Løsningsforslag

UNIVERSITETET I OSLO

Det matematisk-naturvitenskapelige fakultet

Eksamen i :	INF2310 — Digital bildebehandling
Eksamensdag :	Fredag 29. mars 2019
Tid for eksamen :	14:30 – 18:30 (4 timer)
Løsningsforslaget er på :	8 **sider**
Vedlegg :	**Ingen**
Tillatte hjelpemidler:	**Godkjent kalkulator**

- 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 **24** 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. Romlig oppløsning og sampling

a) For å karakterisere/beskrive avbildningssystemer benyttes ofte konseptet punktspredningsfunksjon. Forklar hva som menes med en punktspredningsfunksjon.

SVAR: En punktspredningsfunksjon beskriver hvordan ett enkelt punkt fremstår etter avbildning, altså typisk i bildeplanet før sampling.

b) Denne punktspredningsfunksjonen henger tett sammen med et vanlig mål på den romlige oppløsningen til et avbildningssystem. Forklar målet og beskriv sammenhengen mellom dette målet og bredden/utstrekningen på punktspredningsfunksjonen.

SVAR: Et vanlig mål på romlig oppløsning er hvor nærme to "punkter" (strukturer med nærmest ingen utstrekning) kan være og hvor man fortsatt kan skille de fra hverandre i bildet. Jo bredere punktspredningsfunksjon jo lengre må disse punktene stå fra hverandre for ikke å smøres inn i hverandre.

c) Anta at et avbildningssystem har oppgitt en høyest mulige romlig frekvens på $f_{max} = 5$ mm^{-1}. Hvis man benytter dette systemet til å avbilde "punkter" (eller bittesmå punktlignende strukturer nærmest uten utstrekning), hvor nærme kan man plassere disse punktene og fortsatt være i stand til å skille de fra hverandre i det resulterende bildet?

SVAR: Anta at man plasserer en serie slike punkter med innbyrdes avstand D. Den minste perioden i bildet er T_{min} = 1/f_{max} = 1/5 mm. Vi må ha at D >= T_{min}, siden om D < T_{min} og man kunne skille punktene fra hverandre, så vil jo D være minste periode i bildet og ikke T.

d) La oss så anta at vi vil digitalisere bildene fra avbildningssystemet beskrevet i deloppgave c) ved sampling med et regulært, rektangulært rutenettmønster (grid). Hvor tett må disse samplingspunktene (pikselpunktene) ligge?

SVAR: Den minste perioden i bildet er T_{min} = 1/f_{max} = 1/5 mm. La T_s være samplingsperioden (1/samplingsfreksenen). Samplingsteoremet krever at T_s < 1/2 T_{min} = 1/10 mm.

e) Anta så at vi har samplet bildet med et rektangulært rutemønster med en tetthet som akkurat oppfyller kravet du kom frem til i deloppgave d). Vi ønsker så å redusere bildestørrelsen ved å velge ut annenhver piksel i hver akse. Hva gjør dette siste steget med den effektive samplingsraten? Hva vil en slik reduksjon av bildestørrelsen kunne medføre?

SVAR: Vi ender med en dobling av avstanden mellom sampelpunktene, altså en halvering av samplingsfrekvensen. Vi vil være utsatt for mulig aliasing.

f) Et alternativ til å kun plukke ut annenhver piksel, jfr. deloppgave e), er å heller la de pikslene man beholder bestå av et gjennomsnitt av de nærmeste 3x3 pikslene fra det originale bildet. Hvorfor vil dette være en bedre løsning?

SVAR: Dette er det samme som å først gjøre en lavpassfiltrering (med et 3x3 gjennomsnittsfilter), for så å resample bildet med en halv samplingsrate. Lavpassfiltreringen hjelper til med å dempe/fjerne frekvenser som den nye, lave samplingsraten ikke håndterer. Dette filteret kalles ofte å gjøre en anti-aliasingfiltrering.

g) Anta så at vi må velge mellom to hyllevareløsninger for å kvantisere pikslenes intensitet/gråtone. Den ene løsningen benytter B bits per piksel, mens den andre benytter dobbelt så mange bits per piksel. Hvor mange flere kvantiseringsnivåer vil den andre løsningen kunne gi oss?

SVAR: 2^(2*B)-2^B.

2. Affine, geometriske transformer

Vi har jobbet en del med affine, geometriske transformer. Disse transformene kan skrives på følgende måte, der et koordinat/romlig punkt (x,y) blir transformert til (x',y'):

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Koeffisientene a_0, a_1, a_2 og b_0, b_1, b_2 bestemmer altså transformen.

a) Hvilke vanlige bildeoperasjoner kan man oppnå ved å benytte slike transformer?

