CID No: __________________________

IMPERIAL COLLEGE LONDON

Design Engineering MEng

For Internal Students of the Imperial College of Science, Technology and Medicine
This paper is also taken for the relevant examination for the Associateship or Diploma

DESE71004 – Design of Visual Systems

SAMPLE PAPER 2

THE DURATION OF THIS EXAMINATION IS 90 MINUTES.

This paper contains EIGHT questions.
Full marks of the paper is 100 out of 100.
Attempt ALL questions.

The numbers of marks shown by each question are for your guidance only; they indicate how the examiners intend to distribute the marks for this paper.

Students are allowed to bring to the examination one double sided A4 sheet of handwritten information of their own choosing.

1) The probability of intensities in a greyscale image with 16 levels is given in the table in Figure Q1. This image is to be histogram-equalized with the discrete intensity transformation function T, such that the new image intensity values is given by $s = T[r]$.

List the values of $T[r]$ for $r = 0, 1, \dots., 15$.

[12]

r	p[r]
0	0.02
1	0.02
2	0.02
3	0.04
4	0.04
5	0.04
6	0.04
7	0.08
8	0.08
9	0.08
10	0.08
11	0.08
12	0.08
13	0.10
14	0.10
15	0.10

Figure Q1

Solution to Q1

This question tests student's understanding of the principle of histogram equalization. The test is on the following three key ideas: 1) the probability distribution function (PDF) is the normalised version of the intensity histogram of an image; 2) the cumulative distribution function (CDF) is the intensity transformation required to achieve histogram equalization; 3) the cumulative distribution function is the integral of the probability distribution function (PDF).

By definition, $s = T[r] = 0$ for all values r. outside the range of $[0, 15]$.

Therefore, the histogram equalization transformation $T[r]$ for the intensity range [0, 15] is the summation of the PDF function for $r = 0, 1, \dots., 15$ and scaled up by 15 .

Hence: $s = T(r) = 15 \sum_0^r p[r]$

The solution is:

r	p[r]	sum(p[r])	T[r]
0	0.02	0.02	0
1	0.02	0.04	1
2	0.02	0.06	1
3	0.04	0.10	2
4	0.04	0.14	2
5	0.04	0.18	3
6	0.04	0.22	3
7	0.08	0.30	5
8	0.08	0.38	6
9	0.08	0.46	7
10	0.08	0.54	8
11	0.08	0.62	9
12	0.08	0.70	11
13	0.10	0.80	12
14	0.10	0.90	14
15	0.10	1.00	15

[12]

2) The image f and the filter kernel w are given in *Figure Q2*.

Compute the output image g, which is the filtered version of f with the kernel w using convolution. That is:

$$g = w * f.$$

State any assumption used.

[12]

$$w = \begin{bmatrix} -1 & 0 & 1 \\ 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix} \qquad f = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Figure Q2

Solution to Q2

This question tests student's understanding of filtering of an image with convolution.

We assume that the input image is zero-padded with one extra row and column around the border, and that the output image is trimmed back to the original size.

Since the kernel is NOT symmetrical, we need to flip it before performing convolution:

$$w' = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix}$$

The answer is:

$$g = \begin{bmatrix} -1 & -1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & 1 & 1 \end{bmatrix}$$

[12]

3) Briefly explain the trichromatic theory of perception of colour in a human visual system.

[12]

Solution to Q3

Bookwork. This question tests student's knowledge on how human perceives colour.

Here is a possible short answer from the notes:

In trichromatic theory, perception of colour is achieved by adding or combining stimuli from these three types of cones. R, G, B and known as primary colours. Mixing them in various proportions can create a very large range of colours. This theory also explains how colour blindness occurs in some individuals when one or more types of cones are defective.

This theory fits well with the three types of cones found in the retina: L, M and S for the three regions of wavelengths. However, it does not explain perception of after-image effect.

[12]

4) Figure Q4(a) shows an image A with dark pixel being the foreground and white pixel being the background. Figure Q4(b) shows a 3×3 structuring element B.

(a) Sketch the result of A after the **opening** operation with structuring element B. (Opening is erosion followed by dilation.)

