CBCS SCHEME

USN | | | | | | | | | |

17EC741

Seventh Semester B.E. Degree Examination, Feb./Mar. 2022
Multimedia Communication

Time: 3 hrs. Max. Marks: 100

Note: Answer any FIVE full questions, choosing ONE full question from each module.

Important Note : 1. On completing your answers, compulsorily draw diagonal cross lines on the remaining blank pages.
2. Any revealing of identification, appeal to evaluator and /or equations written eg, 42+8 = 50, will be treated as malpractice.

Module-1

1 a. List the five basic types of communication network that are used to provide multimedia services. Explain with a neat diagram:
(i) Data Networks
(ii) Integrated Services Digital Network **(10 Marks)**

b. Explain the principle of operation of packet switched network with neat diagrams. **(07 Marks)**

c. Derive the maximum block size that should be used over a channel which has BER probability of 10^{-4} if the probability of a block containing an error and hence being discarded is to be 10^{-1}. **(03 Marks)**

OR

2 a. Explain with neat diagrams, Movie on Demand and Near Movie on Demand (MOD/N-MOD) application. **(08 Marks)**

b. Explain the operational modes of multipoint conferencing with neat diagrams. **(06 Marks)**

c. Determine the propagation delay associated with the following communication channels:
(i) A connection through a private telephone network of 1 km
(ii) A connection through a PSTN of 200 km
(iii) A connection over a satellite channel of 50,000 km
Assume velocity of propagation of a signal in the case of (i) and (ii) is 2×10^8 m/sec and in the case of (iii) is 3×10^8 m/sec. **(06 Marks)**

Module-2

3 a. Explain the principle of operation of PCM speech CODEC with a block diagram. Also explain compressor and expander. **(08 Marks)**

b. Explain Interlaced Scanning principle with a diagram. **(06 Marks)**

c. Derive the bit rate and the memory requirements to store each frame that result from the digitization of a 525 line system assuming a 4:2:2 format. Also find the total memory required to store a 1.5 hour movie/video. **(06 Marks)**

OR

4 a. With the aid of diagram, explain the following:
(i) Aspect ratio of display screen
(ii) Raster scan
(iii) 4:2:2 **(08 Marks)**

b. Explain different types of text in detail. **(06 Marks)**

c. Assuming the bandwidth of a speech signal is from 50 Hz through to 10 kHz and that of a music signal is from 15 Hz through to 20 kHz, derive the bitrate that is generated by the 'digitization procedure in each case assuming the Nyquist sampling rate is used with 12 bits per sample for the speech signal and 16 bits per sample for the music signal. Derive the memory required to store a 10 minute passage of stereophonic music. **(06 Marks)**

Module-3

5 a. A message comprising of a string of characters with probabilities e = 0.3, n = 0.3, t = 0.2, w = 0.1, · = 0.1 is to be encoded. The message is "went." Compute the arithmetic code word. **(08 Marks)**

b. With the aid of diagrams, explain JPEG encoder. **(08 Marks)**

c. Explain CPU management in multimedia operating system. **(04 Marks)**

OR

6 a. A message and its probability of occurrence of each character is as follows:
A and B = 0.25, C and D = 0.14, E, F, G and H = 0.055.
(i) Use Shannon's formula to derive the minimum average number of bits per character.
(ii) Construct the Huffman code tree and derive a suitable set of code word. **(08 Marks)**

b. Explain the principle of LZW compression. **(06 Marks)**

c. Explain the main features of distributed multimedia system. **(06 Marks)**

Module-4

7 a. Explain Linear Predictive coding encoder and decoder with neat schematic. **(08 Marks)**

b. A digitized video is to be compressed using the MPEG-1 Standard. Assuming a frame sequence of I BBP BBP BBP BBI… 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 NJSC and PAL formats. **(08 Marks)**

c. Explain different frame types. **(04 Marks)**

OR

8 a. Explain DPCM encoder and decoder with a neat diagram. **(10 Marks)**

b. What do you understand by the terms:
(i) Group of pictures (ii) Prediction span (iii) Motion compensation
(iv) Motion estimation (v) Temporal masking **(10 Marks)**

Module-5

9 a. Explain scalable rate control with a neat block diagram. **(10 Marks)**

b. Explain video streaming architecture with a neat diagram. **(10 Marks)**

OR

10 a. Discuss briefly about Integrated Packet Networks. **(10 Marks)**

b. Explain briefly about errors and losses in ATM. **(10 Marks)**

* * * * *

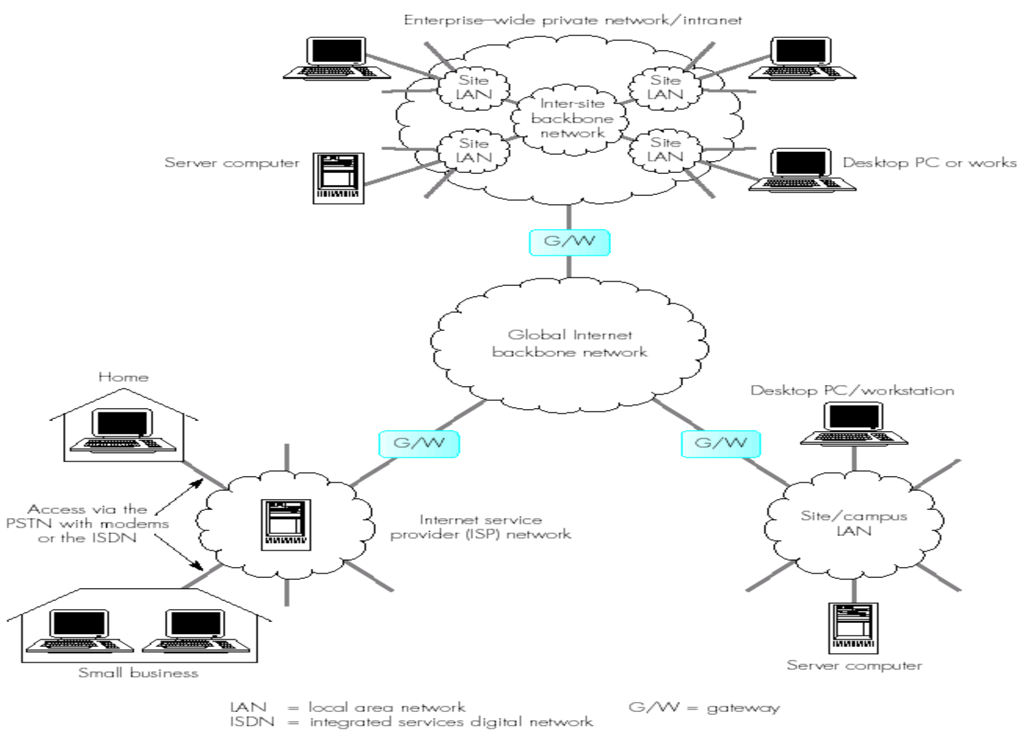
Module 1

1. (a) Five types of Multimedia Networks are:

- Telephone Networks - Telephony
- Data Networks – Data Communications
- Broadcast Television Networks – Broadcast TV
- Integrated services digital Network
- Broadband Multiservice networks

Data Networks

- Designed to provide basic data communication services such as email and general file transfer
- Most widely deployed networks: **X.25 network** (low bit rate data) not suitable for multimedia and the **Internet** (Interconnected Networks)
- ***Communication protocol***: set of rules (defines the sequence and syntax of the messages) that are adhered to by all communicating parties for the exchange of information/data
- ***Packet***: Container for a block of data, at its head, is the address of the intended recipient computer which is used to route the packet through the network
- **Open systems interconnections (OSI)-** is a standard description or "reference model" for how messages should be transmitted between any two points in a telecommunication network
- Access to homes is through an Internet Service provider (ISP)
- Access through PSTN or ISDN (high-bit rate)



- Business users obtain access either through site network or through an enterprise-wide private network (multiple sites)

- Universities with single campus use a network known as the Local Area Network (LAN). However bigger universities with more than one campus use enterprise wide network
- If the communication protocols of the computers on the network are the same as the internet protocols then the network is known as an **intranet** (e.g large companies and universities)
- All types of network are connected using a **gateway** (router) to the **internet backbone network**
- **Router -** a router is a device or, in some cases, software in a computer, that determines the next network point to which a **packet** should be forwarded toward its destination
- **Packet mode –** Operates by transfer of packets as defined earlier
- This mode of operation is chosen because normally the data associated with data applications is in discrete block format.
- With the new multimedia PCs packet mode networks are used to support in addition to the data communication applications a range of multimedia applications involving audio video and speech

Integrated Services Digital Networks

- Started to develop in the early 1980s to provide PSTN users the capability to have additional services

- *Integrated Services Digital Network (ISDN)* in concept is the integration of both analogue or voice data together with digital data over the same network.

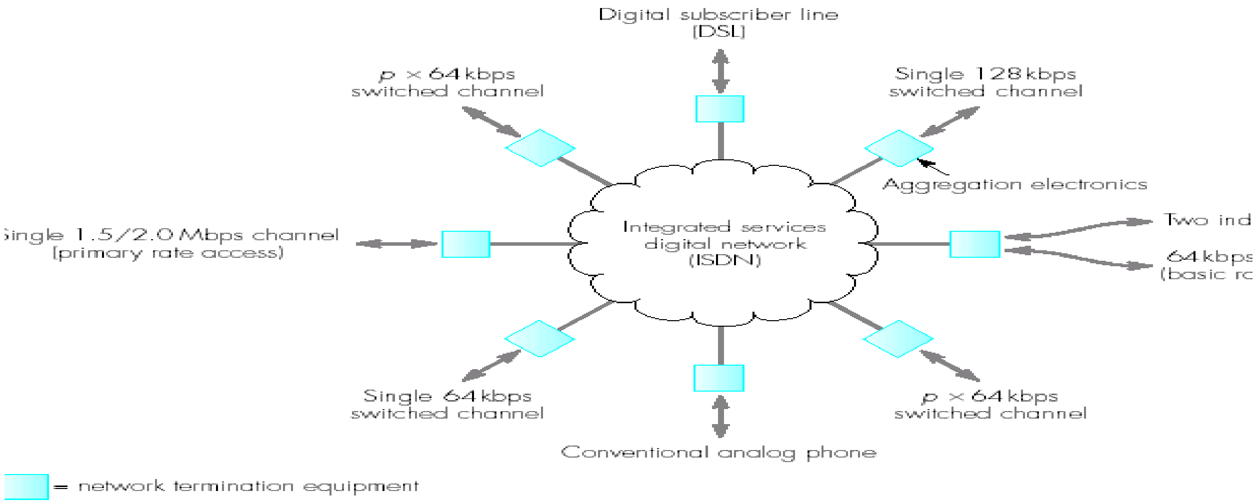
ISDN is a set of ITU standards for digital transmission over ordinary telephone copper wire as well as over other media. Home and business users who install an ISDN adapter (in place of a modem) can see highly-graphic Web pages arriving very quickly (up to 128 Kbps). ISDN requires adapters at both ends of the transmission so your access provider also needs an ISDN adapter. ISDN is generally available from your phone company

- DSL (Digital Subscriber Line) is a technology for bringing high-bandwidth information to homes and small businesses over ordinary copper telephone lines.

- Assuming your home or small business is close enough to a telephone company central office that offers DSL service, you may be able to receive continuous transmission of motion video, audio, and even 3-D effects.

- Typically, individual connections will provide from 1.544 Mbps to 512 Kbps **downstream** and about 128 Kbps **upstream**. A DSL line can carry both data and voice signals and the data part of the line is continuously connected.

- **Access circuit** that allows users either ***two different telephone calls simultaneously or a telephone call and a data network***

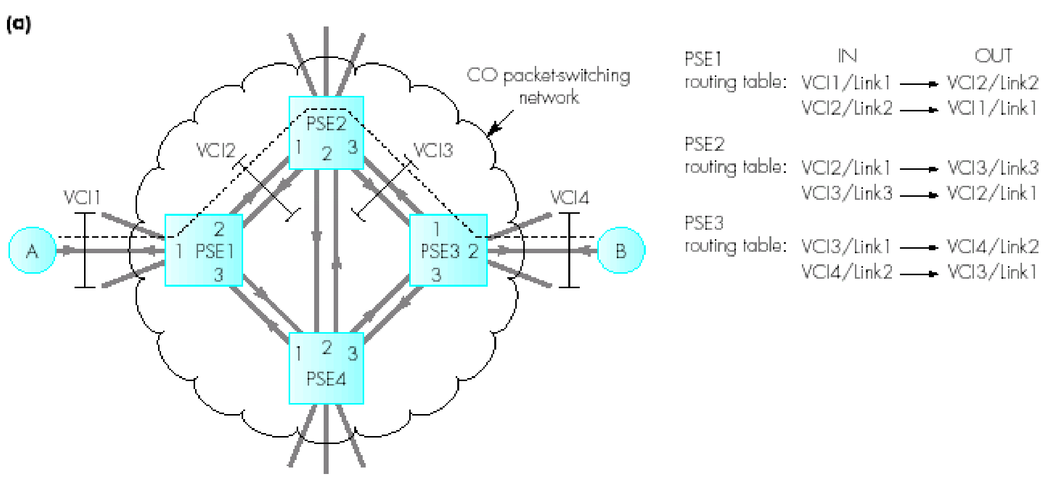


- DSL supports two 64 kbps channels that can be used independently or as a single combined 128kbps channel (additional box of electronics). This is known as the **aggregation** function

(b) **Packet Switched Network**

- There are two types of packet-mode network

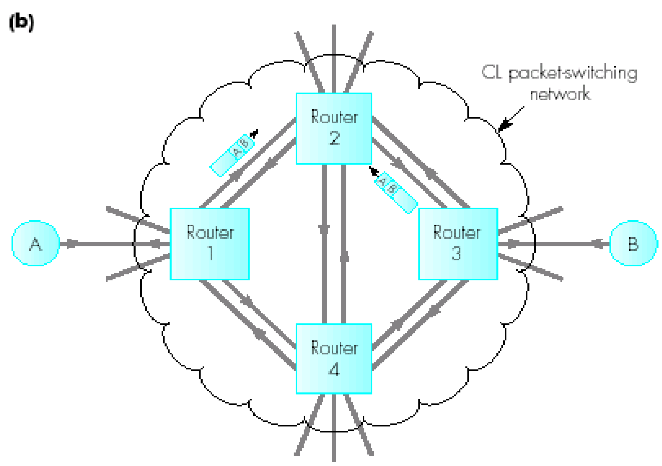
- Connection Oriented (CO)



- As the name implies a connection is established prior to information interchange
- The connection utilizes only a variable portion of the bandwidth of each link and known as virtual circuit (VC)

- To set up a VC the source terminal sends a call request control packet to the local PSE which in addition to the source and destination addresses holds a short identifier known as *virtual circuit identifier* (VCI)
- Each PSE maintains a table that specifies the outgoing link to use to reach the network address
- On receipt of the call request the PSE uses the destination address within the packet to determine the outgoing link
- The next free identifier (VCI) for this link is selected and two entries are made in the *routing table*

Connectionless



- In connectionless network, the establishment of a connection is not required and they can exchange information as and when they arrive
- Each packet must carry the full source and destination address in its header in order for each PSE to route the packet onto the appropriate outgoing link (**router** term used rather than PSE)
- In both types each packet is stored in a memory buffer and a check is performed to determine if any transmission errors are present in the received message. (i.e 0 instead of a 1 or vice versa)
- If an error is detected then the packet is discarded known as **best-effort service**.
- All packets are transmitted at the maximum link bit rate
- As packets may need to use the same link to transfer information an operation known as **store-and-forward** is used.
- The sum of the store and forward delays in each PSE/router contributes to the overall transfer delay of the packets and the mean of this delay is known as the *mean packet transfer delay*.
- The variation about the mean are known as the *delay variation* or *jitter*

- Example of connectionless mode – *Internet*
- Examples of connection oriented network – *X.25 (text) and ATM (multimedia)*

(c)

Derive the maximum block size that should be used over a channel which has a mean BER probability of 10^{-4} if the probability of a block containing an error – and hence being discarded – is to be 10^{-1}.

