LINKÖPING UNIVERSITY

Written Exam in
Image and Audio Coding
TSBK02

28th May 2018   8:00 - 12:00

Location:	TER2
Examiner:	Harald Nautsch
Teacher:	Harald Nautsch, 1361
Department:	ISY
Exam code:	TEN1
Number of problems:	6
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 CELP coding of speech works.

(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) Explain how JPEG coding of still images works.

(2 p)

2 Describe in detail how modern hybrid coders and decoders for video signals work. MPEG4, H.264 and HEVC are examples of such coders.

(4 p)

3 A memoryless source has the alphabet $\mathcal{A} = \{a, b, c\}$ with the symbol probabilities $P(a) = 0.7$, $P(b) = 0.2$ and $P(c) = 0.1$.

a) Construct a Huffman code for the source that gives an average rate of at most 1.2 bit/symbol.

(2 p)

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

(2 p)

4 A random variable X with probability density function

$$f_X(x) = \begin{cases} \frac{\pi}{4} \cos \frac{\pi x}{2} & ; \ -1 \leq x \leq 1 \\ 0 & ; \ \text{otherwise} \end{cases}$$

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)

5 An image is modelled as a stationary twodimensional zero mean normally distributed process $X_{i,j}$ (i and j are coordinates in the image). From a large set of data, the auto correlation function $R_{XX}(k,l) = E\{X_{i,j} \cdot X_{i+k,j+l}\}$ has been estimated as

$$R_{XX}(0,0) = 2209, \quad R_{XX}(0,1) = 2054$$

$$R_{XX}(1,0) = 2002, \quad R_{XX}(1,1) = R_{XX}(1,-1) = 1976$$

The image is coded using a linear predictor of the form

$$p_{ij} = a_1 \cdot \hat{X}_{i-1,j} + a_2 \cdot \hat{X}_{i,j-1}$$

The prediction error is quantized uniformly and then coded using a memoryless arithmetic coder.

How should the predictor coefficients a_1 and a_2 be chosen if we want to minimize the distortion of the coder at a given rate?

What is the lowest rate that can be used if we want to have a signal-to-noise ratio of at least 42 dB?

Compare your result to the lowest rate that can be achived by just using the quantizer and the memoryless arithmetic coder (no predictor) in order to reach 42 dB.

(6 p)

6 The Discrete Sine Transform of type I (DST-I) is a transform suitable for signals with a high-frequency characteristic. For a signal $x_i, i = 0, \ldots, N-1$ of length N, the DST-I coefficents $\theta_j, j = 0, \ldots, N-1$ are defined as

$$\theta_j = \sqrt{\frac{2}{N+1}} \cdot \sum_{i=0}^{N-1} x_i \cdot \sin\left(\frac{(i+1)(j+1)\pi}{N+1}\right)$$

For $N = 4$, this corresponds to using the transform matrix given by

$$\mathbf{A} = \sqrt{\frac{2}{5}} \begin{pmatrix} \sin\frac{\pi}{5} & \sin\frac{2\pi}{5} & \sin\frac{3\pi}{5} & \sin\frac{4\pi}{5} \\ \sin\frac{2\pi}{5} & \sin\frac{4\pi}{5} & \sin\frac{6\pi}{5} & \sin\frac{8\pi}{5} \\ \sin\frac{3\pi}{5} & \sin\frac{6\pi}{5} & \sin\frac{9\pi}{5} & \sin\frac{12\pi}{5} \\ \sin\frac{4\pi}{5} & \sin\frac{8\pi}{5} & \sin\frac{12\pi}{5} & \sin\frac{16\pi}{5} \end{pmatrix} = \begin{pmatrix} a & b & b & a \\ b & a & -a & -b \\ b & -a & -a & b \\ a & -b & b & -a \end{pmatrix}$$

where $a = \sqrt{\frac{2}{5}} \sin\frac{\pi}{5}$ and $b = \sqrt{\frac{2}{5}} \sin\frac{2\pi}{5}$.

A signal is modelled as a stationary gaussian process X_n with zero mean and auto correlation function $R_{XX}(k)$

$$R_{XX}(k) = E\{X_n \cdot X_{n+k}\} = (-0.94)^{|k|}$$

We want to code our signal using a 4-point DST-I and Lloyd-Max quantization of the transform components so that the average rate is 2 bits/sample.

Allocate bits to the four transform components such that the average distortion is minimized and calculate the resulting signal-to-noise ratio.

Compare your result to the signal-to-noise ratio that would have been achieved if the signal had been Lloyd-Max quantized directly to the rate 2 bits/sample, without the transform.

(6 p)