LINKÖPING UNIVERSITY

Written Exam in
Image and Audio Coding
TSBK02

20th August 2018 14:00 - 18:00

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 15:15 and 16.45

1 a) Describe how the following quantization methods work:

- Lloyd-Max quantization
- compander quantization

(2 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 mp3 coding of music signals works.

(2 p)

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

(2 p)

2 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.

(1 p)

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

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

(1 p)

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

(2 p)

4 a) A memoryless time discrete gaussian signal with variance σ^2 and zero mean is quantized using a 12 bit uniform quantizer. The edges of the quantizer are set at $\pm 5\sigma$. The quantized signal is coded using a fixed length code.

What is the resulting signal-to-noise ratio (in dB)?

(2 p)

b) What is the theoretically highest signal-to-noise ratio we can get when coding the signal in problem a (at the same rate) if we are free to choose any types of coding methods?

(2 p)

5 A speech signal is modelled as a one-dimensional stationary gaussian process Y_n. The signal statistics have been estimated as:

$$E\{Y_n\} = 0$$

$$R_{YY}(k) = E\{Y_n \cdot Y_{n+k}\}$$

$$R_{YY}(0) = 9.07\ , \quad R_{YY}(1) = 7.56\ , \quad R_{YY}(2) = 4.09 \quad R_{YY}(3) = 0.08$$

We want to code the signal with no more than 5 bits per sample (on average) and a signal to noise ratio that is at least 35 dB.

Construct a predictive coder that fulfills the requirements. All assumptions and simplifications must be motivated.

(6 p)

6 A three channel audio signal consists of a left channel L_i, a right channel R_i and a center channel C_i. The following mean values have been measured:

$$E\{L_i\} = E\{R_i\} = E\{C_i\} = 0$$

$$E\{L_i^2\} = E\{R_i^2\} = E\{C_i^2\} = 1$$

$$E\{L_i R_i\} = 0.9 \; , \quad E\{L_i C_i\} = E\{R_i C_i\} = 0.93$$

The three channels can be modelled as a three-dimensional normal distribution.

The signal is coded by forming vectors $[L_i \;\; C_i \;\; R_i]^T$ that are transformed using a three point DCT. The transform components are Lloyd-Max quantized such that the average rate is 2 bits/sample/channel.

Allocate bits to the transform components such that that average distortion for the three channels is minimized and calculate this distortion.

What average distortion would we get if we instead did Lloyd-Max quantization of the three channels directly to 2 bits each, without using a transform?

(6 p)