GENERAL SIR JOHN KOTELAWALA DEFENCE UNIVERSITY

Faculty of Engineering

Department of Electrical, Electronic and Telecommunication Engineering

B.Sc. Engineering Degree

Semester 5 Examination – May 2023

(Intake 38 – EE/ET)

ET 3112 – IMAGE PROCESSING AND MACHINE VISION

Time allowed: 2 hours 08 May 2023

ADDITIONAL MATERIAL PROVIDED

None

INSTRUCTIONS TO CANDIDATES

This paper contains 4 questions on 8 pages.

Answer **ALL** questions.

This is a closed book examination.

The symbols used in this paper have their usual meanings.

Neat and orderly presentation is important.

This examination accounts for 70% of the module assessment. The marks assigned for each question and parts thereof are indicated in square brackets.

If you have any doubt as to the interpretation of the wordings of a question, make your own decision, but clearly state it on the script.

Assume 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.

This Page is Intentionally Left Blank

Q1. (a) The following is a list of application areas of image processing and computer vision. Sate two vision algorithms or techniques used in each of the areas. [5]

i. Photo processing
ii. Medical image processing
iii. Autonomous driving
iv. Security and surveillance
v. Robotics

(b) Figure Q1(b) shows four images and four intensity transfer functions. I is the original image. In the transfer functions (f_1, f_2, f_3, and f_4) the horizontal axis represents the value of the input pixel and the vertical axis represents the value of the output pixel. Pick the intensity transfer function that would have produced the images I_1, I_2, and I_3. [6]

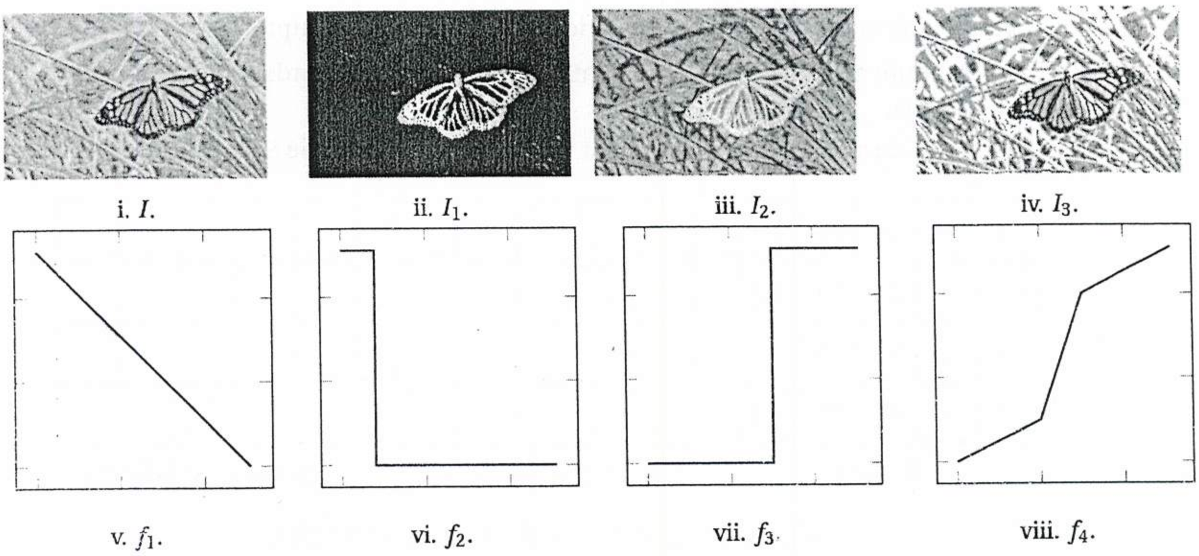


Figure Q1(b): Images and transfer functions.

(c) Fig. Q1(c) shows an image enhancement pipeline. The purpose of this is to increase the vibrance of the facial area. Identify the purpose of each block, and explain the pipeline. [7]

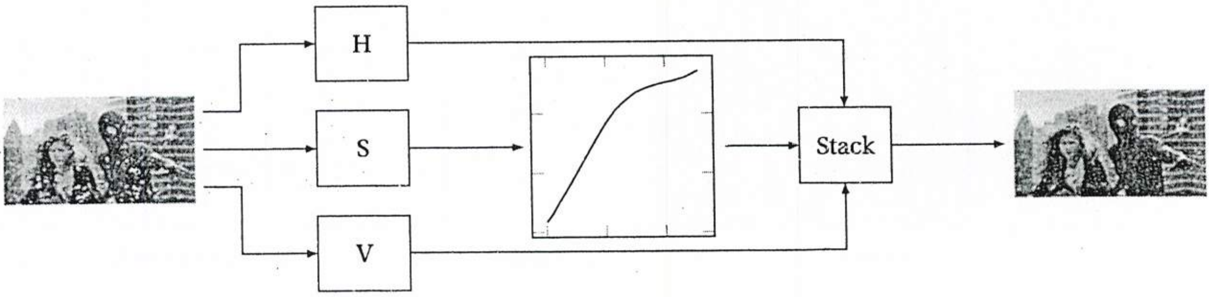


Figure Q1(c): Block diagram of an image enhancement pipelines.

