EGT3
ENGINEERING TRIPOS PART IIB

Tuesday 23 April 2019 9.30 to 11.10

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 script 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.

1 (a) One means of creating a finite support filter from the inverse Fourier transform of an ideal zero-phase frequency response is to use the *windowing method*.

(i) Describe the **product** and **rotation** methods for forming 2D window functions from 1D window functions and explain the effects windowing has on the ideal frequency response. Hence outline desirable properties of a window function. [20%]

(ii) Consider the following 1D window function:

$$w(u) = \begin{cases} 1 - \frac{|u|}{U} & \text{if } |u| < U \\ 0 & \text{otherwise} \end{cases}$$

Now consider the 2D window function, $W(r)$, formed via the *rotation method*, where $r = \sqrt{u_1^2 + u_2^2}$. Sketch $W(r)$. [10%]

(iii) By direct integration, show that the spectrum, $\mathscr{W}(\omega_1, \omega_2)$, takes the form

$$\mathscr{W}(\omega_1, \omega_2) = 2\pi \int_0^U r\left(1 - \frac{r}{U}\right) J_0(r f(\omega_1, \omega_2)) dr$$

and find the form of the function $f(\omega_1, \omega_2)$. J_0 is a 0^{th} order Bessel function defined as $J_0(x) = \frac{1}{\pi} \int_0^{\pi} \cos\{x \sin(\theta)\} d\theta$. [20%]

(iv) Fig.1 plots the two terms of the integral given in part (a)(iii) (taking $U = 1$) as a function of $\omega = \sqrt{\omega_1^2 + \omega_2^2}$ [the first term, $\int r J_0 \, dr$, is the solid line and the second term, $\int r^2 J_0 \, dr$, is the dashed line]. By considering the difference of these two plots, discuss the merits of the circularly symmetric window function W. [15%]

