UNIVERSITETET I OSLO

Det matematisk-naturvitenskapelige fakultet

Eksamen i :	INF2310 — Digital bildebehandling
Eksamensdag :	Onsdag 4. juni 2008
Tid for eksamen :	14:30 – 17:30 (3 timer)
Oppgavesettet er på :	*Dette er et løsningsforslag*
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.

- Merk at alle delspørsmål teller like mye. Det lønner seg derfor å disponere tiden slik at man får besvart alle oppgavene. Hvis dere står fast på enkeltoppgaver, gå videre slik at dere får gitt et kort svar på alle oppgaver.

- **Alle svar skal begrunnes.** Det er ikke tilstrekkelig å gi svar som ”Ja”, ”Nei”, eller bare en tallverdi.

Oppgave 1

Gitt følgende 3 x 3 utsnitt av et 3 bits gråtonebilde med pikselverdier

3	1	7
2	5	7
1	7	3

Hvilken verdi vil senterpikslet her få hvis vi bruker:

a. Et 3x3 homogent/uniformt konvolusjonsfilter hvor summen av filterkoeffisientene er 1?

Svar: (3+1+7+2+5+7+1+7+3)/9 = 36/9 = 4.

b. Et 3x3 kvadratisk medianfilter?

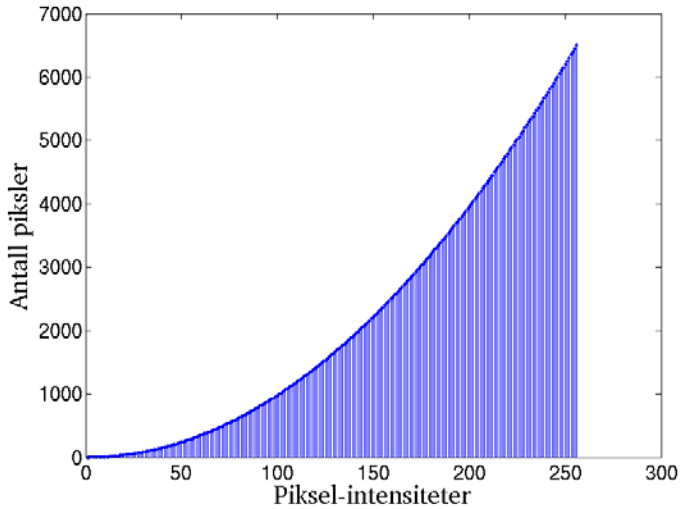
Svar: Sorterer : {1,1,2,3,3,5,7,7,7}. I midten finner vi medianen, M=3.

c. Et 3x3 pluss-formet medianfilter?

Svar: Sorterer og finner {1,2,5,7,7}; Midt i listen finner vi medianen, M=5.

NB! For å få full uttelling kan du ikke bare gi et svar i form av en pikselverdi. Du må også gi en kort begrunnelse, eller en utregning.

Oppgave 2



a. Figuren over viser histogrammet til et gitt 8-bits bilde. Anta at du skulle rekvantisere pikslene i bildet, og at du kun har 8 kvantiseringsnivåer tilgjengelig. Skisser hvor du ville satt disse. Forklar kort prinsippene bak valget.

Den totale (re)kvantiseringsfeilen er hver "søyle" multiplisert med avstanden til det nye kvantiseringsnivået. For å minimere denne feilen settes derfor i dette tilfellet kvantiseringsgrensene gradvis nærmere hverandre etter hvert som pikselintensitetene øker.

b. Anta at vi har et avbildningssystem som gir en punktspredningsfunksjon med bredde 0.5mm. Altså vil det kunne skille punkter som har avstand 0.5mm mellom seg. Hva er den minste samplingsraten (**frekvensen**) vi må benytte ifølge samplingsteoremet? Vær presis med benevningen.

Samplingsteoremet krever $T_s <$ ½T, der T=0.5mm, følgelig må $f_s > 2*1/T = 2*1/0.5\ mm^{-1} = 4mm^{-1}$.

c. Anta den geometriske transformen

$$x' = ax$$

$$y' = by$$

der a og b er konstanter, og hvor x og y er heltallige pikselkoordinater og x' og y' er de transformerte koordinatene. Ved a>1 og/eller b>1 kan man ende opp med aliasingproblemer. Vil man kunne unngå dette ved å benytte høyere ordens interpolasjoner, for eksempel bikubisk interpolasjon? Forklar.