[6]

(b) Sketch the result of A after the **closing** operation with structuring element B. (Closing is dilation followed by erosion.)

[6]

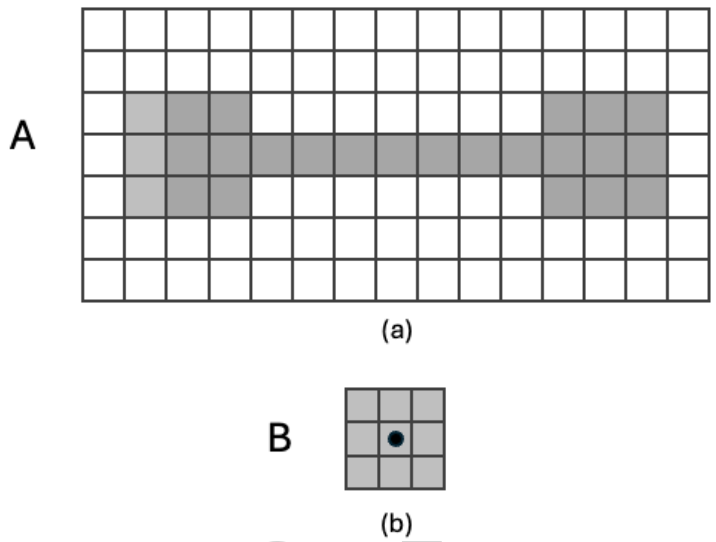
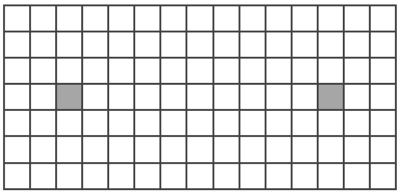


Figure Q4

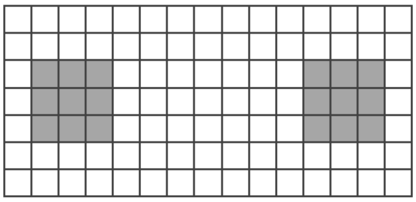
Solution to Q4

This question tests student’s understanding of the two basic morphological operations: erosion and dilation.

a) A eroded by B gives:

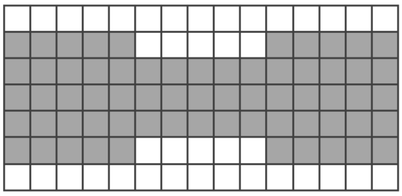


Then, dilated by B = opening gives:

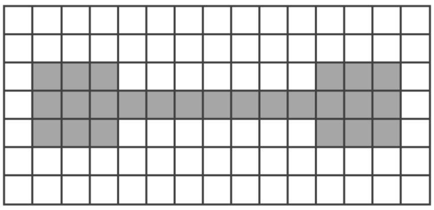


[6]

b) A dilated by B gives:



Then, erosion by B = closing gives:



[6]

5) Hough Transform is a method for detecting features such as straight lines.

a) What is a parameter space in Hough transform for straight line detection?

[3]

b) Figure Q? shows straight line four straight line features with (m, c) parameter space, where m is the gradient and c is the intercept of a line feature. Sketch the Hough transform in the parameter space.

[7]

c) In straight line detection with Hough transform, what is the (ρ, θ) parameter space, and why it is a better space to use than the (m, c) parameter space?

[4]