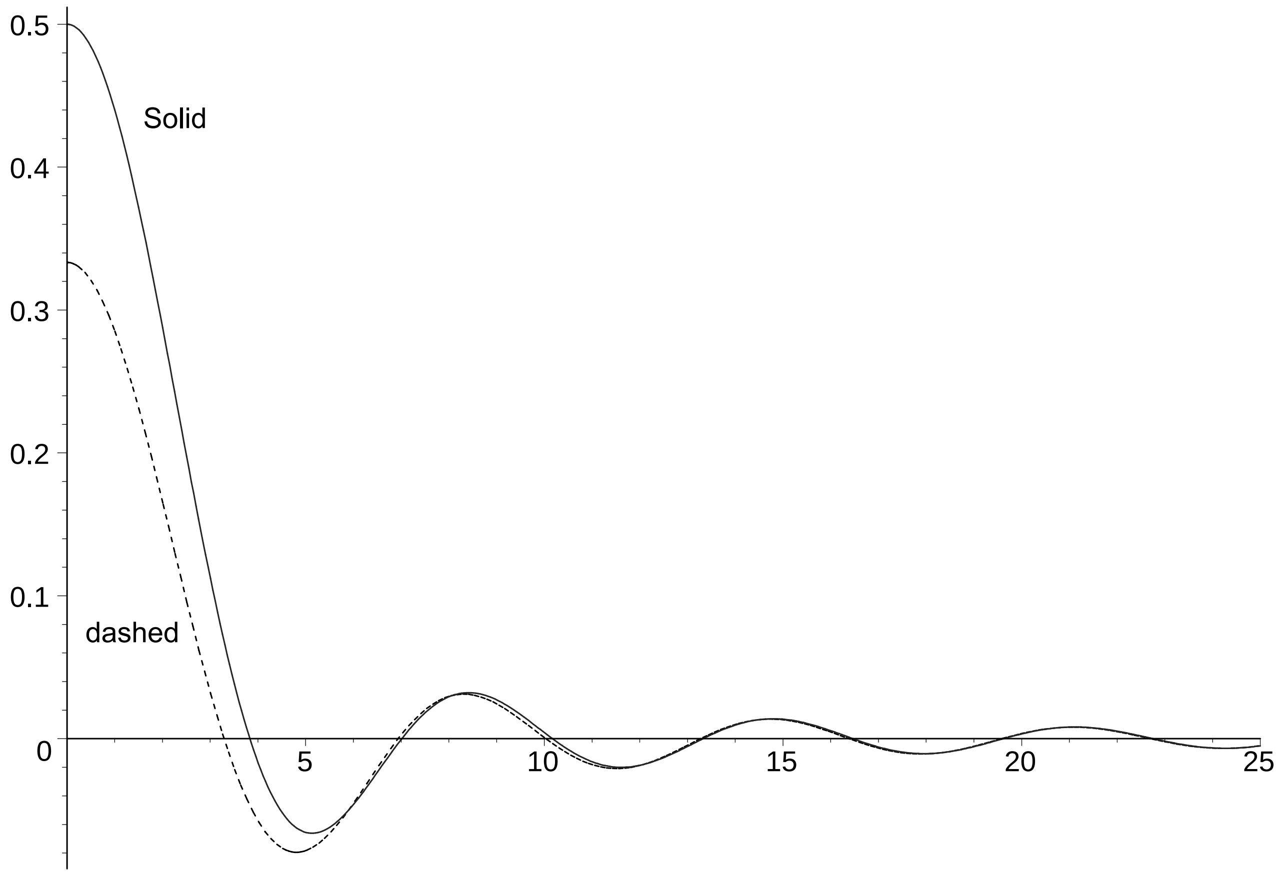

Fig. 1

(b) Consider the ideal bandpass filter shown in Fig 2, with $H(\omega_1, \omega_2) = 1$ in the shaded regions and $H(\omega_1, \omega_2) = 0$ otherwise. Sampling is carried out on a rectangular grid with spacings of Δ_1 and Δ_2 in the u_1 and u_2 directions respectively.

(i) Using standard results or otherwise, show that the ideal impulse response, $h(n_1\Delta_1, n_2\Delta_2)$, of this filter takes the following form

$$h(n_1\Delta_1, n_2\Delta_2) = \Omega^2 \frac{\Delta_1\Delta_2}{\pi^2} \sum_{k=1}^{3} \gamma_k \,\mathrm{sinc}(\alpha_k \Omega n_1 \Delta_1)\,\mathrm{sinc}(\beta_k \Omega n_2 \Delta_2)$$

and find the values of $\alpha_k,\ \beta_k,\ \gamma_k,\ k = 1, 2, 3.$ [30%]

(ii) Describe the effects of a filter with this frequency response acting on an image. [5%]

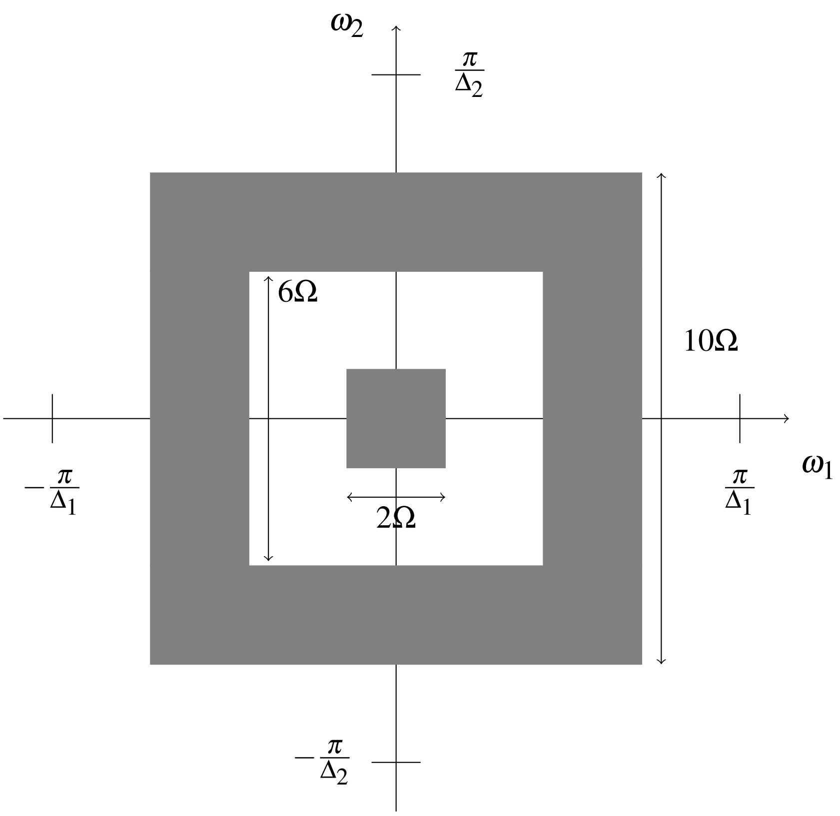

Fig. 2

2 (a) Assume that an observed image **y** can be modelled as a linear distortion of the true image **x** plus additive noise **d**.

(i) Assuming that we can write **y** in terms of a linear distortion of **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. [10%]

(ii) Assuming **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 $P(\mathbf{x}|\mathbf{y})$. Explain, qualitatively, how the *Wiener Filter* result is obtained from $P(\mathbf{x}|\mathbf{y})$. [20%]

(b) A method of image deconvolution and denoising which emphasises sparsity is *compressed sensing*, in which the Gaussian prior on the unknown true pixel values **x** is replaced by an l_1 norm. As a way of comparing the *Wiener filter* with a *compressed sensing* approach, we consider a pure image denoising problem. Assume there is no linear distortion and the noise is Gaussian, and that the observed image, **y**, has a noise covariance matrix which is diagonal with constant entries σ^2. Assume that all the true pixel values **x** are positive or 0.

