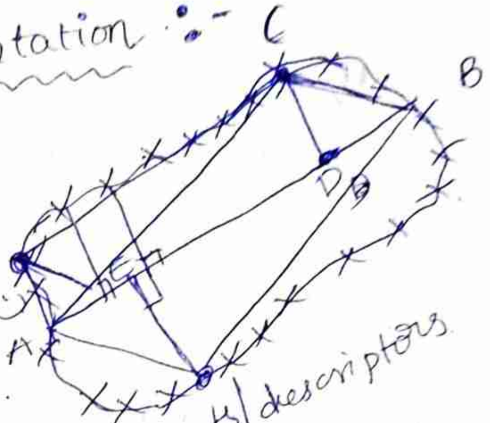
Deep Learning

Feature Descriptors

- Shape Descriptor (or) Boundary
- color Descriptors
  {color / intensity / texture}

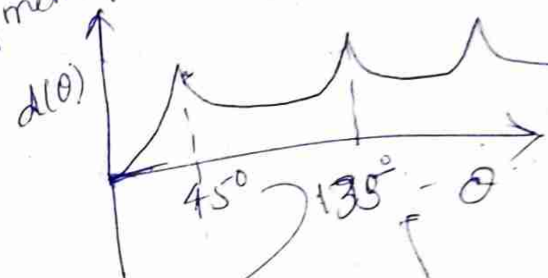
Polygon Representation :-



this is the polygonal Representation

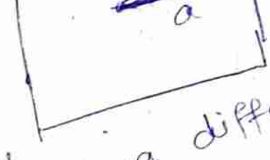
gives information about shape

we can obtain different segments / descriptors.



shape of plot also gives important info about shape of the boundary

Signature :-



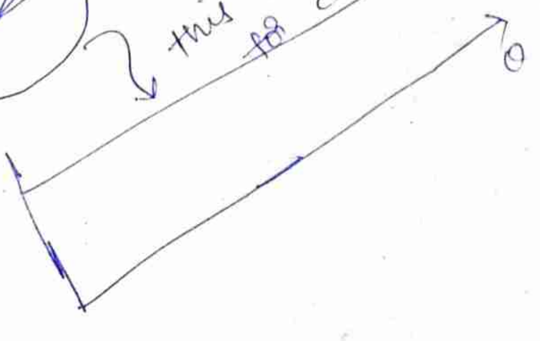
plot of the distance of different Boundary points from centre along different direction

this is in case of square Boundary

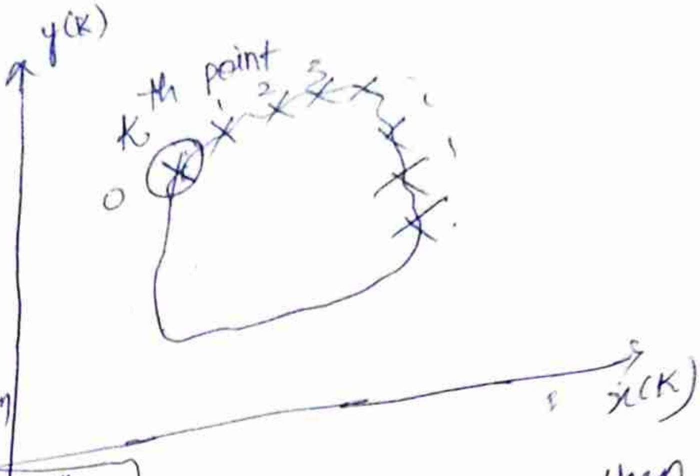
so it is an important thing that tells about Boundary Descriptors

this is for circle

d(θ)



