Pattern Recognition & Image Processing

CSE-4875 & CSE-4876

Pattern Recognition & Image Processing.

By – Shayha Ameen Chy.

Image :– An image can be defined as a two-dimensional function, $f(x,y)$ where, x and y are spatial (plane) co-ordinates and the amplitude of f at any pair of coordinates (x,y) is called the intensity or gray level of the image at that point.

একটা image এ তিনটা info থাকে, সে তিনটা মিলে এক-একটা color হবে।

$f(5,3) = 5$ → amplitude

↑

space/spatial value

Digital Image :– When x,y and the intensity values of f are all finite, discrete quantities then it's called digital image.

example 9/2

Image Processing :— The analysis and manipulation of a digitized image, espacially in order to improve its quality is called image processing.

Digital Image Processing :— The field of digital image processing refers to processing digital images by means of a digital computer. A digital image is composed of a finite number of elements, each of which has a particular location and value. These elements are called picture element, image elements, pels and pixels.

Pixels :— is the term used most widly to denote the elements of a digital image.

Present System of Image Processing:-

→ Virtual Reality
- আমি যে environment এ আছি তার change করতে পারবে। ex:- video game.

→ Cognative Reality
- Real life and digital life একসাথে থাকবে। ex:- microsoft Hololens.

→ Human Detecting
- একটা image থেকে human face detect করতে পারবে।

→ Human Tracking
- একটা video তে কোনটা কোন human specifically detect করতে পারবে।

→ Google translator
- product এর গায়ে অন্য language এ কিছু লিখা থাকলেও camera-র সাথে connect করে only camera Image processing এ লিখাগুলো সিলেক্ট করে নিলে তা আমাদের english language এ output দিবে।

Some types of Image – (Basic)

Binary Image:– A binary image is a digital image that has only two possible values for each pixel. Typically, the two colors used for a binary image are black and white.

object color of binary is foreground color
rest of the color " background color

Grayscale color :– the grayscale color values are available from 0 to 255 which starts with black and ends at white.

RGB Image:– it's a three dimensional Image. which can be written as,

$f(x, y, z)$ → co-ordinates

where, x → reflects the value of R
y → " " " G
z → " " " B

Types of Images:—

→ Based on radiation from EM (Electromegnet spectrum.

Smallest → Gamma Ray
largest → Radio Ray

→ Acoustic / Ultrasonic
→ Electronic
→ Synthetic.

Acoustic:— sound generate এর মাধ্যমে image আনা হয়।

ex: Ultrasonogram

EM (Gamma Ray):—

→ major use in nuclear medicine.
→ Images are produced from the emissions collected by gamma ray detectors.

Uses of UV imaging:—

- Lithography
- electromegnetic isotop used for tumor.
- Fluoscence microscopy.

Uses of Radio Wave:—

- MRI (tumor/ cancer)

See Images from slide.

Electronic — Photo-র base হিসেবে light এর পরিবর্তে electron use করা হয়।

ex:— example scanning electron Microscopy image of normal circulating human blood

SS — See Slide.

Synthetic — Computer generated Image.
SS for example. ex: Avatar Interface.

PET image — photon generate করে light emit করে।

Visual — eyes have limited bandwidth যা পরিধি তা খুবই optimal.

LANSAT-8 (SS)

bands, 1-2-3 → google map.
bands 4-3-2 → নদী খুঁজতে
bands 7-4-2 → গাছ খুঁজতে

Ultrasonic — Sound wave যা
— skin easily pass করতে পারে।
— sculp " " "
— Sound reflect করে যার উপর depend করে image generate হয়।

EM:— electron microscope
→ 10^50 times zoomable from normal
→ in stead of photon electron bin নিক্ষেপ করে।

Computer Generated Image:— / Synthetic :—
→ Realistic view তৈরার জন্য used.
→ real life এ থাকে না, pc generate করে।
→ প্রতিটি image তথ্যের info convey করে।

fundamental steps of Digital Image Processing :—

Image Acquisition :—

— Digital Image retrieve করার বা নেয়ার
— Analog camera দিয়ে তোলা হবি digital
এ convert করতে হবে পারে।
— scaling করা।

Image Enhancement :—

— Image contain একটা space কে enhance করা, blur থাকলে bright করবে।
— it is based on human subjective preference. means human judge করে বলে কোনটা ভালো কোনটা খারাপ।

Image restoration :—

— deals with improving the appearance of an Image.
— objective, thus restoration techniques based on mathematical or probabilistic models of image degradation. (depends করবি কি চাই)
— human tested না।

Color Image Processing :-

- Image কে compressed করা হয়।
- Normally use করি coz time বেচে যায়
- Internet এ upload করার জন্য use হয়

যেমন, ex: fb pp upload.

Webet tansfer :-

- foundation for representing images in various degrees of resolution.
- একটা image কে ছোট করে দেখা।

ex: (1000 X 1000) img কে (25x25) করে দেখা।

- একে pinamid scale বলে। coz original image কে different scale এ দেখি।

Image Compression & Watermarking :-

Compression — *) jpg method

*) size reduce করে।

*) transmit করতে সাহায্য করে।

*) কতটুকু information আমি চাচ্ছি তার উপর depened করছে।

Watermarking:-

→ extra code generate করা হবে copy করলেও same থাকে।

→ change করা possible না

→ owner claim করতে পারবে

→ color change করলেও original data থাকবে।

→ authentic করা।

Morphological:-

→ mathematical process

→ image shape change করার জন্য use হয়।

Segmentation:-

→ আমি যা নিয়ে কাজ করবো তা pic থেকে আলাদা করে নেওয়া।

→ digital license plate detection is one of the process of segmentation.

