UNIVERSITY COLLEGE LONDON

University of London

EXAMINATION FOR INTERNAL STUDENTS

1998-99 and RESIT

For the following qualifications :-

M.Res

COURSE CODE	:	**MV**
TITLE OF EXAMINATION	:	**Machine Vision**
DATE	:	**5 March-1999**
TIME	:	**09.30**
TIME ALLOWED	:	**2 hours 30 minutes**

*Answer **three** questions in total, **one** from each of sections A, B and C. Each question is worth 33 marks. The total time allowed is two and a half hours.*

Ensure that answers to each section are in a separate answer book.

SECTION A

Answer one question from this section.

Question 1

(a) Describe how the geometry of image formation may be represented by means of a projection matrix in homogeneous co-ordinates and use this model to explain what is meant by (i) a Euclidean, (ii) an affine, and (iii) a projective geometric invariant. Describe the conditions under which geometric invariants of each type may be computed from an image and give one qualitative and one quantitative example of each type of invariant.

[12 marks]

(b) Geometric invariants are to be used in a computer controlled system to detect a specially designed transparent plastic wand to which surgical instruments may be attached. The wand is fitted with nine small, spherical light emitting diodes (LEDs) centred at co-ordinates: $(\pm1,0,0)$ $(0,\pm1,0)$, $(0,0,\pm1)$, $(1,1,0)$, $(0,0,4)$ and $(0,0,-3)$ imaged by a wide field of view, high resolution, 2000 x 2000 pixel, monochrome camera fixed above the operating table. When the wand is in use it typically fills approximately half the field of view of the camera, the LEDs may be detected by thresholding the images, and each covers a region of approximately 20 x 20 pixels.

Describe how you would locate the centre of each LED accurately in the images so obtained. Explain how geometric invariants may be computed from this data and used to detect the wand and to verify that it has been detected.

[9 marks]

(c) Initial trials indicate that surgeons are not satisfied with the performance of the above system. The system engineers propose to use the shape of the LEDs in the images to provide additional invariants to improve reliability of the verification process. Critically discuss the extent to which their proposals are viable and propose an alternative solution using only the same images.

[12 marks]

[TURN OVER]

Question 2

(a) Explain what is meant by a "flexible shape model" and describe how such a geometrical model may be constructed from observations of the locations of a set of n distinct landmark points x(i) described by the vector $\underline{x}^T = (x_1, y_1, x_2, y_2, x_3, y_3, ... x_n, y_n)$ in N training images, i=1,…N. Pay particular attention in your explanation to the properties required of the images and landmark points used and to what must be done to ensure that each image used in building the model is correctly registered or aligned. Explain why this alignment is necessary.

[12 marks]

(b) A biologist studying the behaviour of hamsters wishes to automate the process by using a machine vision system in which a perspex animal run is imaged from above by an inexpensive, monochrome CCD camera. The hamsters are brown in colour and easily visible from above as silhouettes against the light coloured background of the run. The shape of the silhouette depends on the animal's posture and behaviour in the run.

Describe how you would define suitable flexible shape models for this application and explain how they could be used as active shape models to classify an animal's behaviour such as "running", "sitting and grooming", "sniffing the air" and "climbing" from a long sequence of images stored off-line.

[12 marks]

(c) Describe how you might design an alternative system based on simpler methods not utilising a flexible shape model and discuss the potential advantages and disadvantages of the two systems.

[9 marks]

[CONTINUED]

SECTION B

Answer one question from this section.

Question 3

(a) A railway company wishes to train drivers by using an augmented reality system in order to simulate situations and scenarios that are too dangerous to be tested in real life. In this training system, computer vision and graphics techniques are to be used to replace the images of railway signals in video data recorded from forward looking cameras mounted on express trains during normal operation There are a small number of different types of signal of interest, all in approximately similar positions at the side of straight sections of track, some carrying two, and others three coloured lights, the latter in a number of different configurations. The stored video data is to be processed off-line to locate the signals of interest and the results written to a database linked to the video for replay in the training system. The quantity of video data available is so large that the processing must be automated.

Describe what computer vision processes you would use to detect the type of railway signals visible in a video clip and to locate them accurately in each frame for potential graphics replacement.

[11 marks]

(b) By using the perspective projection for an idealised camera of focal length, f, write down the relationship between the position $\left(X_0, Y_0, Z_0\right)^T$ of a signal at the side of the track, speed V of the train, and the expected position $\left(x(t), y(t)\right)^T$ of the signal in an image at time t. Hence, show that the position of the signal in a subsequent frame at time $t+1$ may be obtained from a recurrence relation of the form:

$$\begin{pmatrix} x(t+1) \\ y(t+1) \end{pmatrix} = \frac{\tau_0 - t}{\tau_0 - (t+1)} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \qquad (1)$$

where τ_0 is the time at which the train passes the signal. Describe the form of the trajectory of the image of the signal as the train approaches it.

[8 marks]

[Question 3 cont. over page]

[TURN OVER]

[Question 3 cont.]