Fourier Descriptor



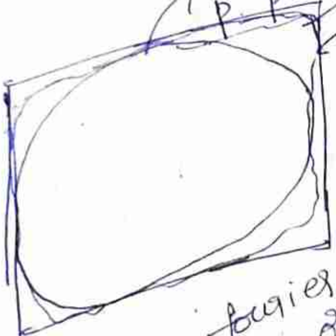
$s(K) = x(K) + j\,y(K)$

$K = 0, 1, \ldots N-1.$

after fourier transform

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

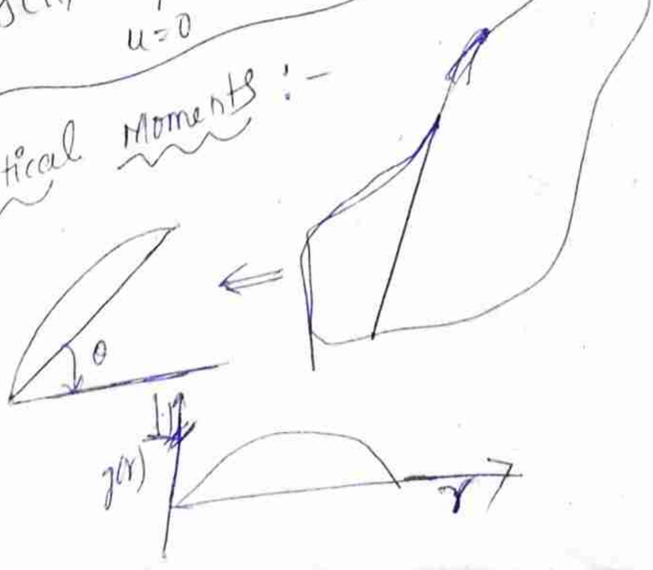
instead of considering all points N
lets say i consider only P no. of points then
$a(u)$ $u = 0, 1, 2, \ldots P-1$.
with P points
$P < N$. $P \uparrow$ then this shape
if no. of Boundary points
N
$\downarrow$
P. (first P no. of points)

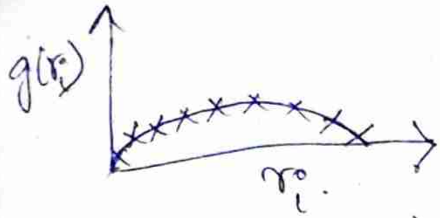


according to fourier transform
$\hookrightarrow$ low order coeff gives some info about trend about signal
$\hookrightarrow$ high order coeff gives detail info of signal.

$$\boxed{s(K) = \sum_{u=0}^{P-1} a(u)\, e^{j2\pi uK/N} \quad K = 0, 1, 2, \ldots N-1}$$

Statistical Moments :-





if we normalize $g(r_i)$ with area under the curve. then it is going to represent histogram.

i.e, at r_i distance what is the frequency of occurance of the points

$$\sum_{\forall i} (r_i - \mu)^K p(r_i) \rightarrow \text{statistical moments} = \sigma_K.$$

so it captures the ~~shape~~ one information

$$\mu = \sum_i r_i p(r_i)$$

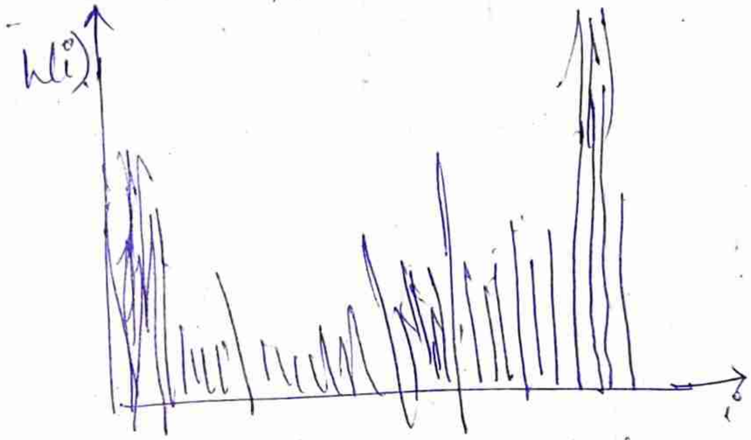
if $K = 2 \rightarrow$ variance $\rightarrow$ tells about spread of distribution.

$K = 3 \rightarrow$ skewness $\rightarrow$ symmetry about mean.