Representation & Description:—

— segmentation এর output নিয়ে কাজ করে almost

— output টা কে define করার, like: △ → এই pic নিয়ে এটাকে নাম দেওয়া।

Recognition:-

— starts with labeling.

— detect করে সে কোন object টা কী;

— △ sign কী define করে তাও বলা।

Feature Extraction

(Representation, Recognition) এই step গুলোর analysis এর output image হতেও পারে না ও পারে।

see fig: 1.23

Component of Image Processing:-

→problem domain থাকবে

→real life থাকবে।

hard copy real life থেকে নিয়ে (image sensor এর মাধ্যমে) camera representation করা। – for this reason hardware needed.

– Image processing software (Abode photoshop / Matlab) needed.

→ display materials

→ storage materials. – অনেক বড় হওয়া needed.

→ network needed / route storage যদি image পাঠাতে হয় connection needed

(SS – See Slide)

S. fig:– 1.24.

Chapter - 02

→ How human eye works.
→ How eye works like camera
→ difference between camera & eye.

Some basics

→ চোখের আইরিস আমাদের চোখে আলোর পরিমাণ নিয়ন্ত্রণ করে।

according to this camera & eye same

Difference

- camera mechanical & eye physical
- Retina light sensitivity বুঝতে পারে।
- কেউ monocular vision হয়
কেউ binocular vision এর হয় সাথে depth vision বের করা যায়।
- monocular - আমাদের চোখ - 180° দেখি and only 1 টা image দেখি।
- binocular - পাখির চোখ - 360° দেখে। 2 টা image দেখি
- আমাদের চোখ একটা নির্দিষ্ট intensity level পর্যন্ত দেখতে পায়।

Light & Electromagnetic Spectrum :—

- বিশ্বের মধ্যে যিনি different colors দেখতে পায়।
visual band এর smallest হলো uv
" " " largest " red.

Thermal Image :—

- object এর heat কে sense করে,
- visible light থাকার প্রয়োজন নেই,
- রাতের অন্ধকারে camera-র মধ্যে object থেকে পরিমাণ energy observe হয় তাকে measure করাই হলো thermal image.

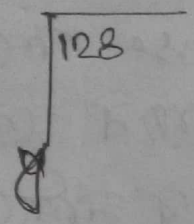
What is color ?:—

visual perceptual property.
অর্থাৎ, দেখার অনুভূতি,

Photo diode :— আলোর উপর depend করে different voltage generate করতে পারে,
photo diode left to right print করে & color generate করে। খুব ছোট portion image generate করে,

Line sensor — 3D object এর pic তোলার জন্য use হয়।

Digital Image:— a function of (x, y) where, x, y co-ordinate value, যেকে spatial co-ordinate বলে, এবং specific value যাকে তার মানকে amplitude বলা।



$f(x,y) = 128$
↳ amplitude
↳ spatial co-ordinate.

Sampling & Quantization:—

If $f(x,y)$ is a function,

(x, y) এর value কে digital value তে convert করা/ finite করাকে বলে sampling

& $f(x,y)$ এর value কে finite বা digitized করা হলো quantization.

like, $f = 2.773054662$
≈ 2 ← this representating is called quantization.

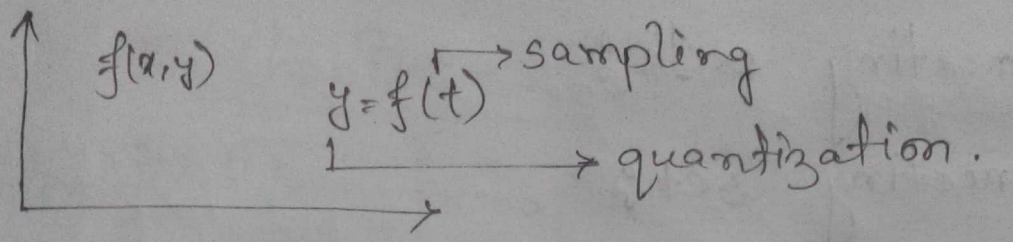


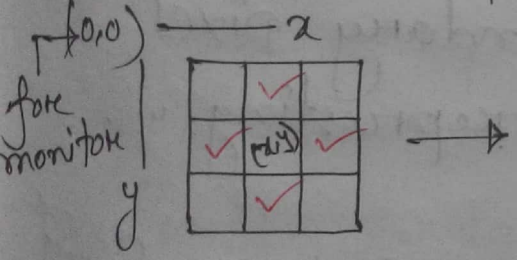
fig: 2.17

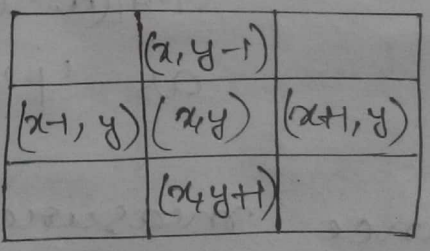
digital Image never be ≠ to) analog image.

→ pixel मूल्य = pic valo.

