GENERAL SIR JOHN KOTELAWALA DEFENCE UNIVERSITY

Faculty of Engineering

Department of Electrical, Electronic and Telecommunication Engineering

BSc Engineering Degree

End Semester Examination – May 2022

Semester 5 - Intake 37 (EE/ET)

ET3122 – IMAGE PROCESSING AND MACHINE VISION

Time allowed: 2 hours 26th May 2022

ADDITIONAL MATERIAL PROVIDED

None.

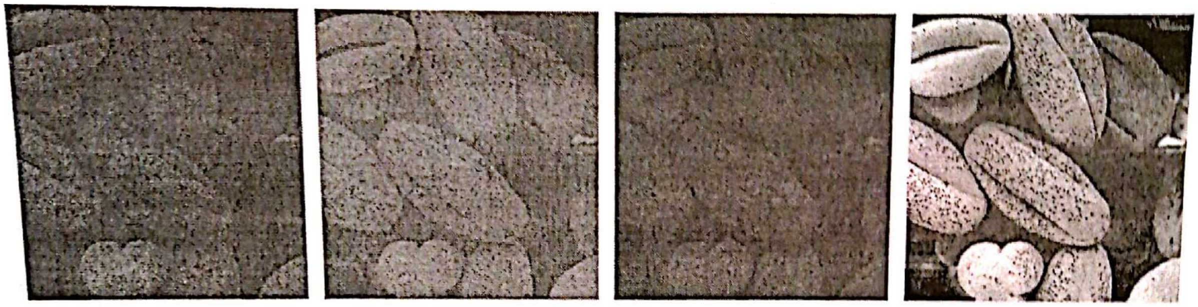
INSTRUCTIONS TO CANDIDATES

- This paper contains 4 questions and answer **all** the questions on answer booklets.
- This paper contains **9** pages with the cover page.
- This is a closed book examination.
- Neat and orderly presentation is important.
- The symbols used in this paper have their usual meanings.
- This examination accounts for 70% of the module assessment. A total maximum mark obtainable is 100. The marks assigned for each questions and parts thereof are indicated in square brackets.
- If you have any doubt as to the interpretation of the wordings of the question, make your own decision, but clearly state it on the script.
- Assume any reasonable values for any data not given in or provided with the question paper, clearly make such assumptions made in the script.
- All examinations are conducted under the rules and regulations of the KDU.

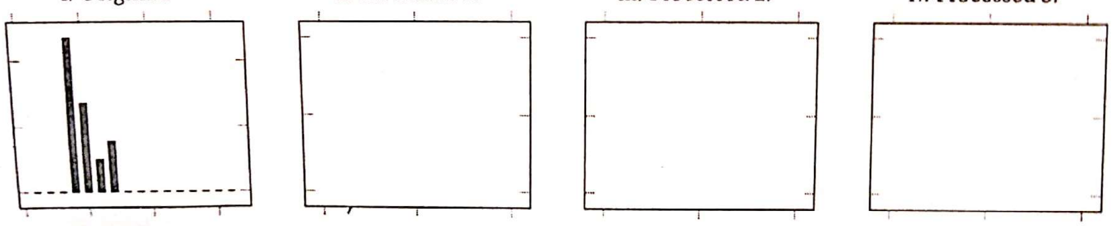
Q1. (a) Fig. Q1(a) shows an image of a shell, its histogram, and three processed versions. [6]

i. Approximately sketch the histograms of the processed images.

ii. State the probable image processing operation done on the source image to produce each of the three results shown in Fig. Q1(a)ii., Q1(a)iii., and Q1(a)iv..



i. Original. ii. Processed 1. iii. Processed 2. iv. Processed 3.



