Section - B

Coding Redundancy :-

Assume that a discrete random variable r_k (kth) in the interval $[0, \ldots L-1]$ is used to represent the intensities of an $M \times N$ image and that each r_k occurs with probability $P_r(r_k)$.

gray level

$$\boxed{P_r(r_k) = \frac{n_k}{MN} \qquad k = 0, 1, 2, \ldots, L-1}$$

L

n_k

Where L is the no. of intensity values,

n_k is the no. of times that k^{th} intensity appears in the image. If the no. of bits used to represent each value of r_k is $l(r_k)$, then the average number of bits required to represent each pixel is

$M \times N$ pixels

$l(r_k)$

r_k

$$\boxed{L_{avg} = \sum_{k=0}^{L-1} l(r_k) P_r(r_k)}$$

avg. length of code.

That is, the average length of the code words assigned to the various intensity value is found by summing the products of bits used to represent each intensity and the probability that the intensity occurs.

avg

The total no. of bits required to represent an $M \times N$ image is

MNL_{avg}.

0 - L

Coding Redundancy :-

If the gray level (intensity) of an image encoded in a way that uses more code symbols than absolutely necessary to represent each gray level, the resulting image is said to contain coding redundancy.

Scanned by CamScanner

Interpixel Redundancy:

In general, coding redundancy is present when the codes assigned to a set of events have not been selected to full advantage of the probabilities of the events.

This almost always present when an image's grey levels are represented with a straight or natural binary code.

4-connected

Interpixel Redundancy

~~Consider the image which have virtually identical histogram~~

The value of any given pixel can be reasonably predicted from the value of its neighbours, the information carried by individual pixels is relatively small. Much of the visual contribution of a single pixel to an image is redundant; it could have been guessed on the basis of the values of its neighbours and this type of data redundancy is known as interpixel redundancy.

Or, we can say,

It is a redundancy corresponding to statistical dependencies among pixels, especially between neighbouring pixels.

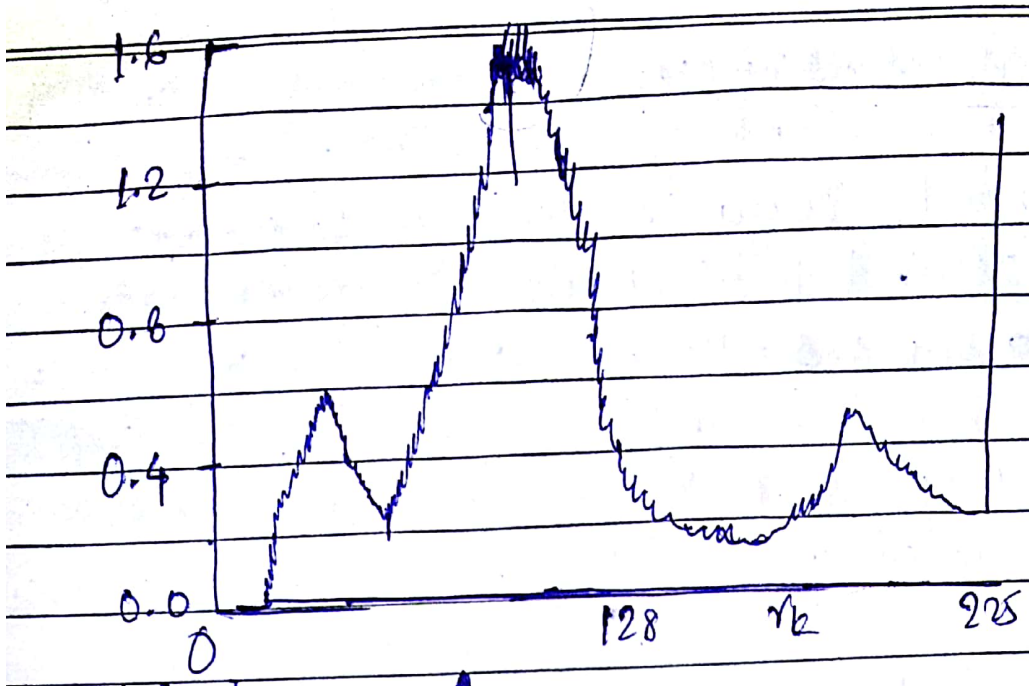
Both histogram are trimodal, indicating the presence of three dominant ranges of grey-level values. Because the grey levels in these images are not equally probable. Variable length coding can be used to reduce the coding redundancy that would result from a straight or natural encoding of their

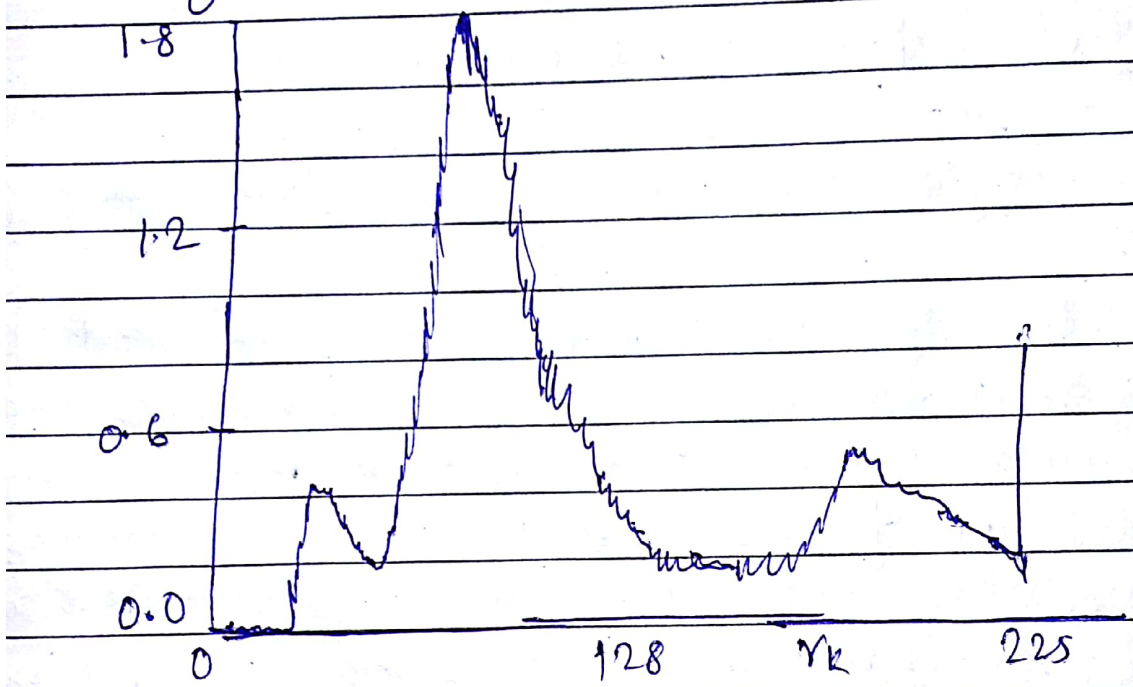
Scanned by CamScanner

Spice

Date:

Page No:





Scanned by CamScanner

Date:
Page No:

Pixels. The coding process, however, would not alter the level of correlation between the pixels within the images.

In other words, the codes used to represent the gray levels of each image have nothing to do with the correlation between pixels. These correlations result from the structural or geometric relationships between the object in the image.

Psychovisual Redundancy:-

[The eye can only resolve about 32 gray levels locally.]

In an image certain information ~~has slightly~~ simply has less relative importance than other information in normal visual processing. This information is said to be ~~psy~~ psycho-visually redundant.

It can be eliminated without significantly impairing the quality of image perception.

That psychovisual redundancies exist should not come as a surprise, because human perception of the information in an image normally does not involve quantitative analysis of every pixel value in the image. In general, an observer searches for distinguishing features such as edges or textural regions and mentally ~~searches for distinguishing features~~ combines them into recognizable groupings. The brain then correlates these groupings with prior knowledge in order to ~~complete~~ complete the image interpretation process.

Scanned by CamScanner

Psychovisual redundancy is fundamentally different from Coding and Interpixel redundancy.

1.) Unlike Coding and Interpixel redundancy, ~~Psycho~~ Psycho. visual redundancy is associated with real or quantifiable visual information. Its elimination is possible only because the information itself is not essential for normal visual processing.

Since the elimination of psychovisually redundant data result in loss of quantitative information, it is commonly referred to as quantization.

Image → pixels → set of intensities f x,y

Information and data

→ Data is not the same thing as information. Data is the means with which information is expressed. The amount of data can be much larger than the amount of information.

→ Redundant data doesn't provide additional information.

→ Image coding or compression aims at reducing the amount of data while keeping the information by reducing the amount of redundancy.

n_1 = data

n_2 = data − redundancy (i.e; data after compression)

Scanned by CamScanner

Compression ratio $= C_R = \frac{n_1}{n_2}$

Relative redundancy : $R_D = \boxed{1 - \frac{1}{C_R}}$

Different type of Redundancy

1. Coding Redundancy 2. Interpixel Redundancy 3. Psycho-visual Redundancy

Image & Binary code

Huffman Coding

One of the most popular technique for removing coding redundancy is due to Huffman. When coding the symbols of an information source individually, Huffman coding yields the smallest possible number of code symbol per source symbol.