Answer:

$$P_B = 1 - (1 - P)^N$$

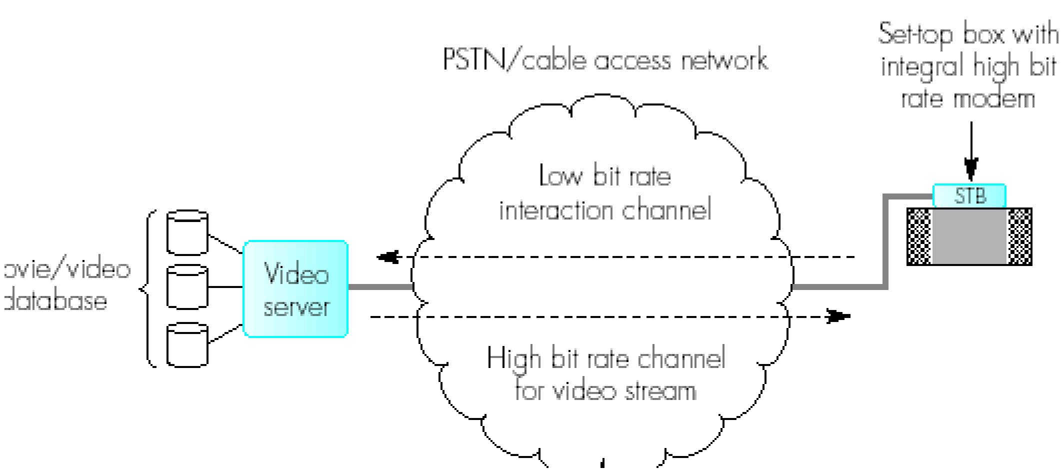
Hence $0.1 = 1 - (1 - 10^{-4})^N$ and $N = 950$ bits

Alternatively, $P_B = N \times P$

Hence $0.1 = N \times 10^{-4}$ and $N = 1000$ bits

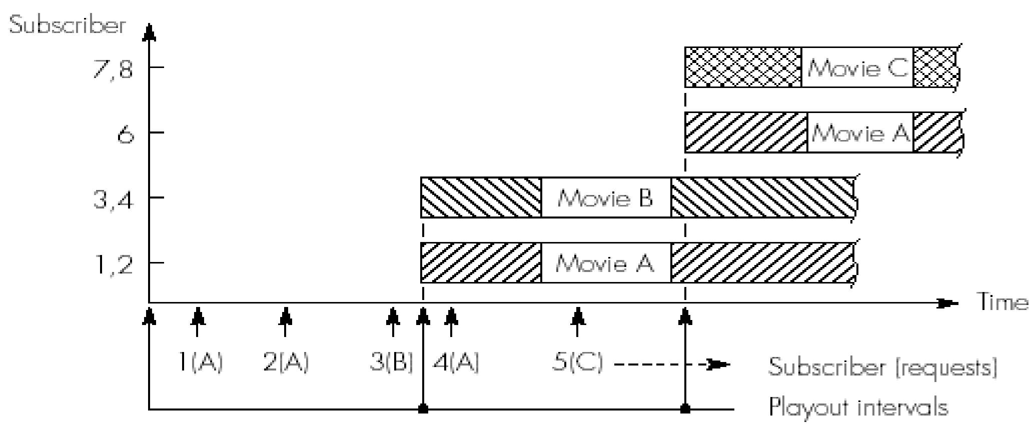
2. (a) Movie on Demand/Near Movie on Demand

- The entertainment applications require higher quality / resolution for video and audio since wide-screen televisions and stereophonic sound are often used



- Normally the subscriber terminal comprises television with a selection deive for interation purposes
- The user interactions are relayed to the server through a set-top-box (STB) which contains a high speed modem
- By means of the menu the user can browse through the movies/videos and initiate the showing of a selected movie. This is known as Movie-on-demand or Video-on-demand.
- *Key features of MOD*
- - Subscriber can initate the showing of a movie from a library of movies at any time of the day or night
- *Issues associated with MOD*
- - The server must be capable of playing out simultaneously a large number of video streams equal to the number of subscribers at any one time
- - This will require high speed information flow from the server (multi-movies + multi-copies)
- In order to avoid the heavy load there is another mode of operation used. In which requests are queued until the start of the next playout time.

N-MOD



MOD = movie-on-demand N-MOD = near movie-on-demand

- This mode of operation is known as the **near movie-on-demand (N-MOD)**

2. (b) Multipoint Conferencing

- Multipoint conferencing is implemented in one of two ways

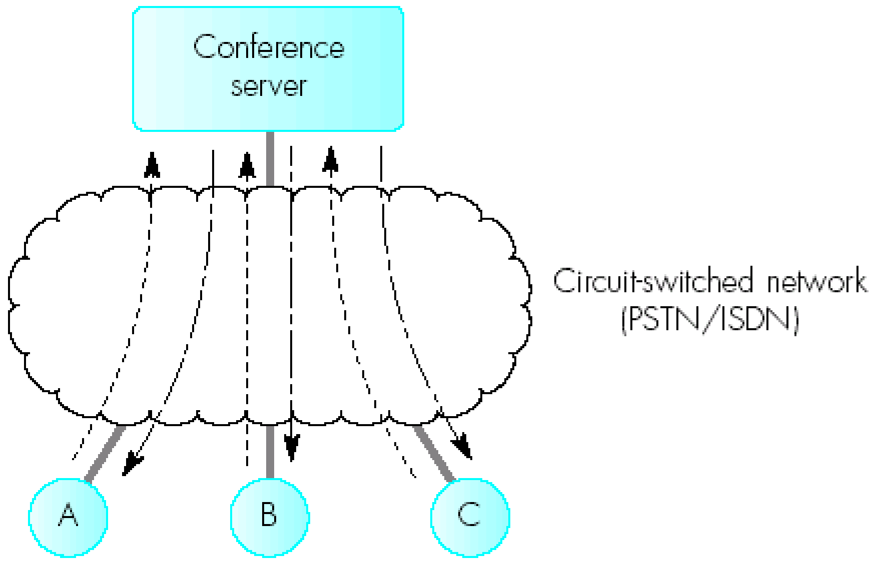
  - *Centralized mode*
  - *Decentralized mode*

Centralized mode

- This mode is used with circuit switched networks such as PSTN and ISDN

(i) Centralized Mode

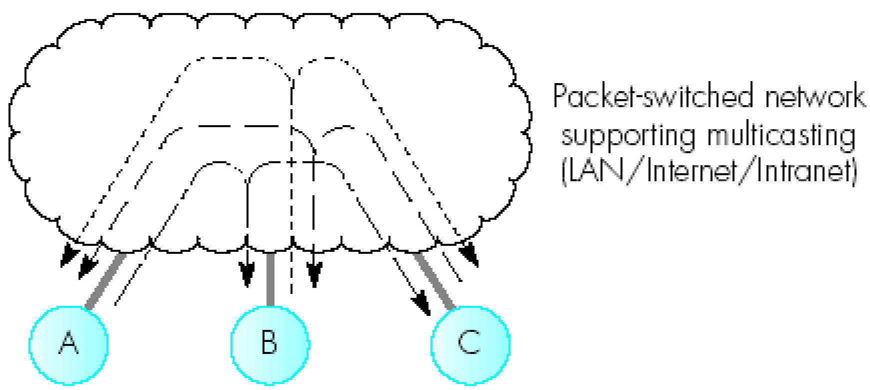
(a)



- With this mode a central server is used
- Prior to sending any information each terminal needs to set up a connection to the server
- The terminal then sends the information to the server.
- The server then distributes this information to all the other terminals connected in the conference

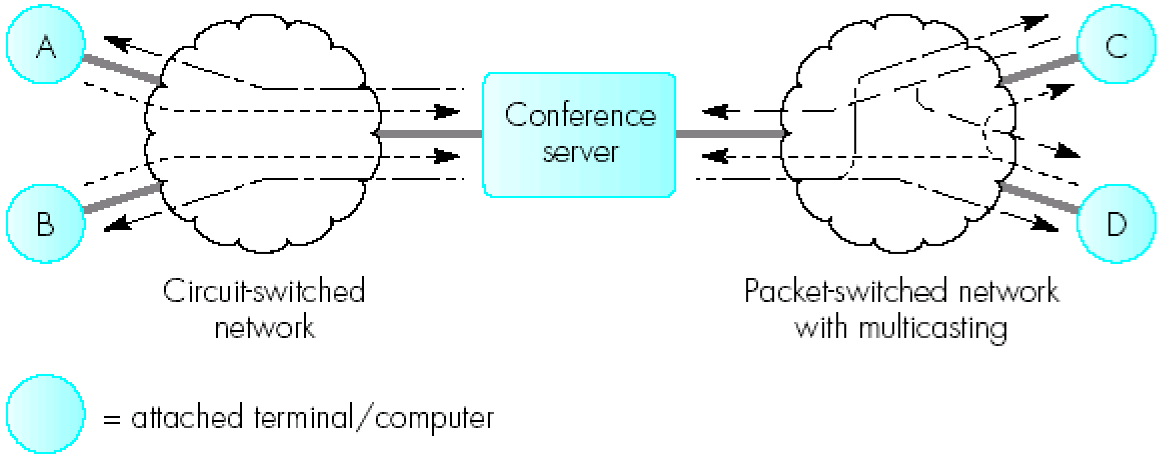
(ii)Decentralized Mode

(b)



- The decentralized mode is used with packet-switched networks that support multicast communications
- E.g – LAN, Intranet, Internet
- The output of each terminal is received by all the other members of the conference/multicast group
- Hence a conference server is not required and it is the responsibility of each terminal to manage the information streams that they receive from the other members

(iii)Hybrid Mode



- This type of mode is used when the terminals are connected to different network types
- In this mode the server determines the output stream to be sent to each terminal

2. (C)

Determine the propagation delay associated with the following communication channels:

(i) a connection through a private telephone network of 1 km,

(ii) a connection through a PSTN of 200 km,

(iii) a connection over a satellite channel of 50 000 km.

Assume that the velocity of propagation of a signal in the case of (i) and (ii) is $2 \times 10^8\ \text{ms}^{-1}$ and in the case of (iii) $3 \times 10^8\ \text{ms}^{-1}$.

Answer:

Propagation delay T_p = physical separation/velocity of propagation

(i) $$T_p = \frac{10^3}{2\times10^8} = 5\times10^{-6}\,\text{s}$$

(ii) $$T_p = \frac{200\times10^3}{2\times10^8} = 10^{-3}\,\text{s}$$

(iii) $$T_p = \frac{5\times10^7}{3\times10^8} = 1.67\times10^{-1}\,\text{s}$$

Module 2

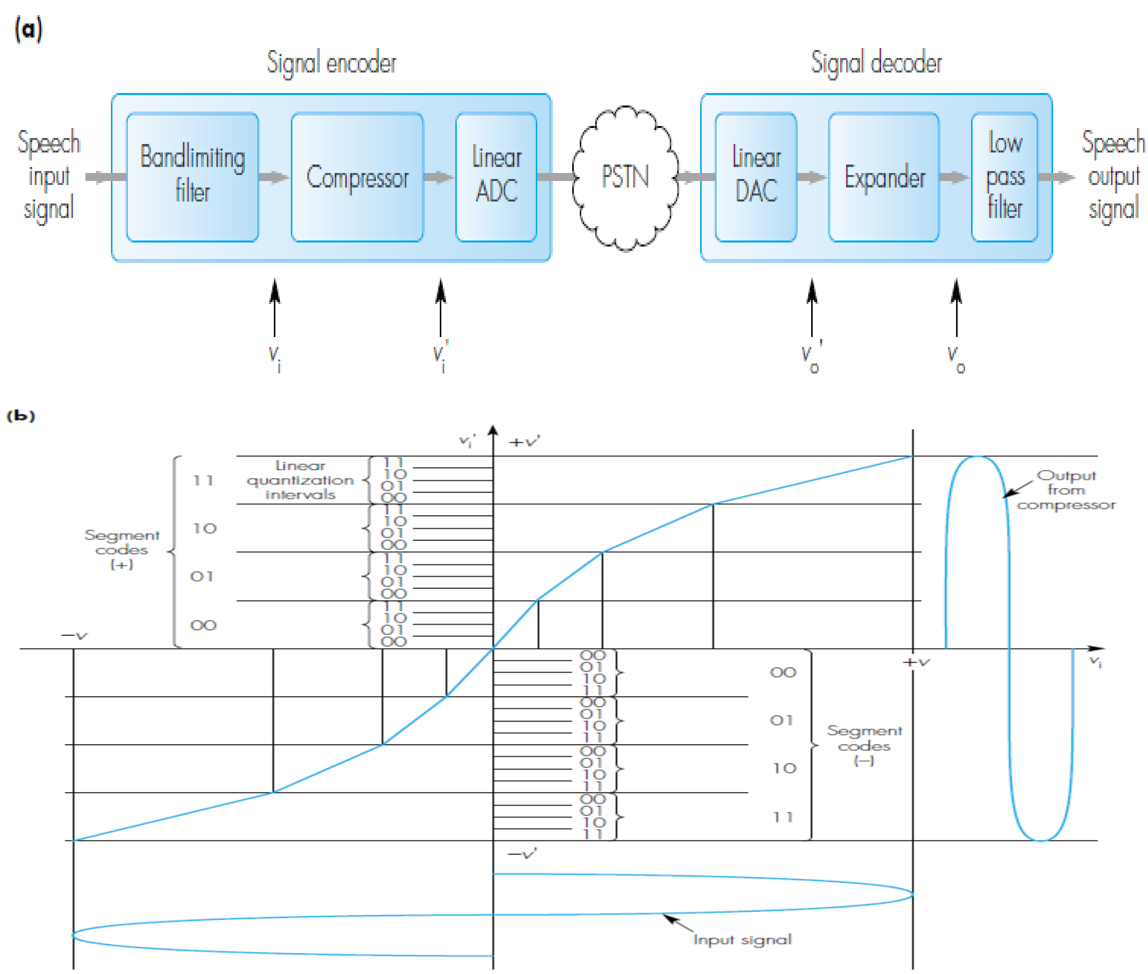
3. (a) PCM Speech CODEC

It is a digitization process. Defined in ITU-T Recommendations G.711.PCM consists of encoder and decoder

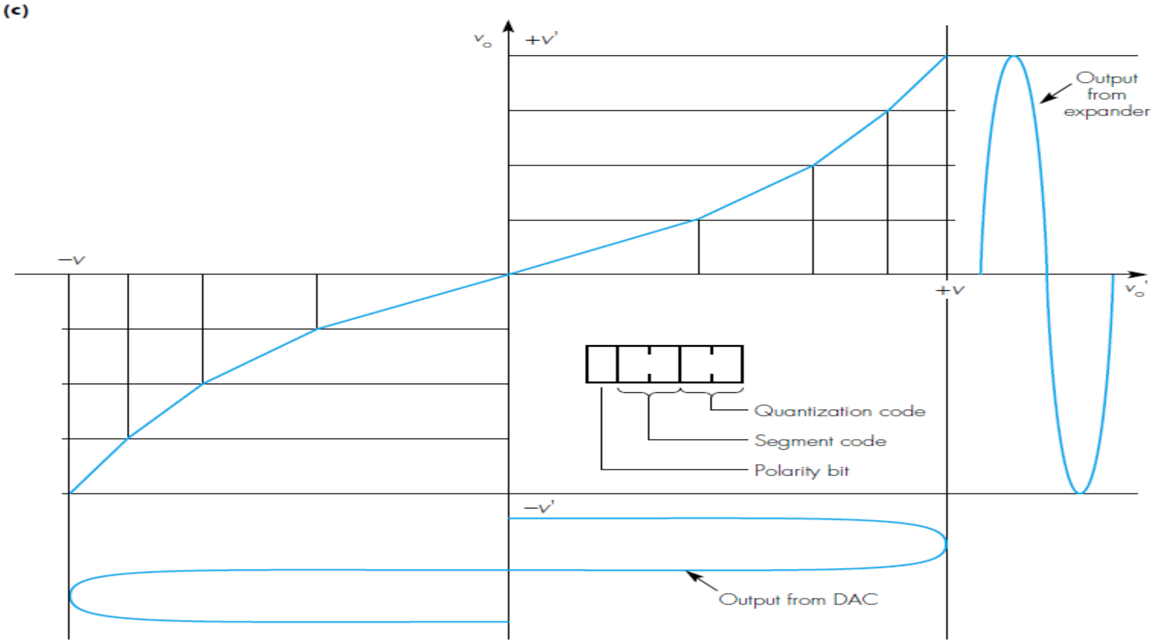
It consists of expander and compressor. As compared to earlier where linear quantization is used – noise level same for both loud and low signals.

