UNIVERSITY OF LONDON

MSci EXAMINATION 2003

For Internal Students of

Royal Holloway

DO NOT TURN OVER UNTIL TOLD TO BEGIN

PH4760A: IMAGE ANALYSIS AND RECOGNITION

Time Allowed: TWO AND A HALF hours

Answer THREE questions

No credit will be given for attempting any further questions

Approximate part-marks for questions are given in the right-hand margin

Only CASIO fx85WA Calculators are permitted

PH4760A/22

GENERAL PHYSICAL CONSTANTS

Permeability of vacuum	μ_0	=	$4\pi \times 10^{-7}$	H m^{-1}
Permittivity of vacuum	ε_0	=	8.85×10^{-12}	F m^{-1}
	$1/4\pi\varepsilon_0$	=	9.0×10^{9}	m F^{-1}
Speed of light in vacuum	c	=	3.00×10^{8}	m s^{-1}
Elementary charge	e	=	1.60×10^{-19}	C
Electron (rest) mass	m_e	=	9.11×10^{-31}	kg
Unified atomic mass constant	m_u	=	1.66×10^{-27}	kg
Proton rest mass	m_p	=	1.67×10^{-27}	kg
Neutron rest mass	m_n	=	1.67×10^{-27}	kg
Ratio of electronic charge to mass	e/m_e	=	1.76×10^{11}	C kg^{-1}
Planck constant	h	=	6.63×10^{-34}	J s
	$\hbar = h/2\pi$	=	1.05×10^{-34}	J s
Boltzmann constant	k	=	1.38×10^{-23}	J K^{-1}
Stefan-Boltzmann constant	σ	=	5.67×10^{-8}	W m^{-2} K^{-4}
Gas constant	R	=	8.31	J mol^{-1} K^{-1}
Avogadro constant	N_A	=	6.02×10^{23}	mol^{-1}
Gravitational constant	G	=	6.67×10^{-11}	N m^2 kg^{-2}
Acceleration due to gravity	g	=	9.81	m s^{-2}
Volume of one mole of an ideal gas at STP		=	2.24×10^{-2}	m^3
One standard atmosphere	P_0	=	1.01×10^{5}	N m^{-2}

MATHEMATICAL CONSTANTS

$e \cong 2.718$ $\pi \cong 3.142$ $\log_e 10 \cong 2.303$

PART MARKS

1. ANSWER ONLY **FOUR** sections of ***Question ONE***.

(a) Determine the result of applying a 3-element median filter to the following 1-dimensional signals:

(i)	0	0	0	0	0	1	0	1	1	1	1	1	1	1
(ii)	2	1	2	3	2	1	2	2	3	2	4	3	3	4
(iii)	1	1	2	3	3	4	5	8	6	6	7	8	9	9

What general lessons can be learnt from the results? In the first case consider also the corresponding situation for a grey-scale edge in a 2-dimensional image. **[5]**

(b) The objects in a binary image are to be labelled according to their number in a forward raster scan over the image. Draw diagrams showing what happens for round blobs. Show also what happens for more complex shapes. Show how any problems may be overcome by use of suitable algorithm strategies. **[5]**

(c) The image of an egg is to be thresholded adaptively to reveal any dark cracks in the shell. Explain why the following formula for determining the local threshold should be effective:

(i) $thr = mean - (max - mean)$

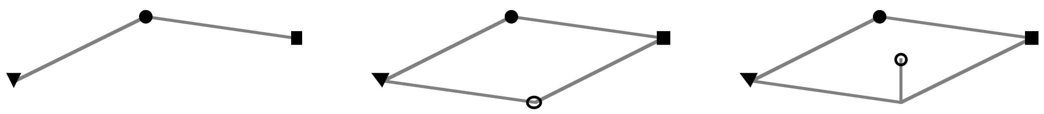
Explain why the following formula should be a useful starting point for reading dark print on paper:

(ii) $thr = 0.5(min + max)$

Indicate why and how it needs to be modified to make it work in practical cases. **[5]**

(d) Explain the concepts *focus of expansion* and *focus of contraction*. Explain how an image can have several such foci. List the various instances of translation and/or rotation of camera and/or objects that can lead to the appearance of such foci. What general cases result in no foci coupled with much greater difficulty in interpretation? **[5]**

(e) Explain briefly each of the following: (i) Why weak perspective projection leads to an ambiguity in viewing an object such as that in Figure 1(a). (ii) Why the ambiguity doesn't disappear for the case of Figure 1(b). (iii) Why the ambiguity *does* disappear in the case of Figure 1(c), if the true nature of the object is known. (iv) Why the ambiguity doesn't occur in the case of Figure 1(b) viewed under full perspective projection. In the last case, illustrate your answer by means of a sketch. **[5]**



a **b** **c**

Figure 1 In this diagram the grey edges are construction lines, *not* parts of the objects. (a) and (b) are completely planar objects.

