National Institute of Technology Rourkela

Department of Computer Science and Engineering

Mid-Term Examination (Feb 2024)

Computer Vision (CS3303), B.Tech. 6th Semester

Time: 90 Minutes **Max. Marks: 30**

Instruction: ***1. In case of any doubt, write your assumptions, write it clearly and continue.***
2. There are 11 questions printed on the paper. Attempt all the questions.
3. Do not use mobile for calculation.

Q1. Write down the two relevant reference books (title, author, publisher, edition, year) for the Computer Vision course. ***[1]***

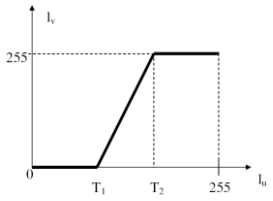
Q2. In a digital camera, the following major components take place while forming an image: ***[4]*** *(1x4)*

a. Aperture
b. Color filter (Demosaicing)
c. Image sensor (CCD/CMOS)
d. Computer Vision (Digital systems)

Explain how the corresponding element for human vision systems works concerning each of the above components.

Q3. Explain the pinhole and perspective projection using a neat diagram. What are the advantages of using lenses instead of pinhole? ***[2]***

Q4. Explain the estimated effect of the following gray level transformation function on the image in the given four scenarios ***[2]*** *(.5x4)*



A. T1=0 and T2 = 128 B. T1=100 and T2 = 200
C. T1=128 and T2 = 255 D. T1=T2=200

Hint: you may write your answer in terms of contrast, brightness, blurring/sharpening, etc.)

Q5. Given the following two input and target images, what will be the resultant image after performing the histogram matching? Write all the steps. ***[4]***

1	7	2	1
2	6	1	2
0	4	3	4
3	2	0	1

Input Image

5	7	2	1
2	6	3	7
6	5	2	4
4	6	0	1

Target Image

Q6. Characterize the three masks by looking at their coefficients. Also describe what the individual masks will do to an image *f* and the applications they normally have. **[2]**

$$A=\frac{1}{6}\begin{bmatrix}0 & 1 & 0\\ 1 & 2 & 1\\ 0 & 1 & 0\end{bmatrix} \quad B=\begin{bmatrix}-1 & 0 & 1\\ -2 & 0 & 2\\ -1 & 0 & 1\end{bmatrix} \quad C=\begin{bmatrix}0 & 1 & 0\\ 1 & -4 & 1\\ 0 & 1 & 0\end{bmatrix} \quad D=\begin{bmatrix}0 & -1 & -1\\ 1 & 0 & -1\\ 1 & 1 & 0\end{bmatrix}$$

Q7. Write down the steps to follow while performing frequency domain filtering ***[2]***

Q8. Consider the three vertical edge models and their corresponding profiles as shown in figure 2. Suppose that we compute the gradient magnitude of each of the three edge models using following operators. Sketch the horizontal intensity profiles of all three gradient images. ***[2]*** (1x2)

A. The Sobel masks.
B. The Laplacian operator (the Laplacian is computed using the 3X3 mask with centre pixel -8)

(Answer the questions without trying to generate the images. Simply provide sketches of the profiles that show what you would expect the profiles of the images to look like.)

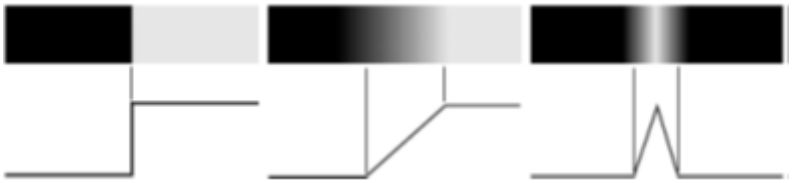


Figure: Step edge, Ramp edge and roof edge

Q9. Consider the convolutional neural network defined by the layers in the left column below. Fill in the shape of the output volume and the number of parameters at each layer. You can write the activation shapes in the format (H, W, C), where H, W, C are the height, width and channel dimensions, respectively. Unless specified, assume padding 1, stride 1 where appropriate. ***[3]***

Notation:

- CONVx-N denotes a convolutional layer with N filters with height and width equal to x.
- POOL-n denotes a n×n max-pooling layer with stride of n and 0 padding.
- FLATTEN flattens its inputs, identical to torch.nn.flatten / tf.layers.flatten
- FC-N denotes a fully-connected layer with N neurons

Layer	Activation Volume Dimensions	Number of parameters
Input	$32 \times 32 \times 3$	0
CONV3-8		
Leaky ReLU		
POOL-2		
BATCHNORM		
CONV3-16		
Leaky ReLU		
POOL-2		
FLATTEN		
FC-10		

Q10. What is the overfitting of the neural network model? Explain popular regularisation methods. ***[1+2]***

Q11. Explain the working principle of the R-CNN model. What is the major limitation of the R-CNN, and how does the fast R-CNN address it? ***[3+1+1]***

End of Paper