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 2018 (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.

ANSWER EACH QUESTION ON A NEW PAGE OF YOUR ANSWER BOOK AND WRITE ITS NUMBER IN THE SPACE PROVIDED.

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

TURN OVER WHEN INSTRUCTED

1. Compulsory Question

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

   i. image processing

   ii. mid-level vision

   iii. horopter

[6 marks]

b. Below are shown a convolution mask, H and an image I.

$$H = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad I = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 2 & 0 & 2 & 0 \\ 2 & 2 & 1 & 1 \\ 2 & 2 & 0 & 0 \end{bmatrix}$$

What is the result of the convolution of mask H with image I? The result should be an image that is the same size as I.

[5 marks]

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

[6 marks]

d. The RGB channels for a 3-by-3 pixel colour image are shown below.

$$R = \begin{bmatrix} 140 & 140 & 150 \\ 150 & 140 & 150 \\ 0 & 10 & 20 \end{bmatrix} \quad G = \begin{bmatrix} 160 & 170 & 255 \\ 170 & 160 & 150 \\ 0 & 0 & 10 \end{bmatrix} \quad B = \begin{bmatrix} 200 & 190 & 180 \\ 210 & 200 & 200 \\ 255 & 200 & 210 \end{bmatrix}$$

   i. What is the colour of the pixel at coordinates (1,3)?

[2 marks]

   ii. What is the colour of the surface in the world shown at coordinates (1,3) in the image? Give reasons for your answer.

[2 marks]

e. Briefly explain the differences between "viewer-centred" and "object-centred" approaches to object recognition?

[4 marks]

2.

a. Draw a cross-sectional diagram showing how a lens forms an image (P') of a point (P). Ensure that you label the optical centre (O), the focal point (F), and the coordinates of the world point (y,z) and the image point (y',z').

[5 marks]

b. Derive the thin lens equation, which relates the focal length of a lens to the depths of the image and object.

[6 marks]

c. If a lens has a focal length of 30mm at what depth should the image plane be placed to bring an object 6m from the camera into focus? Give your answer in millimetres to two decimal places.

[3 marks]

d. Briefly compare the mechanisms used for focusing a camera and an eye.

[4 marks]

e. Derive the equation for the pinhole camera model of image formation relating the coordinates of a 3D point P(y,z) to the coordinates of its image P'(y',f'). Note that in the pinhole camera model, the image plane is located at distance f' from the optical centre.

[4 marks]

f. Use the pinhole camera model to calculate the coordinates (x',y') of the image of a point in 3D space which has coordinates (0.4,0.5,6) measured, in metres, relative to the optical centre of the camera. Assume that the lens has a focal length of 30mm.

[3 marks]

3.

a. To locate intensity discontinuities in an image a difference mask is usually "combined" with a smoothing mask.

　i. How are these masks "combined"?

　ii. Why is this advantageous for edge detection?

[5 marks]

b. Use the following formula for a 2D Gaussian to calculate a 3-by-3 pixel numerical approximation to a Gaussian with standard deviation of 0.46 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]

c. 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. $-\dfrac{\delta}{\delta x}$

　ii. $-\dfrac{\delta^2}{\delta y^2}$

[4 marks]

d. Combine the Gaussian smoothing mask calculated in answer to question 3.b with the difference mask given in answer to question 3.c to produce a 4-by-3 pixel x-derivative of Gaussian mask.

[3 marks]

e. In order to locate intensity discontinuities in both the x and y directions an image can be convolved with an x-derivative of Gaussian mask and a y-derivative of Gaussian mask. Assuming the result of these two convolutions are two images Ix and Iy of equal size, a single image showing intensity discontinuities in all direction can be calculated by taking the L2-norm of corresponding pixels in these two images. Write a MATLAB function $Ixy = l2norm(Ix, Iy)$ that will combine Ix and Iy using the L2-norm.
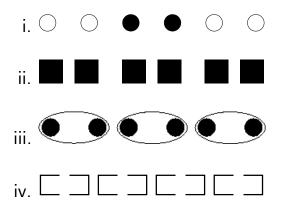
[4 marks]

f. Derivative of Gaussian masks (in the x and y directions) are used by the Canny edge detector. Describe briefly in words, or using pseudo-code, each step performed by the Canny edge detection algorithm.

