IMAGE SEGMENTATION

→ Segmentation is another major step in Digital image processing. Segmentation subdivides an image into its constituent regions or objects.
The level to which the subdivision is carried depends on the problem being solved. i.e segmentation should stop when the objects of interest in an application have been isolated.

→ Image segmentation algorithms generally are based on one of two basic properties of intensity values:
① Discontinuity ② Similarity.
In first category the approach is to partition an image based on abrupt changes in intensity such as edges in an image.
→ The principal approaches in second category are based on partitioning an image into regions that are similar according to a set of ~~de~~ predefined criteria.
Example thresholding, region growing and region splitting and merging are examples of m

① Detection of Discontinuities:-
Several techniques are proposed for detecting. the three basic types of gray-level discontinuities in a digital image
① points
② lines
③ Edges.
The most common way to look for discontinuities is to run a mask through the image. is

→ For Example 3x3 mask. given below

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

The procedure involves computing the sum of products of the coefficient with the gray levels contained in the region encompassed by the mask.

→ The response of the mask at any point in the image is given by

$$R = w_1 z_1 + w_2 z_2 + \cdots\cdots + w_9 z_9 \quad \text{--- (i)}$$

$$= \sum_{i=1}^{9} w_i z_i$$

where z_i is the gray level of the pixel associated with mask coefficient w_i.

Point Detection:-

The detection of isolated points in an image is straight foward in principle.

→ Mask for point processing is

-1	-1	-1
-1	8	-1
-1	-1	-1

A point has been detected at the Location on which the mask is centered if

$$|R| \geq T \quad \text{--- (2)}$$

where T is a non negative threshold and R is given as (1)

This formulation measures the weighted differences between the center point and its neighbors.

→ The idea is that an iso to find the Isolated point.

→ Def: Isolated point – a point whose gray level is significantly different from its background and which is located in a homogeneous area.

→ The mask used for point detection is in connection with Laplacian operations.

→ The mask coefficients sum to zero indicating that the mask response will be zero in areas of constant gray level.

Line Detection:-

The masks for Line Detection are shown below.

Horizontal

-1	-1	-1
2	2	2
-1	-1	-1

+45°

-1	-1	2
-1	2	-1
2	-1	-1

Vertical

-1	2	-1
-1	2	-1
-1	2	-1

-45°

2	-1	-1
-1	2	-1
-1	-1	2

→ If the Horizontal mask were moved around an image it would respond more strongly to lines oriented horizontally.

→ The coefficients in each mask sum to zero, indicating a zero response from the masks in areas of constant gray levels.

→ Let R_1, R_2, R_3 and R_4 denote the response of the masks shown above. Suppose four masks are run individually through an image.

→ If 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. Example

→ Example if $|R_i| > |R_j|$ for $j = 1, 2, 3, 4$ that particular point is said to be more likely associated with a horizontal line.

Edge Detection:-

Edge detection is by far the most common approach for detecting meaningful discontinuities in gray level.

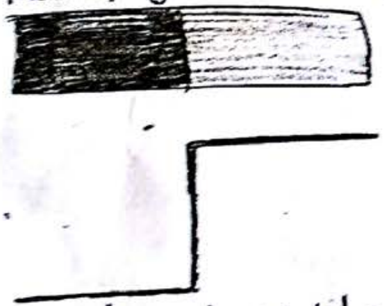
→ For detecting edges in an image we the approaches are first and second-order digital derivatives.

Basic formulation:-

→ An Edge is a set of connected pixels that lie on the boundary between two regions.

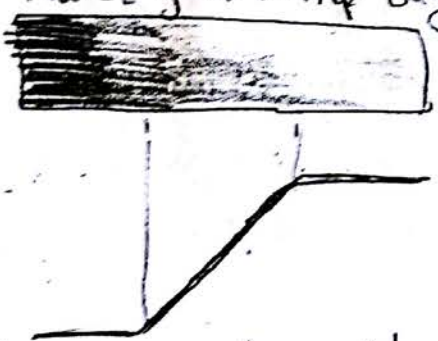
→ An ideal edge has the properties of the model shown in figure below.

