Version NGK/5

EGT3
ENGINEERING TRIPOS PART IIB

Friday 1 May 2015   2 to 3.30

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.*

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.

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

1 (a) Consider the image, $g(u_1,u_2)$, shown in Fig. 1, consisting of 5 vertical stripes of value A (shown dark) separated by 4 equal width vertical stripes of value 0. Each stripe has horizontal width a and vertical height $2b$, where $b = 9a/2$.

(i) Evaluate the 2-D Fourier transform, $G(\omega_1,\omega_2)$, of the image g. Sketch the main peaks of G in the (ω_1,ω_2) plane and discuss how these peaks can be predicted from the original image g. [40%]

(ii) Making any appropriate assumptions, state how this image should be sampled to avoid aliasing. [10%]

(b) 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 2-D window functions from 1-D window functions and explain the effects that windowing has on the ideal frequency response. Hence outline desirable properties for a window function. [15%]

(ii) Consider the following window functions for $i = 1$ and 2:

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

Evaluate the spectrum, $W_1(\omega_1)$, of the 1-D window, $w_1(u_1)$, and sketch both w_1 and W_1. Hence obtain the spectrum, $W(\omega_1,\omega_2)$, of the 2-D window function $w(u_1,u_2)$ formed from the product of w_1 and w_2. [25%]

(iii) By considering the width of the mainlobe and the magnitude of the sidelobes of W, discuss the merits of this window function. [10%]

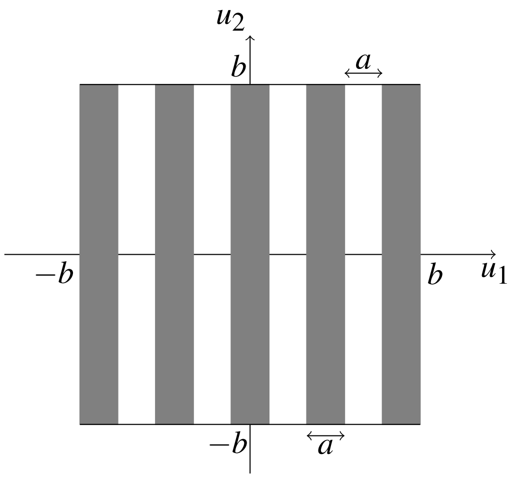


Fig. 1

 (TURN OVER

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) Write, in discrete form, an expression relating the true image to the observed image. Define all quantities used. [10%]

(ii) Explain what is meant by *inverse filtering*, describing how this is used to estimate the true image. Describe also why this method exhibits poor performance and indicate how the *pseudo-inverse* or *generalised inverse* filter can improve performance. [10%]

(iii) All inverse filters perform poorly in the presence of significant noise. Improved performance can be obtained via the *Wiener filter*, g, whose spectrum is given by

$$G(\omega) = \frac{P_{xy}(\omega)}{P_{yy}(\omega)}$$

where P_{xy} is the cross-power spectrum of the true and observed images, and P_{yy} is the power spectrum of the observed image. Derive a form of G in terms of the power spectrum of the true image, P_{xx}, the power spectrum of the noise, P_{dd}, and the spectrum, H, of the point spread function. State any assumptions made. [30%]

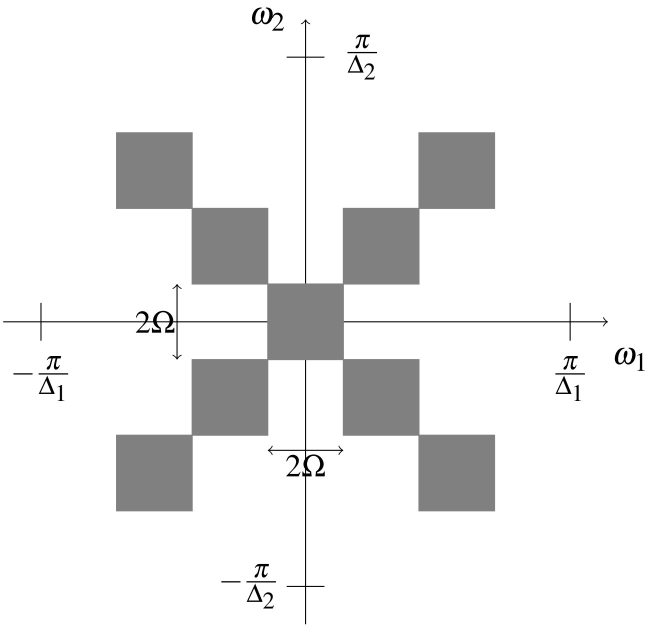
(b) Consider the ideal bandpass filter shown in Fig. 2, with $H = 1$ in the shaded regions and $H = 0$ otherwise. All outer shaded regions are of the same size as the central shaded region. Sampling is carried out on a rectangular grid with spacings of Δ_1 and Δ_2 in the u_1 and u_2 directions respectively.

