ABHYASMITRA.IN

COMPREHENSIVE UNIVERSITY STUDY NOTES

COMPUTER VISION

Unit IV — Feature Extraction and Image Segmentation

Topics Covered in this Unit

Feature Extraction Fundamentals: Low-Level vs High-Level Features & Invariance Requirements
Corner Detection: Structure Tensor Matrix M, Harris Corner Detector (R Response) & Shi-Tomasi Criterion
Texture Analysis: Structural vs Statistical Textures, 1st-Order Metrics & 2nd-Order GLCM Features
Frequency-Based Texture Analysis: Gabor Wavelets & Multi-Scale Texture Characterization
Image Segmentation: Global Thresholding, Local Adaptive Thresholding & Otsu's Optimal Method (Between-Class Variance)
Region-Based Segmentation: Seeded Region Growing, Region Splitting and Merging (Quadtree Partitioning)
Clustering & Morphology: K-Means Segmentation, Watershed Algorithm, Dilation, Erosion, Opening & Closing
Shape Representation & Boundary Descriptors: Freeman Chain Codes, Shape Numbers & Fourier Descriptors
Region Descriptors & Moment Invariants: Compactness, Eccentricity, Euler Number & Hu's 7 Invariant Moments
Shape Matching Techniques: Hausdorff Distance, Chamfer Matching & Structural Deformable Templates
High-Yield University Exam Question Bank with Detailed Model Answers & Unit Summary

Compiled in a structured, easy-to-understand, textbook style format
Specially curated for B.Tech / B.E. / BCA / MCA / B.Sc Computer Science, AI & Data Science Students

COURSE SYLLABUS & MAPPING

[UNIT IV SYLLABUS — PRESCRIBED UNIVERSITY CURRICULUM]

This section contains the official syllabus for Unit IV. Verify with your university curriculum mapping.

Unit IV Syllabus Breakdown:

IV Feature Extraction and Image Segmentation: Feature extraction concepts: Edge-based features, Corner detection (Harris, Shi-Tomasi). Texture analysis: Types of Texture, Statistical Texture Features, Texture Representation Methods, Frequency-Based Texture Analysis (Conceptual), Applications of Texture Analysis. Segmentation techniques: Thresholding (global, adaptive, Otsu), Region-based methods, Clustering Morphological operations. Shape representation & descriptors: Boundary Descriptors Region Descriptors Moment Invariants Fourier Descriptors, Shape Matching Techniques

• Course Code: ____________________ • Semester / Year: ____________________

TABLE OF CONTENTS

4.1 Feature Extraction Concepts & Corner Detection

Definition: Interest Points & Corners

A Corner (or interest point) is an image location where the local gradient exhibits large variations in two distinct orthogonal directions. Corners provide highly distinctive, local 2D spatial information that is invariant to translation, rotation, and illumination changes, making them foundational for feature tracking and image stitching.

The Harris Corner Detector (Harris & Stephens, 1988)

The Harris corner detector considers a local window W centered at (x,y) and evaluates the change in intensity E(u,v) when the window is shifted by a small displacement (u,v):

$E(u,v) = \sum[\ (x,y) \in W\ ]\ w(x,y) \times [\ f(x+u, y+v) - f(x,y)\ ]^2$

Using first-order Taylor Series expansion, E(u,v) simplifies into a quadratic form governed by the Structure Tensor Matrix M:

$E(u,v) \approx [u, v] \times M \times [u, v]^T$

where $M = \sum w(x,y) \times [[\ I_x^2, I_x I_y\ ], [\ I_x I_y, I_y^2\ ]]$

- • **Eigenvalue Analysis of Structure Tensor M (λ_1, λ_2):**
  - **1. Flat Region:** Both λ_1 and λ_2 are very small ($\lambda_1 \approx 0, \lambda_2 \approx 0$). Intensity changes negligibly in all directions.
  - **2. Edge Region:** One eigenvalue is large while the other is small ($\lambda_1 \gg \lambda_2$ or $\lambda_2 \gg \lambda_1$). Intensity varies in only one direction.
  - **3. Corner Region:** Both eigenvalues are distinctly large and of similar magnitude ($\lambda_1 \gg 0$ and $\lambda_2 \gg 0$). Intensity varies sharply in all directions.
- • **Harris Corner Response Measure R:** Avoids expensive explicit eigenvalue computation:

$R = \det(M) - k \times (\mathrm{trace}(M))^2 = (\lambda_1 \lambda_2) - k \times (\lambda_1 + \lambda_2)^2$

where k is an empirical tuning constant (typically $k \in [0.04, 0.06]$). Classification rules:

  - • $R > 0$ (Large Positive): Corner point.
  - • $R < 0$ (Large Negative): Edge pixel.
  - • $|R| \approx 0$ (Near Zero): Flat homogeneous region.

