Løsningsforslag

UNIVERSITETET I OSLO

Det matematisk-naturvitenskapelige fakultet

Eksamen i :	INF2310 — Digital bildebehandling
Eksamensdag :	Tirsdag 20. mars 2018
Tid for eksamen :	14:30 – 18:30 (4 timer)
Løsningsforslaget er på :	**7 sider**
Vedlegg :	**Ingen**
Tillatte hjelpemidler:	**Ingen**

- Det er 10 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 **22** 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. Sampling, geometrisk transform og resampling

Anta at vi har et (kontinuerlig) båndbegrenset bilde med en høyeste frekvens $f_{max} = 2\ \text{mm}^{-1}$. Det som er avbildet er bl.a. noen punktkilder som står så nær hverandre at de så vidt kan skilles fra hverandre i bildet.

a) Gi en grense for hvor tett disse punktkildene står.

Svar: Det står oppgitt hva den høyeste frekvensen i bildet er. Da har vi også den minste perioden i bildet: $T_{min} = 1/f_{max} = 0.5$ mm. Punktkildene kan ikke stå nærmere hverandre enn dette. Om de hadde gjort det, ville de ha vært opphav til en mindre minste periode (høyere høyeste frekvens) enn det som er oppgitt.

b) Hvor tett må vi sample dette bildet for å unngå aliasing?
Gi en nedre grense for *samplingsfrekvensen*, f_s.
La oss videre anta at vi har samplet bildet med en rate så vidt over denne grensen.

Svar: Sampingsteoremet krever $f_s > 2f_{max}$ for å unngå aliasing. I vårt tilfelle må vi altså ha $f_s > 2 \cdot 2\ \text{mm}^{-1} = 4\ \text{mm}^{-1}$. Vi må altså ha mer enn 4 sampler per mm.

c) Etter sampling utføres det en geometrisk transformasjon gitt ved disse likningene:

$$x' = 0.2x + 100,$$
$$y' = 0.2y + 200,$$

der x og y er koordinatene i innbildet, og x′ og y′ er de transformerte koordinatene. La oss videre anta at det benyttes en vanlig resampling ved baklengstransformasjon.
Hva vil den effektive samplingsraten være etter en slik transform, og hvilke (uønskede) effekter vil dette kunne gi opphav til?

Svar: I tillegg til en translasjon, vil transformen skalere bildet med en faktor 0.2. Ved baklengs resampling står vi igjen med de originale romlige sampelposisjonene. Etter transformen så sitter vi igjen med 1/5 av samplene i hver akse. Samplingsraten er altså 1/5 av hva den var. Vi har altså en effektiv samplingsrate som er 1/5 av den originale, og vi bryter dermed samplingsteoremets krav, som igjen setter oss i fare for aliasing.

2. Et avbildnings- og bildebehandlingssystem

Forklar eventuelle prinsipielle problemer med rekkefølgen i dette avbildnings- og bildeanalyse-systemet:

Avbildning → Sampling → Analyse av romlig oppløsning i samplet bilde → Anti-aliasing → Videre strukturanalyse av bildet

Svar: Antialiasing må skje <u>før</u> en evt. sampling/resampling. Etter sampling vil det være umulig (uten annen informasjon) å skille kunstige (grunnet aliasing) og reelle strukturer.

3. Histogrammer

La oss anta at vi har følgende 4 × 4 gråtonebilde:

1	3	2	1
5	4	5	3
4	1	1	2
2	3	2	1

Skissér gråtonehistogrammet, *h*, det normaliserte histogrammet, *p*, samt det kumulative histogrammet, *c*, til bildet.
Gi et generelt uttrykk (matematisk formel) for *p* og *c* basert på *h*.

Svar: h = [5 4 3 2 2], c = [5 9 12 14 16], p = h/16.
Evt: Vi ser at dette er et 3-bits gråtonebilde, der h = [0 5 4 3 2 2 0 0], c = [0 5 9 12 14 16 16 16], p = h/16. Uttrykkene for p og c basert på h blir

$$p(i) = \frac{h(i)}{\sum_{j=0}^{7} h(j)}, \quad c(i) = \sum_{j=0}^{i} h(j)$$

4. Kvantisering

a) Anta at vi i utgangspunktet kvantiserer til B bit per piksel i et digitalt gråtonebilde. Hvis man dobler antall bits per piksel, hvor mange flere kvantiseringsnivåer vil man da ha tilgjengelig?

