous frame and its contents are transferred to second buffer
With fully encoded macroblocks these are decoded directly and contents sent to buffer
For macroblocks with mv, set of difference matrices , they are together with matrices in first buffer are used to define values of macroblocks in second buffer.
For Bframe decoding 3 buffers are used.

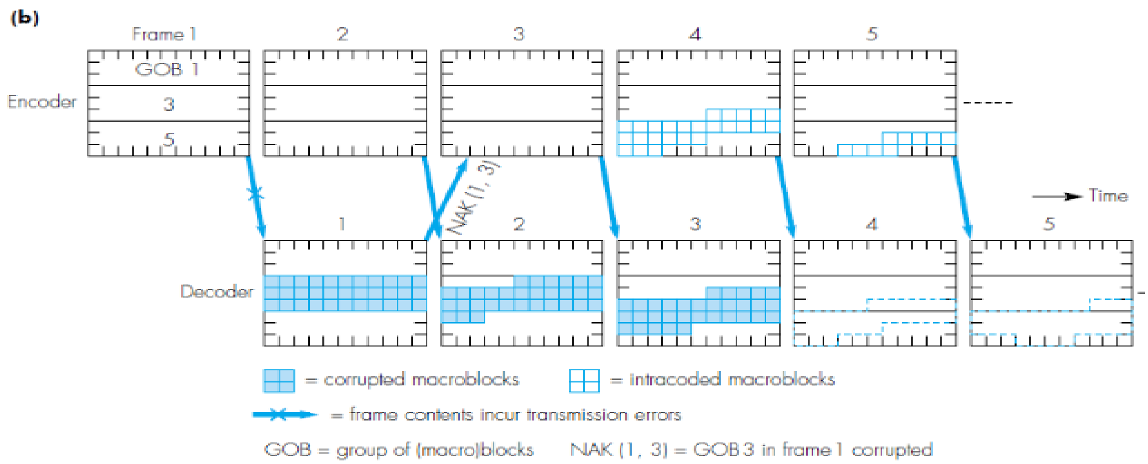
6.Explain the error tracking procedures of H.263 with a neat diagram.

Error resilience
Cause error propagation,show figure4.17(a)
For PSTN , errors present in bitstream is more
Difficulty in finding the error macroblock
-GOB(group of macroblocks) may contain any error macroblock also.
When error in GOB is detected it skips the remaining macroblocks in the affected GOB and finds resynchronization marker.
Masking of error – error concealment scheme.

IAT 3



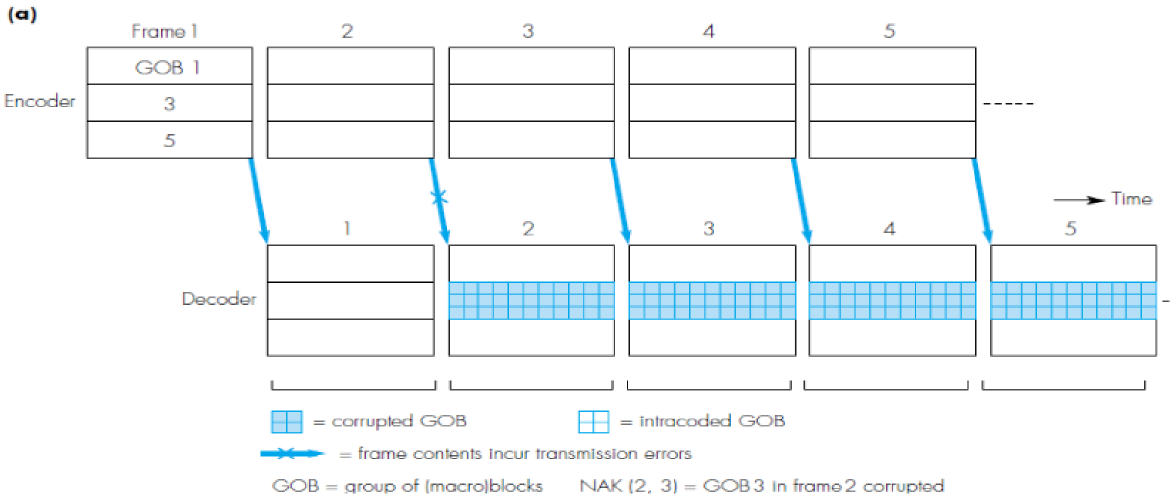
It leads to error propagation to other regions of the frame
To avoid this the schemes used are:
Error tracking
Independent segment decoding
Reference picture selection
Error tracking and resilience,show figure4.17(b)
For the information to encode regarding error in GOB a two way channel communication is used
Error detection types:
One or more out of range mvs
One or more invalid variable length codewords
One or more out of range DCT coefficients
An excessive no of coefficients within a macroblock
It has error prediction information of all GOBs in the recently transmitted frames.
When an error is detected , decoder send NAK(negetive ack) to encoder in the source code with frame no, location of GOB in frame in error.
It identifies the macroblocks to be likely affected in the later frames
Affected macroblocks are intracoded.



