ENGINEERING TRIPOS PART IIB

Thursday April 2006 2.30 to 4

Module 4F12

COMPUTER VISION AND ROBOTICS

Answer not more than ***three*** *questions.*

All questions carry the same number of marks.

The ***approximate*** *percentage of marks allocated to each part of a question is indicated in the right margin.*

There are no attachments.

You may not start to read the questions printed on the subsequent pages of this question paper until instructed that you may do so by the Invigilator

(TURN OVER

1 Consider an algorithm to detect interest points in a 2D image for use in matching. The image is first smoothed with a low-pass filter before image gradients are computed.

(a) Explain why smoothing is necessary. Which filter kernel is used in practice and how is the 2D convolution performed efficiently? [40%]

(b) A typcial interest point (features of interest) is detected and localised by examining the eigenvalues of the 2×2 matrix

$$\begin{bmatrix} \langle I_x^2 \rangle & \langle I_x I_y \rangle \\ \langle I_x I_y \rangle & \langle I_y^2 \rangle \end{bmatrix}$$

evaluated at each pixel, where $\langle\,\rangle$ denotes a 2D smoothing operation, and where $I_x \equiv \partial I / \partial x$ and $I_y \equiv \partial I / \partial y$. Explain what property of the image features is being used for detection and give an algorithm to determine a finite set of interest points. [40%]

(c) Explain how interest points in different images can be matched. [20%]

2 (a) Derive an expression for the *vanishing point* of parallel lines with direction, **b**, when viewed with a pin-hole camera under perspective projection. Why are the vertical lines in early Renaissance paintings usually without vanishing points? [30%]

(b) Outline an algorithm to recover the position, orientation and internal camera parameters of a CCD camera with no non-linear lens distortion from a single perspective image of a known 3D object. You should state clearly the number of image measurements required and how noisy image measurements are processed in practice. [50%]

(c) Under what viewing conditions is the relationship between the CCD image coordinates and the 3D object/world coordinates linear? [20%]

(TURN OVER

3 A planar object is observed by a camera, producing images with pixel positions (u, v) corresponding to positions (X, Y) on the object.

(a) Show that the relationship between the corresponding points is given by a 2D projective transformation:

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix} \begin{bmatrix} X \\ Y \\ W \end{bmatrix}$$

[20%]

(b) Find an expression for the horizon (the vanishing line of the plane) as a line $l_1 u + l_2 v + l_3 = 0$ in the image in terms of the elements of the transformation matrix in (a). Lines can be represented in homogeneous coordinates by vectors $\mathbf{l} = [\, l_1 \;\; l_2 \;\; l_3 \,]^T$. [30%]

(c) Show how the horizon, or vanishing points, can be used to recover the orientation of the camera if the internal parameters are known. [20%]

(d) Derive the relationship between lines on the object (represented in homogeneous coordinates) and their correspondences in the image in terms of the elements t_{ij} of the projective transformation. Show how line correspondences can be used to recover the projective transformation. [30%]

4 Consider a stereo vision system with known projection matrices $P = K[I|0]$ and $P' = K[R|t]$ respectively.

(a) Show how the 3D position of a point can be obtained from measurements in two images. [20%]

(b) What information about two-view camera geometry is encoded by the *essential* and *fundamental* matrices? Give algebraic expressions for these matrices. [20%]

(c) What constraints can be used to find image correspondences? [20%]

(d) Derive an algebraic expression for the *epipolar line* for a point in the left image with pixel coordinates (u, v) in terms of the *fundamental matrix.* [20%]

(e) Can the projection matrices be recovered from the *essential* or *fundamental* matrix? What additional information is required? [20%]

(TURN OVER

5 (a) A hand-held video camera is used to view a small, unknown 3D object from multiple viewpoints under weak perspective. Describe an algorithm to estimate camera positions and the 3D structure from the sequence of uncalibrated images. [50%]

(b) Describe an algorithm to detect faces in images given two large sets of labelled images with faces and background objects. [50%]

END OF PAPER