Model of an ideal digital edge.



(a) Gray-level profile of a horizontal line through the image.

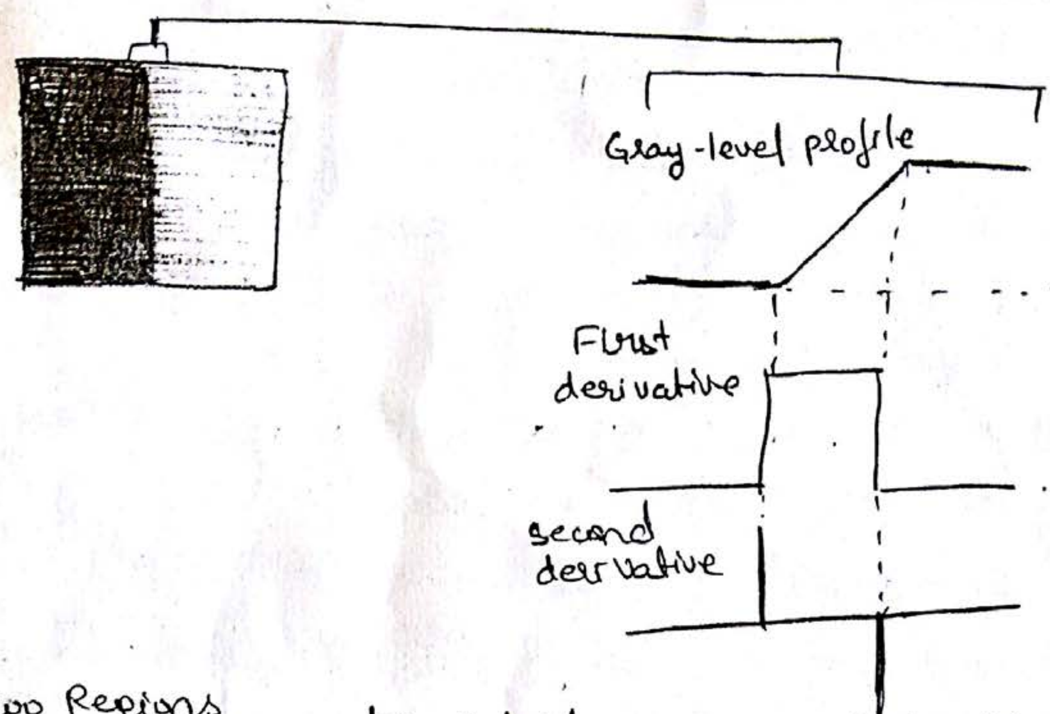
Model of a ramp digital edge



(b) Gray-level profile of a horizontal line through the image.

→ In practice, optics, sampling and other image acquisition imperfections yield edges that are blurred shown (b) figure. The edges are more closely modeled as having a 'ramplike' profile- ~~such~~ as the

→ The slope of the ramp is inversely proportional to the degree of blurring in the edge.



a) Two Regions separated by a vertical Edge

b) Detail near the edge, showing a gray level profile and the first and second derivatives of the profile.

→ The first derivative is positive at the points of transition into and out of the ramp as we move from left to right along the profile. it is constant for points in the ramp and is zero in areas of constant gray level.

→ the second derivative is positive at the transition associated with the dark side of the edge, negative at the transition associated with the ~~of~~ light side of the edge and zero along the ramp and in areas of constant gray level.

→ The signs of the derivatives would be reversed for an edge that transitions from light to dark.

→ From these observations ~~the~~ the magnitude of the first derivative can be used to detect the presence of an edge at a point in an image.

→ Similarly the sign of the second derivative can be used to determine whether an edge pixel lies in the dark or light side of an edge.

two Additional properties of the second derivative around an edge.

① It produces two values for every edge in an image

② An imaginary straight line joining the extreme positive and negative values which cross zero near ~~the~~ ~~two~~ the midpoint of the edge.