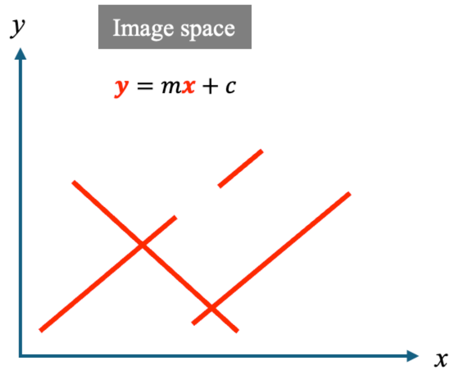


Figure Q5

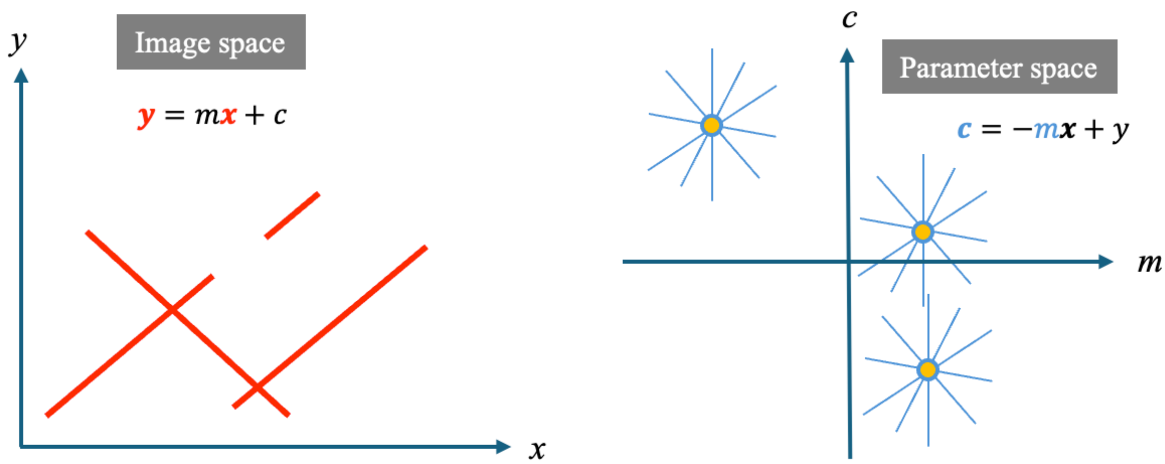
Solution to Q5

This question tests student's understanding of Hough transform.

a) In Hough transform, the parameter space for line detection refers to the space of all possible orientations and locations of lines in the image. Mathematically, it maps each point on the line feature to a line in the (m,c) parameter space. Therefore all the points on the line feature forms a star constellation of lines intersecting at the same (m,c) value.

[3]

b) The plot in of these features in the parameter space is:



Note that two of the lines belong to the same line and hence we only get three intersecting points in the parameter space, and not four.

[7]

c) The Hough transform for straight-line detection typically uses the (ρ, θ) parameter space, where ρ represents the distance from the origin along the line and θ represents the angle of the line with respect to the x-axis.

(m, c) parameter space is very difficult to quantize because c has an infinite range. Further two line segments on the same line maps to the same point in the (m, c) space. In contrast, (ρ, θ) parameter space also provides unique intersections points for line features, but both parameters are within the range of $\pm\pi$.

[4]

6) *Figure Q6* shows data points belonging to two pattern classes A and B in a two-parameter space (x_1, x_2). The stars belong to class A and the squares belong to class B.

a) Construct three straight lines on the diagram to partition the space so that the two classes can be identified.

[5]

b) Hence design a simple neural network architecture with minimum number of layers to classify the two classes of data.

.

[7]