(d) Fig. Q1(d) show a 4-bit image and a mask representing the foreground.

i. Equalize the histogram of the whole image and display the result in a grid. [4]
ii. Equalize the histogram of only the foreground. [7]

5	7	9	9	9	9	9	9	1	1	2	4	10	6	3	2	5	7	7
8	9	9	9	9	9	9	9	5	1	9	6	6	8	1	2	7	7	7
8	8	9	9	9	9	9	9	8	8	7	9	9	2	2	2	7	7	7
8	8	8	8	8	9	9	8	6	4	9	7	6	0	5	4	7	7	6
8	8	8	8	8	8	8	8	4	10	9	1	7	2	7	4	7	7	6
7	8	8	8	8	8	8	8	8	8	1	2	2	2	4	5	7	6	6
7	7	7	7	8	8	8	8	2	4	5	10	4	13	11	7	6	6	6
7	7	7	7	7	7	7	7	7	3	3	9	13	13	7	6	6	6	5
7	7	7	7	7	7	7	7	5	8	13	12	11	11	6	6	6	6	5
6	6	6	6	7	7	6	7	6	6	14	13	8	6	6	6	6	5	5
6	6	6	6	6	6	6	6	6	6	12	6	6	6	6	5	5	5	5

i. Image.

0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0
0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0
0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	0	0	0
0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	0	0	0
0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0
0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0
0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0
0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0

ii. Mask.

Figure Q1(d): Image for histogram equalization and the mask.

Q2. (a) Figure Q2(a) shows an image f, and grids for displaying the output.

i. Compute and display the gradient f_x (vertically downwards) using the central difference formula. [4]

ii. Compute and display the gradient f_y (horizontal) using the Sobel kernel. [4]

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

i. Image f

ii. Output grid for f_x.

iii. Output grid for f_y.

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

(b) A 5×8 image and a kernel are shown in Figure Q2(b). The center of the kernel is indicated by a circle.

145	76	150	21	183	154	135	78
163	194	198	4	170	103	102	162
200	164	153	0	84	95	73	196
105	180	126	68	201	187	202	73
195	194	186	0	195	93	134	83

i. Image

o

ii. Kernel

iii. Grid for result

Figure Q2(b): Image and kernel for median filtering.

i. Carry out median filtering on the image using the kernel to generate a 3×8 image. [3]

ii. By observing the result, identify a special property of this median filtering kernel. [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 $\sigma = 5$, what would be the radius of the corresponding coin? [2]

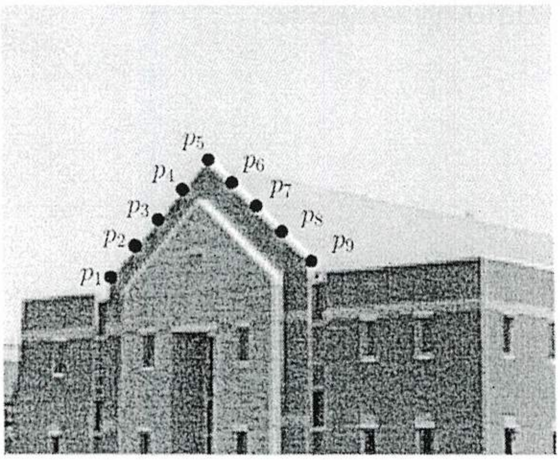
(d) Blurring out the background is an operation used to make pictures, e.g., portraits, attractive.

i. Assuming the availability of a foreground segmentation algorithms such as GrabCuts, design a background blurring algorithm and present it using a block diagram. **[5]**

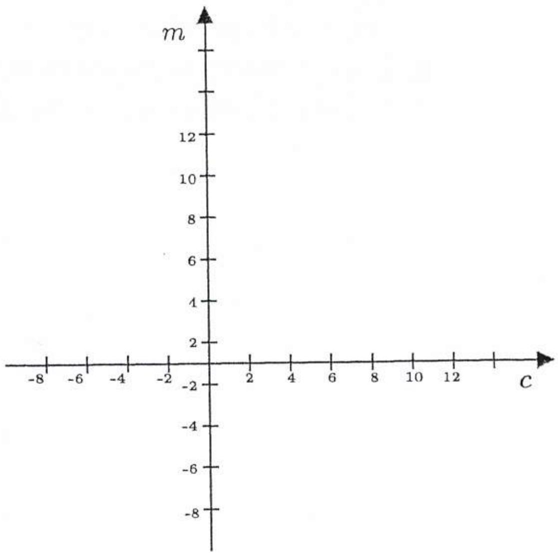
ii. List one important parameter of the algorithm and state its effect on the result. **[1]**