In terms of Shannon's first theorem, the resulting code is optimal for a fixed value of n, subject to the constraint that the source symbols be coded one at a time.

Step 1:

The first step in Huffman's coding approach is to create a series of source reductions by ordering the probabilities of the symbols under consideration and combining the lowest probability symbols into a single symbol that replaces them in next source-reduction.

As shown in Fig 1, At the far left, a hypothetical

Spice
Date:
Page No:

Set of source symbols and their probabilities are ordered from top to bottom in terms of ~~symbols and their probabilities~~ decreasing probability values.

To form the first source reduction, the bottom two probabilities, 0.06 and 0.04 are combined to form a "Compound symbol" with probability 0.1. This compound symbol and their associated probability are placed in first source reduction column so that the probabilities of the reduced source are also ordered from the most to least probable. This process is then repeated until a reduced source with two symbols (at the far right) is reached.

Step 2:

Code each reduced source, starting with the smallest source and working back to the original source. The minimal length binary code for a two symbol source is the symbols 0 and 1.

Original source		Source reduction			
Symbol	Probability	1	2	3	4
a_2	0.4	0.4	0.4	0.4	0.6
a_6	0.3	0.3	0.3	0.3	0.4
a_1	0.1	0.1	0.2	0.3	
a_4	0.1	0.1	0.1		
a_3	0.06	0.1			
a_5	0.04				

Huffman source reductions.

0.07
0.04

original source			source reduction							
Symbol	Probability	Code		1		2		3		4
a_2	0.4	1	0.4	1	0.4	1	0.4	1	0.6	0
a_6	0.3	00	0.3	00	0.3	00	0.3	00	0.4	1
a_1	0.1	011	0.1	011	0.2	010	0.3	01		
a_4	0.1	0100	0.1	0100	0.1	011				
a_3	0.06	01010	0.1	0101						
a_5	0.04	01011								

Average length of this code is

$$L_{avg} = (0.4)(1) + (0.3)(2) + (0.1)(3) + (0.1)(4) + (0.06)(5) + (0.04)(5) \doteq 2.2 \text{ bits/pixel.}$$

and the entropy of the source is 2.14 bits/symbol.

010
011

0 0
0 1

Scanned by CamScanner

Arithmetic Coding:-

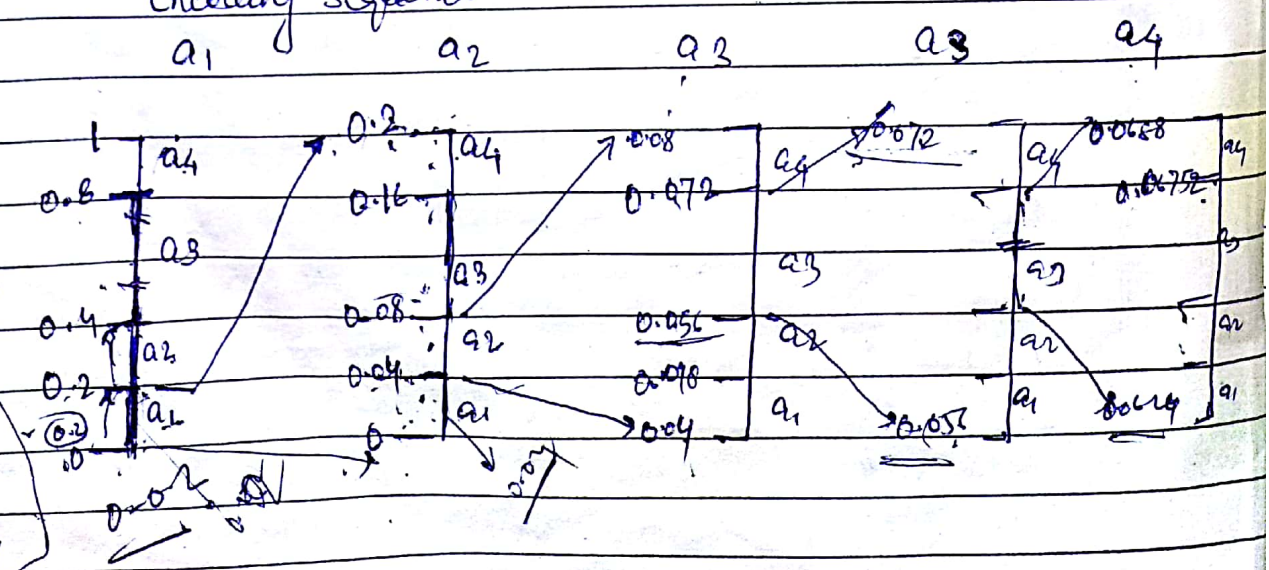
⇒ Unlike Huffman Coding and Golomb coding, arithmetic coding generates nonblocks codes.

In arithmetic coding a one-to-one correspondence between source symbols and code words does not exist. Instead, an entire sequence of source symbols (or message) is assigned a single arithmetic code word. The code word itself defines an interval of real numbers between 0 and 1.

⇒ As the no. of symbols in the message increases the interval used to represent it becomes smaller and the number of information units (say, bits) required to represent the interval becomes larger.

⇒ Each symbol of the message reduces the size of the interval in accordance with its probability of occurence. Because the technique does not require, as does Huffman's approach that each source symbol translate into an integral number of code symbols, it achieves.

Encoding sequence →

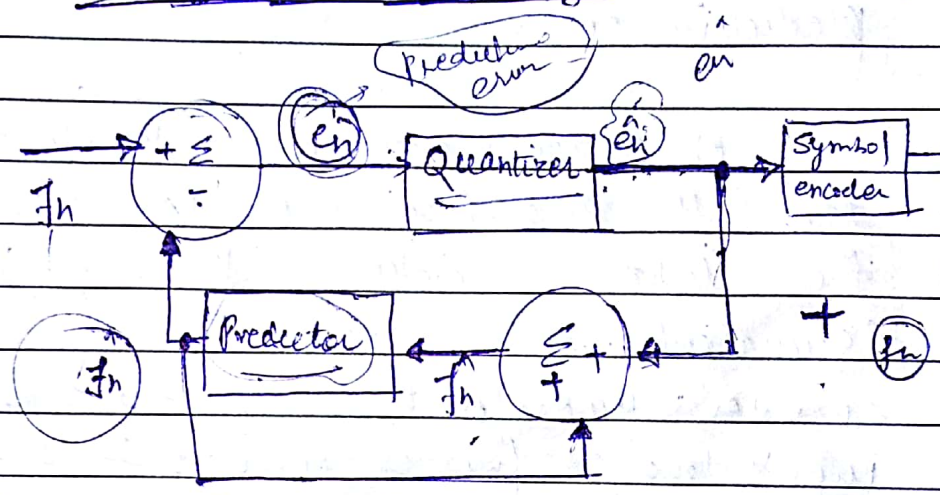


Spice
Date:
Page No:

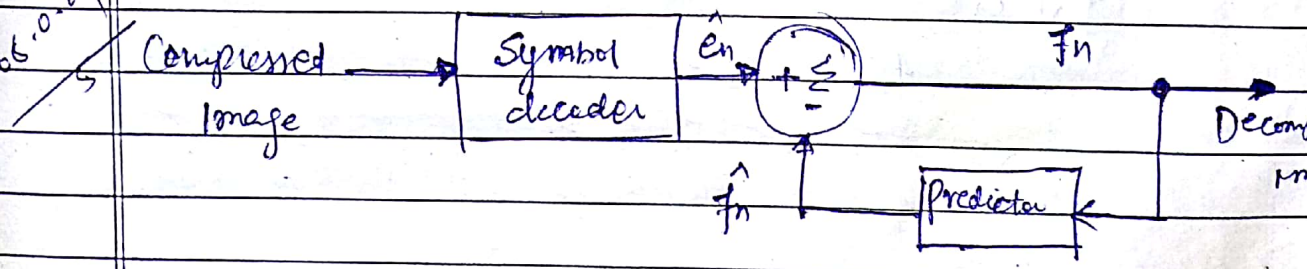
Arithmetic coding example:-

Source symbol	Probability	Initial subinterval
a_1	0.2	[0.0, 0.2)
a_2	0.2	[0.2, 0.4)
a_3	0.4	[0.4, 0.8)
a_4	0.2	[0.8, 1.0)

Lossy Predictive Coding



A lossy predictive coding model: Encoder



A lossy predictive coding model: Decoder

Unlike the error-free approaches, lossy encoding is based on the concept of compromising the accuracy of the reconstructed image in exchange for increased compression.

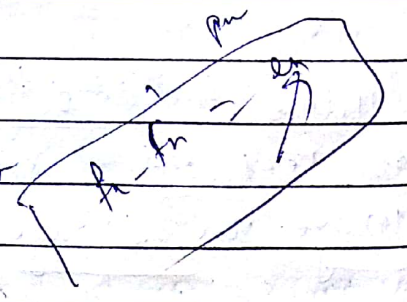
We add a quantizer to the model introduced in lossless compression and examine the resulting trade-off between reconstruction accuracy and compression performance.