SVAR: Translasjon, skalering, refleksjon, "shearing", rotasjon, og et vilkårlig sett med kombinasjoner av disse.

b) Vis hvordan en hel sekvens med slike affine transformer kan uttrykkes som én enkelt affin transform.

SVAR: Jfr. ukeoppgaver uke 2, oppgave 1.

c) Et alternativ til å bestemme en slik affin transform ved å sammenstille (som du beskrev i deloppgave b)) et sett med "grunnleggende" affine tranformer, er å benytte seg av kontrollpunkter. Beskriv hva som menes med kontrollpunkter, og forklar hvordan slike punkter kan benyttes til å bestemme koeffisientene i en affin transform.

SVAR: Kontrollpunktene bestemmer et sett med til og fra-koordinater, altså hvor et sett med punkter i bildet skal flyttes til. De seks koeffisientene i den affine transformen kan entydig bestemmes med tre kontrollpunkter/punktpar. Sett gjerne opp ligningene. Ved flere punktpar enn man trenger for entydig å bestemme transformen, så har vi i dette kurset sett på å finne settet med koeffisienter som minimerer diskrepansen kvadrert.

d) Ved (praktisk) bruk av slike transformer har vi som oftest behov for å interpolere pikselverdier. Hvorfor?

SVAR: Etter en slik transform ønsker man ofte å igjen ha et bilde med sampelpunkter i et regulært rutenettmønster. Vi "resampler" altså bildet vårt, og vi vil ofte trenge å "plukke" nye piksler, og disse nye punktene er som oftest ikke _akkurat_ der "inputbildet" vårt allerede har sampelpunkter. Derfor må vi bestemme koordinatets farge/gråtone basert på kjente piksler i nærheten -> interpolere.

3. Morfologi

a) "Hit / miss"-transformasjonen er definert ved

$$A(*)S = A(*)\,[S_1, S_2] = (A \theta S_1) \cap (A^C \theta S_2)$$

der A er et binært bilde, og strukturelementet S er definert ved et par $[S_1, S_2]$ av binære strukturelementer som ikke har noen felles elementer.
Anta at S_1 og S_2 er symmetriske, har felles origo og er hverandres komplement.
Bruk dualiteten mellom erosjon og dilasjon til å uttrykke A(*)S ved bare A og S_1, altså uten å bruke A^C og S_2.

Svar: Det er bare andre ledd i snittet, $(A^C \theta S_2)$, som vi trenger å gjøre noe med.
Vi har dualiteten $A \oplus S_2 = \left(A^C \,\theta\, \hat{S}_2\right)^C$.
Siden S_2 er symmetrisk kan vi droppe refleksjonen.
Dessuten er komplementet til S_1 lik S_2, og dermed har vi

$$A(*)S = A(*)\,[S_1, S_2] = \underline{\underline{(A \theta S_1) \cap (A \oplus S_1^C)^C}}$$

b) Beskriv hva man kan oppnå med operasjonene lukking og åpning, forklar hvordan den ene operasjonen kan utføres ved hjelp av den andre hvis struktureringselementet er symmetrisk, og forklar hva nytten av dette er.

SVAR: Lukking kan lukke en åpning mellom to strukturer som bare er adskilt med et lite gap, uten at de to strukturene vokser, og fyller mindre hull og innbuktninger.
Åpning kan skape en åpning (mellomrom) mellom to strukturer som bare henger sammen ved en tynn «bro», uten å krympe de to strukturene, og fjerner små og utstikkende strukturer.
Lukking er en **dual** operasjon til **åpning** med hensyn til komplementering og reflektering (180° rotering), og omvendt. Pga symmetri kan vi droppe reflektering.
Lukking kan utføres ved å komplementere bildet f, åpne det med strukturelementet S, og ta komplementet av resultatet: $f \bullet S = (f^C \circ S)^C$
Tilsvarende for åpning: $f \circ S = (f^C \bullet S)^C$
<u>Vi kan altså utføre begge operasjonene med kode bare for den ene, hvis vi har kode for å komplementere et binært bilde</u>.

4. Medianfiltrering

Gitt et 5x5 piksels gråtonebilde F:

1	2	1	8	9
2	3	2	6	8
2	7	8	9	7
1	3	9	3	4
2	1	2	3	3

a) Utfør en medianfiltrering av F med en kvadratisk 3x3 filtermaske med origo i midten. Vis tabellene for hvert ut-piksel, og utfør filtreringen bare der hele filteret ligger innenfor bildet.

