1. Topic: Image Coding (24 marks)

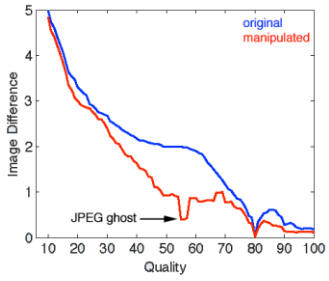
a) An image compressed using JPEG is intended to be viewed by a human observer.

i. Describe how the JPEG algorithm exploits the features of the human visual system to achieve high levels of compression while minimising visual distortion. (6 marks)

- 'Human Visual System Features'.
- Also answers steps taken by JPEG encoder during compression.
- Colour sensitivity - Chroma subsampling is the practice of encoding images by implementing less resolution for chroma information than for luminance information, thus taking advantage of the human visual system's lower acuity for colour differences than for luminance.
- Contrast sensitivity - DCT high frequency: For a typical image, most of the visually significant information about the image is concentrated in just a few coefficients of the DCT. Compression is achieved since the lower right values represent higher frequencies, and are often small - small enough to be neglected with little visible distortion.
- Quantisation - The compressing of image, audio or video files by removing data that does not affect the overall quality of the file i.e. lossy.
- Zig-zag scan - clusters packets of pixel information from low to high frequencies, which changes the 2D matrix into a 1D list. Arrange the coefficients in order of increasing frequency. The higher frequency coefficients are more likely to be reduced to zero after quantisation. This improves the results when using run-length encoding.
- Run Length Coding (lossless)
- Huffman Coding (lossless)

ii. Explain what is meant by the term "JPEG ghosts" and describe how they can be used to detect tampered images. (4 marks)

- This technique can determine what parts of an image underwent double compression.
- A JPEG ghost can be easily uncovered by comparing the image in question to re-saved versions of the image.
- Note that in the former case there are two dips in the graph — an expected dip at 80 (the current quality) and a second dip at 55 (the original compression quality). This second dip is a sign that the image is not an original i.e. the coefficients were previously quantized with a larger quantisation (lower quality).



b) The JPEG File Interchange Format (JFIF) is an image file format for exchanging JPEG encoded files compliant with the JPEG Interchange Format (JIF) standard.

i. Explain why JFIF is needed. (4 marks)

- JFIF enables a JPEG bitstream to be exchanged between a wide variety of platforms and applications.
- A JFIF file consists of JPEG data with a header providing information missing from the JPEG stream i.e. version number, horizontal & vertical pixel density, pixel aspect ratio and an optional thumbnail

ii. Suggest a JFIF extension that allows for transparent JPEG images. (4 marks)

- JFIF allows many extensions, for example the application markers can be utilised in order to incorporate data for transparent images. The JFIF file will contain these extensions and the JPEG data for transparent images, and if the transparent data is not needed by a particular application using the image, this extra data will simply be ignored.

iii. Outline an implementation of this extension that could be used for displaying transparent JPEG images on a web page. (6 marks)

- Mask an image. Use this mask to extract part of an embedded image at particular position. Place in alpha channel. Render alpha channel composite in JPEG image using javascript, css or HTML5 canvas blend modes.
- Produce a canvas element
- Mask the image in question. Use the mask to extract part of the embedded image at a particular position i.e. separate out the JPEG and PNG elements from the image.
- Then you want to re-draw the extracted JPEG and PNG elements in another image.
- This can be accomplished using some elements or a combination of Javascript / CSS / HTML5 canvas element.
    - PRODUCE CANVAS ELEMENT
    - SEPARATE OUT JPEG AND THE PNG
    - USE COMPOSITING MODE IN CANVAS TO ALLOW
    - DRAW JPEG
    - DRAW PNG
    - HOW WOULD YOU EMBED THIS AND MAIN ARCHITECTURAL ELEMNTS - WHAT ELEMENTS USED IN JS AND HTML5 TO CREATE COMPOSITE IMAGE

2. Topic: Audio and Video Coding (28 marks)

a) Describe the advantages and disadvantages of pixel-based and block-based motion representation. (4 marks)

- Blocks allow motion to be estimated between frames however the borders of moving objects rarely coincide with the borders of blocks.
- Pixels allow for a more fine grained search however at a much higher computation cost.
- bb - if block is off more visual and easy to spot, not with pb

b) MPEG-1 video compression uses I-, P- and B- pictures.

i. Explain the advantages of a mixture of picture types. (2 marks)

- The more I frames the MPEG stream has the more editable it is.
- P frames smaller than I frames so reduces the average frame size.
- B frames have the smallest frame size so reduces the average frame size.

ii. Describe a situation where video compression would not be as effective without B-pictures. (6 marks)

Notes
B-frames must be reordered so that “anchor” frames (I & P frames) are available for prediction

