LINKÖPING UNIVERSITY

Solutions to Written Exam in
Image and Audio Compression
TSBK38

26th March 2025

1 Some symbols (or partial sequences of symbols) will be more common than others. If we have a random model for the source this corresponds to having different probabilities. By having short codewords for common symbols and long codewords for uncommon symbols we can get a lower data rate than if we used the same length codewords for all symbols.

Most sources have some kind of memory (dependence between neighbouring symbols in the sequence). This can be used to achieve a lower data rate than if we didn't take the memory into account.

2 See the course literature.

3 See the course literature.

4 See the course literature.

5 a) See the course literature.

b) See the course literature.

c) See the course literature.

6 See the course literature.

7 One Huffman code (there are several other giving the same rate) for pairs is given by

symbols	codeword	codeword length
aa	00	2
ab	10	2
ac	01000	5
ba	11	2
bb	011	3
bc	01010	5
ca	01011	5
cb	010010	6
cc	010011	6

The code has a mean codeword length of 2.78 bits/codeword and a rate of 1.39 bits/symbol.

8 See the course literature.

9 If we assume that the x intervall is always closest to 0, the intervall corresponding to the sequence $xyyyx$ is then $[0.5299\ 0.54691)$. If you ordered your symbols differently, you should at least get an intervall of the same size (0.01701).

We need at least $\lceil -\log_2 0.01701 \rceil = 6$ bits to describe this intervall. If we write the limits as binary numbers we get

$$\begin{aligned} 0.5299 &= 0.100001111010\ldots \\ 0.54691 &= 0.100011000000\ldots \end{aligned}$$

The smallest binary number with six bits in this interval is 0.100010 We can see that six bits will be enough (all numbers starting with these six bits are also inside the interval). The codeword is thus **100010**.

10 a) The distortion is given by

$$\begin{aligned} D &= \sum_{i=1}^{4} \int_{b_{i-1}}^{b_i} (x - y_i)^2 f_X(x)\ dx = \\ &= 2\left(\int_0^{0.5} (x - 0.25)^2 (1 - x)\ dx + \int_{0.5}^{1} (x - 0.75)^2 (1 - x)\ dx \right) = \frac{1}{48} \end{aligned}$$

b) To minimize the distortion, the reconstruction points should be placed in the center of probability mass in each interval, ie

$$y_3 = \frac{\int_0^{0.5} x(1-x)\ dx}{\int_0^{0.5} (1-x)\ dx} = \frac{2}{9}$$

$$y_4 = \frac{\int_{0.5}^{1} x(1-x)\ dx}{\int_{0.5}^{1} (1-x)\ dx} = \frac{2}{3}$$

$$y_1 = -y_4, \quad y_2 = -y_3$$

The distortion is then given by

$$D = 2\left(\int_0^{0.5} \left(x - \frac{2}{9}\right)^2 (1-x)\ dx + \int_{0.5}^{1} \left(x - \frac{2}{3}\right)^2 (1-x)\ dx \right) = \frac{1}{54}$$

11 We assume that the quantization is fine enough so that we can do the calculations as if the predictor worked using the original signal, ie we can disregard the effect of the quantization on the prediction. The variance of the prediction error:

$$\sigma_d^2 = E\{(X_{i,j} - p_{i,j})^2\} \approx E\{(X_{i,j} - a_1 X_{i-1,j} - a_2 X_{i,j-1})^2\}$$

a_1 and a_2 that minimize σ_d^2 are given by

$$\begin{pmatrix} 3.40 & 3.04 \\ 3.04 & 3.40 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 3.08 \\ 3.16 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \approx \begin{pmatrix} 0.3734 \\ 0.5956 \end{pmatrix}$$

$$\Rightarrow \sigma_d^2 \approx 0.3680$$

Uniform quantization followed by entropy coding gives the approximate distortion

$$D \approx \frac{\pi e}{6} \cdot \sigma_d^2 \cdot 2^{-2 \cdot R}$$

The signal-to-noise ratio is given by

$$\text{SNR} = 10 \cdot \log_{10} \frac{\sigma_X^2}{D}$$

If we want an SNR of at least 38 dB, we must choose R so that

$$D \leq \frac{\sigma_X^2}{10^{3.8}} \approx 0.00053886$$

which gives us the smallest possible rate as approximately 4.96 bits/pixel.

12 The transform matrix for a 4-point DCT is

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

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

$$\sigma_0^2 = 16.08,\ \sigma_1^2 \approx 0.6138,\ \sigma_2^2 = 0.3,\ \sigma_3^2 \approx 0.2462$$

To achieve the average rate 2 we should allocate $4 \cdot 2 = 8$ bits. The distribution that minimizes the average distortion is $R_0 = 4, R_1 = 2, R_2 = 1, R_3 = 1$. The resulting average distortion is

$$D \approx \frac{1}{4}(0.009497 \cdot \sigma_0^2 + 0.1175 \cdot \sigma_1^2 + 0.3634 \cdot \sigma_2^2 + 0.3634 \cdot \sigma_3^2) \approx 0.1058$$

The signal to noise ratio is

$$\text{SNR} = 10 \cdot \log_{10} \frac{\sigma_X^2}{D} \approx 10 \cdot \log_{10} \frac{4.31}{0.1058} \approx 16.1 \text{ [dB]}$$