LINKÖPING UNIVERSITY

Solutions to Written Exam in
Image and Audio Compression
TSBK38

18th August 2025

1 See the course literature.

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.
  c) See the course literature.

6 A Huffman code for single symbols will give the rate 1.4 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	1110	4
bc	11110	5
ca	1101	4
cb	111110	6
cc	111111	6

The code has a mean codeword length of 2.67 bits/codeword and a rate of 1.335 bits/symbol.

7 a) Number of quantization levels

$$M = 2^{10} = 1024$$

The step length of the quantizer is

$$\Delta = \frac{10\sigma}{M}$$

Distortion approximation

$$D \approx \frac{\Delta^2}{12} = \frac{100\sigma^2}{12M^2} = \frac{25\sigma^2}{3M^2}$$

Signal-to-noise ratio

$$10\log_{10}\frac{\sigma^2}{D} \approx 10\log_{10}\frac{3M^2}{25} = 10\log_{10} 125829.12 \approx 51.0 \quad [\text{dB}]$$

b) The number of levels is high enough that we can use the fine quantization approximation

$$D \approx \frac{\pi e}{6} \cdot \sigma^2 \cdot 2^{-2\cdot 10}$$

which gives the SNR

$$10\log_{10}\frac{\sigma^2}{D} \approx 10\log_{10}\frac{6 \cdot 2^{20}}{\pi e} \approx 58.7 \quad [\text{dB}]$$

8 See the course literature.

9 The decoded sequence is

$$bahbahbahbahbahbahadddde\ldots$$

and the dictionary looks like

index	word	index	word	index	word
0	a	8	ba	16	aha
1	b	9	ah	17	ad
2	c	10	hb	18	dd
3	d	11	bah	19	ddd
4	e	12	hba	20	de
5	f	13	ahb	21	$e*$
6	g	14	$bahb$	22	
7	h	15	$bahba$	23	

where $*$ in word 21 will be the first symbol in the next decoded word.

10 For this markov source of order 1, the lowest possible rate is given by its entropy rate $H(X_n|X_{n-1})$.

The stationary distribution $\bar{w}$ is given by

$$\bar{w} = (w_a \ w_b \ w_c) = \bar{w} \begin{pmatrix} 0.8 & 0 & 0.2 \\ 0.1 & 0.9 & 0 \\ 0 & 0.3 & 0.7 \end{pmatrix}$$

Substituting one of the equations with $w_a+w_b+w_c = 1$ and solving, we get

$$\bar{w} = \frac{1}{11} \, (3 \ \ 6 \ \ 2)$$

and thus

$$H(X_n|X_{n-1}) = \frac{3}{11} \cdot H_b(0.8) + \frac{6}{11} \cdot H_b(0.9) + \frac{2}{11} \cdot H_b(0.7) \approx 0.6129 \text{ [bits/symbol]}$$

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

$$y_3 = 0$$

$$y_4 = \frac{\int_{0.2}^{0.6} x(1-x)\,dx}{\int_{0.2}^{0.6} (1-x)\,dx} = \frac{17}{45}$$

$$y_5 = \frac{\int_{0.6}^{1} x(1-x)\ dx}{\int_{0.6}^{1} (1-x)\ dx} = \frac{11}{15}$$

$$y_1 = -y_5, \quad y_2 = -y_4$$

The distortion is then given by

$$\begin{aligned} D &= 2\left(\int_0^{0.2} x^2(1-x)\ dx + \int_{0.2}^{0.6}\left(x-\frac{17}{45}\right)^2 (1-x)\ dx + \int_{0.6}^{1}\left(x-\frac{11}{15}\right)^2 (1-x)\ dx\right) \\ &= \frac{409}{3759} \approx 0.01212 \end{aligned}$$

12 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

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

Resulting prediction error variance

$$\sigma_p^2 \approx 0.5531$$

Optimal choices of b_1 and b_2 are given by

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

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

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

13 The variances of the transform components are

$$\sigma_0^2 \approx 3.7089, \quad \sigma_1^2 \approx 0.1927, \quad \sigma_2^2 \approx 0.0617, \quad \sigma_3^2 \approx 0.0368$$

Allocate $1.75 \cdot 4 = 7$ bits intotal to the four components. Optimal bit allocation is $R_0 = 4$, $R_1 = 2$, $R_2 = 1$ and $R_3 = 0$. The average distortion is

$$D \approx 0.009497 \cdot \sigma_0^2 + 0.1175 \cdot \sigma_1^2 + 0.3634 \cdot \sigma_2^2 + \sigma_3^2 \approx 0.0293$$

and the corresponding signal to noise ratio is

$$\mathrm{SNR} = 10 \cdot \log_{10} \frac{\sigma_X^2}{D} \approx 10 \cdot \log_{10} \frac{1}{0.0293} \approx 15.34 \text{ [dB]}$$