TURN OVER

PART MARKS

2. (a) Describe the *Hough transform* approach to object location. Explain its advantages relative to the *centroidal* (r, θ) *plot* approach: illustrate your answer with reference to location of circles of known radius R. **[4]**

(b) Describe how the Hough transform may be used to locate straight edges. **[3]**

(c) Explain what happens if the image contains a square object of arbitrary size and *nothing* else. How would you deduce from the information in parameter space that a square object is present in the image? Give the main features of an algorithm to decide that a square object is present and to locate it. **[3]**

(d) Examine in detail whether an algorithm using the strategy described in (c) would become confused if (i) parts of some sides of the square were occluded; (ii) one or more sides of the square were missing; (iii) several squares appeared in the image; (iv) several of these complications occurred together. **[7]**

(e) How important is it to this type of algorithm to have edge detectors that are capable of accurately determining edge orientation? Describe a type of edge detector that is capable of achieving this. **[3]**

3. (a) Explain the concept of a *distance function*. Illustrate your answer by means of a diagram. **[3]**

(b) *Outline* how propagation of a distance function may be carried out using a parallel algorithm. Give in full a simpler algorithm that operates using two sequential passes over the image. **[4]**

(c) It has been suggested that a four-pass sequential algorithm will be even faster than the two-pass algorithm, as each pass can use just a 1-dimensional window involving at most three pixels. Write down the code for *one* typical pass of the algorithm. **[2]**

(d) Estimate the approximate speeds of these three algorithms for computing the distance function, in the case of an $N \times N$ pixel image. Assume a normal serial computer is used to perform the computation. **[4]**

(e) A certain type of dark insect* is to be located amongst cereal grains* by (i) applying an edge detector which will mark all the edge points in the image as 0's in a 1's background, (ii) applying a propagation algorithm to the *background* region, (iii) locating the local maxima of the distance function, (iv) analysing the values of the local maxima, and (v) carrying out necessary further processing to locate the nearly parallel sides of the insects. Assuming that the image is not so large that the distance function will overflow the byte limit, explain how to design stages (iv) and (v) of the algorithm in order to identify the insects, yet ignore the cereal grains. How robust will this method be if the edge is broken slightly in a few places? **[7]**

---

* In this question, assume that the insects approximate to rectangular bars of dimensions about 20 by 7 pixels, and that the cereal grains are approximately elliptical with dimensions about 40 by 24 pixels.

4. (a) Distinguish between *full perspective projection* and *weak perspective projection*. Explain how each of these projections presents oblique views of the following real objects: (i) straight lines, (ii) several concurrent lines (i.e. lines meeting in a single point), (iii) parallel lines, (iv) mid-points of lines, (v) tangents to curves, (vi) circles whose centres are marked with a dot. **[4]**

(b) Explain the value of using *invariants* in relation to pattern recognition systems. **[3]**

(c) The *cross ratio* of four points (P_1, P_2, P_3, P_4) on a line is defined as the ratio:

$$C(P_1, P_2, P_3, P_4) = \frac{(x_3 - x_1)(x_2 - x_4)}{(x_2 - x_1)(x_3 - x_4)}$$

Explain in what way this embodies the *ratio of ratios* idea, and what makes this a useful type of invariant for objects viewed under full perspective projection. Show that labelling the points in reverse order will not change the value of the cross ratio. **[5]**

(d) Give arguments why the cross-ratio concept should also be valid for weak perspective projection. Work out a simpler ratio-based concept that is valid for straight lines viewed under weak perspective projection. **[4]**

(e) A flat lino-cutter blade has two parallel sides of different lengths: it is viewed under weak perspective projection. Discuss whether it can be identified from any orientation in 3-dimensions by measuring the lengths of its sides. **[4]**

5. (a) Explain the term *shape from shading*. **[3]**

(b) Compare the properties of matt surfaces with those which exhibit 'normal' specular reflection. Matt surfaces are sometimes described as 'Lambertian'. Describe how the brightness of the surface varies according to the Lambertian model. **[3]**

(c) Show that for a given surface brightness, the orientation of any point on a Lambertian surface must lie on a certain cone of orientations. **[2]**

(d) Three images of a surface are obtained on illuminating it in sequence by three independent point light sources. Show with the aid of a diagram how this can lead to unambiguous estimates of surface orientation. Would surface orientation of any points on the surface *not* be estimated by this method? Are there any constraints on the allowable positions of the three light sources? Would it help if *four* independent point light sources were used instead of three? **[5]**

(e) Discuss whether the surface map that is obtained by shape from shading is identical to that obtained by stereo (binocular) vision. Are the two approaches best applied in the same or different applications? To what extent is the application of structured light able to give better or more accurate information than these basic approaches? **[5]**

(f) Consider briefly what further processing is required before 3-dimensional objects can be recognised by any of these approaches. **[2]**

END.

ERD