LINKÖPING UNIVERSITY

Solutions to Written Exam in
Image and Audio Compression
TSBK38

27th May 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 a) See the course literature.

b) See the course literature.

5 a) See the course literature.

b) See the course literature.

6 The entropy rate of the source is approximately 1.2362 bits, so it is possible.

A Huffman code for single symbols will give the rate 1.35 bits/symbol and is therefore not enough. We need to code at least two symbols at a time.

One Huffman code (there are others) for pairs is given by

symbols	codeword	codeword length
aa	0	1
ab	100	3
ac	1100	4
ba	101	3
bb	1101	4
bc	11110	5
ca	1110	4
cb	111110	6
cc	111111	6

The code has a mean codeword length of 2.5025 bits/codeword and a rate of 1.25125 bits/symbol.

7 a) Number of quantization levels

$$M = 2^8 = 256$$

The step length of the quantizer is

$$\Delta = \frac{8\sigma}{M}$$

Distortion approximation

$$D \approx \frac{\Delta^2}{12} = \frac{64\sigma^2}{12M^2} = \frac{16\sigma^2}{3M^2}$$

Signal-to-noise ratio

$$10\log_{10}\frac{\sigma^2}{D} \approx 10\log_{10}\frac{3M^2}{16} = 10\log_{10}12288 \approx 40.9 \quad \text{[dB]}$$

b) The number of levels is high enough that we can use the Lloyd-Max distortion approximation

$$D \approx \frac{\pi\cdot\sqrt{3}}{2}\cdot\sigma^2\cdot 2^{-2\cdot 8}$$

which gives the SNR

$$10\log_{10}\frac{\sigma^2}{D} \approx 10\log_{10}\frac{2^{17}}{\pi\sqrt{3}} \approx 43.8 \quad \text{[dB]}$$

8 See the course literature.

9 The interval corresponding to the sequence is $[0.18792 \quad 0.1944)$.

We need at least $\lceil -\log 0.00648 \rceil$=8 bits in the codeword, maybe one more.

Write the two interval limits in base 2:

$$\begin{aligned} 0.16848 &= 0.001100000001\ldots \\ 0.18792 &= 0.001100011100\ldots \end{aligned}$$

The smallest 8 bit number inside this interval is 0.00110001. However, there are nubers starting with these bits hat are outside of the interval (ie they are larger than the upper limit), which means that we need 9 bits in the codeword

The codeword can then be either **001100001** (if we chose the smallest number) or **001100010**.

10 The distribution is a triangle distribution, symmetric around the origin. This means that the three reconstruction levels are located symmetrically, ie $y_1 = -y_3$, $y_2 = 0$ and y_3. The decision borders should be halfway between the reconstruction levels, ie $b_0 = -2$, $b_1 = -b_2$, $b_2 = \frac{y_3}{2}$ and $b_3 = 2$. Thus, we only have to find one unknown value.

The reconstruction levels should be in the center of probability mass of each area, which gives us the equation.

$$y_3 = \frac{\int_{b_2}^{2} x \cdot f_X(x)dx}{\int_{b_2}^{2} f_X(x)dx} = \frac{(4 - 3b_2^2 + b_2^3)/12}{(4 - 4b_2 + b_2^2)/8} = \frac{2(b_2 - 2)^2(b_2 + 1)}{3(b_2 - 2)^2}$$

We have $y_3 = 2b_2$ and we can also assume that $b_2 < 2$ (ie b_2 is not the same as b_3). We thus arrive at the equation

$$2b_2 = \frac{2}{3}(b_2 + 1) \;\Rightarrow\; b_2 = \frac{1}{2}$$

Thus, the decision borders are -2, $-\frac{1}{2}$, $\frac{1}{2}$ and 2. The reconstruction levels are -1, 0 and 1.

