

Tentamen i
Kompression av ljud och bild
TSBK35

19:e augusti 2022 8:00 - 12:00

| | |
|-------------------------|---|
| Lokal: | SP71 |
| Examinator: | Harald Nautsch |
| Lärare: | Harald Nautsch, 0701718715 |
| Institution: | ISY |
| Modul: | TEN1 |
| Antal uppgifter: | 9 |
| Antal sidor: | 4 + formelsamling |
| Hjälpmedel: | Miniräknare, "Tables and Formulas for Image Coding and Data Compression" |
| Betyg: | 0-13 U 14-19 3 20-25 4 26-30 5 |
| Övrigt: | Läraren är endast tillgänglig via telefon under skrivningstiden |

- 1 Vid kodning av video och stillbilder i färg brukar man använda färgrymden YCbCr istället för RGB. Förklara hur färgrymderna skiljer sig åt och varför man föredrar YCbCr.

(2 p)

- 2 En vanlig kodningsmetod för musik är *MPEG audio layer 3* (även känd som mp3).
 - a) I en typisk mp3-kodare används något som kallas för en *psykoakustisk modell*. Vad har denna för uppgift? Varför behöver man inte använda någon psykoakustisk modell vid avkodningen?

(2 p)
 - b) Hur kan man utnyttja stereoinformation i musiksигnalen för att få en effektivare kodning?

(1 p)
 - c) Förutom utnyttjandet av psykoakustik och stereoinformation, hur går kodningen av ljudsamplerna till i mp3?

(2 p)

- 3 I hybridkodare för videosignaler används rörelsekompenserad prediktion. Förklara hur detta fungerar, både på kodar- och avkodarsidan, och varför det används.

(4 p)

- 4 Två vanliga metoder för stillbildskodning är PNG och JPEG. Förklara kortfattat hur de fungerar, samt ange i vilka sammanhang man föredrar den ena metoden framför den andra.

(3 p)

- 5 Förklara vad *LBG-algoritmen* är och hur den fungerar.

(2 p)

6 En minnesfri källa har alfabetet $\mathcal{A} = \{a, b, c\}$ och symbolsannolikheterna $P(a) = 0.7$, $P(b) = 0.2$ och $P(c) = 0.1$.

a) Vilken är den teoretiskt lägsta medeldataakt man kan få vid en distorsionsfri kodning av källans utsignal?

(1 p)

b) Koda sekvensen

aaaca

med aritmetisk kodning. Ange både det resulterande intervallet och motsvarande kodord. Man kan anta oändlig noggrannhet i alla beräkningar.

(3 p)

7 En minnesfri tidsdiskret process X_n med täthetsfunktion

$$f_X(x) = \begin{cases} 1 - x & ; \quad 0 \leq x \leq 1 \\ 1 + x & ; \quad -1 \leq x \leq 0 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

kvantiseras till tre nivåer.

a) Antag att de tre rekonstruktionspunkterna väljs som $y_1 = -0.5$, $y_2 = 0$ och $y_3 = 0.5$. Hur ska beslutsgränserna väljas för att minimera distorsionen?

(1 p)

b) Givet de beslutsgränser som räknades fram i a), hur ska man välja nya rekonstruktionspunkter för att minimera distorsionen?

(1 p)

- 8 En musiksignal modelleras som en endimensionell stationär gaussprocess Y_n . Signalens statistik har estimerats till

$$E\{Y_n\} = 0$$

$$R_{YY}(k) = E\{Y_n \cdot Y_{n+k}\}$$

$$R_{YY}(0) = 9.07, \quad R_{YY}(1) = 7.56, \quad R_{YY}(2) = 4.09, \quad R_{YY}(3) = 0.08$$

Vi vill koda signalen med en medeldataakt som är högst 5 bitar per sampel och ett signal-brus-förhållande som är minst 35 dB.

Konstruera en prediktiv kodare som uppfyller kraven.

(4 p)

- 9 I den senaste videokodningsstandarden HEVC kan man som transform bland annat använda en diskret sinustransform av typ VII (DST-VII). För en signal $x_i, i = 0, \dots, N - 1$ av längd N , så definieras DST-VII-koefficienterna $\theta_j, j = 0, \dots, N - 1$ enligt

$$\theta_j = \sqrt{\frac{2}{N + 1/2}} \cdot \sum_{i=0}^{N-1} x_i \cdot \sin\left(\frac{\pi}{N + 1/2}(i + 1)(j + \frac{1}{2})\right)$$

För $N = 3$ svarar detta mot att man använder transformmatrisen

$$\mathbf{A} = \frac{2}{\sqrt{7}} \begin{pmatrix} \sin \frac{\pi}{7} & \sin \frac{2\pi}{7} & \sin \frac{3\pi}{7} \\ \sin \frac{3\pi}{7} & \sin \frac{6\pi}{7} & \sin \frac{9\pi}{7} \\ \sin \frac{5\pi}{7} & \sin \frac{10\pi}{7} & \sin \frac{15\pi}{7} \end{pmatrix}$$

En signal modelleras som en stationär gaussprocess X_n med medelvärde noll och autokorrelationsfunktion $R_{XX}(k)$

$$R_{XX}(k) = E\{X_n \cdot X_{n+k}\} = 4 \cdot 0.97^{|k|}$$

Vi vill koda signalen med en 3-punkters DST-VII och Lloyd-Maxkvantisera transformkomponenterna så att medeldataakten blir 2 bitar/sampel.

Allokera bitar till de tre transformkomponenterna så att medel-distorsionen minimeras och beräkna det resulterande signal-brus-förhållandet.

(4 p)