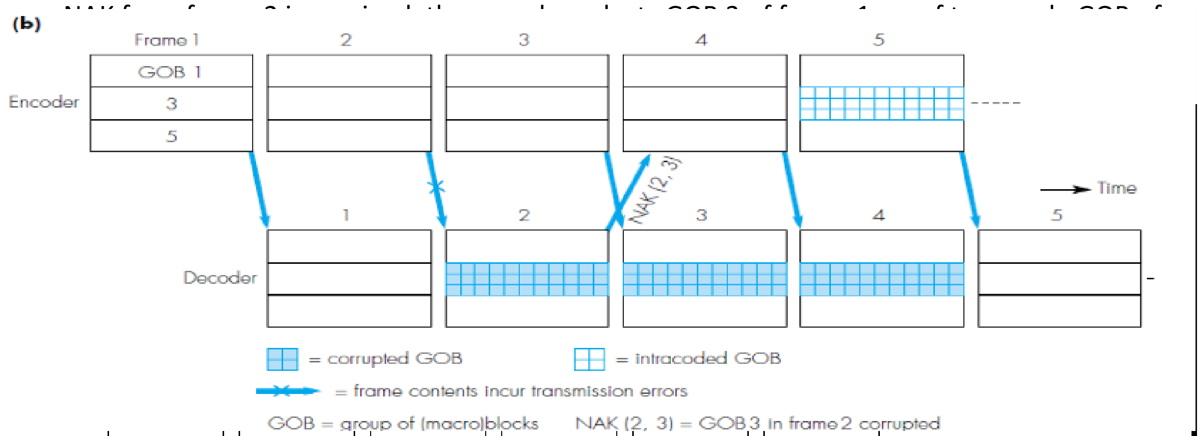
Independent segment decoding
Prevent these errors from affecting neighboring GOBs in succeeding frames
Show figure 4.18
Motion estimation and compensation is with ref boundary pixels of GOB.