As shown in block diagram, the quantizer, which absorbs the nearest integer function of the error-free encoder, is inserted between the symbol encoder and the point at which the prediction error is formed. It maps the prediction error into limited range of outputs, denoted $\hat{e}_n$, which establish the amount of compression and distortion associated with lossy predictive coding.

In order to accommodate the insertion of the quantizer step, the error-free encoder must be altered so that the predictions generated by the encoder and decoder are equivalent. This is accomplished by placing the lossy encoder's predictor within a feedback loop, where its input denoted $\hat{f}_n$, is generated as a function of past predictions and the corresponding quantized errors.

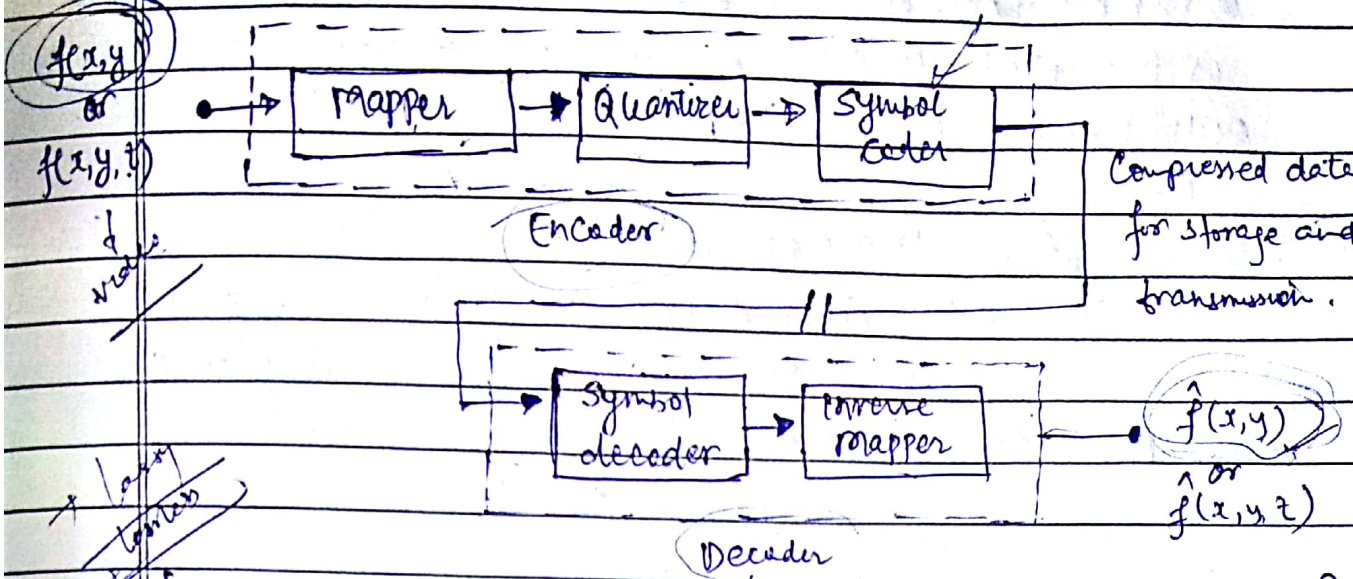
That is

$$\boxed{f_n = \hat{e}_n + \hat{f}_n}$$



Scanned by CamScanner

Image Compression Model :-



[Functional block diagram of a general image Compression method.]

Image Compression system is composed of two distinct functional Components:-

1. Encoder
2. Decoder

Codec

The Encoder performs Compression, and decoder performs the complementary operation of decompression.

Both operations can be performed in software, as is the case in web browsers and many Commercial image editing programs, Or in Combination of hardware, and firmware, as in Commercial DVD Players

A codec is device or program that is capable of both encoding and decoding.

Procedure:-

Input Image f(x,...) is fed into the encoder, which creates a compressed representation of input. This representation is stored for later use, or transmitted for storage and use.

Scanned by CamScanner

a remote location. When the compressed representation is presented to its complementary decoder, a reconstructed output image $\hat{f}(x,\ldots)$ is generated.

In still-image applications, the encoded input and decoder output are $f(x,y)$ and $\hat{f}(x,y)$ respectively.

In video applications, they are $f(x,y,t)$ and $\hat{f}(x,y,t)$

where, discrete parameter t specifies time.

Compression model

- Error free, lossless, or information preserving
- Lossy

If $\hat{f}(x,\ldots)$ is exact replica of $f(x,\ldots)$ then compression system is called lossless.

If the reconstructed output image is distorted then compression system is called lossy.

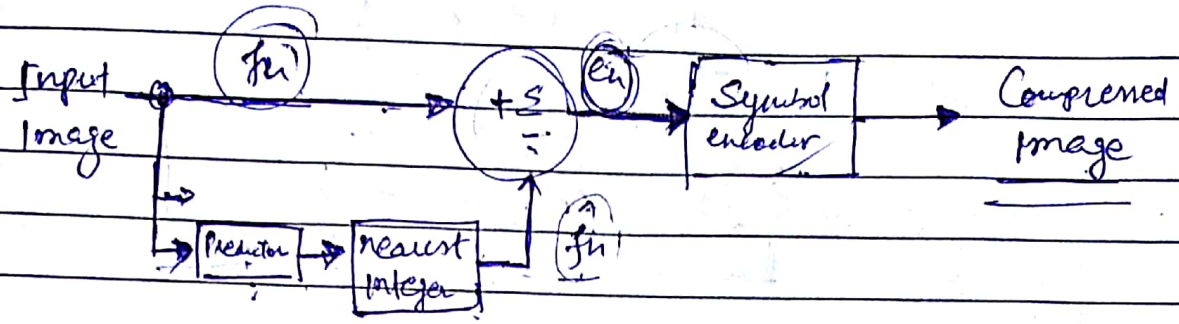
① Lossless predictive coding:-

An error-free compression approach that does not require decomposition of an image into a collection of bit planes. The approach, commonly referred to as lossless predictive coding, is based on eliminating the interpixel redundancies of closely spaced pixels by extracting and coding only the new information in each pixel.

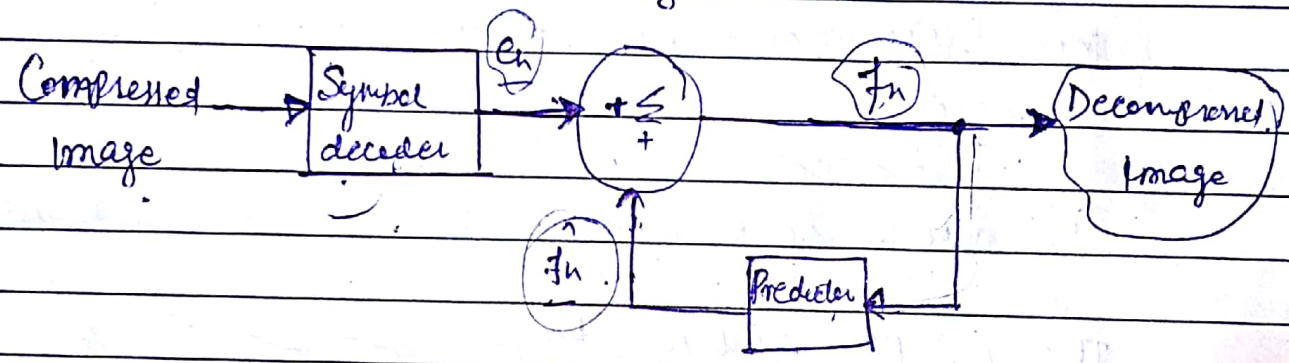
Scanned by CamScanner

The new information of a pixel is defined as the difference between the actual and predictive value of the pixel.

Basic components of a lossless predictive coding system is as follows:-



A lossless Predictive coding model: Encoder



A lossless predictive coding model: Decoder

The above system consists of an encoder and a decoder, each containing an identical predictor. As each successive pixel of the input image, denoted f_n, is introduced to the encoder, the predictor generates the anticipated value of that pixel based on some number of past inputs. The output of the predictor is then rounded to the nearest integer, denoted $\hat{f}_n$, and used to form the difference or prediction error

$$\boxed{e_n = f_n - \hat{f}_n}$$

Scanned by CamScanner

which is coded using a variable-length code (by the symbol encoder) to generate the next element of the compressed data stream.

The decoder reconstructs e_n from the received variable-length code words and performs the inverse operation

$$f_n = e_n + \hat{f}_n$$

JPEG compression

One of the most popular and comprehensive continuous tone, still frame compression standards is the JPEG standard.

CTSFC

~~Standards~~

Continuous tone, still frame compression standards :-

The CCITT and ISO have defined several continuous tone (as opposed to binary) image compression standards. These ~~plus~~ standards, which are in various phases of the adoption process, address both monochrome and color image compression.

In contrast to the binary compression standards continuous tone standards are based primarily on the lossy transform coding techniques.

* JPEG defines three different coding systems :-

1.) A lossy baseline coding system, which is based on the DCT and is adequate for most compression applications.

2.) An Extended coding system for greater compression.

Scanned by CamScanner

higher precision, or progressive reconstruction applications.

3) A lossless independent Coding system for reversible Compression.

