UNIVERSITY OF LONDON

MSci EXAMINATION 2004

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*** only

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 or CASIO fx85MS Calculators are permitted

PH4760A/30

2003-04

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

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

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

(a) Explain the difficulty of forming correspondences between two images in order to use binocular vision to estimate the depths of feature points in a scene. Outline how the epipolar line concept permits correspondences to be found more efficiently, with fewer false correspondences. Consider the situation for aliens with three or four eyes. Would they have any advantages over humans or two-eyed robots? **[5]**

(b) Explain why the median filter is very effective at removing impulse noise from digital images. 2-D median filters are sometimes implemented as two 1-D median filters applied in sequence in order to improve the speed of processing. Why is this necessary? Estimate how much a 3×3 median filter could be speeded up in this way. What about a 7×7 median filter? Would any loss of effectiveness result from use of this technique? **[5]**

(c) A morphological gradient binary edge enhancement operator is defined by the formula:

$$G = (A \oplus B) - (A \ominus B)$$

where A is an image and B is a morphological structuring element.

Using a 1-D model of an edge, or otherwise, show that this will give wide edges in binary images. If grey-scale dilation ($\oplus$) is equated to taking a local *maximum* of the intensity function within a 3×3 window, and grey-scale erosion ($\ominus$) is equated to taking a local *minimum* within a 3×3 window, sketch the result of applying the operator G. Show that it is similar in effect to a Sobel edge enhancement operator, if edge orientation effects are ignored by taking the Sobel magnitude:

$$g = (g_x^2 + g_y^2)^{1/2}$$ **[5]**

(d) What is meant by *pattern recognition*? Explain why it is not practical to perform pattern recognition by comparing incoming patterns with a dictionary of all possible patterns. How does the nearest neighbour technique help to overcome this problem? **[5]**

(e) What are the main parameters that can be used to characterise a texture? Use the parameters you have given to help you describe a number of textures that commonly occur in the natural and man-made worlds. Devise a simple algorithmic process which could help to find faults in textured surfaces, with a view to producing an industrial inspection system. **[5]**

PART MARKS

2. (a) Explain the *connectedness paradox* in a binary image, and how it is resolved. **[3]**

(b) Define *thinning*, making clear the role that connectedness plays in the concept. Explain the importance of *crossing number* in the design of thinning algorithms. Write down a formula for crossing number, giving examples of 3×3 windows which show why the formula has to be more complex than might *a priori* have been expected. **[4]**

(c) Give a basic 'fill' procedure that can be used as part of a convex hull algorithm. Show that connectedness plays a part in adding to the complexity of the fill procedure. **[3]**

(d) Give a simple one-pass algorithm for labelling the objects appearing in a binary image, making clear the role played by connectedness. Give examples showing how this basic algorithm goes wrong with real objects: illustrate your answer with clear pixel diagrams, which show the numbers of labels that can appear with objects of different shapes. **[4]**

(e) Show how a table-orientated approach can be used to eliminate multiple labels in objects. Make clear how the table is set up and what numbers have to be inserted into it. Are the number of iterations needed to analyse the table similar to the number that would be needed in a multi-pass labelling algorithm taking place entirely within the original image? Discuss where the *real* gain in using a table to analyse the labels is coming from. **[6]**

PART MARKS

3. (a) Describe the main stages in the application of the Hough transform to locate objects in digital images. What are the particular advantages offered by the Hough transform technique? Give reasons why they arise. **[3]**

(b) It is said that the Hough transform only leads to *hypotheses* about the presence of objects in images, and that they should all be checked independently before making a final decision about the contents of any image. Comment on the accuracy of this statement. **[2]**

(c) Outline each of the following two methods for locating ellipses using the Hough transform: (i) the *diameter–bisection* method; (ii) the *chord–tangent* method. Explain the principles on which these methods rely. Determine which is the more robust and compare the amounts of computation each requires. **[6]**

(d) For the diameter–bisection method, searching through lists of edge points with the right orientations can take excessive computation. It is suggested that a two-stage approach might speed up the process: (i) load the edge points into a table which may be addressed by orientation; (ii) look up the right edge points by feeding appropriate orientations into the table. Estimate how much this would be likely to speed up the diameter–bisection method. **[4]**

(e) It is found that the diameter–bisection method sometimes becomes confused when several ellipses appear in the same image, and generates false 'centres' that are not situated at the centres of any ellipses. It is also found that certain other shapes are detected by the diameter–bisection method. Ascertain in each case quite what the method is sensitive to, and consider ways in which these problems may be overcome. **[5]**