As ear is more sensitive to noise on quite signals than loud signals, PCM system consists of non-linear quantization with narrow intervals through compressor. At the destination expander is used The overall operation is companding. Before sampling and using ADC, signal passed through compressor first and passed to ADC and quantized. At the receiver, codeword is first passed to DAC and expander.

Two compressor characteristics – A law and mu law



(c)

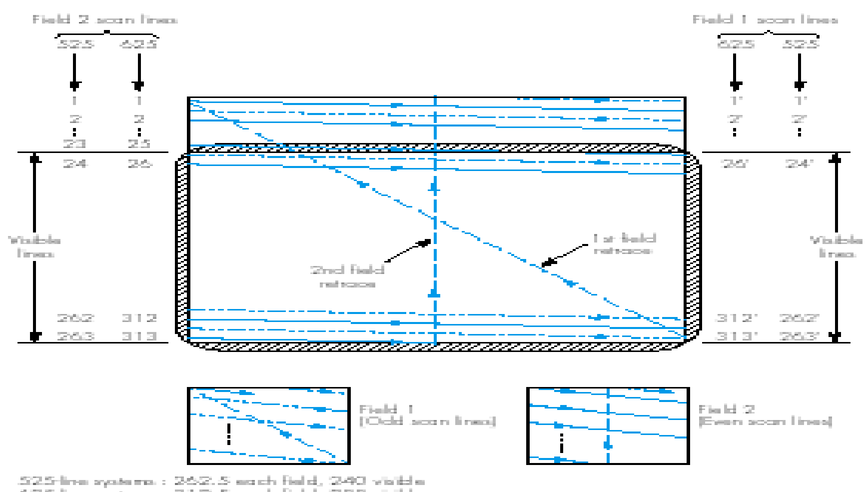


Note that in the G.711 standard a 3-bit segment code and 4-bit quantization code are used.

3. (b) **Interlaced Scanning**

- It is necessary to use a minimum refresh rate of 50 times per second to avoid flicker
- A refresh rate of 25 times per second is sufficient
- Field: The first comprising only the odd scan lines and the second the even scan lines The two field are then integrated together in the television receiver using a technique known as

Figure 2.19 Interlaced scanning principles.



3.(c)

Derive the bit rate and the memory requirements to store each frame that result from the digitization of both a 525-line and a 625-line system assuming a 4:2:2 format. Also find the total memory required to store a 1.5 hour movie/video.

Answer:

525-line system: The number of samples per line is 720 and the number of visible lines is 480. Hence the resolution of the luminance (Y) and two chrominance (C_b and C_r) signals are:

$$Y = 720 \times 480$$
$$C_b = C_r = 360 \times 480$$

Bit rate: Line sampling rate is fixed at 13.5 MHz for Y and 6.75 MHz for both C_b and C_r, all with 8 bits per sample.

Hence: Bit rate $= 13.5 \times 10^6 \times 8 + 2\ (6.75 \times 10^6 \times 8) = 216$ Mbps

Memory required: Memory required per line $= 720 \times 8 + 2\ (360 \times 8)$ $= 11\,520$ bits or 1440 bytes

Hence memory per frame, each of 480 lines $= 480 \times 11\,520$ $= 5.5296$ Mbits or 691.2 kbytes

and memory to store 1.5 hours assuming 60 frames per second:
$= 691.2 \times 60 \times 1.5 \times 3600$ kbytes
$= 223.9488$ Gbytes

625-line system: Resolution: $Y = 720 \times 576$
$C_b = C_r = 360 \times 576$

Bit rate $= 13.5 \times 10^6 \times 8 + 2\ (6.75 \times 10^6 \times 8) = 216$ Mbps

Memory per frame $= 576 \times 11\,520 = 6.635\,55$ Mbits or 829.44 kbytes

and memory to store 1.5 hours assuming 50 frames per second:
$= 829.44 \times 50 \times 1.5 \times 3600$ kbytes
$= 223.9488$ Gbytes

It should be noted that, in practice, the bit rate figures are less than the computed values since they include samples during the retrace times when the beam is switched off. Nevertheless, as we can deduce from the computed values, both the bit rate and the memory requirements are very large for both systems and it is for this reason that the various lower resolution formats have been defined.

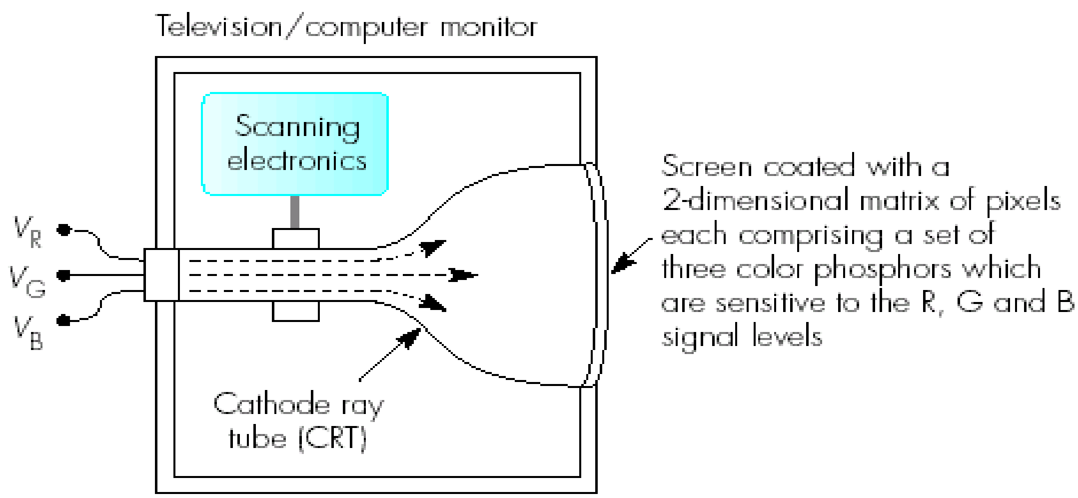
4. (a)

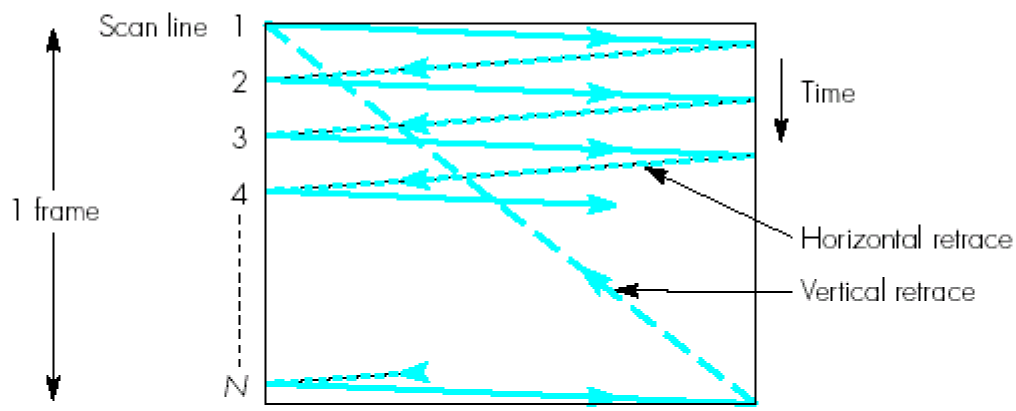
(i) Aspect Ratio

- This is the ratio of the screen width to the screen height (television tubes and PC monitors have an aspect ratio of 4/3 and wide screen television is 16/9)

(ii) Raster Scan

The picture tubes used in most television sets operate using what is known as a **raster-scan**; this involves a finely-focussed electron beam being scanned over the complete screen





N = 525 (NTSC) and 625 (PAL/CCIR/SECAM)
Frame refresh rate = 60 times per second (NTSC)
= 50 times per second (PAL/CCIR/SECAM)

- *Progressive scanning* is performed by repeating the scanning operation that starts at the top left corner of the screen and ends at the bottom right corner follows by the beam being *deflected back* again to the top left corner
- *Frame:* Each complete set of horizontal scan lines (either 525 for North & South America and most of Asia, or 625 for Europe and other countries)
- *Flicker*: Caused by the previous image fading from the eye retina before the following image is displayed, after a low refresh rate (to avoid this a refresh rate of 50 times per second is required)
- *Pixel depth*: Number of bits per pixel that determines the range of different colours that can be produced
- *Colour Look-up Table* **(CLUT):** Table that stores the selected colours in the subsets as an address to a location reducing the amount of memory required to store an image

(iii) 4:2:2 Standard

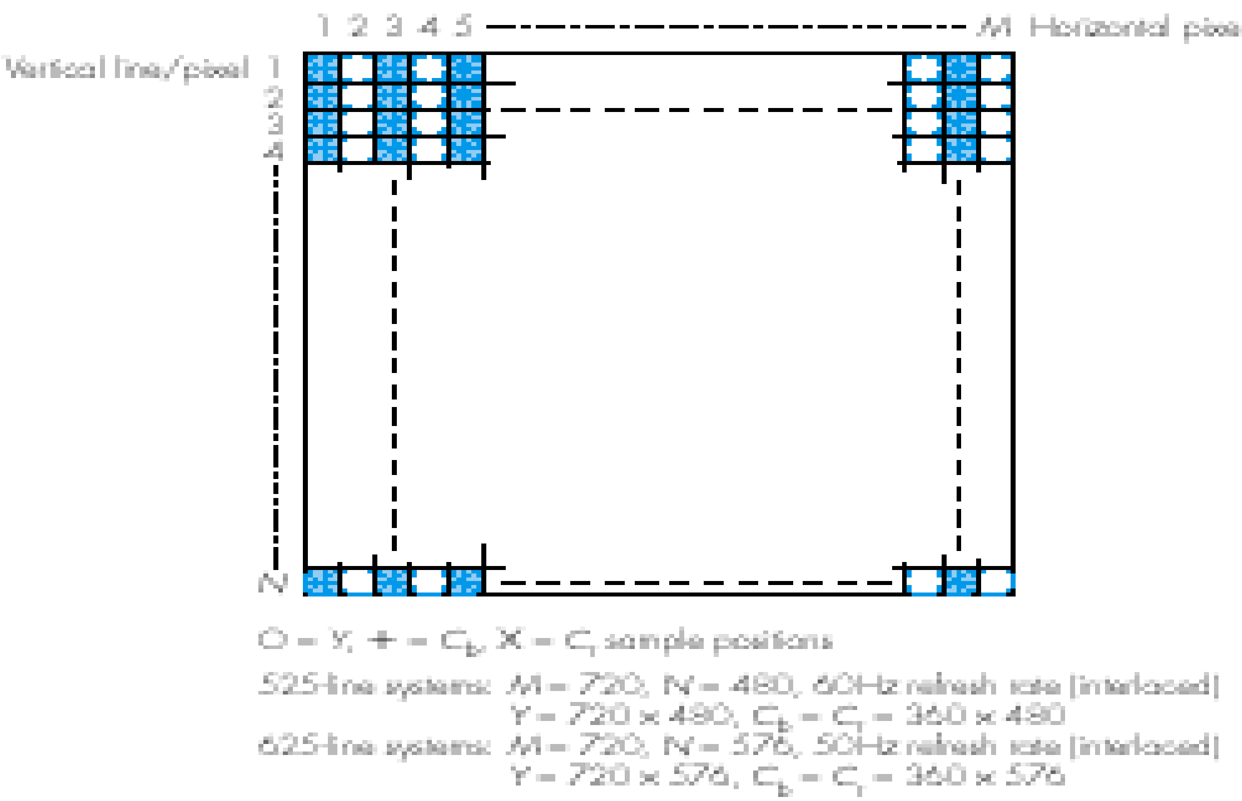
Eye have shown that the resolution of the eye is less sensitive for color than it is for luminance

The original digitization format used in Recommendation CCIR-601

A line sampling rate of 13.5MHz for luminance and 6.75MHz for the two chrominance signals

The number of samples per line is increased to 720

- The corresponding number of samples for each of the two chrominance signals is 360 samples per active line
- This results in 4Y samples for every 2Cb, and 2Cr samples
- The numbers 480 and 576 being the number of active (visible) lines in the respective system



4. (b**) Different Types of Texts**

- *Unformatted text*: Known as plain text; enables pages to be created which comprise strings of fixed-sized characters from a limited character set
- *Formatted Text*: Known as rich text; enables pages to be created which comprise of strings of characters of different styles, sizes and shape with tables, graphics, and images inserted at appropriate points
- *Hypertext*: Enables an integrated set of documents (Each comprising formatted text) to be created which have defined linkages between them
- **Unformatted Text – The basic ASCII character set**
- Control characters
  (Back space, escape, delete, form feed etc)
- Printable characters
  (alphabetic, numeric, and punctuation)

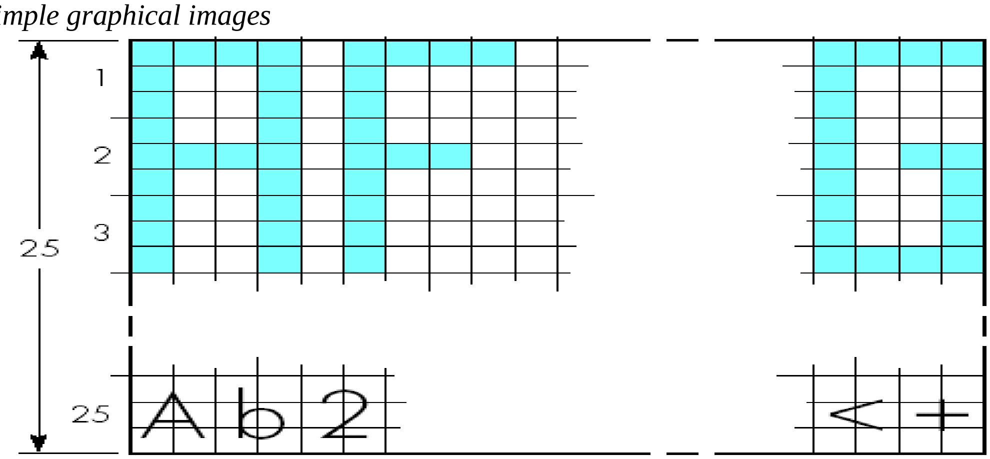
- The American Standard Code for Information Interchange is one of the most widely used character sets and the table includes the binary *codewords* used to represent each character (7 bit binary code)

Bit positions 4	3	2	1	7 6 5: 0 0 0	0 0 1	0 1 0	0 1 1	1 0 0	1 0 1	1 1 0	1 1 1
0	0	0	0	NUL	DLE	SP	0	@	P	\	p
0	0	0	1	SOH	DC1	!	1	A	Q	a	q
0	0	1	0	STX	DC2	"	2	B	R	b	r
0	0	1	1	ETX	DC3	#	3	C	S	c	s
0	1	0	0	EOT	DC4	$	4	D	T	d	t
0	1	0	1	ENQ	NAK	%	5	E	U	e	u
0	1	1	0	ACK	SYN	&	6	F	V	f	v
0	1	1	1	BEL	ETB	'	7	G	W	g	w
1	0	0	0	BS	CAN	(	8	H	X	h	x
1	0	0	1	HT	EM	)	9	I	Y	i	y
1	0	1	0	LF	SUB	*	:	J	Z	j	z
1	0	1	1	VT	ESC	+	;	K	[	k	{
1	1	0	0	FF	FS	,	<	L	\	l	|
1	1	0	1	CR	GS	–	=	M	]	m	}
1	1	1	0	SO	RS	.	>	N	^	n	~
1	1	1	1	SI	US	/	?	O	—	o	DEL

Unformatted Text – Supplementary set of Mosaic characters

Bit positions 4	3	2	1	7 6 5: 0 0 0	0 0 1	0 1 0	0 1 1	1 0 0	1 0 1	1 1 0	1 1 1
0	0	0	0					@	P		
0	0	0	1					A	Q		
0	0	1	0					B	R		
0	0	1	1					C	S		
0	1	0	0					D	T		
0	1	0	1					E	U		
0	1	1	0					F	V		
0	1	1	1					G	W		
1	0	0	0					H	X		
1	0	0	1					I	Y		
1	0	1	0					J	Z		
1	0	1	1					K	[		
1	1	0	0					L	\		
1	1	0	1					M	]		
1	1	1	0					N	^		
1	1	1	1					O	—		

The characters in columns 010/011 and 110/111 are replaced with the set of mosaic characters; and then used, together with the various uppercase characters illustrated, to *create relatively simple graphical images*



Note: Grid only included as a template.

