University of Toronto
Faculty of Applied Science and Engineering

FINAL EXAMINATION
ECE462H1S, Multimedia Systems

April 29, 2019, 9:30 am
Instructor: D. Hatzinakos

Instructions:

1. Type A exam
2. Non-programmable calculators are allowed
3. Please solve all four problems.
4. All answers must be written in the examination booklet. Do not write any answers in this problem handout.

QUESTION 1. (10 points)

An image f(x,y) has the following wavelet transform.

21	**6**	**15**	**5**
-9	**3**	**6**	**3**
3	**-3**	**0**	**-1**
1	**0**	**0**	**0**

You are asked to apply EZW coding in compressing the image under the following assumptions

- We have a budget of 30 bits.
- We can send the initial threshold value T_0 separately without effecting the bit budget
- We use the following codes

Zerotree root	zr	00
Significant positive	sp	11
Significant negative	sn	01
Isolated zero	iz	10

1. Show all the steps of the EZW coding and the final transmitted bitstream .
2. Show the original and reconstructed final image assuming the following wavelet transform matrices.

$$A_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}, \quad A_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

QUESTION 2 (10 points)

The autocorrelation sequence R(k)=E{x(n)x(n+|k|)} of a short voiced speech segment x(n), n=0,1,2, …,7 , takes the values

k :	0	1	2	3	4	5	6	7
R(k):	1.0	0.5	0.3	0.1	0.8	0.3	0.2	0.1

An LPC coder uses this information to generate the excitation signal

$$e(n) = \delta(n) + \delta(n-P) + \delta(n-2P) + ...$$

where *P* is the pitch period in samples and $\delta(n)$ is the delta function. It also uses the following model for the vocal tract filter

$$x(n) = a_1 x(n-1) + a_2 x(n-2) + e(n)$$

where *G* is an appropriate gain.

1. Estimate the Pitch value P .
2. Estimate the coefficients a_1, a_2 of the vocal track filter
3. How many bits do you save by using this LPC process?
4. Using the above speech model estimate the first 8 values of the synthesized speech.

QUESTION 3 (20 points)

Answer all of the following questions by providing sufficient explanation:
(2 points for each question)

1. Given a speech signal, what is the difference between coding it using a 10th order predictor, compared to coding it using LPC 10?

2. Say you have a video with a frame rate of 30 frames per second. Your motion prediction system achieves a good level of performance using logarithmic motion search with p=16. If you reduce the frame rate to 20 frames per second , how this will affect the choice for p? Justify your answer.

3. Using the same scenario as in the previous question, but instead of changing the frame rate, you decrease the resolution by a factor of 2 both horizontally and vertically. How the value of p is affected in this case? Justify your answer.?

4. What does it mean to have a separable transform? Why is it useful?

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

6. A video sequence uses the following GOP display order: IBBBPBBPBB.. What is the transmission order?

7. What is the difference between the DCT and MDCT transforms?

8. What are the various forms of redundancy used in compression theory?

9. Design a huffman code for the word SISIFOS. What is the rate of the transform and how it relates to the entropy?

10. What are the different forms of redundancy used in compression theory?

QUESTION 4 (10 points)

We wish to compress the following image via JPEG

f(x,y)

$$\begin{matrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 4 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{matrix}$$

1. Apply a 4x4-point DCT (given the transform matrix T below) on f(x,y) and then quantize by applying the quantization table Q and rounding the result.

A second 4x4 image g(x,y) is obtained by filtering f(x,y) with the 2x2 filter h(x,y) and then by truncating the resulting 5x5 image to 4x4 pixels.

h(x,y)

$$\begin{matrix} 0.25 & 0.25 \\ 0.25 & 0.25 \end{matrix}$$

g(x,y)

$$\begin{matrix} 0 & 0 & 0 & 0 \\ 0 & 0.25 & 0.75 & 0.5 \\ 0 & 1.25 & 2.5 & 1.25 \\ 0 & 1 & 1.75 & 0.75 \end{matrix}$$

2. Apply a 4x4-point DCT on g(x,y) and then quantize by applying the quantization table Q and rounding the result.

3. Based on the above results can we conclude which of the two images will be better compressed by JPEG?

4. If you repeat the process by filtering g(x,y) with h(x,y) will JPEG compression of the resulting image improve further or not ? Explain your answer.

$$T = \begin{pmatrix} 0.25 & 0.25 & 0.25 & 0.25 \\ 0.4 & 0.2 & -0.2 & -0.4 \\ 0.25 & -0.25 & -0.25 & 0.25 \\ 0.2 & -0.4 & 0.4 & -0.2 \end{pmatrix}, \qquad Q = \begin{pmatrix} 10 & 10 & 50 & 100 \\ 10 & 10 & 50 & 100 \\ 50 & 50 & 50 & 100 \\ 100 & 100 & 100 & 100 \end{pmatrix}$$