

COMPUTER VISION
(CSEN 4236)

Time Allotted : 2½ hrs

Full Marks : 60

Figures out of the right margin indicate full marks.

Candidates are required to answer Group A and any 4 (four) from Group B to E, taking one from each group.

Candidates are required to give answer in their own words as far as practicable.

Group – A

1. Answer any twelve:

12 × 1 = 12

Choose the correct alternative for the following

- (i) What is the process of moving a filter mask over the image and computing the sum of products at each location called as?
 - (a) Convolution
 - (b) Correlation
 - (c) Linear spatial filtering
 - (d) Nonlinear spatial filtering
- (ii) What is the difference between Convolution and Correlation?
 - (a) Image is pre-rotated by 180 degree for Correlation
 - (b) Image is pre-rotated by 180 degree for Convolution
 - (c) Image is pre-rotated by 90 degree for Correlation
 - (d) Image is pre-rotated by 90 degree for Convolution
- (iii) Which of the following is the first step in Canny Edge Detection Algorithm?
 - (a) Noise Reduction
 - (b) Finding Intensity Gradient of the Image
 - (c) Non-maximum Suppression
 - (d) Hysteresis Thresholding
- (iv) What is the formula to calculate the angle in the Gradient calculation of canny Edge detector?
 - (a) $\theta(x, y) = \arctan(I_y - I_x)$
 - (b) $\theta(x, y) = \arctan(I_y * I_x)$
 - (c) $\theta(x, y) = \arctan(I_y / I_x)$
 - (d) $\theta(x, y) = \arctan(I_y + I_x)$
- (v) Which technique is used to detect and track objects in a sequence of frames?
 - (a) Optical flow
 - (b) Template matching
 - (c) Harris corner detection
 - (d) Scale-invariant feature transform (SIFT)
- (vi) Which technique is used to detect and recognize objects based on their geometric shapes?
 - (a) Contour detection
 - (b) Texture analysis
 - (c) Blob detection
 - (d) Edge detection
- (vii) The 8-point algorithm is used for
 - (a) Estimating depth from stereo images
 - (b) Computing the fundamental matrix from point correspondences
 - (c) Rectifying images before processing
 - (d) Reducing noise in image sequences

- (viii) Which of the following is an intrinsic camera parameter?
 (a) Rotation matrix (b) Translation vector
 (c) Focal length (d) Camera pose
- (ix) The image brightness constancy equation is a key assumption in optical flow estimation. It states that:
 (a) The brightness of a pixel does not change between consecutive frames
 (b) Objects closer to the camera always appear brighter
 (c) Motion estimation works only for grayscale images
 (d) The gradient of an image is always constant
- (x) Statistical filtering in motion tracking is used to
 (a) Reduce the number of frames needed for object tracking
 (b) Estimate the true state of a moving object from noisy observations
 (c) Detect motion blur in an image sequence
 (d) Segment objects based on color information

Fill in the blanks with the correct word

- (xi) The measure of the amount of light energy per unit area reaching the image sensor is called _____.
- (xii) The Eigenfaces approach for face recognition is based on _____, which reduces the dimensionality of image data while preserving important features.
- (xiii) The Harris corner detector identifies corners by looking for regions in an image where the _____ changes significantly in multiple directions.
- (xiv) The intrinsic parameters of a camera include focal length, principal point, and _____, which accounts for axis misalignment.
- (xv) The line connecting the projection centers of two cameras is called the _____, and the points where it intersects the image planes are known as epipoles.

Group - B

2. (a) A 3×3 sharpening filter (high-pass filter) is applied to the following grayscale image patch using convolution:

$$\begin{bmatrix} 10 & 20 & 30 \\ 40 & 50 & 60 \\ 70 & 80 & 90 \end{bmatrix}$$

Apply the following sharpening filter to the given 3×3 image patch and compute the output pixel value at the center.

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

[[C05](Apply/IOCQ)]

- (b) If we instead apply a 3×3 averaging (smoothing) filter, what would be the output pixel value at the center? Use the following kernel.

$$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

[[C05](Apply/IOCQ)]

- (c) Explain how the difference between the outputs of the sharpening and smoothing filters affects edge detection in images.