Neighbours of pixel :—

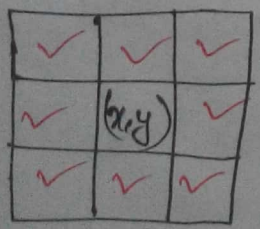
4-connected neighbours, if (x,y) is the selected pixel then from fig we get,





→ this are the 4 neighbour pixels of (x,y)

8 connected neighbour,

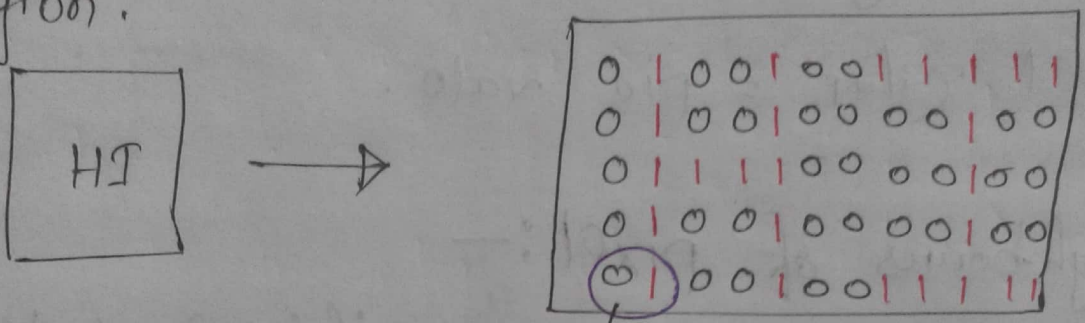


(x-1, y-1)	(x, y-1)	(x+1, y-1)
(x-1, y)	(x, y)	(x+1, y)
(x-1, y+1)	(x, y+1)	(x+1, y+1)

BJ-225 Connecting path:

Connecting set:

Connecting Region: কোন image এর ছোট কোন portion যদি connected থাকে তাহলে তাকে region.



এখানে, HI কে 1 দ্বারা replent করা হয়েছে

এখানে, '0' boundary pixel as '1' is resting.

Distance measure – দুটো pixel এর মধ্যবর্তী দূরত্ব।

euclidean distance

$$D_e(p,q) = [(x_2-x_1)^2 + (y_2-y_1)^2]$$

Manhatan distance / city block distance.

(Book – 2.5.3)

Chapter - 03

"Intensity Transformation & Spatial filtering"

Image Enhancement:- Specific image এর information কে enhance করা।

ex: dark image কে bright করা।

gray scale image এ আমরা যে value টা পাই তা হলো intensity.

Intensity transform:- এক value কে অন্য value তে repment করা।

like- 50 কে 180 দ্বারা represent করা।

most of case intensity transform করা হয় subjective purpose এ।

Transformation function:- একটা image কে read করে অন্য একটা image output দেওয়া।

Output = f(input)

$y = f(x)$ → input img

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

output img ↳ transformation function.

→ image কে represent করা হয় co-ordinate spatial co-ordinate and function দ্বারা।

→ TF (transformation function) most case এ fixed fixel এর উপর কাজ করে, adjucent এর উপর কাজ করে না।

Normalization :—

5	10	7	3
11	105	4	1
12	13	14	7
10	4	5	7

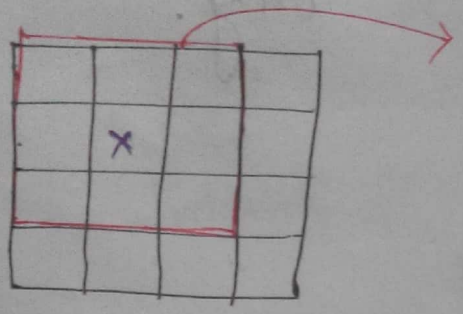
→ এই value টি হলো noise কেননা, এটি অত্যন্তাতীপূর্ণ, thus, normalization needed.

সবচেয়ে easy way হলো neighbour value average নিয়ে replace করা।

for this, $(5+10+7+4+14+13+12+11)/8 = 57$.

normaliz করতে 3×3, 4×4, 7×7 যে কোন filtering করতে পারবো।

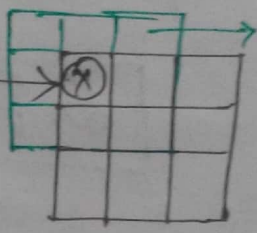
Spatial filtering :- specific অঞ্চল portion নিয়ে কাজ করা।



this is called 3×3 window/filter/kernel/mask.

filter/mask → most used.

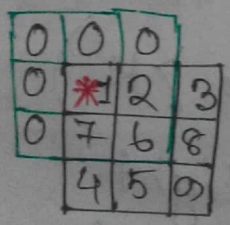
কিন্তু point টা যদি এই position এ থাকে তখন padding করতে হয়। দুভাবে করা যায়।



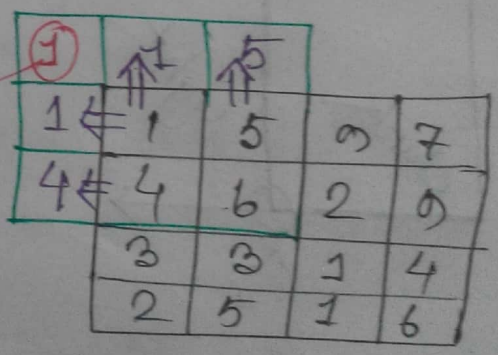
padding.