~~Feature~~ Region Descriptors — Intensity, Texture, colour.

Intensity Descriptor :- (Black & white images)

assume picture is black & white with shaded portion as black & unshaded portion as white



intensity $\rightarrow$ represented by a number from 0 to 255

$\rightarrow$ every pixel contains a value from 0 to 255 { 0 $\rightarrow$ Black, 255 $\rightarrow$ white.

$i \rightarrow$ intensity value from 0 to 225 then $h(i)$ tells about the occurance of that intensity value in the image

$$h(i) = N_i \Rightarrow h(i) = \frac{N_i}{N}$$

$N_i \rightarrow$ no. of pixels with intensity i.

$N \rightarrow$ total. no. of pixels in image

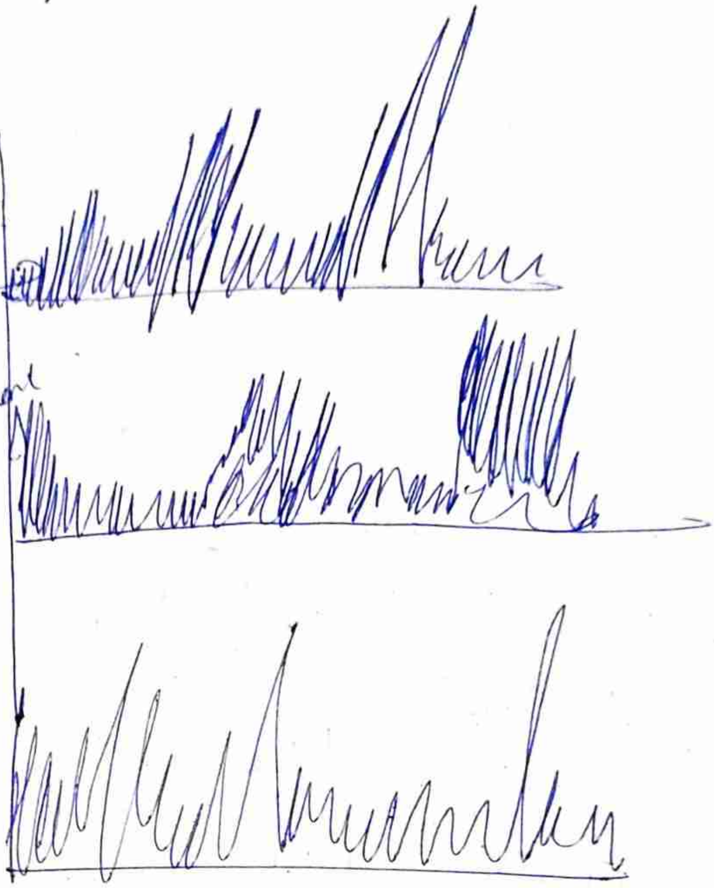
after normalization

what is the probability of occurance of pixel i

for colour images → there are 3 channels.
(1) Red (2) Blue (3) Green.
So for each colour we will get one histogram

