CMR INSTITUTE OF TECHNOLOGY, BENGALURU.
ACCREDITED WITH A++ GRADE BY NAAC

USN

Internal Assessment Test 2 – May 2025								
Sub:	**MULTIMEDIA COMMUNICATION**			Sub Code:	**BEC613A**	Branch:	ECE	
Date:	05-2025	Duration:	90 minutes	Max Marks:	50	Sem/Sec:	6^{th} (A,B,C,D)	OBE

		MARKS	CO	RBT
	ANSWER ANY 5 FULL QUESTIONS	MARKS	CO	RBT
1	**A digitized video is to be compressed using the MPEG-1 standard. Assuming a frame sequence of : IBBPBBPBBPBBI, and average compression ratio of 10:1 (I), 20:1 (P) and 50:1 (B), derive the average bit rate that is generated by the encoder for both NTSC and PAL digitization formats.**	10	CO 2	L3
2	**Explain principle of linear predictive coding with a neat block schematic.**	10	CO 2	L2
3	**Explain MPEG-4 coding principles with the help of a neat diagram.**	10	CO 2	L2
4	**Explain the following terms : 1) Group of Pictures, 2) Prediction Span 3) Motion Compensation, 4) Motion Estimation, 5) Temporal Masking.**	10	CO 2	L1
5	**Explain the working of CSMA/CD protocol with suitable diagrams.**	10	CO 3	L2
6	**Explain IEEE 802.3 frame format with suitable figure**	10	CO 3	L2
7	a) **Assuming signal propagation delay in the fiber of 5 microseconds per 1 km, derive the latency of the following FDDI ring configurations in both time and bits assuming a usable bit rate of 100 Mbps. 1) 2 km ring with 20 stations, 2) 20 km ring with 200 stations, 3) 100 km ring with 500 stations.** b) **Explain principle of operation of LZW algorithm**	4 + 6	CO 3	L1

Solution:

Q 1. A digitized video is to be compressed using the MPEG-1 standard. Assuming a frame sequence of: IBBPBBPBBPBBI, and average compression ratio of 10:1 (I), 20:1 (P) and 50:1 (B), derive the average bit rate that is generated by the encoder for both 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 = $1.013760 \times 1/29.24$
= 34.670 Kbits per frame

Hence bit rate generated at 30fps = 1.040 Mbps

PAL frame size:

Without compression= $352 \times 288 \times 8 + 2\,(176 \times 144 \times 8)$
= 1.216512 Mbits per frame

With compression = $1.216512 \times 1/29.24$
= 41.604 kbits per frame

Hence bit rate generated at 25 fps = 1.040Mbps

Normally, allowing for packetization and multiplexing overheads, a bandwidth of 1.2 Mbps is allocated for the video. Hence, assuming a maximum bit rate of 1.5 Mbps, this leaves 300 kbps for the compressed audio stream.

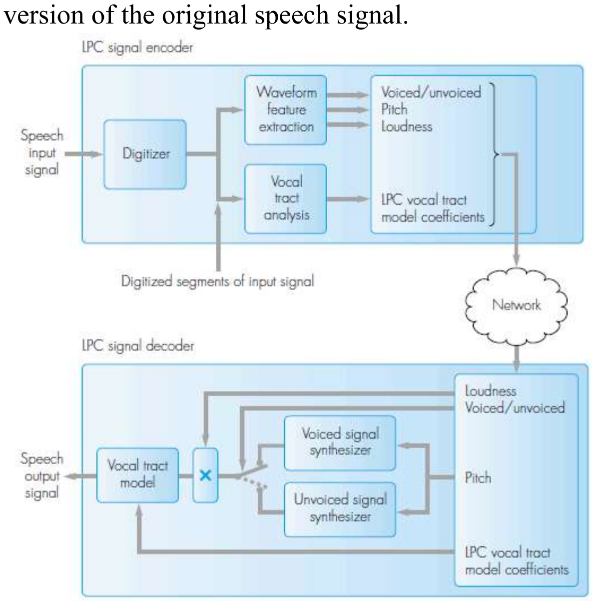
Q2. Explain principle of linear predictive coding with a neat block schematic.

