Written Exam in

Image and Audio Compression
TSBK38

14th August 2023 14:00 - 18:00

Location:	TER3
Examiner:	Harald Nautsch
Teacher:	Harald Nautsch, 1361
Department:	ISY
Module:	TEN1
Number of problems:	11
Number of pages:	5 + formula collection
Permitted equipment:	Calculator, "Tables and Formulas for Image Coding and Data Compression"
Grades:	3 : 15+ points from part I 4 : 15+ points from part I, 25-32 total points 5 : 15+ points from part I, 33-40 total points
Other:	Answers can be given in English or Swedish.

Exam structure

The exam is split into two parts, with maximum 20 points in each. In order to get a passing grade (3) you will need to get at least 15 out of 20 points from part I.

In addition, 25-32 total points gives grade 4 and 33-40 total points gives grade 5.

Part I

1 Describe the difference between *lossy* and *lossless* compression. In what situations would we prefer one type of compression over the other?

(3 p)

2 Make a comparison between using *Lloyd-Max quantization* and *uniform quantization*. Describe how the two quantization methods work and describe what advantages and disadvantages they have.

(2 p)

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

(2 p)

b) In most modern video coders, individual frames can be coded as either I, P or B frames. Explain the differences between the different types of frames and what type of information that needs to be transmitted for each frame.

(2 p)

4 a) When coding audio signals, a modified cosine transform (MDCT) is often used. Describe how the MDCT differs from a regular DCT and explain why the MDCT is a better choice for audio coding.

(1 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) Most music is stored in a stereo format, using two channels (left and right). Describe how the stereo information in an audio signal can be utilized to get a more efficient coding.

(1 p)

5 Describe how the coding works in each of these still image coding standards

a) PNG

(2 p)

b) JPEG

(2 p)

6 A memoryless source has the alphabet $\mathcal{A} = \{a, b, c, d, e, f\}$ with the symbol probabilities

$$P(a) = 0.46,\ P(b) = 0.19,\ P(c) = 0.12$$

$$P(d) = 0.10,\ P(e) = 0.09,\ P(f) = 0.04$$

a) What is theoretically lowest rate (in bits/symbol) we can get if we we want to code the output of the source without distortion?

(1 p)

b) Construct a Huffman code for the source and calculate the average rate (in bits/symbol) of the code.

(2 p)

Part II

7 Explain how Lempel-Ziv coding works. Describe both major variants (LZ77 and LZ78).

(2 p)

8 A stationary memoryless source has the alphabet $\mathcal{A} = \{a, b, c\}$. The symbol probabilities are $P(a) = 0.5$, $P(b) = 0.3$ and $P(c) = 0.2$.

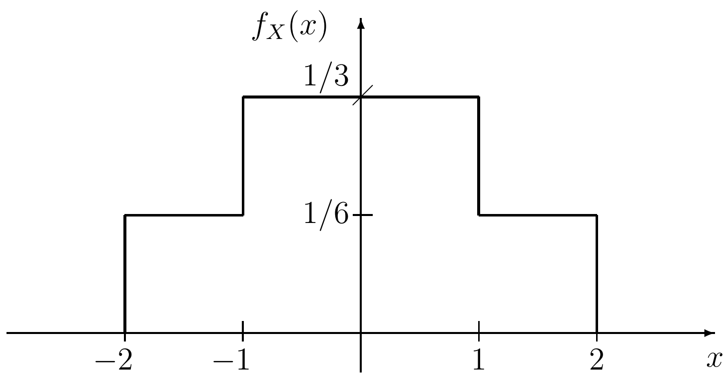
Code the sequence

$$baaac$$

using arithmetic coding. Give both the resulting interval and the corresponding codeword.

(4 p)

9 A stationary memoryless amplitude continuous and time discrete signal X_n has the probability density funcion $f_X(x)$ given by the following figure



X_n is quantized uniformly with the stepsize $\Delta = 2^{-k}$, where k is a non-negative integer. The quantized signal $\hat{X}_n$ is coded using a perfect source coder, ie the rate is given by $R = H(\hat{X}_n)$.

Give the rate of the coder as a function of the mean square error D.

(4 p)

10 An image is modelled as a two-dimensional stationary Gaussian signal $X_{i,j}$ with the following statistics:

$$E\{X_{i,j}\} = 0$$

$$E\{X_{i,j}^2\} = 8.70$$

$$E\{X_{i,j} \cdot X_{i-1,j}\} = E\{X_{i,j} \cdot X_{i,j-1}\} = 8.27$$

$$E\{X_{i,j} \cdot X_{i-1,j-1}\} = E\{X_{i,j} \cdot X_{i-1,j+1}\} = 8.09$$

$$E\{X_{i,j} \cdot X_{i-1,j-2}\} = E\{X_{i,j} \cdot X_{i-1,j+2}\} = 7.65$$

The signal should be coded using linear prediction. We can choose between the following two predictors:

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

$$q_{i,j} = b_1 \cdot \hat{X}_{i,j-1} + b_2 \cdot \hat{X}_{i-1,j+1}$$

Given that we choose the predictor coefficents so that the prediction error variance is minimized, which of the two predictors give the lowest distortion at a given rate?

Calculate the resulting signal-to-noise ratio for the best predictor, given that the prediction error is quantized using uniform quantization followed by arithmetic coding and where the quantiztion step is chosen such that the rate is 6 bits/sample.

State all assumptions and approximations that you make.

(5 p)

11 A sampled audio signal is modelled as a stationary time-discrete gaussian process X_n with mean zero and auto correlation function $R_{XX}(k) = E\{X_n \cdot X_{n+k}\}$. From a large set of data we have estimated the auto correlation function as

$$R_{XX}(0) = 3.50,\ R_{XX}(1) = 3.22,\ R_{XX}(2) = 2.94,\ R_{XX}(3) = 2.73$$

The signal is transform coded using a 4-point DCT and then Lloyd-Max quantized so that the average rate is 1.75 bits/sample.

Allocate bits to the quantizers of the different transform components so that the average distortion is minimized and calculate the resulting signal to noise rate (in dB).

(5 p)