INDIAN INSTITUTE OF TECHNOLOGY, KHARAGPUR

**Time** 3 Hrs

*Date* **Full Marks:** 100

Autumn Semester, 2023-24 **UG students**

Sub No. CS40019 Sub. Image Processing

**No. of Students** 56

Section A (All questions)

1. Name the geometric entities sufficient to define parallel and perspective projections, respectively from a 3D real space to a 2D real space. How many independent parameters are required to uniquely define each of these projections. 3+2

   Ans. Parallel projection: Direction of parallel rays and image plane.

   Perspective Projection: Center of projection and image plane

   For parallel projection: Direction: 2 (two direction cosines, the other will be obtained from them) + image plane (3, ax+by+cz+d=0): Total: 5

   For perspective projection: Center of projection: 3 (Coordinates in 3-D) and the image plane (3): Total: 6

2. Consider the following transformation matrix of color spaces (from RGB to XYZ)

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.49 & 0.31 & 0.2 \\ 0.18 & .81 & .01 \\ 0 & 0.01 & .99 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

   Answer the following.

   (a) Convert a color vector $(60, 50, 20)$ in the RGB space to a vector in the normalized $x - y$ chromaticity space. 2

   (b) Explain how hue can be measured of the above color vector. 1

   (c) Suppose you have another color vector $(50, 60, 40)$. Which one of these two vectors is more reddish? Justify. 2

   Ans.: (a)

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.49 & 0.31 & 0.2 \\ 0.18 & .81 & .01 \\ 0 & 0.01 & .99 \end{bmatrix} \begin{bmatrix} 60 \\ 50 \\ 20 \end{bmatrix}$$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 48.9 \\ 51.5 \\ 20.3 \end{bmatrix}$$

Hence,

$$\begin{bmatrix} nX \\ nY \end{bmatrix} = \begin{bmatrix} 0.41 \\ 0.43 \end{bmatrix}$$

(b) My measuring the angle w.r.t. a reference axis for example X-axis.

(c) Measures the angles w.r.t. reference axis. The point with lower angle is more reddish, as the reference direction makes closer angle with the pure red chromatic point. For (50,60,40) its normalized chromaticity point is: (0.34, 0.39).

For (60,50,20), hue angle: $tan^{-1}\frac{.43}{.41} = 0.81$ rad For (50,60,40), hue angle: $tan^{-1}\frac{.39}{.34}$ =0.90 rad

Hence, (60,50,20) is more reddish.

3. Consider the octagonal digital distance defined by the periodic neighborhood sequences $\{1, 1, 2\}$, where in the sequence 1 denotes type-1 (or 4-Neighbor) neighborhood and 2 denotes type-2 (or 8-Neighbor) neighborhood. Given a radius of 12 compute the perimeter of the convex polygon formed by the vertices of the disc of this distance function. 5

   Ans. The vertices of the disc of a radius are given by: permutaions of $(\pm R, \pm\frac{b}{a+b}R)$, where $a$ and $b$ are numbers of type1 and type2 neighborhoods, respectively. In this case, $a = 2$, and $b = 1$.

   Hence, the vertices of radius 12 are given by $(\pm 12, \pm 4)$.

   Take two consecutive sides between (i) $(4, 12)$ and $(12, 4)$, and (ii) $(12, 4)$ and $(12, -4)$. Their lengths are respectively $8\sqrt{2}$, and 8. So the perimeter is ;$4 \times (8\sqrt{2} + 8) = 77.2548$.

Note: In the slide there was a mistake. Instead of $\pm\frac{b}{a+b}R$, it is written as $\pm\frac{a}{a+b}R$. Full marks to be given if someone answers using this expression.

4. Given a sequence $x(n), n = 0, 1, 2, .., N-1$, of length $N$ define its Discrete Fourier Transform (DFT). Define the corresponding Inverse Discrete Fourier Transform (IDFT). Suppose the sequence is circularly shifted by 3 samples, derive the ratio of the $k$th transform coefficient of the modified sequence with that of $x(n)$ 2+2+2

   Ans.

   DFT: $X(k) = \sum_{n=0}^{N-1} x(i)e^{-j2\pi\frac{k}{N}n}$, for $k = 0, 1, ..N-1$.

   IDFT: $x(n) = \frac{1}{N}\sum_{k=0}^{N-1} X(k)e^{j2\pi\frac{k}{N}n}$, for $n = 0, 1, ..N-1$.

