LINKÖPING UNIVERSITY

Written Exam in
Image and Audio Coding
TSBK02

25th October 2017 8:00 - 12:00

Location:	TER2
Examiner:	Harald Nautsch
Teacher:	Harald Nautsch, 1361
Department:	ISY
Exam code:	TEN1
Number of problems:	5
Number of pages:	4 + formula collection
Permitted equipment:	Calculator, "Tables and Formulas for Image Coding and Data Compression"
Grades:	0-13 U 14-19 3 20-25 4 26-30 5
Other:	Answers can be given in English or Swedish. The teacher will visit around 9:15 and 10:45.

1 a) Explain how coding using *analysis by synthesis* works and give an example of a coding method where this is used.

(2 p)

b) Two psychoacoustic phenomena are *frequency masking* and the *hearing threshold*. Explain what these are and how they can be utilized when coding audio signals.

(2 p)

c) Describe how mp3 coding of audio signals works.

(2 p)

d) In hybrid coding of video, motion compensated prediction is used. Explain how this works, both on the coder and the decoder side.

(2 p)

e) Explain how JPEG coding of still images works.

(2 p)

2 A memoryless source has the alphabet $\mathcal{A} = \{a, b, c\}$ with the symbol probabilities $P(a) = 0.8,\ P(b) = 0.15$ and $P(c) = 0.05$.

a) Construct a Huffman code for the source that gives an average rate of at most 1 bit/symbol.

(2 p)

b) Code the sequence a, c, a, a from the source using arithmetic coding. Give both the resulting interval and the binary codeword.

(2 p)

3 A random variable X with probability density function

$$f_X(x) = \begin{cases} \frac{3}{2}(1-x)^2 & ; \ 0 \le x \le 1 \\ \frac{3}{2}(1+x)^2 & ; \ -1 \le x < 0 \\ 0 & ; \ \text{otherwise} \end{cases}$$

is quantized to two levels.

Find the decision borders and reconstruction points such that the resulting distortion is minimized.

Calculate the resulting distortion.

(4 p)

4 A stereo music signal is modelled as a stationary gaussian process (V_i, H_i), where V_i are the samples of the left channel and H_i are the samples of the right channel.

The signal statistics have been estimated as

$$E\{V_i\} = E\{H_i\} = 0$$

$$E\{V_i^2\} = E\{H_i^2\} = 0.170$$

$$E\{V_i \cdot V_{i+1}\} = E\{H_i \cdot H_{i+1}\} = 0.144$$

$$E\{V_i \cdot H_i\} = 0.165$$

When coding the signal, V_i is predicted from $\hat{V}_{i-1}$. The prediction error is quantized uniformly and coded using arithmetic coding. The stepsize of the quantizer is chosen such that the rate is 8 bits per sample. Then H_i is predicted from $\hat{V}_i$. The prediction error is quantized uniformly and coded using an arithmetic coder. The stepsize is chosen such that the rate is R_H bits per sample.

Find the optimal predictor coefficients for the two predictors.

How should R_H be chosen if we want to have the same distortion for the right channel that we have for the left channel?

Calculate the resulting signal to noise ratio in dB.

State all assumptions that you make.

(6 p)

5 An N point transform matrix is constructed by taking polynoms of degree 0 to $N-1$ and then orthogonalizing and normalizing them. For a four point transform the matrix looks like

$$\mathbf{A} = \begin{pmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 3/\sqrt{20} & 1/\sqrt{20} & -1/\sqrt{20} & -3/\sqrt{20} \\ 1/2 & -1/2 & -1/2 & 1/2 \\ 1/\sqrt{20} & -3/\sqrt{20} & 3/\sqrt{20} & -1/\sqrt{20} \end{pmatrix}$$

Suppose that we want to code a onedimensional signal X_i using this transform. X_i is modelled as a gaussian process with mean 0 and autocorrelation function

$$R_{XX}(k) = E\{X_i \cdot X_{i+k}\} = 0.91^{|k|}$$

We want to quantize the transform components with scalar Lloyd-Max quantizers so that the average rate is 2 bits/sample and the average distorsion is minimized.

Allocate bits to the transform components to meet the demands and calculate the resulting signal to noise ratio in dB.

How much lower would the signal to noise ratio be if we instead used a four point Hadamard transform?

(6 p)