Date
Page

The digitized brightness value is called gray level. Row and column in 2-D array identify a point in the image, and the corresponding matrix element value identifies the gray level at that point.

* Application of Digital Image Processing :

Almost inevery field, digital image processing puts a live effect on things and is growing with time to time and with new technologies.

1. Image sharpening and restoration :
It refers to the process in which we can modify the look and ~~feel~~ fill of an image. It basically manipulates the images and achieves the desired output. It includes conversion, sharpening, blurring, detecting edges, retrieval, and recognition of images.

2. Medical field :
There are several applications under medical field which depends on the functioning of digital image processing.

- Gamma-ray imaging
- PET Scan
- X-ray imaging
- Medical CT Scan
- UV imaging

3. Robot Vision

: There are several robotic machines which work on the digital image processing. Through image processing techniques, robot finds their ways, for example, hurdle detection robot and line follower robot.

4. Pattern Recognition

: It involves the study of image processing, it is also combined with artificial intelligence such that computer-aided diagnosis, hand-writing recognition and images recognition can be easily implemented. Nowadays, image processing is used for pattern recognition.

5. Video Processing

: It is also one of the applications of digital image processing. A collection of frames or picture are arranged in such a way that it makes the fast movement of pictures. It involves frame rate conversion, motion detection, reduction of noise and color space conversion etc.

the image $f(x,y)$ to obtain the processed image $g(x,y)$. The point pixel does not consider the pixel values at neighbouring location. The value of each pixel is recalculated independently.

b) contrast stretching
:
It is simply making darker portions more darker and brighter portion more brighter. This process is appropriate for ~~low~~ low contrast image or dark image.

Low contrast image can result from ~~po~~ poor illumination, lack of dynamic range in image sensor or even ~~wro~~ wrong setting of lens ~~apere~~ aperture, During Image acquisition. The Idea behind contrast stretching is to increase the dynamic range of gray level in image being processed. The contrast of gray scale image can be adjusted by multiplying all pixel values by a constant gain, 'a'.

$$g(x,y) = a\, f(x,y)$$

c) Thresholding
:
It is the simplest method of segmentation where the resultant image becomes binary. The Simplest thresholding method replace each pixel in an image with a black pixel if the image intensity (I_{ij}) is less than some fixed constant T. Thresholding produces a binary image that can be used to locate

object in background.

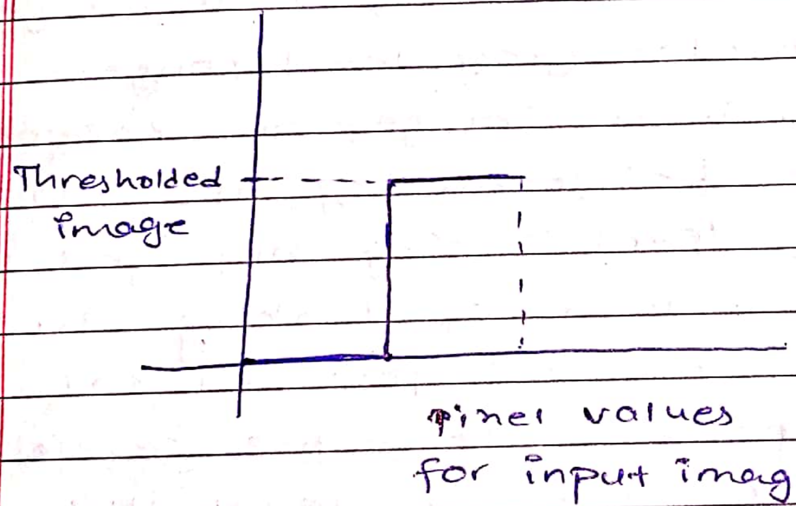
A threshold image $g(x,y)$ is defined as:

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

where, 1 is object

0 is background

T is constant



d) Digital Negative

A positive image is a normal image. A negative image is a total inversion in which light area appear dark and vice versa.

A negative color image is additionally color reversed with red areas appearing cyan, green area appearing magenta and blue appearing yellow.

A negative image is typically obtained by subtracting each pixel from maximum pixel value. Thus, an 8 bit image representation has transformation as; $255 - f(x,y)$.

b) Transmission

: Interference can be added to an image during transmission.

* Noise Model

: We can consider the noisy image to be modelled as follows:-

$$g(x,y) = f(x,y) + N(x,y)$$

where, $f(x,y)$ is the original image pixel
$N(x,y)$ is noise term
$g(x,y)$ is resulting noisy pixel

If we can estimate the model, the noise in an image is based on, we can figure out how to restore the image.

