

CID No: _____

IMPERIAL COLLEGE LONDON

Design Engineering MEng EXAMINATIONS 2025

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

Date: 1 May 2025
Time: 10.00am – 11.30am
Duration: One hour thirty minutes

*This paper contains SEVEN 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. (a) What is bit-plane slicing of an image and why would this provide insight about the image? [5]

(b) A 4 x 4 image has 4-bit intensity values shown in Figure Q1. Derive the most significant bit-plane for this image. [10]

$$\begin{bmatrix} 0 & 1 & 8 & 6 \\ 2 & 2 & 1 & 1 \\ 1 & 15 & 14 & 12 \\ 3 & 6 & 9 & 10 \end{bmatrix}$$

Figure Q1

Solution to Q1

This question tests student's understanding pixel bit-plane and bit-plane slicing.

(a) In a greyscale image, each pixel stores its intensity level which is usually 4 to 8 bits. Bit-plane slicing is a technique where by the image is decomposed into its individual bit planes. Each bit-plane now stores only 1 bit of the corresponding image from MSB to LSB. Bit-plane slicing allows a user to compress the image by dropping those bit-planes that are not significant. It can also help to analyse an image where the higher significant bit-plane can provide structural information about the image. [5]

(b) Convert the image into 4-bit binary intensities:

$$\begin{bmatrix} 0000 & 0001 & 1000 & 0110 \\ 0010 & 0010 & 0001 & 0001 \\ 0001 & 1111 & 1110 & 1100 \\ 0011 & 0110 & 1001 & 1010 \end{bmatrix}$$

From this, the four bit-planes can easily be derived as:

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

[10]

2. a) With the help of suitable diagrams or graphs, briefly explain how a person without visual impairment perceives colour.

[10]

b) *Figure Q2* shows a colour pallet perceived by a person with normal vision and another perceived by a person with colour blindness. Explain with justifications what could cause such colour blindness?

[5]

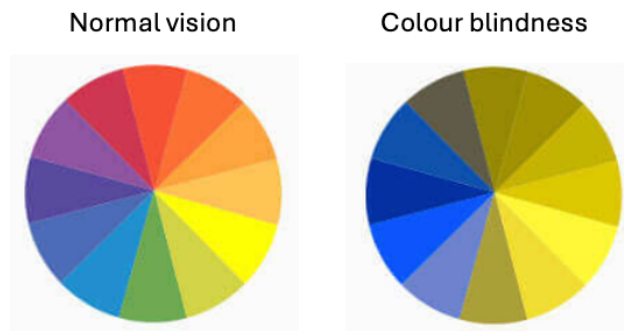


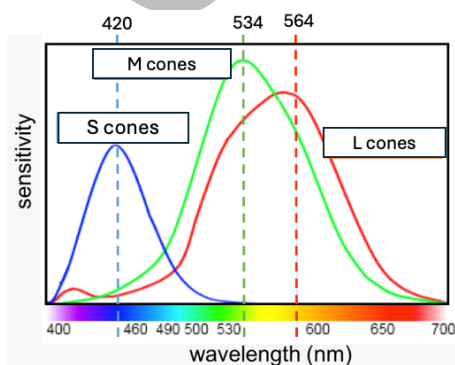
Figure Q2

Solution to Q2

This question tests students understanding about perception of colour in human visual pathway and colour blindness. Such understanding helps to design colour-blind friendly products.

a) Retina contains about 6 million photoreceptors known as cones which are sensitive to light at different wavelengths. There are three types of cones: S cones (short wavelength at around 420nm) that perceives light near the blue spectrum. Similar, M cones (medium wavelength at around 530nm) and L cones (long wavelength at around 560nm) that perceive light near the green and red spectrum respectively.

The sensitivity vs wavelength characteristics of these three types of cones are shown below:



Cones activates neurons connected to them and depending on the relative activation of each cone type, the brain perceives different the different colours. For example, if the M cones and L cones are strongly activated, but S cones are not activated at all, the brain could be “trained” to recognise the colour to be yellow.

