INF 2310 – Digital bildebehandling

Kompresjon og koding – Del I

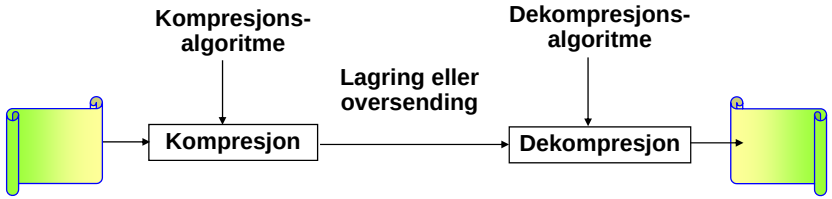
Tre steg i kompresjon
Redundanser
Transformer
Koding og entropi
Shannon-Fano og Huffman

Kompendium: Frem t.o.m. 18.7.2 + Appendiks B

Anvendelser

- Kompresjon og koding benyttes for å redusere antall biter som skal til for å beskrive bildet (eller en god approksimasjon til bildet).
- En mengde anvendelser innen data-lagring og data-overføring
  - Televideo-konferanser
  - Fjernanalyse / meteorologi
  - Overvåking / fjernkontroll
  - Telemedisin / medisinske arkiver (PACS)
  - Dokumenthåndtering / FAX
  - Multimedia / nettverkskommunikasjon
  - Mobil kommunikasjon
  - MP3-spillere, DAB-radio, digitalkameraer, …
- **Tidsforbruket ved "offline" kompresjon er ikke særlig viktig.**
- **Dekompresjons-tiden er langt viktigere.**
- **Ved "sanntids" dataoverføring er tidsforbruket kritisk.**

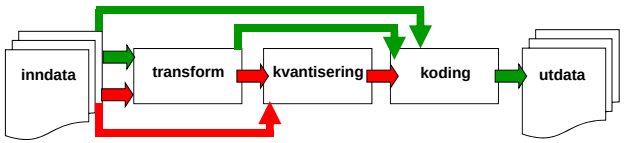
Noen begreper



- **Bildekompresjon** består i å pakke informasjonsinnholdet i bildet ved å ikke lagre redundant informasjon.
- Dataene **komprimeres**, deretter lagres eller overføres de.
- Når de senere skal leses, må vi **dekomprimere** dem.
- Koding er en del av kompresjon, men vi koder for å lagre effektivt, ikke for å hemmeligholde eller skjule informasjon.

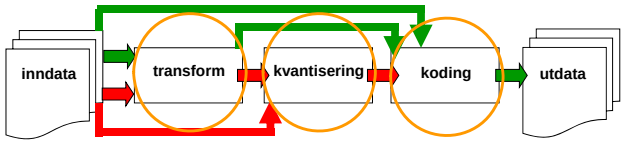
Kompresjon

- Kompresjon kan deles inn i tre steg:
  - **Transform** - representer dataene mer kompakt.
  - **Kvantisering** - avrunding av representasjonen.
  - **Koding** - produksjon og bruk av kodebok.



- Kompresjons kan gjøres
  - **Eksakt / tapsfri** ("loss-less") – følg de grønne pilene
    - Her kan vi rekonstruere den originale meldingen eksakt.
  - **Ikke-tapsfri** ("lossy") – følg de røde pilene
    - Her kan vi ikke rekonstruere meldingen eksakt.
    - Resultatet kan likevel være "godt nok".
- Det finnes en mengde ulike metoder for begge kategorier kompresjon.

De tre stegene i kompresjon



- Mange kompresjonsmetoder er basert på å representere dataene på en annen måte, altså **transformer** av original-dataene.
  - Eks.: Differansetransform, løpelengde-/"run-length"-transform
- Hvis vi **kvantiserer** (original eller de transformerte) dataene, så kan ikke dette reverseres ⇒ ikke-tapsfri kompresjon.
- Til slutt **kodes** dataene, dvs. transformeres til binærrepresentasjon.
  - Bygger ofte på normaliserte histogrammer.
- **Transformer og koding er alltid reversible.**
- **Kvantisering gir alltid et tap av presisjon.**

Eksempler - plassbehov

- Digitalt RGB-bilde:
  - 512 x 512 x 8 biter x 3 farger = 6 291 456 biter
  - 3264 x 2448 x 8 biter x 3 farger = 191 766 528 biter ≈ 24 MB
- Radarbilde fra Radarsat-1 satellitten:
  - 400MB, 300 km × 300 km, 16 biter pr. piksel.
  - Miljøovervåkning av havområder:
    - Trenger 28 bilder for å dekke hele Middelhavet
- 3D-seismikk data fra 60.000 km^2 i Nordsjøen
  - 4 000 GB = 4 Terabyte...
- Røntgen-bilde:
  - 7 112 x 8 636 piksler à 12 biter: 737 030 784 biter