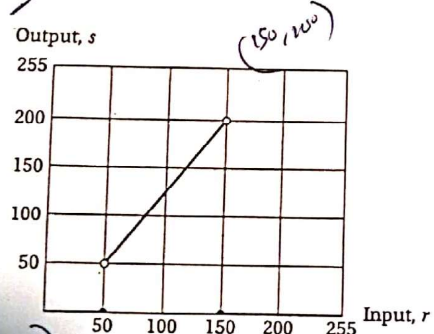
v. Histogram. vi. Answer. vii. Answer. viii. Answer.

Figure Q1(a): Images and histograms.

(b) Figure Q1(b) shows an intensity transfer function and an input image.

i. Apply this transfer function to the input image and show the resulting pixel values in the grid shown. [6]

ii. Suggest a use of this operation. [2]



(50,50)

i. Transfer function.

150	200	255
100	100	200
50	120	100
10	10	50

ii. Input image.

0	0	0
125	125	0
0	155	125
0	0	0

iii. Grid for output.

Figure Q1(b): Intensity transfer function, input image, and grid for output.

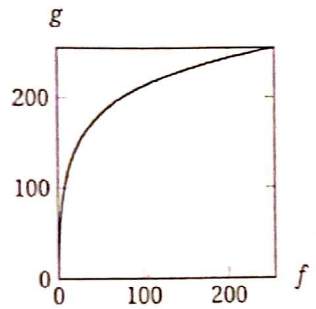
(c) Consider intensity transfer function designated by

$$g = \max(f)\frac{\log(1+f)}{\log(1+\max(f))},$$

where f is the input image, and g is the output image. $g, f \in [0,255]$. Figure Q1(c) shows this transfer function, an input image, and a grid for showing the output.

i. What is the effect of this transformation? [2]

ii. Show the output due to this transformation in the gird shown. [4]



i. Transfer function.

150	200	255
100	100	200
50	50	100
10	10	50

ii. Input image $f(i, j)$.

iii. Grid for output.

Figure Q1(c): Intensity transfer function, input image, and grid for output.

(d) Table Q1(d) shows the pixel counts of a 3-bit image.

i. Compute the transfer function (look-up table) that would equalize the histogram of a this image. [4]

ii. Compute the corresponding values of the resultant image for input pixel value 0, 3, and 7. [1]

Table Q1(d): Histogram of a 3-bit image.

k	0	1	2	3	4	5	6	7
n_k	5	10	16	20	10	8	6	5

Q2. (a) Figure Q2(a) shows an image, a kernel, and a grid for displaying the output.

i. Carry out filtering using the kernel shown in the same figure. Show the output in the given grid. **[4]**

ii. In the above Q2(a)i, the output image size is smaller than the input image. Use zero padding to produce a filtered image of the same size as the input image. **[4]**

0	0	0	0	0	0	0
0	20	60	200	180	240	0
0	20	60	200	220	220	0
0	40	160	180	200	220	0
0	120	180	180	200	240	0
0	0	0	0	0	0	0

i. Image

0	−1	0
−1	4	−1
0	−1	0

ii. Kernel

	200	140	80	
	180	-20	-20	

iii. Output without padding.

0	40	360	60	560
-40	-20	140	80	200
-140	180	-20	-10	220
-60	260	160	160	540

iv. Output with padding.

Figure Q2(a): Image, kernel, and output grids for ??.

(b) The Laplacian of Gaussian is

$$l(x,y) = -\frac{1}{\pi\sigma^4}\left[1 - \frac{x^2+y^2}{2\sigma^2}\right] e^{-\frac{x^2+y^2}{2\sigma^2}}.$$

Figure Q2(b) shows a part of the un-normalized 3×3 Laplacian of Gaussian kernel with $\sigma = 1.4$.

