UNIVERSITY OF LONDON
IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 2003

MEng Honours Degree in Information Systems Engineering Part IV
MEng Honours Degrees in Computing Part IV
MSc in Advanced Computing
PhD
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 2 May 2003, 10:00
Duration: 120 minutes

Answer THREE questions

Paper contains 4 questions
Calculators required

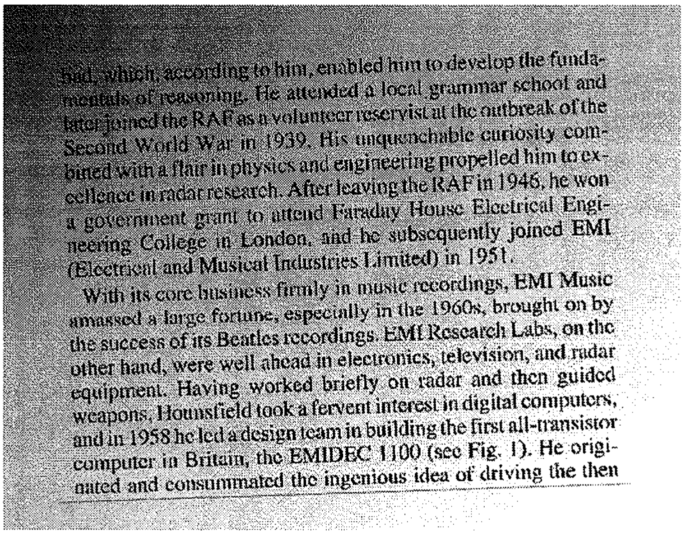
1. Image Segmentation

(a) Explain the terms *low-level processing* and *high-level processing* in visual recognition. What is the primary difference between the two?

(b) Explain why it is important to incorporate *heuristics* and *a priori knowledge* in computer vision systems. What heuristics can be used in using dynamic programming for linking edge segments into boundary contours?

(c) Given Fischer's criterion: $\dfrac{|\mu_1 - \mu_2|}{\sqrt{\gamma_1^2 + \gamma_2^2}}$, where μ_i and γ_i $(i=1,2)$ are the mean and variance of the pixels within each region, explain how it can be used to determine whether two regions should be merged or separated.

(d) Optical Character Recognition (OCR) systems use scanned documents to segment out and recognise individual characters. In practice, it is difficult to control the background illumination and the exact orientation of the document. Based on the scanned image shown below, describe a system that is suitable for image segmentation and character recognition. Explain and illustrate the processing steps involved.



had, which, according to him, enabled him to develop the fundamentals of reasoning. He attended a local grammar school and later joined the RAF as a volunteer reservist at the outbreak of the Second World War in 1939. His unquenchable curiosity combined with a flair in physics and engineering propelled him to excellence in radar research. After leaving the RAF in 1946, he won a government grant to attend Faraday House Electrical Engineering College in London, and he subsequently joined EMI (Electrical and Musical Industries Limited) in 1951.

With its core business firmly in music recordings, EMI Music amassed a large fortune, especially in the 1960s, brought on by the success of its Beatles recordings. EMI Research Labs, on the other hand, were well ahead in electronics, television, and radar equipment. Having worked briefly on radar and then guided weapons, Hounsfield took a fervent interest in digital computers, and in 1958 he led a design team in building the first all-transistor computer in Britain, the EMIDEC 1100 (see Fig. 1). He originated and consummated the ingenious idea of driving the then

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

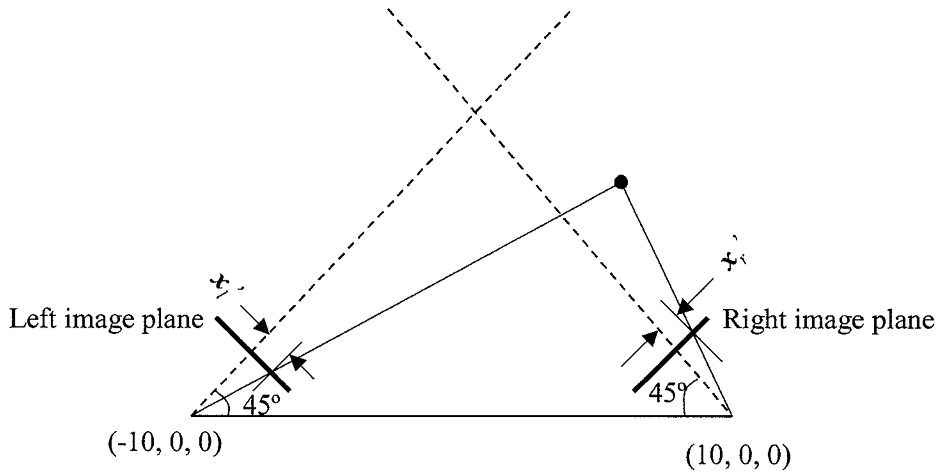
2. Computational Stereo

(a) Explain briefly what is meant by an *epipolar* line in Computational Stereo.

(b) Two identical cameras are placed at positions $C_l = [-10,0,0]$ and $C_r = [10,0,0]$, both facing in direction (0, 0, 1). They use a normal pinhole projection with their plane of projection lying in the plane $z = 10$. An object point has position $P_i = [x_i, y_i, z_i]$. Derive an expression for the epipolar line that appears in the right hand camera.

(c) Based on your answer to 2(b), sketch for the right hand camera the epipolar lines that correspond to the points [10,0,20] and [-10,20,20].

(d) Rather than both pointing towards the same direction, the two cameras used above are re-oriented such that their optical axes intersect at a point in the plane y=0, as shown in the following figure. Derive the equation for a surface in 3D space that has zero disparity, *i.e.*, for all points on the surface, their projections onto the two image planes have the same local image coordinate values.



The four parts carry, respectively, 20%, 30%, 20%, 30% 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. Shape from Motion

(a) Given that an object is moving with velocity components $[u, v, w]$ in 3D space and is being filmed by a camera placed at the origin, derive an expression for the focus of expansion of the object in the image. You can assume that the camera has a focal length of 1 and is pointing towards the Z direction.

(b) The time-to-adjacency equation is normally written as:

$$\frac{D(t)}{V(t)} = \frac{z(t)}{w(t)}$$

explain what is meant by the four variables $D(t)$, $V(t)$, $z(t)$ and $w(t)$, and why it is called the time-to-adjacency equation.

(c) Show how the time-to-adjacency equation might be used to estimate relative depth of an object moving with constant velocity? Discuss whether a practical algorithm could be based on this approach.

(d) By expanding the intensity function $I(x+dx, y+dy, t+dt)$, show that the change in intensity at a pixel is related to the velocity of the object seen from that pixel by the equation:

$$\frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v + \frac{\partial I}{\partial t} = 0$$

Clarify the assumptions that need to be made about the illumination if this is to be made into an algorithm for determining object motion.

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