Løsningsforslag

# UNIVERSITETET I OSLO

## Det matematisk-naturvitenskapelige fakultet

| | |
|---|---|
| Eksamen i: | INF2310 — Digital bildebehandling |
| Eksamensdag: | Onsdag 28. mai 2014 |
| Tid for eksamen: | 09:00 – 13:00 |
| Løsningsforslaget er på: | **8 sider** |
| Vedlegg: | **Ingen** |
| Tillatte hjelpemidler: | **Ingen** |

- Det er 7 oppgaver i dette oppgavesettet.

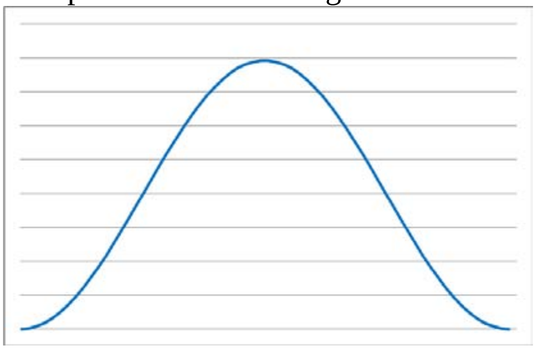
- Les gjennom hele oppgavesettet før du begynner å løse oppgavene ! Kontroller at oppgavesettet er komplett før du begynner å besvare det. Dersom du savner opplysninger i en oppgave, kan du selv legge dine egne forutsetninger til grunn og gjøre rimelige antagelser, så lenge de ikke bryter med oppgavens "ånd". Gjør i såfall rede for forutsetningene og antagelsene du gjør.

- Det er tilsammen 18 deloppgaver. **Hver deloppgave teller like mye.** Det lønner seg derfor å disponere tiden slik at man får besvart alle deloppgavene. Hvis du står fast på en deloppgave, gå videre slik at du får gitt et kort svar på alle deloppgavene.

- **Alle svar skal begrunnes.** Gjør rede for bruken av eventuelle teoremer, prinsipper eller forutsetninger slik at en tredjeperson kan følge dine resonnementer.

## 1. Histogramtransform

Anta at du har et analogt bilde $f$ med gråtoner $0 \leq z \leq 2\pi$ der bildets en-toppede histogram er gitt ved en én-perioders sinusoid og ser slik ut:

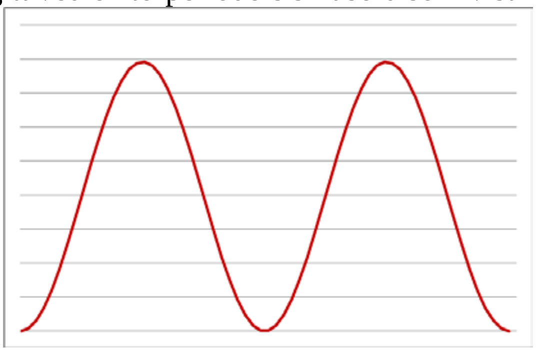


**a)** Forklar hvordan vi finner gråtonetransformen som histogramutjevner bildet uten å endre intervallet gråtonene dekker, dvs. $[0, 2\pi]$.
Hva blir uttrykket for denne transformen?

Løsningsforslag:

1. Finn bildets normaliserte histogram p(z)
2. Lag det normaliserte kumulative histogrammet c(z).
3. Gråtonetransformen er $T(z) = 2\pi\, c(z)$

**b)** Forklar hvordan vi finner en histogramtransform som gir et utbilde med et to-toppet histogram gitt ved en to-perioders sinusoid som vist i figuren nedenfor.



Løsningsforslag: Ofte skal vi gå fra et vilkårlig histogram. For eksempel et to-toppet, til et en-toppet, for eksempel et Gaussisk histogram. Men vi kan selvsagt gjøre det omvendt. Oppskriften er den samme:

1. Vi har fått spesifisert ønsket nytt histogram g(z)
2. Finn den transformen $T_g$ som histogramutjevner g(z).
3. Inverstransformer det histogramutjevnede bildet fra deloppgave a ved $s=T_g^{-1}(T(z))$

**c)** Hva skjer med gråtone-entropien til bildet ved histogramutjevningen i deloppgave a? Forklar!

Løsningsforslag: Siden vi her er i det kontinuerlige domenet, vil vi få en perfekt histogramutjevning, altså et helt flatt histogram, og vi er garantert at entropien øker.

(I det diskrete domenet ville vi hatt søyler i histogrammet som enten ble flyttet eller slått sammen. Flytting av søyler ville ikke endret entropien, mens sammenslåinger ville minsket entropien. Men dette har vi ikke spurt om her.)

