Lösningsförslag till tentamen i

Kompression av ljud och bild
TSBK35

24:e mars 2022

1 Se kurslitteraturen

2 Se kurslitteraturen

3 a) Se kurslitteraturen.

b) Se kurslitteraturen.

4 Se kurslitteraturen.

5 Se kurslitteraturen.

6 a) Den teoretiskt lägsta datatakten ges av källans entropi:

$$H = -0.5 \cdot \log_2 0.5 - 0.4 \cdot \log_2 0.4 - 0.1 \cdot \log_2 0.1 \approx 1.3610 \text{ bitar/symbol}$$

b) Huffmankoden får en kodordsmedellängd som är 2.78 bitar/kodord, vilket ger datatakten 1.39 bitar/symbol. Jämför med den teoretiska gränsen (entropin).

7 Den avkodade sekvensen är

$$gagagagabeeeeedagab\ldots$$

och ordboken ser i detta läge ut som

index	sekvens	index	sekvens	index	sekvens
0	a	8	ga	16	eed
1	b	9	ag	17	da
2	c	10	gag	18	aga
3	d	11	$gaga$	19	$ab*$
4	e	12	ab		
5	f	13	be		
6	g	14	ee		
7	h	15	eee		

8 Prediktorkoefficenterna fås som lösning till ekvationssystemet

$$\begin{pmatrix} E\{X_{i-1,j}^2\} & E\{X_{i,j-1} \cdot X_{i-1,j}\} \\ E\{X_{i,j-1} \cdot X_{i-1,j}\} & E\{X_{i,j-1}^2\} \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} E\{X_{i,j} \cdot X_{i-1,j}\} \\ E\{X_{i,j} \cdot X_{i,j-1}\} \end{pmatrix}$$

$$\begin{pmatrix} R_{XX}(0,0) & R_{XX}(1,-1) \\ R_{XX}(1,-1) & R_{XX}(0,0) \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} R_{XX}(1,0) \\ R_{XX}(0,1) \end{pmatrix}$$

$$\begin{pmatrix} 251 & 205 \\ 205 & 251 \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 227 \\ 222 \end{pmatrix}$$

med lösning

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \approx \begin{pmatrix} 0.5467 \\ 0.4380 \end{pmatrix}$$

Prediktionsfelets varians:

$$\sigma_d^2 = 251 - \begin{pmatrix} a_1 & a_2 \end{pmatrix} \begin{pmatrix} 227 \\ 222 \end{pmatrix} \approx 29.67$$

Givet att kvantiseringen är fin så kommer prediktionsfelet approximativt att vara normalfördelat. Givet att vi kodar tillräckligt långa sekvenser med varje aritmetisk kodord så kan vi få en datatakt som är godtyckligt nära entropin för det kvantiserade prediktionsfelet. Den resulterande distorsionen ges då approximativt av

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

Givet datatakten $R = 4.5$ får vi alltså distorsionen

$$D \approx 0.08249$$

och signal-brusförhållandet

$$10 \cdot \log_{10} \frac{\sigma_X^2}{D} = 10 \cdot \log_{10} \frac{251}{D} \approx 34.8 \quad [\text{dB}]$$

9 Transformmatrisen (basvektorer i raderna) för en 4-punkters DCT ser ut som:

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

där $a = \frac{1}{\sqrt{2}} \cos \frac{\pi}{8} \approx 0.6533$ och $b = \frac{1}{\sqrt{2}} \cos \frac{3\pi}{8} \approx 0.2706$. Notera att $2a^2 + 2b^2 = 1$.

Varianser för de fyra transformkomponenterna:

$$\begin{aligned}
\sigma_0^2 &= E\{\theta_0^2\} = \frac{1}{4} E\{(X_0 + X_1 + X_2 + X_3)^2\} = \\
&= \frac{1}{4}(4R_{XX}(0) + 6R_{XX}(1) + 4R_{XX}(2) + 2R_{XX}(3)) \approx 3.7089 \\
\sigma_1^2 &= E\{\theta_1^2\} = E\{(aX_0 + bX_1 - bX_2 - aX_3)^2\} = \\
&= R_{XX}(0) + (4ab - 2b^2)R_{XX}(1) - 4abR_{XX}(2) - 2a^2 R_{XX}(3) \approx \\
&\approx 0.1933 \\
\sigma_2^2 &= E\{\theta_2^2\} = \frac{1}{4} E\{(X_0 - X_1 - X_2 + X_3)^2\} = \\
&= \frac{1}{4}(4R_{XX}(0) - 2R_{XX}(1) - 4R_{XX}(2) + 2R_{XX}(3)) \approx 0.06169 \\
\sigma_3^2 &= E\{\theta_3^2\} = E\{(bX_0 - aX_1 + aX_2 - bX_3)^2\} = \\
&= R_{XX}(0) - (4ab + 2a^2)R_{XX}(1) + 4abR_{XX}(2) - 2b^2 R_{XX}(3) \approx \\
&\approx 0.03614
\end{aligned}$$

Alternativt kan man räkna ut varianserna som diagonalelementen i $\mathbf{A} \cdot \mathbf{R}_X \cdot \mathbf{A}^T$, där

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

Medeldatatakten ska vara 1.75 bitar/sampel, så vi ska dela ut $4 \cdot 1.75 = 7$ bitar på de fyra transformkomponenterna. Distorsionen minimeras om man ger fyra bitar till komponent 0, två bitar till komponent 1, en bit till komponent 2 och inga bitar till komponent 3. Medeldistorsionen blir

$$D \approx \frac{1}{4}(0.009497 \cdot \sigma_0^2 + 0.1175 \cdot \sigma_1^2 + 0.3634 \cdot \sigma_2^2 + \sigma_3^2) \approx 0.02912$$

Signal-brus-förhållandet blir

$$10 \cdot \log_{10} \frac{\sigma_X^2}{D} = 10 \cdot \log_{10} \frac{1}{D} \approx 15.4 \text{ [dB]}$$