Zero padding :- padding এর '0' value দিয়ে

Mirror ,, :- ঐ ,, ,, mirror ,, ,,



Zero padding.

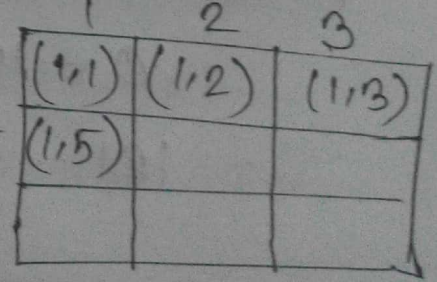


Mirror padding.

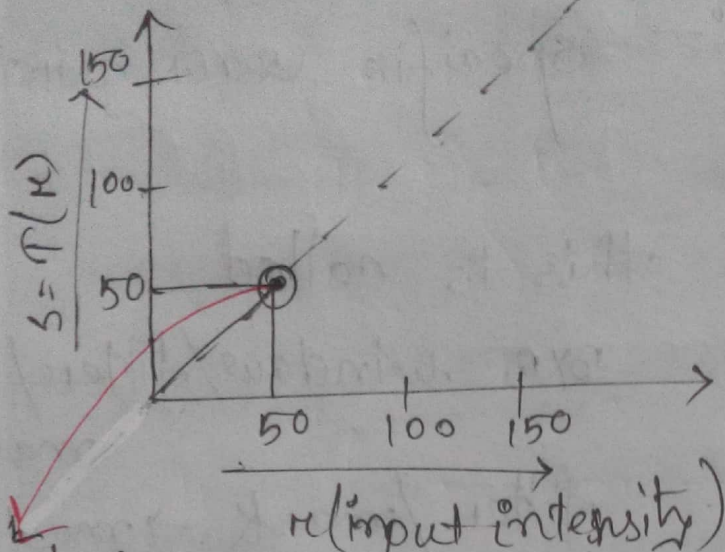
* $g(x,y) = T[f(x,y)]$

$\Downarrow \quad = \quad \Downarrow$

$s \quad = \quad T[r]$



two-co-ordinate দ্বারা একটা co-ordinate কে represent করা।

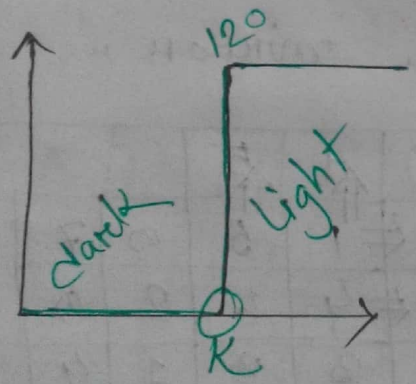


→ mobile filtering এরকম graph only.

(Input = Output
so, no change)

S. F. = see figure = 3.2

Threshold value:— যে point থেকে barrier create করা।



K = threshold value,

let, K = 120

if (value > K)
⇒ light

if (value < K)
⇒ dark

in another words,

K = 120 means,

0 → 120 ⇒ represents black (0)

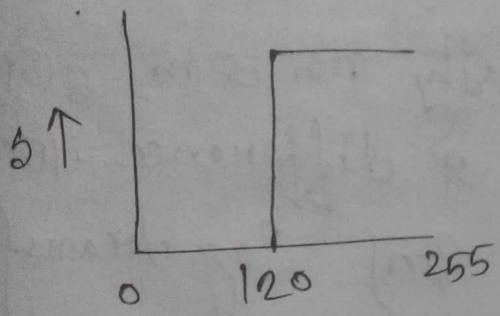
121 – 255 ⇒ " – white (1)

* Intensity Transformation is a subjective process.

$s = T(r)$

s ↓ Output Intensity

r → input intensity



0-120 ⇒ value of $s = 0$, black

121-255 ⇒ " " $s = 1$, white.

Binary Image (output)

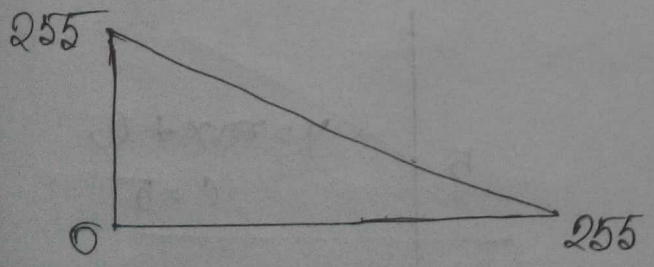
Image Negative:—

input image ⇒ gray scale

output " ⇒ " "

Input Image এর value 0 হলে output image এ 255

" " " " 1 " " " " 254.



$s = 255 - r$

⇓

highest level of color.

