LINKÖPING UNIVERSITY

Solutions to Written Exam in
Image and Audio Compression
TSBK38

14th August 2023

1 See the course literature.

2 See the course literature.

3
a) See the course literature.
b) See the course literature.

4
a) See the course literature.
b) See the course literature.
c) See the course literature.

5
a) See the course literature.
b) See the course literature.

6
a) The theoretical limit is given by entropy rate of the source, which for a memoryless source is given by the entropy of single symbols, in this case approximately 2.1682 bits/symbol.

b) One Huffman code (there are other) is given by

symbol	codeword	codeword length
a	0	1
b	100	3
c	110	3
d	111	3
e	1010	4
f	1011	4

The code has a rate of 2.21 bits/symbol.

7 See the course literature.

8 The interval corresponding to the sequence is $[0.53 \;\; 0.5375)$.

We need at least $\lceil -\log 0.0075 \rceil$=8 bits in the codeword, maybe one more.

Write the two interval limits in base 2:

$$\begin{aligned} 0.53 &= 0.1000011110101\ldots \\ 0.5375 &= 0.1000100110011\ldots \end{aligned}$$

The smallest 8 bit number inside this interval is 0.10001000. All numbers starting with these bits are also inside the interval (ie they are smaller than the upper limit), which means that 8 bits are enough.

The codeword is **10001000**.

9 The number of levels in the quantizer is

$$M = \frac{4}{\Delta} = 4 \cdot 2^k = 2^{k+2}$$

We get $M/2$ symbols with probability $\Delta/3$ and $M/2$ symbols with probability $\Delta/6$. The rate will then be

$$\begin{aligned} R &= H(\hat{X}) = \frac{M}{2}\left(-\frac{\Delta}{3}\log\frac{\Delta}{3} - \frac{\Delta}{6}\log\frac{\Delta}{6}\right) = \\ &= \frac{2}{3}\log(3 \cdot 2^k) + \frac{1}{3}\log(6 \cdot 2^k) = \log 3 + \frac{1}{3} + k \end{aligned}$$

Since the distribution is constant in each quantization interval, the distortion is given by

$$D = \frac{\Delta^2}{12} = \frac{2^{-2k}}{12} \implies k = -\frac{1}{2}\log 12D$$

which finally gives us the rate R as a function of the distortion D

$$R = \log 3 + \frac{1}{3} - \frac{1}{2}\log 12D = \frac{1}{3} + \frac{1}{2}\log\frac{3}{4D}$$

10 Fine quantization means we can use the approximation $\hat{X}_{i,j} \approx X_{i,j}$. The prediction error is approximately gaussian.

We assume that the arithmetic coder gives a rate equal to the entropy of the quantized prediction error.

Optimal choices of a_1 and a_2 are given by

$$\begin{pmatrix} 8.70 & 8.09 \\ 8.09 & 8.70 \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 8.27 \\ 8.27 \end{pmatrix} \;\Rightarrow\; \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \approx \begin{pmatrix} 0.4926 \\ 0.4926 \end{pmatrix}$$

Resulting prediction error variance

$$\sigma_p^2 \approx 0.5531$$

Optimal choices of b_1 and b_2 are given by

$$\begin{pmatrix} 8.70 & 7.65 \\ 7.65 & 8.70 \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 8.27 \\ 8.09 \end{pmatrix} \;\Rightarrow\; \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \approx \begin{pmatrix} 0.5860 \\ 0.4146 \end{pmatrix}$$

Resulting prediction error variance

$$\sigma_q^2 \approx 0.4996$$

Since the second predictor gives a lower prediction error variance, it will give a lower distortion at the given rate.

The resulting distortion is given by

$$D \approx \frac{\pi e}{6} \cdot \sigma_q^2 \cdot 2^{-2\cdot 6} \approx 1.7359 \cdot 10^{-4}$$

and signal-to-noise ratio

$$10 \cdot \log_{10}\frac{8.70}{D} \approx 47.0 \text{ dB}$$

11 The transform matrix for a 4-point DCT is

$$\mathbf{A} = \begin{pmatrix} 0.5 & 0.5 & 0.5 & 0.5 \\ a & b & -b & -a \\ 0.5 & -0.5 & -0.5 & 0.5 \\ b & -a & a & -b \end{pmatrix}$$

where $a = \frac{1}{\sqrt{2}} \cos \frac{\pi}{8} \approx 0.6533$ and $b = \frac{1}{\sqrt{2}} \cos \frac{3\pi}{8} \approx 0.2706$.

The variances for the four transform components are

$$\sigma_0^2 \approx 12.635,\ \sigma_1^2 \approx 0.8962,\ \sigma_2^2 \approx 0.3150,\ \sigma_3^2 \approx 0.1538$$

To achieve the average rate 1.75 we should allocate $4 \cdot 1.75 = 7$ bits in total.

The allocation that minimizes the average distortion is $R_0 = 4, R_1 = 2, R_2 = 1, R_3 = 0$. The resulting average distortion is

$$D \approx \frac{1}{4}(0.09497 \cdot 12.635 + 0.1175 \cdot 0.8962 + 0.3634 \cdot 0.3150 + 0.1538) \approx 0.1234$$

The signal to noise ratio is

$$\mathrm{SNR} = 10 \cdot \log_{10} \frac{\sigma_X^2}{D} \approx 10 \cdot \log_{10} \frac{3.5}{0.1234} \approx 14.5 \text{ [dB]}$$