LINKÖPING UNIVERSITY

Written Exam in
Image and Audio Coding
TSBK02

14th August 2017 14:00 - 18:00

Location:	TER1
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 15:15 and 16.45

1 Explain how the coding works in each of these still image coding standards

a) GIF

(1 p)

b) PNG

(1 p)

c) JPEG

(1 p)

d) JPEG-2000

(1 p)

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

(2 p)

b) Draw a block schedule over a typical hybrid coder for video signals and explain how the parts work.

(2 p)

c) Explain what the LBG algorithm is and how it works.

(2 p)

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

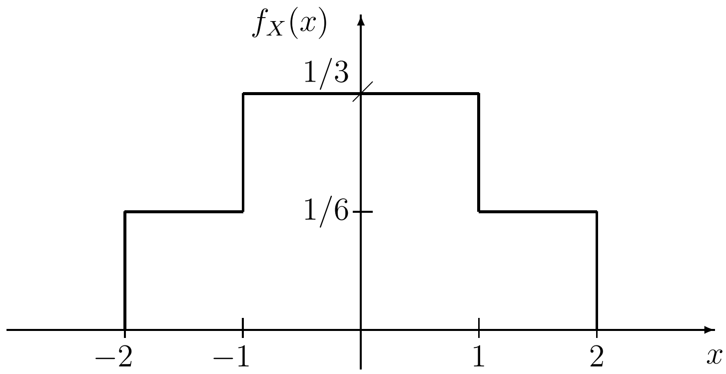
Code the sequence

$$aaabb$$

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

(4 p)

4 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)

5 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)?

(6 p)

6 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 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 36 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 36 dB.

(6 p)