Image Processing and Analysis

Unit - 5

Feature Analysis and Object Recognition

Boundary representation - boundary descriptors; regional descriptors - texture; relational descriptors; Object recognition: patterns and pattern classes - recognition based on decision - theoretic measures - case study: model based tracking of animals

* Boundary Representation

→ Segmentation techniques yield data in the form of pixels along a boundary, or pixels in a region

→ Schemas are used to compact the segmented data into representations that facilitate the computation of descriptors

→ Some representation schemes include:

(i) Boundary Following
(ii) Chain codes
(iii) Polygonal approximations
(iv) Signatures
(v) Boundary segments
(vi) Skeletons

Method1 - Boundary Following - Moore's boundary tracking algorithm.

→ Assume that we are working with binary images in which the object and the background points are labeled as 1 and 0.

→ The images are padded with a border of 0s to eliminate the possibility of an object merging with the image border

Algorithm : 1. Let the starting point b_0 be the uppermost, left most point in the image that is labelled 1. Denote c_0 by the west neighbor of b_0. Examine the 8 neighbors of b_0 starting at c_0 and proceed in the clockwise direction.

2. Let b_1 denote the first neighbor encountered whose value is 1, and let c_1 be the background point immediately preceding b_1.

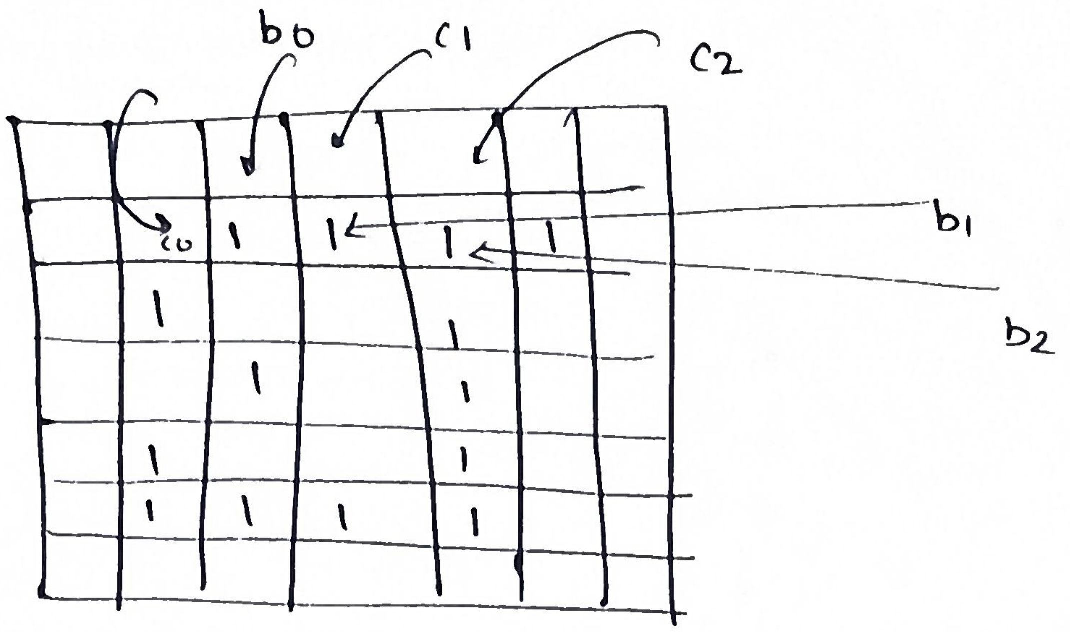
3. Let $b = b_0$, $c = c_0$

4. Let the 8 neighbors of b starting at c, proceeding in a clockwise manner by $n_1, n_2 \dots n_8$. Find the first neighbor labeled 1 and denote it as n_k.