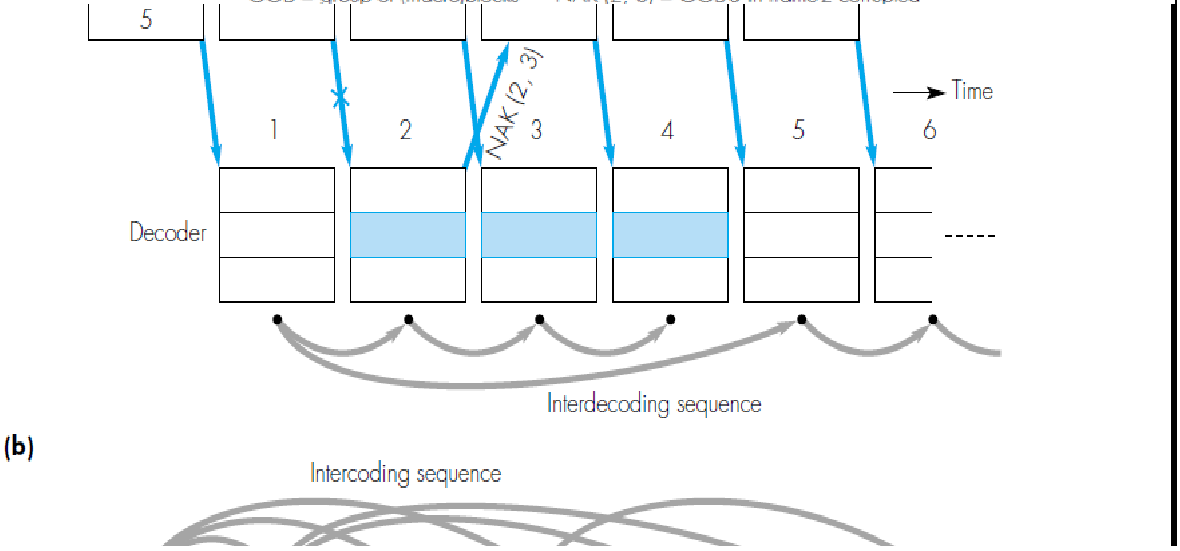
IAT 3

An error in GOB will affect the same GOB in successive frames till anew intracoded GOB is sent by the encoder.
Used in conjunction with other schemes,.

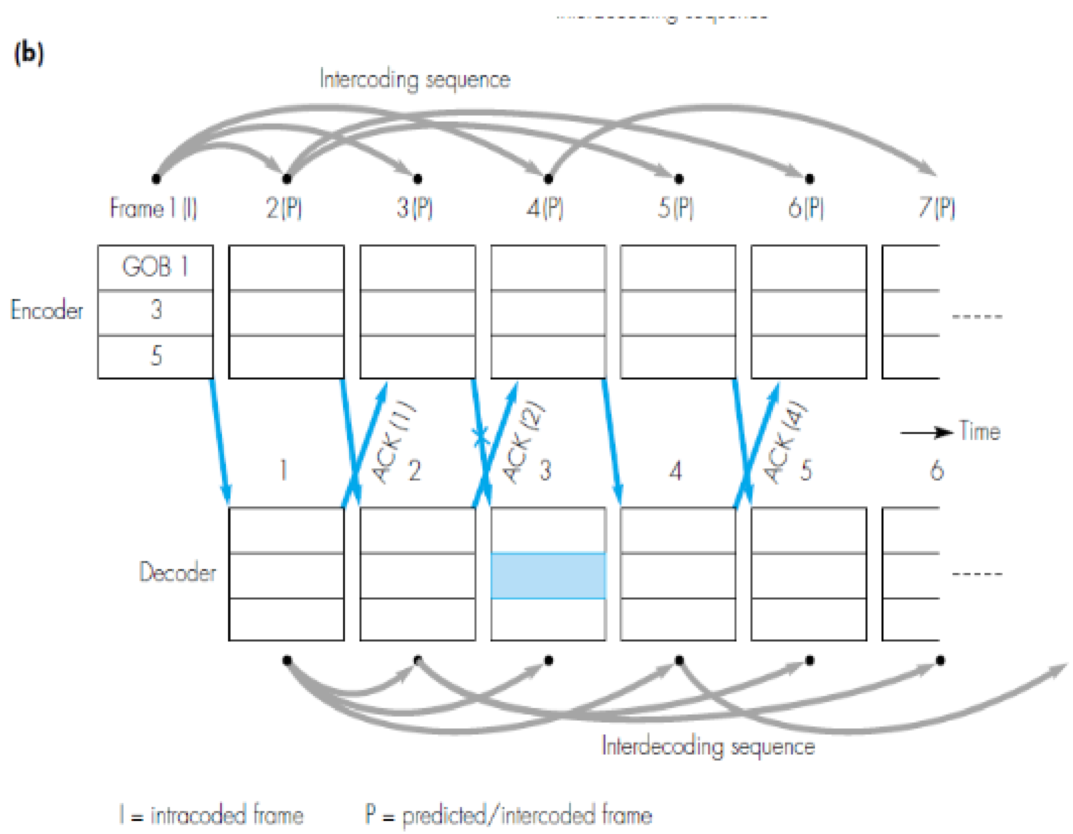


Similar to error tracking scheme
 DECODER sends ack messages to avoid error propagation
During encoding of intercoded frames a copy of preceding frame is retained in the encoder
The encoder can select any of the previously decoded frame as ref.





