LINKÖPING UNIVERSITY

Written Exam in

Image and Audio Compression TSBK38

18th August 2025   14:00 - 18:00

Location:	TER1
Examiner:	Harald Nautsch
Teacher:	Harald Nautsch, 1361
Department:	ISY
Module:	TEN1
Number of problems:	13
Number of pages:	7+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 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)

2 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)

3 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)

4 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) Briefly describe what improvements that have been made to make a modern hybrid coder (eg HEVC) much better than an old hybrid coding standard (eg MPEG2).

(1 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 A memoryless source has the alphabet $\mathcal{A} = \{a, b, c\}$ with the symbol probabilities $P(a) = 0.6,\ P(b) = 0.3$ and $P(c) = 0.1$.

Construct a Huffman code for the source that gives an average rate of at most 1.35 bits/symbol.

(2 p)

7 Given a gaussian input signal with zero mean and variance σ^2. We want to quantize and code the signal such that the resulting rate is 10 bits/sample.

a) First we try a coder where we use uniform quantization and fixed length coding. The quantizer stepsize is chosen such that the working range of the quantizer becomes $\pm 5\sigma$. What is the resulting signal-to-noise ratio (in dB)?

(2 p)

b) Next we add a source coder to our coding setup. The source coder can be assumed to be perfect. The stepsize is adjusted such that the rate after source coding is still 10 bits/sample. What is the resulting signal-to-noise ratio (in dB)?

(1 p)

Part II

8 Explain how arithmetic coding works.

(2 p)

9 A source has the alphabet $\{a, b, c, d, e, f, g, h\}$. A sequence from the source is coded using LZW and gives the following index sequence:

$$1, 0, 7, 8, 10, 9, 11, 14, 9, 0, 3, 18, 3, 4, \ldots$$

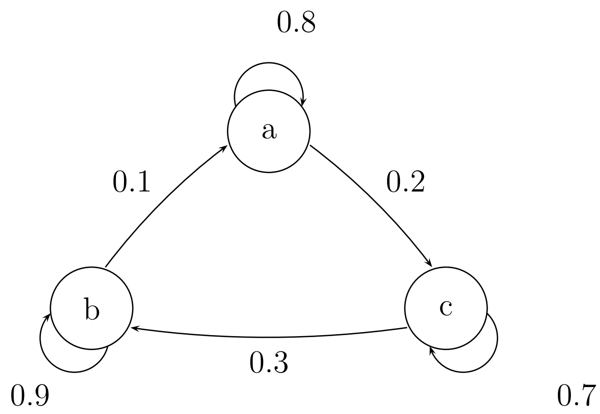
The starting dictionary is:

index	sequence	index	sequence
0	a	4	e
1	b	5	f
2	c	6	g
3	d	7	h

Decode the index sequence. Also give the resulting dictionary.

(3 p)

10 A Markov source of order 1 X_n with alphabet $\{a, b, c\}$ is given by the following state diagram



What is, on average, the lowest possible rate (in bits/symbol) that the output from the source can be coded to?

(3 p)

11 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 five levels. The decision borders are chosen as

$$b_0 = -1, \quad b_1 = -0.6, \quad b_2 = -0.2, \quad b_3 = 0.2, \quad b_4 = 0.6, \quad b_5 = 1$$

How should we place the recconstruction points in order to minimize the distortion? Also calculate this distortion.

(3 p)

12 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 quantization step is chosen such that the rate is 6 bits/sample.

State all assumptions and approximations that you make.

(5 p)

13 In the video coding standard H.264 the following transform is used

$$\mathbf{A} = \begin{pmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 2/\sqrt{10} & 1/\sqrt{10} & -1/\sqrt{10} & -2/\sqrt{10} \\ 1/2 & -1/2 & -1/2 & 1/2 \\ 1/\sqrt{10} & -2/\sqrt{10} & 2/\sqrt{10} & -1/\sqrt{10} \end{pmatrix}$$

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

$$R_{XX}(k) = E\{X_i \cdot X_{i+k}\} = 0.94^{|k|}$$

We want to quantize the transform components using Lloyd-Max quantizers such that the average rate is 1.75 bits/sample and the average distortion is minimized.

How should the bits be allocated among the transform components and what is the resulting signal to noise ratio (in dB)?

(4 p)