Answer: DSP circuits help in analyzing the signal based on the required features (perceptual) and then quantized to send to the destination. The **destination** uses them together with a sound synthesizer to **regenerate a sound** that is perceptually comparable with the source audio signal. This is the **basis of the linear predictive coding technique.** With this generated sound, **very high levels of compressions** achieved.

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

- **Pitch:** refers to the fundamental frequency, or the rate of vibration of the vocal cords, which determines the perceived highness or lowness of a sound. Ear is sensitive to frequencies in the range 2-5KHz as compare to outside this range.
- **Period:** duration of the signal
- **Loudness :** this is determined by the amount of energy in the signal
- **Vocal tract excitation parameters**: origins of the sound.
- These are classified as:
- **Voiced sounds**: generated thro' the vocal chords and examples include sounds relating to letters **m, v** and **l**
- **Unvoiced sounds**: vocal chords are open example Sounds relating to **f** and **s**

These are obtained from the speech waveform and once obtained shall be used with a suitable model of Vocal tract to generate synthesized version of the original speech signal.



The i/p waveform is first sampled and quantized at a defined rate. A block of digital samples-segments is analyzed to determine the various perceptual parameters of the speech that it contains. 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 the encoder is a set of frames; each frame consists of fields for pitch and loudness.

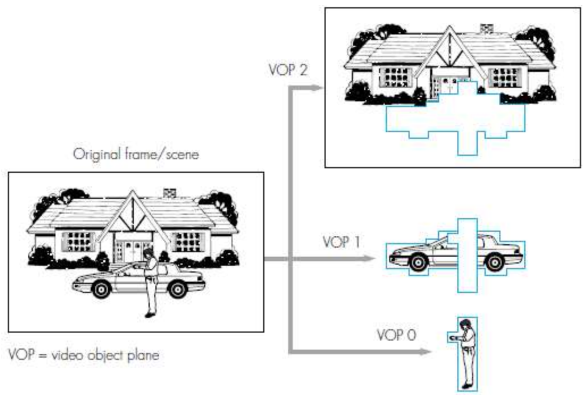
The period is determined by the sampling rate - a notification of whether the signal is voiced or unvoiced and a new set of computed model coefficients. Some LPC encoders use **up to ten set of previous model** coefficients to predict the output sound and use bit rates as low as 2.4kbps or even 1.2 kbps. Generated sound is very synthetic.

Application: **military applications** where BW is all important

Q3. Explain MPEG-4 coding principles with the help of a neat diagram.

Answer:

- MPEG-4 has:
  - **Content-based functionalities-**each scene is defined in the form of background and one or more foreground objects **-Audio-visual object(AVOs)**
  - **Each AVO-** is defined as one or more audio and video objects
  - **Object descriptor-** each audio and video object origin description is required for manipulation
  - **Binary format for scenes (BIFS)-** language used for modifying objects(example for video objects changing shape, color, appearance and for audio objects changing volume etc..)
  - **Scene descriptor-** contains composition of a scene.
  - Defines how the various AVOs are related to each other in the context of the complete scene
  - Figure 4.23- A frame/scene is defined in the form of a number of AVOs
  - Each video frame is segmented into a number of **Video object planes(VOPs)**, each of which corresponds to an AVO of interest



- In the example frame is shown as consisting of three VOPs:

➢ VOP 0 – To represent the person approaching the car

➢ VOP 1 – the remainder of the car parked outside the house

➢ VOP 2 – the remainder of the background

➢ Each VOP is encoded separately based on its shape, motion and texture

➢ Each VOP is encapsulated within a rectangle. It is chosen so that it completely covers the related AVO using the minimum number of macroblocks

➢ EACH VOP in the rectangle which refers the related AVO WITH min MB.