* To be JPEG Compatible, a product or system must include support for baseline system. No particular file format, spatial resolution or color space model is specified.

Baseline System:-

In the baseline system, often called the sequential baseline system, the input and output data precision is limited to 8 bits, whereas the quantized DCT values are restricted to 11 bits.

The Compression itself is performed in three Sequential steps:-

1.) DCT Computation
2.) Quantization
3.) Variable-length code assignment.

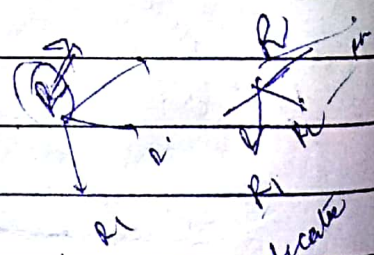
The image is first subdivided into pixel blocks of size 8x8, which are processed left to right, top to bottom. As each 8x8 block or submage is encountered its 64 pixels are level shifted by subtracting the quantity 2^n, where 2^n is the maximum no. of grey levels. The 2-D discrete cosine transform of the block is then computed, quantized in accordance with Eq.:-

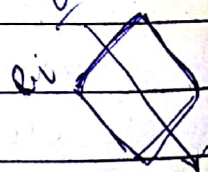
$$\hat{T}(u,v) = \text{round}\left[\frac{T(u,v)}{Z(u,v)}\right]$$

Scanned by CamScanner

and reordered, using the zigzag pattern to form a 1-D sequence of quantized coefficients.

Since the one-dimensionally reordered array generated under the zigzag pattern is qualitatively arranged according to increasing spatial frequency, the JPEG coding procedure is designed to take advantages of the long runs of zeros that normally result from the recording.





Scanned by CamScanner

Image Segmentation & Representation

Spice Date: Page No:

* Image Segmentation:-

Segmentation subdivides an image into its constituent regions or objects. The level of detail to which the subdivision is carried depends on the problem being solved.

That is, segmentation should stop when the objects or regions of interest in an application have been detected.

Let R represent the entire spatial region occupied by an image. We may view image segmentation as a process that partitions R into n subregions $R_1, R_2, R_3, \ldots, R_n$, such that

(a) $\bigcup_{i=1}^{n} R_i = R$.

(b) R_i is a connected set, $i = 1, 2, 3, 4, \ldots, n$.

(c) $R_i \cap R_j = \phi$, for all i and j, $i \neq j$.

(d) $Q(R_i) = \text{TRUE}$ for $i = 1, 2, \ldots n$.

(e) $Q(R_i \cup R_j) = \text{FALSE}$ for any adjacent regions R_i and R_j.

~~where,~~

~~$Q(R_K)$ is a predicate defined over the points in set R_K.~~

where,

$Q(R_K)$ is a logical predicate defined over the points in set R_K, and ϕ is the null set.

The symbol $\cup$ and $\cap$ represent set union and

Intersection respectively.

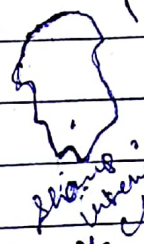
Condition (a) indicates that segmentation must be complete; that is, every pixel must be in a region.

Condition (b) requires that points in a region be connected in some predefined sense

Condition (c) indicates that the regions must be disjoint.
Condition (d) deals with the properties that must be satisfied by the pixel in a segmented region.
for ex:-
$Q(R_i)$ = TRUE if all pixels in R_i have the same intensity level.

Condition (e) indicates that two adjacent regions R_i and R_j must be different in the sense of predicate Q.

Point Line and Edge Detection



The focus of this section is on segmentation methods that are based on detecting sharp, local changes in intensity.

* The three types of image features in which we are interested are
1. Isolated points,
2. Lines
3. Edges.

Edge and Edge pixels

=> Edge pixels are pixels at which the intensity of an image function changes abruptly, and edge are sets of connected edge pixels.

Scanned by CamScanner

20

Edge detection

edge segment

Camy

Spice

Date:

Page No:

Edge detectors are local image processing methods designed to detect edge pixels.

Line

A line may be viewed as an edge segment in which the intensity of the background on either side of the line is either much higher or much lower than the intensity of the line pixels.

Detection of Isolated points

$f(x,y)$

Point detection based on second derivative.

Using Laplacian:-

$f(x,y)$

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

$$\boxed{\nabla^2 f(x,y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}}$$

$$\frac{\partial^2}{\partial x^2} f(x,y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

where the partials are obtained using Eq:-

$f(x)$

$$\frac{\partial^2 f}{\partial x^2} = f''(x) = f(x+1) + f(x-1) - 2f(x)$$

$$= f(x+1) + f(x-1) - 2f(x)$$

$$\frac{\partial^2 f(x,y)}{\partial x^2} = f(x+1, y) + f(x-1, y) - 2f(x,y) \quad (1)$$

and

$$\frac{\partial^2 f(x,y)}{\partial y^2} = f(x, y+1) + f(x, y-1) - 2f(x,y) \quad (2)$$

The Laplacian is then

$$\nabla^2 f(x,y) = f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x,y)$$

Scanned by CamScanner

Date:
Page No:

Using the Laplacian mask for detecting point

	x_1	x_2	x_3
y_1	1	1	1
y_2	1	-8	1
y_3	1	1	1

$w_2 = 1$. (x_1, y_3) (x,y)

$w_1 z_1 + w_2 z_2$

We say that a Point has been detected at the location (x,y) on which the mask is centered if the absolute value of the response of the mask at that point exceeds a specified threshold.

Such points are labeled 1 in the output image and all others are labeled 0, thus producing a binary image.

In other words, the output is obtained using the following eqn:-

$$g(x,y) = \begin{cases} 1 & \text{if } |R(x,y)| \geq T \\ 0 & \text{otherwise} \end{cases}$$

Where g is the output image, T is a nonnegative threshold, and R is given by

$$R = w_1 z_1 + w_2 z_2 + \dots + w_9 z_9 = \sum_{k=1}^{9} w_k z_k$$

$R \rightarrow$ threshold intensity

where z_k is the intensity of the pixel whose spatial location corresponds to the location of the kth coefficient in the mask.

This formulation simply measures the weighted differe

between a pixel and its 8-neighbours.

Intuitively, the idea is that the intensity of an isolated point will be quite different from its surroundings and thus will be easily detectable by this type of mask.

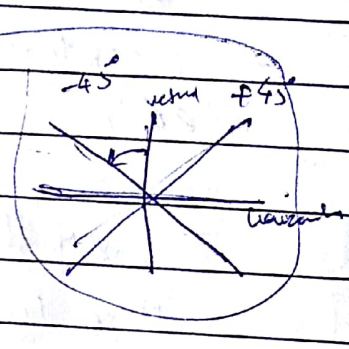
* The only differences in intensity that are considered of interest are those large enough (as determined by T) to be considered isolated points.

Line detection

- Line Masks

Line detection masks angles are with respect to the axis system.

Horizontal line
$$\begin{bmatrix} -1 & -1 & -1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix}$$



45° line
$$\begin{bmatrix} -1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & -1 \end{bmatrix}$$

vertical line
$$\begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix}$$

$|R_i| > |R_j|$

-45° line
$$\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

$|R_i| > |R_j|$

- If, at a certain point in the image, $|R_i| > |R_j|$ for all $j \neq i$, that point is said to be more likely associated with a line in the direction of mask i.

23

Edge detection

→ It locates sharp changes in the intensity function.

→ Edges are pixels where brightness changes abruptly.

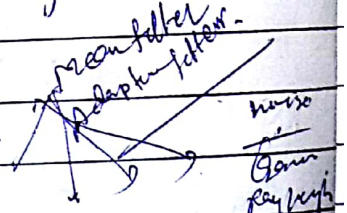
→ A change of the image function can be described by a gradient that points in the direction of the largest growth of the $f(x,y)$ image function.

→ An edge is a property attached to an individual pixel and is calculated from the image function behavior in a neighborhood of the pixel.

→ magnitude of the first derivative detects the presence of the edge.

→ Sign of the second derivative determines whether the edge pixel lies on the dark ~~side~~ side or light side.

There are three fundamental steps performed in edge detection :-



1. Image Smoothing for noise reduction :-

2. Detection of edge points :- This a local operation that extracts from an image all points that are potential candidates to become edge points.

3. Edge localization :- The objective of this step is to select from the candidate edge points that are true members of the set of points comprising an edge.

Scanned by CamScanner

Spice
Date:
Page No:

Basic Edge detection :-

⇒ Detecting changes in intensity for the purpose of finding edges can be accomplished using first - or second order derivatives.

⇒ The image gradient and its properties :-

The tool of choice for finding edge strength and direction at location (x,y). of an image, f, is the gradient, denoted by ∇f, and defined as the vector

$$\nabla f \equiv \text{grad}(f) = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}.$$

This vector has the important geometrical property that it points in the direction of the greatest rate of change of f at location (x,y).

⇒ Magnitude of vector $\nabla f(x',y')$:

$$|\nabla f(x',y')| = \left[\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2\right]^{1/2}\Bigg|_{x',y'}$$

The magnitude (length) of vector ∇f, denoted as $M(x,y)$, where

$$M(x,y) = \text{mag}(\nabla f) = \sqrt{g_x^2 + g_y^2}$$

is the value of rate of change in the direction of the greatest ~~rate of change of f at location (x,y)~~ vector.