   The ratio for 3 sample circular shift: $e^{-j2\pi\frac{k}{N}3} = e^{-j\frac{6\pi k}{N}}$.

5. Derive the mask of size $3 \times 3$ implementing the Laplacian operator, where the gradients along horizontal and vertical directions are weighed $\sqrt{2}$ times with respect to those in the diagonal directions. Compute the Laplacian operated value at the central pixel (having the value 78 and shown underlined) of $3 \times 3$ image block as given in the following. 4+2

| | | |
|---|---|---|
| 85 | 80 | 100 |
| 140 | <u>78</u> | 46 |
| 130 | 27 | 68 |

   Ans. Laplacian Horiziontal-Vertical:

| | | |
|---|---|---|
| 0 | 1 | 0 |
| 1 | $-4$ | 1 |
| 0 | 1 | 0 |

   Laplacian Diagonal:

|  |  |  |
|---|---|---|
| 1 | 0 | 1 |
| 0 | $-4$ | 0 |
| 1 | 0 | 1 |

|  |  |  |
|---|---|---|
| 1 | $\sqrt{2}$ | 1 |
| $\sqrt{2}$ | $-4(1+\sqrt{2})$ | $\sqrt{2}$ |
| 1 | $\sqrt{2}$ | 1 |

A weighted combination in the ratio of $\sqrt{2}:1$, will provide the following mask (the values to be divided by $(2+1)$.

The value of Lapalcian operation: $\frac{\sqrt{2}(80+46+27+140))+(85+100+130+68)-4(1+\sqrt{2})78}{1+\sqrt{2}}$ $= 18.28$.

6. Suppose that an image $g(x,y)$ of size $M \times N$ has been captured under degradation by a linear filter whose Fourier spectrum is given by $H(u,v), 0 \leq u \leq M-1, 0 \leq v \leq N-1$. Express the power spectra of the undegraded image as a function of Fourier Spectra of $g(x,y)$ and the filter $H(u,v)$. Supposing that the system is further contaminated by a random additive noise with the power spectra of $S_\eta(u,v)$, how this relationship would be modified. 3+2

   Ans.

   Let the Fourier transform of undegraded image $f(x,y)$ be $F(u,v)$ and that of $g(x,y)$ be $G(u,v)$.

   Hence, $G(u,v) = H(u,v)G(u,v)$. So, $||G(u,v)||^2 = G(u,v)^*G(u,v) = (H(u,v)F(u,v)) * (H(u,v)F(u,v)) = ||H(u,v)||^2||F(u,v)||^2$. Hence,

$||F(u,v)||^2 = \frac{||G(u,v)||^2}{||H(u,v)||^2}$.

For additive noise, it would be $||F(u,v)||^2 = \frac{||G(u,v)||^2 - S_\eta(u,v)}{||H(u,v)||^2}$

7. Compute the mean shift at the central pixel (shown underlined) assuming that the kernel is a truncated Gaussian function (zero mean and s.d.=1 along both horizontal and vertical directions) with width 2 following the computation of mean shift segmentation algorithm. 8

| | | | | |
|---|---|---|---|---|
| 60 | 110 | 12 | 110 | 60 |
| 50 | 100 | 11 | 100 | 50 |
| 40 | 90 | <u>10</u> | 90 | 40 |
| 30 | 80 | 9 | 80 | 30 |
| 20 | 70 | 8 | 70 | 20 |

Ans.

The proportional Gaussian mask of 5$x$5 with zero mean and s.d. 1 with origin at the center is given by $G(x,y) = e^{-\frac{x^2+y^2}{2}}$.

| | | | | |
|---|---|---|---|---|
| .02 | .08 | .14 | .08 | .02 |
| .08 | .37 | .61 | .37 | .08 |
| .14 | .61 | <u>1</u> | .61 | .14 |
| .08 | .37 | .61 | .37 | .08 |
| .02 | .08 | .14 | .08 | .02 |

Next the weight for each position is obtained by multiplying the Gaussian weights with the functional values.