The Shi-Tomasi Corner Detector (Good Features to Track, 1994)

- • **Mathematical Criterion:** Shi and Tomasi demonstrated that tracking stability is strictly governed by the minimum eigenvalue of matrix M:

$R_{ShiTomasi} = \min(\lambda_1, \lambda_2)$

- • **Decision Rule:** If $\min(\lambda_1, \lambda_2) > \lambda_{min}$ (user-specified threshold), the pixel is marked as an optimal corner. Shi-Tomasi corners provide superior repeatability in optical flow tracking (Lucas-Kanade) and SLAM systems.

4.2 Texture Analysis & Statistical Descriptors

Definition: Texture in Computer Vision
Texture refers to the spatial arrangement, repetition, and distribution of brightness variations, roughness, regularity, and granular patterns across an image region. Textures are categorized into Structural Textures (repeating regular geometric primitives or texels) and Statistical Textures (random, stochastic visual noise like grass, sand, fabric).

1. First-Order Histogram Texture Statistics

Measures global intensity distribution without considering relative spatial positioning:

- • **Mean (μ):** $\mu = \sum_{i=0}^{L-1} z_i \times p(z_i)$ (Average brightness)
- • **Standard Deviation / Variance (σ^2):** $\sigma^2 = \sum (z_i - \mu)^2 p(z_i)$ (Measure of texture contrast/roughness)
- • **Skewness (μ_3):** $\mu_3 = (1/\sigma^3) \sum (z_i - \mu)^3 p(z_i)$ (Asymmetry of intensity histogram)
- • **Kurtosis (μ_4):** $\mu_4 = (1/\sigma^4) \sum (z_i - \mu)^4 p(z_i)$ (Peakiness / flatness of distribution)

2. Second-Order Statistics: Gray-Level Co-occurrence Matrix (GLCM)

The GLCM $P(i,j \mid d, \theta)$ counts how often a pixel with intensity i occurs in a specific spatial relationship (distance d, angle $\theta \in \{0°, 45°, 90°, 135°\}$) to a pixel with intensity j:

GLCM Texture Metric	Mathematical Formulation	Physical & Perceptual Interpretation
Contrast	$\sum\sum (i - j)^2 P(i,j)$	Measures local intensity variations; high for sharp, distinct textures.
Energy (Angular Second Moment)	$\sum\sum P(i,j)^2$	Measures textural uniformity; equals 1 for a completely homogeneous image.

GLCM Texture Metric	Mathematical Formulation	Physical & Perceptual Interpretation
Entropy	$- \sum \sum P(i,j) \log_2(P(i,j))$	Measures statistical randomness/disorder; high for complex stochastic textures.
Homogeneity (Inverse Difference)	$\sum \sum P(i,j) / (1 + \|i - j\|)$	Measures spatial closeness of elements to GLCM diagonal.
Correlation	$\sum \sum (i - \mu_x)(j - \mu_y) P(i,j) / (\sigma_x \sigma_y)$	Measures linear dependency of gray levels of neighbor pixels.

3. Frequency-Based Texture Analysis (Gabor Filters)

- • **Concept:** Gabor filters act as localized multi-scale, multi-orientation band-pass filters mathematically modeled after mammalian primary visual cortex (V1) receptive fields. They convolve a 2D Gaussian envelope with a sinusoidal plane wave, isolating specific spatial frequencies and structural orientations.

4.3 Image Segmentation Techniques

Formal Definition: Image Segmentation

Image Segmentation is the process of partitioning a digital image into multiple disjoint, non-overlapping homogeneous regions $\{R_1, R_2, ..., R_n\}$ *such that* $\cup R_i = R$, $R_i \cap R_j = \emptyset$ *for* $i \neq j$, *and a logical uniformity predicate* $P(R_i) = TRUE$ *for each region while* $P(R_i \cup R_j) = FALSE$ *for adjacent regions.*