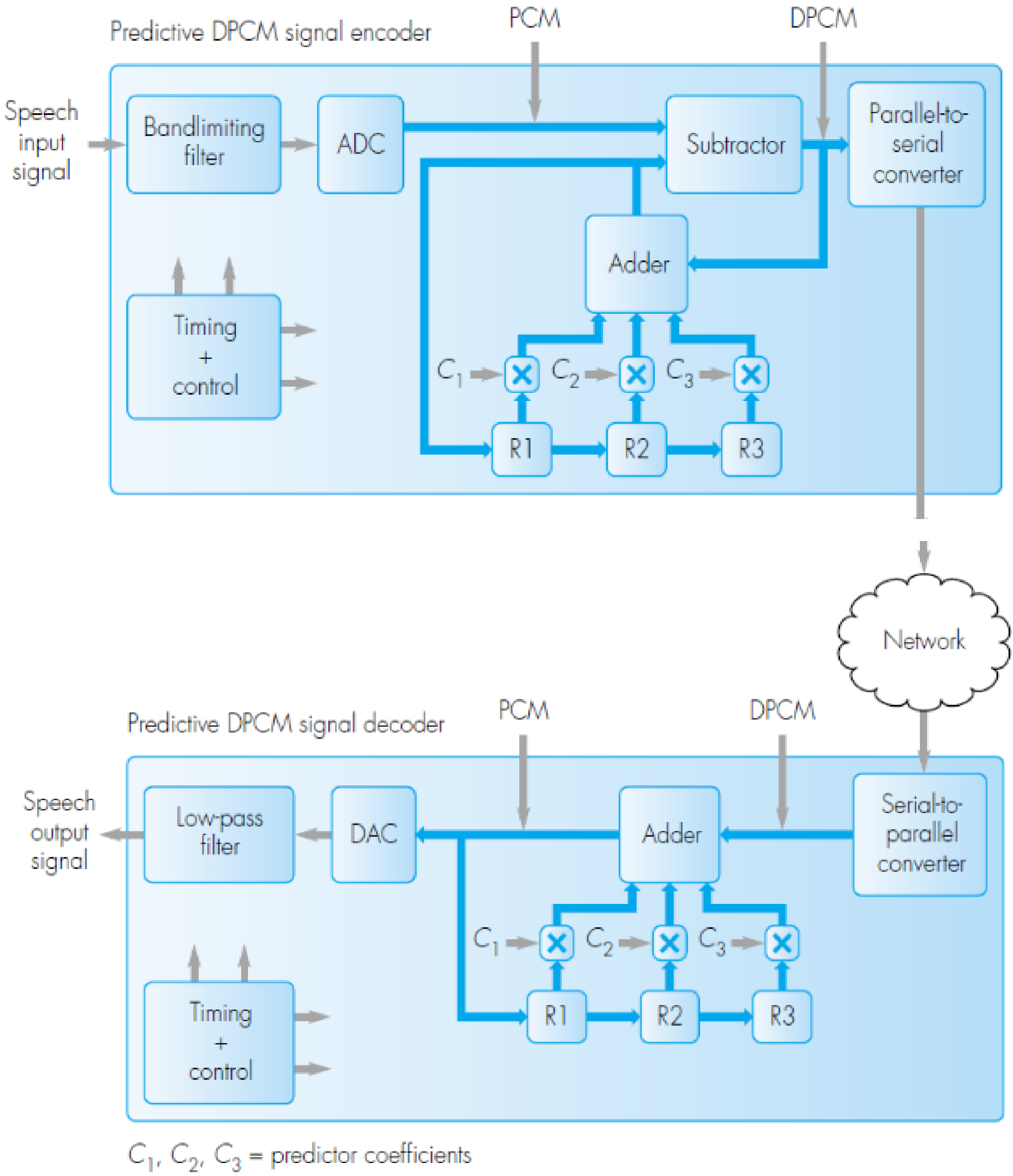
IAT 3



7.Explain the 3rd order predictive DPCM signal encoder and decoder

for most audio signals, the range of the differences in amplitude between successive samples of the audio waveform is less than the range of the actual sample amplitudes
The previous digitized sample value is held in reg R
Difference signal is by subtracting (Ro) from the digitized sample of ADC
Reg R is updated with the difference signal
The decoder adds the DPCM with previously computed signal in the reg
The o/p of ADC is also known as residual
There are schemes to predict the more accurate previous signal
The proportions used are determined by predictor co-efficients

IAT 3



8.Illustrate with a neat diagram how error and losses are handled in ATM networks.