The different types of noise models are:

a) Gaussian Model

Eg: Electronic circuit noise, sensor due to poor illumination

b) Rayleigh

: Image ranging

c) Earlang / Gamma

: Laser imaging

d) Exponential Model

: Laser imaging

e) Impulse Model

: Faulty switching

f) Uniform Model

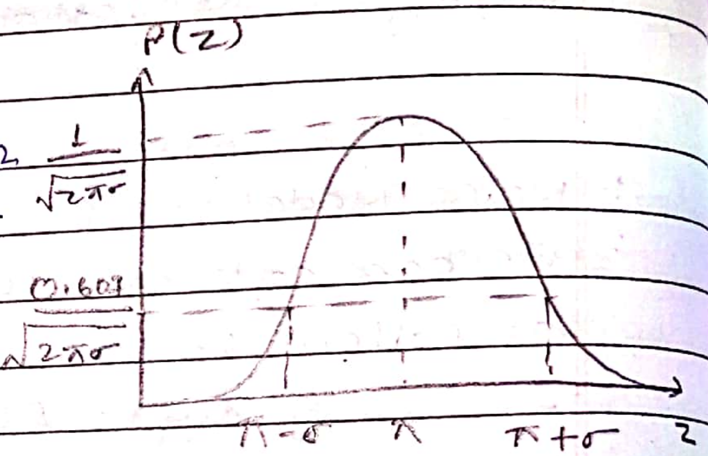
: Least descriptive

(Graph & equation)

* Resultant Nature Graph of different noise ~~models~~ models :

a) Gaussian Model :

$$p(z) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-(z-\mu)^2}{2\sigma^2}}$$



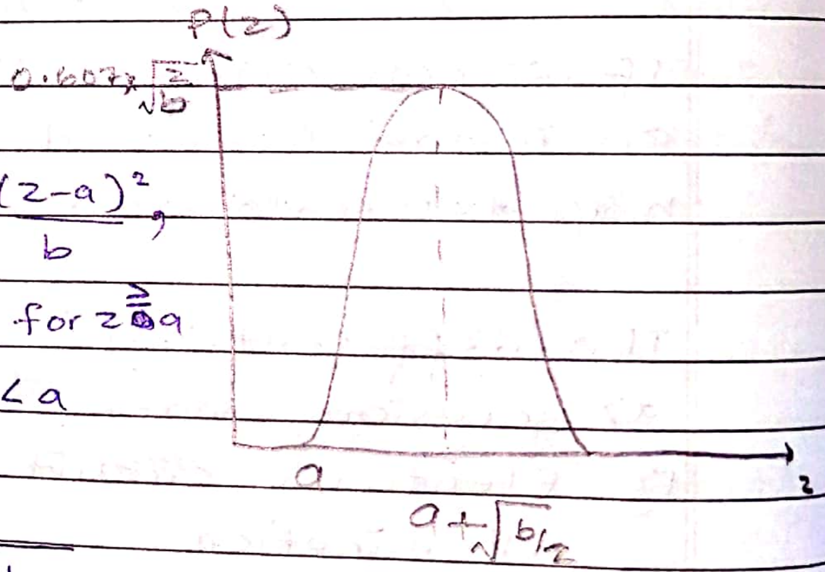
z: Gray level

μ : Mean of random variable z

σ^2 : variance of z

b) Rayleigh Model :

$$p(z) = \begin{cases} \frac{2}{b}(z-a)e^{\frac{-(z-a)^2}{b}}, & \text{for } z \geq a \\ 0, & \text{for } z < a \end{cases}$$



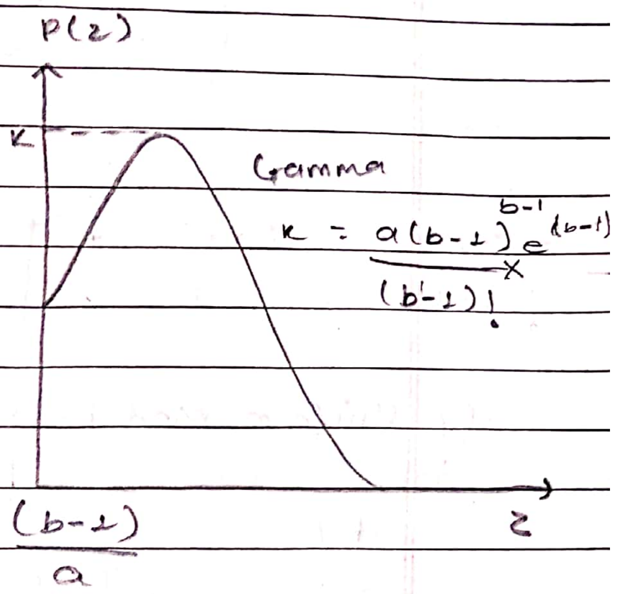
Mean: $\mu = a + \sqrt{\frac{\pi b}{4}}$

