Version RET/5

EGT3
ENGINEERING TRIPOS PART IIB

Friday 25 April 2014   2 to 3.30

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.*

Write your candidate number ***not*** *your name on the cover sheet.*

STATIONERY REQUIREMENTS
Single-sided script paper

SPECIAL REQUIREMENTS TO BE SUPPLIED FOR THIS EXAM
CUED approved calculator allowed
Engineering Data Book

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

1 (a) Many feature detection algorithms in computer vision smooth the image using a low pass filter. Give an expression for computing the smooth pixel intensities $S(x,y)$ of an image $I(x,y)$ and explain why the smoothing step is often necessary. [10%]

(b) Describe how to implement smoothing efficiently on a computer. Your answer should include the form of the low pass filter and it should also state the computational saving over a naïve implementation. [25%]

(c) Describe an algorithm to localise the position and recover the orientation of an edge in an image using spatial derivatives of the smoothed image. [25%]

(d) Describe an algorithm for detecting corners in an image using spatial derivatives of the smoothed image. [30%]

(e) Discuss whether edges or corners are more appropriate interest points for an object tracking application. [10%]

2 (a) Starting from the pin-hole camera, show that the relationship between a point measured in camera centred coordinates (X_c, Y_c, Z_c) and its corresponding image coordinates (x, y) can be expressed as follows:

$$\begin{bmatrix} sx \\ sy \\ s \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda X_c \\ \lambda Y_c \\ \lambda Z_c \\ \lambda \end{bmatrix}.$$

Your answer should explain why it is useful to express the relationship between the two coordinate systems in this way. [10%]

(b) By considering a pair of parallel planes in camera centred coordinates, derive an expression for the *horizon line* in image coordinates. [40%]

(c) A camera observes an ellipse that lies on the plane $Y_c = Y_0$. The ellipse takes the following form in camera centred coordinates:

$$\left(\frac{X_c - X_0}{a}\right)^2 + \left(\frac{Z_c - Z_0}{b}\right)^2 = 1$$

where the centre of the ellipse is (X_0, Z_0) and the shape of the ellipse is determined by the parameters (a, b).

(i) Derive the expression for the *perspective projection* of the ellipse in image coordinates. Under what conditions is this projection a circle? [40%]

(ii) Under what conditions would the *weak perspective* camera model provide a good approximation to your answer in (i). What form does the ellipse take in the image in this case? [10%]

 (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) State the conditions that must hold for the correspondences in the two views be described by a 2D *projective transformation*:

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

Your answer should include a description of how the transformation depends on the intrinsic and extrinsic camera parameters. [30%]

(b) A large number of interest points (blob-like features) are detected in each image and candidate matches from one image to the other are found by comparing their scale-invariant feature transform (SIFT) descriptors.

(i) How many correspondences are required to estimate the transformation shown in part (a)? [10%]

(ii) A candidate match returned by SIFT can sometimes be incorrect. Give reasons why these spurious matches occur. [20%]

(iii) How are consistent matches obtained in the presence of spurious matches? Your answer should include pseudo-code for the Random Sample Consensus (RANSAC) algorithm. [20%]

(iv) A user requires the RANSAC algorithm to return a good solution with probability P_0. The probability that a candidate match selected at random is spurious is known to be ρ. Estimate how many iterations, T, of the RANSAC algorithm are required to yield a good solution with probability P_0. Explain any assumptions and simplifying approximations used in your answer. [20%]

4 In a stereo vision system a point has 3D coordinates $\mathbf{X}_c$ and $\mathbf{X}'_c$ in the left and right camera centred coordinate systems respectively. The rigid body transformation between the two coordinate systems is represented using a rotation matrix $\mathbf{R}$ and a translation vector $\mathbf{T}$ so that $\mathbf{X}'_c = \mathbf{R}\mathbf{X}_c + \mathbf{T}$.

(a) Derive the *epipolar constraint* on $\mathbf{X}_c$ and $\mathbf{X}'_c$ and show how the *essential matrix* depends on $\mathbf{R}$ and $\mathbf{T}$. [40%]

(b) By considering a pair of rays from the optic centres to the image planes, derive an expression for the location of the left and right *epipoles* in camera-centred coordinates. [10%]

(c) Consider the special case when $\mathbf{R} = \mathbf{I}$ and $\mathbf{T} = (-d, d, 0)$.

(i) Give algebraic expressions for the epipolar lines. [10%]

(ii) Where are the epipoles located? [10%]

(d) The internal camera calibration matrices of the left and right cameras are represented by matrices $\mathbf{K}$ and $\mathbf{K}'$ respectively. Derive an expression for the *fundamental matrix* and explain how it can be used to find point correspondences in a stereo pair. [30%]

END OF PAPER

THIS PAGE IS BLANK