Answer
- B-frames allow a ‘summary’ to be made of its surrounding frames. B-frames are different to I-frames where the entire image with all its pixel values are encoded. A much smaller amount of pixels are encoded in B-frames. Thus B-frames have the smallest frame size so reduces the average frame size. This is especially important in motion video compression for example in a video clip or even a movie. If the entire movie were to be compressed using only I-frames and P-frames, its files size would be huge.
- ** B-frames are very important in reducing the file size of compressed video.

c) In telephony, the usable voice frequency band ranges from approximately 300Hz to 3400Hz. When implementing a voice codec

i. Explain why a band-pass filter would be applied to the input audio signal. (2 marks)

- This reduces the amount of data by screening out lower and higher frequencies by decomposing the signal into subbands which adds a further DCT step.

ii. Describe how the discrete cosine transform could be used to compress the speech signal. (5 marks)

- The DCT would be used to filter the speech audio signal to remove unwanted frequencies.

- The frequencies that are kept depend on the application, however in the case of speech, typically from 50Hz to 10kHz is retained. All other frequencies are blocked by the use of a band-pass filter that screens out lower and higher frequencies.

iii. Comment on the effectiveness of this approach. (3 marks)

- At the player high frequencies may reappear because of the staircase form of the signal.
- MUST FINISH WITH MILANS EXAMPLE ANSWER

d) Describe the main steps in MPEG-1 Layer 3 audio coding. (4 marks)

- MP3 coding uses a hybrid filter bank consisting of the polyphase filter bank and a Modified Discrete Cosine Transform (MDCT). The polyphase filter bank has the purpose of making Layer-3 more similar to Layer-1 & Layer-2.
- The subdivision of each polyphase frequency band into subbands increases the potential for redundancy removal.
- An estimate of the actual (time and frequency dependent) masking threshold is computed using rules known from psychoacoustics.
- The spectral components are quantized and coded with the aim of keeping the noise, which is introduced by quantizing, below the masking threshold.
- A bitstream formatter is used to assemble the bitstream, which typically consists of the quantized and coded spectral coefficients and some side information, e.g. bit allocation information.

3. Topic: Media Delivery and Presentation (28 marks)

a) ISO/IEC developed the MPEG-DASH standard allowing for dynamic adaptive streaming over HTTP.

- **Describe the main elements of an MPEG-DASH player. (6 marks)**
    - MPEG DASH streams media using small chunks of media data requested using HTTP and spliced together by the client.
    - DASH presents available content to the media player in a manifest (index) file, the Media Presentation Description (MPD), which uses an XML format.
    - Client uses HTTP to download each media segment as a sequence of files that is played back continuously.
    - DASH does not prescribe any client-specific playback functionality.

- **Describe a typical deployment architecture for MPEG-DASH. (6 marks)**

- In a typical deployment, a DASH server provides segments in several bitrates and resolutions for example, low bitrate and high bitrate, which are encoded in video streams.
- The video streams are then segmented into HTTP resources and a MPD file is generated for the video files, after which a URL is generated for the MPD file.
- The DASH client may then access the MPD file based on its URL and so makes request for appropriate video files.
- The DASH client continuously monitors and adjusts media rate based on network conditions.

b) Explain how Forward Error Correction (FEC) can be used to combat errors in wired and wireless links. Describe the conditions in which the scheme will not be effective. (4 marks)

- FEC can be used to combat errors in wired and wireless links.
- Column FEC (for Bursty Losses) where each column has a repair packet (overhead = 1/D (Depth))
- Row FEC (for Random Losses) where each row has a repair packet (overhead = 1/L (Length))
- 2D (Column x Row) FEC where overhead = (D+L) / (D x L)
- As the row and column FEC's are effectively 1 dimensional, situations in which two packets are lost side by side in the same column or same row or both would result in packet loss.

c) Explain how a decoder can conceal the loss of texture data. (4 marks)

- Spatial Concealment: the decoder can try to estimate missing pixels, blocks or groups of blocks based on the correctly received neighbouring pixels or blocks.
- Temporal Concealment: the decoder can try to estimate missing blocks or groups of blocks based on what was in the specific 'area of error' before or after the error occurred.

d) In a media streaming system

- **Explain why a client playout buffer is a key component. (2 marks)**
    - The client playout buffer is needed to account for variable network delay (jitter). As data does not always arrive at a constant rate, this buffer needs to be able to deal with times too much or too little data is being received in order to ensure smooth playback.

- **In MPEG-DASH the client controls the delivery of data. Describe an approach that an MPEG-DASH player could use to schedule media data from the server to ensure smooth playback. (6 marks)**

- Optimise transmission rates: Here the server can plan a transmission rate so that the media can be viewed without interruption and also minimise the amount of bandwidth used.
- To achieve this, current bandwidth consumption needs to be measured and recent behaviours need to be taken into accounts i.e. how the network has responded to similar operations recently. Used to identify whether to increase or reduce quality. The overall idea is to avoid either too much or too little data.