This zero-crossing property of the second derivative is useful ~~to~~ for locating the center

Gradient Operators:-

The gradient of an image $f(x,y)$ at location (x,y) is defined as the vector

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

the magnitude of this vector denoted ∇f where

$$\nabla f = mag(\nabla f) = \left[G_x^2 + G_y^2 \right]^{1/2}$$

The direction of the gradient vector is given as:

Let $\alpha(x,y)$ represent the direction angle of the vector ∇f at (x,y) then

$$\alpha(x,y) = \tan^{-1}\left(\frac{G_y}{G_x}\right)$$

where angle is measured with respect to the x-axis.

→ The direction of an edge at (x,y) is perpendicular to the direction of the gradient vector at that point.

→ For 3x3 area the ~~g~~ simplest way to implement a first order partial derivative at point z_5 is using the Roberts cross-gradient operators

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

Roberts

-1	0
0	1

0	-1
1	0

$$G_x = z_9 - z_5 \qquad G_y = z_8 - z_6.$$

Mask of size 2x2 are awkward to implement because they do not have a clear center. so we prefer 3x3 mask.

An approach using 3x3 masks is given by.

G_x

-1	-1	-1
0	0	0
1	1	1

G_y

-1	0	1
-1	0	1
-1	0	1

$$G_x = (z_7 + z_8 + z_9) - (z_1 + z_2 + z_3)$$

$$G_y = (z_3 + z_6 + z_9) - (z_1 + z_4 + z_7)$$

The masks shown above are called Prewitt operators.

→ A slight variation of the above equations using a weight of 2 in the center coefficient:

-1	-2	-1
0	0	0
1	2	1

-1	0	1
-2	0	2
-1	0	1

$$G_x = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)$$

$$G_y = (z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)$$

A weight value of 2 is used to achieve smoothing. and the operators are called the Sobel operators.

The Prewitt masks are simpler to implement.
→ Sobel masks have slightly superior noise-suppression characteristics.

→ An approach used frequently is to approximate the gradient by absolute values:

$$\nabla f \approx |G_x| + |G_y|.$$

→

0	1	1
-1	0	1
-1	-1	0

-1	-1	0
-1	0	1
0	1	1

Prewitt

0	1	2
-1	0	1
-2	-1	0

-2	-1	0
-1	0	1
0	1	2

Sobel

Fig Prewitt and Sobel masks for detecting Diagonal Edges.

The Laplacian:-

The Laplacian of a 2-D function $f(x,y)$ is a second-order derivative defined as

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

0	-1	0
-1	4	-1
0	-1	0

-1	-1	-1
-1	8	-1
-1	-1	-1

Laplacian Masks

$$\nabla^2 f = 4z_5 - (z_2 + z_4 + z_6 + z_8)$$

$$\nabla^2 f = 8z_5 - (z_1 + z_2 + z_3 + z_4 + z_6 + z_7 + z_8 + z_9)$$

→ The Laplacian generally not used in its original form for edge detection for several reasons like sensitive to noise, produces double edges, are unable to detect edge direction.

→ So the Laplacian is combined with smoothing as a precursor to find edges via zero-crossing.

Consider the function

$$h(r) = -e^{-\frac{r^2}{2\sigma^2}} \quad —①$$

where $r^2 = x^2 + y^2$

σ is the standard deviation

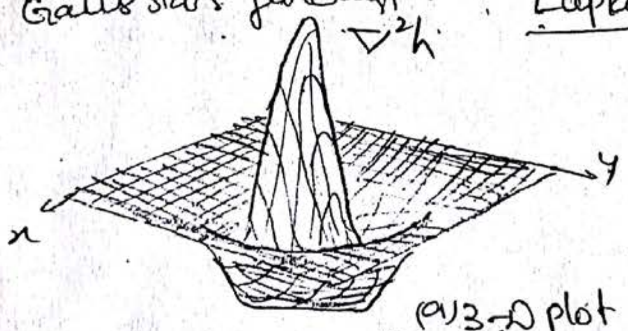
Convolving this function with an image blurs the image, with degree of blurring is determined by the value of σ.