	-1	0	1
-1	0.0	0.3233	0.6
0	0.0478	1	0.0478
1	0.6	0.3317	0.6

Figure Q2(b): A part of the 3×3 Laplacian of Gaussian kernel.

i. Fill the other entries in the kernel. **[4]**

ii. Why are all the entries positive? **[2]**

(c) In order to detect circular coins of different sizes, scale-space extrema detection (blob detection) cane be used.

i. Obtain a relationship between the radius r of the detected circular coin and the σ value of the Laplacian of Gaussian kernel. **[3]**

ii. If a scale-space extremum is found with $\sigma = 10$, what would be the radius of the corresponding coin? **[2]**

(d) Harris corder detection is a widely used algorithm.

i. List the steps of detecting Harris corners. **[4]**

ii. In a typical Harris corner detection algorithm applied for a 600×800 image, what would be the size of the Harris corniness measure R? **[2]**

Q3. (a) A camera calibration software gave the following parameters for a particular camera:

$$\alpha_x = 650, \quad \alpha_y = 650, \quad \beta_x = 402, \quad \beta_y = 297.$$

The camera rotation matrix and the translation matrix with respect to the world coordinate system are

$$R = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and,} \quad \tilde{C} = [0 \quad 0 \quad 2]^T.$$

i. Compute the camera matrix P. [2]

ii. Compute the Cartesian coordinates of the image of the point $X = [2 \quad 3 \quad 10 \quad 1]^T$. [2]

iii. What will happen to the size of the image of an object if the camera is translated such that $\tilde{C} = [0 \quad 0 \quad 4]^T$? [2]

(b) Fig. Q3(b) shows an image with two earrings and the corresponding connected components analysis result. Following segmentation, morphological closing, and connected component analysis, the statistics returned the following three pixel counts as area of objects: 930135, 59192, 59249. Pixel size is 2.2 μm. The focal length of the lens is 8 mm. The distance between the lens and the imaging surface is 720 mm.

i. Earrings image. ii. Colormapped connected components

Figure Q3(b): Earrings and colormapped connected components.

i. Compute the area of each earring in mm^2. [4]

ii. What is the reason for having to apply morphological closing? [2]

(c) Segment the image in Figure Q3(c)i. using thresholding. Choose 200 as the threshold. [6]

(d) Figure Q3(d) shows a set of noisy points, which are known to form a line. The point set is

$$\begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \\ x_4 & y_4 \\ x_5 & y_5 \end{bmatrix} = \begin{bmatrix} 0 & 4 \\ 1 & 4 \\ 2 & 7 \\ 3 & 8 \\ 4 & 9 \end{bmatrix}.$$

i. Show that the normal equations,

$$X^T X B = X^T Y,$$

give the parameters of the least-squares-fit line. [3]

ii. Use all the points to find the least-squares-fit line and show graphically in the same scatter plot. [3]

iii. What must be done if there are outliers in the dataset? [1]

(c) Consider a liner classifier that takes in CIFAR10 images (size $32 \times 32 \times 3$, 10 classes). For a (flattened) image $\boldsymbol{x}$, the scoring function is [5]

$$f(\boldsymbol{x}) = \mathrm{W}\boldsymbol{x} + \boldsymbol{b}.$$

i. Stat the sizes of $\boldsymbol{x}$.
ii. Stat the sizes of W.
iii. Stat the sizes of $\boldsymbol{b}$.
iv. Write the expression of a simple loss function that can be used to optimize this classifier.
v. After the iterations that learns W and $\boldsymbol{b}$, the plot of the loss history was noisy. Explain the reason for this observation.

(d) Consider a convolutional neural network that classifies the digits in the MNIST dataset. Each image in MNIST is a 28×28 grayscale image. There are 10 classes representing the 10 digits. The network comprises the following o/p

- A convolutional layer of 6 kernels of size 5×5.
- 2×2 average pooling with stride 2.
- A convolutional layer of 16 kernels of size 3×3.
- 2×2 average pooling with stride 2.
- A flattening layer.
- A dense layer of 120 nodes.
- A dense output layer of 10 nodes.

i. Sketch the network. [2]
ii. What is the layer most heavy in terms of the computations? convolution. [2]
iii. Compute the number of learnable parameters clearly showing the work. [4]