- Although in practice the total page is made up of a matrix of symbols and characters which all have the same size, some simple graphical symbols and text of larger sizes can be constructed by the use of groups of the basic symbols
- **Formatted Text**

```
<B><FONT SIZE=4><P>Formatted Text</P>
</B></FONT>
<P>This is an example of formatted text, it includes:</P>
<FONT SIZE=2>
</FONT><I><P>Italics,</I> <B>Bold</B> and <U>Underlining</P>
</U>
<FONT FACE="French Script MT"><P>Different Fonts</FONT> and <FONT
SIZE=4>Font Sizes</P>
```

Formatted text

This is an example of formatted text, it includes:

Italics, **Bold** and Underlining

Different fonts and Font Sizes

- It is produced by most word processing packages and used extensively in the publishing sector for the preparation of papers, books, magazines, journals and so on..
- Documents of mixed type (characters, different styles, fonts, shape etc) possible.
- Format control characters are used
- **Hypertext – Electronic Document in hypertext**

- Hypertext can be used to create an electronic version of documents with the index, descriptions of departments, courses on offer, library, and other facilities all written in hypertext as pages with various defined hyperlinks

Page 2

Note:
- Page 2 is displayed after clicking the cursor on •Admissions of Page 1
- Selected images can be used as a background.
- Hyperlinks can be either underlined (as shown) or in a different color

- An example of a hypertext language is HTML used to describe how the contents of a document are presented on a printer or a display; other mark-up languages are: Postscript, SGML (Standard Generalized Mark-up language) Tex, and Latex.

4. (c)

Assuming the bandwidth of a speech signal is from 50 Hz through to 10 kHz and that of a music signal is from 15 Hz through to 20 kHz, derive the bit rate that is generated by the digitization procedure in each case assuming the Nyquist sampling rate is used with 12 bits per sample for the speech signal and 16 bits per sample for the music signal. Derive the memory required to store a 10 minute passage of stereophonic music.

Answer:

(i) Bit rates: Nyquist sampling rate = $2 f_{max}$

Speech: Nyquist rate = 2×10 kHz = 20 kHz or 20 ksps

Hence with 12 bits per sample, bit rate generated = $20\text{ k} \times 12 = 240$ kbps

Music: Nyquist rate = 2×20 kHz = 40 kHz or 40 ksps

Hence bit rate generated = $40\text{ k} \times 16 = 640$ kbps (mono)

or $2 \times 640\text{ k} = 1280$ kbps (stereo)

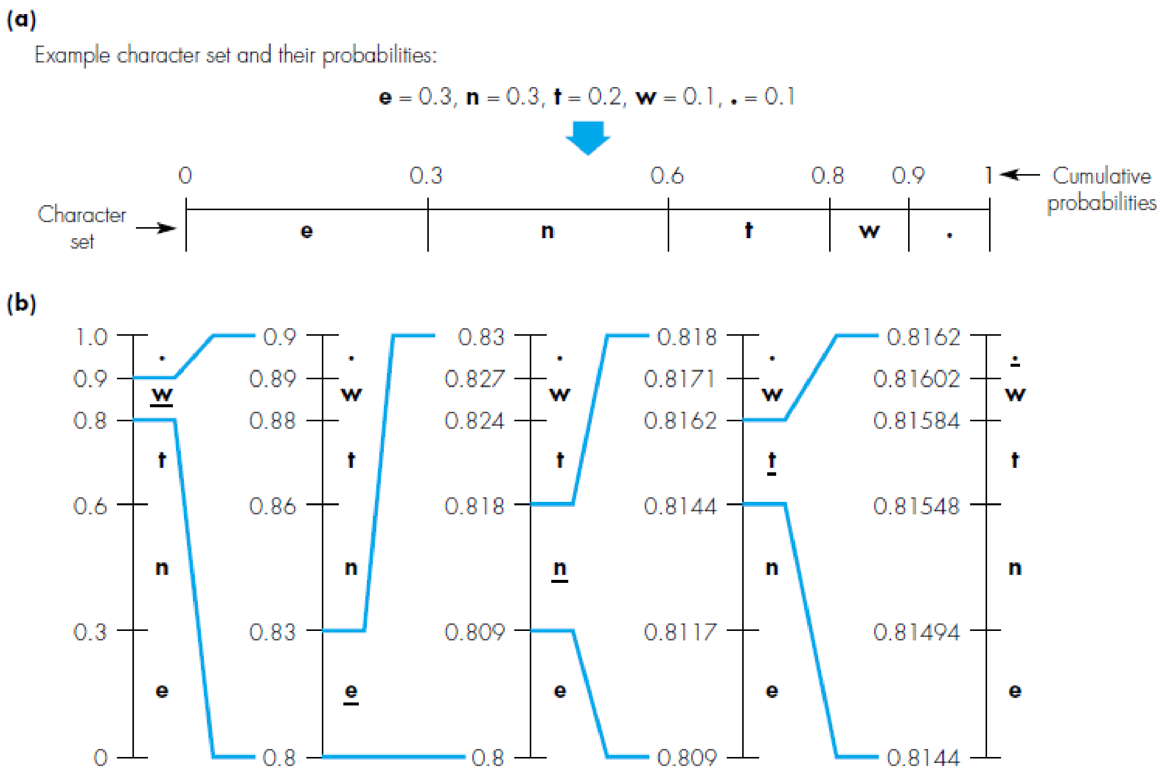
(ii) Memory required: Memory required = bit rate (bps) × time (s)/8 bytes

Hence at 1280 kbps and 600 s,

$$\text{Memory required} = \frac{1280 \times 10^3 \times 600}{8} = 96 \text{ Mbytes}$$

Module 3

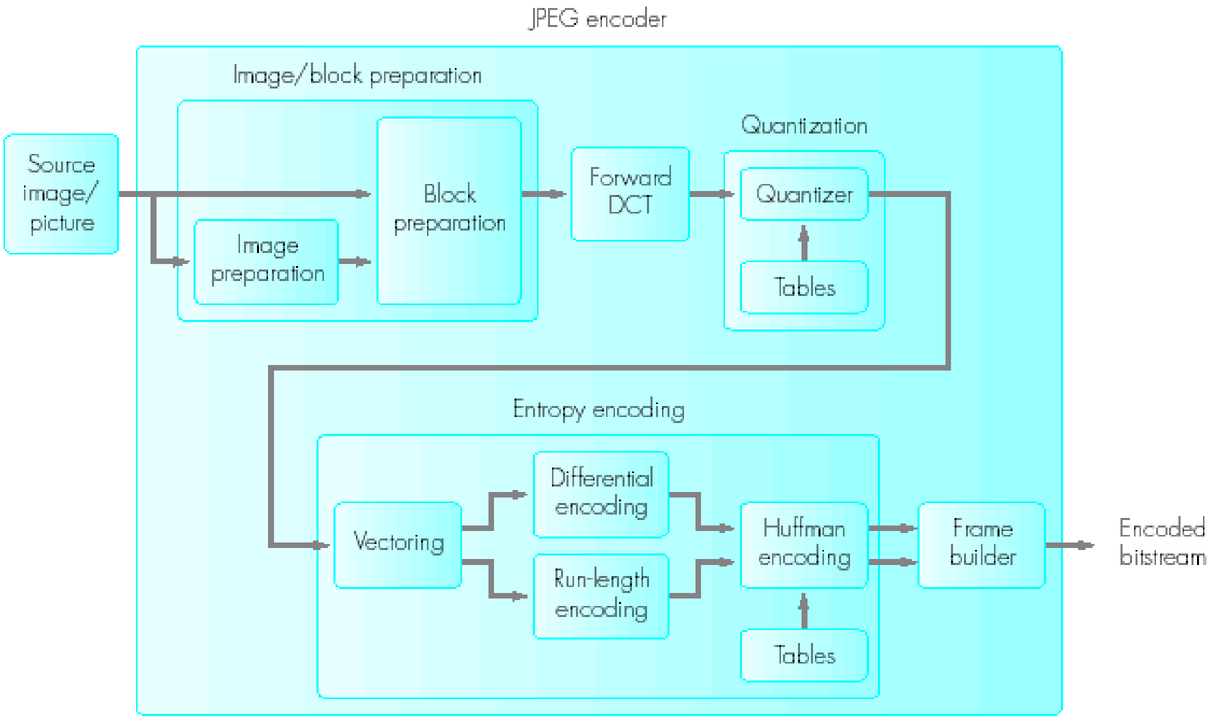
5. (a)



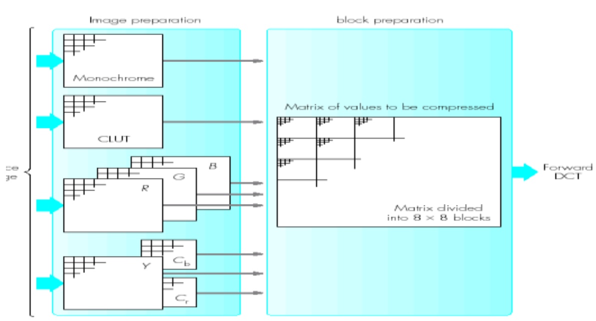
Encoded version of the character string **went.** is a single codeword in the range 0.816 02 ≤ codeword < 0.8162

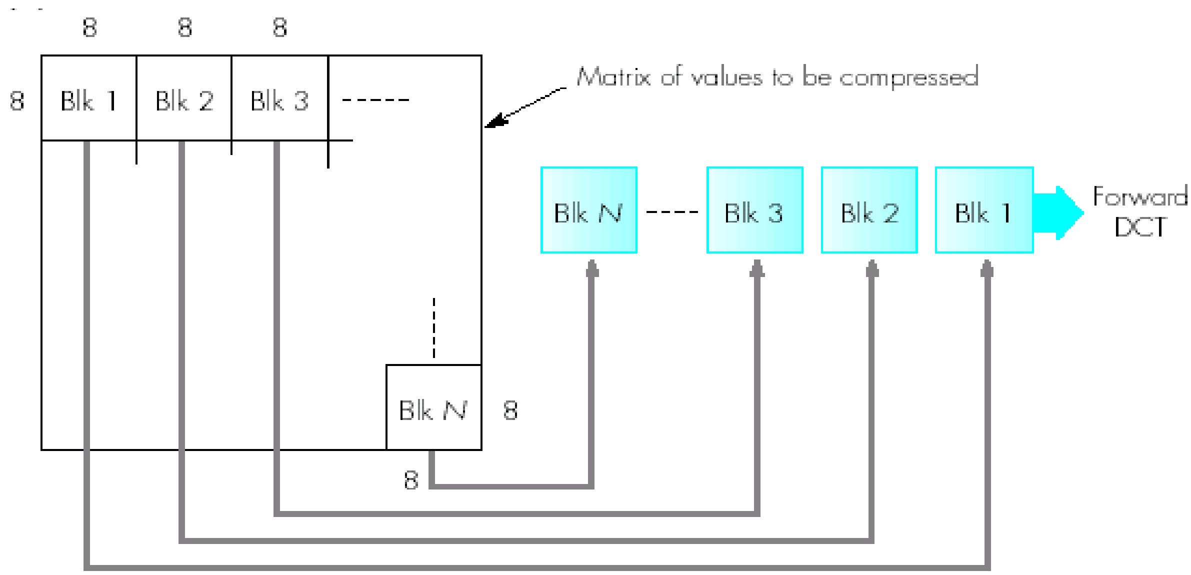
5.(b) JPEG Encoder

- The Joint Photographic Experts Group forms the basis of most video compression algorithms
- 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)



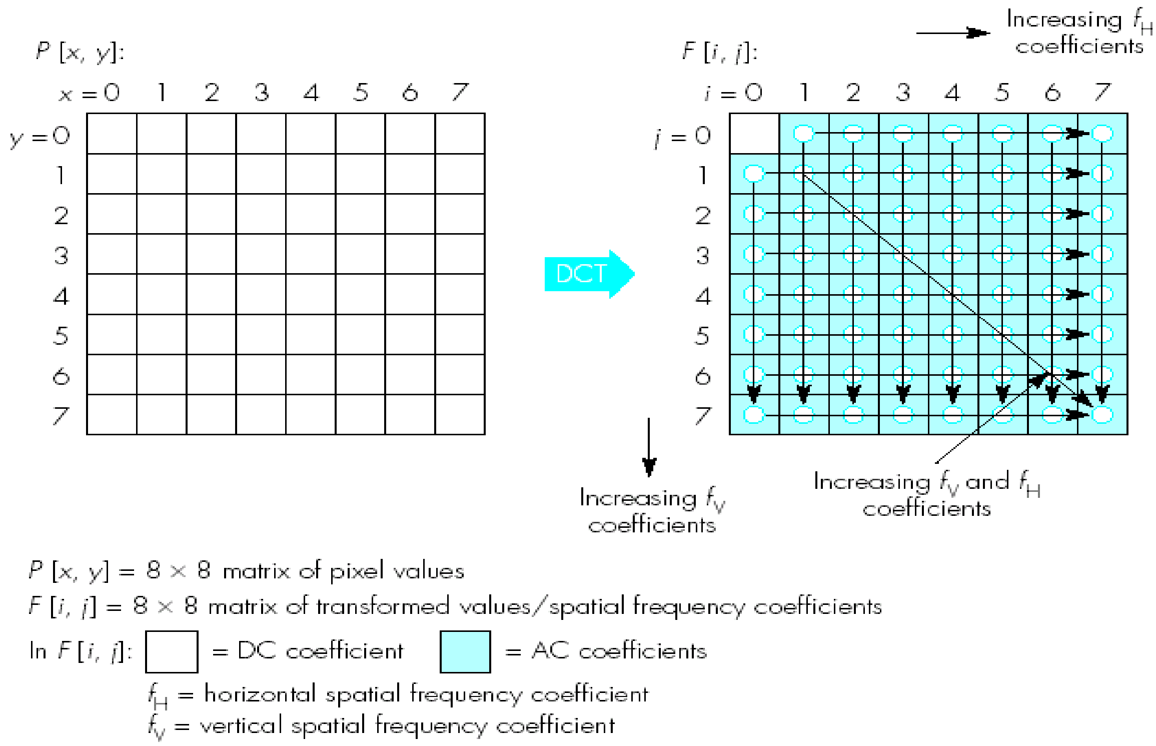
- Once the image format is selected then the values in each matrix are compressed separately using the DCT
- In order to make the transformation more efficient a second step known as **block preparation** is carried out before DCT
- In block preparation each global matrix is divided into a set of smaller 8X8 submatrices (block) which are fed sequentially to the DCT
- Once the source image format has been selected and prepared (four alternative forms of representation), the set values in each matrix are compressed separately using the DCT)