Sum of weights: 324

Sum of columns: At x= -1 and 1, Sum=135, At x=1: Sum=135, x=2 and -2, sum=13.6

Hence mean of x-cooridnate: 0

| | | | | |
|---|---|---|---|---|
| 1.2 | 8.8 | 1.1 | 8.8 | 1.2 |
| 4 | 37 | 5.5 | 37 | 4 |
| 5.6 | 54.9 | $\underline{10}$ | 54.9 | 5.6 |
| 2.4 | 29.6 | 6.7 | 29.6 | 2.4 |
| .4 | 5.6 | 1.7 | 5.6 | .4 |

Sum of rows: y=1, Sum=87.5; y=-1, Sum=70.7; y=2 Sum=21.1, y=-2, sum= 13.7

Mean of y-cooridnate: ((87.5 - 70.7) +2 (21.1-13.7))/324 = 32.4/324 = 0.1

Hence mean shift vector is : (0, 0.1)

Section B (Any three)

8. Consider an image of size $4 \times 4$ represented in the form of a 2-D matrix $A$. The coordinates of pixels of the image are given by respective row and column numbers of the matrix. The row index of the first row starts with 1 and similarly the column index also starts from 1.

$$A = \begin{bmatrix} 50 & 50 & 51 & 50 \\ 51 & 58 & 58 & 70 \\ 50 & 70 & 60 & 60 \\ 50 & 58 & 51 & 68 \end{bmatrix}$$

Answer the following:

(a) Compute the local binary pattern (LBP) at $A(3, 2)$ assuming the scanning order of neighboring pixels anticlockwise and starts with the leftmost neighbor. Write the algorithm to compute the pattern. 2+3

Ans. LBP centering $A(3, 2)$ is shown here:

$$\begin{bmatrix} * & * & * & * \\ 0 & 0 & 0 & * \\ 0 & \mathbf{70} & 0 & * \\ 0 & 0 & 0 & * \end{bmatrix}$$

Hence the value of LBP at A(3,2): 0

The algorithm should be properly written, ensuring the scanning order as specified in the question.

(b) Explain how a feature vector could be formed to represent the above image using LBPs. State two transformations (with justification) for which this feature vector is invariant. 3+2

Ans. Formation of histogram of LBP values should be discussed. It is invariant to background illumination (addition of a constant value to each of them) and also with the variation of contrast (scaling of values).

(c) Considering formation of the co-occurrence matrix from the above image, where two pixels are at relative offset of a vector $(1, -1)$. List the non-zero bins of the 2-D histogram of the co-occurrence matrix with their frequency of occurrences. 5

Ans. Offset vector is downward diagonal with unit changes. Hence the pairing of values and the respective bins in the array are as follows:

$\{(50, 58) : 3; (51, 70) : 2; (58, 60) : 2; (70, 50) : 1; (60, 68) : 1\}$

(d) Explain how a feature vector could be formed exploiting the statistics of the co-occurrence matrix for representing the image. 5

Ans. Explanation should be provided by mentioning different measures over the normalized distribution of coccurrence pairs: Such as mean, s.d.s, energy, etc.

9. (a) Given a set of $N$ complex basis vectors $B = b_1, b_2, \ldots, b_N$ (called *base*) for linear expansion of a vector of dimension $N$, form the forward transformation matrix and answer the following. 2

Ans. Let the complex conjugate of a basis vector $b_i$ be denoted by $b_i^*$. The transformation matrix formed as follows:

$$\begin{bmatrix} b_1^* \\ b_2^* \\ . \\ b_N^* \end{bmatrix}$$

i. State the .properties related to the basis vectors that required to be satisfied for the the following cases. $5 \times 2$

A. The base is orthogonal.

Ans. $b_i^*.b_j = 0$, for $i \neq j$, else *positive*.

B. The base is complete.

Ans. The transformation matrix is invertible. So that any arbitrary vector is fully reconstructible from its forward transform.

C. The base is non-orthogonal but complete.

Ans. The transformation matrix is invertible. It is sufficient to fully reconstruct the vector from its transformation, though there may exist a pair of distinct basis vectors for which the inner product is non-zero.

D. The base is orthonormal.

Ans. The base is orthogonal and magnitude of each of them is unity.

E. The transformation matrix is invertible.

Ans. The determinant of the square matrix is non-zero.

ii. Prove that, $B$ is complete if $B$ is orthogonal. 2

Ans. Since there are exactly $N$ basis vectors and none of the rows of the transformation matrix can be expressed as linear combination of the remaining rows (due to orthogonality), the determinant of the square matrix is not zero and hence the matrix is invertible.

iii. Explain how computation of inverse transforms gets simplified if $B$ is orthonormal. 2

Ans. Inversion of transformation matrix is the same as its transposition (of a Real matrix) or Hermitian Transposition for a complex matrix. Hence inversion becomes simpler.