Gray scale এ total color 256,

এদের L দ্বারা represent করা হয়।

$L = 2^l$ ⟶ কত bit color এ represent করা

So, $s = (L-1) - r$

$= (256-1) - r$

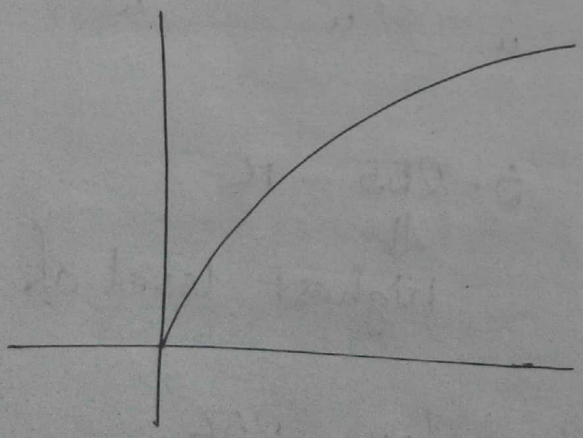
$= 255 - r$

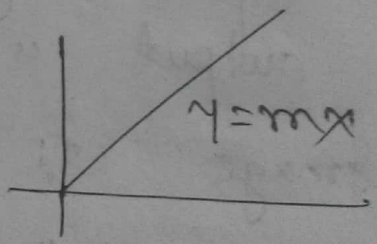
Log transformation:—

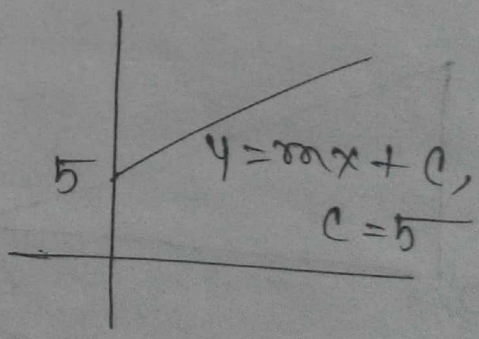
দুটো color এর intensity কম হলে, দুটোর difference কম হয়। তখন ঐ difference টিকে বড় করে represent করার জন্য log transformation use করা হয়।

ex: 76, 80 } ⟹ 70, 90
diff: 4 } ⟹ diff = 20.

$s = c + \log(1+r)$

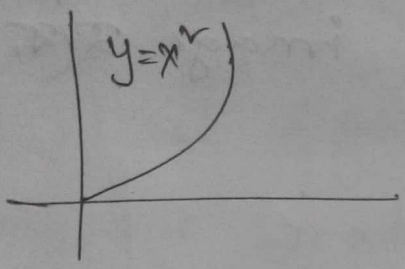


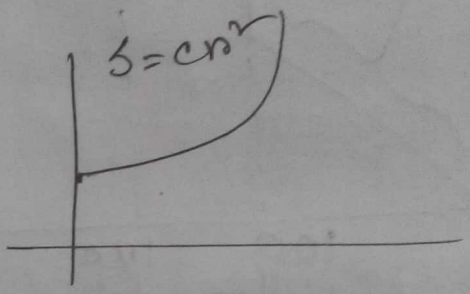




Power-law (Gamma) Transformation:—

CRT monitor এর plot গুলো এর input image থেকে output image টা dark হয়ে যায়। এটা reduce করতে input image কে γ power include করা হয়। তাই একে gamma transformation বলে।





→ in general image এর উপর filter করা হয়।

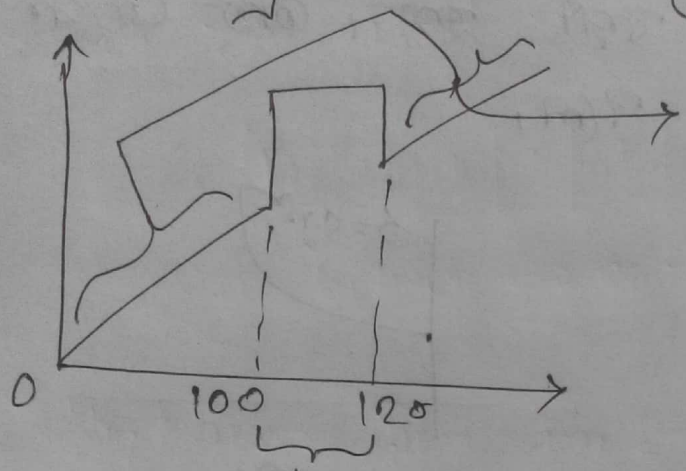
→ γ value manupulate করে full image কে dark or light করা হয়।

→ Contrast → এলাকা image এর highest and lowest point এর পার্থক্য

Piece wise transformation:-

Contrast Streetching:- এটার threshold value থাকে,

Intensity level slicing:-



এই value গুলোতে থাকলে input = output image হবে,

এ value-র মাঝে থাকে input থেকে output বড় হবে

Bit plane Slicing:- bit wise করা হবে,

According to youtube:-

if an image matrix is—

6	2	3	2
1	5	0	7
4	3	2	1
2	5	7	6

এখানে সবচেয়ে বড় num = 7 থাকে bit এ represent করলে হবে,

7 = 111

bit wise represent করে matrix টি পাই,

110	010	011	010
001	101	000	111
100	011	010	001
010	101	111	110

এখানে, b = 010
এই 3bit এর

0 1 0 → least significant bit (LSB)
↓ most significant bit (msb)
↓ middle bit

বলে।

এখন, সবগুলো value-র most significant, middle and least significant bit গুলো নিয়ে আলাদা আলাদাভাবে matrix represent করবো সাবেরবলে bit plane scaling.

msb নিয়ে scaling →
(সব bit-এর 1st value নিয়ে)