(c) The time τ_0 is recorded on the video tape by an alarm set off as the train passes the signal. The video tape is then most conveniently processed in reverse order beginning with the last frame in which the signal is visible just before the train passed it. Show that equation (1) may then be used to construct a linear Kalman filter for estimation of the position of the signal in the image in *preceeding* frames as the tape is processed backwards in time. Describe how the filter and equation (1) may both be used to help locate the signal in preceding frames, and describe what further benefits you would expect to result from processing the video sequence in reverse.

[14 marks]

Question 4

(a) Explain what is meant by the terms: "optical flow", "motion constraint equation", and "aperture problem" in the computation of visual motion. Discuss the extent to which the motion constraint equation may be used to compute visual motion. Describe what you would do in practice in order to:

(i) improve the reliability and accuracy of the results obtained,

(ii) be able to use this technique to compute large flows,

and comment on the extent to which you would expect to be able to obtain a dense flow field.

[13 marks]

(b) It is proposed to use computer vision techniques in the film industry to detect when a camera is panning *or* zooming and to estimate the pan and zoom rates. Show that, when the camera is panning about the Y axis with angular velocity Ω the motion in the image is given by:

$$\frac{dx}{dt} = -\mathrm{f}\Omega\left(1 + \frac{\mathrm{x}^2}{f^2}\right) \qquad \frac{dy}{dt} = -\mathrm{f}\Omega\left(\frac{\mathrm{xy}}{f^2}\right) \quad , \qquad (1)$$

and that when it is zooming at a rate $\alpha = \frac{1}{f}\frac{df}{dt}$, the motion in the image is given by:

$$\frac{dx}{dt} = \alpha x \qquad \frac{dy}{dt} = \alpha y \quad . \qquad (2)$$

[question 4 cont. over page]

[CONTINUED]

[Question 4 (b) cont'd]

Explain the origin of the terms in these equations and the significance of the fact that in both (1) and (2) the motion depends only on the image co-ordinates and external parameters.

[Hint: you may use the fact that, when a co-ordinate frame of reference is rotated at an angular velocity $\mathbf{W}$, the co-ordinates of points in space move with velocity $\frac{d\mathbf{R}}{dt} = \mathbf{W} \times \mathbf{R}$.]

[8 marks]

(c) Discuss the extent to which the optical flow obtained as in (a) above is related to the image motion described in (b). Show, if it is assumed that they are the same, how use of the motion constraint equation in a least squares sense leads to an estimate for the zoom rate of:

$$\alpha = -\frac{\sum_{i>} \frac{\partial \mathrm{I}}{\partial \mathrm{t}} \nabla \mathrm{I}.\mathbf{x}}{\sum_{i>} (\nabla \mathrm{I}.\mathbf{x})^2} \qquad , \qquad (3)$$

where $\mathbf{x} = (x,y)^T$ and $\sum_{i>}$ indicates summation over pixels i where $|\nabla I|$ is large enough to ensure a reliable estimate of the flow. Show also that, when the camera is panning, substitution of (1) in the motion constraint equation yields:

$$\frac{\partial \mathrm{I}}{\partial \mathrm{t}} = f\Omega\left(\nabla I.\hat{\mathbf{x}} + (\nabla I.\mathbf{x})\left(\mathbf{x}.\hat{\mathbf{x}}\right) \Big/ f^2 \right) = f\Omega F(\mathbf{x}) \quad \text{, say,}$$

with $\hat{\mathbf{x}}$ a unit vector in the x direction and hence, by analogy with (3), write down a formula for the least squares estimate of $f\Omega$.

[12 marks]

[TURN OVER]

SECTION C

Answer one question from this section.

Question 5

(a) Define the 3 different types of edge detectors which have been developed to date. Describe how the accuracy of these edge detectors can be measured given an industrial inspection environment.

[8 marks]

(b) Given an image of a natural outdoor scene containing buildings and vegetation, describe which feature extraction technique should be applied to try to obtain invariant features which could be used for camera calibration and discuss what difficulties you are likely to encounter. Suggest ways in which these difficulties could be overcome.

[10 marks]

(c) Explain how edge operators can be used with Hough transforms in order to extract building roof-lines from an oblique aerial image for automated building detection. Suggest ways in which a CAD modelling system could be used to identify particular structures within the scene for model-based identification of particular buildings.

[15 marks]

Question 6

(a) Define the two different types of stereo matcher which can be applied for the automated retrieval of three-dimensional co-ordinates emphasizing the strengths and weaknesses of each.

[10 marks]

(b) Describe two different methods for quantifying the accuracy of different stereo matching techniques when applied to a scene of wooden blocks viewed at close-range by CCD cameras.

[10 marks]

(c) Discuss two different methods by means of which measurements of 3D co-ordinates can be fused with monoscopic cues derived from feature extraction operators to assist an autonomous vehicle equipped with a stereo CCD imaging system to navigate around a complex industrial environment.

[13 marks]

[END OF PAPER]