UNIVERSITETET I OSLO

Det matematisk-naturvitenskapelige fakultet

Eksamen i :	INF2310 — Digital bildebehandling
Eksamensdag :	Onsdag 3. juni 2009
Tid for eksamen :	14:30 – 17:30
Løsningsforslaget er på :	**7** sider
Vedlegg :	Ingen
Tillatte hjelpemidler :	Ingen

- Les gjennom hele oppgaven før du begynner å løse oppgaven.
  Kontroller at oppgavesettet er komplett før du begynner å besvare det.
  Dersom du savner opplysninger i oppgaven, 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 lønner seg å disponere tiden slik at man får besvart alle oppgavene.
  Hvis du står fast på enkeltoppgaver, så gå videre slik at du i hvert fall får gitt et kort svar på alle oppgavene.

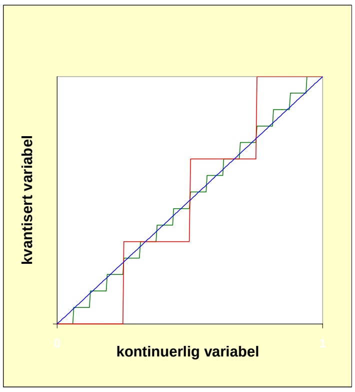
Dette er et løsningsforslag

1. Kvantisering og histogramoperasjoner

a) Anta at vi i utgangspunktet har B biter per piksel i et digitalt bilde Hvis vi så reduserer antall biter per piksel med to (fra B til B-2), hvor mange kvantiseringsnivåer vil vi miste?

Med B biter per piksel har vi 2^B kvantiseringsnivåer. Etter endringen i kvantiseringen har vi 2^(B - 2). Altså får vi 2^(B)-2^(B - 2) færre kvantiseringsnivåer. Vi mister halvparten for hver bit vi tar bort, og sitter igjen med ¼. Eksempel: 2^8 = 256, 2^6 = 64. Vi mister 256-64= 192 nivåer.

b) Anta at vi i utgangspunktet kvantiserer til $B = 4$ biter per piksel, med en kvantisering som illustrert i figuren til høyre, der 16 gråtoneintervaller mappes til 16 nivåer. Vi antar en uniform fordeling av gråtoner.
Anta så at vi reduserer antall biter per piksel med to.
For hvor stor andel av pikselverdiene vil det ikke bli noen endring i kvantiserings-feilen? Illustrer med figur!

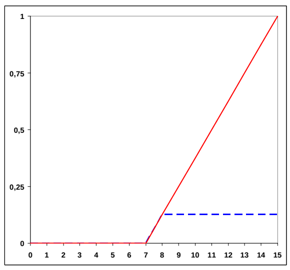


For hvert bit vi fjerner halveres antall nivåer. For ¼ av pikselverdiene vil vi ikke få noen endring i kvantiseringsfeilen, og siden vi antar en uniform fordeling, gjelder dette også ¼ av pikslene.

c) Anta at vi har et bilde f med gråtoneskala $0 \leq i \leq (G-1)$, med $G = 16$, og at det normaliserte histogrammet er gitt ved

$$p(i) = \begin{cases} 0 & 0 \leq i < G/2 \\ 2/G & i \geq G/2 \end{cases}$$

Skisser det kumulative normaliserte histogrammet til bildet f.



c[i] stiger lineært med 1/8 per gråtone fra c=0 ved i=7 til c=1 ved i=15.

d) Anta at vi har et bilde $f(x,y)$ med B biter per piksel. Vi produserer et binært bilde $g(x,y)$ som består av det *mest* signifikante bitplanet i bildet f. Hvilken enkel bildebehandlingsoperasjon svarer dette til? Beskriv gjerne med en formel!

Med B biter har vi G gråtoner nummerert fra 0 til G-1 : 0, ..., $2^B - 1$.
Det mest signifikante bitplanet er en terskling av f(x,y) ved T = (G/2) - 1:

$$g(x,y) = \begin{cases} 0 \text{ hvis } f(x,y) \leq T = 2^{B-1} - 1 \\ 1 \text{ hvis } f(x,y) > T \end{cases}$$

2. Konvolusjon

a) Du får oppgitt en kombinert konvolusjonsoperator:

$$O = -\left(A * A * B^T * B^T \ + \ A^T * A^T * B * B\right)$$
$$A = \begin{bmatrix}1 & 0 & -1\end{bmatrix}, \qquad B = \begin{bmatrix}1 & 1 & 1\end{bmatrix}$$

Hvilken operator er O, og hva bruker vi den til?
Bruk egenskaper ved konvolusjon til å begrunne svaret uten å utføre konvolusjonen!