- Block preparation is necessary since computing the transformed value for each position in a matrix requires the values in all the locations to be processed
- Each pixel value is quantized using 8 bits which produces a value in the range 0 to 255 for the R, G, B or Y and a value in the range –128 to 127 for the two chrominance values C_b and C_r
- If the *input matrix* is *P[x,y]* and the *transformed matrix* is *F[i,j]* then the DCT for the 8X8 block is computed using the expression:
- $$F[i,j]=\frac{1}{4}C(i)C(j)\sum_{x=0}^{7}\sum_{y=0}^{7}P[x,y]\cos\frac{(2x+1)i\pi}{16}\cos\frac{(2y+1)j\pi}{16}$$

  All 64 values in the input matrix *P[x,y]* contribute to each entry in the transformed matrix *F[i,j]*
- For *i = j = 0* the two cosine terms are 0 and hence the value in the location *F[0,0]* of the transformed matrix is simply a function of the summation of all the values in the input matrix
- This is the mean of all 64 values in the matrix and is known as the **DC coefficient**
- Since the values in all the other locations of the transformed matrix have a frequency coefficient associated with them they are known as **AC coefficients**
- for *j = 0* only the horizontal frequency coefficients are present
- for *i = 0* only the vertical frequency components are present
- For all the other locations both the horizontal and vertical frequency coefficients are present



- The values are first centred around zero by subtracting 128 from each intensity/luminance value
- 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
-

DCT coefficients

120	60	40	30	4	3	0	0
70	48	32	3	4	1	0	0
50	36	4	4	2	0	0	0
40	4	5	1	1	0	0	0
5	4	0	0	0	0	0	0
3	2	0	0	0	0	0	0
1	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Quantizer

Quantized coefficients

12	6	3	2	0	0	0
7	3	2	0	0	0	0
3	2	0	0	0	0	0
2	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0

10	10	15	20	25	30	35	40
10	15	20	25	30	35	40	50
15	.20	25	30	35	40	50	60
20	25	30	35	40	50	60	70
25	30	35	40	50	60	70	80
30	35	40	50	60	70	80	90
35	40	50	60	70	80	90	100
40	50	60	70	80	90	100	110

Quantization table

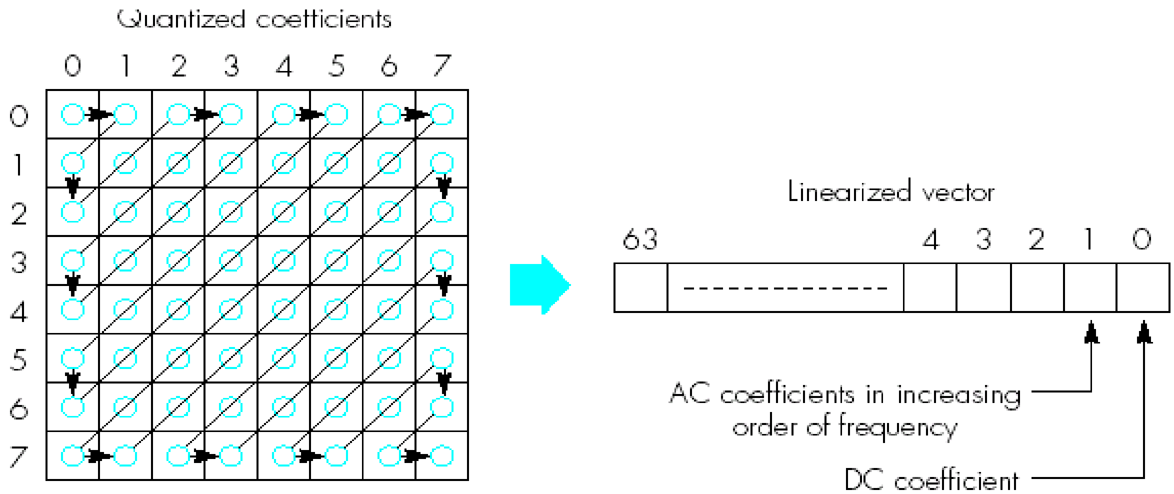
- From the *quantization table* and the *DCT and quantization coefficents* number of observations can be made:
  - The computation of the quantized coefficients involves rounding the quotients to the nearest integer value
  - The threshold values used increase in magnitude with increasing spatial frequency
  - The DC coefficient in the transformed matrix is largest
  - Many of the higher frequency coefficients are zero
- *Entropy encoding consists of four stages*

  Vectoring – The entropy encoding operates on a one-dimensional string of values (vector). However the output of the quantization is a 2-D matrix and hence this has to be represented in a 1-D form. This is known as vectoring

  Differential encoding **– In this section only the difference in magnitude of the DC coefficient in a quantized block relative to the value in the preceding block is encoded. This will reduce the number of bits required to encode the relatively large magnitude**

  The difference values are then encoded in the form *(SSS, value) SSS indicates the number of bits needed and actual bits that represent the value*

  e.g: if the sequence of DC coefficients in consecutive quantized blocks was: 12, 13, 11, 11, 10, --- the difference values will be 12, 1, -2, 0, -1



- In order to exploit the presence of the large number of zeros in the quantized matrix, a zig-zag of the matrix is used
- The remaining 63 values in the vector are the AC coefficients
- Because of the large number of 0's in the AC coefficients they are encoded as string of pairs of values
- Each pair is made up of (*skip, value*) where *skip* is the number of zeros in the run and *value* is the next non-zero coefficient

63		12	11	10	9	8	7	6	5	4	3	2	1	0
0	------------------------	0	0	0	2	2	2	2	3	3	3	7	6	12

- The above will be encoded as
  (0,6) (0,7) (0,3)(0,3)(0,3) (0,2)(0,2)(0,2)(0,2)(0,0)
  Final pair indicates the end of the string for this block
- Significant levels of compression can be obtained by replacing long strings of binary digits by a string of much shorter codewords
- The length of each codeword is a function of its relative frequency of occurrence
- Normally, a table of codewords is used with the set of codewords precomputed using the Huffman coding algorithm
- In order for the remote computer to interpret all the different fields and tables that make up the bitstream it is necessary to delimit each field and set of table values in a defined way
- The JPEG standard includes a definition of the structure of the total bitstream relating to a particular image/picture. This is known as a ***frame***
- The role of the frame builder is to *encapsulate* all the information relating to an encoded image/picture

5.(c) CPU Management in Multimedia Operating System

CPU Management

Real-time processing can be achieved through efficient real-time scheduling. In the context of continuous media, a deadline can be the acceptable playback time of each frame. Therefore, it is a soft deadline and appears periodically. The challenges of multimedia scheduling are due to two conflicting goals: non-real-time processes and real-time processes. Non-real-time processes should not suffer from the execution of real-time processes, because multimedia applications equally depend on discrete and continuous media data. Real-time processes should be allowed to pre-empt non-real-time processes or other real-time processes with lower priorities.

The most important real-time scheduling approaches include Earliest Deadline First (EDF) and rate monitoring scheduling [4.23]. With EDF, each task is preemptive and is assigned a priority according to the deadline. The highest priority is assigned to the job with the earliest deadline, and tasks are executed in a priority order. When a new task arrives, the scheduler recomputes the priorities of all pending tasks and then reorganizes such that the order of the task being executed is preempted and the new task gets served immediately. The interrupted process is resumed later from the interruption point. Otherwise, the new task will be put in an appropriate position.

With rate-monotonic scheduling, each task is pre-empted and is assigned a priority according to the request rate. The highest priority is assigned to the job with the highest rate. In contrast to EDF, such assignments are performed only at the connection establishment time and are maintained through the lifetime of the connection. For preemptive periodic tasks, rate-monotonic scheduling is optimal in the sense that no other-static algorithm can schedule a task that the rate-monotonic algorithm cannot also schedule [4.24].

Comparing these two algorithms, EDF is more dynamic. It has to be executed frequently and thus incurs higher scheduling overhead. The advantage is that it can achieve processor utilization up to 100%. On the other hand, a rate-monotonic algorithm is static because the priority assignment is only calculated once. Because the priorities are assigned according to the request rate, more context switches occur in rate-monotonic scheduling than EDF. The worst-case upper bound of the process use is about 69% even though, on the average, the use is suitable for continuous media applications because it has no scheduling overhead and is optimal for periodic jobs.

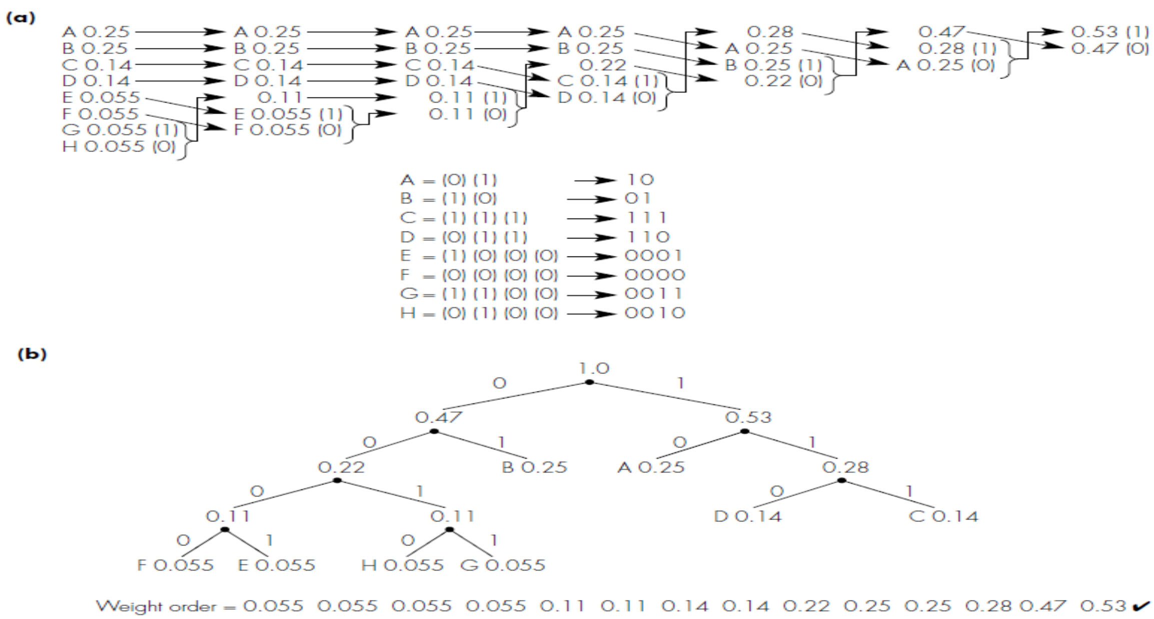
6. (a)

A series of messages is to be transferred between two computers over a PSTN. The messages comprise just the characters A through H. Analysis has shown that the probability (relative frequency of occurrence) of each character is as follows:

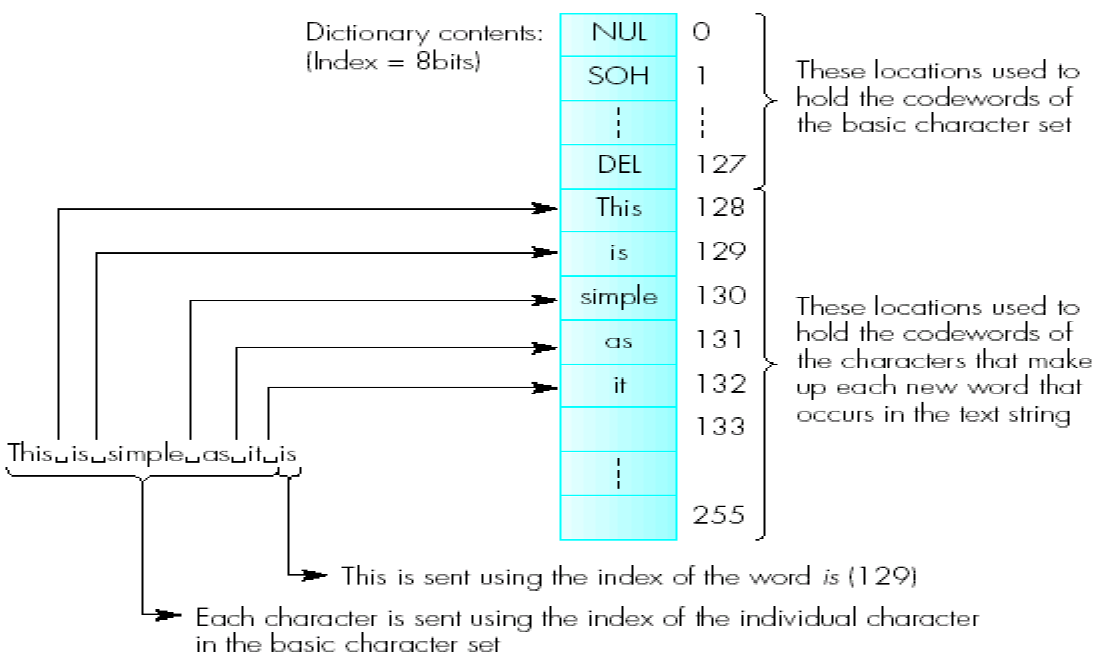
A and B = 0.25, C and D = 0.14, E, F, G, and H = 0.055

(a) Use Shannon's formula to derive the minimum average number of bits per character.

(b) Use Huffman coding to derive a codeword set and prove this is the minimum set by constructing the corresponding Huffman code tree.



6.(b) LZW Compression

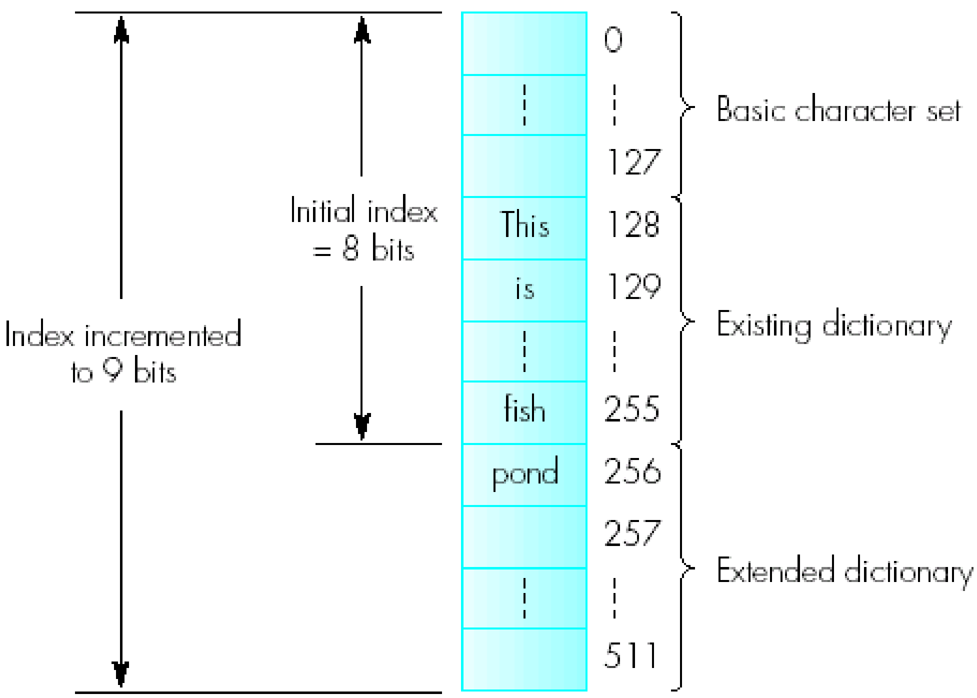


- The principle of the *Lempel-Ziv-Welsh* coding algorithm is for the encoder and decoder to build the ***contents of the dictionary dynamically*** as the text is being transferred

- Initially the decoder has only the character set – e.g ASCII. The remaining entries in the dictionary are *built dynamically* by the encoder and decoder

- Initially the encoder sends the index of the four characters T, H, I, S and sends the space character which will be detected as a non alphanumeric character

- It therefore transmits the character using its index as before but in addition interprets it as terminating the first word and this will be stored in the next free location in the dictionary
- Similar procedure is followed by both the encoder and decoder
- In applications with 128 characters initially the dictionary will start with 8 bits and 256 entries 128 for the characters and the rest 128 for the words



- A key issue in determining the level of compression that is achieved, is the **number of entries** in the dictionary since this determines the **number of bits** that are required for the index

6. (c) **Features of Distributed Multimedia System**

- ***Technology integration***—Integrates information, communication and computing systems to form a unified digital processing environment.
- ***Multimedia integration***—Accommodates discrete data as well as continuous data in an integrated environment.
- ***Real-time performance***—Requires the storage systems processing systems and transmission systems to have real-time performance. Hence, huge storage volume, high network transmission rate and high CPU processing rate are required.
- ***Systemwide QoS support***—Supports diverse QoS requirements on an end-to-end basis along the data path from the sender, through the transport network and to the receiver.
- ***Interactivity***—Requires duplex communication between the user and the system and allows each user to control the information.
- ***Multimedia synchronization support***—Presents the playback continuity of media frames within a single continuous media stream, and temporal relationships among multiple related data objects.
- ***Standardization support***—Allows interoperability despite heterogeneity in the information content, presentation format, user interfaces, network protocols and consumer electronics.