Svar: Med B biter per pixel har vi 2^B kvantiseringsnivåer. Ved 2*B biter har vi 2^(2B) nivåer. Vi har altså 2^(2B)-2^B ekstra nivåer tilgjengelig ved å gå fra B til 2B biter.

b) Hvis vi antar lik sannsynlighet for alle intensiteter i input-intervallet (uniform fordeling), hvor stor endring i kvantiseringsfeil vil vi forvente hvis vi tar med ett ekstra bit (B+1)?

Svar: Ved å legge til ett ekstra bit, dobler vi antall kvantiseringsivåer. Den forventede kvantiseringsfeilen for en gitt sample vil dermed halveres. Den totale kvantiseringsfeilen, som er summen over alle pikslene, vil følgelig også forventes å bli halvert.

5. Gråtonetransform og bildeseriestandardisering

Anta gråtonetransformen T[i] = ai + b, der a og b er koeffisienter/konstanter, og i er pikselintensitet.

a) Hvilken effekt har parametrene a og b på gråtonehistogrammet til det resulterende bildet? Forklar også effekten av a og b på kontrasten og lysheten til bildet.

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.
a vil også kunne påvirke lysheten i bildet, siden ny middelverdi er gitt ved **μ_T=b+aμ.**

b) Man kan gi en serie bilder lik varians og middelverdi ved å benytte slike gråtone-transformer. Hva prøver man å oppnå ved slik standardisering av varians og middelverdi?

Svar: Variansen/standardavviket er et mål knyttet til kontrast, gjennomsnitt knyttet til bildets lyshet. Ved å standardisere disse to forsøker man altså å gi bildeserien lik kontrast og lyshet.

6. En eksponentiell gråtonetransform

Anta gråtonetransformen T[i] = exp(i).

a) Hva vil denne transformen gjøre med kontrasten i henholdsvis de mørke og lyse intensitetsintervallene? (Som for alle andre oppgaver; forklar!)

Svar: I de mørke områdene har vi et lavt lokalt stigningstall, omvendt i de lyse områdene. Vi vil altså dempe kontrasten i de mørke områdene mens vi øker kontrasten i de lyse.

7. Histogramutjevning

Hva er histogramutjevning? For et generelt bilde, beskriv hvordan man finner den gråtonetransformen som utfører en histogramutjevning.

Svar: Ved histogramutjevning benytter vi en gråtonetransform som forsøker å gjøre resultathistogrammet mest mulig flatt. På denne måten maksimerer vi kontrasten mens vi beholder gråtonerikheten.
Gråtonetransformen er en skalert versjon av bildets kumulative histogram.

8. Konvolusjon

Anta at gråtoneverdiene omkring en vilkårlig pikselposisjon (x,y) i et digitalt bilde kan modelleres som kvadratisk i x-retningen men lineær i y-retningen:

$$f(x,y) = k_1+k_2x+k_3y+k_4xy+k_5x^2$$

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

Y →

X ↓

$k_1-k_2-k_3+k_4+k_5$	$k_1-k_2+k_5$	$k_1-k_2+k_3-k_4+k_5$
k_1-k_3	k_1	k_1+k_3
$k_1+k_2-k_3-k_4+k_5$	$k_1+k_2+k_5$	$k_1+k_2+k_3+k_4+k_5$

a) Vis at filtermasken

$$\frac{-1}{4}\begin{bmatrix} 1 & 2 & 1 \\ 2 & -12 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

gir et korrekt estimat av Laplace-verdien, gitt denne gråtone-modellen.