Variance: $\sigma^2 = \frac{b-(4-\pi)}{4}$

c) Erlang Model

:

$$p(z) = \begin{cases} \dfrac{a^b z^{b-1}}{(b-1)!} e^{-az}, & \text{for } z \geq 0 \\ 0, & \text{for } z < 0 \end{cases}$$



$a > 0$, $b \in I+$

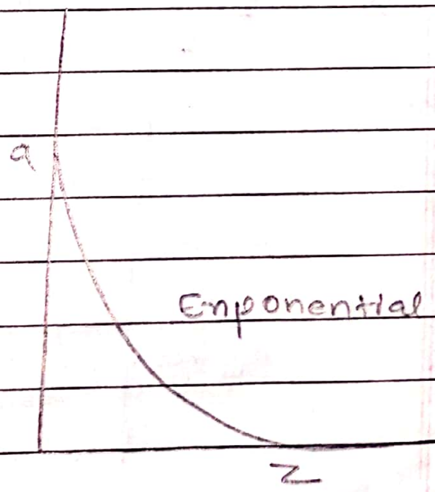
Mean: $\mu = b/a$,

Variance: $\sigma^2 = b/a^2$

d) Exponential Model

:

$$p(z) = \begin{cases} ae^{-az}, & \text{for } z \geq 0 \\ 0, & \text{for } z < 0 \end{cases}$$



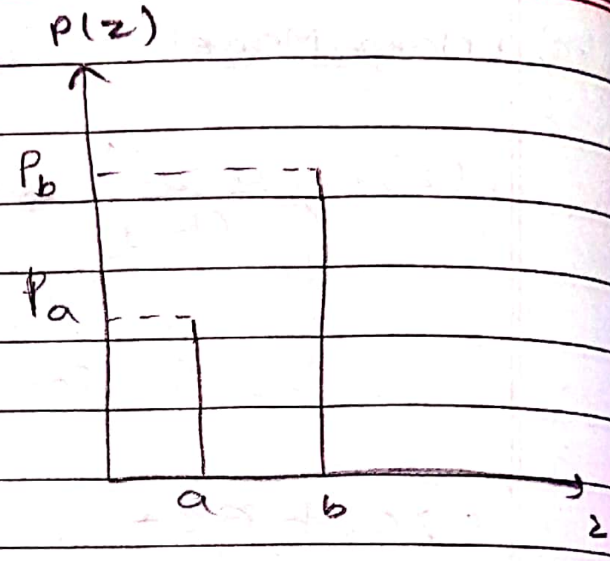
where $a > 0$,

Mean $=$ $\mu = 1/a$

Variance : $\sigma^2 = 1/a^2$

e) Impulse Model :

$$p(z) = \begin{cases} P_a & \text{for } z = a \\ P_b & \text{for } z = b \\ 0 & \text{otherwise} \end{cases}$$

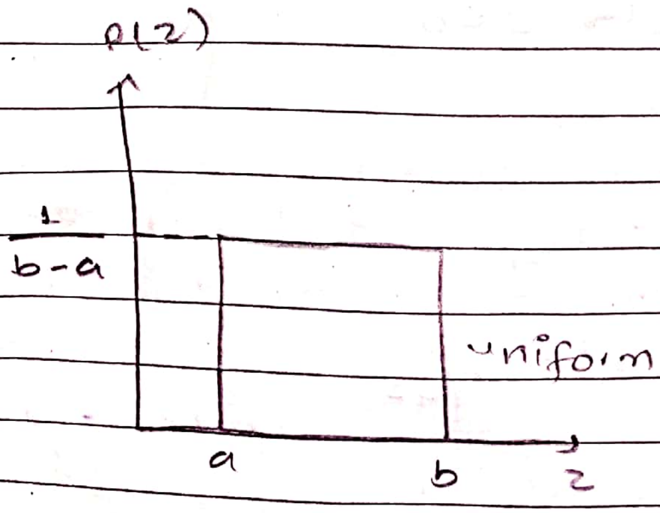


f) Uniform Model :

$$P(z) = \begin{cases} \frac{1}{b-a}, & \text{if } a \le z < b \\ 0, & \text{otherwise} \end{cases}$$

The mean & variance are given by,

$$\mu = \frac{a+b}{2}, \quad \sigma^2 = \frac{(b-a)^2}{12}$$



* Compression

~~90 minut~~ One 90 minute color movie, each second playing 24 frames, when is digitized with each frame of 512 X 512, each pixel having 3 components ~~(R,G,B)~~ R, G, B (8 bits) has total size

$$\frac{90 \times 60 \times 24 \times 8 \times 3 \times 512 \times 512}{8}$$

~~No. of pixel~~ $= 1.0192 \times 10^{11}$ bytes

$= 99532800$ KB

$= 97200$ MB

$= 94.92$ GB

Image compression addresses the problem of ~~amount~~ of reducing the amount of data required to represent the digital image. The basis for image reduction process is, removal of redundant data. The compressed image is decompressed to reconstruct the original images and approxiamation of it. Thus, image compression is, the reduction of number of bits required to transmit the image probably without any remarkable loss of ~~informatn~~ information.

