LINKÖPING UNIVERSITY

Solutions to Written Exam in **Image and Audio Coding** **TSBK02**

1st June 2017

1 a) See the course literature.

b) See the course literature.

c) See the course literature.

2 See the course literature.

3 a) The theoretical limit is given by the entropy rate, which for this memoryless source is given by:

$$H = -0.8 \cdot \log 0.8 - 2 \cdot 0.1 \cdot \log 0.1 \approx 0.9219 \text{ bits}$$

b) You need to extend the source to two symbols. The Huffman code will have an average codeword length of 1.92 bits/codeword, which gives a rate of 0.96 bits/symbol.

4 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

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

5 Quantization to 6 bits/sample can be seen as fine quantization $\Rightarrow$ assume that the predictor works with original samples:

$$p_n = a_1 \cdot \hat{X}_{n-1} + a_2 \cdot \hat{X}_{n-2} \approx a_1 \cdot X_{n-1} + a_2 \cdot X_{n-2}$$

The prediction error variance is

$$\sigma_d^2 = E\{(X_n - p_n)^2\} \approx E\{(X_n - a_1 \cdot X_{n-1} - a_2 \cdot X_{n-2})^2\}$$

The predictor that minimizes σ_d^2 is given by

$$\mathbf{A} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \mathbf{R}^{-1} \cdot \mathbf{P}$$

where

$$\mathbf{R} = \begin{pmatrix} R_{XX}(0) & R_{XX}(1) \\ R_{XX}(1) & R_{XX}(0) \end{pmatrix} \quad \text{and} \quad \mathbf{P} = \begin{pmatrix} R_{XX}(1) \\ R_{XX}(2) \end{pmatrix}$$

This gives the optimal predictor **A** and the prediction error variance σ_d^2

$$\mathbf{A} \approx \begin{pmatrix} 0.9270 \\ -0.8382 \end{pmatrix} \qquad \sigma_d^2 \approx 1.5436$$

Uniform quantization followed by perfect source coding gives the approximative distortion

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

To get an SNR of at least 37 dB, we must choose R such that $D < \frac{\sigma_X^2}{10^{3.7}}$. This gives us approximately $R > 5.31$

6 A two point DWHT is given by the transform matrix

$$\mathbf{H}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

Performing a separable transform on blocks of size 2×2 pixels is equivalent with applying the transform

$$\mathbf{A} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

on vectors given by

$$\mathbf{x} = \begin{pmatrix} X_{i,j} \\ X_{i+1,j} \\ X_{i,j+1} \\ X_{i+1,j+1} \end{pmatrix}, \qquad \theta = \mathbf{A} \cdot \mathbf{x}$$

Variances for the four transform components are

$$\begin{aligned} \sigma_0^2 &= 2361 \\ \sigma_1^2 &= 527 \\ \sigma_2^2 &= 143 \\ \sigma_3^2 &= 9 \end{aligned}$$

Optimal bit allocation when doing Lloyd-Max quantization ($4 \cdot 2.25 = 9$ bits to allocate) is

$$\begin{aligned} R_0 &= 4 \\ R_1 &= 3 \\ R_2 &= 2 \\ R_3 &= 0 \end{aligned}$$

Resulting distortion

$$D \approx \frac{1}{4}(0.009497 \cdot \sigma_0^2 + 0.03454 \cdot \sigma_1^2 + 0.1175 \cdot \sigma_2^2 + \sigma_3^2) \approx 16.6069$$

and corresponding SNR

$$\text{SNR} \approx 10 \cdot \log_{10} \frac{760}{16.6069} \approx 16.61 \text{ [dB]}$$