5. Let $b = n_k$ and $c = n_{k-1}$

6. Repeat steps 4 & 5 until $b = b_0$.

Example

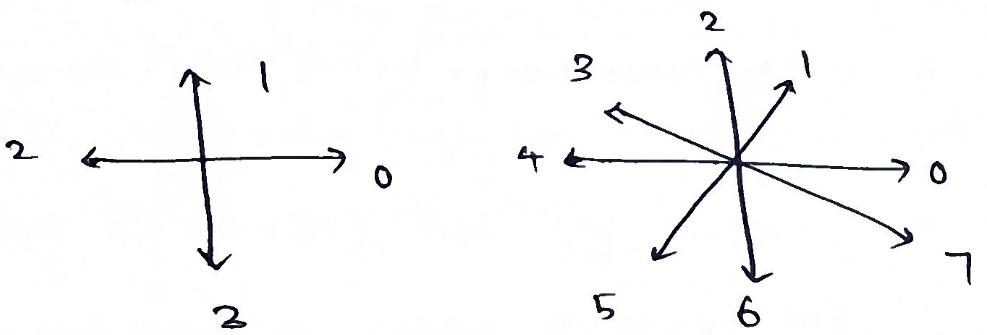


Method 2 : Chain Codes (Freemain chaincode)

→ Chain codes are used to represent a boundary by a connected sequence of straight line segments of specified length and direction

→ This representation is based on the 4 or 8 connectivity of the segments.

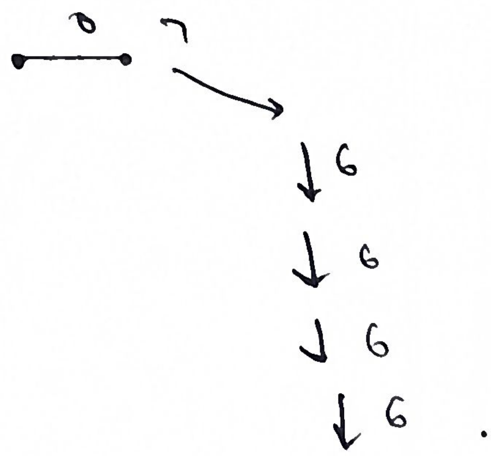
→ The direction of each segment is coded using a numbering scheme



→ The chain code depends on the starting point. It can be normalized by:

(i) Treat the chain code as a circular sequence of direction numbers and redefine the starting point so that the resulting sequence of numbers formed is the smallest integer.

→ Normalise the rotation by using the difference in chain code.



Method 3 : Polygonal Approximation

→ A digital boundary can be approximated with arbitrary accuracy by a polygon

→ For a closed boundary, the approximation becomes exact when the no. of segments of the polygon = no. of points in the boundary.

→ The goal of the polygon approximation algo. is to capture the essence of the shape in a given boundary, using the fewest number of segments

Minimum Perimeter Polygon - An approach where a boundary is enclosed by a set of concatenated cells

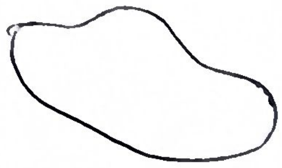
→ Allow the boundary to shrink like a rubber band. This shrinking produces the shape of a polygon of minimum perimeter.

→ The size of the cells determines the accuracy of the polygonal approx.

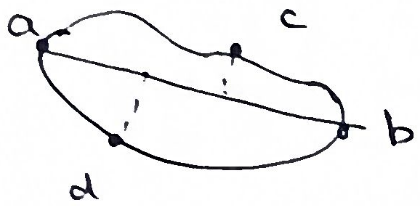
Method 4 : Signatures

→ Signatures are a 1D functional representation of a boundary.

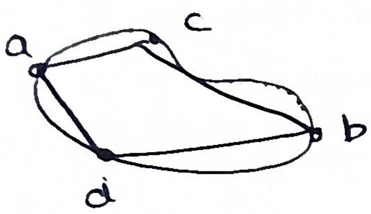
→ They are plotted as the distance from the centroid to the boundary as a function of an angle.



