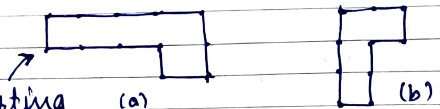
Q: Show that images are translation and rotational invariate.



For (a)

Chain code →

1 0 0 0 0 3 3 2 1 2 2 2

Taking the minimum integer by shifting

⇒ 0 0 0 0 3 3 2 1 2 2 2 1

Calculating difference clockwise

⇒ 0 0 0 1 0 1 1 3 0 0 1 1

Taking the minimum integer by shifting

⇒ No shifting as it is already minimum

0 0 0 1 0 1 1 3 0 1 1

Shape number → min (I^st difference clockwise)

COMPRESSION

PSNR   SNR ↓, image quality ↓

Compression
- Lossless
  1. Binary Huffman
  2. Non-Binary Huffman (Ternary)
  3. Shanno-Fano Code
  4. Arithmetic Code
- Lossy

Some values and terms :

Average length $(\bar{L}) = \sum_{k=0}^{N} p_k l_k$, where l_k is length after coding.

Entropy $H(S) = -\sum_{k=0}^{N} p_k \log_2 p_k = \frac{1}{\log_{10} 2} \sum -p_k \log_{10} p_k$

Efficiency $(\eta) = \frac{H(S)}{(\bar{L})}$

7	6	5
3	4	3
0	1	2

→ Total no. of bits required in original image, $n_1 = 3 \times 3 \times 3 = 27$ bits

Total no. of bits required in compressed image, n_2.

Now on the basis of n_1 and n_2 calculate the compression ratio. That is,

$$C_R = n_1/n_2$$

Saving % $= \left(1 - \frac{1}{C_R}\right) \times 100\%$

Example → original image = 10 bit
compressed image = 7 bit

$$\therefore C_R = \frac{10}{7}$$

$$\% \text{ saving} = \left(1 - \frac{1}{\frac{10}{7}}\right) \times 100\% = 30\%$$

Binary Huffman coding

1) Different intensity values are given in the image we need to calculate the frequency of each pixel intensity and make the following table :

GL	10	20	30	40	50
f_k	4	6	9	4	2
p_k	0.16	0.24	0.36	0.16	0.08

2) Now sort the GLs based on p_k in decreasing order.

30	20	10	40	50
0.36	0.24	0.16	0.16	0.08

3) Start combining the values from the right-most end two at a time and place them as the leaf node of a binary tree, and their combined value as their parent. Repeat step 2 and 3 until the parent becomes 1 that is we reach root node.

(i)

30	20	10	40	50
0.36	0.24	0.16	0.16	0.08

(ii)

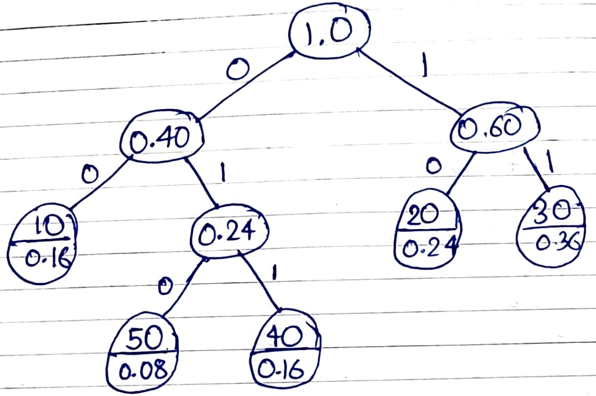
30	20		10
0.36	0.24	$\boxed{0.24}$	0.16

(iii)

	30	20
$\boxed{0.40}$	0.36	0.24

(iv) $\boxed{0.60}$ $\boxed{0.40}$

(v) 1.0



	code	l_k
10	00	2
20	10	2
30	11	2
40	011	3
50	010	3

$$\text{Average length} = \sum_{k=1}^{N} p_k \times l_k = 0.16\times 2 + 0.24\times 2 + 0.36\times 2 + 0.16\times 3 + 0.08\times 3$$

$$= 2.24 \text{ bits/symbol}$$

Let us assume that 256 different intensity levels are there

$$n_1 = 5\times 5\times 8 = 200 \text{ bits/symbol}$$
$$n_2 = 5\times 5\times 2.24 = 56 \text{ bits.}$$
$$C_R = \frac{n_1}{n_2} = \frac{200}{56} = 3.57$$

$$\%\ \text{saving} = \left(1 - \frac{1}{C_R}\right)\times 100\% = 72\%$$

This means image is represented in 28% only.