Svar: Laplace-ligningen er ikke gitt i oppgaven, men dette bør kandidatene beherske, Slik at man kan derivere f(x,y) to ganger mhp x og y, og finne at den korrekte lokale Laplace-verdien av denne modellen er gitt ved :

$$-\nabla^2(f(x,y)) = -\frac{\partial^2 f(x,y)}{\partial x^2} - \frac{\partial^2 f(x,y)}{\partial y^2} = -2\,k_5$$

Den gitte filtermasken lagt over de oppgitte intensitetene gir også $-2k_4$, som altså er et korrekt estimat.

b) Forklar hvordan vi kan gjøre Laplace-estimatet som vi får med dette filteret 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:
  (matrisene nedenfor er en del av løsningsforslaget)

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

c) Vis at du kan lage et 5x5 LoG-lignende filter (LoG = «Laplacian-of-Gaussian») ved hjelp av DoG ("Difference of Gaussians"), ved å bruke to enkle tilnærminger til en 2D Gauss-profil med forskjellig størrelse.

Svar: Her er det mest naturlig å bruke en 5x5 og en 3x3 tilnærming til Gauss-profilen, null-utvide og skalere den sistnevnte, og få følgende filter:
(matrisene nedenfor er en del av løsningsforslaget)

$$DoG = \frac{1}{256}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 1 \\ 1 & 2 & 1 \end{bmatrix} - \frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

$$= \frac{1}{256}\begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix} - \frac{1}{256}\begin{bmatrix} 16 & 32 & 16 \\ 32 & 64 & 32 \\ 16 & 32 & 16 \end{bmatrix} = \frac{1}{256}\begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 0 & -8 & 0 & 4 \\ 6 & -8 & -28 & -8 & 6 \\ 4 & 0 & -8 & 0 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$$

d) Forklar hva vi mener med at et 2D filter er "separabelt", og vis hvordan dette DoG-filteret kan gjøres separabelt.

Svar: Et filter er separabelt hvis filtreringen kan utføres som to sekvensielle filtreringer, eller hvis vi bruker assosiativitets-egenskapen ved konvolusjon ($f*h=(f*h_1)*h_2$, hvis $h=(h_1*h_2)$).

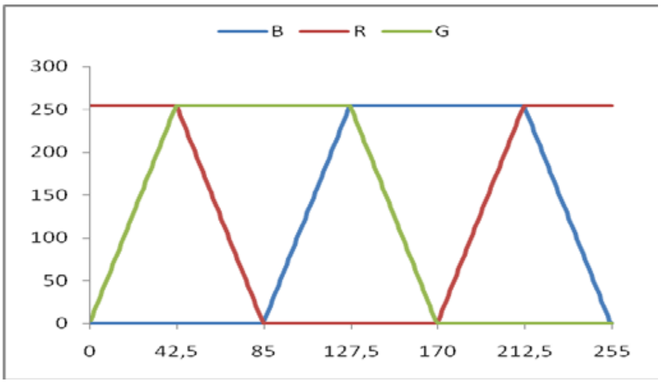
Selve DoG-filteret er ikke separabelt, men hver av komponentene som ble brukt til å produsere det, kan separeres. Dette bør kunne besvares selv om man ikke har fått resultatet ovenfor riktig. (matrisene nedenfor er en del av løsningsforslaget)

$$DoG = \frac{1}{256}\begin{bmatrix}1\\4\\6\\4\\1\end{bmatrix} * \begin{bmatrix}1 & 4 & 6 & 4 & 1\end{bmatrix} - \frac{1}{16}\begin{bmatrix}1\\2\\1\end{bmatrix} * \begin{bmatrix}1 & 2 & 1\end{bmatrix}$$

9. Fargerom

Dette var tema for siste forelesning og siste ukeoppgavesett før midtveiseksamen.

Anta at indeks (i) i en RGB pseudo-fargetabell løper fra 0 til 255, og at RGB-verdiene ligger mellom 0 og 255, som vist i figuren nedenfor.



a) Anta at vi nå skal konvertere fra RGB til HSI. Velg ut noen punkter langs i-aksen, og lag en skisse som viser hvordan intensiteten I vil forløpe.