[10]

b) Red colour is not perceived, therefore either L cones are missing or the L cone sensitivity characteristics overlaps with that of the M cones.

[5]

3. The image f and the filter kernel w are given in *Figure Q3*.

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

$$g = w * f.$$

You should zero pad the image as necessary, and the output image should not be trimmed to the original size.

[15]

$$w = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad f = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Figure Q3

Solution to Q3

This question tests student's understanding of filtering of an image with convolution, how neighbourhood processing is performed, zero padding of images, and that correlation and convolution are NOT the same if the filter kernel is non-symmetrical (as is the case here).

First, we need to form the 'flipped' version of the kernel, i.e. mirrored in both x and y direction around the centre to get:

$$w' = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

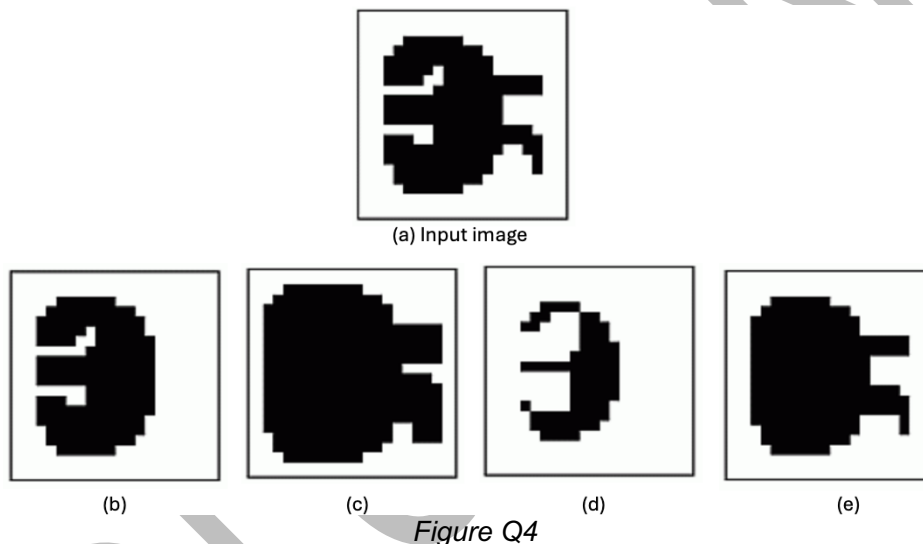
Assume that we zero-padded the input image with one extra row and column around the border, and that the final image is not trimmed back to the original size.

$$f_{padded} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$g = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

[15]

4. a) Explain in words the following morphological operations as applied to image processing and their applications:
- i) erosion;
 - ii) dilation.
- [10]
- b) How can dilation and erosion can be combined to form opening and closing operations?
- [4]
- c) Figure Q4 (a) shows an input image subjected to dilation, erosion, opening and closing operations with some suitable structuring element. The results are the output images shown in Figure Q4 (b)-(e).
- List in your answer book which of these four morphological operations would produce the corresponding outputs.
- [6]



Solution to Q4

This question tests student's understanding of morphological operations.

- a) Book work. Morphological operation uses a structuring element on a binary or a greyscale image to produce outputs that is expanded or contracted. The structuring element is a small matrix of certain shape (could be square, rectangular, circular or in a shape of a cross) which determines how the input image is modified.

In erosion, the structuring element is placed over each pixel of the input image. The output pixel is set to foreground if ALL pixel under the structuring element is a foreground pixel (i.e. 1). This is a logical AND operation. Erosion can be used for disconnection of thin objects, boundary smoothing and feature reduction or thinning. Together with dilation, erosion can also remove noise.

$$A \ominus B = \{z | (B)_z \subseteq A\}$$

[5]

In dilation, the structuring element is again placed over each pixel of the input image. The output pixel is only set to foreground if ANY pixel under the structuring element is a foreground pixel (i.e. 1). This is a logical OR operation. Dilation can be used for hole and gap filling, edge smoothing and feature enhancement. Together with erosion, dilation can also remove noise.

$$A \oplus B = \{z | (\hat{B})_z \cap A \neq \emptyset\} \quad [5]$$

- b) Opening operation: erosion followed by dilation.
Closing operation: dilation followed by erosion.

