UNIVERSITY OF DUBLIN

TRINITY COLLEGE

FACULTY OF ENGINEERING, MATHEMATICS & SCIENCE

SCHOOL OF ENGINEERING

Electronic & Electrical Engineering

Senior Sophister
Engineering
Annual Examinations

Trinity Term, 2014

Digital Media Processing (4C8)

Date: 19th May **Venue: Luce Lower** **Time: 14:00 – 16:00**

Dr. David Corrigan

Answer FOUR questions, including
Question ONE and any THREE of the remaining four questions.

All questions carry equal marks

Permitted Materials (Complete List):

Calculator
Drawing Instruments
Mathematical Tables
Graph Paper

SECTION A – COMPULSORY QUESTION

Q.1

(a) Name the 3 primary stages in a lossy image compression encoder. Explain how each stage contributes to improving compression efficiency. In your answer, clearly state which stages are lossy and which are lossless.

[7 marks]

(b) Explain how the JPEG encoder takes advantage of the variation in sensitivity to spatial frequency of the Human Visual System to reduce the file size without a noticeable drop-off in visual quality under standard viewing conditions.

[5 marks]

(c) Encode the following quantised DCT 8×8 block where the dc coefficient of the previously encoded block was 14. Your answer should indicate the symbols that would be provided to the Huffman encoder as well as any other uncoded bits. However, you do not have to apply Huffman encoding to the stream. The table provided in Fig. Q.1 can be used as a reference for your answer.

$$\begin{matrix} 30 & 6 & 1 & 0 & 0 & 0 & 0 & 0 \\ 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{matrix}$$

[13 marks]

DC Coef Difference	Size	Typical Huffman codes for Size	Additional Bits (in binary)
0	0	00	–
$-1, 1$	1	010	$0, 1$
$-3, -2, 2, 3$	2	011	$00, 01, 10, 11$
$-7, \cdots, -4, 4, \cdots, 7$	3	100	$000, \cdots, 011, 100, \cdots, 111$
$-15, \cdots, -8, 8, \cdots, 15$	4	101	$0000, \cdots, 0111, 1000, \cdots, 1111$
$\vdots$	$\vdots$	$\vdots$	
$-1023, \cdots, -512, 512, \cdots, 1023$	10	1111 1110	$00\,0000\,0000, \cdots, 11\,1111\,1111$
$-2047, \cdots, -1024, 1024, \cdots, 2047$	11	1 1111 1110	$000\,0000\,0000, \cdots, 111\,1111\,1111$

Fig. Q.1

SECTION B – ANSWER ANY THREE QUESTIONS

Q.2

(a) The image shown in Fig. Q.2a is filtered with three different filters and the output of each filter is shown in Fig. Q.2b. By examination of the output images, determine what type of filter is responsible for each of the output images. Explain your decisions and give an example of a filter mask that represents that type.

Note: In the top right and bottom images for Fig. Q.2b an intensity of 0 is represented by mid-grey. Lower intensities correspond to negative pixel values and higher values correspond to positive pixel values.

[12 marks]

(b) Explain why the concept of Perfect Reconstruction is key in the use of the Discrete Wavelet Transform (DWT) in image compression.

[3 marks]

(c) Explain how rate distortion curves can be used to compare the performance of different sets of DWT filters in image compression systems. In your answer, also outline how the rate distortion curves can be calculated.

[10 marks]

