LINKÖPING UNIVERSITY

Written Exam in
Image and Audio Coding
TSBK02

1st June 2017 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 is available on phone until 11.30

1 a) Describe how the following quantization methods work and what advantages and disadvantages the different methods have.

- Uniform quantization
- Lloyd-Max quantization
- Compander quantization

(3 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) Describe 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 probabilities $P(a) = 0.1,\ P(b) = 0.8$ and $P(c) = 0.1$.

a) What is the theoretically lowest rate we can get when coding the source losslessly?

(1 p)

b) Construct a Huffman code for the source that gives an average rate of no more than 1 bit/symbol.

(2 p)

4 A memoryless time discrete process with probability density function

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

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)

5 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 $R_{XX}(k) = E\{X_n X_{n+k}\}$.

$$R_{XX}(0) = 6.96, \quad R_{XX}(1) = 3.51, \quad R_{XX}(2) = -2.58, \quad R_{XX}(3) = -5.34$$

Find the linear predictor of order two that minimizes the prediction error variance. Calculate the prediction error variance.

Assuming that the prediction error is quantized with a scalar quantizer and that the quantized prediction error is coded with a memoryless source coder, what is the lowest rate that we can use if we want a signal-to-noise ratio of at least 37 dB?

(6 p)

6 We want to code still images using a simple transform coder.

By examining a large number of images, it has been shown that they can be modelled as a twodimensional stationary gaussian process $X_{i,j}$, with the following statstics

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

$$R_{XX}(k,l) = E\{X_{i,j} \cdot X_{i+k,j+l}\}$$

$$R_{XX}(0,0) = 760, \quad R_{XX}(1,0) = 492$$

$$R_{XX}(0,1) = 684, \quad R_{XX}(1,1) = R_{XX}(1,-1) = 425$$

We want to do transform coding by taking small blocks consisting of 2×2 pixels and performing a separable discrete Walsh-Hadamard transform of the blocks.

The transform components are Lloyd-Max quantized such that the average rate is 2.25 bits/pixel.

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

(6 p)