Module 4

7. (a) **Linear Predictive Coding**

All algorithms – sampling, digitization and quantization using DPCM / ADPCM

DSP circuits 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:
  - Pitch
  - Period
  - Loudness

Basic feature of an LPC encoder/decoder:

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

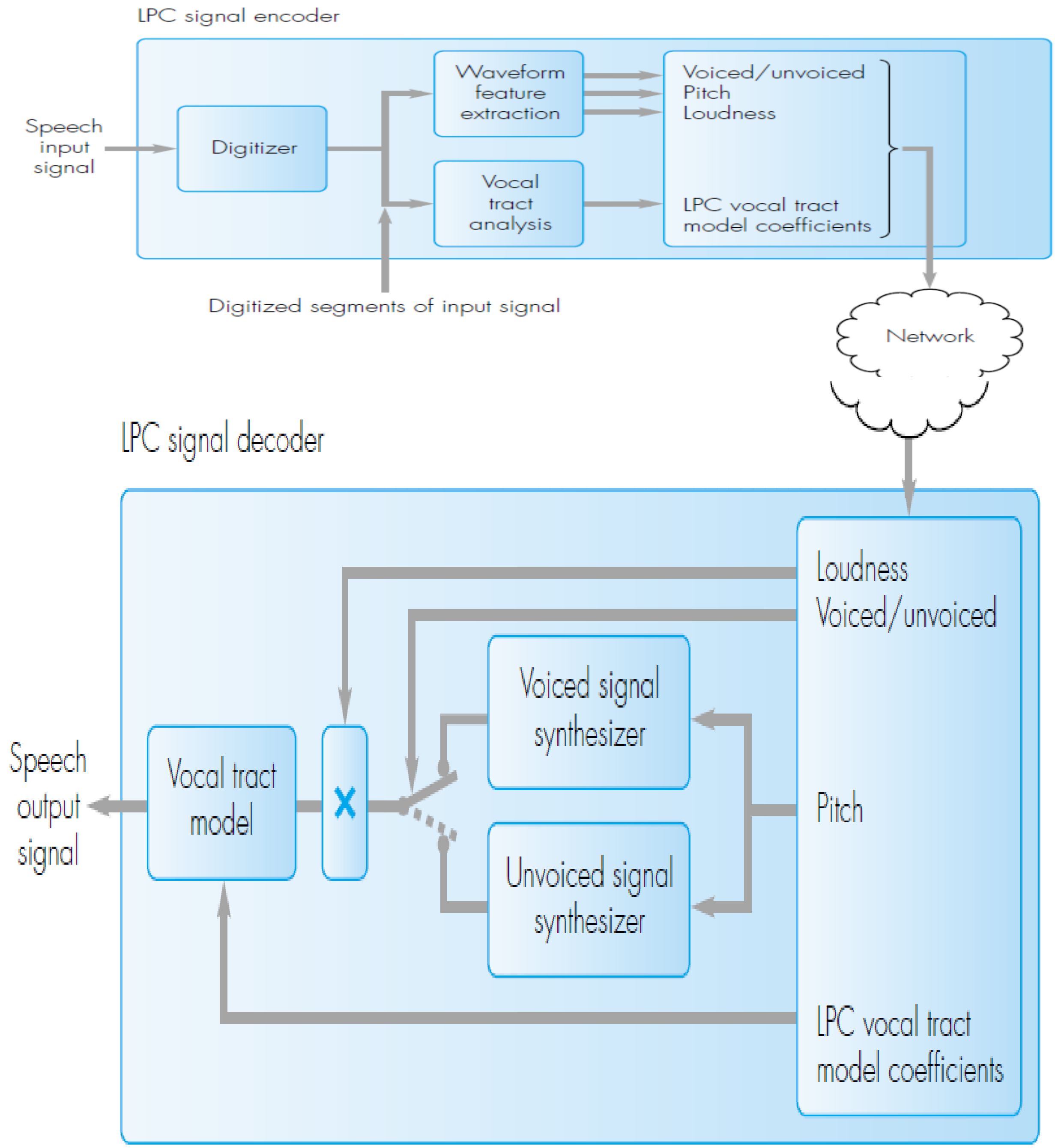
The speech signal generated by the vocal tract model in the decoder is the present o/p signal of speech synthesizers and linear combination of previous set of model coefficients

Hence the vocal tract model is adaptive

Encoder determines and sends a new set of coefficients for each quantized segment

The output of encoder is a set of frames, each frame consists of fields for pitch and loudness

Bit rates as low as 2.4 or 1.2 kbps. Generated sound at these rates is very synthetic and LPC encoders are used in military applications, where bandwidth is important


LPC signal encoder
Speech input signal
Digitizer
Waveform feature extraction
Vocal tract analysis
Voiced/unvoiced
Pitch
Loudness
LPC vocal tract model coefficients
Digitized segments of input signal
Network
LPC signal decoder
Loudness
Voiced/unvoiced
Pitch
LPC vocal tract model coefficients
Voiced signal synthesizer
Unvoiced signal synthesizer
X
Vocal tract model
Speech output signal

7. (b)

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

7. (c) Different types of Frames

- Frame type
  - I-frame- Intracoded

- I-frames are encoded without reference to any other frames
- GOP:The number of frame between successive I-frames

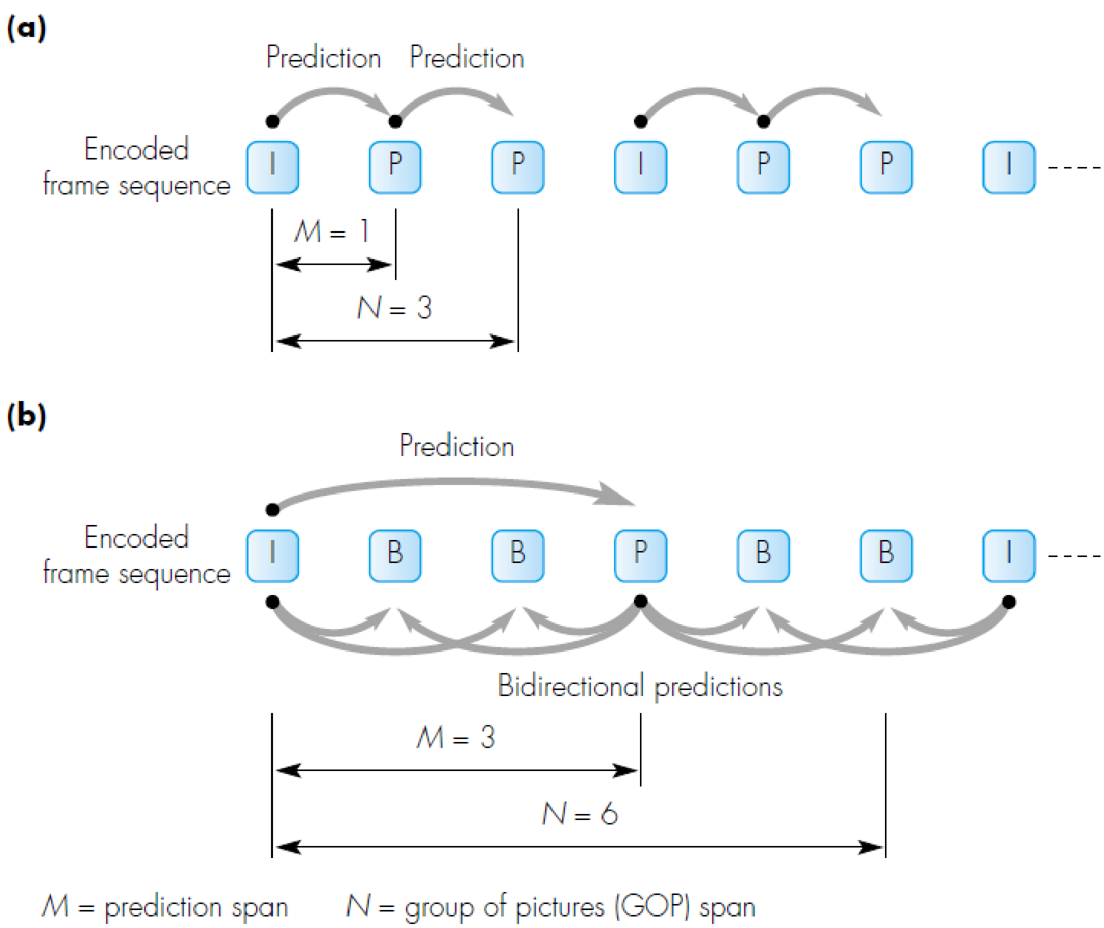
– P-frame:intercoded

- 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

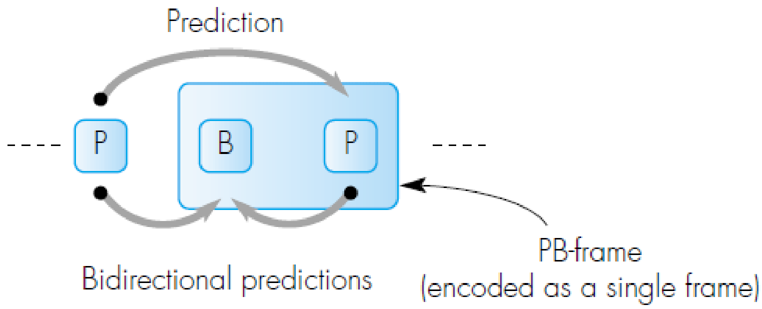
-B-frame:their contents are predicted using search regions in both past and future frames

-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



(c)

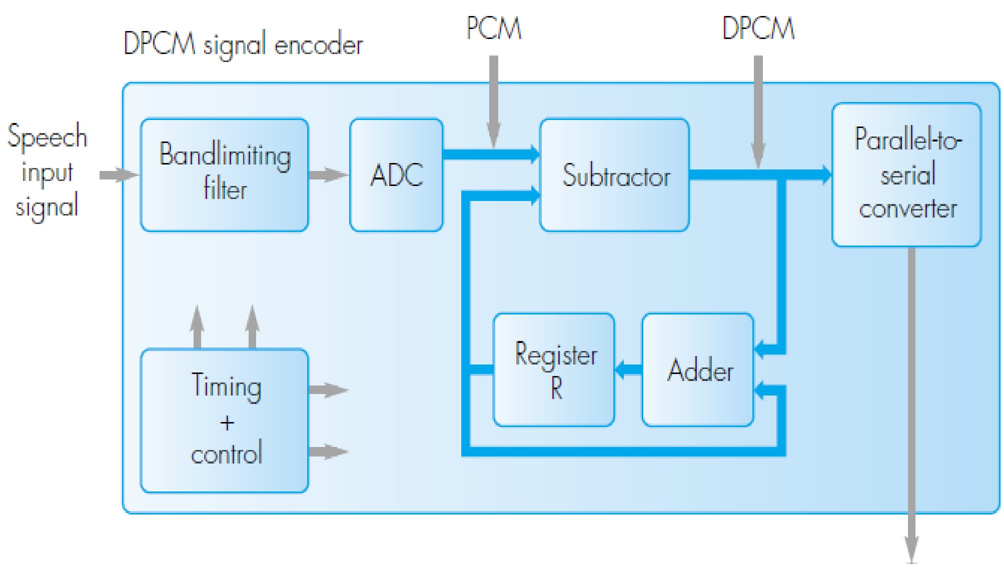


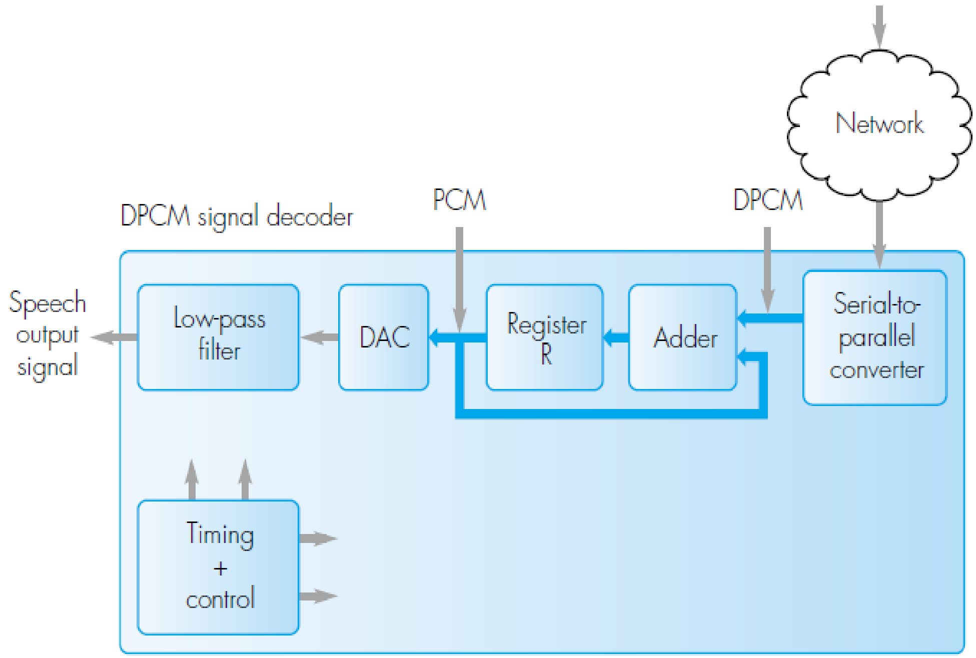
8. (a) **DPCM Encoder and Decoder**

- DPCM is a derivative of standard PCM

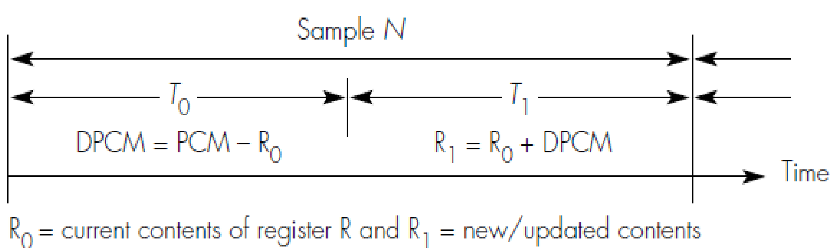
For most audio signals, the range of the differences in amplitude between successive samples of the audio waveform is less than the range of the actual sample amplitudes

(a)





(b)