1	0	0	0
0	1	0	1
1	0	0	0
0	1	1	1

middle bit নিয়ে →

1	1	1	1
0	0	0	1
0	1	1	0
1	0	1	1

LSB নিয়ে →

0	0	1	0
1	1	0	1
0	1	0	1
0	1	1	0

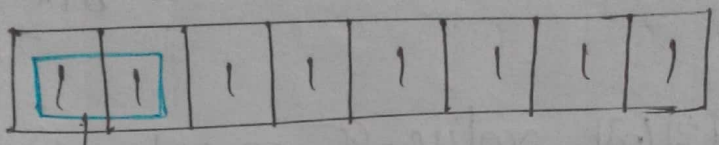
0 → 1 bit

255 → 8 bit

8 bit নিম্ন কলর বললে white value দেরকে বুঝায়

1,2 " " " " black " " "

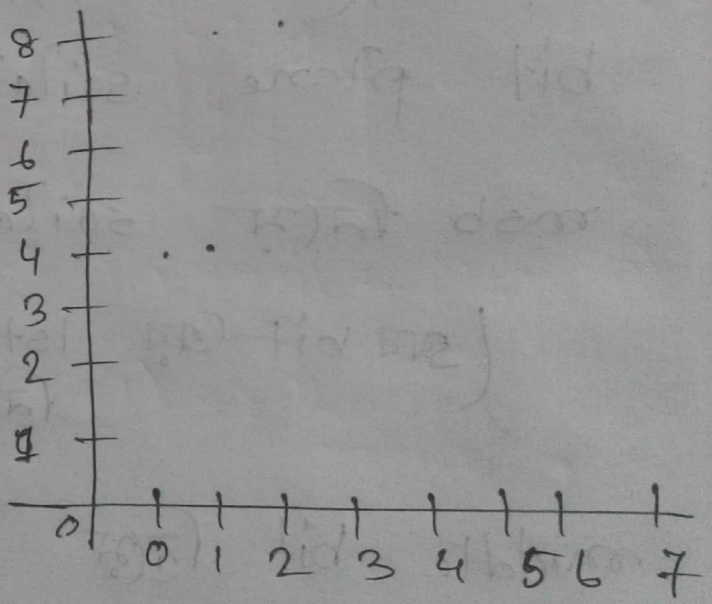
8 bit কে 4 bit এ convert করতে হলে



দুটো bit একসাথে করা

Histogram Processing :—

7	4	4	3	2
1	7	7	5	2
2	2	1	3	3
4	4	3	3	2
1	1	2	2	2



Normalized histogram/Probability density function:—

$$PDF = \frac{\text{Number of Pixel}}{\text{Total intensity}}$$

ex: for intensity $2 = \frac{7}{30} = 0.23 \approx 23\%$

Advantage:— যে কোন intensity (like-02) হওয়ার সম্ভাবনা বের করা যায়।

Fundamental of Spatial filtering:—

↓
space related

image এর উপর portion আর্থে.
— x, y co-ordinate

$f(x,y) = 50$ → amplitude
↓
Spatial co-ordinate

10	10	15	16
17	70	18	12
9	8	11	13
14	15	16	17

noise remove,

new value $= \frac{10+10+15+17+70+18+9+8+11}{9}$

$= 18.67$.

≈ 19.

It's one way.

Another way is—

Low pass filter / smooth filter / average filter:-

এখানে দুইরকম kernel করা যায়।

	1	
1	−1	1
	1	

1	2	1
2	0	2
1	2	1

for 1st one if an image is

10	10	15
15	16	15
5	10	10

	1	
1	−1	1
	1	

Here, $\frac{(10*1)+(15*1)+(10*1)+(15*1)+(16*(-1))}{(1+1+1+1)}$ ⇒ there are red mark 1

$= \frac{10+15+15+16+10}{4}$

$= 34/4 = 8.5$.

10	10	15
15	16	15
5	10	10

1	2	1
2	0	2
1	2	1

Here, $$\frac{(10*1)+(10*2)+(15*1)+(15*2)+(10*1)+(10*2)+(5*1)+(15*2)+(16*0)}{1+2+1+2+1+2+1+2}$$

$$= \frac{10+20+15+30+10+20+5+30}{12}$$

$$= \frac{140}{12} = 11.67.$$

average করলে always value কম হয় তাই কারনে একে বলে, low pass filter/smooth filter/ average filter.

fundamental of spatial filtering :—

is a subjective process.

low pass filtering :— কিছু value কে specific কিছু value দিয়ে করবো rather than taking all values.

- easy way to remove noise.

Ex:—

10	10	10
10	100	10
10	10	20

total pixel.

10	10	10
10	20	10
10	10	10

$$\text{value} = \frac{10+10+10+10+10+10+10+10+10}{9}$$

$$= \frac{180}{9} = 20$$

Using kernel on latest matrix,

10	10	10
10	20	10
10	10	10

1	2	1
2	0	2
1	2	1

⟹

10	10	10
10	10	10
10	10	10

$$\text{value} = \frac{(10*1)+(10*20)+(10*1)+(10*20)+(10*1)+(10*2)+(10*1)+(20*2)+(20*0)}{(1+2+1+2+4+2+1+2+1)}$$

$= \frac{120}{12} = 10$

subjective process — noise remove করার পর output image টা কি better or not তা human define করে।

→ averaging image filter use করলে তা blur হয়ে যায়।

→ Kernel usually odd size এর নেয়া হয়। (ex: 3×3 , 5×5)

Integration represents working with continuous value.

