LINKÖPING UNIVERSITY

Written Exam in
Image and Audio Compression
TSBK38

29th May 2023 8:00 - 12:00

Location:	FE249
Examiner:	Harald Nautsch
Teacher:	Harald Nautsch, 1361
Department:	ISY
Module:	TEN1
Number of problems:	11
Number of pages:	5 + formula collection
Permitted equipment:	Calculator, “Tables and Formulas for Image Coding and Data Compression”
Grades:	3 : 15+ points from part I 4 : 15+ points from part I, 25-32 total points 5 : 15+ points from part I, 33-40 total points
Other:	Answers can be given in English or Swedish.

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 15 out of 20 points from part I.

In addition, 25-32 total points gives grade 4 and 33-40 total points gives grade 5.

Part I

1 Describe the difference between *lossy* and *lossless* compression. In what situations would we prefer one type of compression over the other?

(3 p)

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

(2 p)

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

(4 p)

4 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)

5 Describe how the coding works in each of these still image coding standards

a) GIF

(2 p)

b) PNG

(2 p)

c) JPEG

(2 p)

6 A memoryless source has the alphabet $\mathcal{A} = \{a, b, c, d, e, f\}$ with the symbol probabilities

$$P(a) = 0.48, \ P(b) = 0.18, \ P(c) = 0.11$$

$$P(d) = 0.10, \ P(e) = 0.09, \ P(f) = 0.04$$

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

(1 p)

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

(2 p)

Part II

7 When coding speech signals, a relatively simple model of how human speech is generated is often used. Describe this model and how it can be used in the coding and decoding process.

(3 p)

8 A source has the alphabet $\{p,\ r,\ s,\ t,\ u,\ v\}$. A long sequence of symbols from the source is coded using LZW. The resulting index sequence starts as

$$0,\ 4,\ 1,\ 6,\ 8,\ 7,\ 9,\ 12,\ 11,\ 14,\ 10,\ 16,\ 3,\ \ldots$$

The starting dictionary is:

index	sequence	index	sequence
0	p	3	t
1	r	4	u
2	s	5	v

Decode the given index sequence as far as possible. Also give the resulting dictionary.

(3 p)

9 A random variable X with probability density function

$$f_X(x) = \begin{cases} \frac{3}{2}(1-x)^2 & ;\ 0 \leq x \leq 1 \\ \frac{3}{2}(1+x)^2 & ;\ -1 \leq x < 0 \\ 0 & ;\ \text{otherwise} \end{cases}$$

is quantized to two levels.

Find the decision borders and reconstruction points such that the resulting distortion is minimized.

Calculate the resulting distortion.

(4 p)

10 An image is modeled as a twodimensional gaussian process $Z_{i,j}$ (i and j are image coordinates) with the following statistics

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

$$E\{Z_{i,j} \cdot Z_{k,l}\} = 29 \cdot 0.91^{|i-k|+0.5\cdot|j-l|}$$

Construct a predictive coder for the image that gives an average rate of no more than 5 bits/pixel and a signal to noise ratio of at least 40 dB.

(5 p)

11 In the HEVC video coding standard, the following approximation of a discrete cosine transform is used

$$\mathbf{A} = \frac{1}{128}\begin{pmatrix} 64 & 64 & 64 & 64 \\ 83 & 36 & -36 & -83 \\ 64 & -64 & -64 & 64 \\ 36 & -83 & 83 & -36 \end{pmatrix}$$

Assume we want to code a signal X_n using this transform. X_n is modelled as a zero mean stationary Gaussian process with auto correlation function

$$R_{XX}(k) = E\{X_n X_{n+k}\} = 0.94^{|k|}$$

We want to quantize the transform coefficients using Lloyd-Max quantization so that the average rate is 1.75 bits/sample and the distortion is minimized.

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

(NOTE: As can be seen, the transform is not fully normalized. However, the lengths of the basis vectors are close enough to 1 that this can be ignored when allocating bits and calculating the distortion.)

(5 p)