NOTE :

- Greyscale image → 8 bit
- If nothing is mentioned then check from the given ~~image~~ image for bits.

- **Arithmetic Code**

Dictionary Based
- LZ - 77 (Past data)
- LZ - 78 (Future data)
- LZW

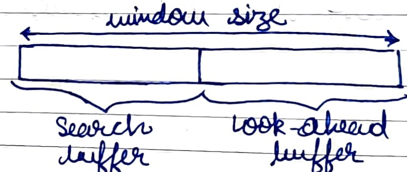
→ Static → Before I/P dictionary & data inside the dictionary

→ Adaptive → Maintain the dictionary (update the dictionary to data)

① Run length coding
② Difference coding
③ Bit plain coding

} (Self study)

④ Lossless Predictive Coding (DPCM) → If taught then not in exam



Triplet $\langle o, l, (c) \rangle$

o → offset, l → length, (c) → code

cabeac
cabeacadab

1. Run length coding:

Direction from left to right

0 0 0 0
1 1 1 2
1 2 2 2
3 3 0 0

(0,4) (1,3) (2,1) (1,1) (2,2)
(3,2) (0,2)

Direction like following (ziczac)

0 0 0 0
2 1 1 1
2 2 2 3
3 3 0 0

(0,4) (1,3) (2,4) (3,3) (0,2)

LZ-77

sliding window = 7

Look ahead = 6

							c	a	b	e	a	c	adabeaee aceead
						c	a	b	e	a	c	a	<0,0,c(c)>
					c	a	b	e	a	c	a	d	<0,0,c(a)>
				c	a	b	e	a	c	a	d	a	<0,0,c(b)>
			c	a	b	e	a	c	a	d	a	b	<0,0,c(e)>
		c	a	b	e	a	c	a	d	a	b	e	a	<3,1,c(c)>
c	a	b	e	a	c	a	d	a	b	e	a	e	e	<5,1,c(d)>
8	a	d	a	b	e	a	e	e	a	e	e	a	d	<7,4,c(e)>
	e	a	e	e	a	e	e	a	d				<3,3,c(e)>
	e	e	a	e	e	a	d						<6,1,c(d)>

decoding

Wsize = 13 Look ahead = 6 sliding window = 7

after decoding
cabeacad abeaeeaeead

<0, 0, c(d)>

LZ-78

c a b e a c a d a b

Dictionary

Index	String	encode
1	c	(0, c(c))
2	a	(0, c(a))
3	b	(0, c(b))
4	e	(0, c(e))
5	ac	(2, c(c))
6	ad	(2, c(d))
7	ab	(2, c(b))
8	ea	(4, c(a))
9	ee	(4, c(e))
10	ae	(2, c(e))
11	ead	(8, c(d))

LZW

Index	Dictionary value
1	a
2	b
3	c

Index	string	encoded o/p
1	a	
2	b	
3	c	
4	ab	1
5	b a	2
6	abb	4
7	bab	5
8	b c	2
9	ca	3
10	aba	4
11	abba	6
	–	

Recieve sequence		Index	String
		1	A
		2	B
		3	C
1	a	–	–
2	b	4	ab
4	ab	5	ba
5	ba	6	abb
2	b	7	bab
3	c	8	bC
4	ab		
6			
1			

