

— SOLUTIONS —

# 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**

© 2018 King's College London

# — SOLUTIONS —

January 2018

7CCSMCVI

## 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]

### Answer

- i) image processing = signal processing applied to an image, with another image as the resulting output
- ii) mid-level vision = a range of processes that group together related image elements, and to segment them from all other image elements
- iii) horopter = an imaginary surface on which all points have zero disparity

### Marking scheme

2 marks for each correct definition.

# — SOLUTIONS —

January 2018

7CCSMCVI

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

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

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]

Answer

$$\begin{bmatrix} 0 & 3 & 1 & 3 \\ 2 & 2 & 5 & 2 \\ 2 & 6 & 3 & 3 \\ 2 & 4 & 2 & 1 \end{bmatrix}$$

Marking scheme

5 marks. Partial marks are possible for partially correct answers.

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

[6 marks]

Answer

Camera:

- Has sensing elements sensitive to three wavelengths (RGB).
- Sensing elements occur in a fixed ratio across the whole image plane.
- The sampling density is uniform across the whole image plane.

Eye:

- Has sensing elements sensitive to four wavelengths (RGBW).
- Sensing elements occur in a variable ratios across the image plane (cone density highest at fovea, rod density highest outside fovea).

# — SOLUTIONS —

January 2018

7CCSMCVI

- The sampling density is non-uniform across the image plane (density is highest at the fovea).

## Marking scheme

3 marks for each part.

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]

## Answer

Blue

## Marking scheme

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]

## Answer

Unknown. The RGB values of the image will depend both on the properties of the surface (including its colour) and the properties of the light it is reflecting. Without knowledge of the latter, we can't know the former.

## Marking scheme

2 marks

# — SOLUTIONS —

January 2018

7CCSMCVI

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

[4 marks]

## Answer

In the viewer-centred approach, the 3D object is modelled as a set of 2D images, showing different views of the object.

In the object-centred approach, a single 3D model is used to describe the object.

## Marking scheme

4 marks.

# — SOLUTIONS —

January 2018

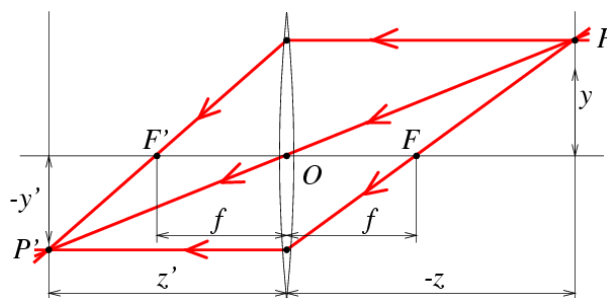
7CCSMCVI

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]

Answer



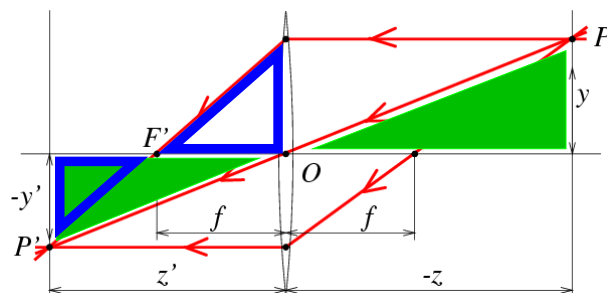
Marking scheme

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]

Answer



From similar triangles:

$$\frac{y'}{z'} = \frac{y}{z} \quad \implies y' = \frac{z'y}{z}$$

and

$$\frac{y'}{z'-f} = \frac{y}{f} \quad \implies y' = \frac{(z'-f)y}{f}$$

# — SOLUTIONS —

January 2018

7CCSMCVI

equating for  $y'$ :  $\frac{z'y}{z} = \frac{(z'-f)y}{f}$   
 $y$  cancels, hence:  $\frac{z'}{z} = \frac{(z'-f)}{f} = \frac{z'}{f} - 1$   
 dividing both sides by  $z'$ :  $\frac{1}{z} = \frac{1}{f} - \frac{1}{z'} \implies \frac{1}{z} + \frac{1}{z'} = \frac{1}{f}$

Marking scheme

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]

Answer

$$\frac{1}{f} = \frac{1}{\|z\|} + \frac{1}{\|z'\|} \implies \frac{1}{\|z'\|} = \frac{1}{f} - \frac{1}{\|z\|}$$

For object at 6m:

$$\frac{1}{\|z'\|} = \frac{1}{30} - \frac{1}{6000} \implies z' = 30.15mm$$

Marking scheme

3 marks.