(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

$$h(n_1\Delta_1, n_2\Delta_2) = \Omega^2 \frac{\Delta_1\Delta_2}{\pi^2} \sum_{k=0}^{2} \alpha_k \,\mathrm{sinc}([2k+1]\Omega n_1\Delta_1)\,\mathrm{sinc}([2k+1]\Omega n_2\Delta_2) + \text{ other terms}$$

and find the values of α_k, $k = 0, 1, 2$. [40%]

(ii) Describe the effects when a filter with this frequency response acts on an image. [10%]



ω_2
$\frac{\pi}{\Delta_2}$
2Ω
$-\frac{\pi}{\Delta_1}$
$\frac{\pi}{\Delta_1}$
ω_1
2Ω
$-\frac{\pi}{\Delta_2}$

Fig. 2

3 (a) The matrices for transforming image pixels between RGB and YUV colour spaces are given by

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

Describe key reasons for representing pixels in YUV space for image coding and explain why the top row of **C** should sum to unity and why its other two rows should sum to zero? [20%]

(b) Explain how a 4×4 two-dimensional discrete cosine transform (DCT) is typically applied to a monochrome (single colour) image and how the result can be viewed as 16 small sub-images. [20%]

(c) A colour image of size 2048×1536 pixels is converted to YUV space and the U and V image planes are each subsampled 4:1 to give U' and V' by taking the mean of each non-overlapping 2×2 block of pixels from U and V respectively. The Y, U' and V' image planes are then transformed, using the 4×4 DCT, to produce 16 sub-images for each plane. The entropies of the sub-images in bits/pixel are approximately modelled by

$$H_Y(i,j) = \frac{6}{i+j-1} \quad \text{and} \quad H_{UV}(i,j) = \frac{6}{(i+j-1)^2}$$

where $i = 1 \ldots 4$ and $j = 1 \ldots 4$ are the indices of the sub-images in each plane in order of increasing DCT frequency in the vertical and horizontal directions. H_Y refers to entropies of the Y sub-images, while H_{UV} refers to entropies of the U and V sub-images. Estimate the number of bits needed to code each 2048×1536 colour image, stating your assumptions. [40%]

(d) Comment on the bit-rate cost of including colour in the coded image data and explain how this relates to known properties of the human visual system. [20%]

4 (a) The basic Haar transform matrix for 2-element vectors is given by

$$\mathbf{T} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Show how this can be employed to perform a 2-dimensional transformation on a 2×2 matrix of pixels, and also how an inverse transformation can be used to recover the pixels from a 2×2 matrix of transform coefficients. What property of $\mathbf{T}$ allows the inverse transform to be computed efficiently? [15%]

(b) Explain how the 2-D Haar transform may be applied to images in a multi-level (or multi-scale) way for image coding. How many output subbands would be produced from a 4-level transform and how large would they each be when the input image is of size $N \times N$ pixels? [20%]

(c) When a multi-level Haar transform is used to encode an image at relatively low bit rates, blocking artifacts tend to be produced. Describe the nature of these artifacts and explain how the Haar transform may be extended to become a wavelet transform that can make the coding artifacts less visible. [25%]

(d) A relatively simple pair of wavelet analysis filters for 1-D signals are defined in the z-domain by

$$H_0(z) = \tfrac{1}{8}(-z^2 + 2z + 6 + 2z^{-1} - z^{-2}) \quad \text{and} \quad H_1(z) = \tfrac{1}{2}z^{-1}(-z + 2 - z^{-1})$$

The corresponding perfect-reconstruction filters are

$$G_0(z) = \tfrac{1}{2}(z + 2 + z^{-1}) \quad \text{and} \quad G_1(z) = \tfrac{1}{8}z(-z^2 - 2z + 6 - 2z^{-1} - z^{-2})$$

Calculate the filter products $H_0(z)\,H_1(z^2)$ and $G_0(z)\,G_1(z^2)$. Explain why these represent level-2 wavelet basis functions (for 1-D signals). [25%]

(e) Discuss how the smoothness of these functions would affect the visibility of artifacts produced in an image coder that is based on these filters, and hence determine whether we should swap the H filters with the G filters to achieve improved performance. [15%]

END OF PAPER

THIS PAGE IS BLANK

4F8 2015 Answers:

1 (a) (i) $G(\omega_1,\omega_2) = 2Aab\ \text{sinc}\, b\omega_2\ \text{sinc}\frac{a\omega_1}{2}\ [1 + 2\cos 2a\omega_1 + 2\cos 4a\omega_1]$

(ii) Sampling period should be $\leq a/6$ (approx).

(b) (i) –

(ii) $W_1(\omega_1) = U_1\ \text{sinc}^2\frac{\omega_1 U_1}{2}$; $W(\omega_1,\omega_2) = U_1U_2\ \text{sinc}^2\frac{\omega_1 U_1}{2}\ \text{sinc}^2\frac{\omega_2 U_2}{2}$

(iii) –

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

(ii) –

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

(b) (i) $\alpha_0 = 2,\ \alpha_1 = 18,\ \alpha_2 = 25$

(ii) –

3 (a) –

(b) –

(c) 7.33 . 10^6 bits

(d) It costs 23% more bits to include colour.

4 (a) –

(b) 13 subbands; 3 subbands of size $\frac{N}{2}\times\frac{N}{2}$, 3 subbands $\frac{N}{4}\times\frac{N}{4}$, 3 subbands $\frac{N}{8}\times\frac{N}{8}$, and 4 subbands $\frac{N}{16}\times\frac{N}{16}$.

(c) –

(d) The coefs of the filters are:
$H_0(z)\,H_1(z^2)$: $\frac{1}{16}\{1\ \ -2\ \ -8\ \ 2\ \ 14\ \ 2\ \ -8\ \ -2\ \ 1\}$
$G_0(z)\,G_1(z^2)$: $\frac{1}{16}\{-1\ \ -2\ \ -3\ \ -4\ \ 4\ \ 12\ \ 4\ \ -4\ \ -3\ \ -2\ \ -1\}$

(e) Do not swap the filters.