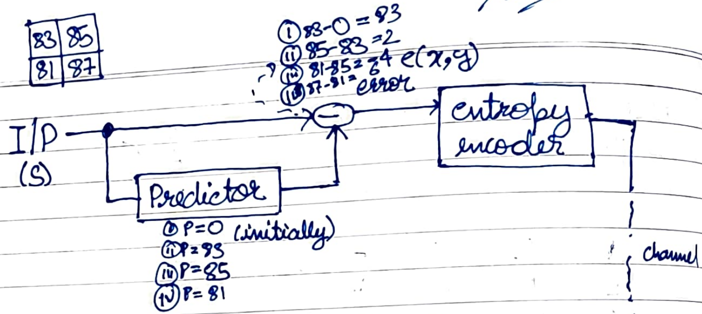
- Predictive Coding

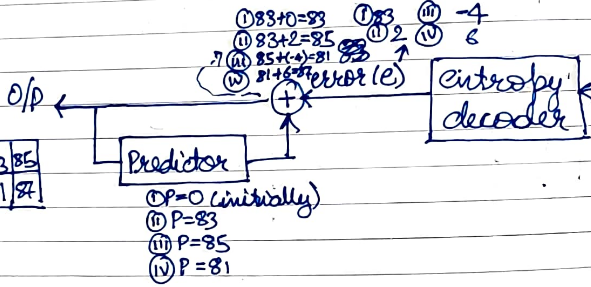
1) a) Differential pulse Code Modulation (DPCM) without quantization [Lossless]
   b) DPCM with quantization [Lossy]
2) Block ~~coding~~ truncate coding

Predicting the current pixel value based on the value of previously processed pixel

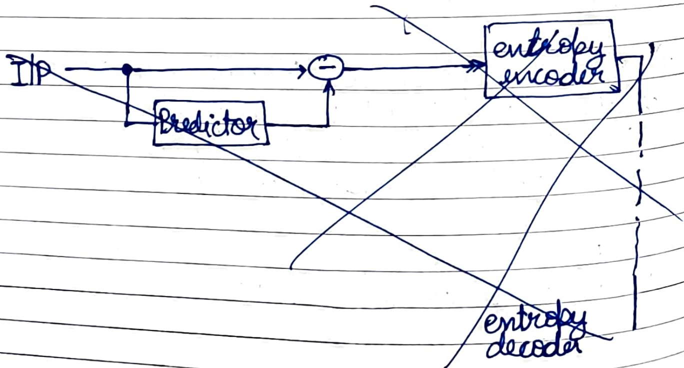
Lossless

83	85
81	87



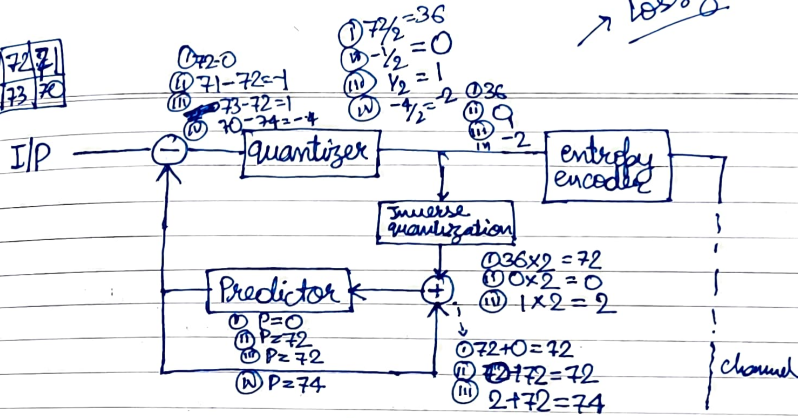


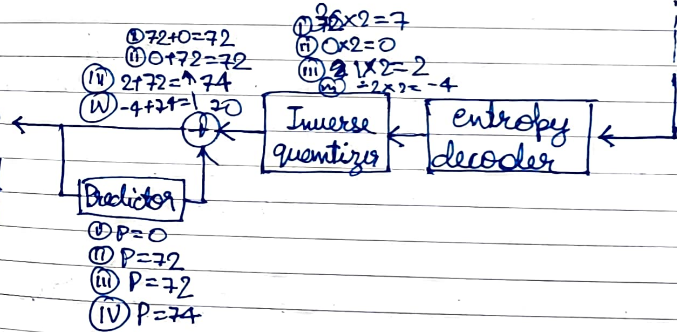
83	85
81	87



Lossy

72	71
73	70





72	72
74	70

PSNR value calculation GLCM matrix

Block Truncate coding (Lossy)

A			B
65	75	80	70
72	75	82	68
84	72	62	65
66	68	72	86

C D

Steps:

① Partition the image into set of non overlapping block.

$X(m,n) \rightarrow f(m,n)$ ~~sub~~ block → block

A:

$\bar{f} = 71.75$

$\bar{f^2} = 5164.75$

$\sigma^2 =$ ~~5088~~ 16.6875

$\sigma =$ ~~4.08~~ 4.08

$B = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ q=~~4~~

a = 71.75

② For each block calculate mean, mean square and variance by using formula.