The image encoding process through the ~~ckt~~ channel is called image compression & Image decoding process is known as image ~~dea~~ decompression.

Q. Explain the block diagram of image compression.

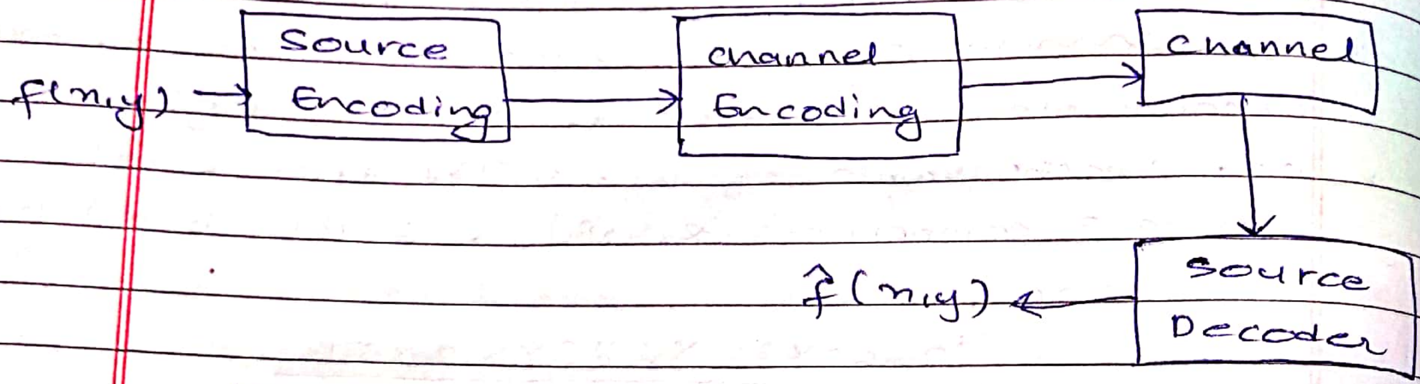


Fig: Block diagram of image compression

A typical image compression ~~sys~~ system comprises of two main blocks an encoder (compressor) and decoder (decompressor). The image $f(x,y)$ is fed to the encoder which encodes the image so as to make it suitable for transmission. The decoder receives the transmitted signal and reconstructs the output image $\hat{f}(x,y)$. If the system is an error free one $\hat{f}(x,y)$, will be the replica of $f(x,y)$.

↳ The encoder and decoder are made up of two blocks each. The encoder is made up of a source encoder removes the input redundancies while the Channel encoder removes ~~the input~~ redundancies ~~while the channel encoder~~ increases the noise immunity of the source encoders.

Date ____
Page ____

* Application of Data compression :

a) Broadcasting TV
b) Remote sensing through satellite
c) Military communication
d) Facsimile transmission and telecommunication

* Data Redundancy :

Image contains data that either provide no relevant information or simply restate that, which is already known. This is called data redundancy. In digital image processing, 3 basic data redundancy can be identified as: i) Coding Redundancy
ii) Interpixel Redundancy
iii) Psycho-visual Redundancy

Coding Redundancy :-

We know that, average no. of bits require to represent each pixel in a given image is

$$L_{avg} = \sum_{k=0}^{L-1} l(k) \cdot p(k)$$

Where, l = length

Hence, no. of bits required to represent the image is ~~max~~ $M \times N \times L_{avg}$.

Where, $M \times N$ in total no. of pixels in the given image.

