EGT3
ENGINEERING TRIPOS PART IIB

Tuesday 2 May 2023 2.00 to 3.40

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.*

Write your candidate number ***not*** *your name on the cover sheet.*

STATIONERY REQUIREMENTS
Single-sided paper.

SPECIAL REQUIREMENTS TO BE SUPPLIED FOR THIS EXAM
CUED approved calculator allowed.
Engineering Data Book

10 minutes reading time is allowed for this paper at the start of the exam.

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

You may not remove any stationery from the Examination Room.

1 (a) In order to create a *finite support* filter from the inverse Fourier transform of an *ideal zero-phase* frequency response, one of the most common approaches is to use the *windowing method*.

(i) Describe two methods for forming 2D window functions from 1D window functions and explain how the properties of the window function cause the ideal frequency response to be modified. [15%]

(ii) Consider the following 1D window functions $w_i(u_i)$ for $i = 1, 2$

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

Sketch the 2D window function, $\mathrm{W}(u_1, u_2) = w_1(u_1)w_2(u_2)$. [10%]

(iii) The Fourier transform of $w_i(u_i)$ is $\mathcal{W}_i(\omega_i)$, where

$$\mathcal{W}_i(\omega_i) = \frac{U_i\pi^2 \operatorname{sinc} U_i\omega_i}{(\pi^2 - U_i^2\omega_i^2)}$$

Sketch $\mathcal{W}$. By looking at the width of the mainlobe and the approximate relative amplitudes of the first two sidelobes, comment on the merits of this windowing function. [20%]

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

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

(i) Using this model, explain how the methods of *inverse filtering* and *generalised inverse filtering* can be used for deconvolution when the psf is known. How do these methods behave in the presence of significant noise? [15%]

(ii) The *Wiener filter*, $G(\omega)$, takes account of the noise when attempting deconvolution. The form of G is given by

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

Explain the nature of all quantities in the above equation and describe how one might estimate P_{xx} and P_{dd} if they are unknown. [15%]

(iii) We can write our model in matrix terms as $\mathbf{y} = L\mathbf{x} + \mathbf{d}$. Write down the matrix form of the Wiener filter, W, describing all quantities used (W is such that our estimate of $\mathbf{x}$ is given by $W\mathbf{y}$). [20%]

(iv) Explain how, in your expression for W in Part (b)(iii), the term derived from the Gaussian prior, acts as a *regularizer*. [5%]

2 (a) Consider a continuous image, $g(u_1, u_2)$, which is sampled on a rectangular grid (with spacings Δ_1 and Δ_2). The sampled signal is $g_s(u_1, u_2)$.

(i) Write down the Fourier transform, $G_s(\omega_1, \omega_2)$, of g_s, in terms of $G(\omega_1, \omega_2)$, the Fourier transform of g. Use this result to explain the phenomenon of *aliasing*. [15%]

(ii) The signal $g(u_1, u_2)$ is now sampled on the regular hexagonal grid, shown in Fig.1, to produce a sampled signal, $\tilde{g}_s(u_1, u_2)$. Each hexagon has sides of length d. By writing the grid as the sum of two or more shifted rectangular grids and using the result in Part (a)(i) for a rectangular grid with grid-points at $n_1\Delta_1$ in the u_1 direction and $n_2\Delta_2$ in the u_2 direction (n_1 and n_2 integers), obtain the spectrum $\tilde{G}_s(\omega_1, \omega_2)$ of the sampled signal in terms of the spectrum $G(\omega_1, \omega_2)$ of the original signal and d. [30%]