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

[4 marks]

Answer

A camera lens has a fixed shape, and hence, a fixed focal length. Focusing is achieved by moving the lens to change the distance to the image plane.

An eye lens has an adjustable shape, and hence, a variable focal length. Whereas the distance between the lens and the image plane (the retina) is fixed. Focusing is achieved by changing the focal length of the lens.

Marking scheme

# — SOLUTIONS —

January 2018

7CCSMCVI

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]

Answer

From similar triangles (as before):

$$\frac{y'}{z'} = \frac{y}{z} \quad \implies \quad y' = \frac{z'y}{z}$$

substituting  $z' = f'$  gives:  $y' = \frac{f'y}{z}$

Marking scheme

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]

Answer

$$\begin{aligned} x' &= \frac{f'x}{z} & \implies & x' = \frac{30 * 400}{6000} = 2mm \\ y' &= \frac{f'y}{z} & \implies & y' = \frac{30 * 500}{6000} = 2.5mm \end{aligned}$$

Marking scheme

1 mark each, plus 1 additional mark for getting the units correct.



# — SOLUTIONS —

January 2018

7CCSMCVI

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]

Answer

- i) Masks are combined using convolution.

Marking scheme

2 marks

- ii) A difference mask is sensitive to noise as well as other intensity-level discontinuities.

A smoothing mask suppresses noise.

The combination of the two produces a mask that is sensitive to intensity-level discontinuities that are image features rather than noise.

Marking scheme

3 marks

# — SOLUTIONS —

January 2018

7CCSMCVI

- 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]

Answer

Gaussian mask

$$= \begin{bmatrix} 0.01 & 0.07 & 0.01 \\ 0.07 & 0.75 & 0.07 \\ 0.01 & 0.07 & 0.01 \end{bmatrix}$$

Marking scheme

1 mark each for the 3 different values.

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

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

[4 marks]

Answer

i)  $-\frac{\delta}{\delta x} \approx \begin{bmatrix} -1 & 1 \end{bmatrix}$

ii)  $-\frac{\delta^2}{\delta x^2} \approx \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}$

Marking scheme

2 marks for each correct definition.

# — SOLUTIONS —

January 2018

7CCSMCVI

- 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]

Answer

To calculate the x-derivative of Gaussian mask:

$$G_x = G * [-1, 1] = \begin{bmatrix} -0.01 & -0.06 & 0.06 & 0.01 \\ -0.07 & -0.68 & 0.68 & 0.07 \\ -0.01 & -0.06 & 0.06 & 0.01 \end{bmatrix}$$

Marking scheme

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  $I_x$  and  $I_y$  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  $I_{xy} = l2norm(I_x, I_y)$  that will combine  $I_x$  and  $I_y$  using the L2-norm.

[4 marks]

Answer

```
function Ixy = l2norm(Ix,Iy)
Ixy=sqrt(Ix.^2+Iy.^2);
```

Marking scheme

4 marks.

# — SOLUTIONS —

January 2018

7CCSMCVI

- 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]

## Answer

1. convolve the image with each derivative of Gaussian mask, to generate  $I_x$  and  $I_y$ .
2. calculate the magnitude and direction of the intensity gradient ( $M = \sqrt{I_x^2 + I_y^2}$ ,  $D = \tan^{-1} \left( \frac{I_y}{I_x} \right)$ ).
3. perform non-maximum suppression (thin multi-pixel wide edges down to a single pixel by setting M to zero for all pixels that have a neighbour, perpendicular to the direction of the edge, with a higher magnitude).
4. perform hysteresis thresholding (pixels above high thresholds set to one, pixels below low threshold set to zero, pixels with values between low and high thresholds set to one if they are connected to a pixel with a magnitude over the high threshold, and set to zero otherwise).

## Marking scheme

6 marks

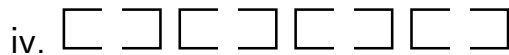
# — SOLUTIONS —

January 2018

7CCSMCVI

4.

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



[8 marks]

Answer

- i) similarity
- ii) proximity
- iii) common region
- iv) closure

Marking scheme

2 marks for each correct definition.

# — SOLUTIONS —

January 2018

7CCSMCVI

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

[5 marks]

Answer

1. Assign each data point to a unique cluster
2. Compute the similarity between each pair of clusters (store this in a proximity matrix)
3. Repeat
  4. Merge the two closest clusters
  5. Update the proximity matrix
6. Until only a single cluster remains (or an earlier stopping criterion has been met)

Marking scheme

5 marks.