➢ The motion and texture of each VOP is also encoded and the bitstream is multiplexed together with related object and screen descriptor Information .

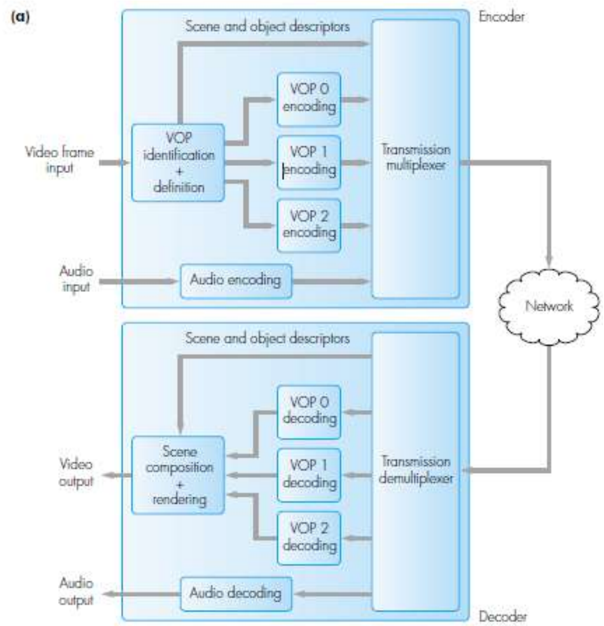
➢ At the receiver the bitstream is demultiplexed and individual stream decoded.

➢ Decompressed information , screen descriptor and object information together creates the video frame

➢ **AUDIO AND VIDEO COMPRESSION**

➢ Audio associated with an AVO is compressed using one of the audio compression algorithms and it depends on the available bandwidth/ bit rate of the transmission channel and the sound quality required.

➢ Different algorithms like G.723.1, DOLBY AC3, MPEG LAYER 2 USED for different applications



Q4. Explain the following terms : 1) Group of Pictures, 2) Prediction Span 3) Motion Compensation, 4) Motion Estimation, 5) Temporal Masking.

Answer:

1) Group of Pictures: It is a sequence of I,P or B pictures / frames. GOP has a start code , time stamp for synchronization, parameter frame – sequence of frame types

2)Prediction span is the distance between P frame and immediately preceding I or P frame.

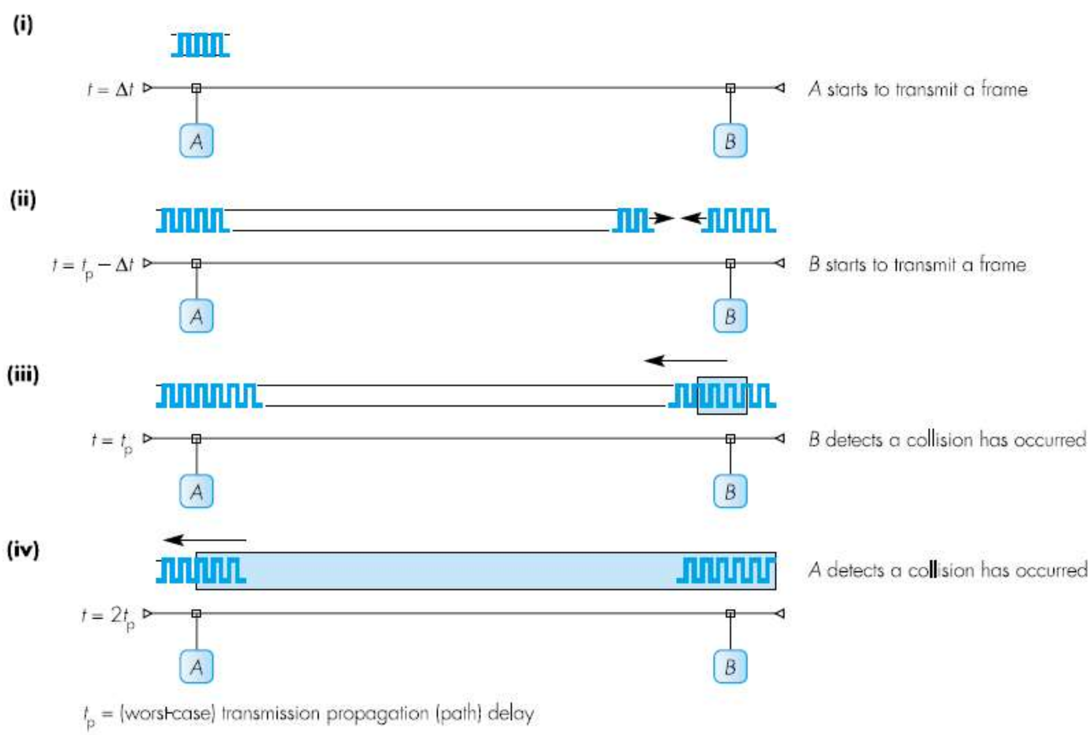
3)Motion Estimation: The process of finding motion vectors that describe the transformation from one frame to another.

