Løsningsforslag

UNIVERSITETET I OSLO

Det matematisk-naturvitenskapelige fakultet

Eksamen i: INF2310 — Digital bildebehandling
Eksamensdag: Tirsdag 29. mai 2018
Tid for eksamen: 09:00 – 13:00
Løsningsforslaget er på: **8 sider**
Vedlegg: **Ingen**
Tillatte hjelpemidler: **Ingen**

- **Det er 8 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 så fall rede for forutsetningene og antagelsene du har gjort.

- **Det er tilsammen 26 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. Histogramtilpasning

I en av ukeoppgavene denne våren implementerte vi histogramtilpasning.

a) Forklar hva som menes med histogramtilpasning.

Løsningsforslag: For et gitt bilde danner og utfører vi en gråtonetransform slik at resultathistogrammet blir mest mulig likt et ønsket sådann. (Intet krav om å definere «likhet» her, selv om det strengt tatt er veldig nærliggende.)

b) Gi en skisse til en implementasjon (pseudokode) av histogramtilpasning.

Løsningsforslag: Noe ala det presentert på slide 14 i forelesningsnotatene:

```
% Finn gråtonetransformen T (en array)
c     <- inn-bildets normaliserte kumulativt histogram
cq    <- ønskede normaliserte kumulative histogram
T     <- gråtonetransformarray, initiert til feks 0-ere
for all intensities i
   [~, argmin] = min(abs(c(i)-cq));
   T(i) = argmin;
end
% Benytt så transformen T på alle pikselverdiene
for all i,j
  g(i,j) = T(f(i,j));
end
```

2. Filtrering

Anta at vi har et 3 x 3 filter beregnet på aritmetisk veiet median-filtrering, der origo ligger sentrert i filteret, og der filterkoeffisientene ***a, b*** og ***c*** angir hvor mange ganger den underliggende filter-verdien skal gjentas ved beregningen av median-verdien.

a	b	a
b	c	b
a	b	a

a) Finn et uttrykk for verdien til koeffisienten **c**, slik at 1 piksel brede horisontale og vertikale objekter i binære bilder bevares.

Løsningsforslag: Det må være flere (repeterte) linje-piksler enn (repeterte) bakgrunns-piksler.
Altså c+2b > 2b+4a => <u>c> 4a.</u>

b) Finn et uttrykk for ***c*** slik at rettvinklede hjørner i rektangulære binære objekter med størrelse $\geq$ 2x2 bevares.

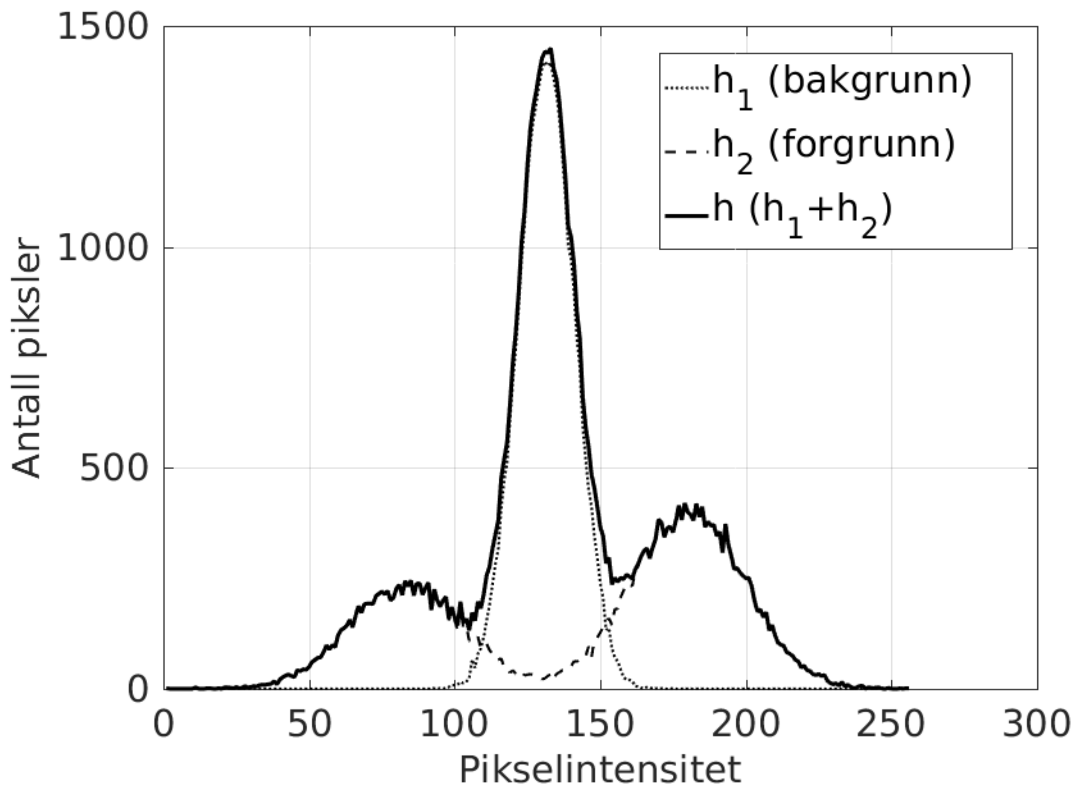
Løsningsforslag: Det må være flere (repeterte) forgrunns- enn bakgrunns-piksler.
Altså c+2b+a > 3a+2b => <u>c > 2a.</u>