The Laplacian of h is

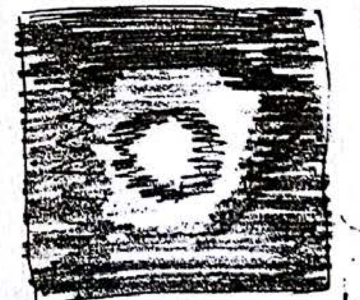
$$\nabla^2 h(r) = -\left[\frac{r^2 - \sigma^2}{\sigma^4}\right] e^{-\frac{r^2}{2\sigma^2}} \quad —②$$

This function is commonly referred to as Laplacian of a Gaussian (LoG) because the eqn ① is in the form of a Gaussian function.

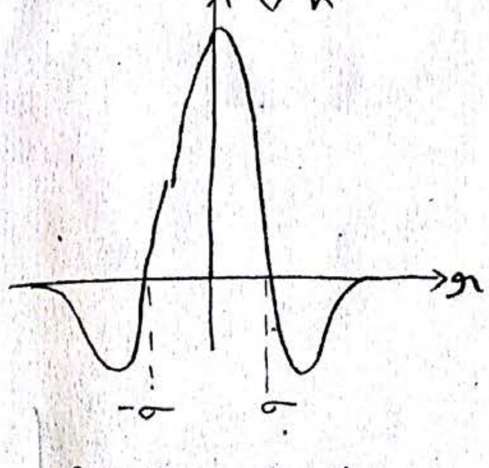
Laplacian of a Gaussian (LoG)



(a) 3-D plot



(b) Image



(c) Cross section

0	0	-1	0	0
0	-1	-2	-1	0
-1	-2	16	-2	-1
0	-1	-2	-1	0
0	0	-1	0	0

(d) 5x5 mask

(2) Edge Linking and Boundary Detection :-

The methods discussed till now yield pixels only on edges.

→ But due to Noise, breaks in the edges form nonuniform illumination and other effects that introduce intensity discontinuities.

→ To remove these intensity discontinuities the edge detection algorithms are followed by Linking procedures to assemble edges pixels into meaningful edges.

Local Processing :-

→ One of the simplest approaches for Linking edge points is to analyze the characteristics of pixels in a small neighborhood (say 3×3, or 5×5) about every point (x,y) in an image that has been labeled an edge point

→ All points that are similar according to a set of predefined criteria are linked forming an edge of pixels that share those criteria.

→ The two principal properties used for establishing similarity of edge pixels are :

(1) The strength of the response of the gradient operator used to produce the edge pixel

(2) The direction of the gradient vector.

→ from first property the given values are of ∇f. ~~defined~~

→ Thus an edge pixel with coordinates (x_0, y_0) in a predefined neighborhood of (x,y) is similar in magnitude to the pixel at (x,y) if

$$\boxed{\left| \nabla f(x,y) - \nabla f(x_0,y_0) \right| \leq E}$$

where E is a non negative threshold.

→ The direction (angle) of the gradient vector is given by

$$\boxed{\alpha(x,y) = \tan^{-1}\left(\frac{G_y}{G_x}\right)}$$

An edge pixel at (x_0, y_0) in the predefined neighborhood of (x,y) has an angle similar to the pixel at (x,y) if

$$\boxed{|\alpha(x,y) - \alpha(x_0, y_0)| < A}$$

where A is a non negative angle threshold.

→ A point in the predefined neighborhood of (x_0, y_0) is linked to the pixel at (x,y) if both magnitude and direction criteria are satisfied.

This process is repeated at every location in the image.

(3) Thresholding :-

Image thresholding plays an important role in applications of image segmentation.

Foundation :-

Suppose the gray-level histogram of an image $f(x,y)$ is taken which is composed of light objects on a dark background.

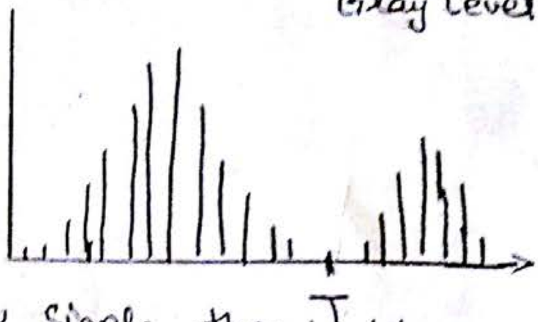
→ In such cases the object and background pixels have gray levels in two dominant modes.