Scanned by CamScanner

25

The direction of the gradient vector is given by the angle

$$\boxed{\alpha(x,y) = \tan^{-1}\left[\frac{g_y}{g_x}\right]}$$

measured with respect to x-axis.

Sobel operator:-

It provides both a differentiating and a smoothing effect, which is particularly attractive as derivatives typically enhance noise.

$$G_x : \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \qquad G_y : \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

Canny Edge detector

Scanned by CamScanner

Thresholding

Because of its intuitive properties, simplicity of implementation and computational speed, image thresholding enjoys a central position in applications of image segmentation.

In thresholding we discuss the techniques for partitioning images directly into regions based on intensity values and/or properties of these values.

* The basics of intensity thresholding:-

Suppose an image, $f(x,y)$, composed of light objects on a dark background, in such a way that object and background pixels have intensity values grouped into two dominant modes.

One obvious way to extract the objects from the background is to select a threshold, T, that seperates these modes. Then, any point (x,y) in the image at which $f(x,y) > T$ is called an object point; otherwise the point is called a background point.

In other words, the segmented image, $g(x,y)$ is given by

$$\boxed{g(x,y) = \begin{cases} 1 & \text{if } f(x,y) > T \\ 0 & \text{if } f(x,y) \le T \end{cases}}$$

Global thresholding → When T is a constant applicable over an entire image, the process given in this equation is referred to as global thresholding.

Variable Thresholding: When the value of T changes over an image, then this equation is referred as variable thresholding.

Scanned by CamScanner

Local or Regional thresholding:-

It is the type of variable thresholding in which the value of T at any point (x, y) in an image depends on properties of a neighbourhood of (x, y).

Dynamic or adaptive thresholding:-

If T depends on the spatial coordinates (x, y) themselves, then variable thresholding is often referred to as dynamic or adaptive thresholding.

* The Role of noise in image thresholding:-

If noise is mixed with image then we have little hope of finding a suitable threshold for segmenting this image.

* Automatic threshold technique / Algorithm:-

1.) Select an initial estimate for the global threshold, T.
2.) Segment the image using T, using Eq:-

$$g(x,y) = \begin{cases} 1 & \text{if } f(x,y) > T \\ 0 & \text{if } f(x,y) \leq T \end{cases}$$

This will produce two groups of pixels : G_1 and G_2

G_1: Consisting of all pixel with intensity values $> T$, and G_2 consisting of pixels ~~without~~ with values $\leq T$.

3.) Compute the average (mean) intensity values m_1 and m_2 for the pixels in G_1 and G_2, respectively.

Scanned by CamScanner

4) Compute a new threshold values.

$$T = \frac{1}{2}(m_1 + m_2)$$

5) Repeat step 2 through 4 until the difference between values of T in successive iterations is smaller than a predefined parameter ΔT.

Optimal Global Thresholding Using Otsu's method

Scanned by CamScanner

29

Region-based segmentation :-

- Partition an image into regions.
- Segmentation techniques that are based on finding the regions.

Region Growing

Region growing is a procedure that groups pixels or subregions into larger regions based on predefined criteria for growth. The basic approach is to start with a set of "seed" points and from these grow regions by appending to each seed those neighboring pixels that have predefined properties similar to the seed (such as specific ranges of intensity or color).

Region-growing algorithm with 8-Connectivity :-

Let $f(x,y)$:- Input image array

$S(x,y)$:- seed array containing 1's at the location of seed points and 0's elsewhere.

Q :- denote a predicate to be applied at each location (x,y) predefined property

* Arrays f and S are assumed to be of same size.

A Basic Region growing algorithm are as follows:-

1. find all connected components in $S(x,y)$ and erode each connected component to one pixels, label all such pixel found as 1. All other pixels in S are labeled 0.

Scanned by CamScanner

Spice
Date:
Page No:

(2) Form an image f_Q such that, at a pair of coordinates (x,y), let

$$f_Q(x,y) = 1$$

if the input image satisfies the given predicate, Q, at those coordinates;

otherwise,

let $f_Q(x,y) = 0$.

(3) Let g be an image formed by appending to each seed point in S all the 1-valued points in f_Q that are 8-connected to that seed point.

(4) Label each connected component in g with a different region level (e.g., 1, 2, 3 ---). This is the segmented image obtained by region growing.

Region splitting and merging:-

(1) Subdivide an image initially into a set of arbitrary, disjoint regions and then merge and/or split the regions in an attempt to satisfy the condition of segmentation.

Let R represent the entire image region and select a predicate Q. One approach for segmenting R is to subdivide it successively into smaller and smaller quadrant regions so that, for any region R_i,

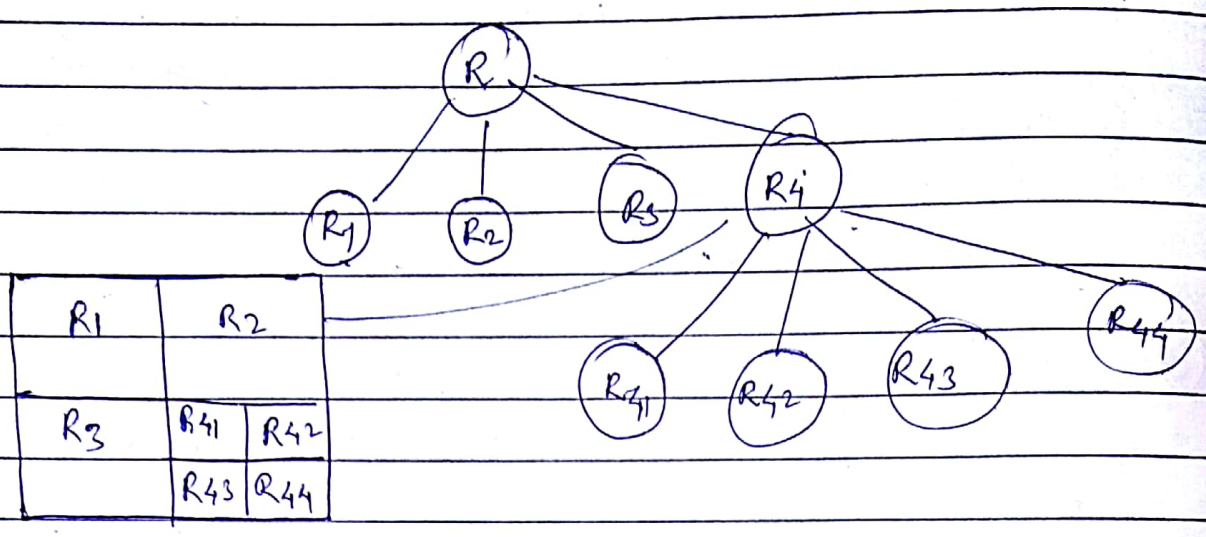
$$\boxed{Q(R_i) = \text{TRUE}}$$

We start with the entire region. If $Q(R) = \text{FALSE}$, we divide the image into quadrants.

$Q(R) = \text{FALSE}$

Scanned by CamScanner

If Q is FALSE, for any quadrant, we subdivide the quadrant into subquadrants and so on.



Region splitting and merging algorithm :-

1. Split into four ~~dyy~~ disjoint quadrants any region R_i for which $Q(R_i) = FALSE$.

2. When no further splitting is possible, merge any adjacent regions R_j and R_k for which

$$\boxed{Q(R_j \cup R_k) = TRUE}$$

3. Stop when no further merging is possible.

It is customary to specify a minimum quadregion size beyond which no further splitting is carried out.

Scanned by CamScanner

Representations :-

Boundary (Border) Following :-

Algorithm :- (Moore boundary tracking algorithm).

Assumptions :-

(1) We are working with binary images in which object and background points are labeled 1 and 0, respectively.

(2) Images are padded with a border of 0s to eliminate the possibility of an object merging with the image border.

Output :- Ordered sequence of points.

For convenience, we limit the discussion to single regions. The approach is extended to multiple, disjoint regions by processing the regions individually.

Given a binary region R or its boundary, an algorithm for following the border of R, or the given boundary, consists of following steps :-

(1) Let the starting point, b_0, be the uppermost, leftmost point in the image that is labeled 1. Denoted by c_0 the west neighbour of b_0. Clearly, c_0 always is a background point. Examine the 8-neighbour of b_0, starting at c_0 and proceeding in a clockwise direction.

Let b_1 denote the first neighbour encountered whose value is 1, and let c_1 be point immediately preceding b_1 in the sequence. Store the locations of b_0 and b_1 respectively.

Scanned by CamScanner

(2) Let $b = b_1$ and $c = c_1$.

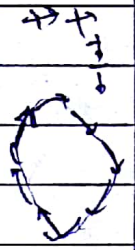
(3) Let the 8-neighbours of b, starting at c and proceeding in a clockwise direction, be denoted by $n_1, n_2, \ldots, n_8$. Find the first n_k labeled 1.

(4) Let $b = n_k$ and $c = n_{k-1}$.