Selv ved "perfekt" interpolasjon vil man kun ende opp med det kontinuerlige originalbildet før man plukker nye samples.
Så, hvis resamplingen gir aliasingproblemer i det kontinuerlige (originale) bildet, vil man følgelig ha aliasingproblemer uansett hvor god interpolasjonsmetode man har. Aliasingproblemene knyttet direkte til interpolasjonen blir derimot redusert ved mer "ideelt" filter (høyere ordens interpolasjon).

Oppgave 3

a. Forklar hvorfor man ved bruk av diskret histogramutgjevning generelt ikke får et helt flatt histogram.

En ren gråtonetransform kan kun ”flytte på histogramsøylene”, og ikke dele søylene opp, som ville vært nødvendig for å danne et flatt histogram. For å ”splitte opp en søyle” må noen piksler gjennomgå en annen gråtonetransform enn andre, og altså bryte med at vi benytter en global transform.

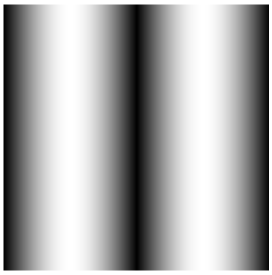
Oppgave 4

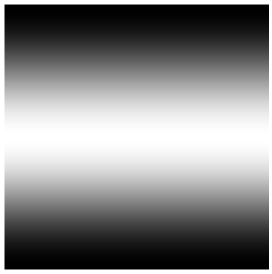
a) Hva sier konvolusjonsteoremet?

Konvolusjon i bildedomenet ⇔ multiplikasjon i frekvensdomenet (og omvendt)

b) Under ser dere to konvolusjonskjerner og deres frekvensrespons (spektre). Fremvisningen er slik at null-frekvensen er midt i bildet og lyst indikerer høy verdi mens mørkt indikerer lav verdi. Hvilke konvolusjonskjerner og spektre hører sammen? Begrunn svaret.

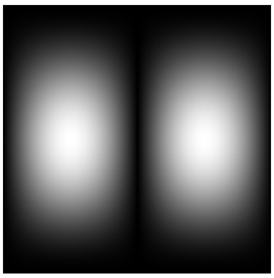
$$\begin{bmatrix}1\\2\\1\end{bmatrix} \qquad \begin{bmatrix}1 & 0 & -1\end{bmatrix}$$





Vertikalt lavpass, og horisontal båndpass. Kjerne 1 hører til spekter 2.

c) Hvis vi multipliserer de to filtrenes Fourier-transformer element for element får vi spekteret:



Hvilken konvolusjonskjerne svarer dette spekteret til?

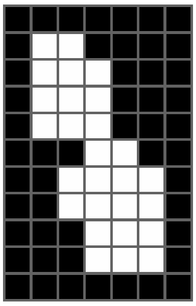
Fra konvolusjonsteoremet vil denne multiplikasjonen tilsvare en konvolusjon av filterkjernene, altså ender vi opp med en horisontal Sobel-operator.

d) Hvordan vil gråtonetransformen g(x,y) = f(x,y)+c, der c er en konstant, virke inn på spekteret til bildet?

Siden transformen kan sees på som et ortogonalt basis-skifte, og en konstant er uttrykt ved én av de nye basisbildene, er alt som endres koeffisienten til dette basisbildet, altså (den reelle) koeffisienten til frekvens (0,0). Alternative forklaringsmåter, evt regning, er selvfølgelig ok.

Oppgave 5

La hvitt være 1 og svart være 0 i bildet under. Anta at vi benytter et 3x3 kvadratisk strukturelement med origo i midten.

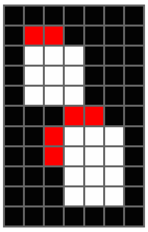


a. Hva vil slutt-resultatet være ved gjentatte morfologiske dilasjoner av bildet over?

Svar: Forgrunnen vil vokse til all struktur i bildet forsvinner. Kun bestående av 1-ere. (I 2007-eksamenen spurte vi om gjentatte erosjoner)

b. Utfør og vis resultatet av en morfologisk åpning. (Morfologisk åpning er morfologisk erosjon etterfulgt av morfologisk dilasjon.) Hva blir sluttresultatet ved gjentatte morfologiske åpninger?

Svar: Piksler merket med rødt vil forsvinne.



(Figuren over er en del av løsningsforslaget)

Morfologisk åpning er idempotent, så gjentagelser av morfologisk åpning gir ingen endring. Hvis de svarer at de har utført en åpning til og sett at det gir uendret resultat, og derav slutter at gjentatte åpninger ikke endrer resultatet er det også ok. (Må ikke bruke ordet idempotent).

