Brief Solutions

University of Toronto
Faculty of Applied Science and Engineering

FINAL EXAMINATION
ECE462H1S, Multimedia Systems

April 24, 2023, 6:30 – 9:00 pm
Instructor: D. Hatzinakos

Instructions:

1. The exam counts for 50% of overall mark.
2. Please solve all problems. Do not show only final answers. You should demonstrate how the answer has been obtained by including intermediate results and explanations wherever needed.
3. All answers must be written only at the examination booklet.

QUESTION 1. (10 points)

Assume that in an EZW based compression of a 4x4 image f(x,y), the decoder receives the following information:

2 level wavelet transform
To=16
Do=11000000   p ttt   4
So=1
D1=10000011110000   zttpptt   7
S1=010
D2=11011111111101000011000000   pnpppnttptt t   13
S2=10101100100

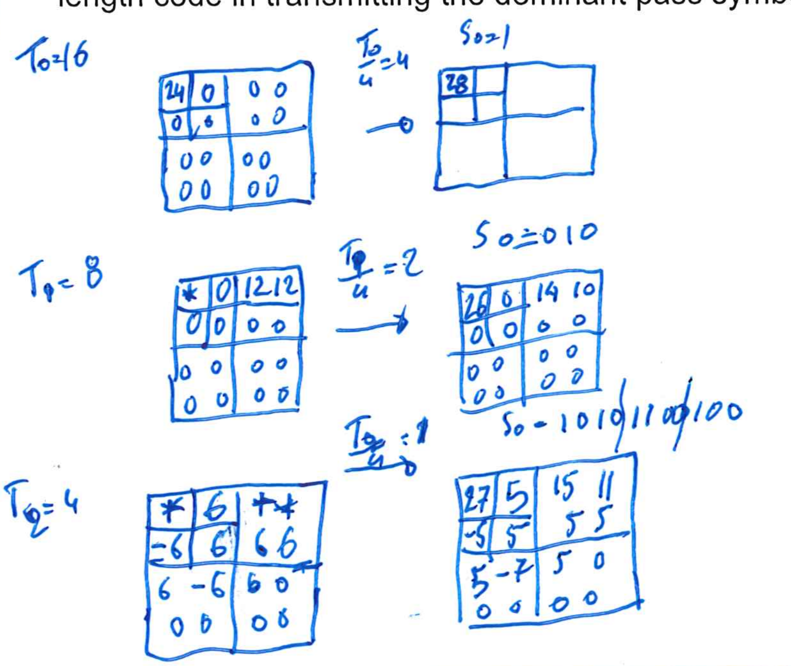
Assuming that the following codes have been used

Zerotree root	t	00
Significant positive	p	11
Significant negative	n	01
Isolated zero	z	10

$\log_2 x = \frac{\log_{10} x}{\log_{10} 2} = \frac{\log_{10} x}{0.3}$

1. What is the reconstructed 3-pass wavelet transform?
2. What is the entropy of the dominant pass symbols? Design a Huffman code for the dominant pass symbols. Will there be any benefits by using a variable length code instead of a fixed length code in transmitting the dominant pass symbols?

(1)



(2.) Counting p, n, t, z in D_0, D_1, D_2

9 2 12 1 = 24 in total.

probabilities $\frac{9}{24}$ $\frac{2}{24}$ $\frac{12}{24}$ $\frac{1}{24}$

$$\text{Entropy} = \frac{9}{24}\log_2\frac{24}{9} + \frac{2}{24}\log_2\frac{24}{2} + \frac{12}{24}\log_2\frac{24}{12} + \frac{1}{24}\log_2 24$$

$$= \frac{3}{8}\log_2\frac{8}{3} + \frac{1}{12}\log_2 12 + \frac{1}{2}\log_2 2 + \frac{1}{24}\log_2 24$$

$$= \frac{3}{8}\log_{10}\frac{(8/3)}{0.3} + \frac{1}{12}\log_{10}\frac{12}{0.3} + \frac{1}{2}\log_2 2 + \frac{1}{24}\log_{10}\frac{24}{0.3} =$$

$$= 0.35 + 0.13 + 0.5 + 0.08 = 1.06$$

Huffman Code

t	p	n	z
$\frac{12}{24}$	$\frac{9}{24}$	$\frac{2}{24}$	$\frac{1}{24}$

t	0
p	10
n	110
z	111

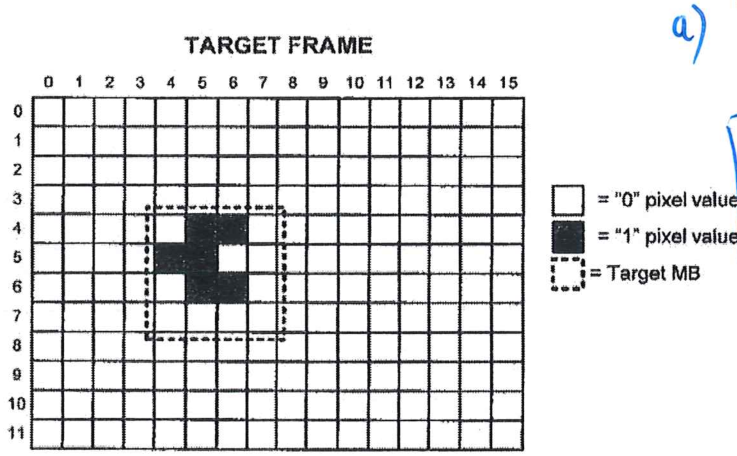
$\text{rate} = \frac{12}{24} + \frac{18}{24} + \frac{6}{24} + \frac{3}{24} = 1.625$

Yes HC is better that fixed length code.