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

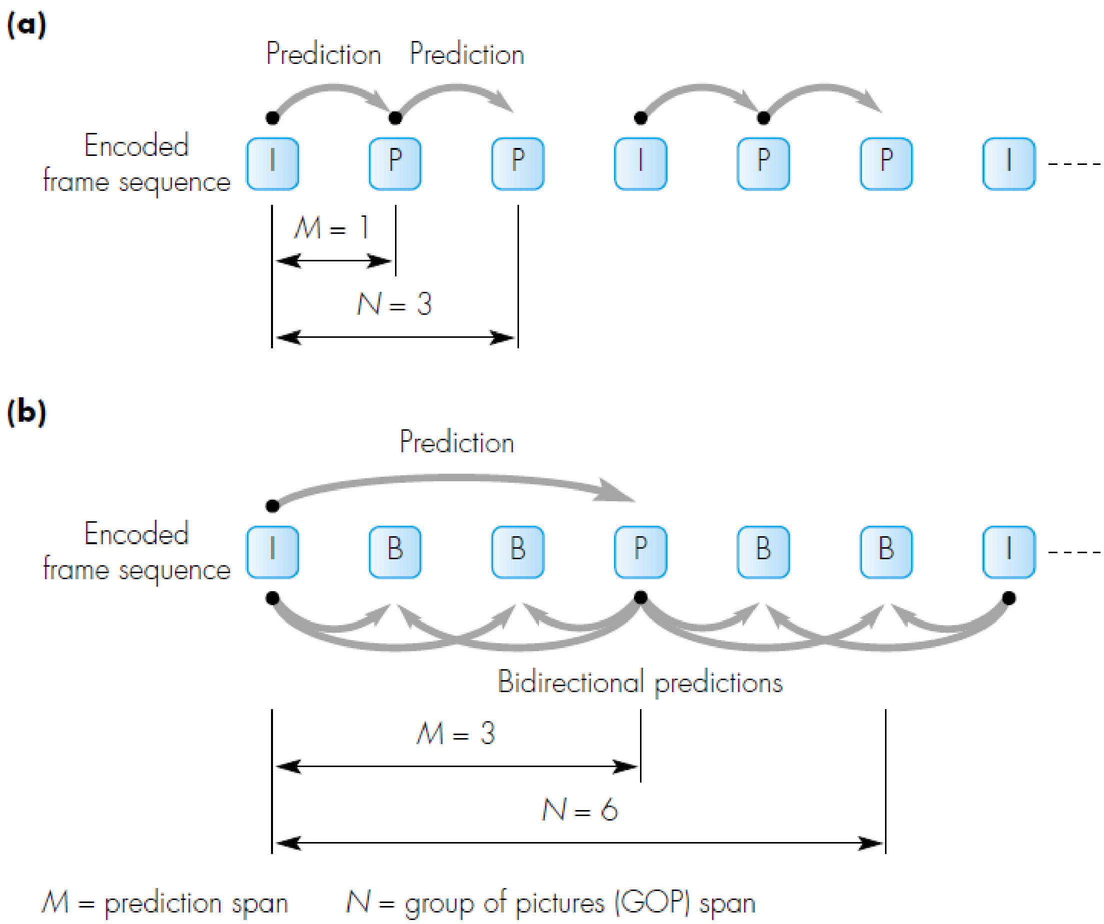
There are schemes to predict the more accurate previous signal

The proportions used are determined by predictor co-efficients

8. (b)

(i) Group of Pictures

The number of frame between successive I-frames



(ii) Prediction Span

The number of frames between a P-frame and the immediately Preceding I or P frame.

(iii) Motion Compensation

Motion compensation uses the knowledge of object motion so obtained to achieve data compression

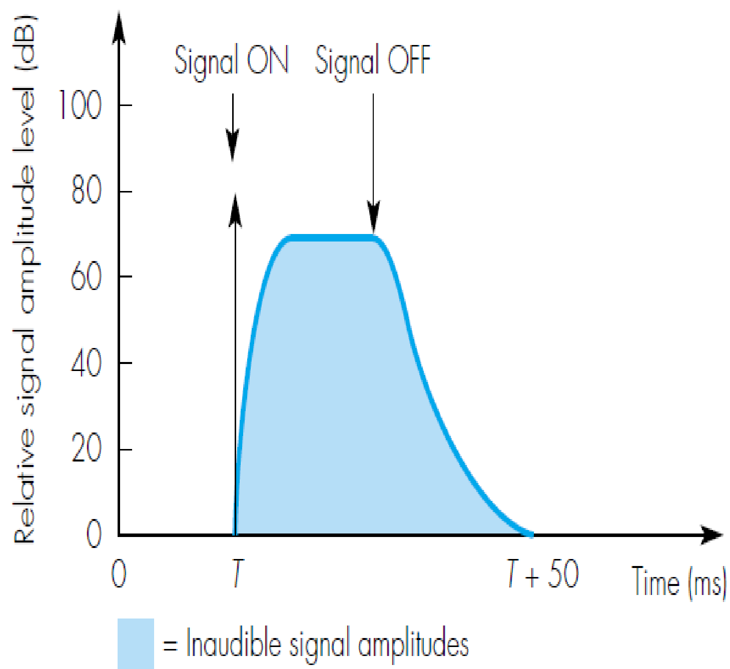
(iv) Motion Estimation

Motion estimation examines the movement of objects in an image sequence to try to obtain vectors representing the estimated motion.

(V) Temporal Masking

- When the ear hears a loud sound,it takes a short but finite time before it can hear a quieter sound
- Masking effect varies with freq-

– effect of temporal masking – signal amplitude decays after a time period after the loud sound ceases and at this time signal amplitude less than decay envelope will not be heard.



–

Module 5

9. (a) Scalable Rate Control

Challenge in multimedia application- how to deliver multimedia streams to users with minimal replay jitters

Network based multimedia system-Layered structure system:

- Application Layer(top)
- Compression Layer
- Transport Layer
- Transmission Layer

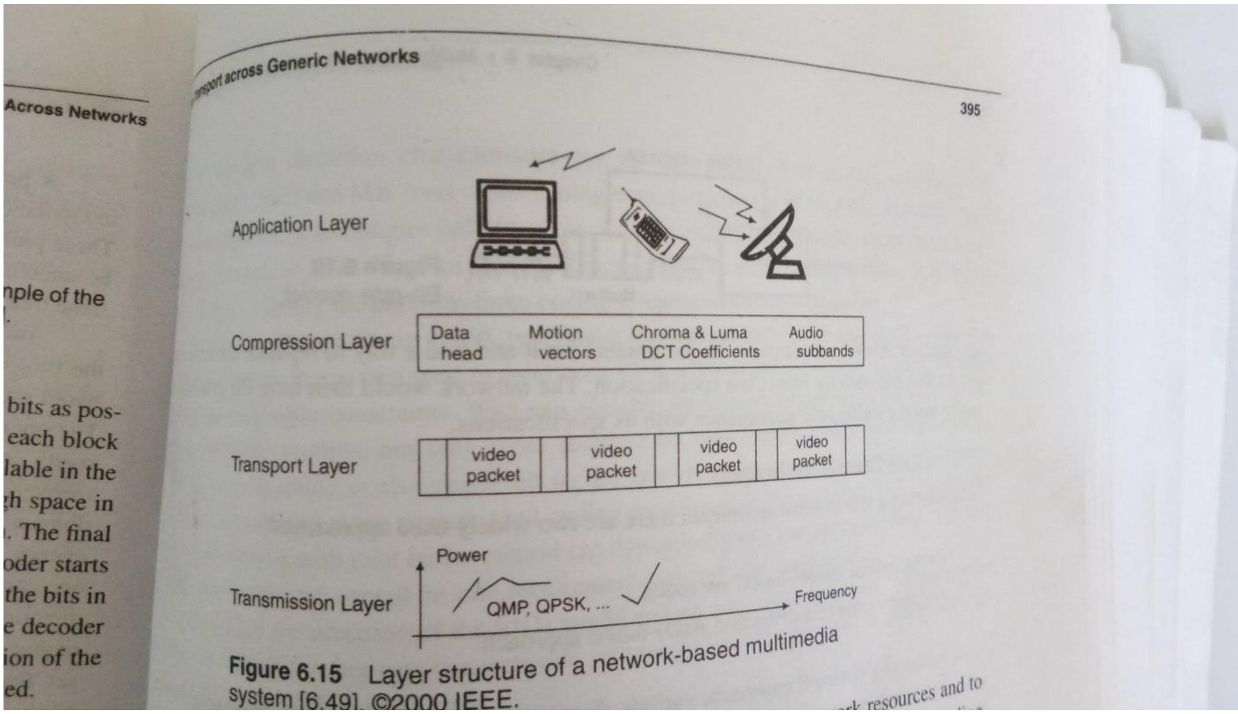


Figure 6.15 Layer structure of a network-based multimedia system [6.49]. ©2000 IEEE.

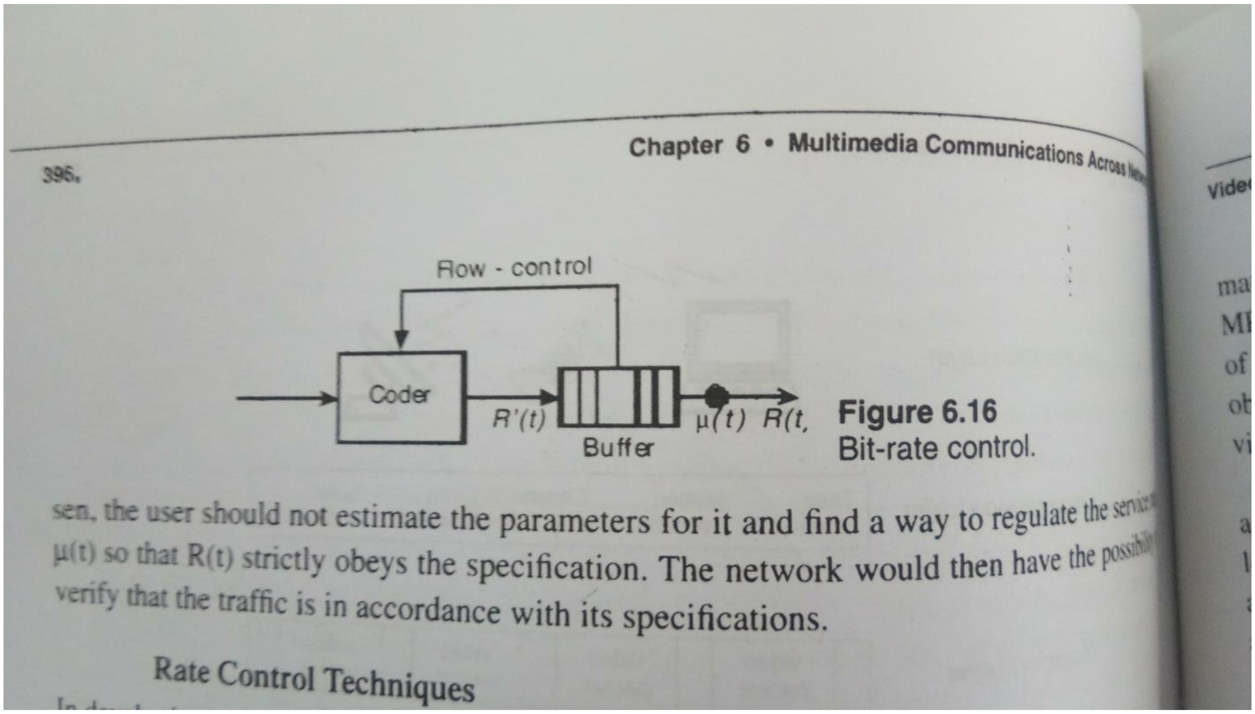
Two techniques to reduce the impact of Jitter on Video Quality:

- Traffic Shaping-Transport Layer approach

  Traffic Pattern is shaped with desired characteristics such as maximal delay bounds, peak rate etc.

- Scalable Rate Control(SRC)-Compression Layer approach

  source video sequence is compressed as per application's requirement and available network resource


396.

Chapter 6 • Multimedia Communications Across Ne

Flow - control

Coder

R'(t)

Buffer

μ(t) R(t,

Figure 6.16
Bit-rate control.

sen, the user should not estimate the parameters for it and find a way to regulate the servic μ(t) so that R(t) strictly obeys the specification. The network would then have the possibil verify that the traffic is in accordance with its specifications.

Rate Control Techniques

In the fig shown in previous slide:

bit stream from the coder is fed into a buffer at a rate R'(t), served at some rate μ(t),so that the output R(t) meets the specified behavior

Bit stream is smoothed by the buffer whenever the service rate is below the input rate

Size of the buffer-determined by delay and implementation constraints

Traffic shaping and SRC together finds an appropriate way of bit stream description such that output R(t) will meet the specification required

Two techniques of Rate Control:

- Analytical model-based approach

  Various distribution characteristics of the signal are considered. Leads to a theoretical Optimization solution which is difficult to implement

- Operational rate distortion R(D) based approach

  Practical coding technique. Optimization solutions are developed using dynamic programming or Lagrangian Multipliers

In the R(D) model-Distortion is measured in terms of quantization parameter

Rate control consists of 4 stages:

Initialization, Pre-encoding, Encoding and Post-encoding

Initialization:

- Buffer size based on Latency
- Subtracting the bit counts of the first frame from the total bit counts
- Buffer fullness in the middle level

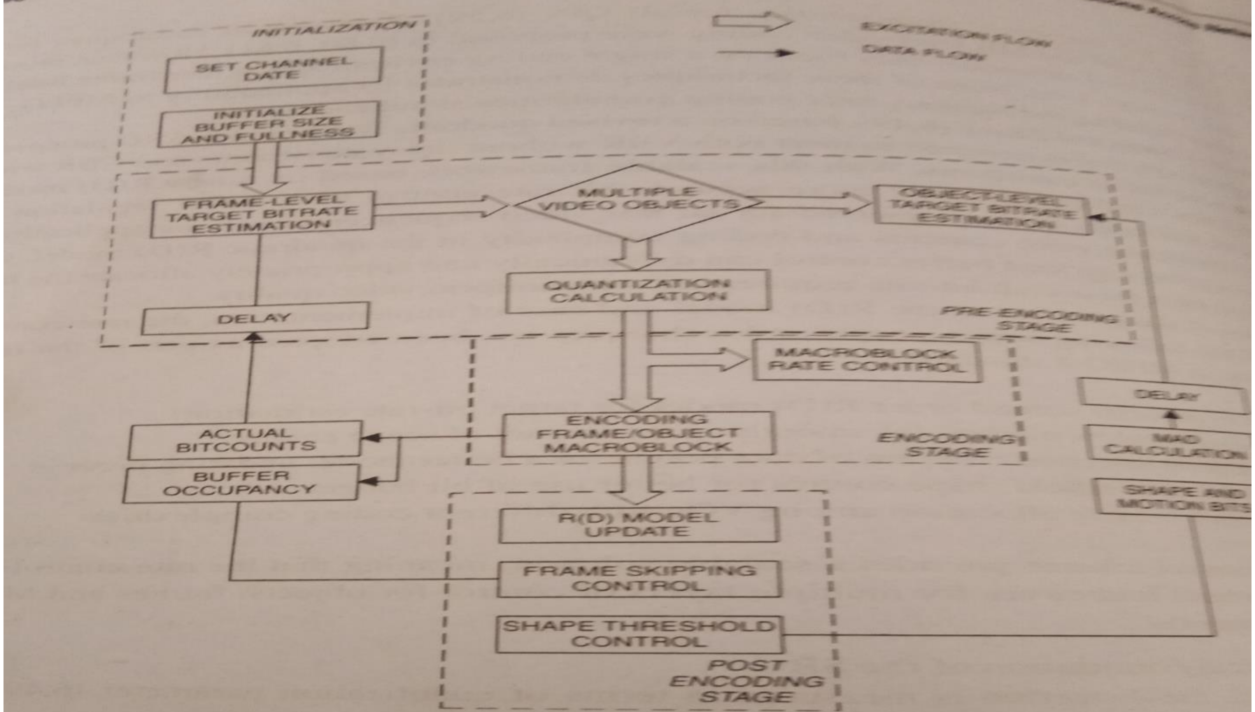


Figure 6.17 Block diagram of the SRC [6.49]. ©2000 IEEE.

ere is a closed-form solution for the R(D) functions derived in Viterbi

$$R(D) = \ln\left(\frac{1}{\alpha D}\right)$$

2

Assuming that the source statistics are Laplacian distributed,

$$P(x) = \frac{\alpha}{2} e^{-\alpha|x|}, \text{ where } -\infty < x < \infty \longrightarrow (1)$$

The distortion measure is defined as,

$$D(x, \tilde{x}) = |x - \tilde{x}| \longrightarrow (2)$$

R(D) is given by,

$$R(D) = \ln\left(\frac{1}{\alpha D}\right) \longrightarrow (3)$$

where

$$D_{min} = 0, \quad D_{max} = \frac{1}{\alpha}, \quad 0 < D < \frac{1}{\alpha} \longrightarrow (4)$$

Based on the above observations, new model is formulated as

$$R_i = \alpha_1 Q_i^{-1} + \alpha_2 Q_i^{-2} \longrightarrow (5)$$

$R_i \rightarrow$ total bits used for encoding the current frame i

$Q_i \rightarrow$ Quantization level used for the current frame i

$\alpha_1, \alpha_2 \rightarrow$ first and second order coefficients.