Operatoren O kan omformes til

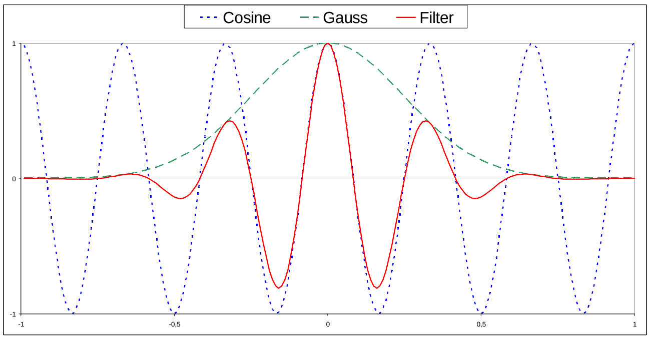
$$O = -\left((A * B^T) * (A * B^T) \ + \ (A^T * B) * (A^T * B)\right)$$

*($A*B^T$) og (A^T*B) er de to konvolusjons-filtrene i Prewitt-operatoren som vi bruker til å estimerer gradienten (førstederiverte) i hhv x- og y-retning. ($A*B^T$) * ($A*B^T$) og (A^T*B)* (A^T*B) gir oss da estimater av den andre-deriverte i hhv x- og y-retning. Så operatoren O svarer til en Laplace-operator realisert ved hjelp av Prewitt-operatoren.*

b) Hvor stor er filtermatrisen O ovenfor?
Begrunn svaret uten å utføre konvolusjonen!

Operatoren O = -($A*A*B^T*B^T$ + A^T*A^T*B*B) kan skrives som
O = -(($A*A$)(B^T*B^T) + (A^T*A^T)*($B*B$)).*
($A*A$) og ($B*B$) er radvektorer som er [5x1].
(A^T*A^T) og (B^T*B^T) er kolonnevektorer som er [1x5].
O er da en matrise som er [5x5].

c) Hvis vi lager oss et filter ved å multiplisere et 2-D lavpass Gauss-filter G(x,y) med en 1-D cosinus-funksjon Cos(x), så vil et snitt gjennom filtret langs x-aksen kunne se slik ut:



Filtret vil altså ha flere minima og maksima med avtagende amplitude på hver side av den sentrale toppen.

Hvilket teorem fra pensum kan du bruke til å forklare at Fourier-transformen av dette filtret bare har to maksima?
Forklar hvordan du bruker teoremet i dette tilfellet.

Konvolusjonsteoremet sier at en multiplikasjon i bildedomenet svarer til en konvolusjon i Fourier-domenet. Fourier-transformen av filtret er da en konvolusjon av Fourier-transformen av Gaussfiltret - som er en 2-D Gauss-funksjon – og Fourier-transformen av en Cosinus-funksjon, som er to singulariteter på u-aksen ved ± frekvensen til sinusoiden. Når det bare er disse to singularitetene som skal konvolveres med en 2-D Gauss, får vi bare to maksima.

3. Median-filtrering og Huffman-koding

Gitt et 3-biters gråtonebilde med 5 x 5 piksler:

2	2	2	2	2
2	1	1	1	2
2	1	5	5	6
2	1	5	5	6
2	2	6	6	6

a) Dette bildet skal filtreres med et 3x3 piksels medianfilter.
Vi antar at bare piksler der bildet og filterkjernen overlapper helt skal beregnes.
Hva blir resultatet?
Svar: Vi får et 3 x 3 utbilde:

2	2	2
2	1	5
2	5	6

b) Gjør en medianfiltrering av hvert bitplan i bildet, og lag et nytt gråtonebilde.
Blir resultatet det samme?
De tre bitplanene blir (MSB til venstre, LSB til høyre)

0	0	0	0	0
0	0	0	0	0
0	0	1	1	1
0	0	1	1	1
0	0	1	1	1

1	1	1	1	1
1	0	0	0	1
1	0	0	0	1
1	0	0	0	1
1	1	1	1	1

0	0	0	0	0
0	1	1	1	0
0	1	1	1	0
0	1	1	1	0
0	0	0	0	0

3x3 medianfiltrering vil fjerne hjørner i alle bitplanene.

0	0	0
0	0	1
0	1	1

1	0	1
0	0	0
1	0	1

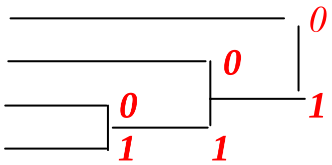
0	1	0
1	1	1
0	1	0

Og resultatbildet blir IKKE det samme som i oppgave a::

2	1	2
1	1	5
2	5	6

c) Vis hvordan du vil gå fram for å Huffman-kode gråtonebildet, og finn det gjennomsnittlige antall biter per piksel i det kodede bildet.

