

Tentamen i
Kompression av ljud och bild
TSBK35

25:e mars 2020 14:00 - 18:00

Lokal:	Hemma
Examinator:	Harald Nautsch
Lärare:	Harald Nautsch, 0701718715
Institution:	ISY
Provkod:	TEN1
Hjälpmedel:	Allt kursmaterial. Kontakt med andra är inte tillåtet.
Inlämning:	Kortfattade svar skickas till <code>harald.nautsch@liu.se</code> senast 18.00. Fullständiga lösningar skickas till samma adress senast 19.00, antingen genom att scanna/fotografera handskrivna lösningar eller som digitala textdokument.

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 Förklara kortfattat hur Lempel-Ziv-kodning fungerar. Beskriv både LZ77 och LZ78. (2 p)

- 2 Förklara vad *LBG-algoritmen* är och hur den fungerar. (2 p)

- 3 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)

- 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 $P(a) = 0.1 + M/100$, $P(b) = 0.75$ och $P(c) = 0.15 - M/100$.

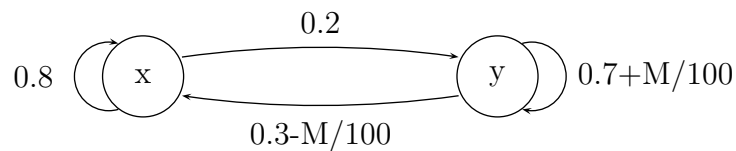
a) Vad är den teoretiskt lägsta medeldataakt som man kan få när man kodar källan distorsionsfritt?

(1 p)

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

(2 p)

7 Betrakta följande markovkälla med alfabet $\mathcal{A} = \{x, y\}$.



Koda följande sekvens med aritmetisk kodning. Källan kan antas vara i tillstånd x när kodningen startar.

$xyyyxx$

Ange både intervallet som hör till sekvensen och motsvarande kodord.

(3 p)

- 8 En bildkälla modelleras som en tvådimensionell normalfördelad process $X_{i,j}$ (där i och j är koordinater i bilden) med följande statistik

$$E\{X_{i,j}\} = 0$$
$$R_{xx}(k, l) = E\{X_{i,j} \cdot X_{i+k,j+l}\} = 0.9^{|k|+|l|}$$

Vi vill koda bilder med linjär prediktiv kodning. Vi kan välja mellan att använda en fix trepunkters prediktor

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

eller att använda en tvåpunkters prediktor

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

där prediktorkoefficienterna a och b väljs så att kodarens medeldistorsion minimeras.

Antag att prediktionsfelet kvantiseras likformigt och sedan kodas med en aritmetisk kodare. Kvantiseringssteget väljs så att medeldataakten blir 5 bitar/bildpunkt för båda prediktorerna. Beräkna vilka signal-brus-förhållanden man får för de två prediktorerna.

(5 p)

9 En 4-punkters transform ges av transformmatrisen

$$\mathbf{A} = \begin{pmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 3/\sqrt{20} & 1/\sqrt{20} & -1/\sqrt{20} & -3/\sqrt{20} \\ 1/2 & -1/2 & -1/2 & 1/2 \\ 1/\sqrt{20} & -3/\sqrt{20} & 3/\sqrt{20} & -1/\sqrt{20} \end{pmatrix}$$

Antag att vi vill koda en endimensionell signal X_n med denna transform. X_n modelleras som en gaussprocess med medelvärde 0 och autokorrelationsfunktion

$$R_{XX}(k) = E\{X_n \cdot X_{n+k}\} = 0.92^{|k|}$$

Transformkomponenterna kvantiseras med skalära Lloyd-Max-kvantiserare så att medeldataakten blir 1.75 bitar/sampel och så att medeldistorsionen minimeras.

Allokera bitar till transformkomponenterna så att kraven uppfylls och beräkna det resulterande signal-brus-förhållandet i dB.

(4 p)