Plass og tid

- Digitale data kan ta stor plass
  - Spesielt lyd, bilder og video
- Eksempler :
  1. Digitalt bilde:
     512 x 512 x 8 biter x 3 farger = 6 291 456 biter
  2. Røntgenbilde:
     7112 x 8636 x 12 biter pr. piksel = 737 030 784 biter
- Overføring av data tar tid:

Linje med 64 kbit/sek:
1. **ca. 1 min. 38 s.**
2. **ca. 3 timer 12 min.**

Linje med 1 Mbit/sek:
1. **ca. 6 s.**
2. **ca. 12 min.**

Overføringskapasitet og bps

- Filstørrelser er oftest gitt i binære enheter, gjerne i potenser av 1 024:
  - Kibibyte (KiB = 2^{10} byte = 1 024 byte),
  - Mebibyte (MiB = 2^{20} byte = 1 048 576 byte),
  - Gibibyte (GiB = 2^{30} byte = 1 073 741 824 byte)
- Overføringshastigheter og linjekapasitet angis **alltid** i titallsystemet, oftest som antall biter per sekund:
  - 1 kbps = 1000 bps = 10^3 biter per sekund.
  - 1 Mbps = 1000 kbps = 10^6 biter per sekund.
  - 1 Gbps = 1000 Mbps = 10^9 biter per sekund.
- Kapasitet for noen typer linjer:
  - GSM-telefonlinje: 9.6 kbps
  - Analogt modem: f.eks. 56 kbps
  - ADSL (Asymmetric Digital Subscriber Line): 1-2 Mbps

Melding, data og informasjon

- Vi skiller mellom data og informasjon:

- **Melding:** teksten, bildet eller signalet som vi skal lagre.
- **Data:** strømmen av biter som lagres på fil eller sendes.
- **Informasjon:** et mål som kvantifiserer mengden overraskelse/uventethet i en melding.
  - Et varierende signal har mer informasjon enn et monotont signal.
  - I et bilde: kanter rundt objekter har høyest informasjonsinnhold,
    - spesielt kanter med mye krumning.

Redundans

- Vi kan bruke ulike mengder data på samme melding.
  - Et bilde av et 5-tall (50 x 50 piksler a 8 biter = 20 000 bits)
  - Teksten "fem" i 8-biters ISO 8859-1 (24 biter)
  - ISO 8859-1 tegnet "5" (8 biter)
  - Et binært heltall "1 0 1" (3 biter)
- **Redundans** sier noe om hvor stor del av datamengden vi kan fjerne uten at vi mister informasjonen i dataene.
- Eks: Vi skal lagre tallet 0, men hvordan lagres det faktisk:
  - Binært i en byte: 00000000 - 8 biter med 0.
  - Standard 7-biters ASCII-kode for 0.
  - Skal vi bare lagre tallene 0 og 1, trenger vi bare 1 bit.
- Ved smart komprimering kan vi fjerne redundante biter.

Ulike typer redundans

- Psykovisuell redundans
  - Det finnes informasjon vi ikke kan se.
    - Kan kanskje subsample eller redusere antall bit per piksel.
- Interbilde redundans
  - Det er en viss likhet mellom nabobilder i en tidssekvens
    - Vi kan kode noen bilder i sekvensen og ellers bare differanser.
- Intersampel redundans
  - Nabopiksler ligner på hverandre eller er like.
    - Hver linje i bildet kan "run-length"-transformeres.
- Kodings-redundans
  - Gjennomsnittlig kodelengde minus et teoretisk minimum.
    - Velg en metode som er "grei" å bruke og gir liten kodingsred.

Kompresjonsrate og redundans

- **Kompresjonsraten :**

$$CR = \frac{i}{c}$$

der i er antall bit pr. sampel originalt,
og c er antall bit pr. sampel i det komprimerte bildet.