Maximum compression ratio is ~~achiee~~ achieved when Lavg is minimised. But, coding the gray level in such a way that Lavg is not minimised then resulting image is ~~sat~~ said to be coding redundant.

ii) Interpixel Redundancy

:- It is related to interpixel correlation within an image. usually the value of certain pixel in the image can be resonably predicted from the values of its neighbour in the image. Thus, the values of ~~its~~ Individual pixel carries relatively small amount of information and much more information about pixel value that can be inferred on the ~~basis~~ bacis of its neighbour value. These dependencies between pixel values in the image is known as interpixel dependency.

iii) Psycho-visual Redundancy

:- The eye doesnot respond with equal sensitivity to all visual information. Human perception searches for Important features (edges, pattern/texture) and doesn't perform quantitative analysis of every pixel in the image. So, psycho-visual redundancy takes into ~~perp~~ perception of human visual system. ~~Image comp~~

Date
Page

* Image compression techniques

Compression techniques

- Pixel coding
  - PCM / Quantization
  - Run length Coding
  - Bit plane coding
- Predictive coding
  - Delta Modulation
  - Line by Line DPCM
  - 2-D PCM
  - Interface Technique
  - Adaptive

Compression Technique

- Transform coding
  - Threshold
  - Multi-dimension
  - Adaptive
- Others
  - Hybrid
  - Color
  - vector

Pixel coding Technique

In this technique, pixel is processed independently ignoring the Interpixel dependencies. This techniques firstly removes the interdependencies between neighbour pixels and then remove the redundant data.

This technique is further divided into following 3 types:-

a) Pulse code modulation (PCM)
:

$v(t) \rightarrow \boxed{\text{Sampling}} \rightarrow \boxed{\text{Quantization}} \rightarrow \boxed{\text{Digital Modulation}} \rightarrow$ Transmit

In PCM technique, the incoming video signal ($v(t)$) is sampled, quantized and coded by suitable code word, before inputting it to the digital modulator for transmission.

The quantizer output is coded by a fixed length binary code word, having usually 8 bits. The formula for achieving rate of compression of PCM is given by,

$$R_{PCM} = \frac{1}{2} \log_2 \frac{\sigma v^2}{\sigma q^2}$$

where, σv^2 is variance of quantized input
σq^2 is mean square deviation

b) Run length coding

: It is a very simple form of data compression which represents each sub-sequence of identical symbols by a ~~pos~~ ~~pai~~ pair like (l, a) where ~~the~~ 'l' is the subsequence and 'a' is the recurring symbol.

For eg:

aaabbcccd is encoded as, 3a2b3c1d

↳ ~~Inter entropy~~ Entropy coding

: Entropy coding encodes the given set of symbols with ~~min~~ minimum no. of bits required to represent them.

The most important entropy coding is Huffman coding. The theoritical minimum average of bits that are required to transmit a particular source string is known as entropy. And it can be computed by following formula.

$$\text{Entropy}(H) = -\sum_{i=1}^{N} P_i \log_2 P_i$$

It has 2 major parts. They are:

a) Construction of probability tree

b) Assigning code to each node

It gives the variable length code-word & highest probability assigns short-path and lowest probability assigns longest-path, i.e. longest code length.

~~Chapter-7~~

~~Image Segmentation~~

* Lossy and Lossless predictive coding

:

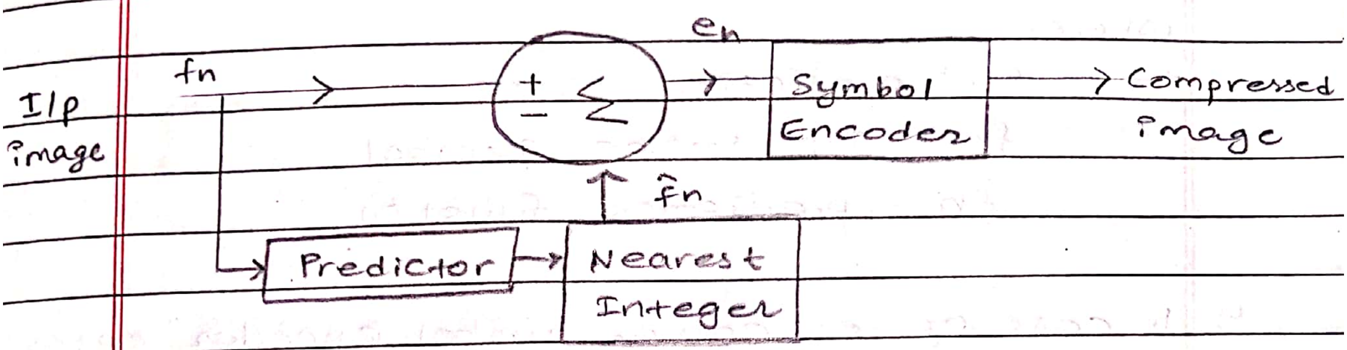


Fig: a) Encoder

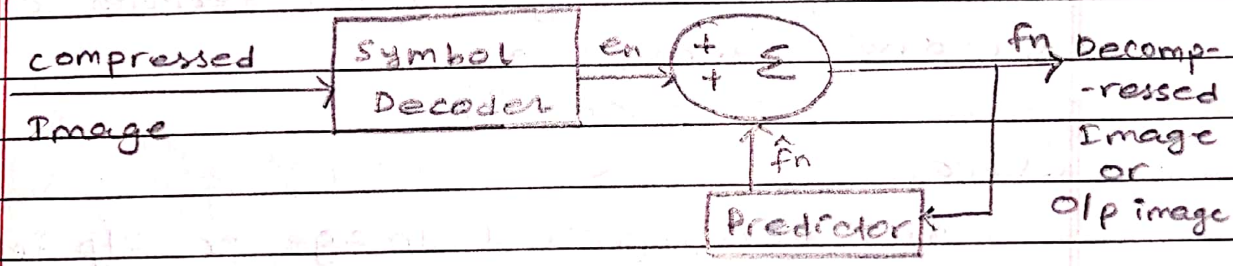


Fig: b) Decoder

Fig: Lossless Predictive Coding

- There is no information loss and image can be reconstructed exactly the same as the original.
- So, compression process is reversible.
- The system consists of an Encoder and a Decoder each containing an identical predictor.
- The predictor generates anticipated value of that pixel based on some number of past

values. The output is then rounded to the nearest integer.

- Then the differences or prediction error is calculated as:

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

where,

e_n = prediction error

f_n = I/p Image symbol

$\hat{f}_n$ = prediction function

- In case of encoding, symbol encoder eases variable length coding techniques to encode symbol.
- In case of decoding, symbol decoder cases decoding function as:

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

where,

f_n = Decompressed Image or O/p image

- In must cases, the predictor is formed by a linear combination of 'm' previous pixel. i.e.