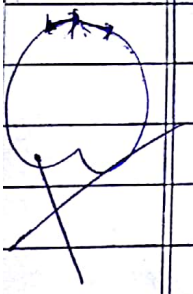
(5) Repeat steps 3 and 4 until $b = b_0$ and the next boundary point found is b_1.

The sequence of b points found when the algorithm stops constitutes the set of ordered boundary points.

(2) * Chain Codes:-

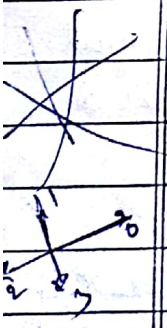


Chain codes are used to represent a boundary by a connected sequence of straight-line segments of specified length and direction. Typically this representation is based on 4 or 8-connectivity of the segments. The direction of each segment is coded by using a numbering scheme.



Freeman Chain Code:-

A boundary code formed as a sequence of such directional numbers is referred to as a Freeman chain code.



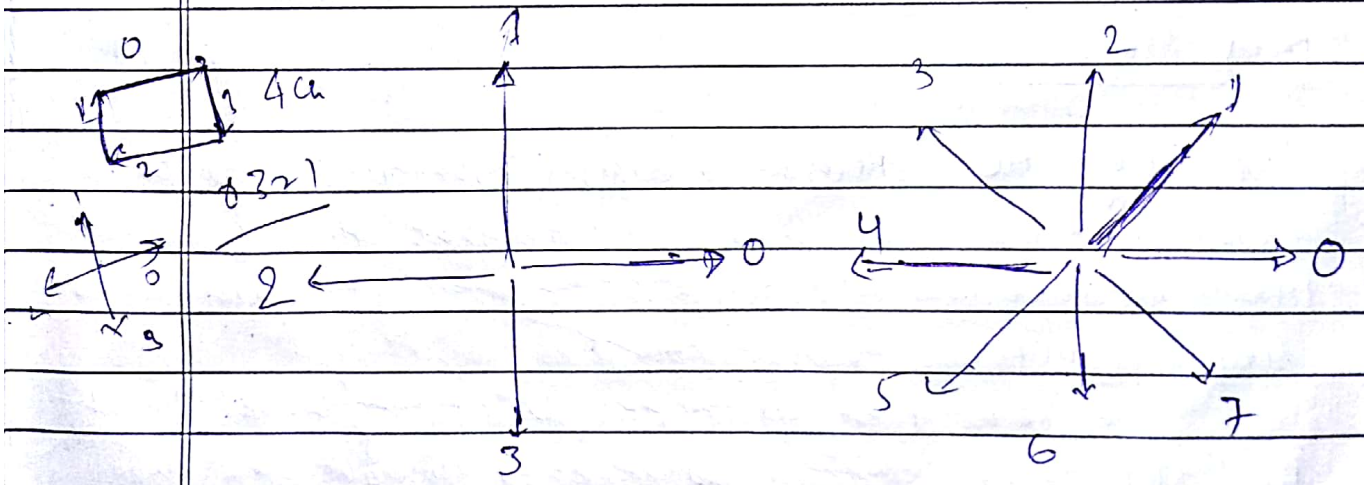
Digital images usually are acquired and processed in a grid format with equal spacing in the x- and y-directions. So a chain code can be generated by following a boundary in, say, a clockwise direction and assigning a direction to the segments connecting every pair of pixels.

This method is generally unacceptable due to two principal reasons

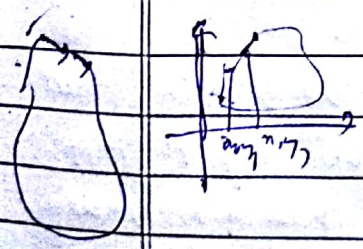
1) The resulting chain tends to be quite long
2) Any small disturbances along the boundary due to noise or imperfect segmentation cause changes in the code that may not be related to the principal shape features of the boundary.

An approach frequently used to circumvent these problems is to resample the boundary by selecting a larger grid spacing. Then as the boundary is traversed, a boundary point is assigned to each node of the large grid, depending upon the proximity of the original boundary to that node.

4 ans The resampled boundary obtained in this way then can be represented by a 4- or 8- code.



(a) 4-directional chain code (b) 8-directional chain code.



36

MPP

* Polynomial Approximations using Minimum-Perimeter Polygons:-

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 is equal to the no. of points in the boundary so each pair of adjacent points defines a segment of the polygon.

The goal of polygonal approximation is to capture the essence of the shape in the a given boundary using the fewest possible no. of segments.

One of the most powerful approximation technique is representing a boundary by a minimum-perimeter polygon (MPP).

MPP algorithm:-

The set of cells enclosing a digital boundary, ~~this~~ is called a cellular complex. We assume that the boundaries under consideration are not self intersecting, which leads to simply connected cellular complexes.

Based on these assumptions, and letting White (W) and Black (B) denote convex and mirrored concave vertices, respectively, we state the following observations:-

1.) The MPP bounded by a simply connected cellular complex is not self intersecting.

2) Every convex vertex of the MPP is a W vertex, but not every W vertex of a boundary is a vertex of the MPP.

3.) Every mirrored concave vertex of the MPP is a B vertex, but not every B vertex of a boundary is a vertex of MPP.

Scanned by CamScanner

Spice
Date:
Page No:

(4) All B vertices are on or outside the MPP, and all W vertices are on or inside the MPP.

(5) The uppermost, leftmost vertex in a sequence of vertices contained in a cellular complex is always a W vertex of the MPP.

Scanned by CamScanner

Boundary Descriptors:-

- length
- Diameter
- Minor axis
- eccentricity
- Spice
- Curvature
- Shape no.
- Fourier

Date:
Page No:

37

Several approaches to describing the boundary of a region.

Some simple Descriptors:-

The length of a boundary is one of its simplest descriptors. The number of pixels along a boundary gives a rough approximation of its length.

For a chain coded curve with unit spacing in both directions, the no. of vertical and horizontal components plus $\sqrt{2}$ times the no. of diagonal components gives its exact length.

length of boundary

The diameter of boundary B is defined as

no. of pixel $n(V)+n(H)+\sqrt{2}(d)$

$$\boxed{Diam(B) = \max_{i,j} \left[D(P_i, P_j) \right]}$$

where D = distance measure ~~and~~
P_i, P_j = points on the boundary.

Imp:- The value of diameter and the orientation of a line segment connecting the two extreme points that comprise the diameter are useful descriptor of a boundary.

Minor Axis:- Minor axis is defined as the line perpendicular to the major axis, and of such length that a box passing through the outer four points of intersection of the boundary with the two axis completely encloses the boundary.

Scanned by CamScanner

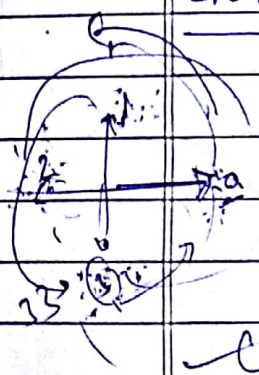
Eccentricity of boundary = major axis (a)
minor axis . (b)

Curvature :- Rate of change of slope.

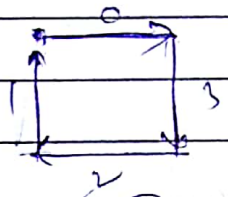
Shape numbers :-

The first difference of a chain-coded boundary depends on starting point. The shape no. of such a boundary based on 4-directional code, is defined as the first difference of smallest magnitude.

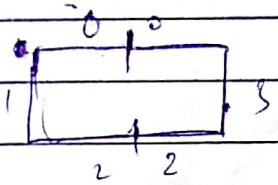
The order n of shape no. is defined as the number of digits in the representation. Moreover, n is even for closed boundary, and its value limits the no. of possible different shapes.

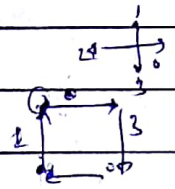


Order 4



Order 6

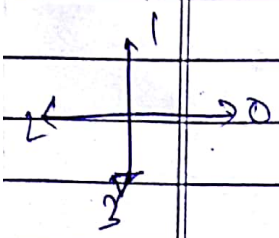


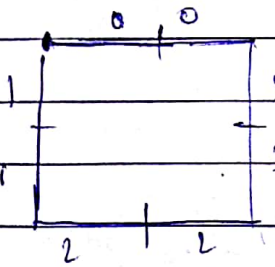


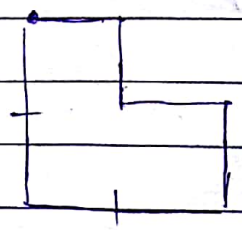
0321
1032

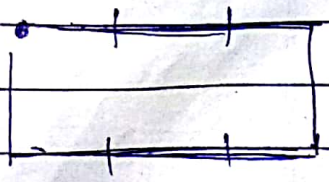
Chain code: 0 3 2 1 | 0 0 3 2 2 1
Difference: 3 3 3 | 3 0 3 3 0 3
Shape no: 3 3 3 3 | 0 3 3 0 3 3