- **Relativ redundans :**

$$R = 1 - \frac{1}{CR} = 1 - \frac{c}{i}$$

- **"percentage removed" :**

$$PR = 100\left(1 - \frac{c}{i}\right)\%$$

Bildekvalitet I

- Hvis vi velger irreversibel kompresjon må vi kontrollere at kvaliteten på utbildet er "god nok".
- Gitt en M×N innbilde f(x,y) og et komprimert / dekomprimert utbilde g(x,y). Feilen vi gjør blir da:

  e(x,y) = g(x,y) – f(x,y)
- RMS-avviket (kvadratfeilen) mellom de to bildene blir da:

$$e_{RMS} = \sqrt{\frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N} e^2(x,y)}$$

- Vi kan også betrakte feilen som "støy" og se på midlere kvadratisk signal-støy-forhold (SNR):

$$(SNR)_{MS} = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N} g^2(x,y)}{\sum_{x=1}^{M}\sum_{y=1}^{N} e^2(x,y)}$$

Bildekvalitet II

- RMS-verdien av SNR er da

$$(SNR)_{RMS} = \sqrt{\frac{\sum_{x=1}^{M}\sum_{y=1}^{N} g^2(x,y)}{\sum_{x=1}^{M}\sum_{y=1}^{N} e^2(x,y)}}$$

- Kvalitetsmålene ovenfor slår sammen alle feil over hele bildet.
- Vårt synssystem er ikke slik !
- Fler-komponent kvalitetsmål er bedre! Del opp bildet etter lokal aktivitet!
- Fra Fourier-analyse:
  - For å representere en skarp kant kan vi trenge mange koeffisienter.
  - Homogene områder kan representeres ved få koeffisienter.
- Kompresjonsgraden bør kanskje variere rundt i bildet:
  - Komprimer homogene områder kraftig, men med relativt lite tap.
  - Mindre kompresjon langs kanter og linjer, men med relativt mer tap.
  - Applikasjonsavhengig! Dersom det vi leter etter ligger i heterogene områder bør disse områdene komprimeres med mindre tap enn homogene områder.

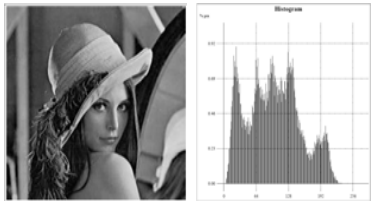
Kompresjon av bilder

- Det er en mengde redundant informasjon i bilder:
  - Mellom piksler på samme linje
  - Mellom ulike linjer i bildet
  - Objekter kan representeres mer kompakt enn med piksler.
  - Vi kan også ta i betraktning psykovisuelle effekter.
- Symmetrisk kompresjon: komprimering og dekomprimering tar like lang tid.
- Assymetrisk kompresjon, komprimering og dekomprimering har ulik regnetid, ofte tillater man langsom komprimering, men ønsker rask dekomprimering.

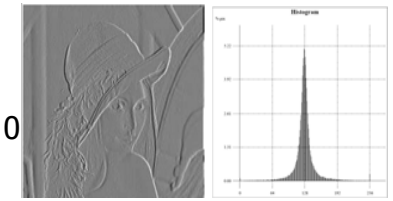
Differansetransform

- Gitt en rad i bildet med gråtoner $f_1,.....f_N$, $0 \le f_i \le 2^b-1$.
- Transformer (reversibelt) til

$g_1 = f_1$, $g_2 = f_2-f_1$, ... , $g_N = f_N-f_{N-1}$

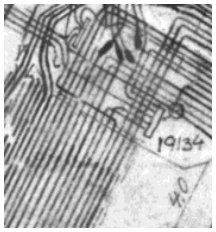


- Vi trenger nå b+1 biter hvis vi skal tilordne like lange binære koder til alle mulig verdier.
- I differansehistogrammet vil de fleste verdiene samle seg rundt 0
- En naturlig bit-koding av differansene er ikke optimal.

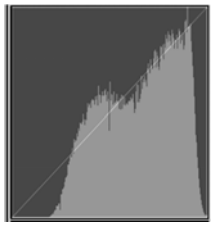


Differansebilder og histogram

Original:



Originalt Histogram:



Differanse-bilde (linje for linje) :

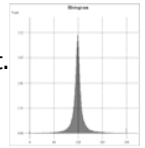
Histogram til Differansebildet:

Mulig koding av differansetransform

- Lag f.eks. en 16 ords naturlig kode
  - c_1=0000, c_2=0001,.... c_{16}=1111
- Tilordne de 14 kodeordene i midten:
  - c_2,.... c_{15} til differansene -7, -6, ..., -1, 0, 1, 2, ..., 5, 6
- Kodene c_1 og c_{16} kan brukes til å indikere om differansen $\Delta x < -7$ eller om $\Delta x \geq 7$ (to-sidet shift-kode)
  - $\Delta x = -22 \Rightarrow c_1c_1c_{15}$ $\Delta x = 21 \Rightarrow c_{16}c_{16}c_2$

…	-22	-21	…	-8	-7	…	0	…	6	7	…	20	21	…
…	$c_1c_1c_{15}$	c_1c_2	…	c_1c_{15}	c_2	…	c_9	…	c_{15}	$c_{16}c_2$	…	$c_{16}c_{15}$	$c_{16}c_{16}c_2$	…

- Her øker kodeordets lengde trinnvis med 4 biter.
  - Neppe helt optimalt i forhold til differansehistogrammet.