[[C01](Understand/LOCQ)]

5 + 4 + 3 = 12

3. (a) A 2D point P(4,6) undergoes the following transformations. The point is rotated 90° counterclockwise about the origin. Compute the new coordinates of the point after rotation. [[CO5](Apply/IOCQ)]
- (b) The rotated point is then scaled by a factor of 2 in the x-direction and 3 in the y-direction and finally translated by (5, -2). [[CO5](Apply/IOCQ)]
- (c) Explain why the rigid transformation preserves distances while the affine transformation does not. [[CO1](Understand/LOCQ)]

4 + 5 + 3 = 12

Group - C

4. (a) Illustrate the algorithm for Region Growing based segmentation technique. [[CO2](Understand/IOCQ)]
- (b) A computer vision system is used to detect circular and elliptical objects in an image using conic section fitting. The detected objects follow the general conic equation $Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$ where A, B, C, D, E and F are the coefficients defining the conic section. How can the values of A, B, C determine whether the detected shape is a circle, an ellipse, a parabola, or a hyperbola? Provide the mathematical condition for each case. [[CO1](Understand/LOCQ)]
- (c) Explain the differences between algebraic distance and Euclidean distance in the context of ellipse and conic sections fitting. [[CO1](Understand/LOCQ)]
- (d) Why is Euclidean distance often preferred in high-precision applications such as medical imaging and object recognition? [[CO1](Understand/IOCQ)]

5 + 3 + 2 + 2 = 12

5. (a) Explain the difference between edge detection and corner detection. What role do these techniques play in feature extraction? Provide an example where corner detection is preferable over edge detection. [[CO1](Understand/IOCQ)]
- (b) What are SIFT and HOG descriptors? Compare them based on their robustness to scaling, rotation, and illumination changes. Which one is more suitable for pedestrian detection, and why? [[CO1](Understand/LOCQ)]
- (c) What are active contours (snakes) in computer vision? How do they help in detecting object boundaries? Also, explain the concept of shape context descriptors and how they assist in object recognition. [[CO5](Apply/IOCQ)]

4 + 4 + 4 = 12

Group - D

6. (a) Explain the concept of epipolar geometry in stereo vision. What is the role of the epipolar constraint in computing depth information? [[CO1](Understand/LOCQ)]
- (b) Describe the 8-point algorithm for estimating the fundamental matrix from a set of corresponding points in two images. [[CO2](Understand/IOCQ)]
- (c) Why is perspective projection more accurate than weak perspective or affine projection in real-world applications? Provide two practical examples where perspective projection is necessary. [[CO2](Understand/IOCQ)]

3 + 5 + 4 = 12

7. (a) What is the role of the fundamental matrix in epipolar geometry? Explain how it helps in finding corresponding points between two images. [[C03](Understand/LOCQ)]
- (b) Two cameras observe the same 3D point P from different viewpoints. The corresponding image points in the two images are given as

$$x = (200, 150, 1)^T, x' = (220, 130, 1)^T$$

The fundamental matrix for this stereo camera system is

$$F = \begin{bmatrix} 0.0001 & -0.002 & 0.3 \\ 0.002 & 0.0001 & -0.4 \\ -0.25 & 0.35 & 1 \end{bmatrix}$$

Verify if the given points satisfy the epipolar constraint $x'^T F x = 0$

[[C04](Analyse/HOCQ)]

- (c) Compute the epipolar line equation in the second image corresponding to the point x in the first image using $l' = Fx$. [[C05](Apply/LOCQ)]

5 + 4 + 3 = 12

Group - E

8. (a) What is optical flow, and how does it help in motion estimation? Explain the assumptions made in computing optical flow. [[C04](Analyse/IOCQ)]
- (b) Two well-known methods for computing optical flow are Lucas-Kanade and Horn-Schunck. Compare these two methods in terms of assumptions, advantages, and limitations. In which scenarios is Lucas-Kanade preferred over Horn-Schunck, and vice versa? [[C04](Analyse/IOCQ)]

- (c) The optical flow constraint equation (OFCE) is given by

$$I_x u + I_y v + I_t = 0$$

where I_x, I_y, I_t are partial derivatives of the image intensity, and (u, v) represent the flow velocities. Explain the significance of this equation in motion estimation. What is the aperture problem, and how does it limit the computation of optical flow?

[[C05](Apply/IOCQ)]

3 + 4 + 5 = 12

9. (a) Explain the difference between linear motion models (used in Kalman filters) and non-linear motion models (used in Extended Kalman Filters). Provide one real-world example where EKF is necessary. [[C02](Understand/LOCQ)]
- (b) A drone is moving in 2D space with an initial position of $(x, y) = (0, 0)$ and velocity $(v_x, v_y) = (5, 5)$ m/s. Using a constant velocity model, predict its position after 2 seconds using the Kalman filter's state transition matrix.

$$A = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where $\Delta t = 2$ seconds.

[[C04](Analyse/LOCQ)]

- (c) Discuss three key limitations of Kalman filtering in real-world motion tracking applications. [[C04](Analyse/IOCQ)]

4 + 5 + 3 = 12

Cognition Level	LOCQ	IOCQ	HOCQ
Percentage distribution	36.46	59.37	4.17