University of Toronto
Faculty of Applied Science and Engineering

FINAL EXAMINATION
ECE462H1S, Multimedia Systems

April 25, 2018
2:00-4:30 pm
Instructor: D. Hatzinakos

Instructions:

1. Type A exam (closed book)
2. All type 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)

a) An image f(x,y) has the following 2-level Haar wavelet transform.

53	-22	21	-9
14	-12	13	-11
15	-8	9	7
34	-2	-6	10

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

- 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

i.) Show all the steps of the EZW coding and the final transmitted bitstream

ii.) Assuming that we have a budget of 16 bits at the most what will be the reconstructed image?

The 4-point Haar matrices (you may derive similarly a 2-point Haar matrix) are defined below:

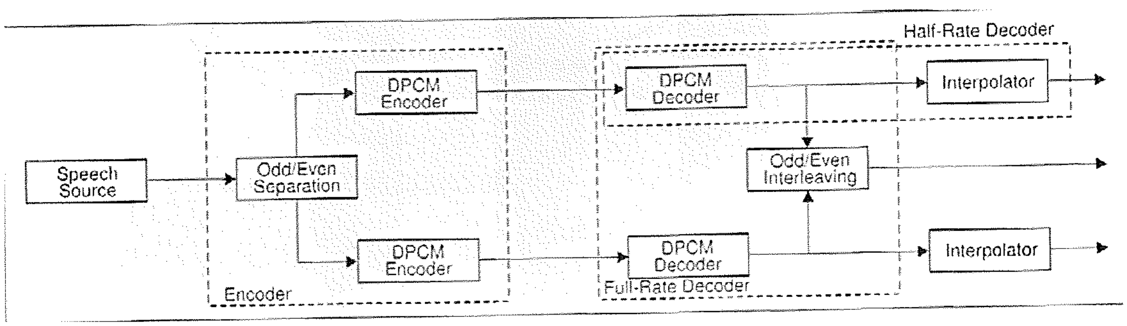
$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}, \qquad H^{-1} = H^T$$

QUESTION 2 (10 points)

The following sequence x(n), n=0, 1, ..., 19 of real numbers has been obtained by sampling a speech segment

{ 2.3, 2.1, 3.2, 1.2, 1.3, 2.3, 2.5, 3.8, 3.8, 2.5, 2.0, 1.4, 1.0, 0.6, 0.0, -0.3, -0.5, -0.8, -1.2,-1.4}

The speech is split into odd and even streams , that is e(n)=x(2n) and o(n)=x(2n+1), n=0,1,... which are separately encoded by DPCM and then are sent to a decoder via different channels. The decoder reconstructs the original sequence either by interleaving (assuming both streams are received and are reliable) or by interpolation (assuming only one of the two streams is ether available or reliable). This process which is depicted in the figure below is called "channel splitting encoding method" and is used to improve average speech quality over lossy channels.



a) Generate and encode via DPCM the odd and even data streams by using 4 bits per sample to encode the differential errors. Provide the values of the quantized error sequences.
b) Reconstruct the original sequence by using DPCM decoding of both streams and then odd/even stream interleaving; (full rate decoder case). Calculate the MSE between original and reconstructed sequences.
c) Reconstruct the original sequence by DPCM decoding only the even stream and by linearly interpolating over two samples, that is by taking the average of two even samples to estimate the odd sample in between; (half rate decoder case). Calculate the MSE(even) between the original and reconstructed sequences. Repeat the process with the odd stream and find the MSE(odd). Then find the average MSE for this case.
d) Compare the performances of the full rate vs half rate decoders and comment both qualitatively and quantitatively on the nature and effect of errors in both cases.