Løpelengde-transform

- Ofte inneholder bildet objekter med lignende gråtoner, f.eks. svarte bokstaver på hvit bakgrunn.
- Vi kan benytte oss av kompresjonsteknikker som tar hensyn til at nabopiksler på samme rad ofte er like.
- Løpelengde-transform er reversibel.
- Først et eksempel:
  333333555555555544777777 (24 tall)
- Når tallet 3 forekommer 6 ganger etter hverandre, trenger vi bare lagre tallparet (3,6).
- Tilsammen trenger vi her 4 tallpar, (3,6), (5,10), (4,2), (7,6) altså 8 tall til å lagre hele sekvensen 24 tall ovenfor.
- Hvor mange biter vi bruker pr. tall, avhenger av den videre kodingen.

Løpelengder i binære bilder

- I to-nivå bilder trenger vi bare å angi løpelengden for hvert "run".
  - Må også vite/definere om raden starter med hvitt eller svart "run".
- Vi trenger kodeord for EOL og EOI.
- "Run-length"-histogrammet er ofte ikke flatt.
  - Benytter da en kode som gir korte kodeord til de hyppigste løpelengdene.
- I ITU-T (tidligere CCITT) standarden for dokument-overføring pr. fax Huffman-kodes løpelengdene.
  - Forhåndsdefinerte Huffman-kodebøker, én for svarte og én for hvite "runs".
- Mer effektiv dersom kompleksiteten i bildet er lav.

Koding

- Et *alfabet* er mengden av alle mulige symboler (gråtoner)
- Hvert symbol får et *kodeord*.
- Tilsammen utgjør kodeordene *kodeboken.*

- Koding skal være reversibel
  - fra koden skal vi kunne rekonstruere det originale symbolet
  - vi kaller denne egenskaper ved koder for unik dekodbarhet; en sekvens kodeord kan dekodes på én og bare én måte.

- Instantant dekodbare koder kan dekodes uten skilletegn.

Naturlig binærkoding

- Alle kodeord er like lange.
  - Kjenner vi noen eksempler på dette?
- Eks: En 3-biters kode gir 8 mulige verdier:

Symbol nr.	1	2	3	4	5	6	7	8
Symbol	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8
Kode c_i	000	001	010	011	100	101	110	111

- Naturlig binærkoding er bare optimal hvis alle verdiene i sekvensen er like sannsynlige.

"Gray code"

Er den konvensjonelle binære representasjonen av gråtoner optimal?

- La oss se på et gråtonebilde med b bitplan.
- Ønskelig med minst mulig kompleksitet i hvert bitplan
  - Da blir løpelengde-transformasjonen mest effektiv.

- Konvensjonell binær representasjon gir høy bitplan-kompleksitet.
  - Hvis gråtoneverdien fluktuerer mellom 2^k-1 og 2^k vil k+1 biter skifte verdi: eksempel: 127 = 01111111 mens 128 = 10000000

- I "Gray Code" skifter alltid bare én bit når gråtonen endres med 1.

- Overgangen fra naturlig binærkode til "gray code" er en transform, men både naturlig binær kode og "gray code" er koder.

Gray Code tranformasjoner

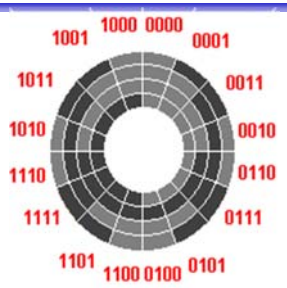
- Transformasjon fra "Binary Code" til "Gray Code":
  1. Start med MSB i BC og behold alle 0 inntil du treffer 1
  2. 1 beholdes, men alle følgende bits komplementeres inntil du treffer 0
  3. 0 komplementeres, men alle følgende bit beholdes inn du treffer 1
  4. Gå til 2.

- Fra "Gray Code" til "Binary Code":
  1. Start med MSB i GC og behold alle 0 inntil du treffer 1
  2. 1 beholdes, men alle følgende bits komplementeres inntil du treffer 1.
  3. 1 komplementeres, men alle følgende bits beholdes inntil du treffer 1.
  4. Gå til 2

”Binary reflected gray code”

4 bits Gray kode og binær

Gray	Binær	Desimal
0000g	0000b	0d
0001	0001	1
0011	0010	2
0010	0011	3
0110	0100	4
0111	0101	5
0101	0110	6
0100	0111	7
1100	1000	8
1101	1001	9
1111	1010	10
1110	1011	11
1010	1100	12
1011	1101	13
1001	1110	14
1000	1111	15



”Gray code shaft encoder”

Gir sikker avlesing av vinkel.

