King's College London

This paper is part of an examination of the College counting towards the award of a degree. Examinations are governed by the College Regulations under the authority of the Academic Board.

Degree Programmes	MSc, MSci
Module Code	7CCSMCVI
Module Title	Computer Vision
Examination Period	January 2014 (Period 1)
Time Allowed	Three hours
Rubric	ANSWER QUESTION ONE AND ANY THREE OTHER QUESTIONS. All questions carry equal marks. If more than four questions are answered, the answer to the first four questions in exam paper order will count.
Calculators	Calculators may be used. The following models are permitted: Casio fx83 / Casio fx85.
Notes	Books, notes or other written material may not be brought into this examination

PLEASE DO NOT REMOVE THIS PAPER FROM THE EXAMINATION ROOM

DO NOT TURN OVER UNTIL INSTRUCTED TO DO SO

1. Compulsory Question

a. Give a brief definition of each of the following terms.

i. computer vision

ii. exposure

iii. epipolar line

[6 marks]

b. The array below shows the intensity values in a 3-by-3 pixel greyscale image.

$$I = \begin{bmatrix} 3 & 6 & 21 \\ 6 & 18 & 9 \\ 3 & 9 & 18 \end{bmatrix}$$

Calculate the result of convolving this image with a 3-by-3 pixel box mask (or mean filter). Pad the image with zeros where necessary to generate a 3-by-3 pixel filtered image.

[4 marks]

c. Convert image I, as defined in question 1.b, to a binary image by applying a threshold of 8.

[3 marks]

d. List four scene properties that might cause an intensity discontinuity (or edge) in an image.

[4 marks]

e. List eight cues to depth.

[8 marks]

2. a. Write down the thin lens equation, which relates the focal length of a lens to the depths of the image and object. Be sure to define each term in this equation.

[4 marks]

b. If a lens has a focal length of 50mm at what depth should the image plane be placed to bring an object 6m from the camera into focus?

[3 marks]

c. For the situation described in question 2.b, where should the image plane be placed to form a focused image if the object is moved to be 3m from the camera?

[1 marks]

d. For the pinhole camera model of image formation, what would be the effect on the size of the image of decreasing the relative distance between the camera and the object from 6m to 3m?

[2 marks]

e. Briefly compare the mechanisms used for sampling the image in a camera and in an eye.

[6 marks]

f. Briefly describe how a colour image is sampled by a camera using

i. Three CCD arrays,

ii. One CCD array.

[4 marks]

QUESTION 2 CONTINUES ON NEXT PAGE

g. In order to locate red regions of the image it has been decided to model the response properties of a red on/green off (R+/G-) colour-opponent retinal ganglion cell. A simple difference of Gaussians model is to be used. Design a MATLAB function $R = f(I)$ that will calculate the response of a R+/G- cell at all locations in the image I. Assume I is a colour image in RGB format and that you have access to a function $g = gaussian(sigma, d)$ that returns a d-by-d pixel Gaussian filter with standard deviation sigma. Use a Gaussian with standard deviation 2 for the centre, and a Gaussian with standard deviation 3 for the surround.

[5 marks]

3. a. Convolution masks can be used to provide a finite difference approximation to first and second order directional derivatives. Write down the masks that approximate the following directional derivatives:

i. $-\frac{\delta}{\delta x}$

ii. $-\frac{\delta^2}{\delta x^2}$

iii. $-\frac{\delta^2}{\delta x^2} - \frac{\delta^2}{\delta y^2}$

[6 marks]

b. For edge detection, a difference mask is usually "combined" with a smoothing mask.

i. How are these masks "combined"?

ii. Why is this advantageous for edge detection?

iii. Name the mathematical function that is usually used to define the smoothing mask.

[7 marks]

c. An image, I, is to be convolved with a 1-dimensional mask, h_1, and the result is to be convolved with another 1-dimensional mask, h_2. Calculate the single, 2-dimensional, mask H that would produce the same result, when: $h_1 = [-1 \quad 1 \quad -1]$ and $h_2 = [1 \quad 2 \quad 1]^T$.

[4 marks]

d. What is the advantage of sequentially performing two convolutions with 1D masks, rather than a single convolution with a 2D mask, to obtain the same result?

[2 marks]

QUESTION 3 CONTINUES ON NEXT PAGE

e. Describe briefly in words, or using pseudo-code, each step required to create:

i. a Gaussian image pyramid,

ii. a Laplacian image pyramid.

[6 marks]

4. **a.** Briefly describe what is meant by "grouping and segmentation" in mid-level vision.

[4 marks]

b. One method of image segmentation is the split and merge algorithm. Write pseudo-code for the split and merge algorithm.