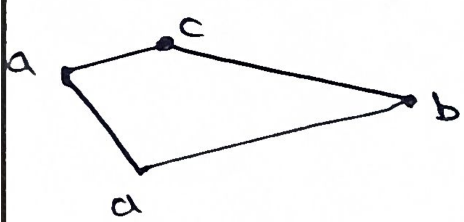
→ original boundary



→ boundary divided into segments based on extreme points



→ join the vertices



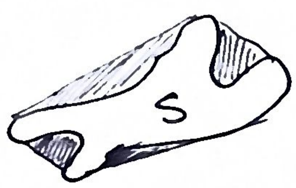
→ resulting polygon

Method 5 : Boundary Segments

→ The boundary can be decomposed into segments - It is useful to extract information from the concave parts of objects.

→ This is achieved by calculating the convex hull of a region

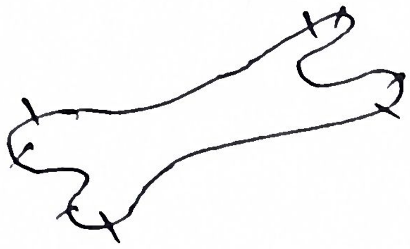
The convex hull H of an arbitrary set S, is the smallest set containing S → H - S = convex deficiency



→ The convex deficiency is the shaded region

→ Convex deficiency can be followed to mark

Features - This is done by following the contour of S and marking points that transition into or out of the convex deficiency



Method 6 : Skeletons

→ The structural shape of a plane region can be represented by reducing it to a graph

→ The reduction may be accomplished by obtaining the skeleton of the region via thinning (skeletonizing) algorithm.

→ Skeleton of a region may be defined as the medial axis transformation (MAT).

→ The MAT of a region R with border B is as follows:

(→) For every point p in R, we find the closest neighbor in B. If p has more than one neighbor, it is said to belong to the medial axis (skeleton)

→ MAT is based on the prairie fire concept.

* Boundary Descriptors

→ Some of the kinds of boundary descriptors include:

A. Simple Descriptors

(i) length of the contour

(ii) boundary diameters

(iii) curvature

(iv) shape number

(v) eccentricity.

B. Fourier Descriptors

C. Statistical Moments

A. Simple Descriptors

(i) Length of contour -

→ count the number of pixels along the contour

→ For a chain code curve : length = no. of H + V + $2^{1/2}$ × no. of components

(ii) boundary diameter

$$\text{Diam}(B) = \max_{i,j} \left[D(p_i, p_j) \right]$$

D = distance measure - Euclidean distance or D_4 distance

(iii) curvature :

→ rate of change of slope

→ Find the difference between the slopes of adjacent boundary segments, at the point of intersection

(iv) shape number - smallest magnitude of the first difference of chain code representation

Order of a shape = no. of digits in its representation

shape order is even for a closed boundary.

(v) eccentricity - ratio of the major axis to the minor axis

B. **Fourier Descriptors**

→ For a set of points that make up the outline of an object in an image - the Fourier transform is applied on them.

→ This would yield a set of complex numbers z_n.

→ Use the real & imaginary parts of these complex numbers to get the magnitude

$$F_n = \sqrt{Re(z_n)^2 + Im(z_n)^2}$$

These F_n values are the Fourier descriptor.

(reduces 2D to 1D)

→ Fourier descriptors have the properties of translation invariance, rotation invariance & scale invariance.

C. **Statistical Moments**

→ Statistical moments capture information about the distribution of pixel intensities along the boundary of an object

→ For all the pixel intensity values along the boundary of the image:

(i) Mean = First Order moment – compute avg. brightness

(ii) Variance = Second Order Moment – measures how much the pixel intensities vary around the mean

$$var = \frac{1}{N} \sum_{k=1}^{N} (I_k - mean)^2$$

(iii) Skewness – Third Order Standardized Moment – describes the asymmetry of the intensity distribution.

* Regional Descriptors

Some simple regional descriptors include:

(i) area
(ii) perimeter
(iii) compactness $= \dfrac{(perimeter)^2}{area}$.
(iv) mean / median or min / max of gray level values

* Topological Descriptors

→ Topology refers to the study of properties of a figure that are unaffected by any deformation as long as there is no tearing or joining of the figure (rubber sheet distortions)

→ Some topological descriptors include:

(i) no. of holes in a region

(ii) no. of connected components - C

(iii) Euler no: $E = C - H$

(iv) For regions represented by straight line segments - the Euler formula would be $\underset{\text{vertex}}{V} - \underset{\text{edges}}{Q} + \underset{\text{face}}{F} = C - H = E$

* Texture

→ Texture has no formal definition - it is a set of measures calculated to quantify the perceived property of an image.