2:	*11*	***0***
1:	*5*	***10***
6:	*5*	***110***
5:	*4*	***111***

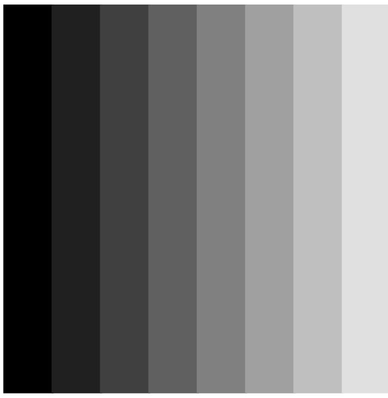


Og det gjennomsnittlige antall bits pr piksel etter koding blir
(11x1 + 5x2 + (5+4)*3)/25 = 48/25 = **1.92 biter/piksel.**

d) Er det mulig å få samme kodetre og kodebok med en Shannon-Fano koding?
Begrunn svaret!
Ja, kodetreet i Shannon-Fano koding lages ved gjentatt oppdeling i "omtrent like store grupper", og ***kan*** *gi samme kodetre og kodebok som Huffman.*

4. Terskling av gråtonebilde

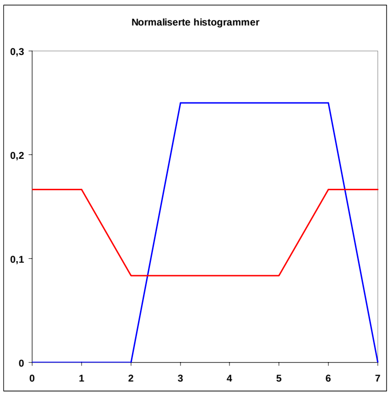
I et 3 biters gråtonebilde med størrelse 64 x 64 piksler består bakgrunnen av vertikale striper som er 8 piksler brede. Intensiteten i stripene øker fra venstre mot høyre, fra 0 til 7, som illustrert i figuren til høyre, som bare viser bakgrunnen i bildet.



På hver av de fire midterste stripene legges det objekter som utgjør 1/2 av stripens areal, og som er ett gråtonetrinn lysere enn bakgrunnen.

a) Skisser de normaliserte histogrammene til bakgrunnspikslene og objektpikslene i bildet.

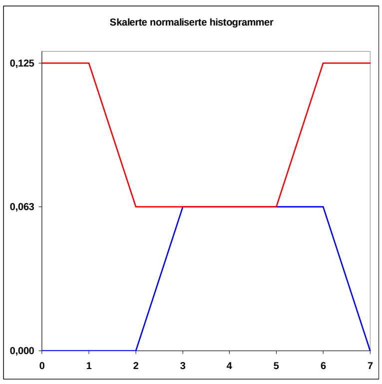
De to normaliserte histogrammene ser ut som i figuren til høyre.



Det er 3 biter i hvert piksel, så gråtoneskalaen går fra 0 til 7.
Bakgrunnspikslene i de to første og de to siste bakgrunnsstripene har samme sannsynlighet, de fire stripene i midten bare halvparten, og tilsammen skal disse sannsynlighetene være 1. pikselverdier. Altså er det normaliserte bakgrunns-histogrammet 1/6, 1/6, 1/12, 1/12, 1/12, 1/12, 1/6, 1/6.
Det normaliserte forgrunnshistogrammet er 0,0, 0, ¼, ¼, ¼, ¼, 0.

b) Skisser de normaliserte histogrammene skalert med a priori sannsynlighet.

Forgrunnspikslene utgjør ½ av hver av de fire midterste stripene, altså ¼ av bildet.
Bakgrunnspikslene utgjør da ¾ av bildet.
Det skalerte normaliserte bakgrunnshistogrammet blir da
1/8, 1/8, 1/16, 1/16, 1/16, 1/16, 1/8, 1/8.
Det skalerte normaliserte forgrunnshistogrammet blir:
0, 0, 0, 1/16, 1/16, 1/16, 1/16, 0.

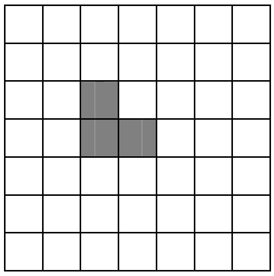


c) Forklar hvorfor vi her ikke har noen entydig løsning på om vi må ha en eller to terskler for å få minst mulig total feil-segmentering ved terskling, og heller ikke en entydig løsning på hvor terskelverdien(e) ligger.

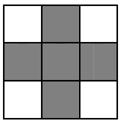
De to skalerte fordelingene skjærer ikke hverandre, de faller sammen for tre gråtoneverdier. I beste fall vil ¼ av bildet feilklassifiseres for t =[3,4,5].