The distortion is given by

$$D = 2\left(\int_0^{1/2} x^2(\frac{1}{2} - \frac{x}{4})dx + \int_{1/2}^{2} (x-1)^2(\frac{1}{2} - \frac{x}{4})dx\right) = \frac{5}{48}$$

11 We assume that the quantization is fine enough so that we can do the calculations as if the predictor is 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

$$\begin{pmatrix} 1104 & 988 \\ 988 & 1104 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1001 \\ 1027 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \approx \begin{pmatrix} 0.3726 \\ 0.5968 \end{pmatrix}$$

$$\Rightarrow \sigma_d^2 \approx 118.11$$

Uniform quantization followed by entropy coding to the rate R gives the approximate distortion

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

or alternatively, with rate as a function of the distortion

$$R \approx \frac{1}{2} \log_2 \frac{\pi e \sigma_d^2}{6D}$$

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 43 dB, we must choose R so that

$$D \leq \frac{\sigma_X^2}{10^{4.3}} \approx 0.0553$$

which gives us the smallest possible rate as approximately 5.78 bits/pixel.

If we didn't use the predictor, we would have needed to use the rate

$$R \approx \frac{1}{2} \log_2 \frac{\pi e \sigma_X^2}{6D} \approx 7.40$$

to reach 43 dB.

12 The transform matrix $\mathbf{A}$ can be written as

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

where $a = \frac{2}{\sqrt{7}} \sin \frac{\pi}{7}$, $b = \frac{2}{\sqrt{7}} \sin \frac{2\pi}{7}$ and $c = \frac{2}{\sqrt{7}} \sin \frac{3\pi}{7}$. We also have $a^2 + b^2 + c^2 = 1$. Since X_n is gaussian, the transform components will also be gaussian. The transform component variances are given by

$$\begin{aligned}
\sigma_0^2 &= E\{\theta_0^2\} = E\{(aX_0 + bX_1 + cX_2)^2\} = \\
&= R_{XX}(0) + 2b(a+c)R_{XX}(1) + 2acR_{XX}(2) \approx 10.7036 \\
\sigma_1^2 &= E\{\theta_1^2\} = E\{(cX_0 + aX_1 - bX_2)^2\} = \\
&= R_{XX}(0) + 2a(c-b)R_{XX}(1) - 2bcR_{XX}(2) \approx 1.0930 \\
\sigma_2^2 &= E\{\theta_2^2\} = E\{(bX_0 - cX_1 + aX_2)^2\} = \\
&= R_{XX}(0) - 2c(a+b)R_{XX}(1) + 2abR_{XX}(2) \approx 0.2034
\end{aligned}$$

(NOTE: The average value of the variances is equal to the signal variance $\sigma_X^2 = R_{XX}(0) = 4$ since the transform is orthonormal.)

Alternatively we can calculate the variances as the diagonal elements of $\mathbf{R}_\theta = \mathbf{A} \cdot \mathbf{R}_X \cdot \mathbf{A}^T$, where

$$\mathbf{R}_X = 4 \cdot \begin{pmatrix} 1 & 0.97 & 0.9409 \\ 0.97 & 1 & 0.97 \\ 0.9409 & 0.97 & 1 \end{pmatrix}$$

The desired average rate is 2 bits/sample, ie

$$R = \frac{R_0 + R_1 + R_2}{3} = 2 \;\Rightarrow\; R_0 + R_1 + R_2 = 6$$

The rate allocation that minimizes the distortion is given by $R_0 = 4, R_1 = 2, R_2 = 0$ which gives the distortion

$$D \approx (0.009497 \cdot \sigma_0^2 + 0.1175 \cdot \sigma_1^2 + \sigma_2^2)/3 \approx 0.1445$$

The resulting signal to noise ratio is

$$\text{SNR} = 10 \cdot \log_{10} \frac{\sigma_X^2}{D} \approx 10 \cdot \log_{10} \frac{4}{0.1445} \approx 14.4 \text{ [dB]}$$