→ One way to extract the objects from background is to select a threshold T that separates these modes.

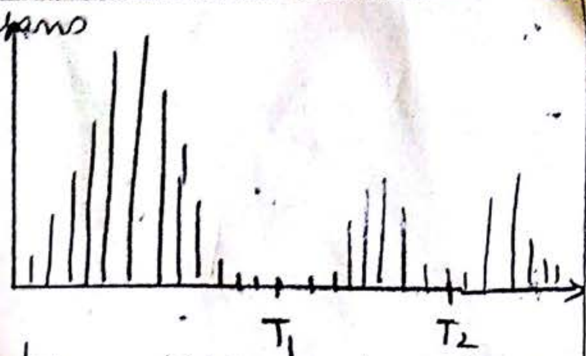
→ Any point (x,y) from which $f(x,y) > T$ is called an object point.

otherwise the point is called a background point.

Gray level histograms



a) single threshold



b) multiple threshold.

In fig(b) we see there are three dominant modes

→ Multiple thresholding classifies a point (x,y) as belonging to one object class if $T_1 < (x,y) \leq T_2$.

other object class if $f(x,y) > T_2$

background if $f(x,y) \leq T_1$.

→ Thresholding involves a test against a function T of the form

$$T = T[x, y, p(x,y), f(x,y)]$$

where $f(x,y)$ is the gray level of point (x,y)

$p(x,y)$ denotes some local properties of this point

A thresholded image $g(x,y)$ is defined as

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

The pixels labeled 1 are objects

pixels labeled 0 are background.

When T depends only on $f(x,y)$ the threshold is called global.

→ If T depends on both $f(x,y)$ and $p(x,y)$ the threshold is called Local.

→ If T depends on the spatial coordinates x and y the threshold is called dynamic or adaptive.

(c) indicates that regions R_i and R_j are different in the sense of predicate P.

Region Growing :-

Region Growing is a procedure that ~~pto~~ groups pixels or sub regions into Larger regions based on predefined criteria.

→ The basic approach is to start with a set of "seed" points and from these grow regions by appending to each seed those neighbouring pixels that have properties similar to the seed (ex gray level or color)

→ ~~selecting~~ Selecting a set of one or more starting points often can be based on the nature of the problem.

→ when a priori information is not available, the procedure is to compute at every pixel the same set of properties that ultimately will be used to assign pixels to regions during the growing process.

→ The result of these computation shows clusters of values. where the centroid of these cluster can be used as seeds.

→ The selection of seed point depend also on type of image.

→ Formulation of a stopping rule is another important criteria in region growing.

→ Basically Region growing should stop when no more pixels satisfy the criteria for inclusion in that region.

(4) Region -Based Segmentation :-

The objective of Segmentation is to partition an image into regions.

→ Till Now the methods discussed are to find boundaries ~~based on~~ between regions based on discontinuities.

→ But now we discuss segmentation techniques that are based on finding the regions directly.

→ Basic Formulation.

Let R represent the entire Image region.

The segmentation is a process that partitions R into n Sub regions $R_1, R_2 --- R_n$ such that

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

b) R_i is a connected region $i = 1, 2, --- n$.

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

d) $P(R_i) =$ TRUE for $i = 1, 2, ----- n$.

e) $P(R_i \cup R_j) =$ FALSE for $i \neq j$.

Here $P(R_i)$ is a Logical predicate defined over the points in set R_i and ϕ is the null set.

Conditions:-

a) Indicates that the segmentation must be complete.

b) requires that points in a region must be connected in some predefined sense.

c) indicates that the regions must be ~~dis~~ disjoint

d) deals with the properties that must be satisfied by the pixels in a segmented region.

For Example. $P(R_i)$ = True if all pixels in R_i have the same gray level.

Region Splitting and Merging :-

An alternative procedure ~~to~~ of region growing is to subdivide an image initially into a set of arbitrary, disjointed regions and then merge and /or split the regions in an attempt to satisfy the condition.

