LINKÖPING UNIVERSITY

Solutions to Written Exam in
Image and Audio Compression
TSBK38

25th March 2026

1 See the course literature.

2 See the course literature.

3 a) See the course literature.

b) See the course literature.

4 See the course literature.

5 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 2.1345 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 2.17 bits/symbol.

7 a) Number of quantization levels

$$M = 2^{11} = 2048$$

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

$$10\log_{10}\frac{\sigma^2}{D} \approx 10\log_{10}\frac{3M^2}{25} = 10\log_{10} 503316.48 \approx 57.0 \quad [\text{dB}]$$

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

$$D \approx \frac{\pi \cdot \sqrt{3}}{2} \cdot \sigma^2 \cdot 2^{-2\cdot 11}$$

which gives the SNR

$$10\log_{10}\frac{\sigma^2}{D} \approx 10\log_{10}\frac{2^{23}}{\pi\sqrt{3}} \approx 61.9 \quad [\text{dB}]$$

8 See the course literature.

9 The decoded sequence is

$$purpurpurpurpurpurpurpurpurpurt\ldots$$

and the dictionary looks like

index	sekvens	index	sekvens	index	sekvens
0	p	8	rp	16	$rpur$
1	r	9	pur	17	$rpurt$
2	s	10	rpu	18	$t*$
3	t	11	urp		
4	u	12	$purp$		
5	v	13	$purpu$		
6	pu	14	$urpu$		
7	ur	15	$urpur$		

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

10 For symmetry reasons, the decision borders must placed as

$$b_0 = -1, \quad b_1 = 0, \quad b_2 = 1$$

The reconstruction point y_2 is given by

$$y_2 = \frac{\int_0^1 x \cdot f_X(x)dx}{\int_0^1 \cdot f_X(x)dx} = \frac{1/8}{1/2} = \frac{1}{4}$$

Also for symmetry reasons, $y_1 = -y_2$.

The distortion is given by

$$D = \int_{-1}^{0} (x + \frac{1}{4})^2 \cdot f_X(x)dx + \int_{0}^{1} (x - \frac{1}{4})^2 \cdot f_X(x)dx = \frac{3}{80}$$

11 5 bits/pixel can be considered as fine quantization. We make the approximation $\hat{Z}_{i,j} \approx Z_{i,j}$ and assume that the prediction error is also gaussian. One good predictor (there are other predictors that also solve the problem) is

$$p_{i,j} = a_1 \cdot \hat{Z}_{i-1,j} + a_2 \cdot \hat{Z}_{i,j-1} \approx a_1 \cdot Z_{i-1,j} + a_2 \cdot Z_{i,j-1}$$

The variance of the prediction error is given by

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

a_1 and a_2 that minimize σ_d^2 are given by

$$\begin{pmatrix} 1 & 0.91^{1.5} \\ 0.91^{1.5} & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0.91 \\ 0.91^{0.5} \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \approx \begin{pmatrix} 0.3323 \\ 0.6654 \end{pmatrix}$$

$$\Rightarrow \sigma_d^2 \approx 1.8206$$

Since we are free to choose, we use uniform quantization followed by source coding at the rate 5 bits/pixel, which gives the distortion

$$D \approx \sigma_d^2 \cdot \frac{\pi e}{6} \cdot 2^{-2 \cdot 5} \approx 0.002531$$

$$\mathrm{SNR} = 10 \cdot \log_{10} \frac{29}{D} \approx 40.6 \text{ [dB]}$$

12 The correlation matrix for the input signal is

$$\mathbf{R}_X = \begin{pmatrix} 1 & 0.94 & 0.94^2 & 0.94^3 \\ 0.94 & 1 & 0.94 & 0.94^2 \\ 0.94^2 & 0.94 & 1 & 0.94 \\ 0.94^3 & 0.94^2 & 0.94 & 1 \end{pmatrix}$$

The correlation matrix for the transformed signal is

$$\mathbf{R}_\Theta = \mathbf{A}\mathbf{R}_X\mathbf{A}^T \approx \begin{pmatrix} 3.7089 & 0 & -0.0547 & 0 \\ 0 & 0.1931 & 0 & 0.0014 \\ -0.0547 & 0 & 0.0617 & 0 \\ 0 & 0.0014 & 0 & 0.0361 \end{pmatrix}$$

The variances for the transform components can be found in the diagonal

$$\sigma_0^2 \approx 3.7089, \quad \sigma_2^2 \approx 0.1931, \quad \sigma_2^2 \approx 0.0617, \quad \sigma_3^2 \approx 0.0361$$

The variances can of course be found by calculating the variance individually for each transform component. For example,

$$\begin{aligned} \sigma_1^2 &= E\{\theta_1^2\} = \frac{1}{128^2} E\{(83 \cdot X_0 + 36 \cdot X_1 - 36 \cdot X_2 - 83 \cdot X_2)^2\} = \\ &= \frac{1}{16384}(16370 \cdot R_{XX}(0) + 9360 \cdot R_{XX}(1) - 11952 \cdot R_{XX}(2) - 13778 \cdot R_{XX}(3)) \approx \\ &\approx 0.1931 \end{aligned}$$

and similarly for the three other components.

Allocating bits, we find that θ_0 should be quantized with 4 bits, θ_1 with 2 bits, θ_2 with 1 bit and θ_3 with 0 bits. The resulting average distortion will be

$$D \approx \frac{0.009497 \cdot \sigma_0^2 + 0.1175 \cdot \sigma_1^2 + 0.3634 \cdot \sigma_2^2 + \sigma_3^2}{4} \approx 0.02911$$

and the corresponding signal to noise ratio is

$$\text{SNR} = 10 \cdot \log_{10} \frac{1}{D} \approx 15.36 \text{ [dB]}$$