[4]

- c) Figure 4 (b) – Opening.
Figure 4 (c) – Dilation.
Figure 4 (d) – Erosion.
Figure 4 (e) – Closing.

[6]

SOLUTIONS

5. Your company asks you to develop a lane detection algorithm as a component in an autonomous driving system. You decided to propose the use of the Canny edge detection algorithm. You are preparing a presentation to explain to your manager and other colleagues how the principles behind the Canny algorithm.

To help your preparation, list in bullet points the main principles behind this edge detection method and why it is suitable for solving the lane detection problem. You may choose to use sketches to illustrate the ideas where appropriate.

[10]

Solution to Q5

This question tests student's understanding of either the Canny edge detection method.

Canny edge detection method:

- Multi-step algorithm that identifies edges by discovering parts of the image that has high rate of intensity variations.
- Step 1: Reduce noise with Gaussian filter.
- Step 2: Find 1st derivative of the filtered image including intensity gradient magnitude and direction.
- Step 3: Thin the edges by applying non-maximum suppression.
- Step 4: Apply double threshold to determined strong and weak edges.
- Step 5: Tracking edges using hysteresis.

Advantages of Canny detection method:

- High accuracy
- Robust against noise
- Edges are thin and very localized as compared to other algorithms

[10]

6. (a) What is image segmentation and what are the advantages and disadvantages of using a K-means clustering algorithm to perform image segmentation? [8]
- (b) List, with brief explanations, the steps required when applying the K-means clustering method to perform segmentation of a colour image? [4]
- (c) What are the three key parameters in K-means that need the user's intervention? [3]
-

Solution to Q6

This question tests student's understanding of image segmentation and K-means as one methods that performs image segmentation based on colour or texture, or similar properties.

- (a) Image segmentation is one of the steps in image analysis, understanding or recognition that divides an image into multiple distinct regions (or segments) that share some characteristics such as colour or texture. This simplifies the image so that it becomes easier for further analysis or interpretation. There are many technical that could be employed to perform image segmentation including thresholding, feature (such as edge) detection, clustering, region identification and neural networks.

The advantages of K-means are: 1) simple to apply and understand; 2) efficient to apply; 3) can handle high-dimensional data (such as RGB clour values).

The disadvantage of K-means are: 1) sensitive to parameters chosen by user; 2) assume that clusters are spherical and similar in size; 3) can be slow if dealing with large images and/or high-dimensional data.

[8]

- (b) The steps in K-means clustering are:

Step 1: choose the number of clusters and randomly initialise the centroids of each cluster in the feature space (such as pixel intensity or colour).

Step 2: Classify each pixel to the nearest centroid using some distance metric (such as Euclidean or some of absolute distance).

Step 3: Recompute the centroids of all pixels after classification.

Step 4: Iterate step 2 and 3 until the centroids is not changing significantly between iterations.

[4]

- (c) The three parameters required user's interventions are: 1) number of clusters, 2) the distance metric and 3) the initial centroid locations.

[3]

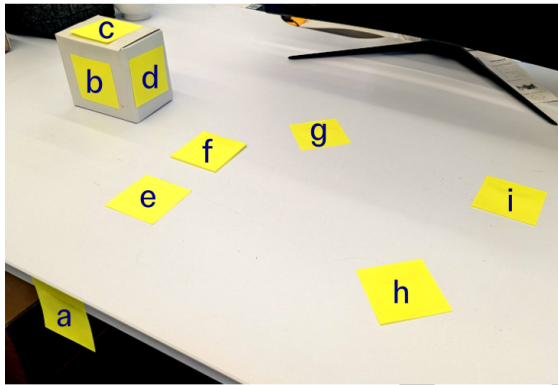
7. Figure Q7 (i) Is an annotated photograph of square post-it notes on a desk. Next to it (Figure Q7 (ii)) is the same photo after a homography transformation, T , is applied.

a) Explain the properties of a homography transformation and how this explains the features of the image in Figure Q7 (i) & (ii). You can refer to the labelled post-it notes.

