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 2017 (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. down-sampling

iii. transduction

[6 marks]

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

$$I = \begin{bmatrix} 5 & 11 & 21 \\ 1 & 6 & 9 \\ 3 & 9 & 8 \end{bmatrix}$$

Calculate the 2-by-2 pixel image that would result if this image was convolved with a 2-by-2 pixel box mask (or mean filter).

[5 marks]

c. Define a MATLAB function B = binarize(I, t) that will convert an image, I, to a binary image, B, by applying a threshold of t.

[4 marks]

d. The binary image that results from applying a threshold of 8 to image I, as defined in question 1.b, is:

$$B = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Write down the new binary image that results from performing dilation on B, assuming each pixel has 4 neighbours (horizontal and vertical).

[4 marks]

QUESTION 1 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]

2. The arrays below show the values of the red, green and blue pixels in a Bayer masked image.

$R =$

0.2		0.3	
0.4		0.3	

$G =$

	0.7		0.6
0.1		0.6	
	0.8		0.9
0.4		0.1	

$B =$

	0.2		0.3
	0.4		0.6

a. Calculate the RGB values for the pixel at coordinates (2,2) in the demosaiced image using bi-linear interpolation.

[4 marks]

b. Calculate the RGB values for the pixel at coordinates (2,2) in the demosaiced image using edge-directed interpolation. Assume that pixels at coordinates (1,1), (1,3), and (3,1) all have green values of 0.5.

[5 marks]

c. The pinhole camera model of image formation relates the coordinates of a 3D point P with coordinates (X,Y,Z) relative to the camera reference frame to the coordinates of its image p=(x,y,z) via the following equation:

$$(x, y, z) = \left(\frac{fX}{Z}, \frac{fY}{Z}, f\right)$$

Given that two identical cameras are mounted so as to have coplanar image planes, such that one camera (the right one) is displaced distance B along the x-axis of the other (the left one), derive an equation for the depth of a 3D point P visible in both cameras.

[6 marks]

d. Describe five constraints that might be applied in trying to determine the locations of the points in the left and right image that correspond to the same point in 3D space.

[10 marks]

3. a. The array below shows a Laplacian mask, L.

$$L = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

Write down a mathematical expression describing the effect of convolving an image I with L.

[2 marks]

b. Write down the mask that approximates the following directional derivative: $-\frac{\delta^2}{\delta y^2}$

[2 marks]

c. Use the following formula for a 2D Gaussian to calculate a 3-by-3 pixel numerical approximation to a Gaussian with standard deviation of 0.5 pixels, rounding values to two decimal places.

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

[3 marks]

d. Using the Laplacian defined in question 3.a and the Gaussian calculated in answer to question 3.c calculate a 3-by-3 pixel Laplacian of Gaussian (or LoG) mask.

[5 marks]

e. What advantage does a LoG mask have over a Laplacian for edge detection?

[2 marks]

QUESTION 3 CONTINUES ON NEXT PAGE

f. An alternative method of performing edge detection is to simulate the responses of orientation tuned neurons in primary visual cortex (V1). Name the mathematical function that is used to model the receptive fields of cortical simple cells, and write down the equation for this function.

[4 marks]

g. Describe how the responses of V1 simple cells can be modelled using convolution.

[4 marks]

h. Describe how the responses of V1 complex cells could be modelled using convolution.

[3 marks]

4. a. Briefly describe the role of mid-level vision.

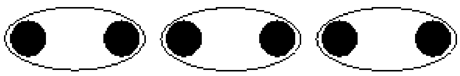
[1 mark]

b. Below are four simple images. For each image identify the "Gestalt Law" that accounts for the observed grouping of the image elements.

i. 

ii. 

iii. 

iv. 

[8 marks]

c. One method of image segmentation is region growing. Write pseudo-code for the region growing algorithm.

[5 marks]

QUESTION 4 CONTINUES ON NEXT PAGE

d. The array below shows 3 element long feature vectors for each pixel in a 2-by-3 pixel image, I_F.

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

Apply the region growing algorithm to image I_F in order to assign each pixel to a region. Assume that (i) the method used to assess similarity is the sum of absolute differences (SAD), (ii) the criterion for deciding if elements are similar is that the SAD is less than 12, (iii) the seed pixel is the top-left corner, (iv) pixels have horizontal, vertical and diagonal neighbours.

[6 marks]

e. An alternative method of image segmentation is Normalised Cuts. Write pseudo-code for the Normalised Cuts algorithm for image segmentation.

[5 marks]

5. **a.** Briefly describe the methodology used in correlation-based and feature-based methods of solving the stereo correspondence problem.

[4 marks]

b. The two arrays below show the intensity values for each pixel in a stereo pair of 5-by-3 pixel images.

$I_{left} =$

40	60	40	20	50
10	50	80	80	30
70	10	70	60	90

$I_{right} =$

20	70	70	20	50
30	20	50	10	50
50	70	40	80	70

Calculate the similarity of the pixel at coordinates (2,2) in I_{left}, to locations in I_{right}, and hence, calculate the disparity at that point. Assume that (i) a 3-by-3 pixel window is used, (ii) similarity is measured using the Sum of Absolute Differences (SAD), (iii) that corresponding points will occur along corresponding scan lines (*i.e.*, on the same row in both images), (iv) disparity is calculated as the translation from right to left.

[4 marks]

c. Write down the equation that defines the measure R used by the Harris corner detector.

[4 marks]

QUESTION 5 CONTINUES ON NEXT PAGE

d. The two arrays below show the derivatives of the image intensities in the x and y directions for a small 3-by-3 pixel image.

$$I_x = \begin{array}{|c|c|c|}\hline 1 & 0 & -2 \\ \hline 0 & -1 & 0 \\ \hline 0 & -3 & 1 \\ \hline \end{array} \qquad I_y = \begin{array}{|c|c|c|}\hline 2 & 1 & -3 \\ \hline 0 & 0 & 0 \\ \hline 0 & -2 & 1 \\ \hline \end{array}$$

Calculate the measure R of the Harris corner detector at each pixel, assuming (i) a value of k=0.05, (ii) that products of derivatives are summed over an equally weighted, 3-by-3 pixel, window around each pixel, padding with zeros when necessary.

[4 marks]

e. For the Harris corner detector, describe what type of image feature will give rise to the following values of R.

(i) R $\approx$ 0

(ii) R $<$ 0

(iii) R $>$ 0.

[6 marks]

f. Briefly explain what role the Harris corner detector might play in solving the stereo correspondence problem.

[3 marks]

6. a. 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, 2, 0, 1)
2	A	(3, 3, 1, 1)
3	B	(1, 2, 3, 2)
4	B	(2, 1, 3, 0)

A new object, of unknown class, has a feature vector (1, 1, 1, 1). Using Euclidean distance as the similarity measure, determine the classification of the new object using:

i. a nearest mean classifier,

[5 marks]

ii. a nearest neighbour classifier,

[5 marks]

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

[4 marks]

b. Describe briefly in words, or using pseudo-code, each step required to train a bag-of-words object recognition system, and then apply it to classifying a new image.

[6 marks]

QUESTION 6 CONTINUES ON NEXT PAGE

c. 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, 3.5, 0.5, 2)

ClassB = (0.5, 0.75, 3.5, 1)

A new image is encoded as follows:

New = (2, 1, 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]