ENGINEERING TRIPOS PART IIB

Monday 27 April 2009 9 to 10.30

Module 4F8

IMAGE PROCESSING AND IMAGE CODING

Answer not more than ***three*** *questions.*

All questions carry the same number of marks.

The ***approximate*** *percentage of marks allocated to each part of a question is indicated in the right margin.*

There are no attachments.

STATIONERY REQUIREMENTS
Single-sided script paper

SPECIAL REQUIREMENTS
Engineering Data Book
CUED approved calculator allowed

You may not start to read the questions printed on the subsequent pages of this question paper until instructed that you may do so by the Invigilator

1 (a) Describe the *windowing method* used to create a finite-support two-dimensional (2d) filter from an ideal zero-phase 2d frequency response, outlining two methods of creating 2d window functions from 1d window functions. [10%]

(b) Find the spectrum of the 2d window function $w(u_1,u_2)$ formed from the product of rectangular window functions $w_1(u_1)$ and $w_2(u_2)$, where, for $i=1,2$

$$w_i(u_i) = \begin{cases} 1 & \text{if } |u_i| < U_i \\ 0 & \text{otherwise} \end{cases}$$

and sketch the spectrum $W(\omega_1,\omega_2) = W_1(\omega_1)W_2(\omega_2)$. [25%]

(c) Now find the spectrum of the 2d window function $w(u_1,u_2)$ formed from the product of cosine window functions $w_1(u_1)$ and $w_2(u_2)$, where, for $i=1,2$

$$w_i(u_i) = \begin{cases} \cos\left(\frac{\pi u_i}{U_i}\right) & \text{if } |u_i| < U_i \\ 0 & \text{otherwise} \end{cases}$$

Sketch the spectrum along the ω_1 axis and comment on why a simple cosine window may be problematic. [25%]

(d) To produce a better main lobe as well as reducing sidelobes, we may take a combination of the above windows, for example:

$$w_i(u_i) = \begin{cases} \alpha + \beta\cos\left(\frac{\pi u_i}{U_i}\right) & \text{if } |u_i| < U_i \\ 0 & \text{otherwise} \end{cases}$$

where α and β are constants. Using the previous results or otherwise, determine the values of α and β such that the following conditions are satisfied

(i) $w_i(0) = 1$ for $i = 1,2$

(ii) $W_i\left(\frac{5\pi}{2U_i}\right) = 0$ for $i = 1,2$

[30%]

(e) Comment on the values found for α and β, compared to those used in the Hamming window. Comment also on the desirability of enforcing the conditions given in Part (d). [10%]

2 (a) An image $g(u_1,u_2)$ is sampled on a rectangular grid (spacings Δ_1 and Δ_2 in the u_1 and u_2 directions respectively). The sampled image $g_s(u_1,u_2)$ may be written as

$$g_s(u_1,u_2) = s(u_1,u_2)g(u_1,u_2)$$

(i) Write down expressions for $s(u_1,u_2)$ in terms of delta functions and for $G_s(\omega_1,\omega_2)$, the Fourier transform of g_s, in terms of the sample spacings and $G(\omega_1,\omega_2)$, the Fourier transform of g. [15%]

(ii) If $g(u_1,u_2)$ is a 2-dimensional sinewave of the following form

$$g(u_1,u_2) = \sin(\Omega u_1)\ \sin(5\Omega u_2)$$

find the minimum sampling frequencies (in the u_1 and u_2 directions) which should be used to avoid aliasing. [15%]

(iii) If the sample spacings used are $\Delta_1 = \frac{3\pi}{2\Omega}$ and $\Delta_2 = \frac{\pi}{6\Omega}$, sketch the spectrum of the sampled signal. [15%]

(b) In a Bayesian derivation of the Wiener filter we maximise the probability of $\mathbf{x}$ given $\mathbf{y}$, $P(\mathbf{x}|\mathbf{y})$, where our true and observed images in vector form are $\mathbf{x}$ and $\mathbf{y}$ respectively.

(i) Assuming that we can write $\mathbf{y}$ in terms of a linear distortion of $\mathbf{x}$ plus noise, $\mathbf{y} = L\mathbf{x} + \mathbf{d}$, give an expression for the *likelihood* $P(\mathbf{y}|\mathbf{x})$, assuming the noise is Gaussian with covariance matrix N. [15%]

(ii) Assuming $\mathbf{x}$ is also a Gaussian random variable, described by a known covariance matrix C, write down the prior, $P(\mathbf{x})$, and hence obtain an expression for $P(\mathbf{x}|\mathbf{y})$. [15%]

(iii) In a conventional derivation of the Wiener filter the optimal filter is given by

$$G(\boldsymbol{\omega}) = \frac{H^*(\boldsymbol{\omega})\,P_{xx}(\boldsymbol{\omega})}{|H(\boldsymbol{\omega})|^2\,P_{xx}(\boldsymbol{\omega}) + P_{dd}(\boldsymbol{\omega})}$$

Explain the meaning of each of the terms in this expression for $G(\boldsymbol{\omega})$ and relate them to the matrices L, N and C. [25%]