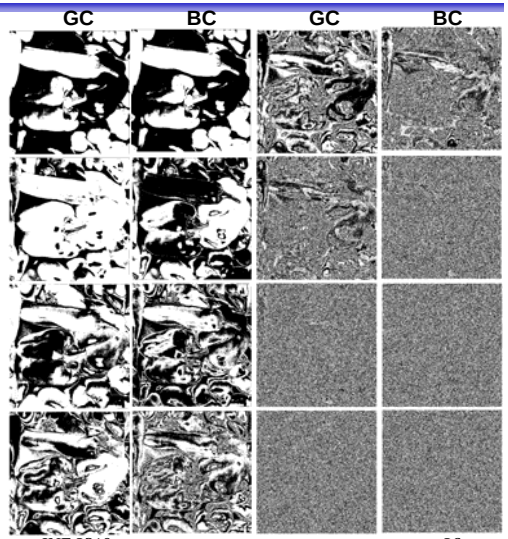
Emilie Baudot’s telegrafskive 1878.

Koden patentert av Gray i 1953.

Brukes i styring av robot-armer etc.

Gray kode i gråtonebilder

- MSB er likt i de to representasjonene.
- Større homogene områder i hvert bitplan i Gray-kode enn i naturlig binærkode.
- Flere bitplan med støy i vanlig binærkode.
- Det er en gevinst i løpelengdekoding av bitplan i Gray kode.



Informasjonsteori og koding

- Koding bygger på sannsynligheter.
- Forekommer en pikselverdi ofte, bør vi lagre den med et lite antall biter for å bruke minst mulig lagerplass totalt.
- Et symbol som forekommer sjeldent, kan vi tillate oss å bruke mange biter på å lagre.
- Vi bruker da et variabelt antall biter pr. symbol.
- Vi skal først se på koding av enkeltpiksler.
- Interpiksel redundans minimeres av transform-steget:
  - Eks.: Differansetransform, løpelengdetransform.

Koder med variabel lengde

- For ulike sannsynligheter er **koder med variabel lengde** på kodeordene bedre enn like lange koder
  - Hyppige symboler ⇨ kortere kodeord.
  - Sjeldne symboler ⇨ lengre kodeord.
  - Dette var forretnings-ideen til Samuel Morse

De vanligste symbolene i engelsk tekst er :

e, t, a, o, i, n, ...

A	. -	F	. . - .	K	- . -	P	. - - .	U	. . -
B	- . . .	G	- - .	L	. - . .	Q	- - . -	V	. . . -
C	- . - .	H	. . . .	M	- -	R	. - .	W	. - -
D	- . .	I	. .	N	- .	S	. . .	X	- . . -
E	.	J	. - - -	O	- - -	T	-	Y	- . - -

Entropi – en liten forsmak

- Entropi er et matematisk mål på gjennomsnittlig informasjonsmengde i en sekvens av tegn eller tall.
- *Har vi en sekvens av N tall eller tegn som lagres med b biter pr. sampel, så kan vi si vi har et alfabet med 2^b mulige symboler.*
- Vi kan finne sannsynligheten for hvert symbol i alfabetet:
  - $P(s_i) = n_i/N$
- Vi er interessert i gjennomsnittlig informasjon pr. symbol.
- Intuitivt har vi at en mindre sannsynlig hendelse gir mer informasjon enn en mer sannsynlig hendelse
  - Informasjon er relatert til mengden overraskelser.

Histogram og normalisert histogram

- Vi har en sekvens med N symboler.
- Tell opp antall ganger symbol s_i forekommer og la n_i være dette antallet.
  - Dette er det samme som histogrammet til sekvensen.
- Sannsynligheten til symbolene finnes da som:
  - $p_i = n_i / N$
  - Dette er det normaliserte histogrammet.

Gjennomsnittlig antall biter pr. piksel

- Vi konstruerer en kode $c_1,....c_N$
  slik at symbol s_i kodes med kodeordet c_i.
- b_i er lengden av kodeordet c_i (angitt i biter).
- Gjennomsnittlig antall biter pr. symbol for denne koden:

$$R = b_1 p_1 + b_2 p_2 + ... + b_N p_N = \sum_{i=1}^{N} b_i p_i$$

- Entropien H gir oss en nedre grense
  for hvor mange biter vi gjennomsnittlig trenger pr. symbol (hvis vi bare koder ett symbol av gangen).