Colour histogram

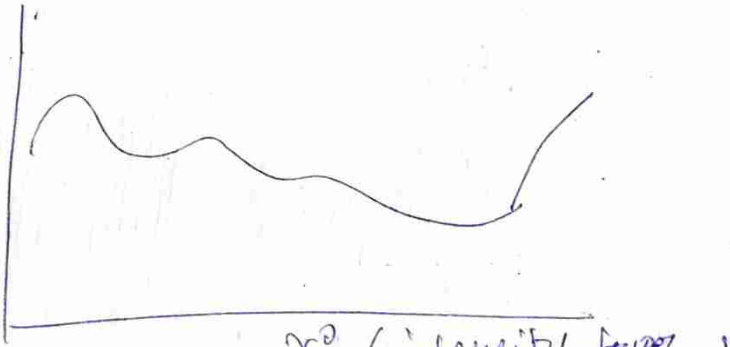
Red



tells about the what is the distribution of the strength of the red component ←

Green

Blue



$h(r_i)$ → histogram

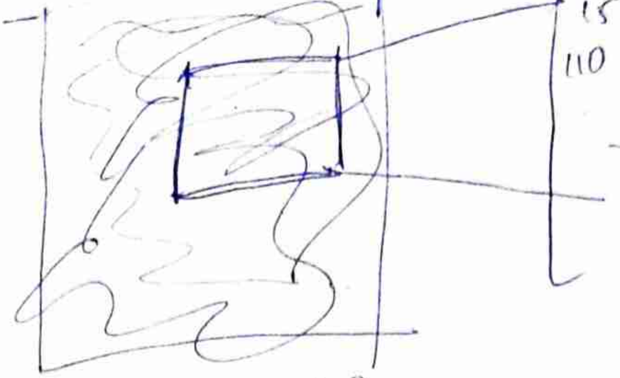
r_i (intensity ~~leve~~ value)

tells about shape of distribution:

$$\mu = \sum_i r_i h(r_i)$$

$$\sigma_K = \sum_{\forall i} (r_i - \mu)^K h(r_i)$$

Texture Descriptor :-



150 100 115 119 100 145 140
110 112 120 135 125 120 132

co-occurance matrix.

$(i,j) \rightarrow$ two intensity values.

$P \rightarrow \langle l, \theta \rangle$ needs to follow this

$a \rightarrow i$

$b \rightarrow j$

lets assume $l = 1$ $\theta = 45^\circ$

10	9	7	9	5	8	11	9
6	5	15	12	4	6	3	2
9	3	2	10	6	8	4	5
8	2	4	3	7	5	6	1
2	0	11	8	10	9	8	2
8	4	7	1	6	0	7	6
2	3	8	9	11	6	3	9
7	2	8	8	6	12	6	7

lets say $i = 4$, $j = 8$

so shape of $A \rightarrow 16 \times 16$

$A_{\langle 1, 45 \rangle}(i, j)$

0–15 0–15

4	x
x	8

$A_{\langle 1, 45^\circ \rangle}(4, 8) = 3$

We will get that what is the co occurance of this intensity values i, j with following $\langle 1, 45^\circ \rangle$ this shape args

Co occurance matrix based descriptors → normalized matrix c_{ij}

Maximum probability = $\max_{ij} (c_{ij})$

element difference moment = $\sum_i \sum_j (i-j)^k c_{ij}$ along the diagonal 0.

inverse element difference moment

$= \sum_i \sum_j \frac{c_{ij}}{(i-j)^k}$ → very large along diagonal

uniformity → $\sum_i \sum_j c_{ij}^2$ → it is maximum when ~~sum of~~ all values are going to be equal

entropy → $-\sum_i \sum_j c_{ij} \log_2 c_{ij}$

pixel to frequency domain transformation

wavelet transformation

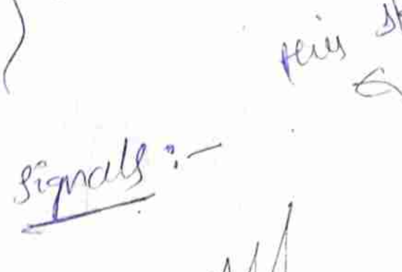
divide pixels into different sub brands & calculate the energy in each sub brand & represent it as frequency

Gabor transformation → A filter that is cosine modulated guassian envelope

variance, frequency, orientation angle

this step for capturing of low band signals compared with other

Audio signals :-



importance is there

this step is due to capturing of low frequency signals as they are important than high frequency

MFCC time representation

Take fourier transform → Spectrum → use triangular overlapping window (to map powers on Mel scale) → mel frequencies → logarithmic powers of frequencies → mel log → take discrete cosine transform → MFCC (amplitude of spectrum)

Traditional machine learning vs Deep learning

signal $\rightarrow$ feature (θ) $\rightarrow$ machine learning (θ) $\rightarrow$ output

signal $\rightarrow$ machine learning (θ) $\rightarrow$ output

tations