

Tentamen i  
**Kompression av ljud och bild**  
**TSBK35**

23:e mars 2023 14:00 - 18:00

<b>Lokal:</b>	SP71
<b>Examinator:</b>	Harald Nautsch
<b>Lärare:</b>	Harald Nautsch,
<b>Institution:</b>	ISY
<b>Modul:</b>	TEN1
<b>Antal uppgifter:</b>	9
<b>Antal sidor:</b>	4 + formelsamling
<b>Hjälpmedel:</b>	Miniräknare, "Tables and Formulas for Image Coding and Data Compression"
<b>Betyg:</b>	0-13 U 14-19 3 20-25 4 26-30 5
<b>Övrigt:</b>	Läraren besöker tentalokalen ca 15.15 och 16.45

- 1 Vad är det för egenskaper hos en datasekvens (med diskret alfabet) som man kan utnyttja för att göra distorsionsfri komprimering?  
(2 p)
  
- 2 När man kodar video med MPEG-2 kan varje bild i sekvensen kodas som en *I-bild*, som en *P-bild* eller som en *B-bild*. Beskriv hur de tre metoderna fungerar. Vilken av de tre metoderna ger den högsta datatakten och vilken den lägsta (vid en viss bildkvalitet)? Vad finns det för anledningar att inte alltid välja den metod som ger den lägsta datatakten?  
(4 p)
  
- 3 a) Beskriv några fördelar och nackdelar med att använda *vektorkvantisering* istället för *skalär kvantisering*.  
(2 p)  
b) Förklara vad *LBG-algoritmen* är och hur den fungerar.  
(2 p)
  
- 4 Det mänskliga hörselsinnet har vissa egenskaper som kan utnyttjas för att få en effektivare kodning av ljudsignaler. Förklara vilka dessa egenskaper är och hur de kan utnyttjas i kodningen.  
(4 p)
  
- 5 I vanlig metod för talkodning är CELP. Ge en kortfattad beskrivning av hur den fungerar, både på kodar- och avkodarsidan.  
(2 p)

6 En minnesfri källa har alfabetet  $\mathcal{A} = \{x, y, z\}$  och symbol-sannolikheterna  $P(x) = 0.5$ ,  $P(y) = 0.4$  och  $P(z) = 0.1$ .

a) Vilken är den teoretiskt lägsta medeldataakten man kan få vid distorsionsfri kodning av källan?

(1 p)

b) Konstruera en huffmankod för par av symboler från källan och beräkna dess medeldataakt (i bitar/symbol).

(2 p)

7 En källa har alfabetet  $\{m, n, o, p\}$ . En lång sekvens från källan kodas med LZW. Den resulterande indexsekvensen börjar som

3, 2, 3, 4, 6, 2, 0, 9, 11, 7, 1, 12, 9, ...

Startordboken är:

index	sekvens
0	$m$
1	$n$
2	$o$
3	$p$

Avkoda indexsekvensen så långt som möjligt. Ange även hur den slutliga ordboken ser ut.

(3 p)

- 8 Linus vill lösa laboration 2 med hjälp av linjär prediktiv kodning. Han modellerar `heyhey.wav` som en tidsdiskret stationär normalfördelad process  $X_n$  med medelvärde 0. Den skattade autokorrelationsfunktionen  $R_{XX}(k) = E\{X_n \cdot X_{n+k}\}$  är

$$R_{XX}(0) = 0.0307, \quad R_{XX}(1) = 0.0301$$

$$R_{XX}(2) = 0.0292, \quad R_{XX}(3) = 0.0289$$

Konstruera en tvåstegsprediktor för signalen som minimerar prediktionsfelets varians samt beräkna denna varians.

Prediktionsfelet kvantiseras likformigt och den kvantiserade signalen kodas med en aritmetisk kodare. Steglängden i den likformiga kvantiseraren väljs så att medeldataakten blir 2.9 bitar/sampel. Vad blir signal-brus-förhållandet (i dB)?

(4 p)

- 9 Linnea vill koda sin stora samling av papegojbilder med hjälp av transformkodning. Som modell för bilderna använder hon en stationär tvådimensionell gaussprocess  $X_{i,j}$  med medelvärde 0 och med en skattad autokorrelationsfunktion  $R_{XX}(k, l) = E\{X_{i,j} \cdot X_{i+k,j+l}\}$  enligt

$$R_{XX}(0, 0) = 1000 \quad , \quad R_{XX}(0, 1) = 950 \quad , \quad R_{XX}(1, 0) = 930$$

$$R_{XX}(1, 1) = 900 \quad , \quad R_{XX}(1, -1) = 910$$

Bilderna kodas genom att ta block om  $2 \times 2$  bildpunkter och utföra en separabel hadamardtransform på blocken. Transformkomponenterna ska sedan Lloyd-Max-kvantiseras.

Den önskade medeldataakten är 1.5 bit/bildpunkt.

Hur ska bitarna fördelas mellan transformkomponenterna så att distorsionen minimeras?

Vad blir det resulterande signal-brus-förhållandet (i dB)?

(4 p)