Informasjonsinnhold

- Definer informasjonsinnholdet $I(s_i)$ i hendelsen s_i ved

$$I(s_i) = \log_2 \frac{1}{p(s_i)}$$

- $\log_2(x)$ er 2-er logaritmen til x
  - Hvis $\log_2(x)=b$ så er $x=2^b$
  - Eks: $\log_2(64)=6$ fordi $64=2^6$ (=2*2*2*2*2*2)
    $\log_2(8)=3$ fordi $8=2*2*2=2^3$
  - $\log_2(\text{tall}) = \log_{10}(\text{tall}) / \log_{10}(2)$
- $\log_2(1/p(s_i))$ gir oss informasjonsinnholdet i hendelsen: "symbolet s_i forekommer en gang", uttrykt i biter.

Entropi

- Hvis vi tar gjennomsnittet over alle symbolene s_i i alfabetet, får vi gjennomsnittlig informasjon pr. symbol.
- Entropi H:

$$H = \sum_{i=0}^{2^b-1} p(s_i) I(s_i) = -\sum_{i=0}^{2^b-1} p(s_i) \log_2(p(s_i))$$

- Entropien setter en nedre grense for hvor kompakt sekvensen kan representeres
  - Dette gjelder hvis vi bare koder hvert symbol for seg.

Øvre og nedre grense for entropi

- Hvis alle symboler like sannsynlige => entropi lik antall biter.
  - Det er 2^b symboler, og sannsynligheten for hvert av dem er $p(s_i)=1/2^b$. Da blir entropien

$$H = -\sum_{i=0}^{2^b-1} \frac{1}{2^b} \log_2(\frac{1}{2^b}) = -\log_2(\frac{1}{2^b}) = b$$

  - NB: Hvis det er 2^b symboler som alle er like sannsynlige, så kan de ikke representeres mer kompakt enn med b biter pr symbol.
- Hvis alle pikslene er like => entropi lik 0.
  - Hvis bare ett symbol forekommer, er sannsynligheten for dette symbolet lik 1, og alle andre sannsynligheter er lik 0.

$$H = -\log_2(1) = 0$$

To eksempler

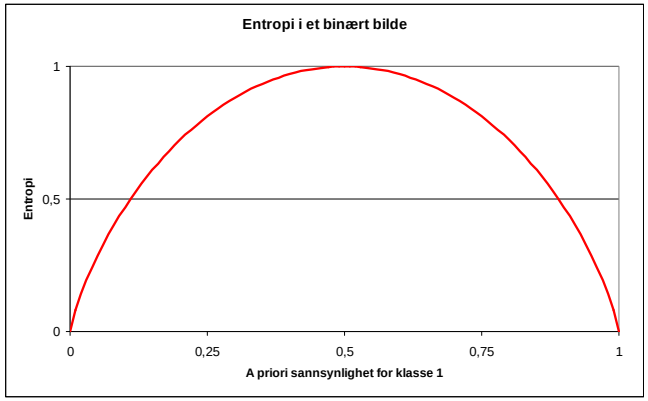
- Et binært bilde med M*N piksler inneholder 1 bit per piksel.
- Det er M*N biter data i bildet, men hvor mye informasjon er det i bildet?
- Hvis det er like mange 0 som 1 i bildet, så er det like stor sannsynlighet for at neste piksel er 0 som 1. Informasjonsinnholdet i hver mulig hendelse er da like stort, og entropien til et nytt symbol er 1 bit.

$$H = \frac{1}{2}\log_2(\frac{1}{1/2}) + \frac{1}{2}\log_2(\frac{1}{1/2}) = \frac{1}{2}*1 + \frac{1}{2}*1 = 1$$

- Hvis det er 3 ganger så mange 1 som 0 i bildet, så er det mindre overraskende å få en 1, og det skjer oftere. Entropien er da mindre:

$$H = \frac{1}{4}\log_2(\frac{1}{1/4}) + \frac{3}{4}\log_2(\frac{1}{3/4}) = \frac{1}{4}*2 + \frac{3}{4}*0.415 = 0.5 + 0.311 = 0.811$$

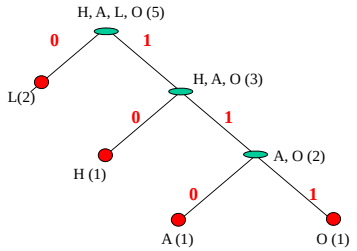
Entropi i binært bilde



- Vi vil alltid måtte bruke 1 bit per piksel i et binært bilde, selv om entropien godt kan bli nær null!
- Kodingsredundansen er null når det er like mange svarte og hvite piksler.

