ENGINEERING TRIPOS PART IIB

Wednesday 27 April 2011 9 to 10.30

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

rc04

1 (a) Consider an algorithm to detect interest points (features of interest) in a 2-D image for use in matching.

(i) The image is first smoothed with a low-pass filter before image gradients are computed. Explain why smoothing is necessary. Which filter kernel is used in practice? [20%]

(ii) Give an expression for computing the intensity of a smoothed pixel in terms of two discrete 1-D convolutions. [20%]

(iii) Show how different resolutions of the image can be represented efficiently in an *image pyramid*. Your answer should include details of the implementation of smoothing within an octave and subsampling of the image between octaves. [20%]

(iv) How can *band-pass* filtering at different scales be implemented efficiently using the image pyramid? Show how image features such as *blob-like* shapes can be localized in both position and scale using band-pass filtering. [20%]

(b) Explain how interest points in different images can be matched. Give details of a suitable descriptor. [20%]

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

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

(a) (i) Under what assumptions is this relationship valid?

(ii) What is the algebraic and geometric significance of the parameter, s ?

(iii) How should the relationship be modified to represent points at infinity in the world? [30%]

(b) Derive the algebraic equations of a ray through the image point (u, v). [20%]

(c) Show how the projection matrix can be decomposed into a product of matrices that contain elements expressed in terms of the internal (focal length, principal point, pixels per unit length) and external (position and orientation) camera parameters. [30%]

(d) The 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). Describe desirable properties of the calibration object. [20%]

3 A 2-D *projective transformation* can be described algebraically with homogeneous coordinates:

$$\begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$$

(a) Under which two viewing conditions will correspondences in two views be described by a 2-D *projective transformation*? [20%]

(b) Consider the degrees of freedom of the 2-D projective transformation.

(i) An object appears as a square in the first image. Describe, using sketches, how it might appear in the second image. [20%]

(ii) What would happen in the limit under *weak* perspective projection? [10%]

(c) A *mosaiced panorama* of a scene is acquired by taking multiple images with a mobile phone camera which is rotated about its optical centre.

(i) Give an expression for the transformation between point correspondences in successive images. [10%]

(ii) How is the transformation estimated in practice? [20%]

(iii) What effect will rotation which is not about the optical centre have on the quality of the mosaic? Describe another common defect in mosaics using a mobile phone camera. [20%]

4 In stereo vision a point has 3-D coordinates $\mathbf{X}$ and $\mathbf{X}'$ in the left and right camera coordinate systems respectively. The rotation and translation between the two coordinate systems are represented by a matrix $\mathbf{R}$ and vector $\mathbf{T}$ with $\mathbf{X}' = \mathbf{R}\mathbf{X} + \mathbf{T}$ and the internal calibration parameter matrices of the left and right cameras are represented by matrices $\mathbf{K}$ and $\mathbf{K}'$ respectively.

(a) Consider the ray through the point in the left image with pixel coordinates (u, v). Derive an algebraic representation of the projection of this ray (also known as the *epipolar line*) in the right image. [30%]

(b) Show how to exploit the *epipolar line* for finding point correspondences in a stereo pair. What other constraints can be used to aid matching? [20%]

(c) Show how to recover the 3D positions of points which are visible in both views. [20%]

(d) Consider a single moving camera instead of a stereo pair. Show how to recover unknown camera motion and state what additional information is required. [30%]

END OF PAPER