c) Hvor stort er det minste kvadratiske objekt i et binært bilde som ikke forsvinner ved gjentatt filtrering med et 3x3 medianfilter der ***a=b=c=1***?
Hva med et medianfilter der ***a=1, b=2*** og ***c=3***?

Løsningsforslag:

Et 1-piksels kvadrat vil opplagt forsvinne ved første filtrering. Det samme gjelder et 2x2 kvadrat. Et 3x3 objekt mister de fire hjørnepikslene i første filtrering. Ved andre filtrering blir 4-naboene til senterpikslet borte, og dette blir så borte i tredje filtrering. Et 4x4 objekt mister hjørnepikslene i første filtrering, men er stabilt i andre filtrering.
Ved medianfiltrering med angitte repetisjoner kan vi enten sjekke manuelt, eller bare bruke resultatet fra b): Her er c>2a, og følgelig bevares hjørnene i et 2x2 objekt, siden bakgrunnen gir 7 pikselverdier = 0, mens forgrunnen gir 0 pikselverdier =1.

3. Segmentering ved terskling



I figuren over vises histogrammet, h(i), der i er pikselintensitet, til et bilde som består av to klasser: forgrunn og bakgrunn. Videre har vi tegnet inn histogrammene til disse to klassene, henholdsvis h_1 og h_2. Vi har altså $h=h_1+h_2$. Vårt mål her er å benytte pikslenes gråtoneverdier til å bestemme om en piksel skal settes til forgrunn eller bakgrunn (klassifisere).

a) Basert på h, h_1 og h_2 i figuren, for hvilke pikselintensitetsintervaller ville du klassifisert til bakgrunn og for hvilke intervaller ville du klassifisert til forgrunn? (Som alltid, forklar og begrunn svaret ditt.)

Løsningsforslag: Minimerer antall feilklassifiserte piksler ved å velge forgrunn for i der h2(i)>h1(i).

b) Ved en slik intervallinndeling som du kom frem til i oppgave a), benytt h_1 og h_2 til å gi et matematisk uttrykk for antall feilklassifiserte piksler.

Løsningsforslag: Ved regelen fra a); antall feilklassifiserte piksler ved hver intensitet summert, altså \sum_{i=0}^{G-1} \min(h_1(i), h_2(i)).

c) La $p_1(i)$ og $p_2(i)$ være det henholdsvis normaliserte bakgrunns- og forgrunnshistogrammet, og la B og F være a-priori sannsynlighet for henholdsvis bakgrunn og forgrunn.
Uttrykk h ved p_1, p_2, B, F og N, der N er antall piksler totalt i bildet.

Løsningsforslag: N*B er antall bakgrunnspiksler, N*F antall forgrunnspikler. Vi har dermed h1 = NBp1, og h2=NFp2, altså er h = h1+h2 = NBp1+NFp2.

d) La B, F, p_1, p_2 være som spesifisert i forrige deloppgave, dog denne gangen kontinuerlig modellerte sådan. (Vi antar kontinuerlige pikselintensiteter.)
Vi har i forelesningene nevnt at intensiteter, i, hvor $Bp_1(i) = Fp_2(i)$ er spesielt interessante. Hvorfor?

Løsningsforslag: Ved disse i vil det være like sannsynlig at vi har en forgrunns- som bakgrunnspiksel. Ved en antagelse om en viss glatthet i histogrammodellene våre vil vi altså her endre sannsynlig klasse; og altså vil disse punktene være tersklingskandidater.

