UNIVERSITY OF LONDON
IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 2001

MEng Honours Degree in Information Systems Engineering Part IV
MEng Honours Degrees in Computing Part IV
MSc in Advanced Computing
for Internal Students of the Imperial College of Science, Technology and Medicine

This paper is also taken for the relevant examinations for the Associateship of the City and Guilds of London Institute

PAPER C418=I4.2

COMPUTER VISION

Friday 18 May 2001, 14:30
Duration: 120 minutes

Answer THREE questions

Paper contains 4 questions
Calculators required

1. Image Segmentation

(a) i) Define a quadtree, and explain how it can be used for splitting and merging in region based image segmentation.

ii) An alternative to using a quadtree would be to divide the image uniformly into squares of say 8 by 8, and computing the properties of each block. What would be the advantages and disadvantages of adopting this approach for region segmentation?

(b) Under certain circumstances, it is possible to characterise a region of an image by the mean and variance of its pixel intensities. Briefly explain the circumstances under which this can be done, and give an example where it would be wrong to do so.

(c) An experiment was carried out in a biochemistry laboratory to analyse the relative abundance and characteristics of the three cell types defined in Figure 1.1. Due to the limitation of the available image acquisition process, the scanned images are usually corrupted by steak noise as shown in Figure 1.2.

i) Propose a suitable technique for removing the background noise.

ii) Define a method for unambiguous segmentation of these three different cell types. The proposed approach should be immune to rotation, scaling, and minor distortion of the cells. To simplify your algorithm, it can be assumed that there is no overlap or connection between cells.

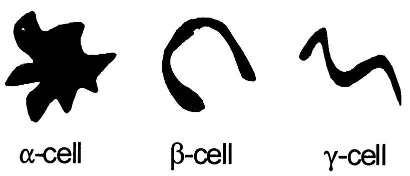


Figure 1.1 Defined cell types

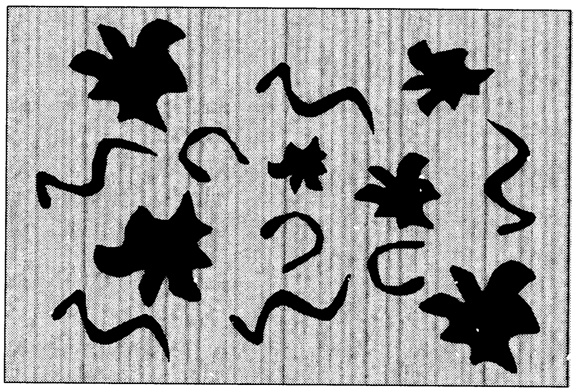


Figure 1.2 An example cell image

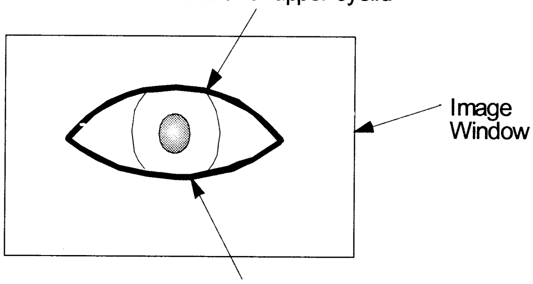
The three parts carry, respectively, 30%, 20%, 50% of the marks.

2. Hough Transform

(a) i) Explain why it is important to incorporate **heuristics** and ***a priori* knowledge** in computer vision systems.

ii) In using the Hough transform for line detection, what heuristics can be used?

(b) As part of a face recognition program the Hough transform is being used to determine the position of the eyes. The idea is to model the eyelids as a pair of parabolas as shown in the diagram



The function used to describe the parabola is

$$y=a(x-x_c)^2+y_c \qquad 2.1$$

It can be observed that the point $[x_c, y_c]$ is the apex of the parabola, and that the parameter a determines the shape of the parabola. Assuming that the Hough transform will be performed in small windows of 32 by 32 pixels, suggest suitable bounds on the histogram array for the parameters.

(c) For Question (b), write an implementation for the generation of the histogram array for this Hough transform in the language (or pseudocode) of your choice. (You may assume that procedures exist to find the edge points in the image.)

(d) For the above problem of face recognition, what constraints could you apply when searching the histogram for peaks that would indicate the presence of an eye?

(e) Propose a method by incorporating edge direction for simplifying the Hough Transform for parabolas defined by Equation 2.1.

The five parts carry, respectively, 20%, 20%, 30%, 10%, and 20% of the marks.

3. Computational Stereo

A vision system uses two cameras placed at different positions viewing at the same scene.

a) Explain briefly what is meant by the *epipolar* line.

b) The cameras are placed at positions Cl = [-10,0,0] and Cr = [10,0,0]. The cameras use a normal pinhole projection with their plane of projection lying in the plane $z=10$. An object point has position Pi=[xi, yi, zi]. Derive an expression for the epipolar line that appears in the ***right*** hand camera.

c) Sketch the epipolar lines that correspond to the point [-10,20,20] on the ***right*** hand camera plane.

d) Explain how the epipolar line can be used in a search algorithm to find corresponding points in the two camera images.

e) How would the algorithm change if there were three, rather than two cameras available?

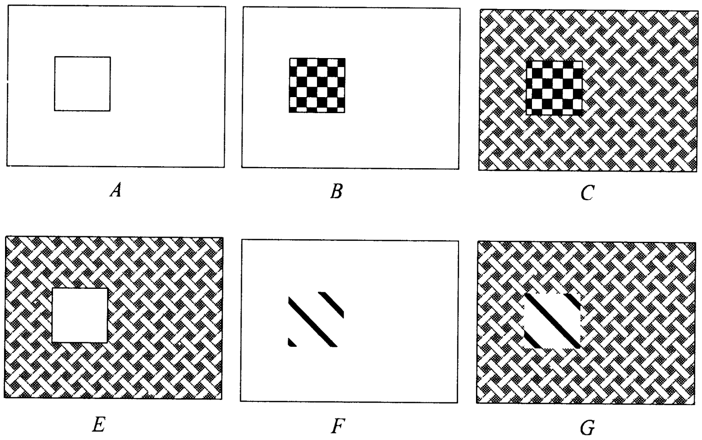
The five parts carry, respectively, 10%, 30%, 20%, 20%, 20% of the marks.

4. Motion Analysis

(a) i) Explain the term **optical flow** and the usual assumptions in using it for motion analysis in an image sequence.

ii) Provide two examples for which the optical flow technical would fail.

(b) Consider the following different scenarios when a square object is moving towards the upper right corner, sketch out the detected optical flow directions at the **object borders**.



(c) Use the basic optical flow equation defined below to explain your findings for question (b).

$$(\partial f/\partial x)\, dx + (\partial f/\partial y)\, dy + (\partial f/\partial t)\, dt = 0$$

(d) Under perspective projection, if the position of a point in the 3D space is (X_0, Y_0, Z_0) at time $t_0=0$, then the position of the same point projected onto the image plane at time t can be determined as follows by assuming unit focal distance of the optical system and constant velocity:

$$(x,y)=\left(\frac{X_0+Ut}{Z_0+Wt}, \frac{Y_0+Vt}{Z_0+Wt}\right) \qquad (4.1)$$

where U,V,W are velocity components along the X, Y, and Z axes respectively.

i) Use Equation (4.1) to explain and derive the term **focus of expansion** (FOE) used in motion analysis.

ii) Explain how FOE can be used to estimate relative depth of an object moving with constant velocity. Discuss whether a practical algorithm could be based on this approach.

The four parts carry, respectively, 20%, 20%, 20%, 40% of the marks.