→ It provides measures of properties like smoothness, coarseness and regularity of an image.

→ There are 3 principal approaches for texture analysis:

(a) Statistical Approach

(b) Structural Approach

(c) Spectral Approach.

A. Statistical Approach

→ A quantitative measure of the arrangement of intensities in a region.

→ Statistical measures compute local features at each point in the image and derive a set of statistics from the distribution of local features.

→ The statistical approach is divided into:

(i) Histogram method

(ii) Co-occurrence matrix method

(i) Histogram method

$\boxed{\text{image}}$ $f \longrightarrow h_f$ $\boxed{\text{histogram}}$

Obtain statistics from the histogram:

$$\text{mean} = \sum_{i=0}^{L-1} i\,h(i)$$ - avg. intensity

$$\text{variance} = \sum_{i=0}^{L-1} (i-\mu)^2\, h(i)$$ - measure of intensity contrast

$$\text{skewness} = \sum_{i=0}^{L-1} (i-\mu)^3\, h(i)$$

$$\text{entropy} = -\sum_{i=0}^{L-1} h(i) \log h(i)$$

(ii) Co-occurrence Matrix Method

→ The co-occurrence matrix of an image is created based on the correlations between image pixels.

→ For a k bit image with $L = 2^k$ brightness levels, an $L \times L$ matrix is created whose elements are the number of occurrences of a pair of pixels with brightness a, b separated by d pixels in a certain direction

→ Generally, the co-occurrence matrix is defined for four main direction, (0, 45, 90 and 135) degrees

→ Statistical metrics can be calculated from the co-occurrence matrix. These include:

- max probability : $\max_{i,j} (p_{ij})$
- correlation : $\sum_{i=1}^{K} \sum_{j=1}^{K} \frac{(i - m_r)(j - m_c) p_{ij}}{\sigma_r \sigma_c}$
- contrast : $\sum_{i=1}^{K} \sum_{j=1}^{K} (i-j)^2 p_{ij}$
- uniformity : $\sum_{i=1}^{K} \sum_{j=1}^{K} p_{ij}^2$
- homogenity : $\sum_{i=1}^{K} \sum_{i=1}^{K} \frac{p_i}{1 + |i-j|}$
- entropy : $- \sum_{i=1}^{K} \sum_{i=1}^{K} p_{ij} \log_2 p_{ij}$

B. $\boxed{\text{Structural Approach}}$ (Relational)

→ A simple texture element can be used to form more complex texture patterns by means of rules

→ · For example, there is a rule that indicates aS : S → aS

Then 3 repetitions of the rule yield the string aaaS.

→ If a represents a circle and has the meaning circles to the right, then aaa would mean OOO.

→ If there are rules like :

$S \to bA$, $A \to cA$

$A \to c$, $A \to bS$, $S \to a$ and if b means circle down, and c means a circle to the left

then aaabccbaa corresponds to a 3 x 3 matrix of circles

c) Spectral Approach

→ The Fourier transform is used to describe the directionality of periodic or almost periodic structure of within the image.

→ 3 Features of the Fourier spectrum are useful in texture description:

(i) Peaks in the spectrum give the principle direction of the texture pattern

(ii) Location of the peaks in the frequency plane gives the spatial period of the pattern

(iii) Eliminating the periodic components other than around the origin. would show that the FT is symmetric around the origin.

→ The spectrum can also expressed in terms of polar coordinates

$S(r,\theta)$

$$S(r) = \sum_{\theta} S(r,\theta)$$

$$S(\theta) = \sum_{r} S(r,\theta)$$

* Object Recognition

→ The automatic recognition of objects or patterns is an image analysis task that a computer program can do.