SVAR: Vi får følgende 9 sorterte tabeller a 3x3=9 tall:

1 1 2 2 **2** 2 3 7 8	1 2 2 3 **6** 7 8 8 9	1 2 6 7 **8** 8 8 9 9
1 2 2 2 **3** 3 7 8 9	2 3 3 3 **6** 7 8 9 9	2 3 4 6 **7** 8 8 9 9
1 1 2 2 **2** 3 7 8 9	1 2 3 3 **3** 7 8 9 9	2 2 2 2 **4** 7 8 9 9

og medianene

2	6	8
3	6	7
2	3	4

b) Ved vektet medianfiltrering forteller vektene i filtermatrisen hvor mange ganger den underliggende pikselverdien skal kopieres over i en endimensjonal tabell før vi finner medianen.
Utfør en vektet medianfiltrering av bildet F med filteret.
Vis tabellene for hvert ut-piksel.
Sammenlign de to medianfiltreringene og beskriv effekten av den vektede medianfiltreringen.

0	1	0
1	3	1
0	1	0

SVAR: Vi får følgende 9 sorterte tabeller a **bare 4+3=7** tall:

2 2 2 **3** 3 3 7	1 2 2 **2** 3 6 8	2 6 6 **6** 8 8 9
2 3 3 **7** 7 7 8	2 7 8 **8** 8 9 9	3 6 7 **8** 9 9 9
1 1 3 **3** 3 7 9	2 3 3 **8** 9 9 9	3 3 3 **3** 4 9 9

og de vektede medianene

3	2	6
7	8	8
3	8	3

Den vektede medianen som er gitt her vil faktisk bare endre piksel-verdien dersom den er ekte større (mindre) enn alle sine 4-naboer. Da blir svaret lik den nærmeste 4-naboen. Derfor vil den bare endre lokale maksima (minima).
Bare to av pikselverdiene i vår F endres.

c) Et ordinært kvadratisk 3x3 medianfilter vil fjerne hjørnepikslene i rektangulære objekter i binære bilder, mens et pluss-formet 5x5 medianfilter vil bevare dem. Anta at et vektet 3x3 filter er gitt ved heltallene a, b og c

a	b	a
b	c	b
a	b	a

Finn et uttrykk for sammenhengen mellom a, b og c slik at rettvinklede hjørner til objekter med størrelse $\geq$ 2x2 piksler i binære bilder bevares.
Blir det noen forskjell mellom konvekse og konkave hjørner? Begrunn gjerne med en liten skisse!

SVAR: Det må være flere repeterte objekt- enn bakgrunns-piksler når c ligger over et hjørnepiksel i et objekt. Altså

$c+2b+a > 3a+2b \Rightarrow c > 2a$, helt uavhengig av hva b er.

Med en liten skisse kan man demonstrere at det ikke blir noen forskjell mellom konvekse og konkave hjørner.
(Et konkavt hjørnepiksel til et objekt er et bakgrunnspiksel.)

5. Kontrastendring ved lineær (affin) transform

Vi skal her jobbe med følgende lineære (affine) gråtonemapping:

$$T[i] = ai + b$$
,

der T[i] er hva pikselintensiteten i mappes til, og a og b er koeffisienter.

Om vi antar at vi har et gråtonebilde med en gjennomsnittspikselverdi, μ, og intensitetsvarians (det kvadrerte av standardavviket) lik σ^2, vil vi etter en slik affin transform ha et nytt pikselgjennomsnitt og varians på henholdsvis (om vi ser bort fra kvantiseringsfeil og mulige intensitetsklippeeffekter) $\mu_T = a\mu + b$ og $\sigma_T^2 = a^2 \sigma^2$.

a) Beskriv først hva parametrene a og b gjør med et resultatbildes "lyshet" og "kontrast", samt hva parametrene gjør med et bildes gråtonehistogram.

SVAR: |a|>1 vil strekke histogrammet, |a|<1 vil «krympe» histogrammet, a<0 vil i tillegg speilvende histogrammet. b vil forskyve histogrammet. |a|>1 vil øke kontrasten, |a|<1 vil minke kontrasten, b>0 vil gjøre bildet lysere, b<0 mørkere, samt at også a<1 i seg selv vil gjøre bildet mørkere og a>1 lysere.

b) Ved disse ligningene oppgitt i oppgaveintroduksjonen kan vi bestemme transformkoeffisientene, a og b, som gir et spesifikt bilde et bestemt intensitetsgjennomsnitt og en spesifikk varians. Dette kan benyttes, som sett i dette kurset, til å gi en hel serie bilder samme slik intensitetsgjennomsnitt og -varians. Hva forsøker man å oppnå ved å gjøre dette? Hva er intensitetsgjennomsnittet og -variansen et mål på?

Man ønsker å gi bildene (nogenlunde) den samme lysheten (jfr. intensitetsgjennomsnittet) og kontrasten (jfr. variansen).

c) Anta at et gråtonebilde har en middelverdi lik 100 og varians lik 400, altså $\mu = 100$ og $\sigma^2 = 400$. Hvis man så ønsker at resultatbildet skal ha $\mu_T = 128$ og $\sigma_T^2 = 100$, hvilken lineær gråtonetransform ville man benyttet?