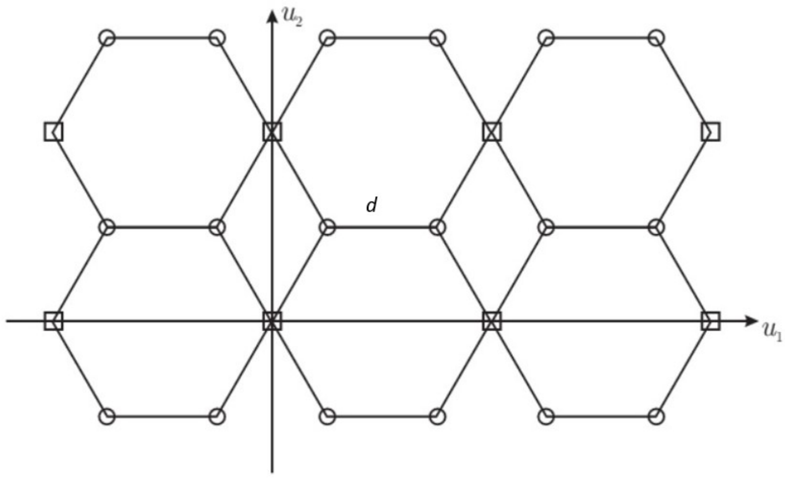

Fig. 1

(b) The 2D Fourier transform of an image produces a 2D array of complex values which can be used to understand the image structure.

(i) Discuss the relative importance of amplitude and phase in the Fourier transform of images. [10%]

(ii) The Fourier transform of $g(u) = [2(H(u) - 0.5)]$ is $\frac{2}{j\omega}$, where H is the Heaviside step function. Sketch the image given by

$$g(u_1, u_2) = [2(H(u_1) - 0.5)]\ [2(H(u_2) - 0.5)] \quad -\infty < u_1, u_2 < +\infty$$

and give its Fourier transform. [15%]

(iii) Now consider the image, $g_r(u_1, u_2)$, shown in Fig. 2, where the shaded regions are $+1$ and the non-shaded regions are -1. The regions are delineated by the lines $u_1 = u_2$ and $u_1 = -u_2$. Using the results in Part (b)(ii), find the Fourier transform, $G_r(\omega_1, \omega_2)$, of g_r. [20%]

(iv) By looking at the phase of G_r, verify that this process has picked out the edges in the image. [10%]

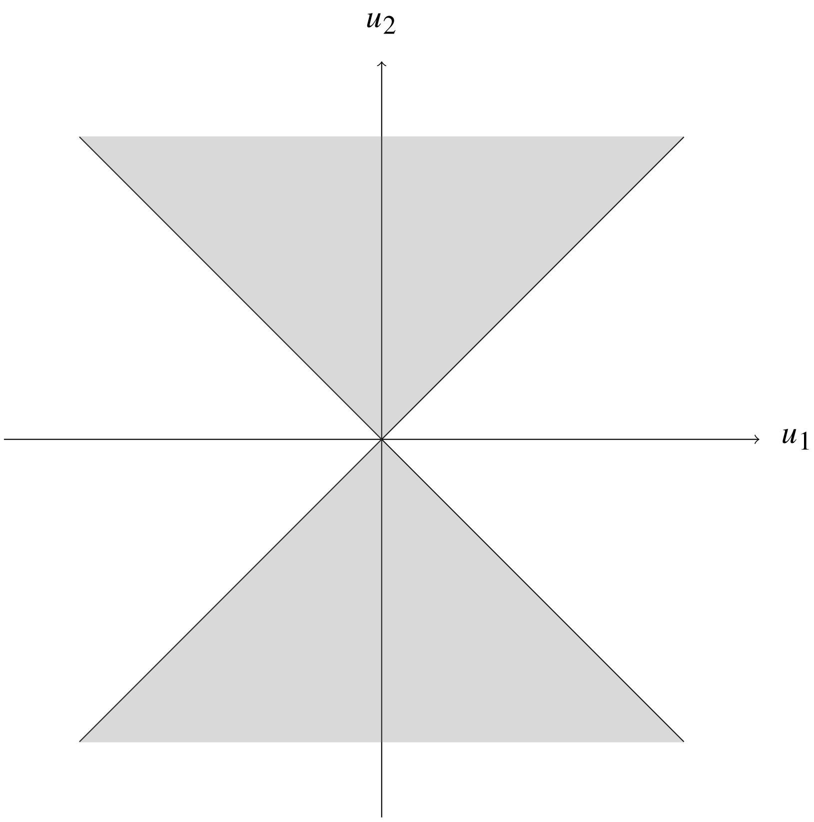

Fig. 2

3 (a) Explain how one would apply the Haar matrix T to an image X (i.e. the 2D Haar transform) and what the elements of the resulting image Y represent in terms of the frequency content of the image, where

