

Tentamen i
Kompression av ljud och bild
TSBK35

29:e maj 2023 8:00 - 12:00

Lokal:	SP71
Examinator:	Harald Nautsch
Lärare:	Harald Nautsch,
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 tentamen.

- 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 Gör en jämförelse mellan att använda *Lloyd-Max-kvantisering* och *likformig kvantisering*. Beskriv kortfattat hur de två kvantiseringsmetoderna fungerar samt vilka för- respektive nackdelar de har.

(3 p)

- 3 Förklara följande begrepp inom psykoakustik och beskriv hur man kan utnyttja dem för att få en effektivare kodning av ljud.
 - a) Hörseltröskeln

(2 p)

 - b) Frekvensmaskering

(2 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 (ganska detaljerat) hur en modern *hybridkodare* för video fungerar. Beskriv både kodning och avkodning.

(4 p)

- 6 En minnesfri källa har alfabetet $\mathcal{A} = \{a, b, c\}$ och symbolsannolikheterna $P(a) = 0.65$, $P(b) = 0.25$ and $P(c) = 0.1$.

a) Vad har källan för entropi?

(1 p)

b) Konstruera en huffmankod för par av symboler och beräkna kodens medeldataakt i bitar/symbol.

(2 p)

- 7 En minnesfri stationär tidsdiskret process X_n med täthetsfunktion

$$f_X(x) = \begin{cases} 1.5 \cdot (x - 1)^2 & ; \quad 0 \leq x \leq 1 \\ 1.5 \cdot (x + 1)^2 & ; \quad -1 \leq x < 0 \\ 0 & ; \quad \text{annars} \end{cases}$$

kvantiseras till två nivåer. Hur ska beslutsgränserna och rekonstruktionspunkterna väljas om man vill minimera den resulterande distorsionen?

(3 p)

- 8 En ljudsignal modelleras som en stationär tidsdiskret gaussprocess X_n med medelvärde 0 och autokorrelationsfunktion $R_{XX}(k) = E\{X_n \cdot X_{n+k}\}$.

$$R_{XX}(0) = 2.00, \quad R_{XX}(1) = 1.86, \quad R_{XX}(2) = 1.62$$

Signalen transformkodas med en 3-punkters DCT. Transformkomponenterna Lloyd-Max-kvantiseras skalärt så att medeldataakten blir 2 bitar/sampel.

Fördela bitar till transformkomponenterna så att medeldistorsionen minimeras och beräkna det resulterande signal-brus-förhållandet.

Vad får man för signal-brus-förhållande om man istället Lloyd-Maxkvantiserar X_n direkt, utan transformen, till samma medeldataakt som ovan?

(4 p)

- 9 En bildkälla modelleras som en tvådimensionell normalfördelad process $X_{i,j}$ (där i och j är koordinater i bilden) med medelvärde noll. Autokorrelationsfunktionen

$$R_{XX}(k, l) = E\{X_{i,j} \cdot X_{i+k,j+l}\}$$

har från en stor mängd testdata skattats till:

$$\begin{aligned} R_{XX}(0, 0) &= 121, & R_{XX}(1, 0) &= 109 \\ R_{XX}(0, 1) &= 112, & R_{XX}(1, 1) &= R_{XX}(1, -1) = 102 \end{aligned}$$

Källan kodas med prediktorn

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

Hitta de prediktorkoefficienter a_1 och a_2 som minimerar prediktionsfelets varians.

Antag att prediktionsfelet kvantiseras skalärt och källkodas minnesfritt. Vad är den lägsta datatakt som kan användas om vi vill att signal-brusförhållandet ska vara minst 36 dB?

(4 p)