$$\hat{f}_n = \text{round}\left[\sum_{i=1}^{m} \alpha_i f_{m-i}\right]$$

where,

α = predictive coefficients

m = order of linear prediction

b) Lossy Predictive coding

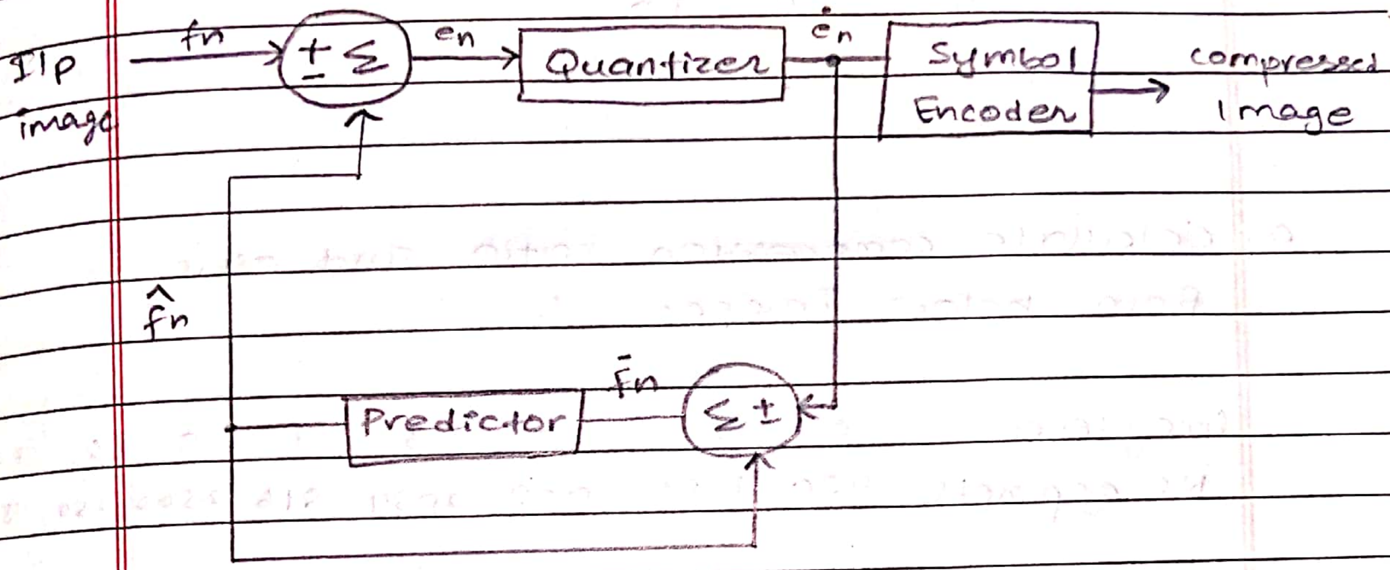


Fig: Encoder

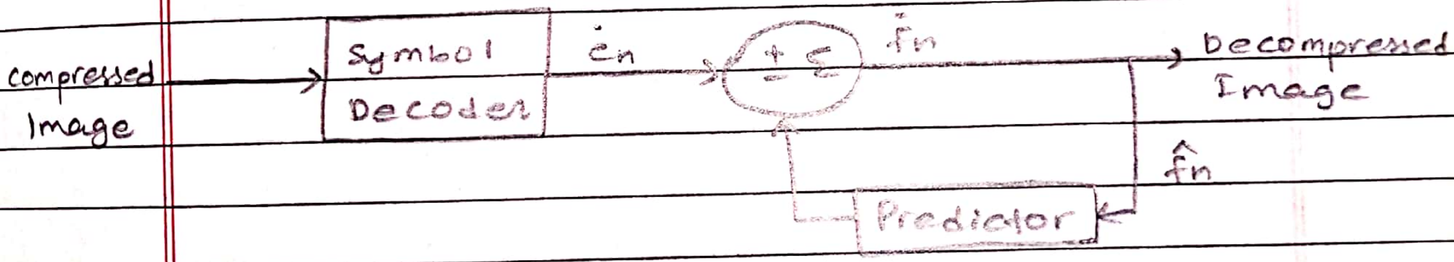


Fig: Decoder

↳ The major difference in this technique is that a quantizer is added between the symbol encoder and the point where the prediction ~~en~~ error ~~en~~ is ~~calculated~~. calculated.

↳ The quantizer absorbs the nearest value of a prediction error and maps it into a limited range of output denoted by $\dot{e}_n$ at encoding and $\dot{f}_n$ at decoding, which is defined as,

$\dot{e}_n$ is given by:

$$\dot{e}_n = \dot{f}_n - \hat{f}_n$$

And,

$\dot{f}_n$ is given by:

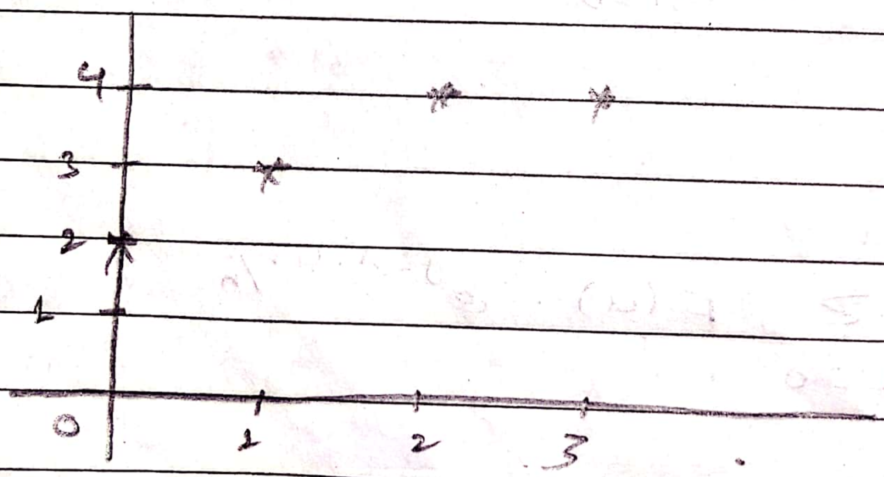
$$\dot{f}_n = \dot{e}_n + \hat{f}_n$$

Q.

q. Calculate compression ratio and efficiency from below image.

Gray level	0	1	2	3	4	5	6	7
No. of pixels	400	1350	659	2034	816	2560	250	1800

Q. Find the fourier spectrum of the function at point 0 and 1 on the basis of given figure below.



Given,

$f(0) = 2$

$f(1) = 3$

$f(2) = 4$

$f(3) = 4$

Now,

$e^{-j\pi/2}$

$= \cos\frac{\pi}{2} - j\sin\pi/2$

~~for u=0~~,

$|F(u)| = \sqrt{R^2(u) + I^2(u)}$

For $u=0$,

$$F(0) = \frac{1}{4}\sum_{n=0}^{3} f(n) \cdot e^{-j2\pi \cdot \frac{0 \cdot n}{4}}$$

$$= \frac{1}{4}\sum_{n=0}^{3} f(n)$$

$$= \frac{1}{4}[f(0) + f(1) + f(2) + f(3)]$$

$$= \frac{13}{4} + j \cdot 0$$

$\therefore$ $|F(u)| = \sqrt{(13/4)^2 + (0)^2} = \frac{13}{4}$

$|P(u)| = \left(\frac{13}{4}\right)^2 = \frac{169}{16}$

$\phi = \tan^{-1}\left(\frac{0}{13/4}\right) = 0$

For $u=1$,

$$F(1) = \frac{1}{4} \sum_{n=0}^{n-1} f(n) \cdot e^{-j2\pi \cdot un/4}$$

$$= \frac{1}{4} \sum_{n=0}^{3} f(n) \cdot e^{-j2\pi \cdot \frac{un}{4}}$$

$\cancel{2} j\pi \cdot \frac{1}{\cancel{4}2}$

$$= \frac{1}{4} \left[f(0) + f(1) * e^{-j\pi \cdot \frac{1 \cdot 1}{2}} + f(2) * e^{-j\pi} + f(3) * e^{-j\pi 3/2} \right]$$

