UNIVERSITY OF CAMBRIDGE

ENGINEERING TRIPOS PART IIB

Thursday 22 April 2004 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 (a) A 512 × 512 grey scale image is smoothed by filtering with a 2D Gaussian kernel:

$$G_\sigma(x,y) = \frac{1}{2\pi\sigma^2} \exp - \left(\frac{x^2+y^2}{2\sigma^2}\right)$$

(i) What property of the Gaussian kernel is exploited when smoothing is implemented by two 1D convolutions? [10%]

(ii) Give an expression for computing the intensity of a smoothed pixel in terms of the two discrete 1D convolutions. [30%]

(b) By first showing that a discrete approximation to the second-order spatial derivative $\partial^2 I/\partial x^2$ can be obtained by convolving $I(x,y)$ with the following kernel

1	-2	1

derive a 3 × 3 kernel that can be used to compute a discrete approximation to the Laplacian of an image:

$$\nabla^2 I = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2}$$ [30%]

(c) Edge contours can be found by convolving the image with a discrete version of the Laplacian of a Gaussian and looking for zero-crossings. What are the advantages and disadvantages of this approach compared to alternative edge detection algorithms? [30%]

2 (a) An image is formed by perspective projection onto an image plane, as shown in Fig. 1. If the image plane is sampled by a CCD array, show that the relationship between a point (X_c, Y_c, Z_c) and its image (u, v) (in pixels) is given by

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} k_u f & 0 & u_0 & 0 \\ 0 & k_v f & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix}$$

and explain the geometric and physical significance of each of the terms. [30%]

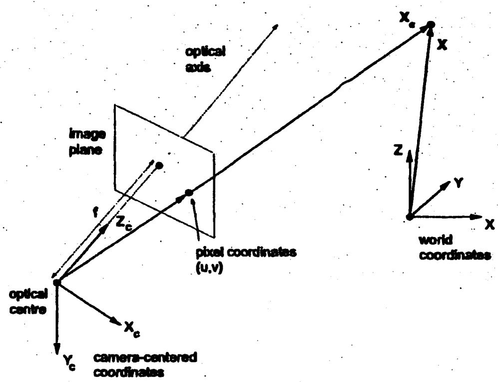

Fig. 1

(b) A *weak perspective* projection comprises an orthographic projection onto the plane $Z_c = Z_A$ followed by perspective projection onto the image plane.

(i) Derive the homogeneous relationship between a point $(X_c, Y_c, Z_c, 1)$ and its image (su_A, sv_A, s) under weak perspective projection and show that the error $(u - u_A, v - v_A)$ introduced by the weak perspective approximation is given by

$$\left((u - u_0)\frac{\Delta Z}{Z_A} \; , \; (v - v_0)\frac{\Delta Z}{Z_A} \right)$$

where $\Delta Z \equiv Z_A - Z_c$. [50%]

(ii) Under what viewing conditions is weak perspective a good camera model? What are its advantages? [20%]

(TURN OVER

3 A static scene is observed twice by the same camera, producing a pair of images with pixel correspondences (u, v) and (u', v').

(a) Show under what viewing conditions are the two images related by a 2D projective transformation:

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

[40%]

(b) How many degrees of freedom does the 2D projective transformation have? If an object appears as a square in the first image, describe, using sketches, how it might appear in the second image. Be sure to account for each degree of freedom of the 2D projective transformation. [20%]

(c) Consider the line $l_1 u + l_2 v + l_3 = 0$ in the first image. The line can be represented in homogeneous coordinates by the vector $\mathbf{l} = [\, l_1 \;\; l_2 \;\; l_3 \,]^T$. Derive an equation for the corresponding line in the second image. [20%]

(d) Repeat part (c) for the conic $au^2 + buv + cv^2 + du + ev + f = 0$. [20%]

4 (a) List four matching constraints which can be used to find point correspondences in stereo vision. Outline an algorithm which uses these constraints to match a large number of features between left and right images. [30%]

(b) A point has 3D coordinates $\mathbf{X_C}$ in the left camera's coordinate system and $R\mathbf{X_C} + \mathbf{T}$ in the right camera's coordinate system.

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

(c) Explain how the *fundamental matrix* can be estimated from point correspondences between the stereo views. What special property must the estimated matrix have? [20%]

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

(TURN OVER

5 (a) Two unbiased sensors give readings for a one-dimensional variable of z_1 and z_2 with errors which are normally distributed with stardard deviations σ_1 and σ_2 respectively. Explain the significance of the *information weighted* mean and additivity of *statistical information* for the fusion of the two sensors and give formulae.

Obtain an expression for the posterior distribution for the fused measurement from the two sensors. [50%]

(b) A video camera is used to detect pedestrians walking in front of it. Outline a template-based vision system that can be used to detect the pedestrians. Explain how the templates can be acquired from sample images, and how hypotheses can be evaluated efficiently using a suitable distance measure computation and preprocessing of the images. The recognition system should be made to work independent of lighting conditions and small changes in viewpoint. [50%]

END OF PAPER