Video sequence is encoded first as an I-frame and subsequently as P-frames

Pre-encoded stage:

- Target bit estimation, adjustment of target bit based on buffer status for each QP and VO

Encoding Stage:

Encoding the video frame,recording all actual bit rates and activating the MB layer rate control

Post-encoding stage:

- Updation of Quadratic model, shape-threshold control

9. (b) Video Streaming Architecture

Streaming video across the Internet (OR) video Streaming Architecture

* 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. Full file transfer in the download mode usually suffers long and 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.
* Streaming video has six main concepts:
(i) video compression (ii) Application-layer QoS control
(iii) continuous media distribution services
(IV) Streaming servers (V) media synchronization mechanisms (VI) Protocols for streaming media.

* Raw video and audio data are precompressed by video compression and audio compression algorithms and saved in storage devices.
* Upon the client's request, a streaming server

retrieves compressed audio/video data from storage devices.

* 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.

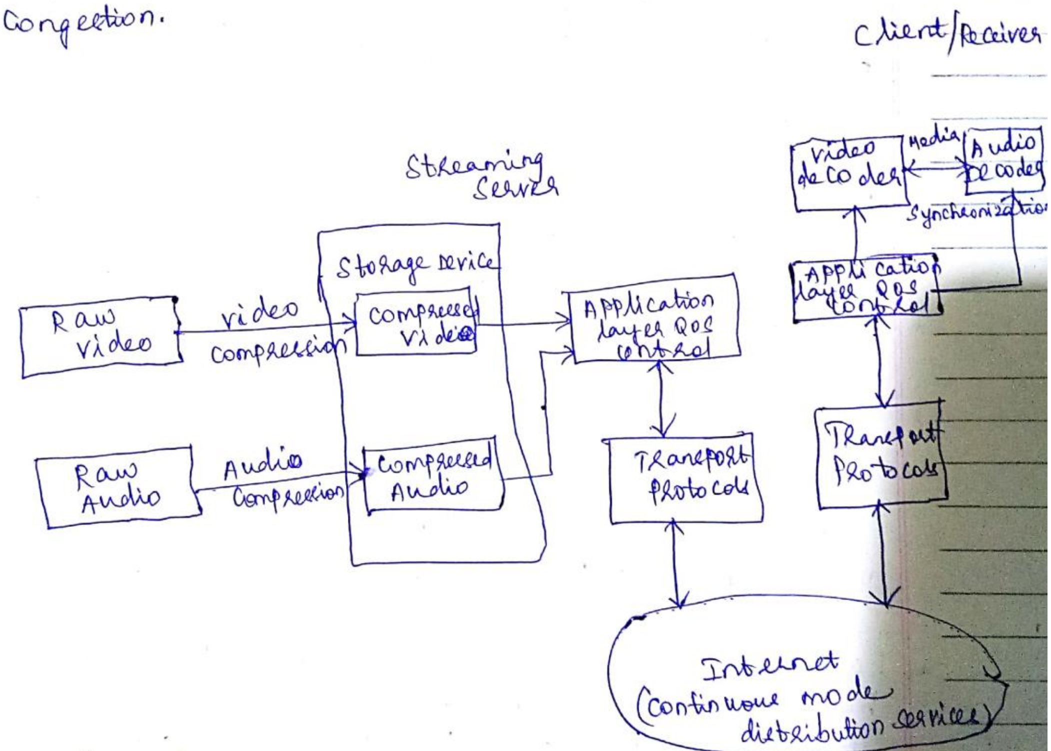


Fig: video streaming Architecture

Date
Page No.

* 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.

10. (a) Integrated Packet Networks

Integrated packet networks

* The effective integration of speech and other signals, such as graphics, image and video into an Integrated packet network (IPN) can rearrange network design properties.
* One of the main goals in IPNs is to construct a model that considers the entire IPN (transmitters, packet multiplexers and receivers) as a system to be optimized for higher speeds and capabilities.

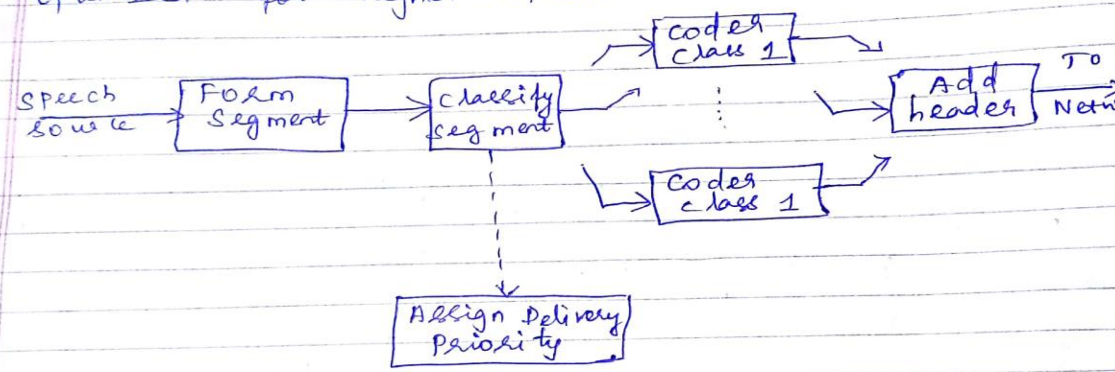


Fig: Transmitter subsystem

* The transmitter first classifies speech segments according to models of the speech production process (voiced sounds, fricatives and plosives).
* After classification, the transmitter removes redundancy from the speech using a coding algorithm based on the determined model.
* After coding, the transmitter assigns a delivery priority to each packet based on the Quality of regeneration possible at the receiver.

* In packets, the delivery priority would be included in the network portion of the packet header. The classification and any coding parameters would be included in the end-to-end portion of the header.

* Packet multiplexers exist at each outgoing link of each network node as well as at each multiplexed network access point.

* Each packet multiplexer monitors local overload and discards packets, according to packet delivery priority.

* In some cases arriving packets, and in some others already queued packets are discarded.

* In addition if error checking is performed by the nodes, any packet found to have an error is discarded.

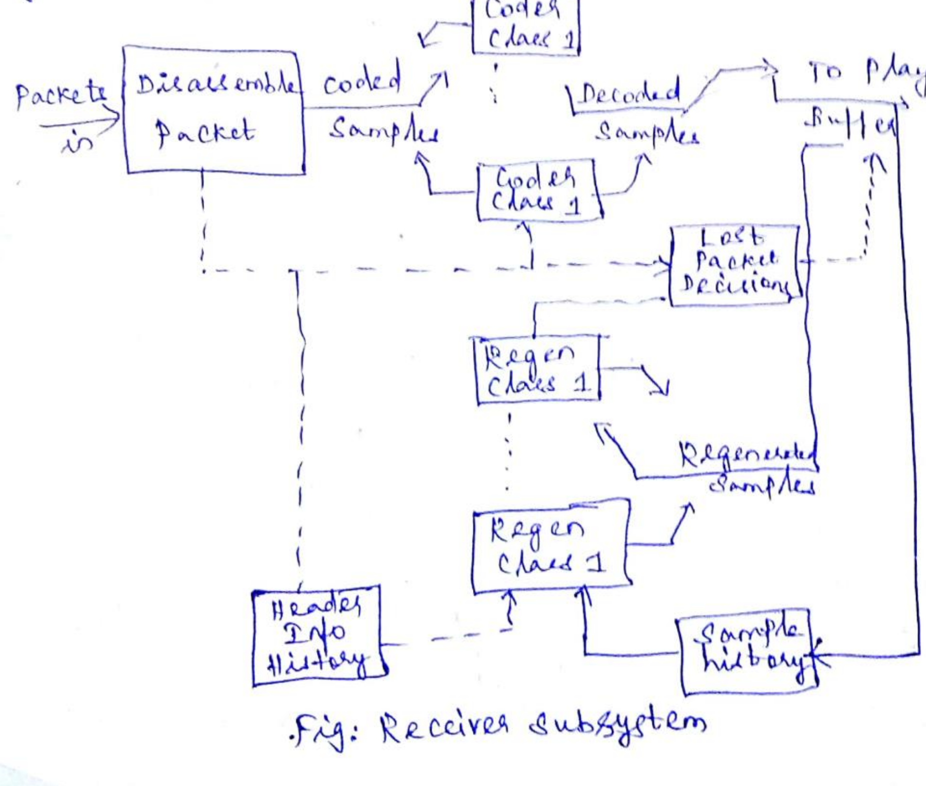


Fig: Receiver subsystem

* The receiver decodes the samples in speech packets delivered to it based on the classification and coding parameters contained in the end-to-end header.
* It also determines the appropriate time to play them out.
* The receiver synchronisation problem requires only packet sequence numbers. If a packet is lost for any reason, the receiver must first detect the loss by inspecting sequence numbers of those packets that have been received.
* In a string of packets with the same class, we can virtually ensure that the first packet will be received by assigning it a high delivery priority. Assuming perfect delivery of these first packets the class of any lost packet will match the class of the last received packet.

* Advantages of Integrated packet network are:
(i) A powerful overload control mechanism is provided.
(ii) Structure of speech is effectively exploited
(iii) only one packet per speech segment is required.
(iv) Receiver speech synchronization is simplified.

10. (b) Errors and Losses in ATM

- In the ATM Networks ,a cell can be lost due to two reasons:

(i)Channel errors

(ii)Limitations of network capacity and Statistical multiplexing

- If an uncorrectable error occurs in the address field of an ATM cell, the cell will not be delivered to the right destination and the cell is considered to be lost.
- A Buffer can be used to absorb the instantaneous traffic peak to some extent, but the buffer overflow in case of congestion.
- In the case of network congestion or buffer overflow, the network congestion control protocol will drop cells.
- Cell discarding can occur on the transmitting side if the no.of cells generated are in excess of allocated capacity or it can occur on the receiving side if a cell has not been arrived within the delay time of the buffer memory.
- In such cases 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 sub sampling and interlacing.
- ATM have cells and packets as multiplexing units that are shorter than a full cell.
- By means of network framing, appropriate control information is added to each multiplexing unit.
- Network framing is used to detect and to possibly correct lost and corrupted multiplexing units
- Errors may be detected by a CRC of sufficient length.
- Loss is detected by means of sequence numbers which turn it into erasures.
- Sequence number is based on the number of transferred data octets.
- Errors and losses can be identified by a CRC on the application frame after reassembly.
- It is important that frame length is known a priori because the frame length of a faulty frame is uncertain.
- Failed CRC could be caused by a bit error.
- A lost or corrupted network frame would be retransmitted

There are complications in retransmission for video:

- The delay requirements might not allow it because it adds another round trip delay which violates the end to end delay requirements.
- The jitter introduced is higher than that induced by queuing.

Several reasons to be cautious of FEC of cell and packet loss:

- Adds a complex function to the system which will reflect in cost.

- The interleaving adds delay.
- Loss caused by multiplexing overload is likely to be correlated as more traffic bursts can occur which the code cannot correct.

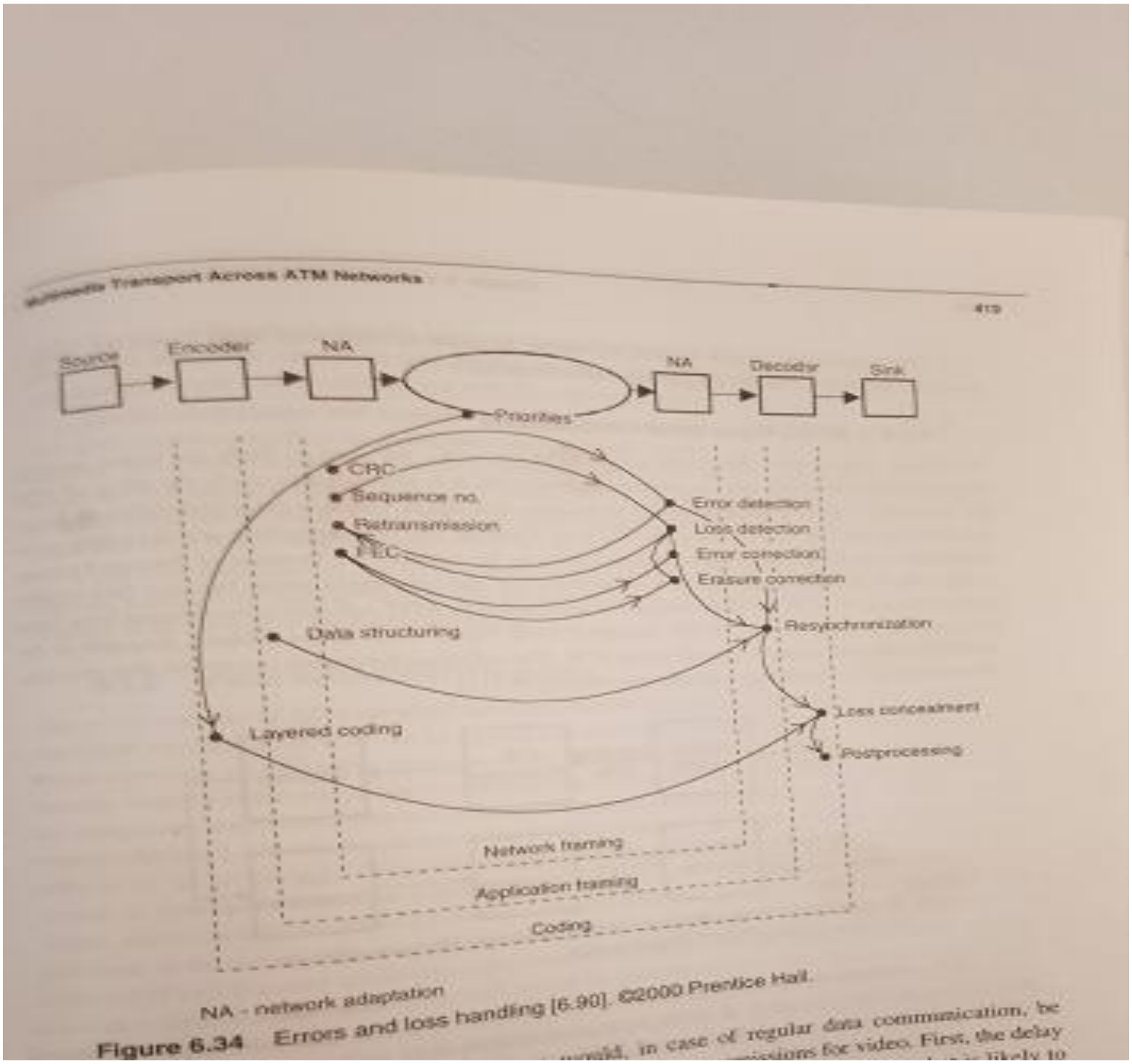


Figure 6.34 Errors and loss handling [6.90]. ©2000 Prentice Hall.