Oppgave 6

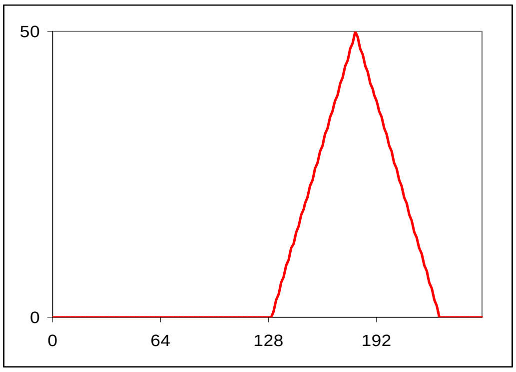
Et 8-bits bilde med størrelse 50x50 piksler består av **diagonale** linjer. Linjene går parallelt med diagonalen som går fra nedre venstre hjørne til øvre høyre hjørne i bildet. Linjene øker lineært i gråtone, første stripe har gråtone 131, neste diagonale stripe har gråtone 132 osv. Dette vil vi kalle bakgrunnsbildet.

Hvis vi zoomer inn i øvre venstre hjørne av dette bildet, ser det slik ut:

131	132	133	134	...			
132	133	134	...				
133	134	...					
134	...						
...							

a. Skisser histogrammet til dette bildet.
   Hvilken gråtone forekommer hyppigst, og hvor mange ganger forekommer denne?

   Histogrammet til bakgrunnsbildet ser slik ut:



Histogrammet stiger lineært fra 0 ved gråtone 130 til 50 ved gråtone 180, og faller lineært til 0 ved gråtone 230.

Diagonalen i bildet har gråtone 180, og det er 50 piksler på diagonalen, altså er gråtone 180 den som forekommer hyppigst: 50 ganger.

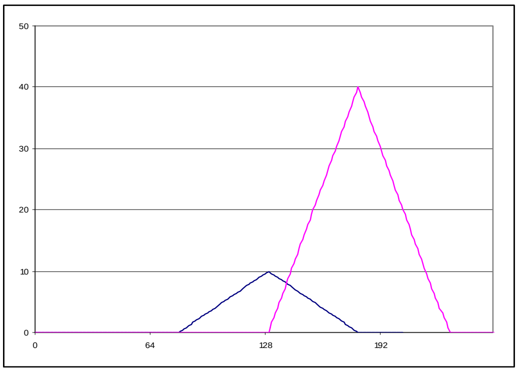
b. Anta at en mørk tekst legges oppå denne bakgrunnen. Tekst-pikslene dekker 1/5 av bildet på en slik måte at for hver diagonale stripe skifter 1/5 av pikslene i bakgrunnen over til forgrunn/tekst.
Hvert tekst-piksel har en gråtone som er L gråtonetrinn lavere enn den lokale bakgrunnen. Vi antar at teksten er fordelt over hele bildet, dvs at for hver stripe er 1/5 av pikslene tekst.

Anta først at L=50. Skisser først histogrammet for teksten. Pass på eventuell normalisering/skalering.

Skisser så histogrammet for de resterende bakgrunnspikslene i samme plott som histogrammet for teksten.

I denne oppgaven ser vi bort fra effektene vi får i hjørnene der stripene er mindre enn 5 piksler lange.

Histogrammet for teksten og den resterende bakgrunnen ser slik ut for L = 50:



Histogrammet til teksten er 0 opp til gråtone 80, har en topp på 10 ved gråtone 180 – L = 130, og faller lineært til 0 ved gråtone 180. Den resterende bakgrunnen stiger fra 0 ved 130 til 40 ved gråtone 180, og faller til 0 ved gråtone 230.

c. Bildet skal terskles med en global terskel T. Også her antar vi at L=50. Skisser i plottet hvor terskelen bør være dersom antall feilklassifiserte piksler skal minimeres. Begrunn svaret ditt.

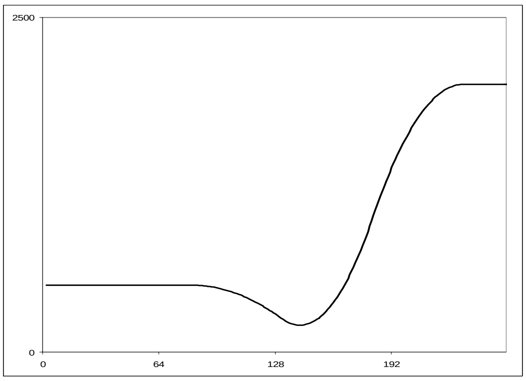
Vi vet at terskelen T som gir minst mulig samlet feil finnes der de to normaliserte fordelingene, skalert med a priori sannsynligheter, er like.

$$F \cdot f(T) = B \cdot b(T)$$

Det normaliserte tekst-histogrammet har en maksimalverdi på 10/500 ved gråtone 130. Skalert med a priori sannsynlighet har vi 10/2 500. Det normaliserte bakgrunnshistogrammet har en maksimalverdi på 40/2000 ved gråtone 180. Skalert med a priori sannsynlighet får vi 160/10 000. Forholdet mellom disse verdiene er det samme som mellom maksimalverdiene i de to histogrammene.