Order 8









Chain code:	0 0 3 3 2 2 1 1	0 3 0 3 2 2 1 1	0 0 0 3 2 2 2 1
Difference:	3 0 3 0 3 0 3 0	3 3 1 3 3 0 3 0	3 0 0 3 3 0 0 3
Shape no:	0 3 0 3 0 3 0 3	0 3 0 3 3 1 3 3	0 0 3 3 0 0 3 3

Scanned by CamScanner

19

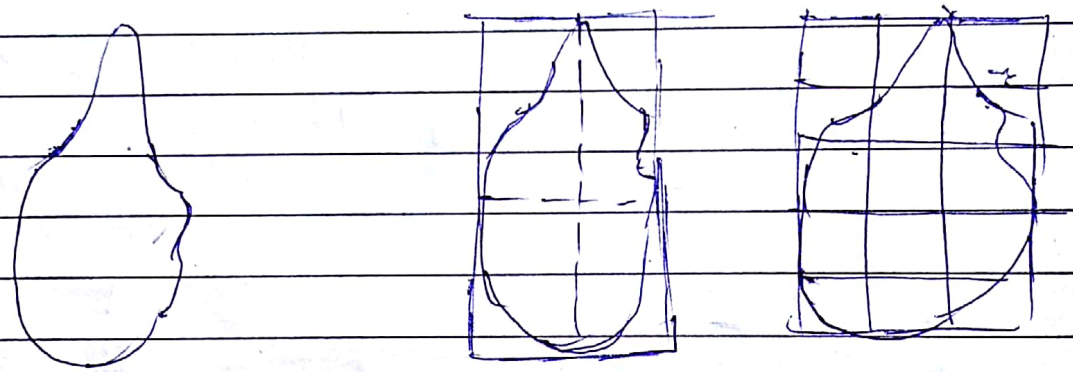
In Practice, for a desired shape order, we find the rectangle of order n whose eccentricity best approximate that of the basic rectangle and use this new rectangle to establish the grid size.

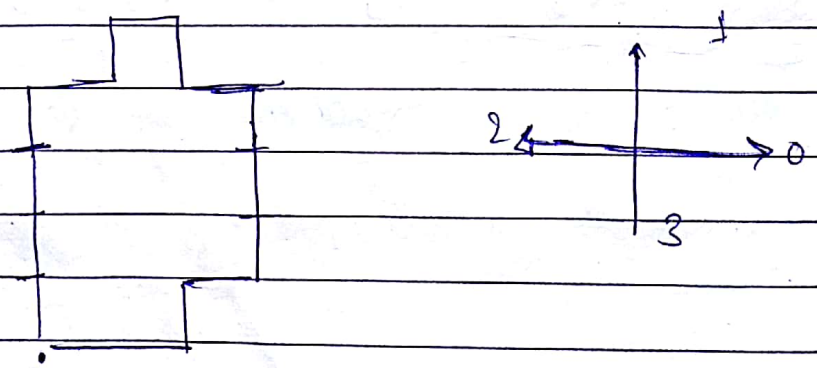
For ex:- If n=12, all rectangles of order 12 are

2x4, 3x3, and 1x5

If the eccentricity of the 2x4 rectangle best matches the eccentricity of the base rectangle for a given boundary, we establish a 2x4 grid centered on the base rectangle and use the procedure to obtain chain code.

Steps in the generation of a shape number





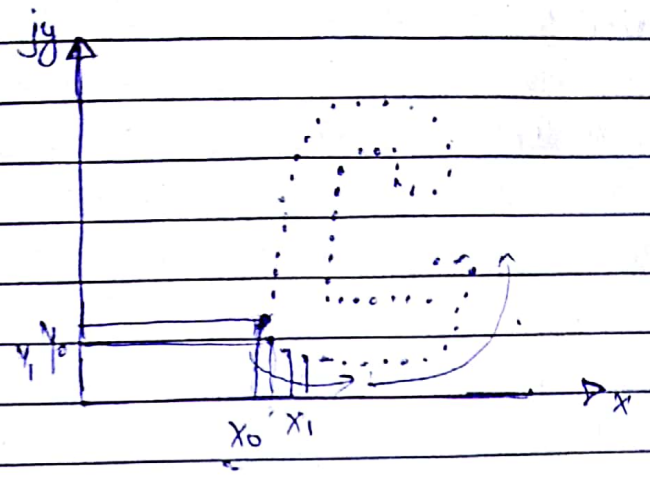
Chain code: 0 0 0 0 3 0 0 3 2 2 3 2 2 2 1 2 1 1

~~3 0 0 0 3 1 0 3 3 0 1 3 0 0~~

Difference code: 3 0 0 0 3 1 0 3 3 0 1 3 0 0 3 1 3 0

Shape no: 0 0 0 3 1 0 3 3 0 1 3 0 0 3 1 1 3, 0, 3

Fourier Descriptors



x_{k-1}, y_{k-1}

The above fig shows a k-point digital boundary in the xy-plane. Starting at an arbitrary point (x_0, y_0), coordinate pairs $(x_0, y_0), (x_1, y_1), (x_2, y_2) \ldots (x_{k-1}, y_{k-1})$ are encountered in traversing the boundary say, in the counterclockwise direction.

These coordinates can be expressed in the form

$x(k) = x_k$
$y(k) = y_k$
$s(k) = [x(k), y(k)]$

$\cancel{x(k) = y_k}$ $x(k) = x_k$ and $y(k) = y_k$.

$x(k) = x_k$
$y(k) = y_k$

With this notation boundary itself represented as the sequence of coordinates

$s(k) = x(k) + j\,y(k)$

$$s(k) = [x(k), y(k)]$$

for $k = 0, 1, 2 \ldots, k-1$.

$x + iy0$

Moreover each coordinate pair can be treated as ~~the~~ a complex no. so that

$$\boxed{s(k) = x(k) + j\,y(k)}$$

for $k = 0, 1, 2 \ldots k-1$.

x axis → Real axis
y-axis → imaginary axis of sequence of complex no.

This representation has one advantage. It reduces 2-D to a 1-D problem.

Scanned by CamScanner

$$f(x,y) = \sum f(u,v)\, e^{-j2\pi(ux+vy)/N}$$

Spice
Date:
Page No:

41

The discrete fourier transform (DFT) of $s(k)$ is :-

$$a(u) = \sum_{k=0}^{K-1} s(k)\, e^{-j2\pi uk/K}$$

$$\sum f(x,y) e^{-j2\pi(ux+vy)/N}$$

$s(k)$

for $u = 0, 1, 2, ---, K-1$

The complex coefficient $a(u)$ are called the fourier descriptors of the boundary. The inverse fourier transform of these coefficients restores $s(k)$.

$$s(k) = \frac{1}{K} \sum_{u=0}^{K-1} a(u)\, e^{j2\pi uk/K}$$

$\frac{1}{K}$

for $k = 0, 1, 2, --- K-1$

Suppose, however, that instead of all the fourier coefficients, only the first P coefficients are used. This is equivalent to setting

$$a(u) = 0 \text{ for } u > P-1$$

The result is following approximation to $s(k)$:

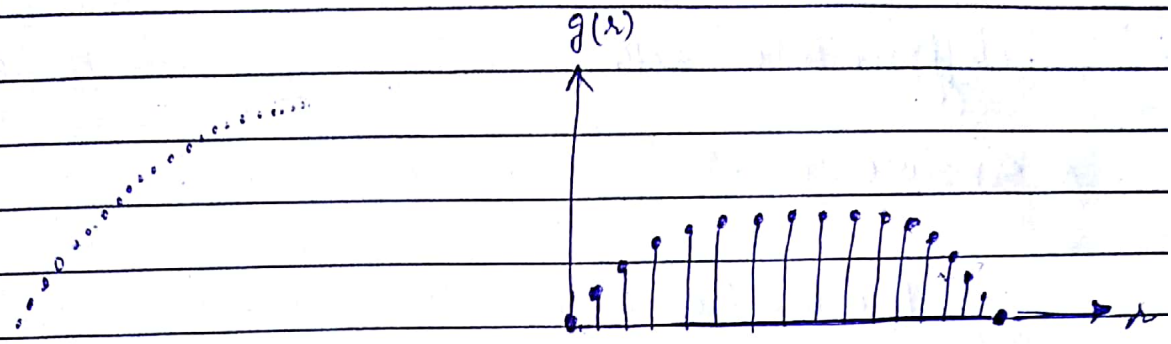
$$\hat{s}(k) = \frac{1}{P} \sum_{u=0}^{P-1} a(u)\, e^{j2\pi uk/P}$$

for $k = 0, 1, 2, ---, K-1$.

Scanned by CamScanner

Statistical Method :-

The shape of boundary segments (and of signature waveforms) can be described quantitatively by using statistical moments. Such as mean, variance and higher order moments.



(a) Segment of the boundary

(b) Segment represented as 1-D function $g(r)$ of an arbitrary variable r.

Suppose,

the amplitude of g as a discrete random variable v and form an amplitude histogram ~~p(v)~~.

$\boxed{p(v_i),\ i = 0, 1, 2, \ldots, A-1}$ where A is the no. of discrete amplitude increments in which we divide the amplitude scale.

