ENGINEERING TRIPOS PART IIB

Thursday 26 April 2007 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.

STATIONERY REQUIREMENTS
Single-sided script paper

SPECIAL REQUIREMENTS
Engineering Data Book
CUED approved calculator allowed

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

1 A grey level image, $I(x, y)$, is to be smoothed by an isotropic low-pass filter by convolution with a discrete approximation to the 2D Gaussian kernel, $G_\sigma(x, y)$, of size $(2n + 1) \times (2n + 1)$ pixels.

(a) What is the effect on the image of filtering with a Gaussian low-pass filter and when is this useful in computer vision? [20%]

(b) Give an expression for computing the intensity of a smoothed pixel, $S(x, y)$. [20%]

(c) Show how the convolution can be performed by two discrete 1D convolutions along image rows and columns and comment on the computational saving this achieves. [20%]

(d) Determine the filter size and coefficients of the 1D Gaussian kernel for $\sigma = 1$. State clearly any assumptions and approximations used. [20%]

(e) Show how an isotropic band-pass filter can be implemented efficiently from the low-pass filter described above. When are band-pass filters used in computer vision? [20%]

2 The relationship between a 3D world point (X, Y, Z) and its corresponding pixel at image coordinates (u, v) can be written using a 3×4 camera projection matrix as follows:

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ W \end{bmatrix}$$

(a) Under what assumptions is this relationship valid? [10%]

(b) A camera is to be calibrated from a single perspective image of a known 3D object from the image measurements (u_k, v_k) of known reference points (X_k, Y_k, Z_k). By first deriving linear equations in the unknown elements p_{ij} of the projection matrix, outline an algorithm to recover the projection matrix. How many reference points, $k = 1 \ldots N$, are required to estimate all the elements p_{ij} and what geometric layout of reference points should be avoided? State clearly how noisy image measurements are processed in practice. [50%]

(c) Derive expressions for the vanishing points in the image plane of lines parallel to the X, Y and Z axes and show how to recover the camera orientation from this information. [20%]

(d) How should the projection matrix be modified when viewing a distant object with little *relief*? Comment on the advantages of this new relationship. [20%]

(TURN OVER

3 A *mosaiced panorama* of a scene is acquired by rotating a camera about its optical centre.

(a) Derive an expression for the transformation between correspondences in two successive images and show how this transformation depends on the camera rotation between the two views. [30%]

(b) Explain how this transformation can be estimated in practice. Your answer should include details of the localisation of image features, suitable descriptors for matching and an algorithm to reject incorrect matches. [50%]

(c) What is the effect on the mosaic if the rotation of the camera is not about the optical centre? Comment on how image features will be transformed. [20%]

4 In stereo vision a point has 3D coordinates $\mathbf{X}$ in the left camera's coordinate system and $\mathbf{X}' = \mathrm{R}\mathbf{X} + \mathbf{T}$ in the right camera's coordinate system. The perspective projection of this point in the left and right images has image coordinates (u, v) and (u', v') respectively.

(a) Describe the *epipolar constraint* and show that it can be expressed algebraically with the *fundamental matrix*, F.

Your answer should include an expression for the fundamental matrix in terms of the rotation matrix R and translation vector $\mathbf{T}$ between the two cameras and the internal calibration parameter matrices of the left and right cameras, K and K' respectively. [30%]

(b) 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%]

(c) What is meant by the *epipole* and show how to compute its position in the left and right images. [20%]

(d) How can the left and right camera projection matrices be recovered from a fundamental matrix? What additional information is required in order to recover 3D positions from image correspondences from a pair of uncalibrated cameras? [30%]

(TURN OVER

5 (a) A camera is equipped to detect faces in the viewfinder to assist in focussing. Describe a simple algorithm that can be used for automatic face detection. Your answer should include details of the detector and how it is trained from sample data. [50%]

(b) Describe a system to localise a mobile phone which is equipped with a camera. Give details of how the database of viewpoints should be registered and how the phone position and orientation can be recovered from a single image. [50%]

END OF PAPER