

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

24:e mars 2021 14:00 - 18:00

<b>Lokal:</b>	Hemma
<b>Examinator:</b>	Harald Nautsch
<b>Lärare:</b>	Harald Nautsch, 0701718715
<b>Institution:</b>	ISY
<b>Provkod:</b>	TEN1
<b>Hjälpmedel:</b>	Allt kursmaterial. Kontakt med andra är inte tillåtet.
<b>Inlämning:</b>	Kortfattade svar skickas till <a href="mailto:harald.nautsch@liu.se">harald.nautsch@liu.se</a> senast 18.00. Fullständiga lösningar skickas till samma adress senast 18.30, antingen genom att scanna/fotografera handskrivna lösningar eller som digitala dokument.

För att göra tentan lite mer personlig använder några av räkneuppgifterna en konstant  $Q$ . Denna konstant beräknas från siffrorna i ditt LiU-ID (format aaabbXYZ). Om  $Z$  är jämnt är  $Q = Y$  och om  $Z$  är udda är  $Q = -Y$ .  $Q$  är alltså ett heltal mellan -9 och 9. Till exempel, om ditt LiU-ID är harna379 så är  $Q = -7$ .

- 1 a) Din vän Linnea, som inte vet särskilt mycket om datakompression, vill göra distorsionsfri kodning av ett stort antal textfiler. Hon förslår att golombkodning kanske är en bra metod för att få hög kompression. Förklara för Linnea varför golombkodning inte är ett bra val i det här fallet.

(1 p)

- b) Vilken av följande metoder skulle du föreslå att Linnea använder istället för att få så hög kompression som möjligt? Motivera ditt svar.

- Minnesfri huffmankodning
- LZSS

(1 p)

- 2 Förklara kortfattat hur aritmetisk kodning fungerar. Vad har aritmetisk kodning för fördelar jämfört med huffmankodning?

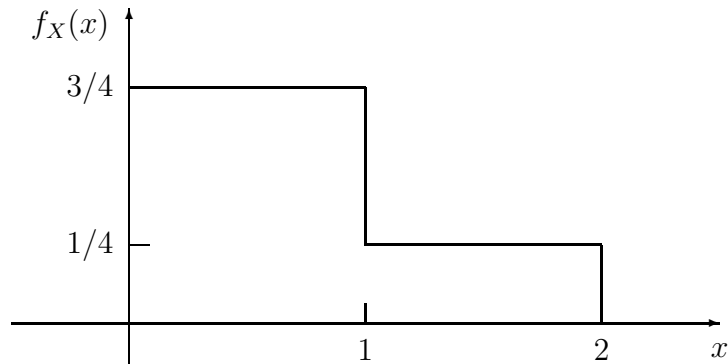
(2 p)

- 3 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)

- 4 a) När man gör transformkodning av musik använder man ofta en transform som kallas MDCT. Förklara hur den skiljer sig från den vanliga DCT:n och varför den är lämplig att använda för musikkodning. (2 p)
- b) Vid musikkodning kan man utnyttja något som kallas *frekvensmaskering* för att få en effektivare kodning. Förklara vad detta är och hur det utnyttjas i kodningen. (2 p)
- c) Normalt när man kodar musik kodar man stereosignaler, dvs musiken har en höger- och en vänsterkanal. Beskriv hur en ljudkodare kan utnyttja stereoinformation i musiksignalen för att få en effektivare kodning. (1 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  $\mathcal{A} = \{a, b, c, d, e, f, g, h\}$  och symbolsannolikheterna
- $$P(a) = 0.4 - Q/100, P(b) = 0.15, P(c) = 0.1, P(d) = 0.1$$
- $$P(e) = 0.11 + Q/100, P(f) = 0.06, P(g) = 0.05, P(h) = 0.03$$
- a) Vad är den teoretiskt lägsta medeldataakt som man kan få när man kodar källan distorsionsfritt? (1 p)
- b) Konstruera en huffmankod för källan och beräkna kodens medeldataakt i bitar/symbol. (2 p)

- 7 Den stokastiska variabeln  $X$  har täthetsfunktionen  $f_X(x)$  enligt figuren nedan.  $X$  kvantiseras till två nivåer. Beräkna kvantiseringsnivåer och beslutsområden så att medelkvadratfelet minimeras. Beräkna det resulterande medelkvadratfelet.



(3 p)

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

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

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

$$R_{YY}(0) = 6.34, \quad R_{YY}(1) = 5.28, \quad R_{YY}(2) = 2.64, \quad R_{YY}(3) = -0.63$$

Vi vill koda signalen med en linjär prediktiv kodare av formen

$$p_n = a_1 \cdot \hat{Y}_{n-1} + a_2 \cdot \hat{Y}_{n-2} + a_3 \cdot \hat{Y}_{n-3}$$

Prediktionsfelet kvantiseras likformigt och kodas med en minnesfri aritmetisk kodare. Steglängden i kvantiseraren väljs så att datatakten blir 6 bitar/sampel.

Finns de  $a_1$ ,  $a_2$  och  $a_3$  som minimerar kodarens distorsion. Beräkna vilket signal-brus-förhållande (i dB) som vi får.

(4 p)

9 Betrakta 4-punkters transformeringer av typen

$$\mathbf{A} = \begin{pmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ a & b & -b & -a \\ 1/2 & -1/2 & -1/2 & 1/2 \\ b & -a & a & -b \end{pmatrix}$$

där  $2a^2 + 2b^2 = 1$ . Både DCT och DWHT tillhör denna klass.

Vi vill transformkoda en endimensionell signal  $X_i$ . Signalen modelleras som en gaussprocess med medelvärde 0 och autokorrelationsfunktion

$$R_{XX}(k) = E\{X_i \cdot X_{i+k}\} = \rho^{|k|}$$

där  $\rho = 0.92 + Q/1000$ .

Transformkomponenterna kvantiseras med skalära Lloyd-Max-kvantiserare så att medeldata-takten blir 2 bits/sampel och medel-distorsionen minimeras.

Antag att vi väljer transformen som ges av  $a = 1/\sqrt{2}$  och  $b = 0$ . Beräkna den optimala bittilldelningen och det resulterande signal-brus-förhållandet i dB.

Hur mycket högre blir signal-brus-förhållandet om vi istället använder en DWHT (det vill säga, om  $a = b = 1/2$ )?

(5 p)