→ The approaches to pattern recognition are divided into 2 principal areas :

(i) Decision Theoretic Methods - deal with patterns described using quantitative descriptors such as length, area & texture. (vectors)

(ii) Structural methods - deals with patterns best described by qualitative descriptors (symbolic information), such as relational descriptors. (patterns in trees)

* Patterns and Pattern Classes

Pattern : an arrangement of descriptors or features

Pattern class: a family of patterns sharing some common properties, denoted by $\omega_1, \omega_2 \dots \omega_n$,
n = no. of classes

Goal of pattern recognition - Assign patterns to their classes with as little human interaction as possible.

* Types of Features / Descriptors

1. Pattern vectors : in the form of $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

   eg. recognition / classification of irises based on length & width of petals

2. String descriptors : describes structural relationships - eg: Finger print recognition.

   → Look for properties that describe fingerprint properties

   eg staircase pattern - alternating ababab

3. Tree descriptors : have a hierarchical ordering

   eg. satellite imagery

   downtown   residential

A. Decision Theoretic Methods

→ They are based on decision (discriminant) functions.

→ Let $X = [x_1, x_2 \dots x_n]^T$ represent a pattern vector.

→ For n pattern classes $\omega_1, \omega_2 \dots \omega_n$, the basic problem is to find ω decision functions

$$d_1(x), \; d_2(x) \; \dots \; d_\omega(x)$$

with the property that x belongs to class ω_i

$$d_i(x) > d_j(x) \quad \text{for } j = 1, 2 \dots n \quad j \neq i$$

Some decision theoretic methods are:

1. **Matching:** an unknown pattern is assigned to the class to which it is closest to with respect to a metric. Some methods used are:

minimum distance classifier - computes Euclidean distance between the unknown pattern and each of the prototype vectors

correlation - can be directly formulated in terms of images

2. **Optimum Statistical Classifiers** -

3. **Neural Networks**

A. Minimum Distance Classifiers

→ calculate Euclidean distance

$D_j(x) = \|x - m_j\| \quad j = 1, 2, \ldots n$

→ Assign x to a given class with the smallest $D_j(x)$

→ works well when the distance between means is large is compared to the spread of each class

B. Matching By Correlation

→ compute the normalized coefficient, it would lie between $[-1, 1]$. Match according to correlation value

c. Optimum Statistical Classifiers

→ a probabilistic approach to recognition

→ A $\boxed{\text{Bayes classifier}}$ can be represented as:

$$r_j(x) = \sum_{k=1}^{W} L_{kj}\, p(x \mid w_k)\, P(w_k)$$

$p(x|w_j)$ = PDF of pattern classes

$P(w_j)$ = probability of occurrence of a class

→ Assign a pattern x to the class with the lowest loss. The classifier that minimizes the total average loss is the Bayes classifier.

→ The loss for a wrong decision is assigned a non-zero value, while that for a correct decision is 0

→ A $\boxed{\text{Bayes classifier for Gaussian patterns}}$ models each class using a Gaussian distribution.

→ The classifier ~~with the~~ considers prior knowledge and updates it with the observed data to make informed decisions (posterior probability).

→ Bayesian classification chooses the class w/ the maximum posterior probability.

D) Neural networks - refer soft computing

B. Structural Methods

Methods include : (i) matching shape numbers

(ii) string matching

Matching Shape Numbers - The degree of similarity between two shapes is defined as the largest order for which their shape numbers still coincide.

String matching - Region boundaries a & b are coded into strings denoted as $a_1 a_2 a_3 \ldots a_n$ and $b_1 b_2 b_3 \ldots b_m$

→ Matches (p) is when $a_k = b_k$

→ no. of symbols that do not match is $q = (\max(|a|, |b|)) - b$

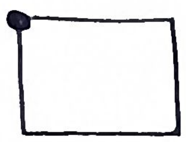
→ A measure of similarity is :

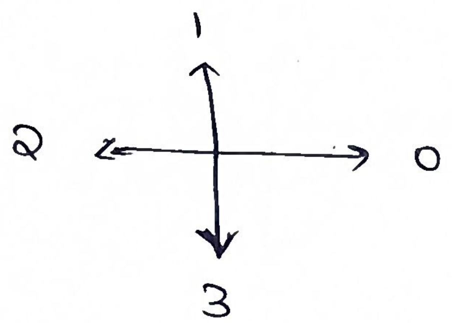
$$R = \frac{\text{no. of matches}}{\text{no. of mismatches}} = \frac{p}{q} = \frac{p}{\max(|a|, |b|) - b}$$