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

[15%]

(b) Explain why the Haar transform is an energy-preserving transform and describe how a 2-level 2D transform is executed on an image. Now apply the level-2 2D Haar transform to a 4×4 matrix Z, whose elements are z_{ij}. Write down the Hi-Lo element (top right of the matrix) of the level 2 decomposition in terms of the $\{z_{ij}\}$. [25%]

(c) For the level-1 2D Haar transform, the majority of the energy (for most natural images) is in the Lo-Lo subimage. Define the *entropy* of a quantised image and explain why this is often a more useful measure for *image compression* purposes. Sketch the typical distribution of entropy for each of the Lo-Lo, Hi-Lo, Lo-Hi, Hi-Hi subimages when the level-1 2D Haar transform is applied to a typical natural image. [20%]

(d) The $n\times n$ *Discrete Cosine Transform* (DCT) has better image compression properties than the Haar transform (for $n > 2$). For an $n \times n$ DCT, T, the elements are given by

$$\begin{aligned} t_{1i} &= \sqrt{\frac{1}{n}} && \text{for } i = 1, .., n \\ t_{ki} &= \sqrt{\frac{2}{n}} \cos\left(\frac{\pi(2i-1)(k-1)}{2n}\right) && \text{for } i = 1, .., n, \quad k = 2, .., n \end{aligned}$$

For $n = 8$ sketch the rows of T and outline how the symmetries present in T enable a fast algorithm for the multiplication $T\mathbf{x}$ where $\mathbf{x}$ is a 8-element column vector. [25%]

(e) Discuss the relative merits of using 4×4, 8×8 and 16×16 DCTs for image compression. Which of these are used in the most common compression standards? [15%]

4 (a) (i) Draw the basic 2-band analysis filter bank with down-samplers, upon which wavelet transforms are based. Also draw the corresponding 2-band reconstruction filter bank with up-samplers. What is the perfect reconstruction requirement and why is it important? [Note: *downsampling* is setting alternate samples to zero, and *upsampling* is introducing zeros for the missing samples]. [20%]

(ii) Show that downsampling by a factor of 2 followed by upsampling by a factor of 2 converts a signal with z-transform $Y(z)$ into a signal with z-transform $[Y(z)+Y(-z)]/2$. [15%]

(iii) Using the result in Part (a)(ii), prove that the operation sequence of *downsample-filter-upsample*, with filter given by $H(z)$, is equivalent to the operation sequences *filter-downsample-upsample* or *downsample-upsample-filter*, with filter given by $H(z^2)$. Explain how this equivalence can be useful for analysis. [25%]

(b) RGB colour space is generally converted to YUV colour space for image compression purposes.

(i) Sketch typical histograms of the Y,U and V components for a standard natural image. [10%]

(ii) Using Part (b)(i) explain why the U and V (chrominance) components can be sampled at a lower rate than the sampling applied to the Y (luminance) component. [5%]

(iii) Appealing to the *visual sensitivity* of the human eye, explain why the U and V components can be quantised more coarsely than the Y component. [10%]

(iv) JPEG compression is applied to a YUV colour image of size 768×1024. The U and V components are subsampled by 2:1 in each direction after which the standard 8×8 DCT-based algorithm is applied to Y and the subsampled U,V images. For a given quantisation step size, the mean entropy of each 8×8 block of Y pixels is 1.2 bits/pixel and that of each 8×8 block of U and V pixels is 0.5 bits/pixel. Estimate the total number of bits that would be needed to encode this image. [15%]

END OF PAPER

THIS PAGE IS BLANK