1. Topic: Images (20 marks)

a) The GIF (8-bit) format uses a colour lookup table to store up to 256 colours. Describe a scheme to select a representative set of 256 colours from the colours in the source image. (7 marks)

- Not covered.

b) JPEG (baseline) is a standard for lossy compression of still images. Describe the steps performed by a JPEG encoder and explain the step where information is lost during the compression. (7 marks)

- Answered in Sample-13.

c) Compound images consist of a combination of text, graphics and natural pictures. Compressing these images with a single algorithm that simultaneously meets the requirements for this mixed content is difficult. Outline an approach for adaptively compressing these images. (6 marks)

- Not covered.

2. Topic: Video (20 marks)

a) The majority of video codecs used in practice are block-based and employ motion compensation. (4 marks)

- **Explain what is meant by the term block-based encoder and why the technique is used.**

    - Block-based encoder refers to encoding carried out on frames that have been broken up into many groups of pixels which are called blocks e.g. 8x8 pixel blocks. These blocks are referred to as macroblocks where each block is equal sized, non-overlapping and rectangular. Ideally the frame dimensions is a multiple of the macroblock size. Each of these blocks can be compared with the contents of other frames for various reasons, such as to balance the effectiveness of approximating commonly seen motions, error concealment and so on.

- **Explain how motion compensation improves codec performance.**

    - Motion compensation is used to compensate inter-frame differences due to motion. The current frame to be compressed is divided into uniform, non-overlapping macroblocks. Each macroblock in the current frame is compared to the content in other frames, which is used to

discover motion vectors. The motion vector detailing the position of the target macroblock's match is then encoded instead of the macroblock itself.

b) MPEG encoders have two encoding modes for individual video frames — intra (I-frames) and inter (P- and B- frames). In the former, a frame of video is encoded as an independent image without reference to other images in the sequence. In the latter, reference frames are used to predict the values. (10 marks)
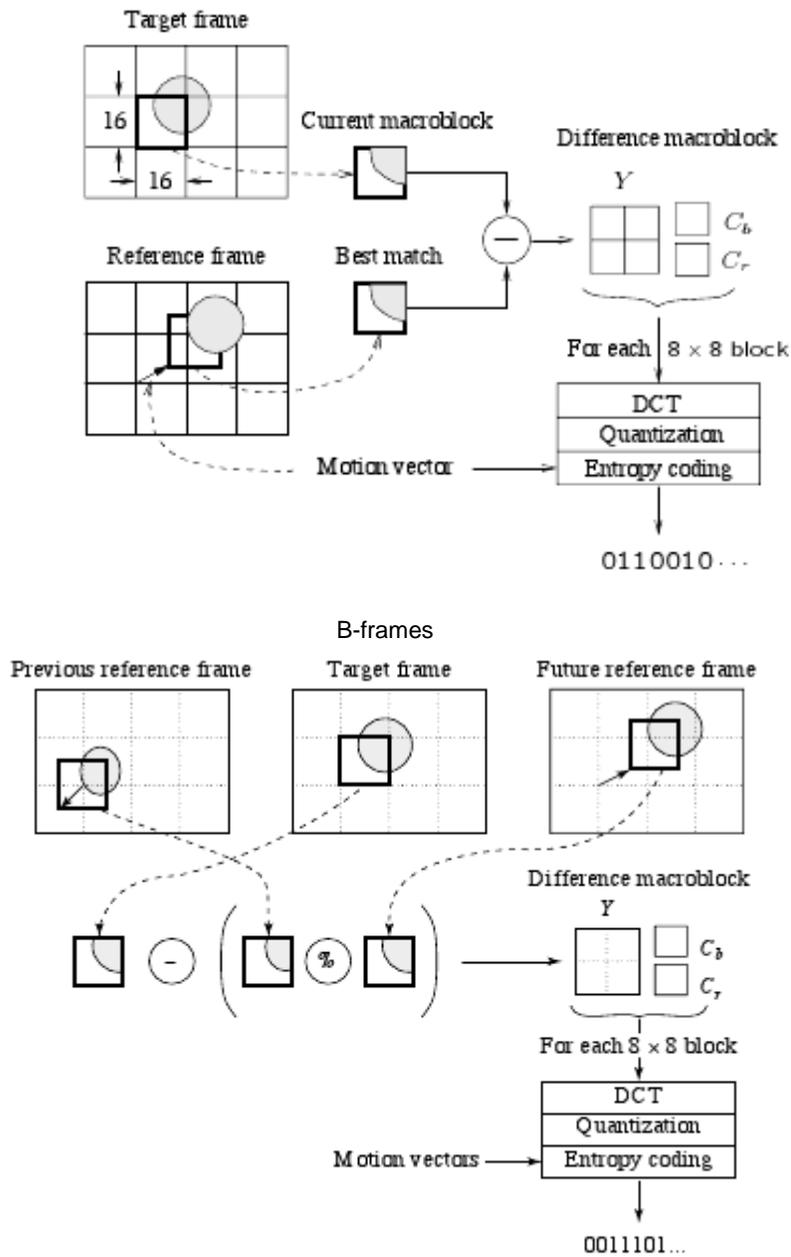
- **Describe using a diagram, the broad steps used in intra frame encoding.**