## 2. Filtrering i Fourier-domenet

**a)** Hva er det vi mener med *ringing* og hvorfor vil det forekomme etter filtrering med et ideelt lavpassfilter i Fourier-domenet?

Løsningsforslag: Ringing er striper/ringer som brer seg ut fra markante kanter og som er forårsaket av filtreringen (de finnes altså i ut-bildet, men ikke i inn-bildet).

Det å filtrere med et ideelt lavpassfilter i Fourier-domenet er ifølge konvolusjonsteoremet akkurat det samme som å (sirkel)konvolvere med filterets romlige representasjon. Siden den inverse 2D diskrete Fourier-transformen av et ideelt lavpassfilter er en trunkert sinc-funksjon vil den ha en stor sentral topp med ringing utenfor. Når man konvolverer et bilde med et slikt filter vil man få ringing.

**b)** Hvorfor vil ikke ringing forekomme etter filtrering med et Gaussisk lavpassfilter i Fourier-domenet?

Løsningsforslag: Den inverse 2D diskrete Fourier-transformen av et Gaussisk lavpassfilter er et Gauss-filter. Siden Gauss-funksjonen er monotont avtagende fra sin maksimalverdi i origo så kan ikke konvolveringen med et slikt filter forårsake ringing.

## 3. Kantdetektering

**a)** Beskriv hvordan man kan beregne gradientmagnitudebildet av et inn-bilde *f* ved bruk av Sobel-operatoren:

$h_x$ =

| | | |
|---|---|---|
| 1 | 2 | 1 |
| 0 | $\boxed{0}$ | 0 |
| -1 | -2 | -1 |

$h_y$ =

| | | |
|---|---|---|
| 1 | 0 | -1 |
| 2 | $\boxed{0}$ | -2 |
| 1 | 0 | -1 |

Cellen med tykk kant angir origo i filteret, dvs. der $x = y = 0$.

Løsningsforslag:

1. Beregn $g_x = f*h_x$
2. Beregn $g_y = f*h_y$
3. Gradientmagnitudebildet er da $(g_x^2 + g_y^2)^{1/2}$

**b)** I Cannys kantdetektoralgoritme blir gradientmagnitudebildet tynnet (før hysterese-tersklingen). Hvordan har vi lært at dette kan bli gjort?

Løsningsforslag: Vi går gjennom gradientmagnitudebildet M piksel for piksel og danner et tynnet gradientmagnitudebilde $g_N$. Hvis piksel (i,j) i gradientmagnitude-bildet har en 8-nabo i eller mot (den 45 grader kvantifiserte) gradientretningen med høyere verdi så settes $g_N(i,j)$ til 0, hvis ikke, altså hvis M(i,j) er større eller lik begge sine to 8-naboer i eller mot gradientretningen, så settes $g_N(i,j)$ lik M(i,j).

**c)** Tynningen i Cannys kantdetektoralgoritme fører grovt sett til at vi står igjen med de lokalt sett største gradientmagnitudene. Hvordan kan vi finne pikslene med lokalt sett størst gradientmagnitude ved bruk av en Laplace-operator?

Løsningsforslag: Vi konvolverer inn-bildet med Laplace-operatoren og leter etter nullgjennomganger i filtreringsresultatet. I pikslene med nullgjennomganger er gradientmagnituden lokalt sett størst.

# 4. Kompresjon og koding

**a)** Beregn differansetransformen av bildet:

| | | | | | | |
|---|---|---|---|---|---|---|
| 7 | 7 | 7 | 7 | 7 | 7 | 7 |
| 7 | 7 | 7 | 6 | 6 | 5 | 5 |
| 7 | 5 | 3 | 2 | 2 | 2 | 2 |
| 3 | 2 | 2 | 2 | 2 | 2 | 2 |

Løsningsforslag: Det differansetransformerte bildet er:

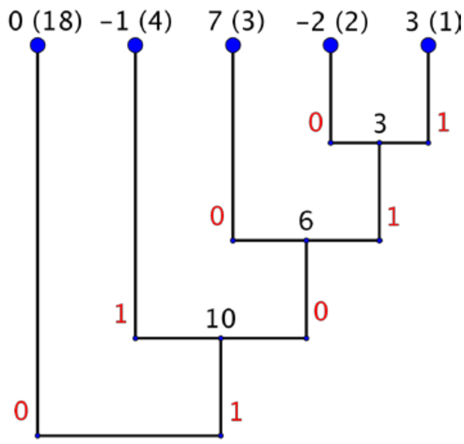
| | | | | | | |
|---|---|---|---|---|---|---|
| 7 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | -1 | 0 | -1 | 0 |
| 7 | -2 | -2 | -1 | 0 | 0 | 0 |
| 3 | -1 | 0 | 0 | 0 | 0 | 0 |

**b)** Finn Huffman-kodeboken av det differansetransformerte bildet.
Dersom du ikke løste deloppgave a så kan du finne Huffman-kodeboken av det opprinnelige bildet i stedet.