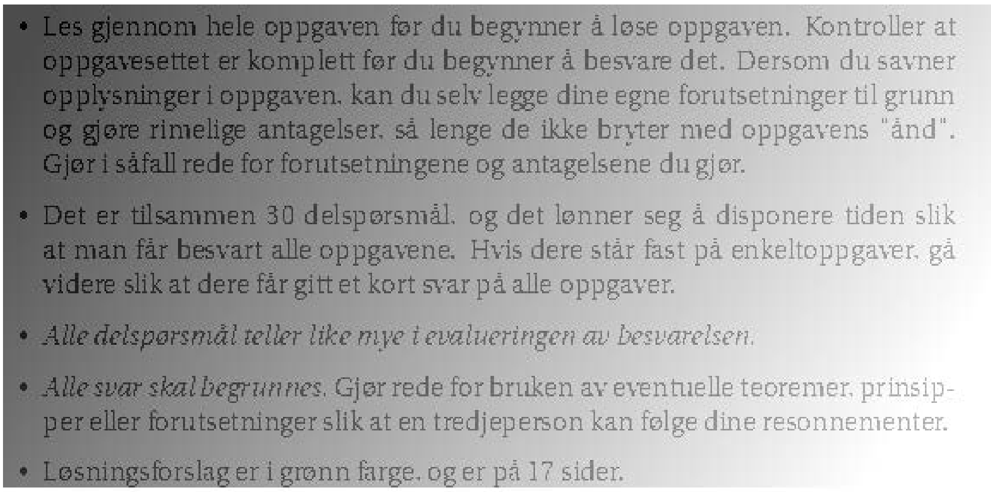
e) For et gitt vilkårlig slikt to-klasse bilde, hvilken tilleggsinformasjon må vi ha for å kunne trekke h_1 og h_2 ut i fra bildet?

Løsningsforslag: Vi må ha riktig segmentering av bildet, det vi egentlig søker å finne.

f) For bildet som omtales i denne oppgavens introduksjonstekst, anta at vi ikke har h_1 og h_2 tilgjengelig, kun bildet og dets histogram, h. Ville Riddler og Calvards algoritme (jfr. k-means) for automatisk terskling trolig gi et tilfredstillende resultat på dette bildet?

Løsningsforslag: Nei. Histogrammene for forgrunn og bakgrunn tilfredstiller ikke på nær antagelsene bak den nevnte algoritmen; som strengt tatt er at hver klasse er normal-fordelt, og med (omtrent) lik varians og a-priori sannsynlighet.

4. Mer segmentering



- 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 er tilsammen 30 delspørsmål, og det lønner seg 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 delspørsmål teller like mye i evalueringen av besvarelsen.*
- *Alle svar skal begrunnes.* Gjør rede for bruken av eventuelle teoremer, prinsipper eller forutsetninger slik at en tredjeperson kan følge dine resonnementer.
- Løsningsforslag er i grønn farge, og er på 17 sider.

Bildet i figuren vist over inneholder tekst med varierende bakgrunns-intensitet.

a) La oss anta at vi er interessert i å klassifisere pikslene i bildet over til enten å være en del av forgrunnen (teksten) eller bakgrunnen, altså segmentere bildet.
Gi (minst) et forslag til hvordan du ville gått frem i dette og lignende tilfeller.

Løsningsforslag: Med tankegangen/notasjonen fra forrige oppgave: h_1 og h_2 overlapper (nærmest fullstendig), så selv med riktig h_1 og h_2 kan vi ikke finne en enkel global terskel.

Studentene bør komme opp med et begrunnet svar hvor lokal/adaptiv terskling og/eller preprosessering er med.

5. Diskret Fourier-transform

a) Anta at vi har et NxN-bilde, f, og at vi har gjort en diskret Fourier-transform (i praksis en FFT) av dette bildet og lagret resultatet i NxN-matrisen F, der F(0,0) gir nullfrekvensen («DC-komponenten»).
Hvis nå alle elementene i F er null bortsett fra F(4,3) og F(N-4, N-3), beskriv hvordan bildet ser ut.

Løsningsforslag: Et rent «bølgebilde» hvor vi i den ene retningen har 4 repetisjoner av en ren sinusbølge og i den andre retningen 3 repetisjoner. Fasen er ukjent..

b) For et gitt NxN-bilde, g, ønsker vi å beregne den diskrete Fourier-transformen G, men vi har ingen programpakker tilgjengelig til å utføre den diskrete Fourier-transformen for oss. Vi har dog de to funksjons-kallene getCosImage(i,j,N) og getSinImage(i,j,N) som gir oss henholdsvis cosinus- og sinus-basisbilder av størrelse NxN for frekvens (i,j) (antall repetisjoner per N-te piksel i hver retning/dimensjon)
Skisser hvordan du ville benyttet disse to funksjons-kallene for å danne den ønskede NxN-matrisen G.