    - Decompose image into three components in RGB space & convert RGB to YCbCr
    - Divide image into several macroblocks (each macroblock has 6 blocks, 4 for Y, 1 for Cb, 1 for Cr)
    - DCT transform each block
    - Quantize each coefficient
    - Zigzag scan
    - Finally carry out Entropy Coding which involves Huffman encoding, run-length coding etc. both of which are lossless.

Compression of I-frames
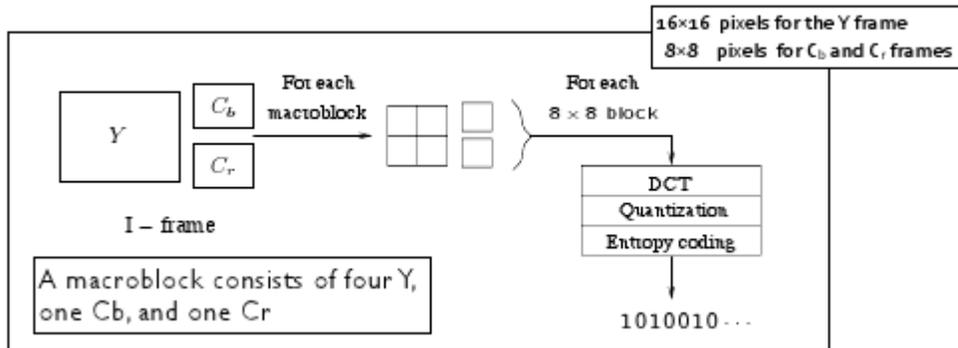
- **Describe using a diagram, the key difference between encoding P- and B- frames.**

    - Each MB (macroblock) in a P-frame may have one motion vector, whereas a B-frame will have up to two motion vectors (one from forward and one from backward). If matching in both directions is successful, both corresponding MBs will be averaged before comparing against the target MB for computing the prediction error. If only one match is made (either forward or backward), its corresponding MB will be used.

P-frames

Target frame

16

16

Current macroblock

Reference frame

Best match

Difference macroblock

Y

C_b

C_r

For each 8×8 block

DCT

Quantization

Entropy coding

Motion vector

0110010...

B-frames

Previous reference frame

Target frame

Future reference frame

Difference macroblock

Y

C_b

C_r

For each 8×8 block

DCT

Quantization

Entropy coding

Motion vectors

0011101...

c) For chromakey content, explain why the presence of flesh colours in the foreground causes difficulties for extracting the uncomposited foreground colour. Describe a technique for overcoming this problem. (6 marks)

- Not part of test (Say's John O'Mullane)

3. Topic: Audio (20 marks)

a) Audio signals are often sampled at different rates. CD quality audio is sampled at 44.1kHz rate while telephone quality audio sampled at 8kHz. What are the maximum frequencies in the input signal that can be fully recovered for these two sampling rates? (2 marks)

- Think just divide frequency by two as per Nyquist.

b) Describe how Pulse Code Modulation (PCM) is used in the coding of audio data. (6 marks)
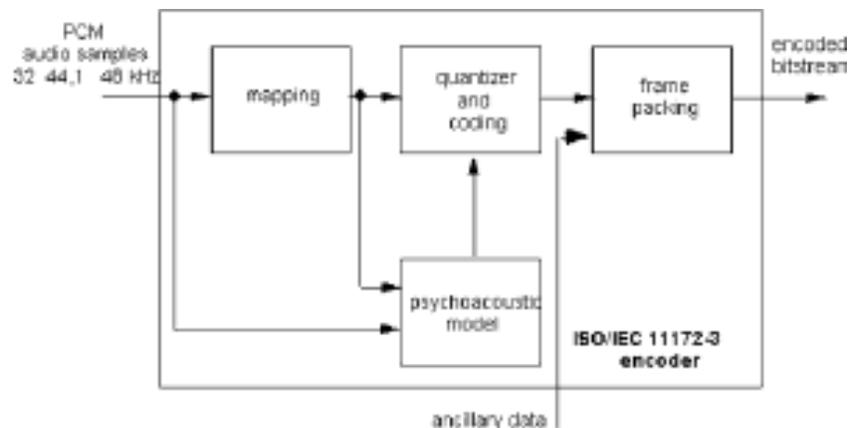
Notes
- The basic techniques for creating digital signals from analogue signals are sampling and digitisation.