[6 marks]

4.

a. 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]

b. One method of grouping image elements is clustering. Write pseudo-code for the agglomerative hierarchical clustering algorithm.

[5 marks]

c. The array below shows feature vectors for each pixel in a 2-by-3 pixel image.

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

Apply the agglomerative hierarchical clustering algorithm to assign pixels into three regions. Assume that (1) the method used to assess similarity is the sum of absolute differences (SAD), and (2) centroid clustering is used to calculate the distance between clusters.

[8 marks]

d. In question 4.c SAD was used to assess the similarity between clusters. It is also possible to perform clustering using a number of other standard metrics. If **a** and **b** represent the feature vectors associated with two clusters, write down the formulae for comparing these two vectors using:

  i. sum of squared differences

  ii. correlation coefficient

[4 marks]

5.

   a. Define what is meant by the "aperture problem" and suggest how this problem can be overcome.

   [4 marks]

   b. Two frames in a video sequence were taken at times t and t+0.04s. The point (110,50,t) in the first image has been found to correspond to the point (95,50,t+0.04) in the second image. Given that the camera is moving at $0.5ms^{-1}$ along the camera x-axis, the focal length of the camera is 35mm, and the pixel size of the camera is 0.1mm/pixel, calculate the depth of the identified scene point.

   [4 marks]

   c. Two frames in a video sequence were taken at times t and t+0.04s. The point (140,100,t) in the first image has been found to correspond to the point (145,100,t+0.04) in the second image. Given that the camera is moving at $0.5ms^{-1}$ along the optical axis of the camera (*i.e.*, the z-axis), and the centre of the image is at pixel coordinates (100,100), calculate the depth of the identified scene point.

   [4 marks]

   d. Give an equation for the time-to-collision of a camera and a scene point which does not require the recovery of the depth of the scene point. Using this equation, calculate the time-to-collision of the camera and the scene point in question 5.c, assuming the camera velocity remains constant.

   [3 marks]

e. In order to calculate depth or time-to-collision using video, it is necessary to determine which image locations in two video frames correspond to the same location in the world. Briefly describe two constraints typically applied to solving this video correspondence problem, and note circumstances in which each constraint fails.

[6 marks]

f. There are many other cues to depth that can be obtained from a *single* image. Name any four of these monocular cues to depth.

[4 marks]

6.

    a. What are "geons", and what is their hypothesised role in biological object recognition?

            [4 marks]

    b. Below are shown three binary templates T_1, T_2 and T_3 together with a patch I of a binary image.

$$T_1 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \qquad T_2 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \qquad T_3 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix},$$

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

Determine which template best matches the image patch using the following similarity measures:

    i. cross-correlation,

            [3 marks]

    ii. normalised cross-correlation,

            [3 marks]

    iii. sum of absolute differences.

            [3 marks]

c. Below are an edge template T and a binary image I which is the result of pre-processing an image to locate the edges.

$$T = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \qquad I = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

Calculate the result of performing edge matching on the image, and hence, suggest the location of the object depicted in the edge template assuming that there is exactly one such object in the image. Calculate the distance between the template and the image as the average of the minimum distances between points on the edge template (T) and points on the edge image (I). Only consider those locations where the template fits entirely within the image.

[5 marks]

d. A production line produces two objects (A and B) which are sorted into separate bins using a computer vision system controlling a robot arm. The two objects have distinct shapes from most viewpoints. However, when object A lies at orientation 1 it is indistinguishable from object B lying at orientation 2.

It is known that the production line produces four times as many of object A than object B. It is also known that the probability of object A lying at orientation 1 is 0.02, while the probability of object B lying at orientation 2 is 0.04.

Use Bayes' theorem to determine the bin into which the robot should sort an object which could be either object A at orientation 1 or object at orientation 2 in order to minimise the number of errors.

[7 marks]

FINAL PAGE