UNIVERSITY OF LONDON
IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 2004

MEng Honours Degree in Electrical Engineering Part IV
MEng Honours Degree in Information Systems Engineering Part IV
MSci Honours Degree in Mathematics and Computer Science 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
This paper is also taken for the relevant examinations for the Associateship of the Royal College of Science

PAPER C418=I4.2

COMPUTER VISION

Friday 7 May 2004, 14:30
Duration: 120 minutes

Answer THREE questions

Paper contains 4 questions
Calculators required

1. Image Segmentation

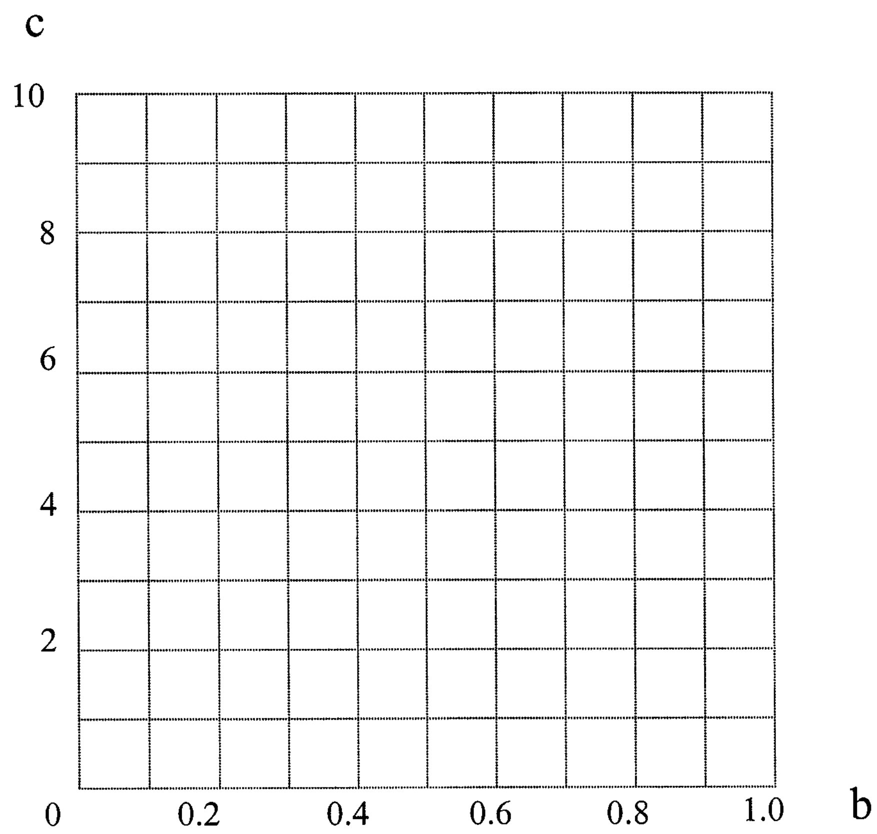
(a) i) Explain the terms *edge-based segmentation* and *region-based segmentation*.

ii) Propose two types of uniformity measures, one with which textured and the other with which untextured regions could be segmented.

(b) Explain why it is important to incorporate *heuristics* and *a priori knowledge* in computer vision systems. In using the Hough transform for line detection, what heuristics can be used?

(c) A system is designed to extract parabolic curve segments from a monochrome shaded image. The curve is defined by the equation $y = ax^2 + bx + c$. Explain how the Hough transform could be used to find these curve segments.

(d) In Question (c), we assume that a = 0.01 and the following edge points (x, y) have been extracted: $P_A = (40, 34)$, $P_B = (0, 6)$, $P_C = (30, 24)$ and $P_D = (10, 10)$. Sketch the corresponding representation in the (b, c) parameter space, shown below, defined by the parabolic equation. Indicate the point where the best-fit parabolic curve that will be extracted by the Hough transform (hint: you can substitute a = 0.01 to simplify the equation first).



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

2. Photometric Stereo

The photometric stereo method used for computing the depth of an image at each pixel is based on inverting Lambert's Cosine law which states:

$$I(x,y) = \rho(x,y)\frac{\mathbf{n} \bullet \mathbf{s}}{|\mathbf{n}| \cdot |\mathbf{s}|} = \rho(x,y)\frac{s_x p + s_y q + s_z}{\sqrt{p^2 + q^2 + 1}\sqrt{s_x^2 + s_y^2 + s_z^2}} \qquad (2.1)$$

where I is the intensity measured at pixel (x,y), $\rho(x,y)$ is the albedo, $\mathbf{n} = (p,q,1)^T$ is the surface gradient vector, and $\mathbf{s} = (s_x, s_y, s_z)^T$ is a vector from the point to the light source.

(a) In photometric stereo, what is the advantage of using three rather than two different illuminating conditions to solve for Equation (2.1)?

(b) Three light sources are to be used in an experiment. These light sources are located far away such that the direction of vectors **s** may be taken as a constant over the image. If the measured data are as follows:

$I(x, y) = I_1$ when $\mathbf{s} = (1,1,0)^T$
$I(x, y) = I_2$ when $\mathbf{s} = (1,0,1)^T$
$I(x, y) = I_3$ when $\mathbf{s} = (0,1,0)^T$

derive expressions for the p, q components of the surface gradient vector in terms of I_1, I_2, and I_3.

(c) From Equation (2.1), a single measurement confines the unknown surface gradient to a conic section (hyperbola, parabola, or ellipse) in pq-space. If we move the light source to obtain two measurements, I_1 and I_2, this confines the surface gradient to the intersection of two conic sections in gradient space. In general, two conic sections can intersect in as many as four distinct places. Prove that there are at most two solutions for p and q when I is represented in the form given by Equation (2.1).

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

3. The Hough Transform

A Hough transform is to be used to extract circles from an image. The image has a resolution of 256 by 256 and each of the circles has a radius of less than 64 pixels. Using a pseudo code of your choice, indicate how the implementation could be carried out in the following three cases:

(a) Using a three-dimensional histogram array for x_c, y_c and r, with x_c, y_c specifying the centre and r the radius of the circle.

(b) Determining the centres with a two-dimensional histogram first, using gradient information as well as position, and then finding the radii.

(c) Determining the centres with a two-dimensional histogram first, using position information only, and then finding the radii.

(d) Comment on the advantages and disadvantages of the three methods used in parts a, b and c.

The four parts carry, respectively, 25%, 25%, 25%, 25% of the marks.

4. Motion Analysis

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

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

ii) Denote image velocity (u, v) as (U, V, W) projected onto the image plane. Derive the expressions for u and v by referring to Equation (4.1), notice that $(u, v) = \left(\frac{dx}{dt}, \frac{dy}{dt} \right)$.

(c) Let $D(t)$ be the distance of a point from the FOE measured in the image plane, and $K(t)$ the magnitude of its measured image velocity, *i.e.*,

$$K(t) = \sqrt{u(t)^2 + v(t)^2} \qquad 4.2$$

Derive the expression for $D(t)$, and calculate the ratio between $D(t)$ and $K(t)$, and explain how it 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 three parts carry, respectively, 20%, 40%, 40% of the marks.