IAT 3

6.3.3 Errors and Losses in ATM

The encoding process introduces controlled amounts of distortion in order to compress the signal. The video signal will also be exposed to bit errors induced in the electronics and in the optics. The probability of bit error is low, below 10^{-8}, but not negligible. More troublesome is the information loss in the ATM network when full stretches of the signal are deleted. The causes of loss are transmission burst errors, loss of cells and packets due to multiplexing overload, misrouting due to inaccurate addresses or entries in address tables and delay greater than the acceptable threshold. Undetected loss in a signal can place encoders and decoders out of phase. Burst errors caused by loss of synchronization and by equipment failures have durations of 20 to 40 ms. Their probability of occurrence has been estimated to be below 10^{-7} [6.94]. Loss, especially due to multiplexing overloads, appears to be the most common signal corruption caused by the ATM network.

Error recovery is based on limited error propagation and correction or concealment of the missing portion of the signal. Error propagation is restricted by proper framing of the bit stream so that errors and loss can be detected [6.99].

Generally speaking, in the ATM network, a cell can be lost due to two reasons:

- Channel errors
- Limitations of network capacity and statistical multiplexing

A communication channel is subject to different impairments. If an uncorrectable error occurs in the address field of an ATM cell, the cell will not be delivered to the right destination. This cell is considered to be lost. This is a rare cause of loss in ATM networks.

An ATM network takes advantage of statistical multiplexing, but also takes the risk of simultaneous traffic peaks of multiple users. Although a buffer can be used to absorb the instan-

taneous traffic peak to some extent, there is still a possibility of buffer overflow in case of congestion. In the case of network congestion or buffer overflow, the network congestion control protocol will drop cells. The malfunction or inefficient network management will also cause the cell loss. For example, loss of synchronization and lack of recovery measures in the physical layer would result in a stream of cell losses in the resynchronization/acquisition phase.

In an ATM network, cell discarding can occur on the transmitting side if the number of cells generated are in excess of allocated capacity, or it can occur on the receiving side if a cell has not been received within the delay time of the buffer memory. Cells can be discarded in the ATM network by the congestion control procedure.

If the incoming traffic exceeds allocated capacity and causes the buffer overflow, the sender could be informed by the network traffic control protocol to reduce the traffic flow or to switch to a lower grade service mode by subsampling and interlacing.

If the network becomes congested and the input buffer overflows, it will drop some cells to reduce the traffic and to assume the normal communication phase.

If the error occurs in the cell header, especially in the address field, the cell may be misdelivered or go astray in the network. In the receiver, if a cell is not received within the maximum time out window, the cell is considered to be lost. The loss of a cell leads to the loss of 384 consecutive bits, which may cause a serious degradation in picture quality for VBR compressed video signals. If the cell loss is caused by network congestion, a few consecutive cells, which contain thousands of bits of information, may be lost. Furthermore, the cell loss may affect the subsequent frames if an interframe coding scheme is employed. Therefore, cell loss is a major problem encountered in VBR coding in the ATM environment. A cell loss may cause the loss of code synchronization. Because a variable number of data is packed into a cell, there is no way of knowing how much information is lost when a cell loss occurs unless some side information is available. Cell loss can occur unpredictably in ATM networks. It is assumed to be random with the probability of cell loss depending only on whether a previous cell of the same priority was lost.

Asynchronously multiplexed networks, such as those based on ATM, have cells and packets as multiplexing units that are shorter than a full cell (session). The multiplexing unit in traditional Time Division

IAT 3

... location and unknown values).

It is important that the sequence number is based on the number of transferred data octets. Knowing that a cell or a packet has been lost does not tell how much data it contained.