Numericals

(1) Chain codes, shape number

1. For each of the following images, write the chain code representation and find the shape number and order

(a)  w/ 4-chain-code representation



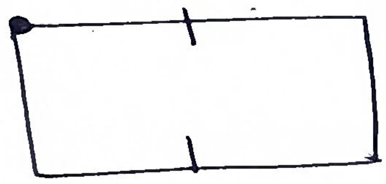
chain code : 0 3 2 1

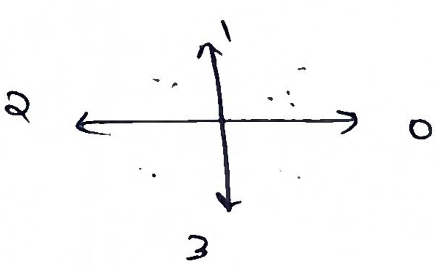
01 - 3
03 - 1

first diff : 3 3 3 3

↳ ~~2nd~~ 1st no to ~~first~~ 2nd number in anti-clockwise direction

shape no = 3 3 3 3

(b)  w/ 4-chain code representation

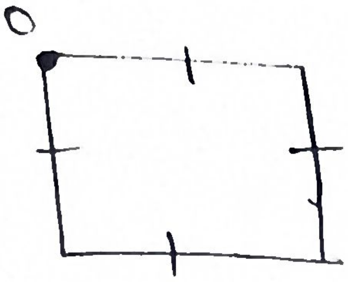


chain code : 0 0 3 2 2 1

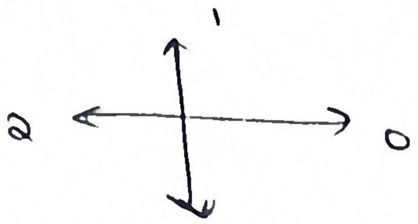
first diff : ~~3 3 3 3 0 3 3~~
3 0 3 3 0 3

shapeno: 0 3 3 0 3 3

(c)



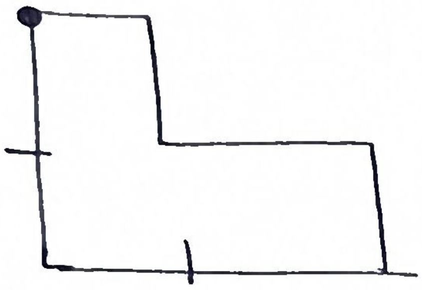
w| 4-chain code

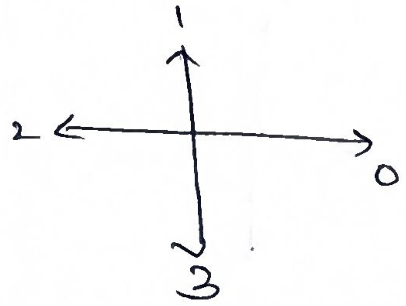


chain code : 0 0 3 3 2 2 1 1 3

First diff : [illegible] 3 0 3 0 3 0 3 0

(d)

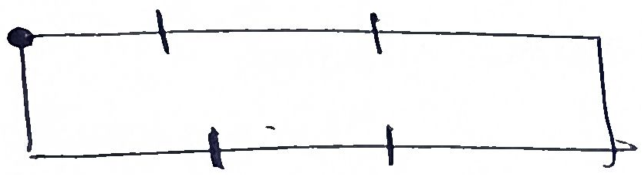




chain code : 0 3 0 3 2 2 1 1

First diff : 3 3 1 3 3 0 3 0

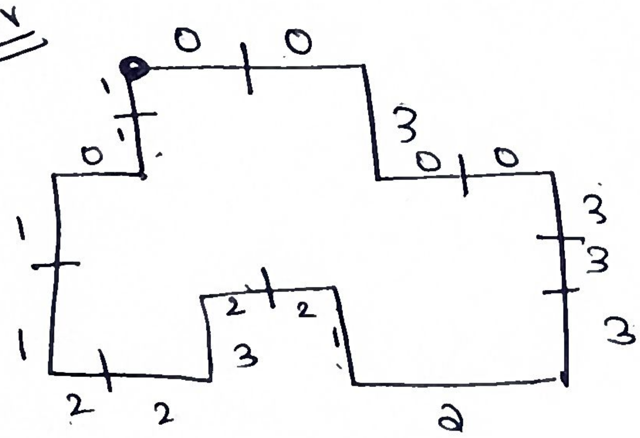
(e)

