ENGINEERING TRIPOS   PART IIB

Monday 23 April 2007   2.30 to 4

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) An image $g(u_1,u_2)$ is sampled using the sampling grid shown in Fig. 1. The sampled image is denoted as $g_s(u_1,u_2)$ and can be written as

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

(i) By viewing the sampling grid as the sum of two rectangular sampling grids, or otherwise, write down a mathematical expression for the sampling function $s(u_1,u_2)$. [15%]

(ii) Find the Fourier coefficients of the periodic sampling function $s(u_1,u_2)$. [20%]

(iii) Hence, or otherwise, derive the Fourier transform, $G_s(\omega_1,\omega_2)$, of the sampled image $g_s(u_1,u_2)$ and sketch this in terms of the modulus of the Fourier transform of $g(u_1,u_2)$. (You can assume any bandlimited form for this). [25%]

(b) A particular form of ideal frequency response, $H(\omega_1,\omega_2)$, is shown in Fig. 2. H takes the value 1 in the shaded regions and zero outside these regions.

(i) Write down a mathematical expression for $H(\omega_1,\omega_2)$. [15%]

(ii) By inverse Fourier transforming $H(\omega_1,\omega_2)$, or otherwise (e.g. quoting standard results for rectangular bandpass filters), find the ideal impulse response, $h(u_1,u_2)$. [25%]

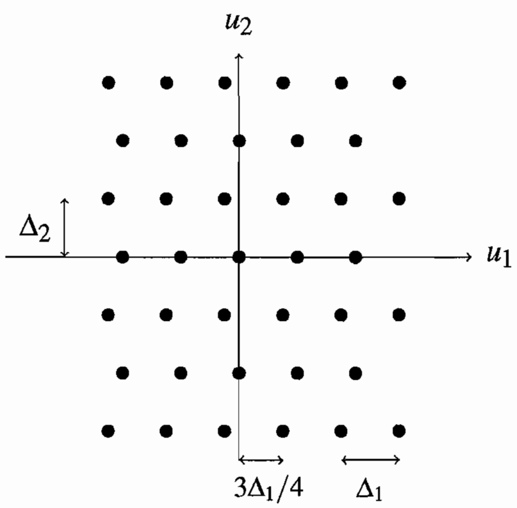

Fig. 1

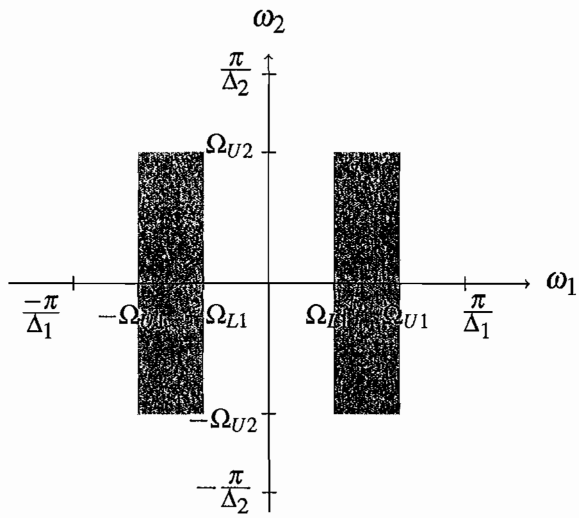

Fig. 2

2 (a) The *windowing method* is used to create finite-support filters from the inverse Fourier transform of some ideal zero-phase 2D frequency response.

(i) The resulting frequency response will be the Fourier transform of the windowed ideal impulse response. Describe the effects the windowing has on the ideal frequency response. [10%]

(ii) Outline the *product* and *rotation* methods for forming 2D window functions from 1D window functions. [15%]

(iii) Consider the following 1D window functions (for $i = 1,2$)

$$w_i(u_i) = \begin{cases} 0.42 + 0.5\cos(\pi \frac{u_i}{U_i}) + 0.08\cos(2\pi \frac{u_i}{U_i}) & \text{if } |u_i| < U_i \\ 0 & \text{otherwise} \end{cases}$$

Find the spectrum of the 2D window function $w(u_1, u_2)$ formed from the product of w_1 and w_2. [35%]

