Lösningsförslag till tentamen i

Kompression av ljud och bild
TSBK35

25:e mars 2020

För att göra tentan lite mer personlig använder några av räkneuppgifterna en konstant M. Denna konstant är lika med månadsiffrorna i ditt personnummer. Till exempel, om ditt personnummer är `941205-9783`, så är $M = 12$.

1 Se kurslitteraturen

2 Se kurslitteraturen.

3 Se kurslitteraturen.

4 Se kurslitteraturen

5 Se kurslitteraturen

6 Lösning för $M = 5$.

a) Den teoretiskt lägsta datatakten ges av källans entropi:

$$H = -0.15 \cdot \log_2 0.15 - 0.75 \cdot \log_2 0.75 - 0.1 \cdot \log_2 0.1 \approx 1.0540 \text{ bitar/symbol}$$

b) Huffmankoden får en kodordsmedellängd som är 2.1375 bitar/kodord, vilket ger datatakten 1.06875 bitar/symbol.

7 Lösning för $M = 5$.

Om delintervallet som hör till x alltid läggs närmast 0, så kommer det intervall som hör till den givna sekvensen att vara $[0.71 \;\; 0.728)$.

Det kommer att gå åt minst $\lceil -\log_2(0.728 - 0.71) \rceil = 6$ bitar i kodordet, eventuellt 7.

Skriv de två gränserna som binära tal:

$$\begin{aligned} 0.71 &= 0.101101011100\ldots \\ 0.728 &= 0.101110100101\ldots \end{aligned}$$

Det minsta talet med 6 bitar i intervallet är 0.101110, men det finns tal som startar på detta sätt som ligger utanför intervallet. Alltså måste vi använda 7 bitar i kodordet.

Kodordet blir **1011011** (även **1011100** fungerar som kodord).

Alternativ lösning: Tag mittpunkten i intervallet $0.719 = (0.101110000001\ldots)_2$ och trunkera till 7 bitar, vilket ger oss kodordet 1011100.

8 Kvantiseringen är relativt fin, så vi gör approximationen att prediktionen sker från originalsamplen istället för från rekonstruerade sampel.

Prediktionsfelets varians för den första prediktorn:

$$\begin{aligned} \sigma_1^2 &= E\{(X_{i,j} - p_{i,j})^2\} \approx \\ &\approx E\{(X_{i,j} - X_{i,j-1} - X_{i-1,j} + X_{i-1,j-1})^2\} = \\ &= 4 \cdot R_{xx}(0,0) - 4 \cdot R_{xx}(0,1) - 4 \cdot R_{xx}(1,0) + 4 \cdot R_{xx}(1,1) = \\ &= 4 \cdot (1 - 0.9 - 0.9 + 0.81) = 0.04 \end{aligned}$$

Antag att den aritmetiska kodaren är perfekt. Distorsionen kommer då approximativt att bli

$$D_1 \approx \frac{\pi e}{6} \cdot \sigma_1^2 \cdot 2^{-2 \cdot 5}$$

Signal-brus-förhållandet blir

$$\mathrm{SNR}_1 = 10 \cdot \log_{10} \frac{\sigma_x^2}{D_1} \approx 42.55 \text{ dB}$$

Prediktionsfelets varians för den andra prediktorn

$$\begin{aligned} \sigma_2^2 &= E\{(X_{i,j} - p_{i,j})^2\} \approx \\ &\approx E\{(X_{i,j} - a \cdot X_{i,j-1} - b \cdot X_{i-1,j})^2\} = \\ &= (1 + a^2 + b^2) \cdot R_{xx}(0,0) - 2a \cdot R_{xx}(0,1) - 2b \cdot R_{xx}(1,0) + 2ab \cdot R_{xx}(1,1) \end{aligned}$$

Derivera uttrycket för σ_2^2 med avseende på a respektive b och sätt lika med 0. Det ger oss ekvationssystemet

$$\begin{pmatrix} R_{xx}(0,0) & R_{xx}(1,1) \\ R_{xx}(1,1) & R_{xx}(0,0) \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} R_{xx}(0,1) \\ R_{xx}(1,0) \end{pmatrix}$$

med lösningen

$$\begin{pmatrix} a \\ b \end{pmatrix} \approx \begin{pmatrix} 0.4972 \\ 0.4972 \end{pmatrix}$$

Prediktionsfelets varians blir $\sigma_2^2 \approx 0.1050$.

Distorsionen blir approximativt

$$D_2 \approx \frac{\pi e}{6} \cdot \sigma_2^2 \cdot 2^{-2 \cdot 5}$$

Signal-brus-förhållandet blir

$$\mathrm{SNR}_2 = 10 \cdot \log_{10} \frac{\sigma_x^2}{D_2} \approx 38.36 \text{ dB}$$

Den fixa trestegsprediktorn är alltså bättre i det här fallet.

9 Varianser för de fyra transformkomponenterna:

$$\begin{aligned}
\sigma_0^2 &= E\{\theta_0^2\} = \frac{1}{4}E\{(X_0 + X_1 + X_2 + X_3)^2\} = \\
&= \frac{1}{4}(4R_{XX}(0) + 6R_{XX}(1) + 4R_{XX}(2) + 2R_{XX}(3)) = 3.615744 \\
\sigma_1^2 &= E\{\theta_1^2\} = \frac{1}{20}E\{(3X_0 + X_1 - X_2 - 3X_3)^2\} = \\
&= \frac{1}{20}(20R_{XX}(0) + 10R_{XX}(1) - 12R_{XX}(2) - 18R_{XX}(3)) = 0.2513408 \\
\sigma_2^2 &= E\{\theta_2^2\} = \frac{1}{4}E\{(X_0 - X_1 - X_2 + X_3)^2\} = \\
&= \frac{1}{4}(4R_{XX}(0) - 2R_{XX}(1) - 4R_{XX}(2) + 2R_{XX}(3)) = 0.082944 \\
\sigma_3^2 &= E\{\theta_3^2\} = \frac{1}{20}E\{(X_0 - 3X_1 + 3X_2 - X_3)^2\} = \\
&= \frac{1}{20}(20R_{XX}(0) - 30R_{XX}(1) + 12R_{XX}(2) - 2R_{XX}(3)) = 0.0499712
\end{aligned}$$

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

$$\mathbf{R}_X = \begin{pmatrix} 1 & 0.92 & 0.92^2 & 0.92^3 \\ 0.92 & 1 & 0.92 & 0.92^2 \\ 0.92^2 & 0.92 & 1 & 0.92 \\ 0.92^3 & 0.92^2 & 0.92 & 1 \end{pmatrix}$$

Medeldatatakten ska vara 1.75 bitar/sampel, så vi ska dela ut $1.75 \cdot 4 = 7$ bitar på de fyra transformkomponenterna. Distorsionen minimeras om man ger fyra bitar till θ_0, två bitar till θ_1, en bit till θ_2 och inga bitar till θ_3. Medeldistorsionen blir

$$D \approx \frac{1}{4}(0.009497 \cdot \sigma_0^2 + 0.1175 \cdot \sigma_1^2 + 0.3634 \cdot \sigma_2^2 + \sigma_3^2) \approx 0.0.35996$$

Signal-brus-förhållandet blir

$$10 \cdot \log_{10} \frac{\sigma_X^2}{D} = 10 \cdot \log_{10} \frac{1}{D} \approx 14.44 \quad [\text{dB}]$$