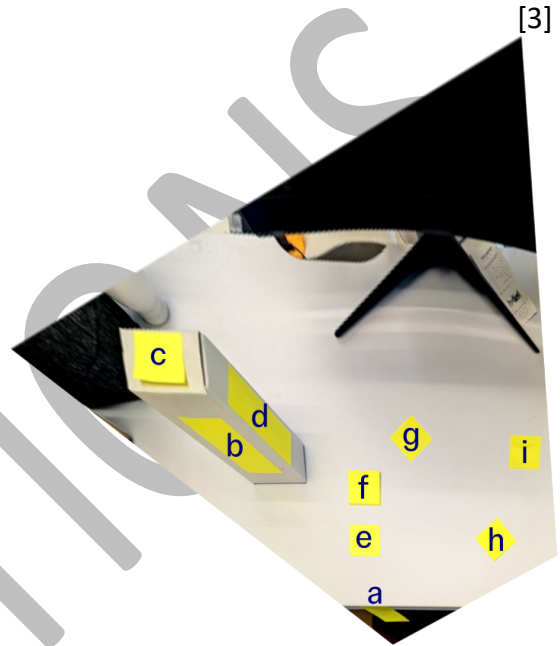
[7]

b) How would you use this homography transformation to edit the original image to include the Imperial logo in the plane of the desk, as in Figure Q7 (iii).

[3]



(i) Original Image



(ii) After Transformation

Figure Q7 (i) and (ii)

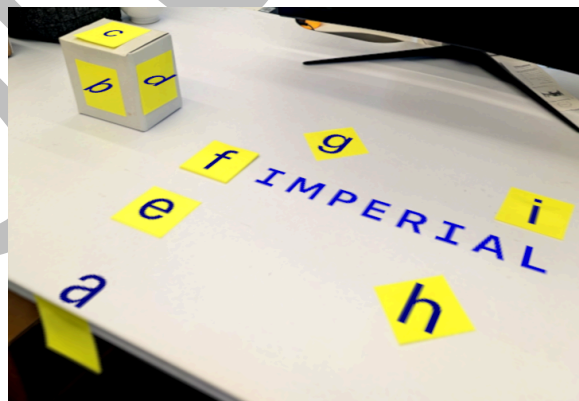


Figure Q7 (iii)

Solution to Q7

- a) Explanation of the transformation might include,
- A homography transform combines a linear transformation in 3D with a projection to the 2D plane.
 - Transformations compose like matrices.
 - The transformation treats all points as coplanar in a 3D space (or projected onto a single plane), and transforms that plane in 3D.
 - Parallel lines in 3D transform to lines that converge to a vanishing point in 2D.
 - This will map trapeziums to trapeziums in the 2D image.

Then relating to the figure,

- Hence, here, the image border is transformed to a trapezium.
- Post-its e, f, g, h , and i are in the plane of the table, and these have each transformed to be square in the final image, and for e, f , and i – axis aligned.
- This transformation has aligned parallel lines in the desk plane in the 3D scene to be parallel in the 2d projection.
- This can be seen in the left edges of e and f are at a different angle in the original image to the left edge of i before being transformed to be parallel.
- Note c is parallel to the desk plane. It transforms to be square, but is larger than the other notes, since the plane it is sitting in is closer to the camera than the desk plane, giving a larger projection on the desk plane.
- Notes b and d are not in the desk plane and are significantly distorted. Their shapes are determined by drawing lines of sight from the camera through their boundary to the desk plane.
- They are closer to the camera than the desk plane, giving an enlarged projection.
- Note a is further away from the camera and as such appears smaller when transformed.

[5]

- b) First, transform the image to put the desk plane as the image plane, using transformation \mathbf{T} . Secondly, compose the graphic onto the transformed image. Then use the inverse transform \mathbf{T} to transform back.

[2]

[END OF PAPER]