Mean. $\bar{f} = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} f(i,j)$

Mean square $\bar{f^2} = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} f^2(i,j)$

Variance, $\sigma^2 = \bar{f^2} - (\bar{f})^2$

$\sigma = \sqrt{\bar{f^2} - (\bar{f})^2}$

③ ~~For~~ Binary allocation matrix B(m, n) is constructed by using

$$B(m,n) = \begin{cases} 1 & \text{if } f(m,n) > \bar{f} \\ 0 & \text{otherwise} \end{cases}$$

Assume q number of pixels greater than mean ~~that~~ meaning no. of 1s ~~ones~~ in B → binary allocation matrix.

④ Two values of variable a and b for each block is calculated by

$a = \bar{f} - \sigma \sqrt{\frac{q}{m-q}}$ → no. of ~~ones~~ 1s in B

→ total number of pixel in each block

$b = \bar{f} + \sigma \sqrt{\frac{m-q}{q}}$

⑤ After calculating the value of a and b ~~reconst~~ reconstruct the image

$$\hat{f}(m,n) = \begin{cases} a & \text{if } B(m,n) = 0 \\ b & \text{if } B(m,n) = 1 \end{cases}$$

Consider the whole image as a block then by calculation

$\bar{f} = 72.25$

$(\bar{f^2}) = 5261.25$

$\sigma^2 = 41.1875$

$\sigma = 6.4177$

$$B = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$a = \bar{f} - \sigma\sqrt{\left(\frac{q}{m-q}\right)} = 67.25 \simeq 67$$

$$b = 80.53 \simeq 80$$

67	81	81	67
67	81	81	67
81	67	67	67
67	67	67	81

Modified Block Truncate Coding

→ Diving the blocks based on certain criteria

→ Binary matrix can be converted into ternary or quaternary allocation table

Q: How to compare the quality of image?

$$MSE = \frac{1}{M \times N} \sum_{m=0}^{m-1} \sum_{n=0}^{n-1} \left(f(m,n) - g(m,n) \right)^2$$

$f(m,n)$ → original image

$g(m,n)$ → output image

$$SNR = 10 \log_{10} \left[\frac{\frac{1}{m \times n} \sum\sum f(m,n)^2}{\frac{1}{m \times n} \sum\sum (f(m,n) - g(m,n))^2} \right]$$

$$PSNR = 10 \log_{10} \left[\frac{(2^{b} - 1)^2}{\frac{1}{mn} \sum\sum (f(m,n) - g(m,n))^2} \right]$$

b → no. of bits required to represent the max intensity

Texture Feature

No numerical
(5 marks) 5-10 lines answer
one question in exam

approx. features (112 to 178):
- GLCM
- HOG
- Wavlet
- LTE
- Tamura

Principle Component Analysis (PCA)
↓
Performes dimensionality reduction
⇓
Feature selection technique

- **GLCM**

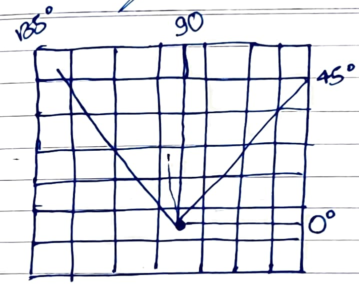
(Grey Level Co-occurence Matrix)

$\frac{i}{j}$ ⇒ 90° j over i ⇒ 45° i over j ⇒ 135° $i \mid j$ ⇒ 0°

Every selection will have different GLCM

① Co-occurence / co-occurence distribution means distribution of co-occuring values at a given offset (angle).
OR
Represents the distance and angular spatial relationship over an image.

Question ↑ rule Morphology operator] Image morphology



distance and angle
Two parameters are must for construction of GLCM.

Eg:

2	1	2	0	1
0	2	1	1	2
0	1	2	2	0
1	2	2	0	1
2	0	1	0	1

$P_{135}(i,j)$
$d = 1$

$C =$ (co-occurence matrix)

	0	1	2
0	0	2	2
1	2	1	2
2	2	3	2

(2) → Total number of distinct grey level and not max. intensity.