→ A split and merge algorithm that ~~that~~ iteratively works toward satisfying these constraints. has to be discussed next.

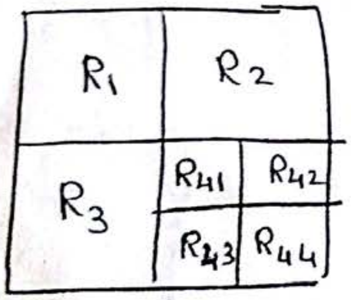
Let R represent the entire image region and select a predicate P.

→ One approach for segmenting R is to subdivide it successively into smaller and smaller of quadrant regions so that for any region R_i, $P(R_i)$ = True.

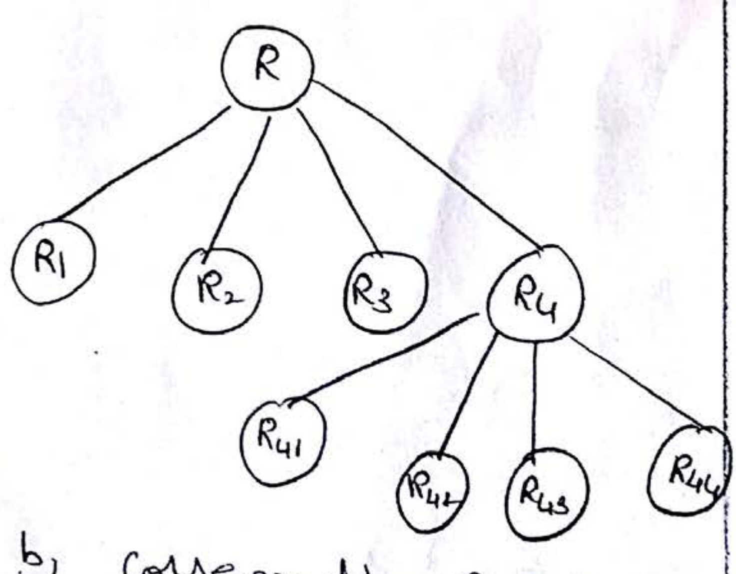
→ We start with the entire region.

If $P(R)$ = False, we divide the image into ~~gra~~ quadrants if P is False for any quadrant, we subdivide that quadrant into subquadrants and so on.

This particular splitting technique is called quad-tree



a, Partitioned Image



b, Corresponding Quadtree.

→ If only splitting were used the final partition likely would contain adjacent regions with identical properties.

→ This can be overcomed by merging as well as splitting..

→ Two adjacent regions R_j and R_K are merged only if

$P(R_j \cup R_K) = \text{True}$.

Steps:-

(1) Split into four disjoint quadrants any region R_i for which $P(R_i) = \text{False}$.

(2) Merge any adjacent regions R_j and R_k for which

$P(R_j \cup R_k) = \text{True}$

(3) Stop when no further merging or splitting is possible.

Global Processing via the Hough Transform:-

Given n points in an image, suppose that we want to find subsets of these points that lie on straight lines.

→ First find all lines determined by every pair of points and then find all subsets of points that are close to particular lines.

→ The problem with this procedure is that it involves finding $n(n-1)/2 \sim n^2$ lines and then performing $n\left(\frac{n(n-1)}{2}\right) \sim n^3$ comparisons of every point to all lines.

→ Hough (1962) proposed an alternative approach, commonly referred to as Hough transform.

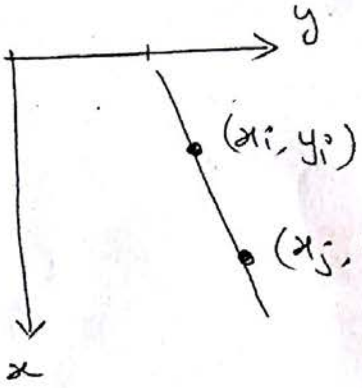
→ consider a point (x_i, y_i)
The general equation of straight line in slope-intercept form $y_i = ax_i + b$. —(1)

→ Infinite lines will pass through (x_i, y_i) satisfying $y_i = ax_i + b$ for varying values of a and b.

writing eqn (1) as $b = -x_i a + y_i$ (ab-plane also called parameter space)

→ consider a second point (x_j, y_j) also has a line in parameter space associated with it and this line intersects the line associated with (x_i, y_i) at (a', b') where a' is the slope and b' the intercept of the line containing both (x_i, y_i) and (x_j, y_j) in the xy-plane.