[7 marks]

c. The array below shows feature vectors for each pixel in a 2-by-2 pixel image.

$$\begin{bmatrix} (10,15,5) & (15,15,15) \\ (5,15,10) & (20,10,15) \end{bmatrix}$$

Apply the split and merging algorithm to assign each pixel to a region. Assume that (1) the method used to assess similarity is the sum of absolute differences (SAD), (2) the criteria for deciding if regions are similar is that the SAD is less than 12, (3) regions have horizontal, vertical and diagonal neighbours.

[6 marks]

d. In the Hough Transform, a point with coordinates (x,y) in the image-plane is mapped onto a curve in the $(\theta,\ r)$-parameter space of the accumulator array. Write down the equation for r in terms of x, y, and θ.

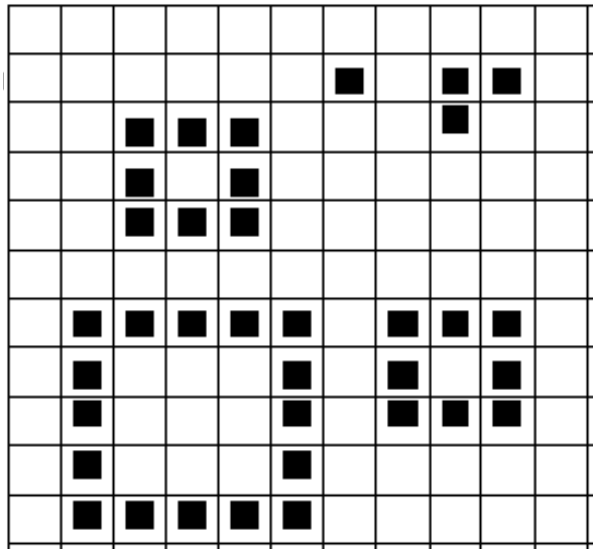
[2 marks]

QUESTION 4 CONTINUES ON NEXT PAGE

e. Below is an 11x11 pixel binary image in which the edge pixels are indicated by the black squares. If we apply the Hough transform to this image, using the following values for θ: [0,45,90,135],

i. what would be the maximum value in the accumulator array?

ii. what would be the (θ, r) coordinates of this location?



[4 marks]

f. Briefly describe the two terms that form the energy function that an active contour (or "snake") attempts to minimise.

[2 marks]

5. Two images (A and B) of the same scene are taken from different viewpoints. Interest points are located in each image and a two element descriptor is calculated for each interest point. The image coordinates and feature vectors for each interest point are given in the table below:

Image A			Image B		
Point	Coordinates	Feature Values	Point	Coordinates	Feature Values
A1	(277, 168)	(10, 4)	B1	(111, 394)	(3, 7)
A2	(193, 290)	(3, 8)	B2	(125, 186)	(1, 1)
A3	(205, 87)	(0, 2)	B3	(145, 128)	(5, 7)
A4	(224, 28)	(6, 9)	B4	(90, 270)	(8, 0)

a. For each interest point in image A, find the best matching interest point in image B assuming that similarity is measured using the sum of absolute differences (SAD).

[8 marks]

b. Write pseudo-code for the RANSAC algorithm.

[6 marks]

c. Apply the RANSAC algorithm to find the true correspondence between the two images. Assume (1) that the images are related by a pure translation in the x-y plane, (2) that the threshold for comparing the model's prediction with the data is 20 pixels, (3) 3 trials are performed and these samples are chosen in the order A1, A2, A3 rather than being randomly chosen.

[11 marks]

6. A simple object recognition system encodes objects using 4-element feature vectors. Four objects from two different classes (A and B) are encoded as follows:

Object	Class	Feature Vector
1	A	(2, 3, 0, 1)
2	A	(3, 3, 1, 3)
3	B	(1, 2, 3, 2)
4	B	(0, 1, 3, 0)

A new object, of unknown class, has a feature vector (2, 2, 2, 2). Using Euclidean distance as the similarity measure, determine the classification of the new object using:

a. a nearest mean classifier,

[5 marks]

b. a nearest neighbour classifier,

[5 marks]

c. a k-nearest neighbour classifier, with k=3.

[4 marks]

QUESTION 6 CONTINUES ON NEXT PAGE

d. In a simple bag-of-words object recognition system classes are represented by histograms showing the number of occurrences of 4 "codewords". The number of occurrences of the codewords in two training images are given below:
ClassA = (2.5, 3.5, 0.5, 2)
ClassB = (0.5, 1.5, 3.5, 1)
A new image is encoded as follows:
New = (1, 2, 2, 1)
Determine the training image that best matches the new image by finding the cosine of the angle between the codeword vectors.

[5 marks]

e. Describe briefly in words, or using pseudo-code, each step required to train the Implicit Shape Model (ISM) for object recognition.

[6 marks]