LINKÖPING UNIVERSITY

Written Exam in
Image and Audio Compression
TSBK38

26th March 2025 14:00 - 18:00

Location:	SP71
Examiner:	Harald Nautsch
Teacher:	Harald Nautsch, 1361
Department:	ISY
Module:	TEN1
Number of problems:	12
Number of pages:	5+formula collection
Permitted equipment:	Calculator, “Tables and Formulas for Image Coding and Data Compression”
Grades:	3 : 14+ points from part I 4 : 14+ points from part I, 24-31 total points 5 : 14+ points from part I, 32-40 total points
Other:	Answers can be given in English or Swedish. The teacher visits the exam room around 15.30.

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 14 out of 20 points from part I.

In addition, 24-31 total points gives grade 4 and 32-40 total points gives grade 5.

Part I

1 What are the properties of a discrete alphabet data sequence that makes it possible to compress it losslessly?

(2 p)

2 What is the purpose of the *LBG algorithm*? Also describe how it works.

(2 p)

3 When coding colour still images and video signals, the colour space used is usually YCbCr instead of RGB. Explain how these colour spaces differ and why the YCbCr colour space is preferred.

(2 p)

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

(4 p)

5 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) When coding general audio signals, a *psychoacoustic model* is often used to get a more efficient coding. Explain how the model works.

(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)

6 Two common methods for still image coding are PNG and JPEG. Explain how they work and in what situations you might prefer to use one method over the other.

(4 p)

7 A memoryless source has the alphabet $\mathcal{A} = \{a, b, c\}$ with the symbol probabilities $P(a) = 0.5,\ P(b) = 0.4$ and $P(c) = 0.1$.

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

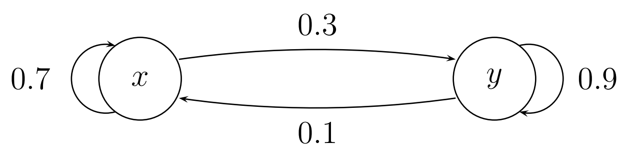
(2 p)

Part II

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

(2 p)

9 The following Markov source is given



Show how arithmetic coding works by coding the sequence $xyyyx$. You should take advantage of the memory between symbols. You can assume that the source is in state x when the coding starts. Give both the resulting intervall and the codeword.

(4 p)

10 A memoryless time discrete random signal with probability density function

$$f_X(x) = \begin{cases} 1-x & ; \quad 0 \leq x \leq 1 \\ 1+x & ; \quad -1 \leq x \leq 0 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

is quantized to four levels. The decision borders are chosen as

$$b_0 = -1, \quad b_1 = -0.5, \quad b_2 = 0, \quad b_3 = 0.5, \quad b_4 = 1$$

a) What distortion do we get if the reconstruction points are placed in the middle of each interval, ie

$$y_1 = -0.75, \quad y_2 = -0.25, \quad y_3 = 0.25, \quad y_4 = 0.75$$

(2 p)

b) What distortion do we get if the reconstruction points are placed such that the distortion is minimized?

(2 p)

11 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

$$R_{XX}(0,0) = 3.40, \quad R_{XX}(0,1) = 3.16$$

$$R_{XX}(1,0) = 3.08, \quad R_{XX}(1,1) = R_{XX}(1,-1) = 3.04$$

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 uniformely and then coded using an 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 38 db?

(5 p)

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

$$R_{XX}(0) = 4.31, \quad R_{XX}(1) = 4.01, \quad R_{XX}(2) = 3.88, \quad R_{XX}(3) = 3.75$$

The signal is transformed using a 4-point DCT and then Lloyd-Max quantized so that the average rate is 2 bits/sample.

Distribute 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)