1. Thresholding Techniques

- • **Global Fixed Thresholding:** A single constant threshold T binarizes the entire image: $g(x,y) = 1$ if $f(x,y) \geq T$, and 0 if $f(x,y) < T$. Fails under non-uniform illumination.
- • **Local / Adaptive Thresholding:** Computes a dynamically varying threshold $T(x,y)$ for every pixel based on the mean or Gaussian-weighted intensity within a local moving window (e.g., 15×15). Highly robust against uneven shadows and document illumination gradients.
- • **Otsu's Optimal Thresholding Algorithm (Nobuyuki Otsu, 1979):** An automated non-parametric method that finds the optimal threshold T^* that maximizes the Between-Class Variance $\sigma_B^2(T)$ (equivalent to minimizing within-class variance):

$$\sigma_B^2(T) = \omega_1(T) \times \omega_2(T) \times [\mu_1(T) - \mu_2(T)]^2$$

where ω_1, ω_2 are probabilities of background/foreground classes and μ_1, μ_2 are their class means. Exhaustively tests all $T \in [0, L-1]$ to find $T^* = \arg\max \sigma_B^2(T)$.

2. Region-Based Segmentation Methods

- • **Seeded Region Growing:** Starts with user-selected or automated seed pixels and iteratively appends neighboring pixels if they satisfy a similarity predicate (e.g., $|f(\text{neighbor}) - \mu_{region}| \leq$ Threshold).
- • **Region Splitting and Merging (Quadtree Algorithm):** Subdivides the entire image into four quadrant blocks. If a quadrant fails the homogeneity predicate $P(R_i)$, it is recursively split into 4 sub-quadrants. Adjacent quadrants satisfying $P(R_i \cup R_j)$ are subsequently merged into unified segments.

3. Clustering & Morphological Segmentation

- • **K-Means Image Segmentation:** Clusters image pixels in feature space (e.g., [L*, a*, b*] color or [R, G, B, x, y] spatial-color) into K distinct centroid clusters by minimizing squared Euclidean error.
- • **Watershed Algorithm:** Treats image gradient magnitude as a 3D topographic relief surface. Simulates water rising from regional minima to build watershed catchment basins, placing dam lines (boundaries) where water from distinct basins meets. Over-segmentation is controlled via marker-controlled watersheds.
- • **Morphological Operations:**
  - • **Dilation:** Expands foreground boundaries; bridges small holes and gaps.
  - • **Erosion:** Shrinks foreground boundaries; strips away isolated noise spikes.
  - • **Opening (Erosion → Dilation):** Eliminates small bright foreground noise while preserving object size.
  - • **Closing (Dilation → Erosion):** Fills small dark interior holes and connects severed boundaries.

4.4 Shape Representation, Descriptors & Moment Invariants

1. Boundary Descriptors

- • **Freeman Chain Codes:** Represents a discrete boundary curve as a sequence of directional unit vectors connecting adjacent boundary pixels. 4-directional (0:East, 1:North, 2:West, 3:South) or 8-directional chain codes. Invariant to translation; rotation invariance is achieved by using the first difference of the chain code normalized to minimum integer value.
- • **Fourier Descriptors:** Represents a 2D closed boundary as a 1D complex coordinate sequence $s(k) = x(k) + j\,y(k)$. The Discrete Fourier Transform yields complex Fourier descriptors. Truncating high-frequency descriptors yields smooth approximations of shape geometry, providing complete invariance to translation, scale, and rotation.

2. Region Descriptors

- • **Compactness (Circularity):** Compactness = $(\text{Perimeter})^2 / \text{Area}$. (Equals 4π for a perfect circle; higher for complex elongated shapes).
- • **Eccentricity:** Ratio of the length of the major axis to the minor axis of an equivalent fitted bounding ellipse.
- • **Euler Number:** $E = C - H$ (where C is the number of connected components and H is the number of interior holes). Topological property invariant to non-affine deformations.

3. Hu's 7 Invariant Moments (Ming-Kuei Hu, 1962)

Definition: Central Moments & Hu's Invariants

The 2D geometric moment of order (p+q) for digital image f(x,y) is:

$m_{pq} = \sum \sum x^p y^q f(x,y)$

Central moments μ_{pq} are computed relative to the centroid $(\bar{x}, \bar{y})$, providing Translation Invariance. Normalized central moments η_{pq} provide Scale Invariance. Hu derived 7 non-linear combinations of normalized central moments (Φ_1 through Φ_7) that are strictly invariant to Translation, Scaling, and 2D Rotation.

4. Shape Matching Techniques

- • **Hausdorff Distance:** Measures the maximum distance of a point in set A to the nearest point in set B: $H(A,B) = \max(h(A,B), h(B,A))$. Highly robust to outliers in shape boundary matching.
- • **Chamfer Matching:** Aligns a template edge contour with a target edge image by minimizing the average distance transform value along the contour path.