Fig. Q.2a

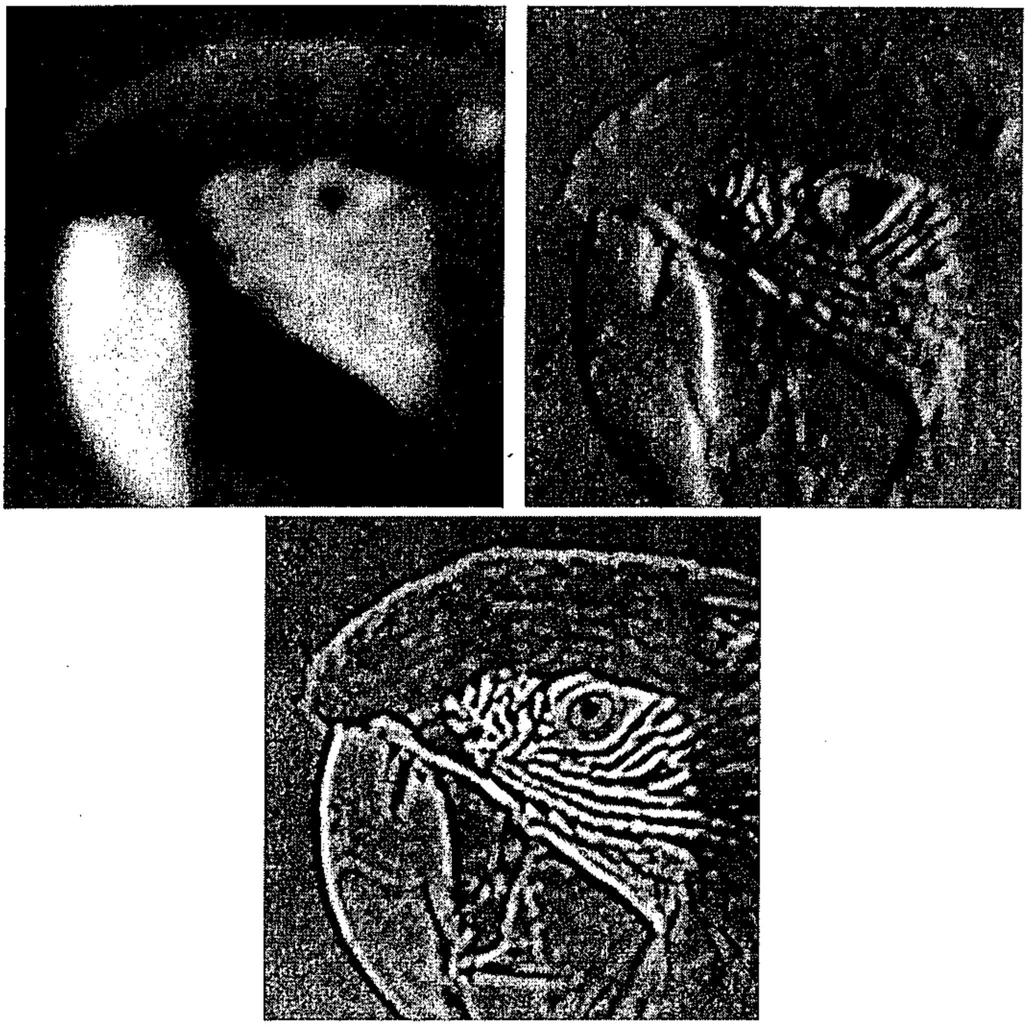

Fig. Q2.b

Q.3

(a) Explain what is meant by the aperture effect in the context of block-based motion estimation. Give examples and explanations of where it occurs and where it does not.

[8 marks]

(b) Describe how the values of the block size and search radius affect the computational efficiency and numerical accuracy of a motion estimator based on block matching. Explain what happens when each parameter is increased and decreased.

[8 marks]

(c) A pixel-based motion detector is used to detect motion on a video sequence where the intensity of the n^{th} frame at pixel $\mathbf{x}$ is $I_n(\mathbf{x})$. The detector applies a thresholding operation to the frame difference $\Delta_n(\mathbf{x}) = I_n(\mathbf{x}) - I_{n-1}(\mathbf{x})$ described by

$$b_n(\mathbf{x}) = \begin{cases} 1 & |\Delta_n(\mathbf{x})| > T_\Delta \\ 0 & \text{otherwise} \end{cases}$$

where a value of $b_n(\mathbf{x}) = \mathbf{1}$ indicates that a motion has been indicated at that pixel.

Assuming that the only source of corruption in the video sequence is additive Gaussian white noise, it is possible to model the distribution of the frame difference in the stationary areas as a Gaussian PDF of variance σ^2.

(i) Show that the probability of a pixel being misclassified as moving is

$$\mathrm{erfc}\left(\frac{T}{2\sigma^2}\right)$$

where

$$\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \exp(-u^2)du$$

[6 marks]

(ii) Describe how this motion detector can be adapted to provide a block-level rather than pixel-level indicator of motion for the block matching algorithm.

[3 marks]

Q.4

(a) Explain what is meant by the terms GOP and I-, P- and B-frames in the context of the MPEG-2 standard. What is a typical GOP structure of an MPEG-2 stream?

Describe the order in which the frames of a GOP are encoded and explain why that is the case.

[12 marks]

(b) Draw a block diagram that explains the operation of how P-frames are encoded in an encoder based on Motion Compensated Prediction Coding (MCPC). The diagram does not have to indicate how motion vectors or any other metadata associated with the frame is multiplexed into the compressed stream.

Give an overview of the structure of the encoder and explain clearly how feedback is used in the encoder. Why is this structure preferred over more simple architectures?

[13 marks]

Q.5

(a) What are the 4 subbands of the Haar Transform? Describe how each of the subbands can be implemented using separable filters and downsampling operations.

[10 marks]

(b) Outline an algorithm in pseudocode that implements the single-level forward Haar Transform and applies a uniform quantisation step size to the subbands. You must implement the transform using matrix multiplications rather than the filtering and downsampling operations referred to in part **(a)**.

Note: You can assume that the typical arithmetic, array handling and image handling operations available in Matlab are available to use and do not have to be defined.

[10 marks]

(c) Explain why applying a multi-level Haar Transform results in further reduction in entropy compared to the single-level transform. Why does the entropy level off after the number of levels increases beyond a given value?

[5 marks]