→ The computational attractiveness of the Hough transform arises from subdividing the parameter space into so-called accumulator cells.



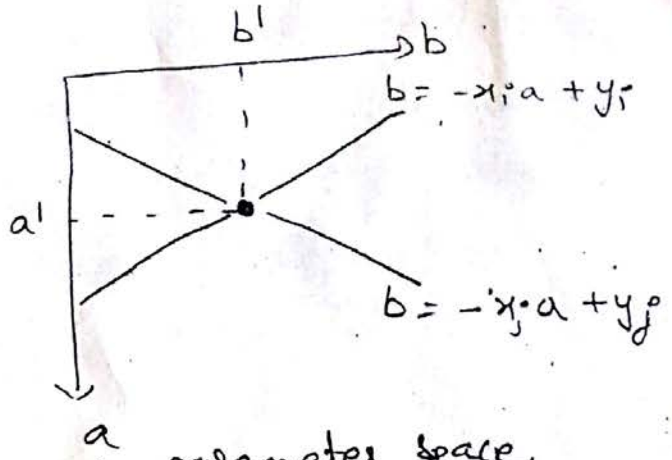


Fig: (a) xy-plane (b) parameter space.

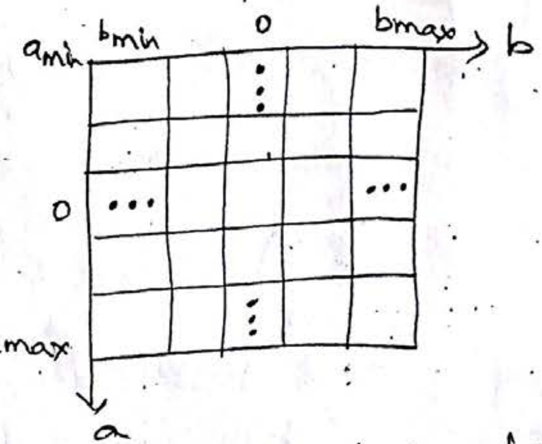
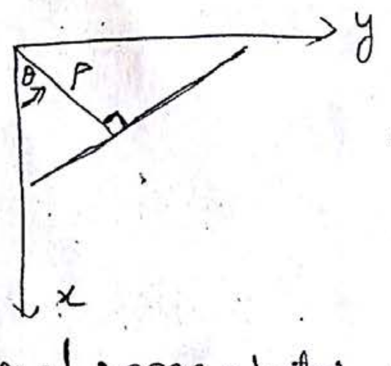


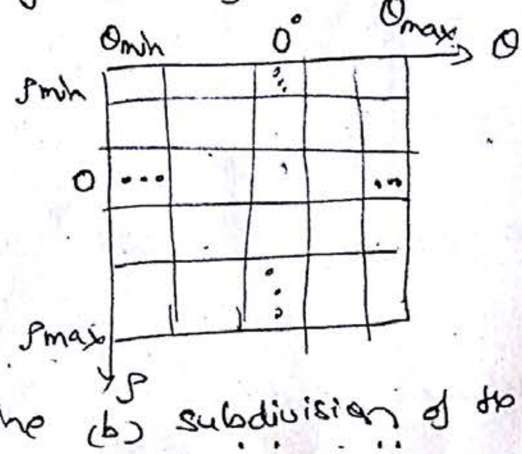
Fig:- Subdivision of parameter plane for use of Hough Transform.

→ A problem with using the equation $y = ax + b$ to represent a line is that the slope approaches infinity as the line approaches the vertical.

→ One way around this difficulty is to use the normal representation of a line:

$$x \cos\theta + y \sin\theta = \rho.$$





(a) Normal representation of a line (b) subdivision of the ρθ-plane