Summation " " " discrete value. নির্দিষ্ট limit এর value.

Spatial Correlation & Convolation:—

Specific কোন value change করা।

→ Smoothing Spatial filters:—

— Sharp transaction reduce করে

— ধারে ধারে blur করবে।

→ median filter use করার হয়

median/max/min filter → Order static filter.

Median filter:—

— Sort the whole pixels

— take the median value.

effective for noise remove করার জন্য

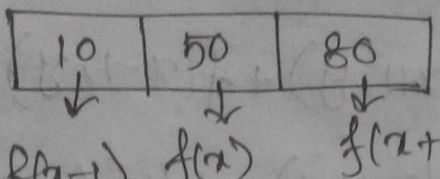
cause image এর original color রেখে দেয়।

Sharping Spatial filters:— edge গুলো highlight করে।

— done by differentiation.

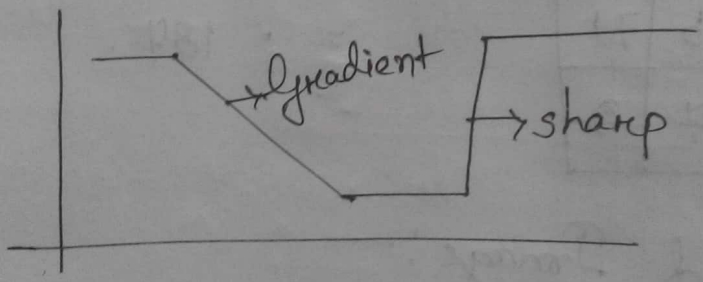
- first order derivative : $f(x+1) - f(x)$
- Second " " : $f(x+1) - f(x) + f(x-1) - f(x)$

$$\Rightarrow f(x+1) + f(x-1) - 2f(x)$$



→ 1st order derivative means distance এর specifically কোথায় বেশি।

gradient :- কার কার change হয়।



→ intensity change করলে 1st order derivative এর value change হচ্ছে।

→ 1st order derivative gradient এ value change করে। it is sensitive to gradient. sharp change এর জন্য value change করে।

→ Second order derivative gradient এ value তেমন একটা change করে না, But it also changes for sharp value.

Using second order deviation for Image Processing —

The laplacian :—

	$(x, y+1)$	
$(x-1, y)$	(x, y)	$(x+1, y)$
	$(x, y+1)$	

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

→ for x direction 2nd deri vatia

→ " y " 2nd "

if,

100	8	31
41	5	71
42	87	9

→ $41 + 8 + 71 + 87 - (4 \times 5)$
$= 187.$

Sharpened Image :—

260	231	255
197	211	255
180	235	233

→ 4 direction ধরে 94
→ 8 neighbour " " 165

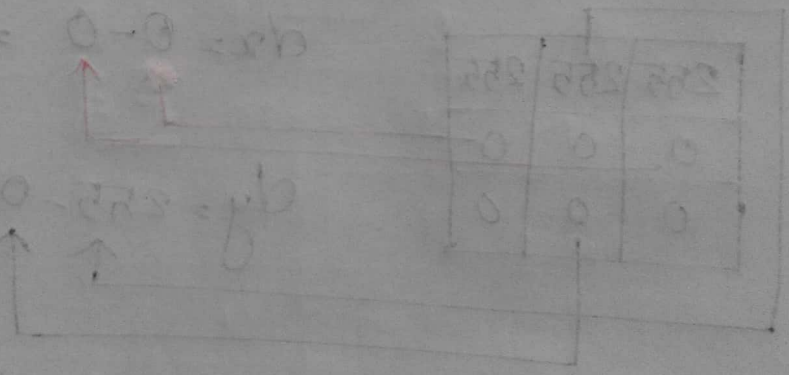
10	200	251
11	200	250
12	200	250

⇒ for 8 neighbour value is = −416 as its neg value we have to take (416) without minus(−) and also have to add the previous value.

So, here finally value will be

(416+200) = 616.

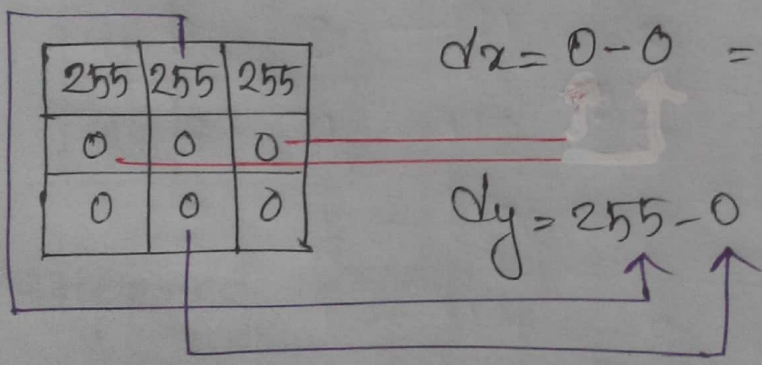
it's also called high pass filter because intensity বাড়ে সবার।



Unsharp masking:-

Using first Order Derivatives - The Gradient:

- Gradient represent করে আমার pixel এর কত change হয়েছে, কোন দিকে কতটা change হয়েছে।