Shannon-Fano koding I

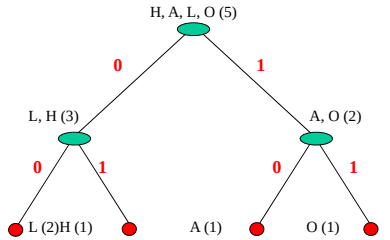
- En enkel metode:
  - Sorterer symbolene etter hyppighet, hyppigst til venstre.
  - Deler symbolene rekursivt i to grupper som forekommer så like hyppig som mulig.
    - Fortsett til hver gruppe inneholder ett symbol.
    - For hver steg, tilordn 0 til venstre gruppe og 1 til høyre gruppe.
  - Traverser treet "fra rot til blad"
    - Finner koden for hvert symbol.
- Eksempel: "HALLO" :



Symbol	Ant.	Kodeord	Lengde	Antall biter
L	2	0	1	2
H	1	10	2	2
A	1	110	3	3
O	1	111	3	3
Totalt antall biter				10

Shannon-Fano koding II

- Oppdeling i to "omtrentlig" like store grupper kan gi et annet binært tre:
- Samme eksempel: "HALLO"
- Selv om treet er annerledes, og kodeboken blir forskjellig, så er koden unikt dekodbar.
- Her har vi altså to likeverdige løsninger.



- Generelt for Shannon-Fano-koding er gjennomsnittlig antall biter per symbol er relatert til entropien:

  $H \leq R \leq H+1$
- Øvre grense for kodingsredundans:

  1 bit per symbol.

Symbol	Ant.	Kodeord	Lengde	Antall biter
L	2	00	2	4
H	1	01	2	2
A	1	10	2	2
O	1	11	2	2
Totalt antall biter				10

Huffman-koding

- Huffman-koding er en algoritme for variabel-lengde koding som er optimal under begrensningen at vi koder symbol for symbol.
- Er basert på at vi kjenner hyppigheten for hvert symbol
  - Dette betyr at vi må lage et histogram.
  - Ofte beregner vi sannsynlighetene
    - Men det holder at vi kjenner hyppighetene.

Oppskrift - Huffman-koding

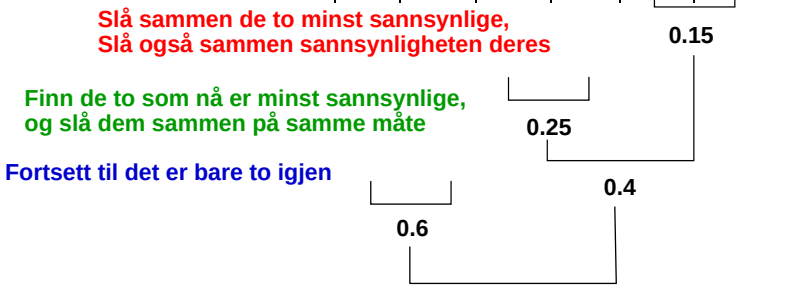
Gitt en sekvens med N symboler:

1. Sorter symbolene etter sannsynlighet, slik at de minst sannsynlige kommer sist.
2. Slå sammen de to minst sannsynlige symbolene i en gruppe, og sorter igjen etter sannsynlighet.
3. Gjenta 2 til det bare er to grupper igjen.
4. Gi koden 0 til den ene gruppen og koden 1 til den andre.
5. Traverser innover i begge gruppene og legg til 0 og 1 bakerst i kodeordet til hver av de to undergruppene.

Eksempel - Huffman-koding

- Gitt 6 begivenheter A, B, C, D, E, F med sannsynligheter

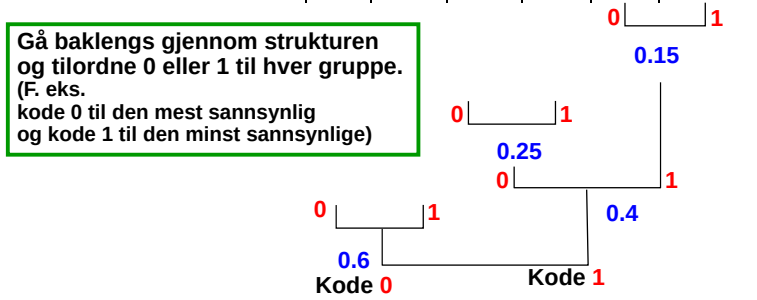
Begivenhet	A	B	C	D	E	F
Sannsynlighet	0.3	0.3	0.13	0.12	0.1	0.05