PART MARKS

4. (a) Figure 1 shows a 2-dimensional view of a widget with four corners. Explain how the maximal clique technique can be used to locate widgets even if they are partly obscured by various types of object including other widgets. **[5]**

(b) Explain why the basic algorithm will not distinguish between widgets that are normally presented from those that are upside down. Consider how the basic method could be extended to ensure that a robot only picks up those that are the right way up. **[5]**

(c) The camera used to view the widgets is accidentally jarred and then reset at a different unknown height above the worktable. State clearly why the usual maximal clique technique will now be unable to identify the widgets. Discuss how the overall program could be modified to make sense of the data and make correct interpretations in which all the widgets are identified. Assume *first* that widgets are the only objects appearing in the scene, and *second* that a variety of other objects may appear. **[5]**

(d) The camera is jarred again and this time is set at a small unknown angle to the vertical. To be sure of detecting such situations and of correcting for them, a flat calibration object of known shape is to be stuck on the worktable. Decide on a suitable shape and explain how it should be used to make the necessary corrections. **[5]**

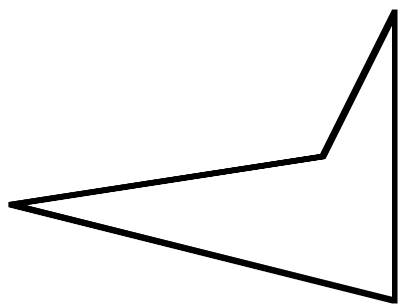

Figure 1 Diagram of a widget

PART MARKS

5. (a) Distinguish between *full perspective projection* and *weak perspective projection*. **[2]**

(b) Explain what is meant by *perspective inversion*, and how it is eliminated in practical instances. **[3]**

(c) Draw up a complete table of pose ambiguities that arise for *weak* perspective projection, for various numbers of object points identified in the image. Your answer should cover both coplanar points and non-coplanar points, and should make clear in each case how much ambiguity would remain in the limit of an infinite number of object points being seen. **[4]**

(d) A geometric calculation based on the sine rule shows that the angles α, β, γ are related to the distances a, b, c in Figure 2 by the equation:

$$\frac{a}{\sin\alpha} \times \frac{c}{\sin\gamma} = \frac{a+b}{\sin(\alpha+\beta)} \times \frac{b+c}{\sin(\beta+\gamma)}$$

Show that this leads to a relation between the cross ratios for various distances on the line and for the sines of various angles. Hence show that this also leads to the constancy of the cross ratios on two lines crossing the pencil of four lines passing though O. **[4]**

(e) Explain the value of the cross ratio in relating distances even in the presence of huge amounts of perspective distortion. **[3]**

(f) Flagstones are observed by a robot as it walks along a pavement. To analyse the scene fully the robot needs to identify the vanishing point V in the image. Show that it is possible to do this quite accurately by assuming that two adjacent flagstones are equal in size: to proceed with the calculation merely apply the cross ratio to the various distances in Figure 3 to find c. **[4]**

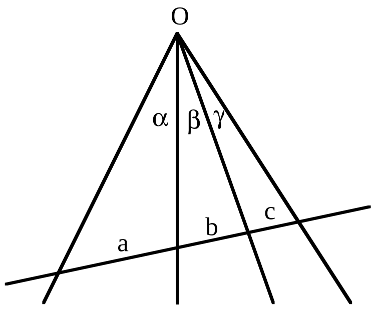

Figure 2 Geometry for cross ratio calculation

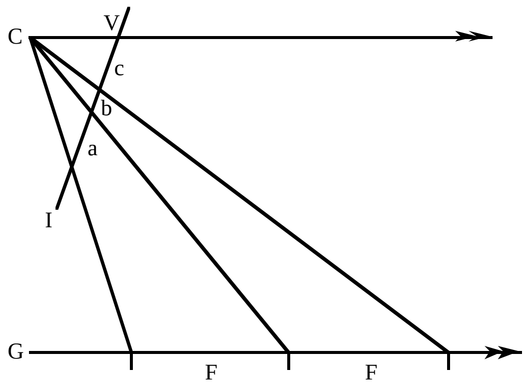

Figure 3 Robot's view of two flagstones. C is the centre of projection. Each F signifies a flagstone on the ground plane G. *a* and *b* are the apparent lengths of the flagstones. V is the vanishing point in the image plane I. the vanishing point V appears on a line parallel to the ground plane G.

END.

ERD