Løsningsforslag: En opptelling av antall forekomster av hvert symbol i differansebildet gir:

| **Differanseverdi** | 0 | -1 | 7 | -2 | 3 |
|---|---|---|---|---|---|
| **Antall forekomster** | 18 | 4 | 3 | 2 | 1 |

Huffman-sammenslåingene blir for dette histogrammet entydige. Hvis vi for hver forgrening tilordner 0 til den mest sannsynlige gruppen og 1 til den minst sannsynlige gruppen (og bare gjør et valg for forgrening i de to like sannsynlige gruppene {7} og {-2, 3}), får vi følgende Huffman-kodetre:




(Figuren er en del av løsningsforslaget.)

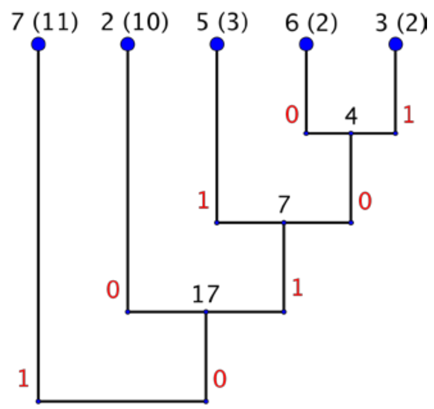
Dette gir følgende Huffman-kodebok:

| **Differanseverdi** | 0 | -1 | 7 | -2 | 3 |
|---|---|---|---|---|---|
| **Huffman-kodeord** | 0 | 11 | 100 | 1010 | 1011 |

Dersom vi i stedet Huffman-koder det opprinnelige bildet får vi:

| **Gråtoneverdi** | 7 | 2 | 5 | 6 | 3 |
|---|---|---|---|---|---|
| **Antall forekomster** | 11 | 10 | 3 | 2 | 2 |

Huffman-sammenslåingene blir igjen entydige. Hvis vi igjen for hver forgrening tilordner 0 til den mest sannsynlige gruppen og 1 til den minst sannsynlige gruppen (og denne gangen bare gjør et valg for forgrening i de to like sannsynlige gruppene {6} og {3}), får vi følgende Huffman-kodetre:




(Figuren er en del av løsningsforslaget.)

Dette gir følgende Huffman-kodebok:

| **Gråtoneverdi** | 7 | 2 | 5 | 6 | 3 |
|---|---|---|---|---|---|
| **Huffman-kodeord** | 1 | 00 | 011 | 0100 | 0101 |

**c)** Hva mener vi med at JPEG-kompresjon kan gi blokk-artefakter?

Løsningsforslag: At vi i det komprimerte og dekomprimerte bildet ser hvordan bildet har blitt oppdelt i blokker under kompresjonen. Dette ser vi fordi feilen som kvantifiseringen har innført, f.eks. glatting- eller ringing-feil, er tydelig forskjell mellom naboblokker, noe som gjør at vi ser grensene mellom naboblokker.

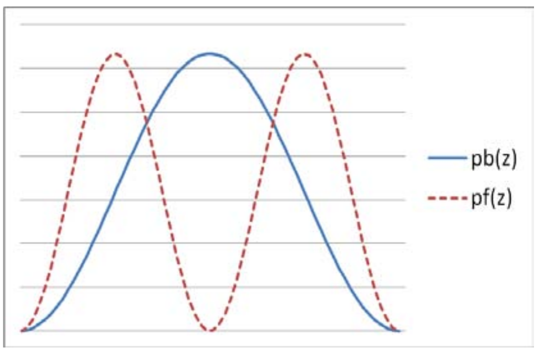
## 5. Segmentering ved terskling

I et analogt bilde med gråtoner $0 \leq z \leq 2\pi$ er de normaliserte histogrammene til bakgrunn og forgrunn gitt som sinusoider med henholdsvis én og to perioder:

$$p(z) = \begin{cases} k[1-\cos(z)] & \textit{bakgrunn} \\ k[1-\cos(2z)] & \textit{forgrunn} \end{cases}$$

der $k$ er en felles konstant.

De normaliserte histogrammene $p_b(z)$ og $p_f(z)$ ser altså slik ut:



**a)** Finn et enkelt uttrykk for den høyeste verdien, $p_{max}$, i de normaliserte histogrammene. Forklar resonnementet!