(iv) By inspection of the form of this spectrum, describe how this window function differs in terms of its mainlobe and sidelobe behaviour, compared to the *Hamming window*. [10%]

(b) Assume that an observed image $y(\mathbf{n})$ can be modelled as the convolution of the true image $x(\mathbf{n})$ and a point-spread function $h(\mathbf{n})$, plus additive noise $d(\mathbf{n})$, i.e.

$$y(\mathbf{n}) = \sum_{\mathbf{m} \in \mathbf{Z}^2} h(\mathbf{m}) x(\mathbf{n} - \mathbf{m}) + d(\mathbf{n})$$

where $\mathbf{n} = (n_1, n_2)$, $\mathbf{m} = (m_1, m_2)$ and $\mathbf{Z}$ denotes the set of integers.

(i) We can obtain an estimate, $\hat{x}(\mathbf{n})$, of the true image via convolving the observations with a linear filter $g(\mathbf{q})$, i.e. $\hat{x}(\mathbf{n}) = \sum_{\mathbf{q} \in \mathbf{Z}^2} g(\mathbf{q}) y(\mathbf{n} - \mathbf{q})$. Describe the cost function we minimise in order to produce an estimate of $g(\mathbf{q})$ which corresponds to the *Wiener filter*. [10%]

(ii) In a Bayesian derivation of the Wiener filter explain what *prior*, $Pr(\mathbf{x})$, is used for the true image $\mathbf{x}$. [10%]

(iii) Outline how one might improve on the performance of the Wiener filter by taking alternative priors for $\mathbf{x}$ and give an example of one such alternative prior. [10%]

3 (a) What are the key characteristics of human vision that are exploited in current image compression systems? [20%]

(b) The basic Haar transform matrix, **T**, and a 2×2 block of image pixels, **X**, are given by

$$\mathbf{T} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad \mathbf{X} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Derive expressions for the coefficients of **Y**, the 2D Haar transform of **X**, in terms of the pixels a, b, c, d. [20%]

(c) If the four coefficient outputs from each block are reordered into four subbands, describe the typical image characteristics selected by each subband and explain why this process is helpful for image compression. [20%]

(d) Explain how a Haar transform may be extended to provide multiple levels of transformation. Discuss the typical type of visual artifacts which appear in such a multilevel system, as a result of severe quantisation of the subband coefficients. [20%]

(e) Discuss briefly how wavelet concepts allow these visual artifacts to be modified so as to reduce their visibility. [20%]

4 (a) The elements in row k of the transform matrix $\mathbf{T}$ for a 1-dimensional N-point discrete cosine transform (DCT) are of the form:

$$t_{ki} = C_k \cos\left(\frac{\pi(2i-1)(k-1)}{2N}\right) \quad \text{for} \quad i,k = 1,2\ldots,N$$

where the C_k are constants. For $N = 4$, show that the rows of $\mathbf{T}$ are orthogonal and calculate the values of C_k which make the transform orthonormal. [30%]

(b) Explain how $\mathbf{T}$ can be used to transform a block of image pixels, represented by the 4×4 matrix $\mathbf{X}$, into a 4×4 matrix of coefficients $\mathbf{Y}$, representing the 2-dimensional DCT of $\mathbf{X}$. Then show how $\mathbf{T}$ may also be used to recover $\mathbf{X}$ from $\mathbf{Y}$, giving reasons why it is important for $\mathbf{T}$ to be an orthonormal matrix. [20%]

(c) Such a DCT is applied to all the 4×4 blocks of pixels in a 1600×1200 pixel monochrome image; and the elements of the $\mathbf{Y}$ matrices are reordered into 16 subbands $\mathbf{Y}_{p,q}$ of size 400×300, each comprising the element $y_{p,q}$ from every 4×4 block $\mathbf{Y}$. For a given image and quantisation strategy, the entropies of the coefficients in the 16 subbands are given by

$$H_{p,q} = \log_2\left(\frac{64}{(p+q)^2}\right) \quad \text{for} \quad p,q = 1,2,3,4$$

Calculate the approximate number of bits needed to code this image. [30%]

(d) If the image were instead a colour image, estimate the increased number of bits needed, compared with that for the monochrome version, stating your assumptions clearly. [20%]

END OF PAPER