QUESTION 3 (20 points)

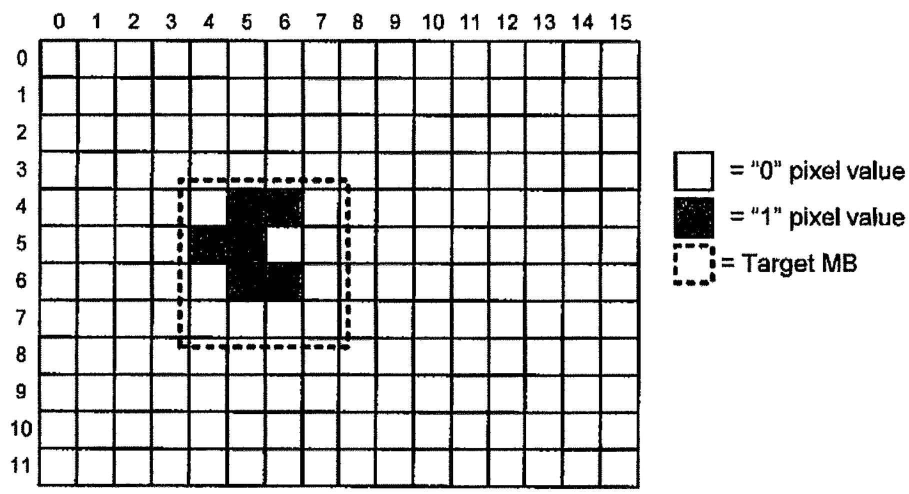
Answer all of the following questions by providing sufficient explanation:
(2 points for each question)

1. In MPEG audio coding, what have the three audio layers in common?
2. What is the most common procedure to assess sound quality in compression standards?
3. A source generates the symbols {1,2,3,4,5,6} with respective probabilities {0.3,0.26, 0.14,0.13,0.09,0.08}. What is the entropy of the source?
4. A signal takes values between 0 and 1 with probability ¾ and between 1 and 2 with probability ¼. By making proper assumptions design a two level quantizer to exhibit minimum MSE performance.
5. A signal has autocorrelation $R(k) = 0.9^{|k|}$, k=-0,1,2,... What type of compression strategy would you recommend and why? Assume that the signal is down-sampled by M and then compressed, would you follow the same compression strategy or not?
6. In an MPEG audio compression process it is found that signal to quantization noise ratio SNR=-10dB and the signal to mask ratio SMR=-16 dB. What does this imply for the number of assigned bits?
7. Under what conditions or assumptions is uniform quantization of a signal a meaningful choice?
8. What is the motivating principle behind transform coding?
9. The bit-stream order in an MPEG GOP is IPBBIBBPBBIBB. What will be the corresponding sequence in display order? Use arrows to indicate dependencies between frames.
10. Given that singing has characteristics of both speech and music which compression algorithms would you expect to be most successful on songs?

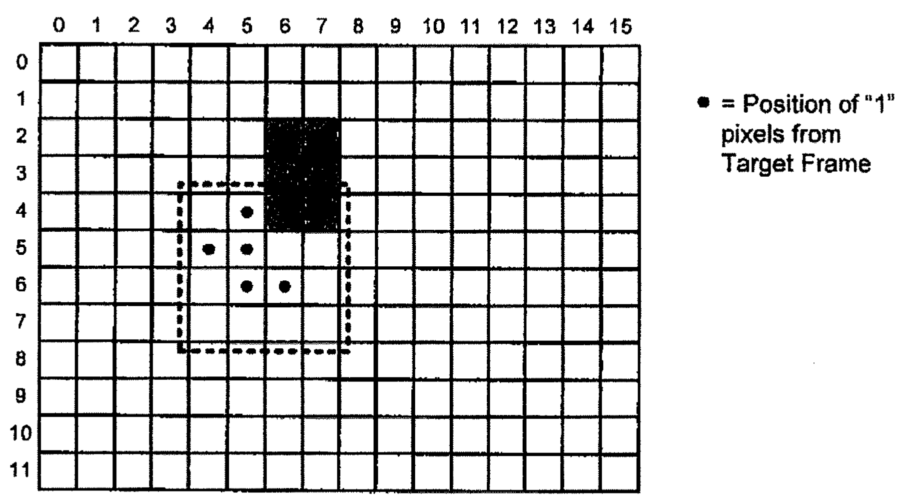
QUESTION 4 (10 points)

Consider the two video frames shown below. Perform a logarithmic motion search for 6the indicated target macroblock (4x4), using p=3 and SAD as the distance measure. Show each SAD calculated for each iteration, and the final motion vector.

TARGET FRAME



REFERENCE FRAME



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

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