Løsningsforslag: De normaliserte histogrammene er histogrammer med heltallig periode, og ligger innenfor et rektangel med sider [0, $p_{max}$] og [0, $2\pi$]. Pga symmetri er arealet under og over de normaliserte histogrammene like, mens arealet under et normalisert histogram alltid er lik 1, slik at (ligningen er en del av løsningsforslaget):

$$p_{\max} \cdot 2\pi = 2 \;\Rightarrow\; \underline{\underline{p_{\max} = 1/\pi}}$$

**b)** Med *a priori* sannsynlighet F for forgrunn, finn en annengradsligning for terskelverdiene som gir minst mulig total feil ved en histogrambasert segmentering. Hint: $\cos(2z) = \cos^2(z) - \sin^2(z) = 2\cos^2(z) - 1$.

Løsningsforslag: Det er skjæringspunktene mellom de to fordelingene veiet med de respektive *a priori* sannsynlighetene som gir oss terskelverdiene som gir minst mulig total feil, altså (ligningene er en del av løsningsforslaget):

$$F \cdot p_f(T) = B \cdot p_b(T)$$
$$\Rightarrow F \cdot [1 - \cos(2T)] = B \cdot [1 - \cos(T)]$$
$$\Rightarrow 2F\cos^2(T) - B\cos(T) - 2F + B = 0$$
$$\Rightarrow \underline{\underline{2F\cos^2(T) + (F-1)\cos(T) + 1 - 3F = 0}}$$

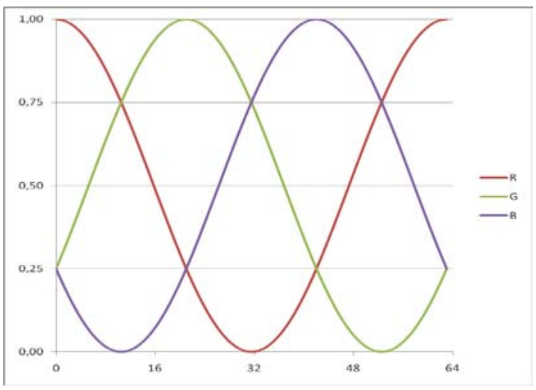
**c)** Anta at du har et gråtonebilde med ganske homogen bakgrunn og en del objekter med ganske homogen intensitet. Finnes det en gruppe filtre fra pensum i dette kurset som vil gi et resultat-bilde der histogrammet er en veiet sum av to histogrammer som har samme form som $p_b$ og $p_f$ ovenfor? Forklar!

Løsningsforslag: Et LoG-filter eller et Laplace-filter.

## 6. Fargebilder og fargetabeller

Et 6-bits gråtonebilde med L = 64 gråtoner vises fram med en RGB-pseudofargetabell der R er en sinusoid, R=0.5(1+cos(2πz/64)), mens G og B-komponentene er tilsvarende sinusoider som er faseforskjøvet med henholdsvis 4π/3 og 2π/3, som vist i figuren nedenfor. Når gråtonen *l* øker, vil fargen (H) øke lineært fra 0 til 1 med denne fargetabellen.
I oppgaven kan du få bruk for formlene for metning (S) og intensitet (I):

$$S = 1 - \frac{3\min(R,G,B)}{R+G+B} \qquad I = \frac{R+G+B}{3}$$



**a)** Hvilke farger vil piksler med gråtoneverdi $l = L/6$, $l = L/2$ og $l = 5L/6$ vises som?

Løsningsforslag:
*l=L/6: B=0, og G=R=0.75 som er en gul farge, og den er mettet ifølge formelen ovenfor.*
*l=L/2: R=0, og G=B=0.75 som er cyan, og den er mettet ifølge formelen ovenfor.*
*l=5L/6: G=0, og R=B=0.75 som er magenta, og også denne fargen er mettet.*

**b)** Vi har i en av ukeoppgavene i kurset sett på en fargetabell med lineær intensitet, og intensiteten er faktisk lineær også i fargetabellen ovenfor. Hvis vi nå konverterer fra RGB til HSI, hvordan vil da intensiteten I endre seg når gråtonen *l* går fra 0 til 63?