Errors and loss can be identified by a CRC on the application frame after reassembly. It is important that frame length is known a priori because the length of a faulty frame cannot be ascertained. The failed CRC could be caused by a bit error, which would not be affected by its length or by a lost packet or cell.

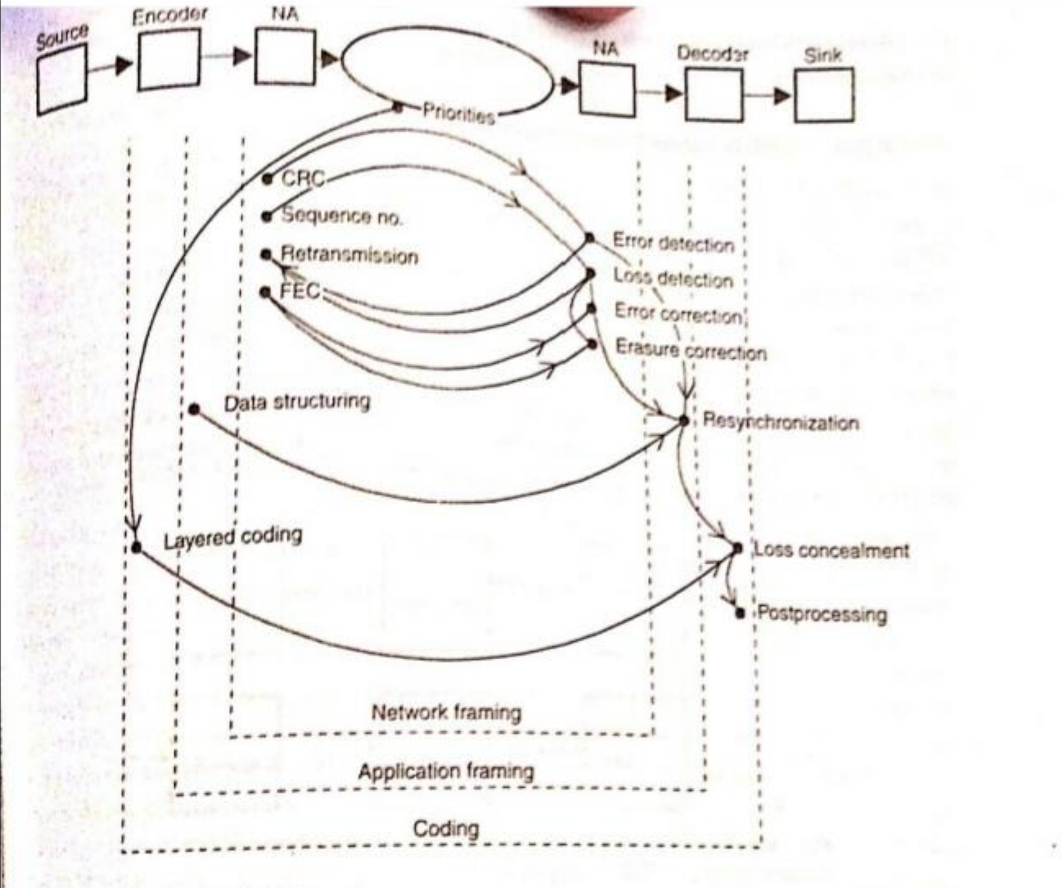


Figure 6.34 Errors and loss handling [6.90]. ©2000 Prentice Hall.

A lost or corrupted network frame would, in case of regular data communication, be retransmitted. There are complications with the use of retransmissions for video. First, the delay requirements might not allow it because it adds at least another round-trip delay that is likely to violate end-to-end delay requirements for conversational services. Second, the jitter introduced is much higher than that induced by queuing. Delay equalization is thus further complicated. Even if this would be acceptable, the continuously arriving datastream must be buffered until the missing frame eventually is received.

There are several reasons to be cautious of FEC of cell and packet loss. First, it adds a fairly complex function to the system, which will be reflected in its cost. Second, the interleaving adds delay. Third, loss caused by multiplexing overload is likely to be correlated because the overload is caused by traffic bursts and more loss may occur than what the code can correct. If an interleaving matrix cannot be corrected, then the full matrix is useless, and the loss situation is in fact made worse.

IAT 3