4) Motion Compensation: The use of motion vectors to reconstruct or predict frames based on reference frames.

5)Temporal Masking.when the ear hears a loud sound,it takes a short but finite time before it can hear a quieter sound.

Q5 Explain the working of CSMA/CD protocol with suitable diagrams.

Answer:



All the stations are attached directly to the same cable/bus, it is said to operate in a multiple access mode. The bus operates in the broadcast mode which means that every frames transmitted is received by all the other stations that are attached to the bus. Because of the broadcast mode, this will result in the contents of the two frames being corrupted and a collision is said to have occurred.

Working:

– **Step 1:** Check if the sender is ready to transmit data packets.

– **Step 2:** Check if the transmission link is idle. If idle then

The sender sends dummy data on the link. If it does not receive any collision signal, it sends the data. Otherwise, it refrains from sending data.

– **Step 3:** Transmit the data & check for collisions.

– During transmission, if a collision signal is received by the node, transmission is stopped.

– The station then transmits a jam signal onto the link and waits for random time intervals before it resends the frame.

– After some random time, it again attempts to transfer the data and repeats the above process.

– **Step 4:** If no collision was detected in propagation, the sender completes its frame transmission and resets the counters.

– Collision Signal is generated by Physical layer.

– Jam signal (collision enforcement): To make sure that all stations involved in the collision will detect collision. This is a pattern of 32 bits.

– Collision backoff and retransmission method (Truncated Binary Exponential Backoff Algorithm, BEBA):

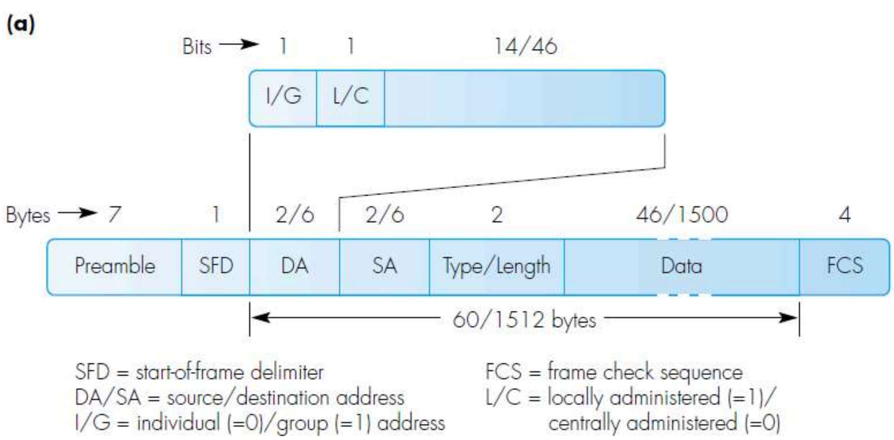
– R : Random delay time (unit: slot time) between $0 <= R < 2^k$

– K : Min (N,10(backoff limit)) -- Truncation

– N : number of collisions experienced (N <= 16)

Q6 Explain IEEE 802.3 frame format with suitable figure.