Løsningsforslag: I = (R+G+B)/3 = 0.5 for alle verdier av *l*, så I er konstant lik ½. Dette kan man enkelt se for alle skjæringspunktene i figuren ovenfor og dermed dedusere ut fra at det oppgis at I er lineær. Se også figuren nederst. Det er lett å vise at dette gjelder for alle gråtoner, men det krever vi IKKE:

$$I = \frac{R+G+B}{3} = \frac{1}{3}\left[\frac{1}{2}(1+\cos x) + \frac{1}{2}\left(1+\cos\left(x+\frac{4\pi}{3}\right)\right) + \frac{1}{2}\left(1+\cos\left(x+\frac{2\pi}{3}\right)\right)\right]$$

$$= \frac{1}{6}\left[1+\cos x + 1 + \cos x\cos\frac{2\pi}{3} - \sin x\sin\frac{2\pi}{3} + 1 + \cos x\cos\frac{4\pi}{3} - \sin x\sin\frac{4\pi}{3}\right]$$

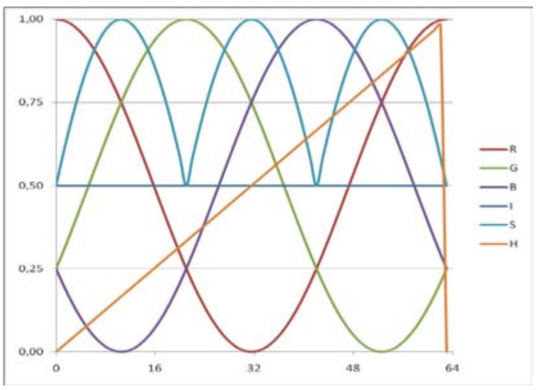
$$= \frac{1}{6}\left[1+\cos x + 1 - \frac{1}{2}\cos x - \frac{1}{2}\sqrt{3}\sin x + 1 - \frac{1}{2}\cos x + \frac{1}{2}\sqrt{3}\sin x\right] = \underline{\underline{\frac{1}{2}}}$$

**c)** For hvilke av gråtonene som svarer til skjæringspunktene i figuren øverst vil denne pseudo-fargetabellen gi halvt mettede farger?

Løsningsforslag: For at vi skal få halv metning, må vi ha 3min(R,G,B)/(R+G+B)=1/2. Vi får følgende tabell:

| *l* | 0 | L/6 | L/3 | L/2 | 2L/3 | 5L/6 |
|---|---|---|---|---|---|---|
| S | 0.5 | 1 | 0.5 | 1 | 0.5 | 1 |

Vi har altså halv metning for $l = 0$, $l = L/3$ og $l = 2L/3$, og høyere metning for alle andre verdier av *l*, se figuren nedenfor.



# 7. Morfologisk lukking

Utfør morfologisk lukking på følgende binære bilde:

| | | | | |
|---|---|---|---|---|
| **1** | 0 | 0 | **1** | 0 |
| 0 | 0 | 0 | **1** | 0 |
| 0 | 0 | 0 | 0 | **1** |
| 0 | 0 | 0 | **1** | 0 |
| 0 | 0 | **1** | 0 | 0 |
| 0 | **1** | 0 | 0 | 0 |
| **1** | **1** | 0 | 0 | **1** |

ved bruk av følgende binære strukturelement:

| | | |
|---|---|---|
| **1** | **1** | **1** |
| **1** | **1** | **1** |
| **1** | **1** | **1** |

Cellen med tykk kant angir origo.

I både bildet og strukturelementet angir 1 forgrunnspiksel og 0 bakgrunnspiksel.

Løsningsforslag: Lukkingsresultatet er:

| | | | | |
|---|---|---|---|---|
| **1** | **1** | **1** | **1** | 0 |
| 0 | 0 | 0 | **1** | 0 |
| 0 | 0 | 0 | **1** | **1** |
| 0 | 0 | 0 | **1** | 0 |
| 0 | 0 | **1** | **1** | 0 |
| 0 | **1** | **1** | **1** | 0 |
| **1** | **1** | **1** | **1** | **1** |

Dette resultatet er enklest å komme frem til vba. den geometriske tolkningen av lukking som vi har gjennomgått mens man husker at vi alltid antar 0 utenfor bildet når vi utfører morfologiske operasjoner. Dersom man skal få samme resultat vba. definisjonen av lukking, dvs. først dilatere og så erodere, så må man huske detaljen om at bilderammen i utgangspunktet ikke skal begrense resultatet av en morfologisk operasjon (dette følger av mengdedefinisjonene til operasjonene), så etter dilasjonen vil bildet være større enn inn-bildet i oppgaveteksten (siden alle forgrunnselementene etter eroderingen er innenfor det opprinnelige bildeutsnittet kan derimot lukkingsresultatet vises som et bilde av samme størrelse som inn-bildet). Ettersom vi ikke har lagt vekt på denne detaljen i forelesningen er den heller ikke lagt stor vekt på i vurderingen av denne oppgaven.

**TAKK FOR OPPMERKSOMHETEN!**