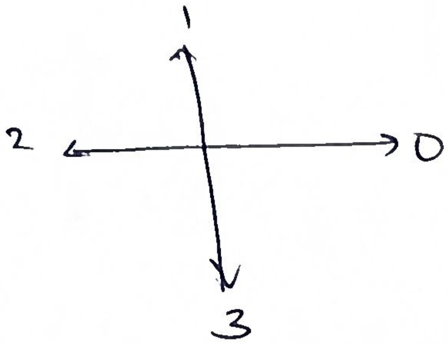


chain code : 0 0 0 3 2 2 2 1

: 3 0 0 3 3 0 0 3

Q paper





Q Paper Descriptors for content-based image retrieval

- color, texture, shape, spatial features

→ geometric properties, boundary discriptors, moments

→ co-occurrence matrices

Q Paper Consider the following image

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

Construct the co-occurrence matrix for the given template.

$$\begin{bmatrix} x & \\ & y \end{bmatrix}$$

Ans

$$\begin{bmatrix} 1 & \boxed{0} & 1 & 0 \\ 0 & 1 & \boxed{0} & 1 \\ 1 & \boxed{0} & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

2 x 2 co-occurrence matrix

$$\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 3 & 0 \\ 1 & 0 & 4 \end{array}$$

for the template (x, y)

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

$$\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 0 & 4 \\ 1 & 5 & 3 \end{array}$$

Unit 2 Paper Numericals

(1) Given an image f of 4×4 and a mask of 2×2 calculate the convolution $f * g$. og = left-lower corner

f =

0	0	0	0
0	1	2	0
0	3	4	0
0	0	0	0

g =

-1	2
1	1

convolve mask

horizontal

1	1
-1	2

⇒ vertical ⇒

1	1
2	-1

size of padded image = $(4+2)-1 \times (4+2)-1 = 5 \times 5$

~~0~~	0	0	0	0	0
~~0~~	0	1	2	0	0
~~0~~	0	3	4	0	0
~~0~~	0	0	0	0	0
~~0~~	0	0	0	0	0

= ~~-1~~

0	0	0	0	0
0	0	0	0	0
0	1	2	0	0
0	3	4	0	0
0	0	0	0	0

$$= \begin{bmatrix} 0 & 0 & 0 & 0 \\ -1 & 0 & 4 & 0 \\ -2 & 5 & 10 & 0 \\ 3 & 7 & 4 & 0 \end{bmatrix}$$

(2) Perform image enhancement using 2nd order derivative Laplacian mask on the given image

1	2	2
4	3	1
1	0	2

Mask =

0	1	0
1	-4	1
0	1	0

0	0	0	0	0
0	1	2	2	0
0	4	3	1	0
0	1	0	2	0
0	0	0	0	0

→

-6	-2	

2-1 4 - 12

and so on

Unit 3 - Q paper numericals

(1) The given img is affected w/ uniform noise. Remove the noise produce the noise free image using statistical order filter of size 3x3

25	40	70	80
60	78	90	32
45	56	30	70
83	78	89	70

use the midpoint filter.

25	40	70	80
60	58	90	32
45	56	30	70
83	78	89	70

(i) $\frac{90+25}{2} = 57.5 = 58$

25	40	70	80
60	58	60	32
45	56	30	70
83	78	89	70

$\frac{90+30}{2} = 60$

25	40	70	80
60	58	60	32
45	60	60	70
83	78	89	70

$\frac{90+30}{2} = 60$

$\frac{90+30}{2} = 60$