Eksempel - Huffman-koding

- Gitt 6 begivenheter A, B, C, D, E, F med sannsynligheter

Begivenhet	A	B	C	D	E	F
Sannsynlighet	0.3	0.3	0.13	0.12	0.1	0.05



Kodeboken

- Dette gir følgende kodebok

Begivenhet	A	B	C	D	E	F
Kode	00	01	100	101	110	111

- Med sannsynlighetene 0.3 0.3 0.13 0.12 0.1 0.05
  blir gjennomsnittlig antall biter pr. symbol (R) for denne koden:
  (se foil 31)

$$R = b_1 p_1 + b_2 p_2 + \ldots + b_N p_N = \sum_{i=1}^{N} b_i p_i = 0.6 * 2 + 0.4 * 3 = 2.4$$

- Entropien H er her mindre enn R (se foil 33):

$$H = -\sum_{i=0}^{2^b - 1} p(s_i) \log_2 (p(s_i)) = 2.34$$

Kodingsredundans for Huffman

- Shannon-Fano metoden har en øvre grensen for kodingsredundans på 1 bit per symbol.
- Kodingsredundansen til Huffman-koden har en øvre grense som er en funksjon av p_0 , der p_0 er sannsynligheten til det hyppigste symbolet.
  - Kodingsredundansen blir større ettersom p_0 nærmer seg 1.

Generelt om Huffman-koding

- **Ingen kodeord danner prefiks i en annen kode**
  - Dette sikrer at en sekvens av kodeord kan dekodes entydig
  - Og at man IKKE trenger endemarkører / skilletegn.
  - **Mottatt kodesekvens er unikt og instantant (øyeblikkelig) dekodbar.**
  - Dette gjelder også Shannon-Fano og naturlig binærkoding.
- **Hyppige symboler har kortere koder enn sjeldne symboler.**
  - Dette gjelder også for Shannon-Fano.
- De to minst sannsynlige symbolene har like lange koder.
  - Siste bit skiller dem fra hverandre.
  - Dette gjelder også for Shannon-Fano.
- **Merk at kodeboken må overføres!**
  - Dette gjelder også for Shannon-Fano.
- **Lengden av en Huffman kodebok er opptil G=2^b kodeord.**
- **Det lengste kodeordet kan ha opptil G-1 biter.**

Kodingsredundans

- La oss bruke de 6 mest sannsynlige symbolene i engelsk tekst:

Symbol	^	e	t	a	o	i
Sannsynlighet	0.34	0.19	0.14	0.12	0.11	0.10
Huffman-kodeord	00	10	010	011	110	111
Ordlengde	2	2	3	3	3	3
Entropi	0.529	0.455	0.397	0.367	0.350	0.332

- Det gjennomsnittlige antall biter per symbol, R, er gitt ved:

$$R = \sum_{i=1}^{N} b_i p_i = 2 \cdot 0.53 + 3 \cdot 0.47 = 2.47$$

- Entropien H er litt mindre enn R:

$$H = -\sum_{i=0}^{2^b-1} p(s_i) \log_2(p(s_i)) \approx 2.43$$

- Kodingsredundansen R-H er dermed:

$$R - H \approx 2.47 - 2.43 = 0.04$$

Ideell og faktisk kodeord-lengde

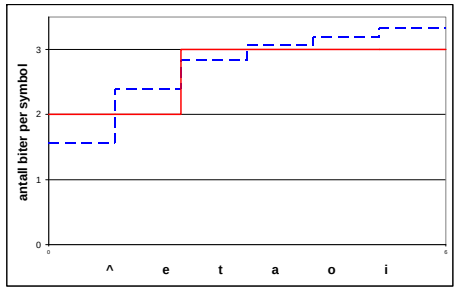
- Vi definerte informasjonsinnholdet I(s_i) i hendelsen s_i ved:

$$I(s_i) = \log_2 \frac{1}{p(s_i)}$$

- Den ideelle binære kodeordlengden for symbol s_i er dermed:

$b_i = - \log_2(p(s_i))$.

- Plotter den ideelle lengden på kodeordene sammen med den faktiske kodeordlengden, som er heltall:



Når gir Huffman-koding ingen kodingsredundans?

- Den ideelle binære kodeord lengden for symbol s_i er

$b_i = - \log_2(p(s_i))$

- Siden bare heltalls ordlengder er mulig, er det bare når

$$p(s_i) = \frac{1}{2^k}$$

for et heltall k som dette er tilfredsstilt.

Eksempel: hvis vi har

Symbol	s_1	s_2	s_3	s_4	s_5	s_6
Sannsynlighet	0.5	0.25	0.125	0.0625	0.03125	0.03125
Huffman-kodeord	0	10	110	1110	11110	11111

blir den gjennomsnittlige ordlengden R = 1.9375 = H.