$$= \frac{1}{4} \left[2 + 3 * \left(\cos\frac{\pi}{2} - j\sin\frac{\pi}{2}\right) + 4 * (\cos\pi - j\sin\pi) + 4 * \left(\cos\frac{3\pi}{2} - j\sin\frac{3\pi}{2}\right) \right]$$

$$= \frac{1}{4} [2 + 3(0 - j \cdot 1) + 4(-1 - j0) + 4(0 - j(-1))]$$

$$= \frac{1}{4} [2 - 3j - 4 + 4j]$$

$$= \frac{1}{4} [-2 + j]$$

$$= \frac{j}{4} - \frac{1}{2}$$

Now,

$$|F(u)| = \sqrt{R^2(u) + I^2(u)}$$

$$= \sqrt{(-1/2)^2 + (1/4)^2} = \sqrt{1/4 + 1/16} = \sqrt{5/16}$$

$$= \sqrt{5}/4$$

$$|P(u)| = R^2(u) + I^2(u)$$

$$= \left(\frac{-1}{2}\right)^2 + \left(\frac{1}{4}\right)^2$$

$$= \frac{1}{4} + \frac{1}{16}$$

$$= 5/16$$

$$\phi = \tan^{-1}\left(\frac{I(u)}{R(u)}\right) = \tan^{-1}\left(\frac{1/4}{-1/2}\right)$$

$$= \tan^{-1}\left(-\frac{1}{2}\right)$$

$$= -26.56$$

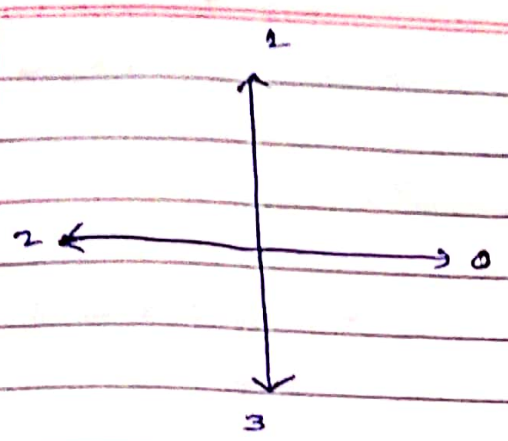
* # Importance

~~DCT~~ DFT is most widely used orthogonal transform in the field of Image processing. However in this transform, it takes n^2 multiplication and $[n*(n-1)]$ addition to calculate 1D discrete fourier transform of n data points. It can also be used as ~~diagona~~ diagonalised circullar matrix.

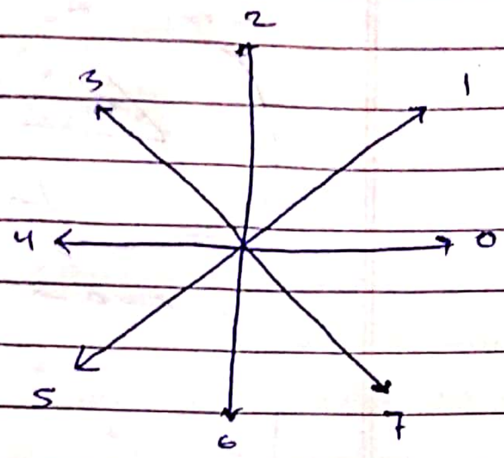
* Discrete cosine Transform

DCT is a fourier related transform similar to DFT but it is only concerned with real numbers only. It can be defined as,

$$DCT = C(u) = \alpha(u) \sum_{n=0}^{N-1} f(n) \cdot \cos\left[\frac{(2n+1)u\pi}{2N}\right]$$



4-directional chain code



8-directional chain code

The chain code of a boundary depends on the starting point. However, chain code can be normalised with respect to starting point by treating it as a circular sequence of direction number & redefining the starting point so that the resulting sequence of numbers is a integer magnitude. we can normalize by using rotation, ie 90° or 45° for the codes using the difference of chain code.

b) Signature

Signature is a one-dimensional function representing the boundary and maybe generated in various ways. One of the simplest method is plotting the distance from the centroid point, to the boundary as a function of angle. Signature generated by this approach is ~~Invarian~~ invariant to the translation but they depend on rotation & scaling.