(i) Let the Gaussian prior on the true image, **x**, in the Wiener case be given by:

$$P(\mathbf{x}) \propto \exp\left(-\frac{1}{2}\lambda\sum_i x_i^2\right)$$

where x_i is the i^{th} element of **x** and λ is real and positive. By maximising $P(\mathbf{x}|\mathbf{y})$ with respect to **x**, obtain the *Wiener* solution for the best estimate of the denoised image,$\{\hat{x}_i\}$. [25%]

(ii) Now, suppose that the l_1 norm prior on the true image, **x**, in the *compressed sensing* case, is given by:

$$P(\mathbf{x}) \propto \exp\left(-\mu\sum_i |x_i|\right)$$

where the x_i are again the elements of **x** and μ is real and positive. As in part (b)(i), maximise $P(\mathbf{x}|\mathbf{y})$ with respect to **x** (via direct differentiation) to show that the solution is now:

$$\hat{x}_i = \begin{cases} y_i - \mu\sigma^2 & \text{if } y_i > \mu\sigma^2 \\ 0 & \text{otherwise} \end{cases}$$

[35%]

(iii) Comment on whether the compressed sensing prior has indeed increased the chances of *sparsity* in the result. [10%]

3 (a) Image compression systems use the characteristics of the *human visual system* (HVS) in their design.

(i) *RGB* (Red-Green-Blue) is the standard system for colour images. Describe the alternative YUV (luminance-chrominance) system, outlining how *RGB* can be transformed to obtain YUV. [15%]

(ii) Using a sketch of the *contrast sensitivity of the eye* as a function of spatial frequency for YUV components, explain why YUV is a good system for image compression. [15%]

(b) Sketch a basic *image coding* system, labelling the blocks of both the *encoder* and *decoder*. [15%]

(c) The **Haar transform** provides a simple choice for the compression block of the image coding system.

(i) Write down the 2×2 Haar matrix, T, and show that it is an *orthonormal* matrix, and thus preserves energy. [15%]

(ii) A 4×4 image, X, is given below. Apply the Haar transform from part (c)(i) to 2×2 blocks of X, to produce transformed blocks Y_i, $i = 1, 2, 3, 4$.

$$X = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 3 & 0 & 1 & 2 \\ 1 & 2 & 3 & 0 \end{bmatrix}$$

[15%]

(iii) Describe the process of taking elements from the blocks Y_i of part (c)(ii) to form a 4×4 image Y, comprising 4 subimages. Give the characteristics of each of these subimages. [15%]

(iv) Explain how we achieve data compression via a multi-level Haar transform. For a typical image, after how many levels of the Haar transform do we expect no significant increase in compression performance? [10%]

4 (a) The standard baseline version of JPEG uses an 8-point DCT (Discrete Cosine Transform), which has better compression characteristics than the Haar transform. The components of an $n \times n$ DCT matrix, T, are given by:

$$
\begin{aligned}
t_{1i} &= \sqrt{\frac{1}{n}} \quad \text{for } i = 1,2,\ldots,n \\
t_{ki} &= \sqrt{\frac{2}{n}} \cos\left(\frac{\pi(i-\frac{1}{2})(k-1)}{n}\right) \quad \text{for } i = 1,2,\ldots,n,\ \ k = 2,3,\ldots,n
\end{aligned}
$$

(i) By taking the dot product of row vectors of T, or otherwise, show that the matrix T is orthonormal (assume n is even). [25%]

(ii) Explain why each row of the DCT matrix, T, can be viewed as a *basis function* when T acts on vectors (i.e. the 1D DCT). Sketch the form of these 1D basis functions. [20%]

(iii) Show how an 8×8 DCT matrix, T, is used to transform an 8×8 image block, X, to an 8×8 matrix of DCT coefficients, Y. What are the *basis functions* of this 2D DCT? [20%]

(b) JPEG compression is applied to a YUV colour image of size 1024×2048 pixels.

(i) If, for a given quantisation step size, the mean entropy of each 8×8 block of Y pixels is 1.2 bit/pixel and that of each 8×8 block of U or V pixels is 0.5 bit/pixel, estimate the total number of bits that would be needed to encode this image. You should apply a sensible initial subsampling of the U and V images. What proportion of the total bits is needed to encode the chrominance (colour) content of the image? [20%]

(ii) Describe (qualitatively) the quantisation and coding stages of the JPEG algorithm that follow the DCT compression stage. [15%]

END OF PAPER

THIS PAGE IS BLANK