Løsningsforslag: En mulig løsning:

```
for each i,j
  c = getCosImage(i,j,N);
  s = getSinImage(i,j,N);
  G(i,j) = sum(g.*c) – complex(0,1)*sum(g.*s);
```

der .* menes elementvis multiplikasjon, og summasjonen er over begge dimensjonene. (Om studenten setter pluss eller minus foran det komplekse leddet er ikke vesentlig her.)

6. Histogram, Huffman-koding og entropi

La oss anta at vi har følgende 4x4 utsnitt av et 3-bit/piksel gråtonebilde:

1	2	3	4
0	5	6	7
7	7	7	6
5	4	3	2

a) Finn det normaliserte histogrammet, *p*, til bildeutsnittet.

Løsningsforslag: p = 1/16 [1 1 2 2 2 2 2 4]

b) Vis hvordan du finner kodeboken for en Huffman-koding av bildeutsnittet.

Løsningsforslag: Kodeboken blir:

7	10
2	010
3	011
4	000
5	001
6	110
0	1110
1	1111

c) Hva blir kompresjonsraten, når vi ser bort fra at kodeboken også må sendes.

Løsningsforslag: Den komprimerte sekvensen blir 1110010011000111100111010101010110001000011010, dvs 46 bits
(Vi får samme resultat ved summering av antall bits per kodeord, veiet med histogrammet 2*4*1 + 5*3*2 + 1*2*4 = 8+30+8 = 46).
Det ukomprimerte, naturlig binærkodede utsnittets 16*3 bits = 48 bits
Dette gir en (beskjeden) kompresjonsrate på C=48/46 ≈ 1.04

d) Vis hvordan du beregner entropien til det originale bildeutsnittet, og forklar hvorfor entropien her er lik det gjennomsnittlige antall bit per piksel etter kompresjon.

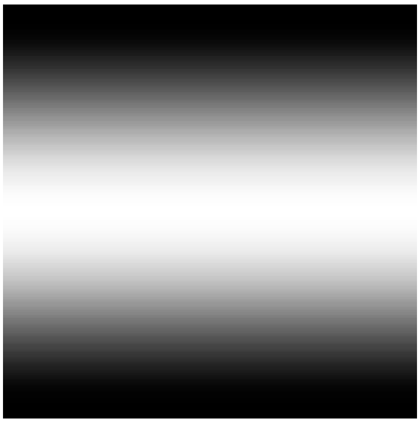
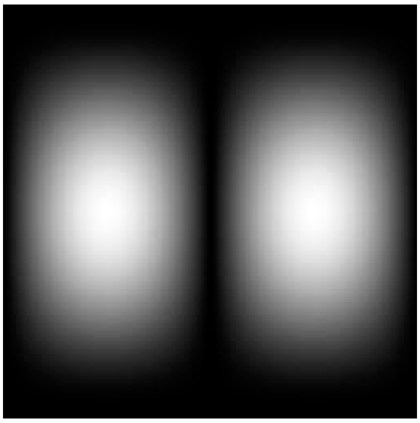
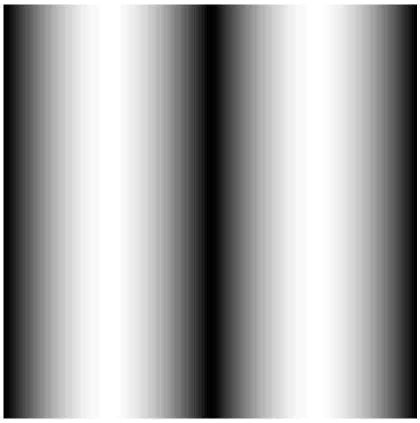
Løsningsforslag: Entropien blir H=2.875, og siden alle sannsynlighetene er inverse toerpotenser, får vi H = G.

7. Konvolusjon og frekvensrespons

a) I lys av diskrete, todimensjonale bilder; hva sier konvolusjonsteoremet?

Løsningsforslag: (Sirkel)konvolusjon i bildedomenet <=> multiplikasjon i frekvensdomenet, og omvendt.

b) I figuren under vises spektrene til tre konvolusjonskjerner. Nullfrekvensen (u=0, v=0) er lagt til midten av bildene. Sort indikerer lav verdi og hvitt høy verdi. En av konvolusjonskjernene er resultatet av en konvolusjon av de to andre – hvilket?