4.5 High-Yield University Exam Question Bank & Model Answers

Part A: Short Answer Questions (2–3 Marks)

- **Q1. What is the Harris Corner Response function R? How are flat, edge, and corner regions classified?**

Answer: The Harris corner response is $R = \det(M) - k(\text{trace}(M))^2 = (\lambda_1 \lambda_2) - k(\lambda_1 + \lambda_2)^2$, where M is the structure tensor. If $R > 0$ (large), it is a Corner; if $R < 0$ (large negative), it is an Edge; if $|R| \approx 0$, it is a Flat homogeneous region.

- **Q2. State the principle of Otsu's Thresholding algorithm.**

Answer: Otsu's algorithm automatically determines the optimal binarization threshold T^* by maximizing the Between-Class Variance $\sigma_B^2(T) = \omega_1 \omega_2 (\mu_1 - \mu_2)^2$ across all candidate intensity levels, effectively minimizing the within-class variance of the foreground and background distributions.

- **Q3. What is the Gray-Level Co-occurrence Matrix (GLCM)? Name four texture features derived from it.**

Answer: GLCM $P(i,j \mid d, \theta)$ is a second-order statistical matrix that tabulates the frequency with which a pixel of gray level i occurs at distance d and angle θ from a pixel of gray level j. Four core GLCM features are Contrast, Energy, Entropy, and Homogeneity.

- **Q4. Define Hu's 7 Invariant Moments.**

Answer: Hu's 7 Invariant Moments are a set of non-linear combinations of normalized 2D central algebraic moments (Φ_1 through Φ_7) that provide mathematical invariance to 2D image Translation, Scale changes, and planar Rotation.

Part B: Long Answer Questions (8–10 Marks)

- **Q5. Explain the Harris Corner Detection Algorithm in complete detail. Derive the Structure Tensor M and discuss the eigenvalue analysis.**

Model Answer Outline:

- 1. Definition of corners as intersection of edges with large multi-directional gradients.
- 2. Formulation of the Sum of Squared Differences (SSD) error function $E(u,v)$ under small spatial shifts.
- 3. First-order Taylor Series approximation leading to quadratic matrix form $E(u,v) \approx [u, v] M [u, v]^T$.
- 4. Detailed breakdown of Structure Tensor M components (Gaussian-weighted spatial gradient outer products I_x^2, I_y^2, $I_x I_y$).
- 5. Geometric eigenvalue interpretation: λ_1, λ_2 ellipse axes representing local curvature.
- 6. Derivation of the Harris Response $R = \det(M) - k(\mathrm{trace}(M))^2$ and comparison with Shi-Tomasi $\min(\lambda_1, \lambda_2)$.

- **Q6. Describe Image Segmentation using (a) Otsu's Thresholding, (b) Region Growing and Region Splitting/Merging, and (c) Morphological Operations.**

Model Answer Outline:

- 1. Formal mathematical definition of Image Segmentation (disjoint partition criteria).
- 2. Otsu's algorithm derivation: Histogram probability p_i, cumulative class probabilities ω_1, ω_2, class means μ_1, μ_2, maximization of between-class variance σ_B^2.

- 3. Region Growing: Seed selection, similarity predicate, stopping criteria.
- 4. Region Splitting and Merging: Quadtree hierarchical decomposition and union merge conditions.
- 5. Morphological Operations: Structuring elements, Dilation, Erosion, Opening (noise cleaning), Closing (hole filling), Morphological Gradient.

4.6 Unit IV Summary & Quick Revision Sheet

Feature / Segmentation Domain	Key Theoretical & Mathematical Takeaways for Exams
Corner Detectors	Harris ($R = \det(M) - k\,\mathrm{trace}(M)^2$; $R>0$ corner, $R<0$ edge). Shi-Tomasi ($\min(\lambda_1, \lambda_2) > \lambda_{min}$).
Texture Features	1st-order (Mean, Variance, Skewness, Kurtosis); 2nd-order GLCM (Contrast, Energy, Entropy, Homogeneity). Gabor filters.
Thresholding	Global, Local Adaptive, Otsu's method (Maximizes Between-Class Variance $\sigma_B^2 = \omega_1\omega_2(\mu_1-\mu_2)^2$).
Region Segmentation	Seeded Region Growing, Quadtree Split-and-Merge, K-Means Clustering, Marker-Controlled Watershed.
Shape Descriptors	Freeman Chain Codes, Fourier Descriptors, Compactness (P^2/A), Euler Number ($E = C - H$), Hu's 7 Moment Invariants.