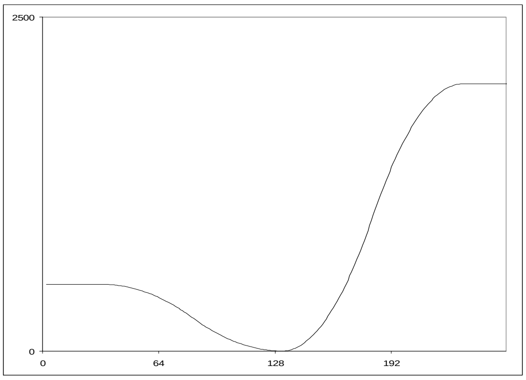
Histogrammet til teksten har en topp på 10 ved gråtone 180 – L = 130, og faller lineært til 0 ved gråtone 180. Den resterende bakgrunnen stiger fra 0 ved 130 til 40 ved gråtone 180.
Skjæringspunktet ligger da ved gråtone 140,
som altså er den terskelverdiene vi er på jakt etter.

Histogrammet over feil-klassifiserte piksler til høyre viser at vi får et minimum på 200 feilklassifiserte piksler ved gråtone 140, med asymptotiske verdier på hhv 500 og 2000 piksler ved lave og høye terskelverdier. Dette gjelder for L = 50.



d. Hvis L kan ha vilkårlige verdier kan det være mulig å finne en global terskel T som ikke gir noen feilklassifiserte piksler. Hvilke verdier for L gjør det mulig å få null feil? Begrunn svaret.

Dersom L > 99 vil de to fordelingene ikke overlappe hverandre. Siden vi summerer feilene over de to fordelingene hhv fra 0 til T og fra (T+1) til (G-1), så holder det med T > 98 for å unngå feilklassifikasjoner. Optimal terskelverdi er da 130, som vist til venstre.
Asymptoteverdiene er de samme som for L= 50.



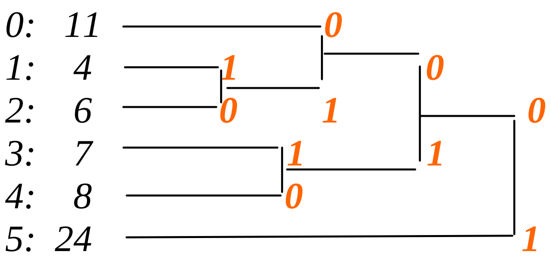
Oppgave 7

Et bilde har følgende histogram:

Gråtone	0	1	2	3	4	5
Antall piksler	11	4	6	7	8	24

a. Finn Huffman-kodene for dette bildet. Vis hele utregningen din.

Vi har et histogram som vist til venstre nedenfor, etter symbolverdien. Slår sammen først 1 og 2, så 3 og 4, så (1+2) med 0, dette med (3+4). Dermed sitter vi igjen med to grupper. Tilordner kodene 0 og 1 til de to gruppene, 0 til den mest og 1 til den minst sannsynlige av de to. Traverserer bakover, og legger til 0 og 1 i kodeordet i hvert steg,som vist nedenfor:



Merk: vi ikke har laget frekvenstabell – bare brukt histogrammet.

Huffman-kodene blir da (her er det flere mulige riktige løsninger!), sortert etter lengden på kodeordene:

5: 1
0: 000
4: 010
3: 011
2: 0010
1: 0011

b. Hvor mange bit trengs for å lagre hele bildet dersom Huffman-koding benyttes (se bort fra selve kokeboken)? Hvor mange bits besparelse ville en slik koding av bildet gitt i forhold til bruk av (vanlig) kortest mulig fastlengdekode?

Det totale antall bit i det Huffmankodete bildet blir

$$B = b_1 h_1 + b_2 h_2 + \ldots + b_N h_N = \sum_{i=1}^{N} b_i h_i$$

Der b_i er antall bit i kodeord i og h_i er histogram-verdi nr i. Vi får altså 1*24 + 3*(11+8+7) + 4*(6+4) bits = 24+78+40 =142 bit. Fastlengdekoding gir 3*60=180 ved å anta 3 bit per piksel. Altså en besparelse på 38 bit.

Takk for oppmerksomheten!