Answer
- Pulse Code Modulation (PCM) is the process of
  - Sampling an analogue signal's amplitude at fixed intervals
  - Converting the amplitude into discrete levels (quantisation)
  - Assigning digital codes to represent those levels
- PCM is also used in predictive coding, where differences are transmitted using a PCM system i.e. do not send the sample but the difference between samples.

c) Describe using a diagram the basic MPEG audio compression algorithm. (6 marks)

PCM audio samples 32 44.1 48 kHz → mapping → quantizer and coding → frame packing → encoded bitstream

psychoacoustic model

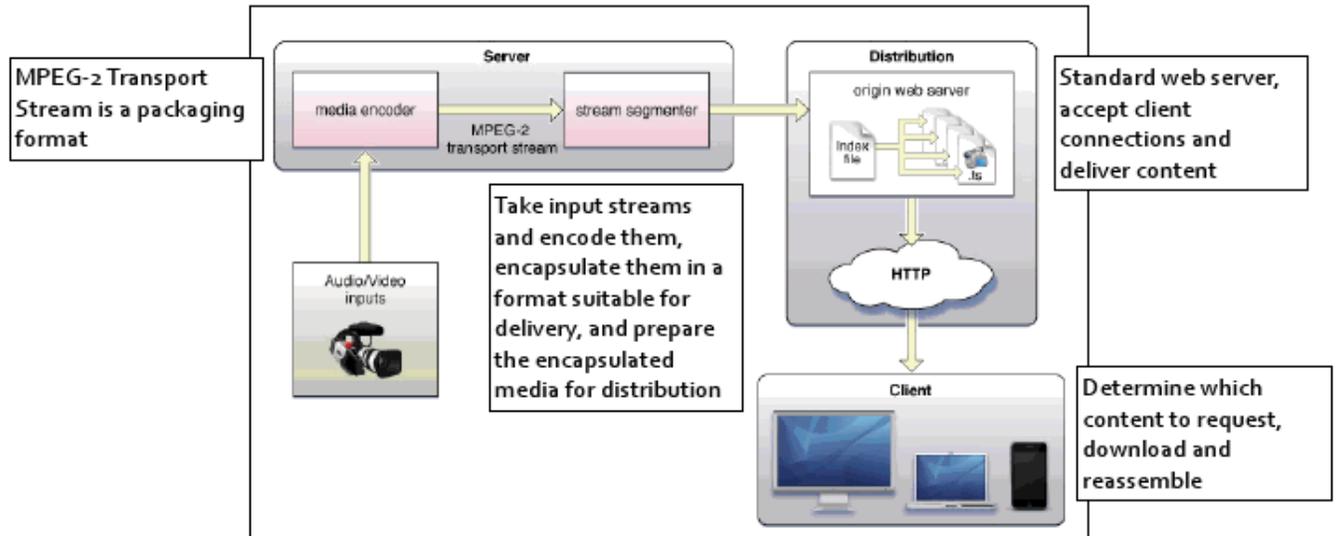ISO/IEC 11172-3 encoder

ancillary data

d) Describe two psychological phenomena that have been exploited in MPEG audio compression. (6 marks)

- Answered in Summer-06. (Frequency & Temporal Masking)

4. Topic: Delivery (20 marks)

a) Describe the essential elements of the HTTP Live Streaming Architecture. (6 marks)

Notes

Answer
- MPEG-2 Transport Stream is a packaging format
- Audio / Video Inputs: The data to be streamed.
- Server (media encoder & stream segmenter): Take input streams and encode them, encapsulate them in a format suitable for delivery, and prepare the encapsulated media for distribution
- Distribution (origin web server & HTTP): Standard web server which accepts client connections and delivers content to the client over HTTP
- Client: Determine which content to request, download and reassemble.

b) Jitter is the variation in the arrival rate of packets on a network connection. Explain how a playout buffer can overcome the effects of jitter. (4 marks)

- Answered in Sample-13.

c) What is meant by the term Forward Error Correction (FEC)? Describe a simple FEC scheme. (4 marks)

- Answered in Sample-13.

d) Explain what makes loss concealment techniques feasible for digital video. (6 marks)

- Pictures are arranged as a group of frames or a group of pictures (GOP).
- Each frame may be similar to its previous or following frame. Motion of an image changes in these frames and so while pixel values may be similar throughout the frames the pixels may be in different places due to motion.
- As a result, various loss concealment techniques such as spatial & temporal concealment are made feasible.