A	B	C
		

Løsningsforslag: Ved konvolusjonsteoremet vil en konvolusjon av filterkjernene tilsvare en elementvis multiplikasjon i frekvensdomenet. Altså, vi søker et spekter som er et elementvis produkt av de to andre. Vi ser at spekter B er et slikt produkt av A og C.

c) En implementasjon av et gitt konvolusjonsfilter kan gjøres i frekvensdomenet ved å transformere både bildet og filterkjernen med en diskret Fourier-transform, gjøre en elementvis multiplikasjon, for så å transformere tilbake til bildedomenet. Hva må vi passe på om vi ønsker et resultat som er mest mulig identisk med en "vanlig" konvolusjonsimplementasjon hvor det har blitt benyttet en implisitt nullutvidelse for å håndtere bilderanden (bildekantene)?

Løsningsforslag: En slik elementvis multiplikasjon tilsvarer en sirkelkonvolusjon i bildedomenet. Altså er vi nødt til å nullutvide både filterkjernen og bildet for å unngå at det implisitt benyttes "symmetrisk padding". (En av de obligatoriske oppgavene våren 2018 gikk eksplisitt på denne problemstillingen.)

d) Når vi designer filtre i frekvensdomenet, har vi nevnt at vi må sørge for at vi har konjugert-symmetriske koeffisienter, altså at H(i,j) = H*(-i,-j), der H er filteret vårt i frekvensdomenet, og i og j er frekvenskomponentene i henholdsvis vertikal og horisontal retning. (Ved kun bruk av reelle koeffisienter har vi dog full symmetri: H(i,j) = H(-i,-j)).
Hvorfor har vi dette kravet?

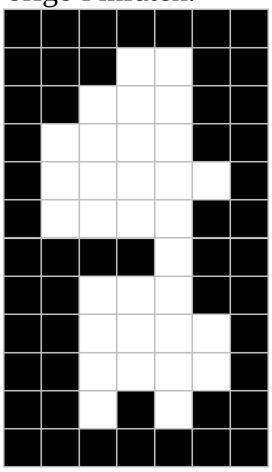
Løsningsforslag: Vi vet at ethvert reelt bilde gir opphav til en Fourier-transform som er konjugert-symmetrisk. Ved konvolusjonsteoremet må H også kunne representeres som en kjerne, h, i et konvolusjonsfilter. Om H ikke er konjugert-symmetrisk vil h måtte inneholde komplekse tall. En konvolusjon med h vil følgelig kunne gi resultatbilder med komplekse pikselverdier.

e) Ved bruk av et «ideelt» lavpassfilter vil det typisk forekomme «ringing». Hva menes med «ringing», og hva skyldes dette fenomenet?

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 i følge 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 typisk få ringing.

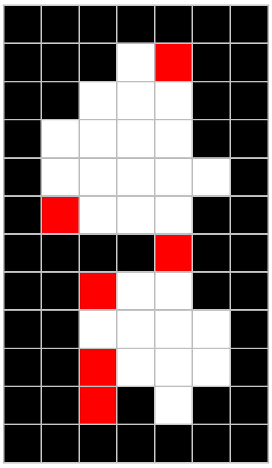
8. Matematisk morfologi

La hvitt være 1 og svart være 0 i bildet under. Anta at vi benytter et 3x3 «pluss-formet» strukturelement med origo i midten.



a) Utfør stegene i og vis resultatet av en morfologisk åpning.

Løsningsforslag: Morfologisk åpning utføres som erosjon fulgt av dilasjon.
Piksler merket med rødt i figuren under vil forsvinne.



(Figuren over er en del av løsningsforslaget)

b) Hvis man etter en segmentering sitter igjen med et binært bilde hvor objektene har litt rufsete kanter og små, uønskede hull spredt rundt omkring, hvilke morfologiske operasjoner ville du da benyttet for å "rengjøre" bildet?

Løsningsforslag: Morfologisk lukking består av dilasjon – som vil fylle igjen små hull i objektene og utvide alle objektene, fulgt av erosjon – som vil krympe objektene og utvide de hullene som ikke ble helt fylt, og dermed er de minste hullene blitt fylt, og de minste rufsene langs objekt-kantene er borte.

c) Hvilken eventuelt repetert morfologisk operasjon vil du bruke for å fjerne små objekter i et binært bilde. Forklar og begrunn svaret!

Løsningsforslag: Erosjon vil krympe alle objektene, og de minste objektene vil da bli helt borte, mens større objekter blir litt mindre. Dilasjon av resultatet vil ikke gjenskape de små objektene som ble borte i erosjonen, men vil stort sett bringe de andre objektene tilbake til den form og størrelse de hadde. Dette er morfologisk åpning.
Gjentatt åpning med samme strukturelement gir ingen endring (idempotens), så hvis man vil repetere åpning må man endre størrelse/form på strukturelementet.

TAKK FOR OPPMERKSOMHETEN!