Answer:



• Bit rate	– 10 Mbps (Manchester encoded)
• Slot time	– 512-bit times
• Interframe gap	– 9.6 microseconds
• Attempt limit	– 16
• Back off limit	– 10
• Jam size	– 32 bits
• Maximum frame size (including FCS)	– 1518 bytes
• Minimum frame size (including FCS)	– 512 bits

Preamble field- sent at the head of all frames, for synchronization of bits.
Start of frame delimiter after preamble, single byte and informs the valid frame start.
Destination and source address – MAC address as used by MAC layer
First bit in the destination address specifies the address is individual or group address
Type of grouping is specified in second bit and can be locally or centrally administered
Group address is used for multicasting
Two-byte type field indicates network layer protocol
Length field indicates number of bytes in the data field
Maximum size of data field – MTU (Maximum Transmission Unit)
FCS – Frame Check Sequence used for error detection

Q7. a) Assuming signal propagation delay in the fiber of 5 microseconds per 1 km, derive the latency of the following FDDI ring configurations in both time and bits assuming a usable bit rate of 100 Mbps. 1) 2 km ring with 20 stations, 2) 20 km ring with 200 stations, 3) 100 km ring with 500 stations.

Answer: Ring latency, T_l = Signal propagation delay, $T_p + N \times$ station latency, T_s where N is the number of stations.

For 100 Mbps rate station latency $\cong$1 µs

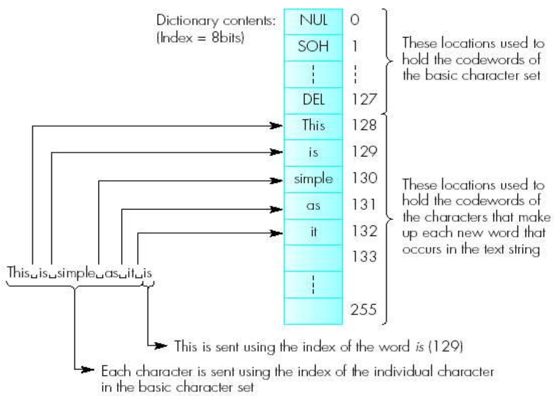
(i) $T_l = 2 \times 5 + 20 \times 1 = 30$µs or 3000 bits

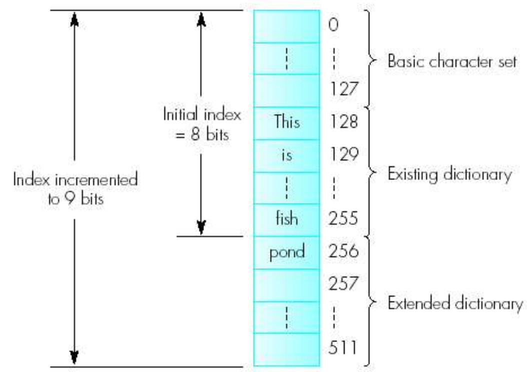
(ii) $T_l = 20 \times 5 + 200 \times 1 = 300$µs or 30000 bits

(iii) $T_l = 100 \times 5 + 500 \times 1 = 1000$µs or 100000 bits

Q7. b) Explain principle of operation of LZW algorithm

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.

The LZW algorithm is a very common compression technique.

LZW compression works best for files containing lots of repetitive data.

This is often the case with text and monochrome images.

Files that are compressed but that do not contain any repetitive information at all can even grow bigger!

LZW compression is fast.

LZW compression can be used in a variety of file formats:

TIFF files

GIF files

- **Problems:**
- Too many bits,
- everyone needs a dictionary,
- only works for English text