(TURN OVER

3 (a) A sampled signal y_n, $n \in \mathbb{Z}$, is down-sampled by 2 and then up-sampled by 2 so that the resulting signal $\widehat{y}_n$ is

$$\widehat{y}_n = y_n \quad \text{if } n \text{ is even,} \qquad \text{and} \qquad \widehat{y}_n = 0 \quad \text{if } n \text{ is odd.}$$

Show that in the z-domain the transforms are related by

$$\widehat{Y}(z) = \frac{1}{2}[Y(z) + Y(-z)]$$

[20%]

(b) The analysis and reconstruction parts of a 2-band filter-bank system are shown in Fig 1(a) and Fig 1(b). Obtain an expression for $\widehat{X}(z)$ in terms of $X(z)$, $X(-z)$ and the filter z-transfer functions, H_0, H_1, G_0 and G_1. Hence derive the anti-aliasing condition and the perfect-reconstruction condition on the filter transfer functions. Why are these conditions important in an image compression system based on wavelets? [25%]

(c) Show that if $H_1(z) = zG_0(-z)$ and $G_1(z) = z^{-1}H_0(-z)$, and if the product filter, given by $P(z) = H_0(z)G_0(z)$, satisfies

$$P(z) + P(-z) = 2$$

then the anti-aliasing and perfect-reconstruction conditions are satisfied. [20%]

(d) In a given filterbank, $G_0(z) = \frac{1}{2}(z + 2 + z^{-1})$. If $H_0(z)$ is of the form

$$H_0(z) = az^2 + bz + c + bz^{-1} + az^{-2}$$

and if we require that $H_0(-1) = 0$, calculate the coefficients a, b and c in order to get a perfect reconstruction filterbank. Hence obtain expressions for $H_1(z)$ and $G_1(z)$. [15%]

(e) Calculate the locations of the zeros of each filter in the 2-band filterbank, and suggest why these locations help to give good compression performance in a wavelet-based image compression system. [20%]

(cont.

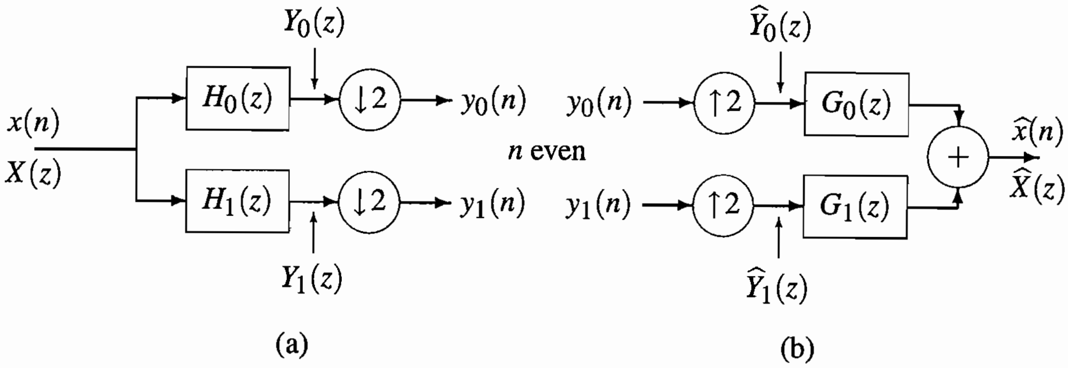

Fig. 1

(TURN OVER

4 (a) Explain why it is helpful to convert colour images from RGB to YUV format prior to performing compression on them. [15%]

(b) The forward and inverse conversion matrices between $[\mathrm{R} \;\; \mathrm{G} \;\; \mathrm{B}]^T$ and $[\mathrm{Y} \;\; \mathrm{U} \;\; \mathrm{V}]^T$ vectors are given by

$$C = \begin{pmatrix} 0.3 & 0.6 & 0.1 \\ -0.15 & -0.3 & 0.45 \\ 0.4375 & -0.3750 & -0.0625 \end{pmatrix} \quad \text{and} \quad C^{-1} = \begin{pmatrix} 1 & 0 & 1.6 \\ 1 & -0.3333 & -0.8 \\ 1 & 2 & 0 \end{pmatrix}$$

Explain why the top row of C is $[0.3 \;\; 0.6 \;\; 0.1]$ and why the left column of C^{-1} is $[1 \;\; 1 \;\; 1]^T$. [25%]

(c) JPEG compression is applied to a YUV colour image of size 768×1024 pixels, by first subsampling the U and V components by 2:1 in each direction (by taking the mean of each 2×2 block of U or V pixels) and then using the standard 8×8 DCT-based compression algorithm on each Y, U and V subimage. If, for a given quantisation step size, the mean entropy of each 8×8 block of Y pixels is 1.3 bit/pixel and that of each 8×8 block of U or V pixels is 0.6 bit/pixel, estimate the total number of bits that would be needed to encode this image, and also the proportion of bits needed to encode the chrominance (colour) content of the image. [30%]

(d) The JPEG encoding standard defines a 2-dimensional Huffman code for representing the AC coefficients in each 8×8 block. Describe how the information for each non-zero coefficient is split into *Run*, *Size* and *Additional Bits*, and how the 2-dimensional Huffman code is applied to this data. In particular, explain how such a code is in theory able to produce an encoded bit rate that is lower than predicted from the basic entropy of the quantised coefficient values. [30%]

END OF PAPER