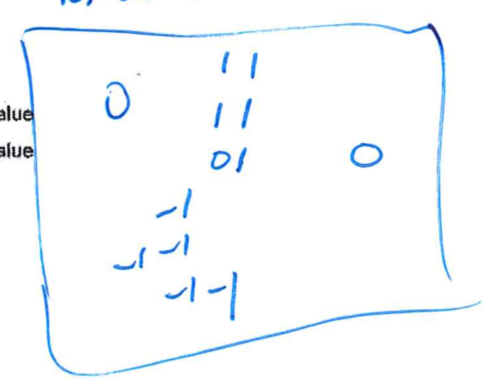
QUESTION 2 (10 points)
Consider the two video frames shown below.

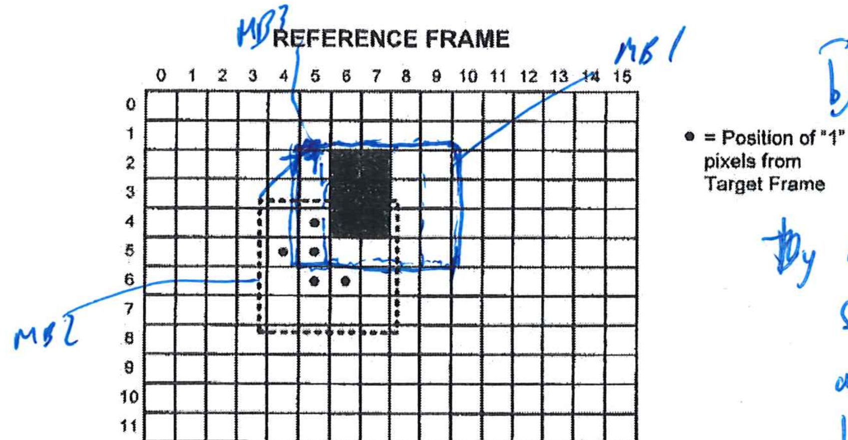
a) A first inspection of the frames suggests that due to the small motion between the two frames, motion compensation may not be necessary and that the difference between the two frames (pixel by pixel) can be transmitted in place of the target frame.
b) Alternatively an exhaustive search is carried out for the indicated target macro block (4x4), using SAD as the distance measure, to estimate the motion vectors. Then a predicted target frame is calculated and the prediction error with the actual frame is obtained.

Which one of the two approaches is better from a compression point of view?
Calculate the relevant signals in both cases and indicate the amount of bits needed for each case.



a) Difference between two frames results in





b) Motion vector
(3,4) → (1,5)
By inspection the minimum value SAD occurs with the above motion vector and the difference after motion compensation will be

Note: the "dots" and the dashed box in the reference frame are aids to show the relative position of the target frame content.

$$SAD(i,j) = \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} \left| T(k,l) - R(k+i, l+j) \right|$$

(b) is better

③ ① Let Y and X be column vectors $Y = AX$

Energy of Y: $P_Y = E\{Y^H Y\} = E\{(AX)^H AX\} = E\{X^H \underbrace{A^H A}_{I} X\} = E\{X^H X\}$

② Transmission order will be $I_0\ P_2\ B_1\ P_5\ B_4\ B_3\ P_8\ B_7\ B_6$

if P_2 is damaged then all of the following frames are affected

if B_6 is damaged no other frame is affected.

③ The logarithmic perceptual property of vision

④ It relates to the entropy of the source

	1	0
probability	$\frac{3}{4}$	$\frac{1}{4}$

so $H = \frac{3}{4}\log_2\frac{4}{3} + \frac{1}{4}\log 4 = 0.225 + \frac{1}{4}\cdot 2 = 0.725$

⑤ The image is completely random. $P(1) = P(0) = \frac{1}{2}$

⑥ 1 0.9 0.8 0.7 0.6 0.5 0.4 0.3 0.2, 0.1

$\hat{X}(n) = a\,x(n-1)$ where $a = \frac{R(1)}{R(0)} = \frac{0.351}{0.385} = 0.91$

where $R(0) = \frac{1}{10}(1^2 + 0.9^2 + 0.8^2 + \dots 0.1^2) = \frac{1}{10}(\quad) = 0.385$

$R(1) = \frac{1}{10}(1\cdot 0.9 + 0.9\cdot 0.8 + 0.8\cdot 0.7 \dots + 0.2\cdot 0.1) = 0.351$

So $\boxed{\hat{X}(n) = 0.91\,x(n-1)}$

⑦ No, because downsampling may cause aliasing

⑧ P frames compute motion estimation and compensation from previous I or P frames

B frames compute motion estimation from previous and next I or P frames

QUESTION 3 (1.5 point each, total 30 points).