# — SOLUTIONS —

January 2018

7CCSMCVI

- 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]

Answer

Each point is a separate cluster initially. Compute the distance between each pair of clusters:

$$\begin{bmatrix} & c1 & c2 & c3 & c4 & c5 & c6 \\ c1 : (10, 15, 5) & - & - & - & - & - & - \\ c2 : (15, 15, 15) & 15 & - & - & - & - & - \\ c3 : (5, 15, 10) & 10 & 15 & - & - & - & - \\ c4 : (20, 10, 15) & 25 & 10 & 25 & - & - & - \\ c5 : (10, 20, 5) & 5 & 20 & 15 & 30 & - & - \\ c6 : (10, 15, 5) & 0 & 15 & 10 & 25 & 5 & - \end{bmatrix}$$

Merge closest clusters (c1 and c6), and update the proximity matrix:

$$\begin{bmatrix} & c1 + c6 & c2 & c3 & c4 & c5 & c6 \\ c1 + c6 : (10, 15, 5) & - & - & - & - & - & - \\ c2 : (15, 15, 15) & 15 & - & - & - & - & - \\ c3 : (5, 15, 10) & 10 & 15 & - & - & - & - \\ c4 : (20, 10, 15) & 25 & 10 & 25 & - & - & - \\ c5 : (10, 20, 5) & 5 & 20 & 15 & 30 & - & - \end{bmatrix}$$

Merge closest clusters (c1+c6 and c5), and update the proximity matrix:

# — SOLUTIONS —

January 2018

7CCSMCVI

	$c1 + c6 + c5$	$c2$	$c3$	$c4$	$c5$	$c6$
$c1 + c6 + c5 : (10, 16.67, 5)$	—	—	—	—	—	—
$c2 : (15, 15, 15)$	16.67	—	—	—	—	—
$c3 : (5, 15, 10)$	11.67	15	—	—	—	—
$c4 : (20, 10, 15)$	26.67	10	25	—	—	—

Merge closest clusters ( $c2$  and  $c4$ ). We have three regions, so stop.  
Regions are:  $c1+c6+c5$ ,  $c2+c4$ , and  $c3$ .

## Marking scheme

8 marks. 3 marks for proximity matrices, 2 for calculating distances correctly, 2 for merging closest clusters correctly, 1 mark for knowing when to stop.

- 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]

## Answer

sum of squared difference =  $\sum_i (a_i - b_i)^2$

correlation coefficient =  $\frac{\sum_i (a_i - \bar{a})(b_i - \bar{b})}{\sqrt{\sum_i (a_i - \bar{a})^2} \sqrt{\sum_i (b_i - \bar{b})^2}}$

## Marking scheme

2 marks for each correct definition.



# — SOLUTIONS —

January 2018

7CCSMCVI

5.

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

[4 marks]

Answer

The aperture problem refers to the fact that the direction of motion of a small image patch can be ambiguous.

Particularly, for an edge information is only available about the motion perpendicular to the edge, while no information is available about the component of motion parallel to the edge.

Overcoming the aperture problem might be achieved by

1. integrating information from many local motion detectors / image patches, or
2. by giving preference to image locations where image structure provides unambiguous information about optic flow (e.g. corners).

Marking scheme

2 marks for describing problem, 1 mark each for possible solutions.

- 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]

Answer The depth is given by:  $Z = -\frac{fV_x}{\dot{x}}$ .

The velocity of the image point is  $\frac{95-110}{0.04} = -375$  pixels/s.

Given the pixel size this is equivalent to  $0.0001 \times -375 = -0.0375$  m/s.

# — SOLUTIONS —

January 2018

7CCSMCVI

Hence, the depth is  $Z = -\frac{0.035 \times 0.5}{-0.0375} = 0.467\text{m}$ .

Marking scheme

2 marks for equation, 2 marks for correct application.

- c. Two frames in a video sequence were taken at times  $t$  and  $t+0.04\text{s}$ . 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  $0.5\text{ms}^{-1}$  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]

Answer The depth is given by:  $Z_2 = \frac{x_1 V_z}{\dot{x}}$ .

The coordinates of the points with respect to the centre of the image are:  $(40,0,t)$  and  $(45,0,t+0.1)$ .

The velocity of the image point is  $\frac{45-40}{0.04} = 125 \text{ pixels/s}$ .

Hence, the depth is  $Z_2 = \frac{40 \times 0.5}{125} = 0.16\text{m}$ .

Marking scheme

2 marks for equation, 2 marks for correct application.

- 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]

Answer time-to-collision  $= \frac{x_1}{\dot{x}}$ .

Hence, time-to-collision  $= \frac{40}{125} = 0.32\text{s}$ .