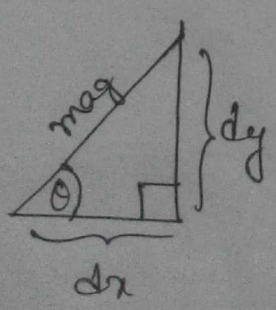


$dx = 0 - 0 = f(x+1, y) - f(x-1, y)$

$dy = 255 - 0 = f(x, y-1) - f(x, y+$

$\theta = \tan^{-1}\left(\frac{255}{0}\right)\frac{180^\circ}{\pi} = 90^\circ$ → degree (ডিগ্রিতে রূপান্তর করার জন্য)

→ 0.0001 নিয়ে কাজ করা হয় main



$mag = \sqrt{dx^2 + dy^2}$

$\theta = \tan^{-1}\left(\frac{dy}{dx}\right)$

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

⇓

vertical value দিচ্ছে

কিসের vertical or horizontal?

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

⇓

horizontal value দিচ্ছে

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

⇓

sovel operator

question এর ২টা ans.

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

⇓

sovel operator

From chapter 4 & 5 some definition (self study)

Imp Math given,

7	0	4	1
2	6	3	7
2	1	0	5
1	2	3	2

from this, we get,

g	f	c(g)	N(g'')
0	2	2	0
1	3	5	2
2	4	9	4
3	2	11	5
4	1	12	5
5	1	13	6
6	1	14	6
7	2	16	7

n=16

g = from given we can see the values are from (0 - 7)

f = কোনটা কতবার এসে matrix G

c(g) = cumulative frequency.

N(g) = max(0, round $\frac{2^l * c}{n}$ - 1)

Ques. total color = 8 (as 0-7)

So, $2^l = 2^3 = 8$

means, we need 3 bit to represent.

So, $N(g) = \max\left(0, \text{round}\ \frac{2^l * c(g)}{n} - 1\right)$

$N(0) = \max\left(0, \text{round}\ \left(\frac{8*0}{16}\right) - 1\right)$

$= \max(0, (\text{round}\ 0) - 1)$

$= 0$

$N(1) = \max\left(0, \text{round}\ \left(\frac{8*5}{16}\right) - 1\right)$

$= \max\{0, \text{round}(2.5) - 1\}$ → take the ceiling

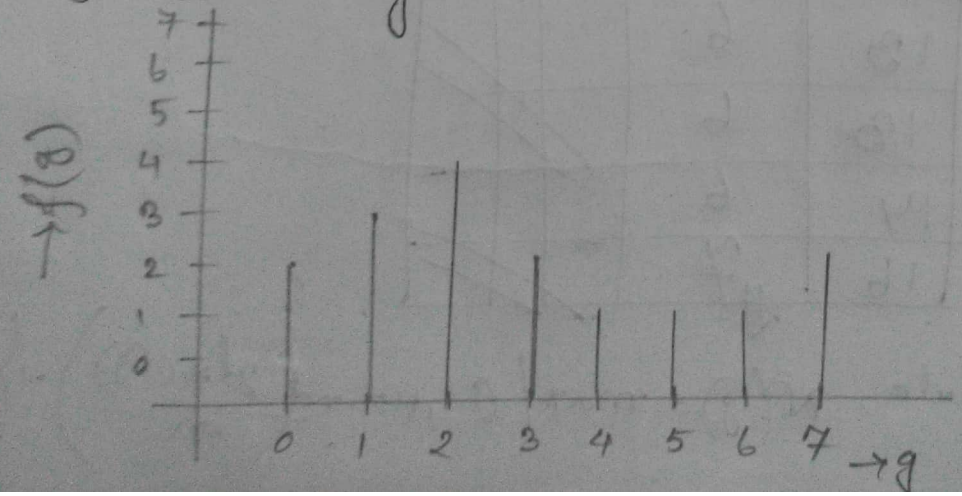
$= \max(0, 3-1)$

$= \max(0, 2) = 2.$

$N(2) = \max\left(0, (\text{round}\ \frac{8*9}{16}) - 1\right)$

$= \max(0, 4) = 4.$

Draw histogram:—



CT-ques

Histogram equalization:—

equalize করতে বললে আমার চাইতে one step কমে যাবে। N(g) এর frequency বের করার লাগবে।

Given,

1	0	1	7
2	7	1	3
2	1	0	2
2	2	1	6

Calculate & draw histogram from the following image & equalize that. ($l=3$).

Solution:—

g	f	C(g)	N(g)	f''
0	2	2	0	2
1	5	7	3	
2	5	12	5	0
3	1	13	6	
4	0	13	6	
5	0	13	6	
6	1	14	6	
7	2	16	7	

Formula, $N(g) = \max\left(0, \text{Round}\left(\frac{2^l * C(g)}{n}\right) - 1\right)$

Here, N(g) values are from (0-7)

N(g)	f'
0	2
1	0
2	0
3	5
4	0
5	5
6	2
7	2

f' বের করতে পারো,
1st table এ N(g) এর value-র frequency থেকে,
Like, '0' 1st table এ
N(g) = 0 তে f = 2 তাই f' = 2
আবার,
N(g) = 1 & N(g) = 2 এর
কোন value নাই তাই 0.
N(g) = 3 তে f = 5 তাই f' = 5
N(g) = 6 আর পাওয়া আছে
So, f = (1+0+0+1) = 2 = f'

এবার, 2nd table এর value থেকে histogram
আঁকো —