iii. Identify an artifact that may result from using the algorithm. **[1]**

Q3. (a) Draw the Hough parameter space that corresponds to the image pixels p_1 to p_9 of Figure Q3(a)i. For the drawing use Figure Q3(a)ii. Note that the pixels p_1 to p_5 fit a line of $y = 2x + 1$ and pixels p_5 to p_9 fit a line of $y = -2x + 10$. [5]



i Image space.



ii Hough parameter space.

Figure Q3(a): Image space and Hough parameter space for **Q3**a.

(b) To calculate the angle of inclination of the roof in the house depicted in Figure Q3(a)i, a dataset is formed by selecting pixels p_5 to p_9. It should be noted that the process of extracting the pixel positions is not perfect and some pixel locations may have been altered. These pixel points are shown in Figure Q3(b) which are known to form a line. The resulting dataset consists of a set of points with their corresponding x and y coordinates are given below.

$$\begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \\ x_4 & y_4 \\ x_5 & y_5 \end{bmatrix} = \begin{bmatrix} 0 & 12 \\ 1 & 10 \\ 2 & 8 \\ 3 & 10.5 \\ 4 & 4 \end{bmatrix}.$$

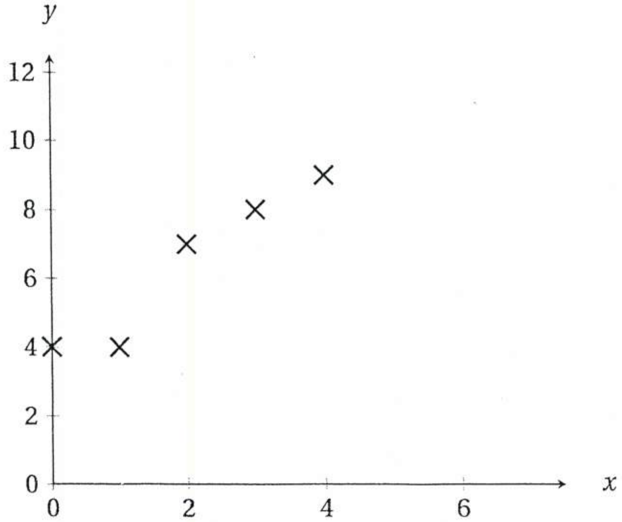


Figure Q3(b): Scatter plot of points for line fitting.

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

(d) Write down the values of the slope (m) and intercept (c) of the line based on your least-squares solution. [2]

(e) To reduce the impact of outliers, a robust estimator is introduced. The robust estimator finds model parameters which minimize the following loss function

$$\sum_i \rho_\sigma(r(x_i;\theta)).$$

Here, $r(x_i;\theta)$ residual of x_i w.r.t. model parameters θ and $\rho_\sigma(r) = \frac{r^2}{r^2+\sigma^2}$ with $\sigma = 1$. Fill Table Q3(e) for two models given below. [10]

i. $y = -2x + 10$,

ii. $y = -1.55x + 12$.

In table Q3(e), $\bar{y} = mx + c$ is the output of the model.

Table Q3(e): Table for question **Q3e**.