1.A square matrix A is unitary. If X and Y are two N-dimensional vectors and Y=A X, show that Y and X have the same energy.

2. A video sequence is MPEG encoded as where the letter represents the type of frame and the number represents the location of the frame in the sequence. The GOP is I0B1P2B3B4P5B6B7P8I9. If the frame *P2* is damaged which of the other frames will be viewable and which ones will not ? What happens if frame B6 is damaged?

3. What property of the human visual system justifies the choice of non-uniform quantization?

4. Given a binary signal with three times as many ones than zeros, what is the maximum lossless compression that we can achieve?

5. A black and white image has entropy equal to 1 bit. What can we conclude for the image?

6. Given the sequence 1,0.9,0.8,0.7,0.6,0.5,0.4,0.3,0.2,0.1 estimate the required autocorrelation lags and then design a linear predictor (MSE optimum) of length 1.

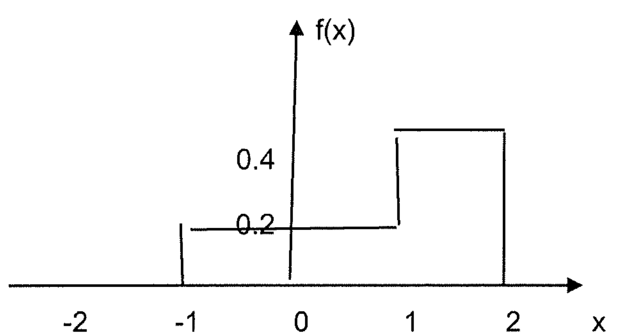
7. An engineer is told that a digital image must be compressed by a factor of 4. After carefully examining the image the engineer decides to down-sample the image by deleting every-other row and column. Do you think this makes sense?

8. What is the difference between B frames and P frames in video encoding?

9. What are different forms of redundancies used in signal compression?

10. Given the same rate (in bits/sample) use of vector quantization results in lower distortion than when scalar quantization is used. Is this statement true or false? Provide some justification for your answer.

11. A signal X has the following pdf *f (x)*

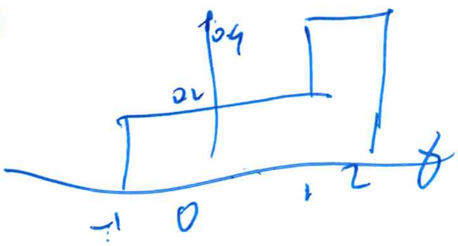


Explain why and under what conditions a uniform quantizer may be sufficient for this signal.

(9) - Spatial redundancy
- time redundancy
- Phycoacoustics
- Frequency redundancy

(10) We cannot say for sure. It depends on the image structure and properties

(11)



A uniform quantizer is justified if pdf is flat over all intervals of Δ.
for example $\Delta = 0.5$ or $\Delta = 1$

(12) x: 31.5 16.2 5.6 3.1 1.1 $\hat{x} = \text{round}\left(\frac{x}{q}\right)$

let $q=8$ then
$\hat{x}$: 4 2 1 0 0
x_{recon} 32 16 8 0 0

errors: -0.5 0.2 1.6 3.1, 1.1
$(x - \hat{x})$

let $q=2$ and reject
$\hat{e}$: 0 0 1 2 1
e_{rec}: 0 0 2 4 2

new reconstruction x_{rect} e_{rec} = 32 16 6 4 2
new error: -0.5, 0.2 -0.4 -0.69 -0.9

→ MSE $= \frac{0.5^2 + 0.2^2 + 0.4^2 + 0.69^2 + 0.9^2}{5} = 0.35!$

(13) MP3. Better reconstruction of audio signals (higher quality)

(14) 3 minute song require (Sampling at 40 kHz) 3×60 (sec) $\times$ 40K samples/sec $\times$ 16 bits/sample = 115,200 K bits/sec = 115.2 Mbits/sec

So $\frac{512 \times 8 \text{ bits}}{115.2} = 35$ songs

(15) Motion compensation, and I, P, B frames are similarities
Advantages of H.264 compared to MPEG-2 is better quality
in signal (video) reconstruction

(16) No bits are wasted or missing.

(17) Simplicity of hardware + good signal quality

(18) $4^3 = 64$ so 3 bits will be saved

(19) In ~~[illegible]~~ uniform quantization Δ is constant
In non-uniform quantization Δ is variable

(20) Celp provides lower compression than LPC10 but better signal quality
Celp uses vector quantization LPC-10 uses scalar quantization