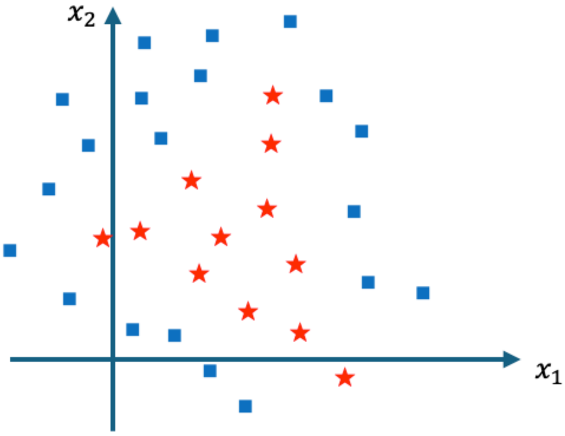


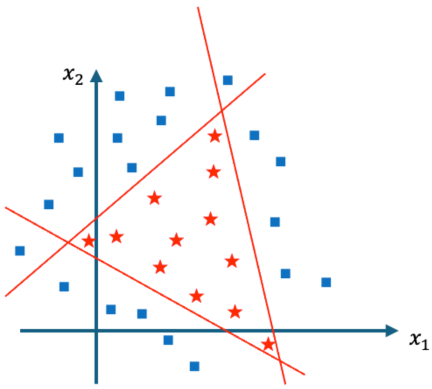
Figure Q6

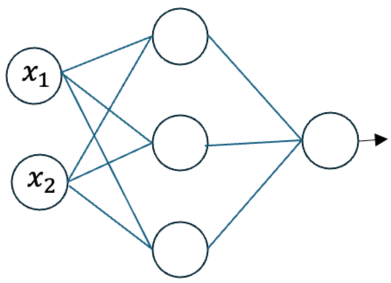
Solution to Q6

This question tests student's understanding why neural network in the form of multilayer perceptron are effective classifiers.

Draw three lines as shown below enclosing all the stars.

[5]





Since a perceptron is a universal linear classifier that separate the two-dimensional feature space into two classes, we can build a neural network with only one hidden layer with three perceptrons (or neurons). The output layer perceptron simply combines all three categories such that on those inside the three linear boundaries belong to class A.

[7]

7) Your classmate asked you to explain, in 2 to 3 minutes, how the retina can help us to see colour images.

Compose, in approximately 200-250 words, the pathway through which light reaches the brain in a human. DO NOT include what the brain does with such information.

[12]

Solution to Q7

This question tests student's knowledge on how neural signals are created by the retina from light. There is no unique answer to this question, but here is a sample answer that includes the main points.

- The retina is the layer of photoreceptor cells located at the back of the eye.
- There are two types of photoreceptors: rods and cones. They are responsible for converting light into electrical signals that are then transmitted to the brain.
- Rods are responsible for detecting low level of light, while cones are responsible for detecting colours.
- The centre of the retina is called the fovea where has a high density of cones and is responsible for detailed colour vision.
- The rest of the retina is full of rods and is responsible for peripheral vision and also particularly sensitive to changes in image.
- When light hits a photoreceptor, it triggers a chemical reaction that generates an electrical signal called an action potential.
- The electrical signal is transmitted to various cells including bipolar cells, and horizontal cells, until it reaches the ganglion cells located in the outer layer of the retina nearer to the source of light.
- The ganglion cells then send the electrical signals to the optic nerve which passes through the hole in the retina (called the blind spot) to the visual cortex in the brain.
- The visual cortex can then process the electrical signals that allow a human to perceive and understand the image generated as electrical signals by the retina.

(211 words)

[12]

8. Affine Transformations allow translations to be encoded in a matrix in addition to linear transformations about the origin.

Given the following transformations that are available to you: a **T**ranslation of (x, y), an in-plane **R**otation (counter-clockwise) about the origin by an angle θ, and a **H**orizontal flip respectively.

$$\mathbf{T}(x,y) = \begin{bmatrix} 1 & 0 & x \\ 0 & 1 & y \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{R}_z(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

What composition of $\mathbf{T}$ and $\mathbf{R}_z$ gives a rotation by an angle θ about a centre (x, y)? Explain this result.

a) Construct, by composition, a flip about a line angled at θ to vertical that passes through the point (x, y).

b) All affine transformations are also homography transformations. Interpreting the translation $\mathbf{T}(x, 0)$ as a homography transformation, what type of transformation does this matrix perform in 3D space?

[14]

Solution to Q8

a) Translate the point at (x, y) to the origin, rotate about the origin, then translate back to (x,y); $\mathbf{T}(x,y)\mathbf{R}_z(\theta)\mathbf{T}(-x,-y)$.

b) Similarly, translate the point at (x, y) to the origin, then rotate such that a line angled at θ to vertical becomes vertical, action the horizontal flip, then undo the rotation and translation; $\mathbf{T}(x,y)\mathbf{R}_z(-\theta)\mathbf{H}\mathbf{R}_z(\theta)\mathbf{T}(-x,-y)$

c) This is a shear transformation. Points in a plane a distance z from the origin are translated proportionate to that distance in the x direction.

[14]

[END OF PAPER]