Marking scheme

2 marks for equation, 1 mark for correct application.

# — SOLUTIONS —

January 2018

7CCSMCVI

- 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]

Answer

- Spatial coherence (assume neighbouring points have similar optical flow). Fails at discontinuities between surfaces at different depths.
- Small motion (assume optical flow vectors have small magnitude). Fails if relative motion is fast or frame rate is slow.

Marking scheme

2 for each, plus 1 each for failure cases

- 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]

Answer

Any four from:

- Interposition/Occlusion
- Size familiarity
- Texture gradients
- Linear perspective
- Aerial perspective
- Shading

Marking scheme

1 mark for each.

# — SOLUTIONS —

January 2018

7CCSMCVI

6.

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

[4 marks]

Answer

Geons are geometrical icons, simple volumes such as cubes, spheres, cylinders, and wedges. There is a hypothesis that object recognition in biological systems is based on the ability to recognise a small set of shapes — geons — from which more complex objects are built up. The visual system breaks down an object into geons and compares this arrangement of geons with arrangements of geons of known objects.

Marking scheme

2 marks for knowing what geons are. 2 marks for explaining how they are used.

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

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

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

- i. cross-correlation,

[3 marks]

# — SOLUTIONS —

January 2018

7CCSMCVI

ii. normalised cross-correlation,

[3 marks]

iii. sum of absolute differences.

[3 marks]

## Answer

i) Cross-correlation.

Similarity =

$$\sum_{i,j} T(i,j)I(i,j)$$

For  $T_1$  Similarity = 8

For  $T_2$  Similarity = 7

For  $T_3$  Similarity = 7

Both  $T_1$  is the best match.

## Marking scheme

2 for correct method, 1 for correct match

ii) Normalised cross-correlation.

Similarity =

$$\frac{\sum_{i,j} T(i,j)I(i,j)}{\sqrt{\sum_{i,j} T(i,j)^2} \sqrt{\sum_{i,j} I(i,j)^2}}$$

For  $T_1$  Similarity =  $\frac{8}{\sqrt{9 \times 8}} = 0.943$

For  $T_2$  Similarity =  $\frac{7}{\sqrt{8 \times 8}} = 0.875$

For  $T_3$  Similarity =  $\frac{7}{\sqrt{7 \times 8}} = 0.935$

$T_1$  is the best match.

## Marking scheme

2 for correct method, 1 for correct match

iii) Sum of absolute differences

Distance =

$$\sum_{i,j} \|T(i,j) - I(i,j)\|$$

# — SOLUTIONS —

January 2018

7CCSMCVI

For  $T_1$  Distance =  $8(1 - 1) + 1(1 - 0) = 1$

For  $T_2$  Distance =  $7(1 - 1) + 1(1 - 0) + 1(0 - 1) = 2$

For  $T_3$  Distance =  $7(1 - 1) + 1(1 - 0) + 1(0 - 0) = 1$

$T_1$  and  $T_3$  match equally well.

Marking scheme

2 for correct method, 1 for correct match

- 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.

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

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]

Answer

At pixel (2,2) Distance =

$$\frac{1}{8} [\sqrt{2} + 1 + 0 + 1 + 0 + \sqrt{2} + 1 + 1] = 0.855$$

At pixel (3,2) Distance =

$$\frac{1}{8} [1 + 0 + 1 + 0 + 0 + 1 + 1 + 0] = 0.5$$

At pixel (2,3) Distance =

$$\frac{1}{8} [1 + 0 + 0 + \sqrt{2} + 1 + 1 + 0 + 0] = 0.552$$

# — SOLUTIONS —

January 2018

7CCSMCVI

At pixel (3,3) Distance =

$$\frac{1}{8} [0 + 0 + 0 + 1 + 0 + 0 + 0 + 0] = 0.125$$

Hence, object at location (3,3).

Marking scheme

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]

Answer

$$p(objA) = 0.8$$

$$p(objB) = 0.2$$

$$p(I|objA) = 0.02$$

$$p(I|objB) = 0.05$$

$$p(objA|I) = \frac{p(I|objA)p(objA)}{p(I)} = k(0.02 \times 0.8) = 0.016k$$

$$p(objB|I) = \frac{p(I|objB)p(objB)}{p(I)} = k(0.05 \times 0.2) = 0.01k$$

Hence, indistinguishable images are most likely to contain object A.

Marking scheme

# — SOLUTIONS —

**January 2018**

**7CCSMCVI**

3 marks for knowing Bayes' theorem, 4 marks for knowing how to correctly apply the theorem to this task.