5. Morfologiske operasjoner

Gitt bildet f der objektpikslene er grå:



og strukturelementet S:

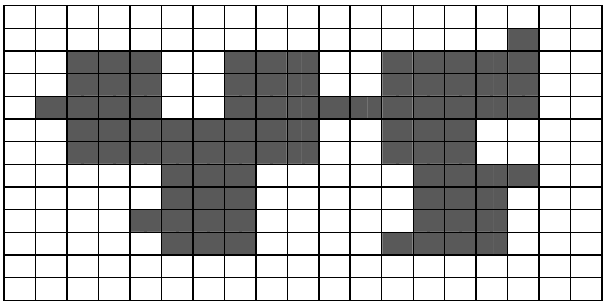


a) Skisser resultatet etter at bildet f er dilatert med S.

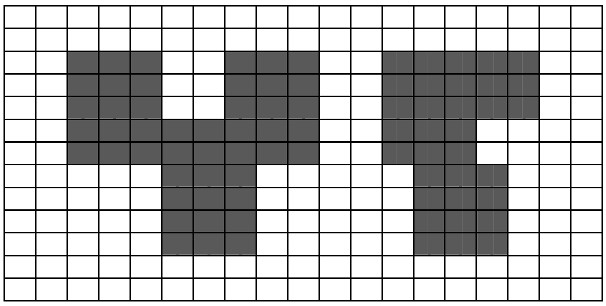
b) Hva blir resultatet etter gjentatte dilasjoner av f med S?

Tilslutt vil hele bildet fylles med objektpiksler.

c) Under ser du to bilder, et originalbilde og et filtrert bilde. Det filtrerte bildet er laget ved å gjøre en morfologisk operasjon på originalbilde med et strukturelement med størrelse 3x3. Hvilken morfologisk operasjon er gjort, og hvordan ser 3x3 strukturelementet ut? Begrunn svaret ditt.



Original

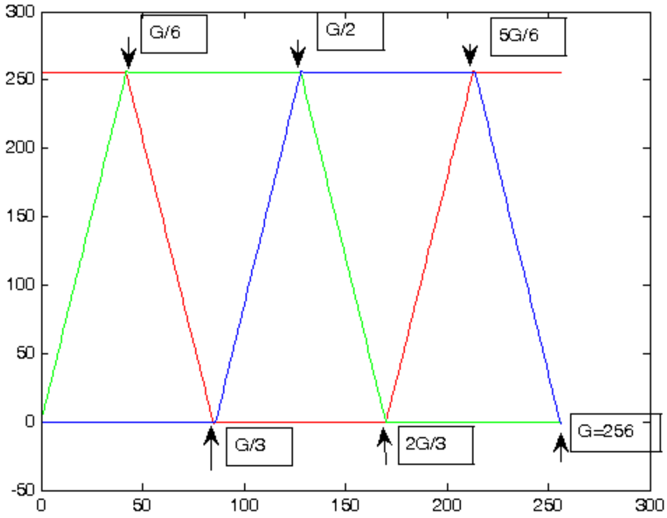


Filtrert bilde.

Det er gjort åpning med 3x3 kvadratisk strukturelement.

6. Fargebilder og fargerom

Et 8-bits gråtonebilde (med G=256 gråtoner) vises fram med en RGB-pseudofargetabell der R, G og B-komponenten er som vist i figuren.

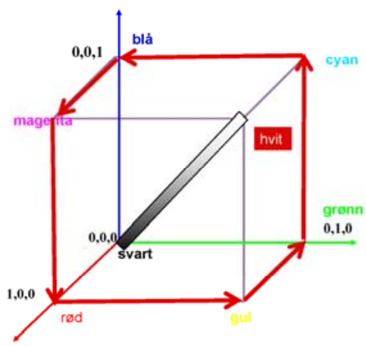


a) Hvilken farge vil piksler med gråtoneverdi G/6 vises som?
G/6 har maksverdi for både rødt og grønt og blir derfor gult.

b) Hvilken farge vil piksler med gråtoneverdi 2G/3 vises som?
2G/3 har maks blå,og rødt og grønt lik null, dvs. blå.

c) Kan du kort beskrive hvordan fargene vil endre seg når gråtonene går fra 0 til 255, dvs. hvordan fargetabellen ser ut. ?
Den vil være som en regnbue, gå fra rødt, gult, grønt, cyan, blå, magenta, og til rødt igjen.

d) Hvis du tegner en RGB-kube, hvilke deler av RGB-kuben vil fargetabellen dekke? Illustrer med en figur!
Den vil dekke seks av ytterkantene som vist på figuren.



Takk for oppmerksomheten!

Anne Schistad Solberg og Fritz Albregtsen