SVAR: a = \sigma_T/\sigma = sqrt(100/400) = sqrt(1/4) = 1/2.
b = \mu_T - a\mu = 128-1/2*100 = 78.

6. Histogramutjevning og histogramtilpasning

Ved histogramutjevning benytter man på et gråtonebilde en global gråtonetransform i et forsøk på å gi et resultatbilde med mest mulig flatt gråtonehistogram. Slike globale gråtonetransformer er gitt ved en funksjon T[i], for pikselintensiteter i.

a) La oss anta vi har et bilde med et gråtonehistogram som ligner på en Gauss-kurve (en klokkeform). Videre ønsker vi å utføre en histogramutjevning på dette bildet. Beskriv og skisser gråtonetransformen, T, du ville benyttet for å oppnå dette.

SVAR: Vi vet at T er gitt som en skalert versjon av bildets kumulative histogram. I vårt tilfelle får vi en "S"-form. Altså, stigningstallet til funksjonen har en klokkeform, så i midten stiger funksjonen raskt, mens den avtar på begynnelsen og slutten (for de veldig mørke og lyse pikselintensitene). De kan gjerne skissere kurven mer detaljert, for eksempel ta med at T vil ligge mellom 0 og G-1, der G er antall gråtoner i bildet.

b) Bekreft eller avkreft følgende påstand (husk at, også her, må svaret begrunnes): Om ikke et bildes gråtonehistogram allerede er helt flatt, vil man *aldri* kunne oppnå et helt flatt gråtonehistogram ved en fremgangsmåte som beskrevet i oppgaveintroduksjonen.

SVAR: Sant. Vi baserer oss på en transform, T, som flytter rundt på _hele_ histogramsøyler; _alle_ pikslene med intensitet i blir flyttet til T[i]. For at histogrammet skal være helt flatt, må det i hvert fall eksistere piksler for alle intensiteter. Med dette som utgangspunkt kan vi i hvert fall ikke få et flatt histogram om ikke alle pikselintensiteter allerede eksisterer i bildet, og om det gjør det, kan vi kun «stokke» om på histogramsøylene, og dette vil jo selvfølgelig aldri kunne gi oss et flatt histogram.

7. Konvolusjon

La intensiteten omkring pikselposisjonen (x,y) være modellert som kvadratisk i y-retningen men lineær i x-retningen:

$$f(x,y) = k_1+k_2x+k_3y+k_4y^2$$

I et lokalt 3x3 område rundt posisjonen (x,y) vil da intensitetene være

$k_1-k_2-k_3+k_4$	k_1-k_2	$k_1-k_2+k_3+k_4$
$k_1-k_3+k_4$	k_1	$k_1+k_3+k_4$
$k_1+k_2-k_3+k_4$	k_1+k_2	$k_1+k_2+k_3+k_4$

a) Vis at filtermasken

1	2	1
2	-12	2
1	2	1

gir et skalert estimat av den korrekte Laplace-verdien av denne modellen.

Svar: Den korrekte lokale Laplace-verdien av denne modellen er gitt ved (ligningen er ikke gitt i oppgaven, men dette bør kandidatene beherske):

$$-\nabla^2(f(x,y)) = -\frac{\partial^2 f(x,y)}{\partial x^2} - \frac{\partial^2 f(x,y)}{\partial y^2} = -2\,k_4$$

Den gitte filtermasken gir $8k_4$, som altså er et skalert estimat.

b) Hva oppnår vi ved å vise at dette filteret gir et estimat av den korrekte Laplace-verdien?

SVAR: Dette forteller oss at hvis innbildet lokalt – dvs innenfor et 3x3 område – ikke er mer komplisert enn at det varierer kvadratisk (som en parabel) i y-retning og samtidig varierer lineært i x-retning, så vil den gitte filtermasken gi et korrekt (skalert) estimat av den sanne andrederiverte av bildeflaten.

c) Forklar hvordan vi kan gjøre Laplace-estimatet som vi får med dette filteret (noe mer) robust for støy, og vis hvordan filtermasken ovenfor eventuelt blir modifisert.
Hva kaller vi et slikt filter?

SVAR:

- Vi kan konvolvere inn-bildet med et lavpassfilter før vi estimerer Laplace-verdien med den samme filtermasken som ovenfor.
- Eller vi kan utnytte kommutativitets-egenskapen til konvolusjonsoperasjonen, og konvolvere filtermasken med lavpassfilteret, og så anvende resultatet som et filter på bildet. F. eks. LoG-filteret:

$$\frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & -12 & 2 \\ 1 & 2 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 1 \\ 1 & 2 & 1 \end{bmatrix} = \frac{1}{16}\begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 0 & -8 & 0 & 4 \\ 6 & -8 & -28 & -8 & 6 \\ 4 & 0 & -8 & 0 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$$

Takk for oppmerksomheten !