C7-R4: DIGITAL IMAGE PROCESSING & COMPUTER VISION

NOTE:

1.	**Answer question 1 and any FOUR from questions 2 to 7.**
2.	**Parts of the same question should be answered together and in the same sequence.**

Time: 3 Hours **Total Marks: 100**

1.

a) Given that $x_i = y_i = 0,1,2,3,4,5,6,7$ and the input probability density function

$$p_r(x_i) = \begin{cases} \dfrac{x_i}{12} & \text{i=0,1,2,3} \\ \dfrac{|7-x_i|}{12} & \text{i=4,5,6,7} \end{cases}.$$

The output probability density function is $p_s(y_i) = \dfrac{y_i}{28}$. Find the transformation for mapping the pixels.

b) Suppose that a digital image is a histogram equalized. What will happen if the image is equalized again? Does the image stay the same or change and why?

c) Compare the Canny edge detector and the Laplacian-of-Gaussian (LoG) edge detector for each of the following -
   i) Which of these operators is/are isotropic and which is/are non-isotropic.
   ii) Describe each operator in terms of the order of the derivatives that it computes.
   iii) What parameters must be defined by the user for each operator?
   iv) Which detector is more likely to produce long, thin contours? Briefly explain.

d) Define the property of energy compaction of a unitary transform and give reasons why this property is useful in image processing.

e) Give examples of 3x3 Prewitt, Sobel and Laplacian spatial masks that approximate local first derivative operators and compare the results arising from their use.

f) Knowing that adding uncorrelated images convolves their histograms, how would you expect the intensity range of the sum of two uncorrelated images to compare with the intensity range of its component images?

g) Explain briefly why uniform quantization of an image may not be optimal.

(7x4)

2.

a) The basic approach used to compute a 5x5 LoG operator involves a mask of the form:

$$h = \begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & -2 & -1 & 0 \\ -1 & -2 & 16 & -2 & -1 \\ 0 & -1 & -2 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{bmatrix}.$$

Find the Fourier transform H (u, v) of this mask in the frequency domain. What type of filter is this LPF, BPF or HPF. And why?

b) Let $f(x,y)$ denote an $M \times N$-point 2-D sequence that is zero outside $0 \le x \le M-1$, $0 \le y \le N-1$. In implementing the even symmetric Discrete Cosine Transform (DCT) of $f(x,y)$, we first relate $f(x,y)$ to a new $2M \times 2N$-point sequence $g(x,y)$.

i) Define the sequence $g(x, y)$.

ii) Comment on the energy compaction property of the DCT as compared to that of the DFT by sketching an example of $f(x, y)$ and $g(x, y)$.

iii) Explain why the even symmetric DCT is more commonly used than the odd symmetric DCT.

(9+9)

3.

a) In the table below you see an example of a three-symbol source with their probabilities.

Symbol	***Probability***
s_1	0.8
s_2	0.02
s_3	0.18

i) Determine the codeword to be assigned to each of the three symbols using Huffman coding.

ii) Determine the entropy, the redundancy and the coding efficiency of the Huffman code for this example.

iii) By observing the above measurements, explain why in this case the extended Huffman code is expected to be more efficient than the conventional Huffman code. Justify your answer without determining the extended Huffman code. What disadvantages, if any, does the extended Huffman code have?

b) Describe briefly what is meant by an inverse filter and how it is related to the Wiener filter.

(12+6)

4.

a) Band pass filters are useful in the enhancement of edges and other high pass image characteristics in the presence of noise. Propose a method to obtain a bandpass filtered version of an image using spatial masks.

b) Suppose that an image is corrupted by random noise. One of the properties of human vision is that the noise is much less visible in the edge regions that in the uniform background regions.

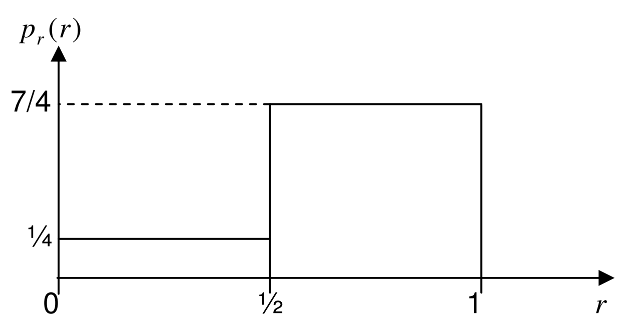
i) Give a possible explanation by considering the local signal-to-noise ratio of the image.

ii) Taking into consideration the above observation, propose a method that uses variable size spatial filters to reduce background noise without blurring the image significantly.

(6+12)

5.

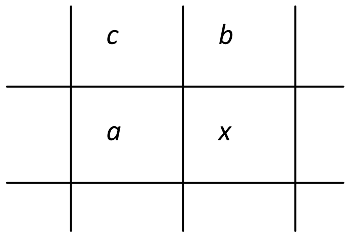
a) Consider an image $f(x, y)$ with intensity r that can be modelled as a sample obtained from the probability density function sketched below:



i) Suppose four reconstruction levels are assigned to quantize the intensity r. Determine these reconstruction levels using a uniform quantizer.

ii) Determine the codeword to be assigned to each of the four reconstruction levels using Huffman coding. Specify what the reconstruction level is for each codeword. For your codeword assignment, determine the average number of bits required to represent r.

b) In lossless JPEG, one forms a prediction residual using previously encoded pixels in the current line and/or the previous line. Suppose that the prediction residual for pixel with intensity x in the following figure is defined as $r = y - x$ where y is the function $y = \frac{a+b}{2}$.



i) Describe the procedure of coding the prediction residual in the Lossless JPEG Standard.

ii) Consider the case with pixel values $a = 100$, $b = 191$ and $x = 180$. Find the codeword of the prediction residual y, knowing that the Huffman code for six is 1110.

(9+9)

6.

a) If a ramp edge (shown below) is used as an edge model, sketch the first and second derivatives

b) Explain how the two types of edge detection methods based on first derivatives and second derivatives work for this type of edge. Describe how these methods work in terms of the edge detector optimality criteria: edge detection, edge localization, and one-response to an edge.

(9+9)

7.

a) An ideal pinhole camera has focal length 5mm. Each pixel is 0.02 mm x 0.02 mm and the image principal point is at pixel (500, 500). Pixel coordinates start at (0, 0) in the upper-left corner of the image. What is the 3x3 camera calibration matrix, K, for this camera configuration?

b) Assuming the world coordinate frame is aligned with the camera coordinate frame (i.e., their origins are the same and their axes are aligned), and the origins are at the camera's pinhole, what is the 3 x 4 matrix that represents the extrinsic, rigidbody transformation between the camera coordinate system and the world coordinate system?

c) Using a projective camera model specialized for this particular scenario, write a general formula that describes the relationship between world coordinates (x), specifying the height of the table top, and image coordinates (u, v), specifying the pixel coordinates where the point of light is detected. Give your answer using homogeneous coordinates and a projection matrix containing variables.

(6+6+6)