~~Then~~ Since $p(v_i)$ is an estimate of the probability of value v_i occuring,

the nth moment of v about its mean is

$$\mu_n(v) = \sum_{i=0}^{A-1} (v_i - m)^n \, p(v_i)$$

Where,

$$m = \sum_{i=0}^{A-1} v_i \, p(v_i)$$

Scanned by CamScanner

m: mean or average value of v

μ_2: variance

Generally, only the first few moments are required to differentiate between signature of clearly distinct shapes.

Alternative approaches-

Normalize $g(r)$ to unit area and treat it as a histogram.

In other words, $g(r_i)$ is now treated as the probability of value r_i occuring.

In this case, r is treated as the random variable and the moments are

$$\mu_n(r) = \sum_{i=0}^{K-1} (r_i - m)^n g(r_i)$$

Where,

$$m = \sum_{i=0}^{K-1} r_i g(r_i)$$

K: no. of points on the boundary

$\mu_n(r)$: Directly related to shape of $g(r)$.

Scanned by CamScanner

Regional Descriptors :-

Some simple descriptors :-

Area: The area of region is defined as the no. of pixels in the region.

Perimeter: - The perimeter of a region is the length of its boundary.

Uses of area and perimeter as descriptors:-

(1) Area and perimeter are sometimes used as descriptors, they apply primarily to the situations in which size of the regions of interest is invariant.

(2) A more frequent use of these two descriptors is in measuring compactness of a region, defined as

$$\text{Compactness} = \frac{(\text{perimeter})^2}{\text{area}}$$

A slightly different descriptor of compactness is the circulatory ratio.

$$\text{circulatory ratio} = \frac{\text{Area of the region}}{\text{Area of the circle (the more compact shape) having the same perimeter}}$$

most compact

The area of circle having perimeter length p is

$$\frac{p^2}{4\pi} \qquad \frac{p(2\pi r)^2}{4\pi}$$

45

Therefore, the circulatory ratio, Rc is given by the expression

$$\boxed{R_c = \frac{4\pi A}{P^2}}$$

where, A: It is the area of the region in question and
P: length of the descriptors.

Topological Descriptors:-

Topological properties are useful for global descriptions of regions in the image plane.

Topology is 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.

* Topological property useful for region description is as follows:-

1) No. of holes in the image.

2) no. of connected components.

Connected Components:-

A Connected Component of a region was defined as follows.

Two pixels p and q are said to be connected in S if there exist a path between them consisting entirely

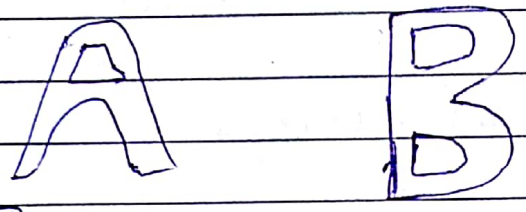
of pixels in S.

The no. of Holes (H) and Connected Components C in a figure can be used to define Euler no E:

$$\boxed{E = C - H}$$

E: Euler no.
C: no. of Connected Components.
H: no. of holes in the image.

Euler no. is also a topological property.



Euler no. $E = C - H$
$= 1 - 1$
$= 0$

$E = C - H$
$= 1 - 2 = -1$

Texture:-

An important approach to region description is to quantify its texture content.

Although no formal definition of texture exists, intuitively this descriptor provides measures of properties such as smoothness, coarseness, and regularity.

* The three approaches used in image processing to describe the texture of a region are
Statistical, structural, spectral.

Scanned by CamScanner

Statistical :- This approaches yields characterizations of textures as smooth, coarse, grainy and so on.

Structural :- This technique deals with the arrangement of image primitives, such as description of texture based on regularly spaced parallel lines.

Spectral :- Spectral techniques are based on properties of the Fourier spectrum and are used primarily to detect global periodicity in an image by identifying high-energy, narrow peaks in spectrum.

Moment Invariants

The 2-D moment of order $(p+q)$ of a digital image $f(x,y)$ of size $M \times N$ is defined as

$$\boxed{m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q f(x,y)}$$

Where $p = 0, 1, 2 \ldots$ and $q = 0, 1, 2, \ldots$

The corresponding central moment of order $(p+q)$ is defined as

$$\mu_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (x-\bar{x})^p (y-\bar{y})^q f(x,y)$$

for $p = 0, 1, 2 \ldots$ and $q = 0, 1, 2 \ldots$, where

$$\bar{x} = \frac{m_{10}}{m_{00}} \text{ and } \bar{y} = \frac{m_{01}}{m_{00}}$$

Scanned by CamScanner

Spice
Date:
Page No:

The normalized central moments, denoted by η_{pq}, are defined as

$$\boxed{\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}}}$$

where,

$$\gamma = \frac{p+q}{2} + 1$$

for $p+q = 2, 3 \ldots$

Scanned by CamScanner

Edge Linking and boundary detection

Set of pixels from edge detecting algorithms, seldom define a boundary completely because of noise, breaks in the boundary etc.

Therefore, edge detecting algorithms are typically followed by linking and other detection procedures, designed to assemble edge pixels into meaningful boundaries.

Types → local
Types → global

Local Processing

Analyse the characteristics of pixels in a small neighbourhood (3x3, or 5x5) about every point that has undergone edge detection.

All points that are similar are linked, forming a boundary of pixels that share common properties.

* Two principal properties for establishing similarity of edge pixels:-

1. Strength of the response of the gradient operator used to produce the edge pixel.
2. Direction of gradient.

A point in the predefined neighbourhood of (x,y) is linked to the pixel at (x,y) if both magnitude and direction criteria are satisfied.

Scanned by CamScanner

50

This process is repeated for every location in the image.

EDGE LINKING AND BOUNDARY DETECTION USING HOUGH TRANSFORM

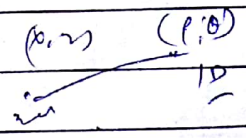
* The Hough Transform is a technique which can be used to isolate features of a particular shape with image.

* Hough Transform is most commonly used for the detection of regular curves such as line, curves, circles, ellipses etc.

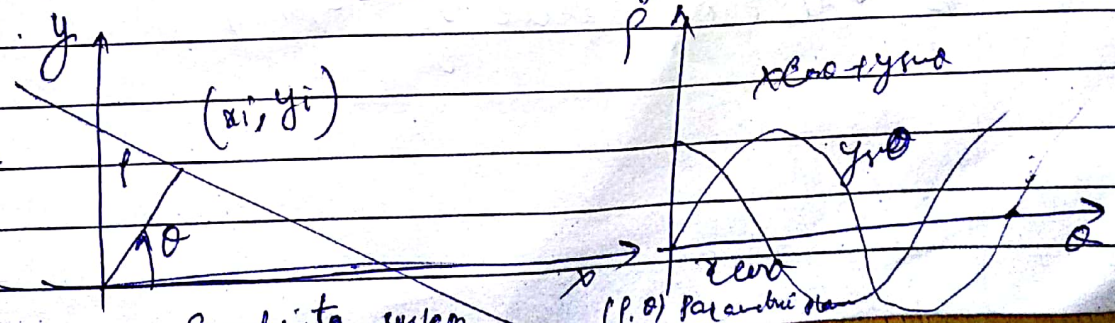
* The main advantage of Hough Transformation technique is that it is tolerant of gaps in feature boundary descriptions and is relatively unaffected by image noise.

* It maps a straight line $y = mx + c$ in a cartesian coordinate system into a single point in the (ρ, θ) plane or

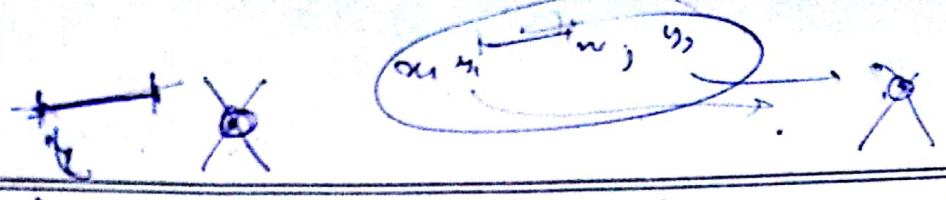
$$\boxed{\rho = x\cos\theta + y\sin\theta}$$



For a point (x, y) in a cartesian coordinate plane, there will be an infinite no. of curves in the (ρ, θ) plane.



Cartesian xy Coordinate system (ρ, θ) Parametric Plane



When two points (x_i, y_i) and (x_j, y_j) lie on the same straight line, the curves in the (ρ, θ) plane which correspond respectively, to the two points (x_i, y_i) and (x_j, y_j) in the cartesian plane will intersect at a point.

This intersection point determines the parameter of the line that joins these two points.

Similar arguments apply for the three collinear points.

This property between cartesian plane and the parametric plane will be useful in finding the line that fits points in the xy plane.

The Hough Transform approach is to find the points of intersection in the curves, each of which corresponds to a line in the cartesian xy plane.

* The Hough transform can be generalized to detect circles as described by

$$\boxed{(x-a)^2 + (y-b)^2 = r^2}$$

Now we get three parameters a, b and r to parametrize a circle. This result in a three dimensional accumulator array, where each accumulator bin is indexed by a discrete value of (a, b, r).

Scanned by CamScanner