Final Exam Data Compression COMP 423

Prof. Michael Langer Thurs. April 24, 2008

Answer all questions in the exam book. Return this exam sheet.

Calculators or other electronic devices are NOT permitted.

This exam has two pages. There are a total of 30 points.

1. **(7 points)**

   Let (X_1, X_2, X_3) be a sequence of three Bernoulli trials with probability $p_0 = \frac{5}{8}$, $p_1 = \frac{3}{8}$.

   (a) What is the tag $T(\vec{x})$ of the *most* probable sequence ?

   (b) What is the tag $T(\vec{x})$ of the *least* probable sequence ?

   (c) What is the codeword of the sequence $(0, 1, 0)$?

   Hint: The equations for computing the arithmetic code for independent variables X_j are:

$$l_{k+1} = l_k + (u_k - l_k)\, F(pred(i_{k+1}))$$

$$u_{k+1} = l_k + (u_k - l_k)\, F(i_{k+1})$$

2. **(6 points)**

   (a) What are the possible values of the quantization function $Q(\ )$ for a *mid-tread* quantizer with $\Delta = 4$ and *no overload* errors ?

   (b) How would your answer change if there were overload errors?

   (c) Using *differential coding* and the midtread quantizer in (a), and given a sequence that begins

$$25,\ 18.1,\ 11.3,\ \ldots$$

   what are the quantization *levels* encoded, and what is the decoder's estimate of this sequence ?

   (d) Using *linear predictive coding* where the prediction for X_{j+1} is $\frac{1}{2}\hat{X}_j$ and using the same quantizer as in (a), give the encoding levels and the decoder's estimate of the sequence.

   (e) Characterize those sequences for which the method in (d) compresses *long sequences* better than the method in (c).

3. **(7 points)**

(a) Sketch out the basis functions of the 1D discrete cosine transform, for the case $m = 8$ and for $k = 0, 1, \ldots, 7$, that is, sketch one basis function for each of the eight k values.
<u>Hint:</u> The discrete cosine transform is given by:

$$\mathbf{C}_{j,k} = \begin{cases} \frac{1}{\sqrt{m}}, & k = 0 \\ \sqrt{\frac{2}{m}} \cos(\frac{\pi}{m} k \ (j + \frac{1}{2}) \ ), & k = 1, \ldots, m-1 \end{cases}$$

and note $\sqrt{\frac{2}{8}} = \frac{1}{2}$.

(b) JPEG uses a 2D DCT to encode blocks of image intensities. Consider the following block of intensities:

$$\begin{bmatrix} 35 & 35 & 35 & 37 & 112 & 110 & 111 & 110 \\ 35 & 35 & 35 & 35 & 111 & 112 & 112 & 111 \\ 35 & 35 & 35 & 35 & 112 & 111 & 112 & 112 \\ 34 & 36 & 35 & 34 & 111 & 112 & 114 & 112 \\ 35 & 35 & 35 & 35 & 112 & 112 & 112 & 113 \\ 35 & 35 & 37 & 38 & 111 & 111 & 112 & 112 \\ 35 & 35 & 35 & 36 & 112 & 112 & 114 & 112 \\ 36 & 35 & 36 & 35 & 112 & 112 & 112 & 115 \end{bmatrix}$$

Give a rough estimate of the corresponding 8×8 matrix of transformed values that you would expect to find. (No calculators allowed and limited time available – hence I am looking for a rough estimate only.)

(c) JPEG uses differential encoding for the DC values, but not the AC values. Briefly explain why. In particular, what assumptions are being made about the AC values, both *within* blocks and *between* blocks ?

4. **(10 points)**

(a) Describe how the DCT is used in MP3 audio coding and decoding.

(b) What is masking? How is it used in MP3 audio coding?

(c) Audio signals often have two channels which are encode the sounds to be played on left and right headphones/speakers. Suppose the various sound events (people singing, instruments playing, etc) represented by the signal occur at slightly different times in the two channels. Invent a *non-DCT-based* method for *lossless* compression of two-channel audio that makes use of these temporal shifts, and *that is inspired by MPEG video.*

In your answer, you do *not* need to consider *why* there are such temporal offsets. But for your interest: the offsets are related to the fact that sound travels at a finite speed. For example, two microphones at different positions in 3D space might record a particular sound event at slightly different times.