① Energy $= \sum_{i,j=0}^{N-1} (C_{i,j})^2$

② Contrast $= \sum_{i,j=0}^{N-1} C_{i,j}(i-j)^2$

③ Entropy $= -\sum_{i,j=0}^{N-1} C_{ij} \log P_{ij}$

(iv) Uniformity

(v) Homogeneity

(vi) Difference Moment (order k) $= \sum \frac{C_{ij}}{(i-j)^k}$

(vii) Inverse difference moment (order k)

(viii) Correlation factor $\xrightarrow{\text{sign}}$ $= \sum_{i,j=0}^{N-1} \frac{(i-\mu)(j-\mu)}{\sigma^3}$

(ix) Shade feature $= \operatorname{Sgn}(A)|A|^{1/3}$

$$A = \sum_{i,j=0}^{N-1} \frac{(i+j-2\mu)^3 P_{ij}}{\sigma^3\left(\sqrt{2(1+C)}\right)^3}$$

C → correlation factor

$$\operatorname{sgn}(x) = \begin{cases} -1, & x<0 \\ 0, & x=0 \\ 1, & x>0 \end{cases}$$

(x) Prominance feature $= \operatorname{sgn}(B)|B|^{1/4}$

$$B = \sum_{i,j=0}^{N-1} \left[\frac{(i-j-2\mu)^4 P_{ij}}{4\sigma^4(1+C)^2}\right]$$

(xi) Mean

(xii) Sum of average

(xii) Variance

(xiv) Sum of square

(xv) Max. probability

(xvi) Min —"—

~~(xvii) Sum entropy.~~ (xvii) Sum of probability

(xviii) Sum of entropy

- Confusion Matrix

Predicted category

		T	F
Actual category	T	TP	FN
	F	FP	TN

↓ confusion matrix

True Positive: Actual category (AC) and Predicted (PC) category are both positive. No. of times model correctly classified.

True Negative: AC and PC both are negative

False Positive: AC is negative PC is positive

False Negative: AC is positive PC is negative

TPR → True Positive Rate
FPR → False Positive Rate

On the basis of CM calculate the parameter known as accuracy, precision, recall or sensitivity, specificity, TPR, FPR.

$$\text{F1 score} = \frac{2 \times P \times R}{P + R}$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

$$TPR = \frac{TP}{TP+FN}$$

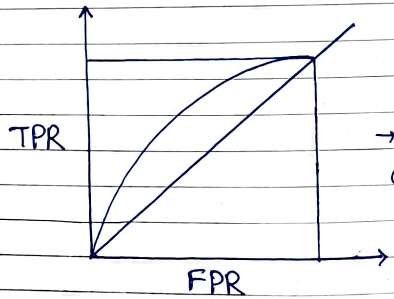
$$\text{(P) Precision} = \frac{TP}{TP+FP}$$

$$FPR = \frac{FP}{TN+FP}$$

$$\text{(R) Recall (sensitivity)} = \frac{TP}{TP+FN}$$

$$\equiv 1 - \text{specificity}$$

$$\text{Specificity} = \frac{TN}{TN+FP}$$



→ After approximation or interpolation we get smooth curve.

(Exam question)

Q: A group of 165 person standing in a line, each person walk up to you and you either predict a person have a disease or not have the disease. In reality each person either has a disease or not a disease.

i) If you tell someone "you have the disease" and they do have the disease they you say "go to room A".

ii) If you tell someone "you donot have disease" and they do not have it then you say "go to room B".

iii) If you tell someone "you have the disease" and they do not have it then you say "go to room C".

iv) If you tell someone "you dont have disease" and they do have it then you say "go to room D".

Consider every person goes in four rooms A, B, C and D. If there a 100 people in room A, 50 people in room B, 10 people in room C and 5 people in room D, then find the sensitivity, specificity, accuracy, and F1-score.

A:

	T	F
T	100	5
F	10	50

TP = 100
TN = 50
FP = 10
FN = 5

$$\text{Accuracy} = \frac{100+50}{165} = \frac{150}{165} = 0.91 = 91\%$$

$$\text{(R) Sensitivity} = \frac{100}{100+5} = \frac{100}{105} = 0.95 = 95\%$$

$$\text{Specificity} = \frac{50}{50+10} = \frac{50}{60} = 0.83 = 83\%$$

$$\text{(P) Precision} = \frac{100}{100+10} = \frac{100}{110} = 0.91 = 91\%$$

$$\text{F1 score} = \frac{2 \times P \times R}{P+R} = 0.92 = 92\%$$