For $y = -2x + 10$					
x	y	$\bar{y} = mx + c$	$r = y - \bar{y}$	L	L_r
$x_1 = 0$	$y_1 = 12$	$\bar{y}_1 =$	$r_1 = y_1 - \bar{y}_1 =$	$l_1 = (r_1)^2 =$	$l_{r,1} = \frac{(r_1)^2}{(r_1)^2+\sigma^2} =$
$x_2 = 1$	$y_2 = 10$	$\bar{y}_2 =$	$r_2 = y_2 - \bar{y}_2 =$	$l_2 = (r_2)^2 =$	$l_{r,2} = \frac{(r_2)^2}{(r_2)^2+\sigma^2} =$
$x_3 = 2$	$y_3 = 8$	$\bar{y}_3 =$	$r_3 = y_3 - \bar{y}_3 =$	$l_3 = (r_3)^2 =$	$l_{r,3} = \frac{(r_3)^2}{(r_3)^2+\sigma^2} =$
$x_4 = 3$	$y_4 = 10.5$	$\bar{y}_4 =$	$r_4 = y_4 - \bar{y}_4 =$	$l_4 = (r_4)^2 =$	$l_{r,4} = \frac{(r_4)^2}{(r_4)^2+\sigma^2} =$
$x_5 = 4$	$y_5 = 4$	$\bar{y}_5 =$	$r_5 = y_5 - \bar{y}_5 =$	$l_5 = (r_5)^2 =$	$l_{r,5} = \frac{(r_5)^2}{(r_5)^2+\sigma^2} =$
				$(\frac{1}{5})\sum_{i=1}^{5} l_i =$	$(\frac{1}{5})\sum_{i=1}^{5} l_{r,i} =$
For $y = -1.55x + 12$					
x	y	$\bar{y} = mx + c$	$r = y - \bar{y}$	L	L_r
$x_1 = 0$	$y_1 = 12$	$\bar{y}_1 =$	$r_1 = y_1 - \bar{y}_1 =$	$l_1 = (r_1)^2 =$	$l_{r,1} = \frac{(r_1)^2}{(r_1)^2+\sigma^2} =$
$x_2 = 1$	$y_2 = 10$	$\bar{y}_2 =$	$r_2 = y_2 - \bar{y}_2 =$	$l_2 = (r_2)^2 =$	$l_{r,2} = \frac{(r_2)^2}{(r_2)^2+\sigma^2} =$
$x_3 = 2$	$y_3 = 8$	$\bar{y}_3 =$	$r_3 = y_3 - \bar{y}_3 =$	$l_3 = (r_3)^2 =$	$l_{r,3} = \frac{(r_3)^2}{(r_3)^2+\sigma^2} =$
$x_4 = 3$	$y_4 = 10.5$	$\bar{y}_4 =$	$r_4 = y_4 - \bar{y}_4 =$	$l_4 = (r_4)^2 =$	$l_{r,4} = \frac{(r_4)^2}{(r_4)^2+\sigma^2} =$
$x_5 = 4$	$y_5 = 4$	$\bar{y}_5 =$	$r_5 = y_5 - \bar{y}_5 =$	$l_5 = (r_5)^2 =$	$l_{r,5} = \frac{(r_5)^2}{(r_5)^2+\sigma^2} =$
				$(\frac{1}{5})\sum_{i=1}^{5} l_i =$	$(\frac{1}{5})\sum_{i=1}^{5} l_{r,i} =$

(f) Using this robust estimator, choose the best-fitted model among the models mentioned in **Q3e** for the given dataset. Justify your answer. [2]

(g) Suppose you have the freedom to remove one data point from the dataset. In that case, which data point would you choose to remove? Next, calculate the least-squares-fit line for the given data (with one data point omitted) and write down the values of the slope (m) and intercept (c) of the line based on the least-squares solution. [2]

(h) What must be done if there are outliers in the dataset? [1]

Q4. (a) List four visual recognition tasks for which machine learning or deep learning is widely used. [4]

(b) Backpropagation and stochastic gradient decent are the two main operations in training a neural network

i. Consider the single-variable function

$$f(w) = 3w^2 - 24w - 4.$$

Stating form $w_0 = 1$ run two iterations of gradient descent with a learning rate of 0.1 and write down the updated values of w and $f(w)$ after two iterations. **[4]**

(c) Figure Q4(c) shows an image, a CNN kernel, and a grid for displaying the output.

i. Apply the convolution operation on the image using the kernel shown in the same figure. Here, both padding and stride is set to 1. Show the output in the given grid. [4]

ii. Apply a ReLU activation function on the result. The ReLU activation function is defined as $f(x) = \max(0, x)$. [2]

0	10	20	30
10	0	20	200
40	20	20	200
190	180	190	200

i Image

$\frac{0}{5}$	$\frac{0}{5}$	$\frac{1}{5}$
$\frac{-1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$
$\frac{-1}{5}$	$\frac{-1}{5}$	$\frac{0}{5}$

ii CNN Kernel

iii Grid for CNN Kernel output

iv Grid for ReLU output

Figure Q4(c): Image, CNN kernel, and output grid for **Q4**c.

(d) Consider a convolutional neural network (CNN) implementation for detecting hand written digits. There are 10 digits from 0 to 9. Here, assume that each digit image is 32×32 RGB (3 color channels). The CNN consists of a convolutional layer with five 3×3 kernels (padding = 0 and stride = 1) followed by a max-pooling layer of 2×2 pooling window (stride = 2). Then, there is a flatten layer of l nodes. Note that the number of nodes in the flatten layer is determined by the size of the output from the preceding layer. Next, the flatten layer follows by a dense layer of k nodes (neurons) which has same dimension as the flatten layer. Finally there is a dense softmax output layer with m nodes (neurons).

i. How many learnable parameters in the convolutional layer ? [2]

ii. How many nodes must be there in the dense layers (k and m)? [4]

iii. Sketch the network. [2]

iv. Compute the total number of learnable parameters of this CNN. [3]