(b) Explain how DC coefficients of blocks of images in JPEG compression standard are computed and encoded. 1+3

Ans. The DC of the starting block followed by differential offsets of subsequent blocks in a predefined scanning order are encoded using Huffman encoding.

10. (a) Consider the problem of separating foreground from background in an image and modeling the probability of a pixel value as a mixture of Gaussians, each expressing class likelihood of a pixel value. Formulate the segmentation problem as a Bayesian classification problem. How the parameters of the distributions are computed and updated iteratively using expectation maximization method? 4+6

Ans. In a Gaussian Mixture Model the probability of a feature vector $x$ is expressed as follows: $p(x) = \sum_{i=1}^{K} p(x|\omega_i)p(\omega_i)$, where $\omega_i$ is the $i$th class in the mixture of $K$ classes. The class likelihoods are modeled by a Gaussian distribution whose parameters are mean vectors and covariance matrices. For separating foreground and background the number of class is 2 and the segmentation problem using GMM is to estimate the parameters of these two classes so that the observed distribution in the feature space closely matches with the model.

For estimating the parameters it is required to estimate the posterior probability of a feature value given a class. The mean and variances of the class are estimated using the posterior probabilities as the weights. Thus the problem is similar to a Bayesian classification problem at every observed feature value.

The description of Expectation-Maximization algorithm should include the expressions for parameter updates.

(b) Consider a color image whose pixels are given in RGB color space. How many possible colors could be there in an image given each component is represented by 8 bit. Write an algorithm by adapting $k$-means clustering algorithm for quantizing the color vectors in $M$ number of representative colors. Formulate a criteria for choosing an appropriate value of $M$ for an image. 2+5+3

Ans. Number of colors: $2^{24}$.

Description of K-Means algorithm in a 3-D color space.

Any reasonable criteria would be okay. For example, keeping the sum of squares of distances from means within a threshold by choosing $M$ iteratively.

11. (a) Suppose an image is contaminated by a periodic noise of spatial frequency of 5 pixels per cycle in the horizontal direction in an image of size $420 \times 640$. Design a zero phase notch filter to restore the image. Provide the computational steps for filtering an image in the transform domain using the filter. 5+5

Ans. Notch filter cut-off spatial frequency: $u_0, v_0) = (0, \frac{5}{640})$. Hence an $n$-th order centered Buuterrworth filter response is given by

$H(u,v) = \frac{1}{1+[\frac{A}{D_0(u,v)}]^{2n}}$, where $D_0(u,v) = \sqrt{(u-210)^2 + (v - \frac{5}{640} - 320)^2}$, and $A$ is a constant.

The filtering algorithm should describe necessary pixel level transformation of input image in the spatial domain for centering its transform in the frequency domain, followed by application of convolution multiplication properties, and then applying IDFT.

(b) Describe the computational steps for adaptive restoration of an image using local statistics of noise and image. How these statistics could be computed? 6+4

Ans. The algorithm should be described by illustrating three conditions on noise and signal variances. Estimation of noise variance from seemingly constant patches to be discussed.

12. (a) Define morphological dilation and erosion operations on gray scale images using a non-flat structuring element. 6

Ans.
$(f \oplus b)(s,t) = \max_{(x,y)}\{f(s-x, t-y) + b(x,y) | (s-x, t-y) \epsilon D_f; (x,y) \epsilon D_b\}$
$(f \ominus b)(s,t) = \min_{(x,y)}\{f(s+x, t+y) - b(x,y) | (s+x, t+y) \epsilon D_f; (x,y) \epsilon D_b\}$
These need to be explained with examples.

(b) Describe an algorithm for sharpening images using above morphological operators. 4

Ans. It should be computed through computation of morphological gradients by subtracting an eroded image from its dilated image using the same structural element. and then adding a part of the computed values with the original image for highlighting edges. This should be described in detail.
$f + \lambda((f \oplus b) - (f \ominus b))$ where $\lambda$ is a positive constant.

(c) Define top hat transformation using dilation and erosion. Explain how it could be used for extracting foreground objects. 5+5

Ans. Top hat transformation -
$T_{hat}(f) = f - (f \circ b)$
The opening operation may be shown with dilation and erosion. Then discuss top-hat thresholding with examples.