Svar: Intensiten er gitt som (R+G+B)/3, og vil variere som en sagtakk-funksjon mellom 85 og 170 (eller 1/3 og 2/3 hvis RGB-verdiene ligger mellom 0 og 1)..

b) Metningen, S, er gitt som 1-3min(R,G,B)/(R+G+B). Velg ut noen punkter langs *i*-aksen, og forklar hvordan S vil forløpe som funksjon av indeksen *i*.

Svar: I alle punktene langs *i*-aksen vil metningen være lik 1, siden vi alltid har min(R.G,B) = 0.

c) Anta at du har et RGB-bilde der det dynamiske området ikke er godt utnyttet, og at du derfor vil gjøre en histogram-transformasjon.
Hvorfor er det en fordel å gjøre dette i et perseptuelt fargerom, og ikke i RGB?

Svar: Histogram-utjevning eller andre histogram-transformasjoner vil endre på forholdet mellom farge-komponentene, og kan gi farge-stikk i ut-bildet. En transformasjon fra RGB til for eksempel IHS, histogram-transformasjon av I-komponenten, og så transformasjon tilbake til RGB vil unngå dette.

10. Canny's kantdetektor

Et av hovedtemaene i årets første obligatoriske oppgave var en implementasjon av Canny's kantdeteksjonsalgoritme.

a) Beskriv hovedpunktene i Canny's kantdeteksjonsalgoritme

Svar:

1. Lavpassfiltrér bildet f(x,y) med et Gauss-filter (med gitt σ).
2. Finn gradient-magnituden g(x,y) og gradient-retningen θ_g, f.eks. med de to maskene i Sobel-operatoren.
3. Tynning av gradient-magnitude ortogonalt på kantene i bildet
   F.eks.: Hvis en piksel i gradient-magnitude-bildet har en 8-nabo i eller mot gradient-retningen med høyere verdi, så settes pikselverdien til 0.
4. Hysterese-terskling (to terskler, T_h og T_l):
   a) Merk alle piksler der $g(x,y) \geq T_h$
   b) For alle piksler der $g(x,y) \in [T_l, T_h)$:
      Hvis (4 eller 8)-nabo til en merket piksel, så merkes denne pikselen også.
   c) Gjenta fra trinn b) til konvergens.

b) Gitt et gråtonebilde f(x,y). Hva inneholder ut-bildet G(x,y) når

$$G(x,y) = \max_{k=0,\ldots,3}\{|g_k(x,y)|\} \quad , \text{ der } \quad g_k(x,y) = f(x,y)*h_k \ , k = 0,1,2,3$$

og $h_0, \ldots, h_3$ er konvolusjonsfiltrene

$$h_0 = \begin{bmatrix} 1 & \sqrt{2} & 1 \\ 0 & 0 & 0 \\ -1 & -\sqrt{2} & -1 \end{bmatrix}, h_1 = \begin{bmatrix} \sqrt{2} & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -\sqrt{2} \end{bmatrix}, h_2 = \begin{bmatrix} 1 & 0 & -1 \\ \sqrt{2} & 0 & -\sqrt{2} \\ 1 & 0 & -1 \end{bmatrix}, h_3 = \begin{bmatrix} 0 & -1 & -\sqrt{2} \\ 1 & 0 & -1 \\ \sqrt{2} & 1 & 0 \end{bmatrix}$$

for retningene $\quad \theta_k = k\frac{\pi}{4}, \quad k = 0,1,2,3.$

Svar: Dette er en 4-retnings "kompass-operator" basert på Frei-Chen-operatoren, produsert ved sirkulær rotasjon mot urviseren. For hvert piksel i bildet finner operatoren absolutt-verdien av den lokale gradient-magnituden i den av de 4 laveste 8-retningene der absolutt-verdien av gradientmagnituden er størst.
Merk at vi trenger ikke å estimere gradientmagnituden i alle 8 retningene, siden fire av resultatene finnes ved å invertere resultatet fra de fire andre (motsatte) retningene.

c) Hvilket trinn i Canny-algoritmen kan denne operatoren erstatte?

Svar: Den kan erstatte beregning av gradient-magnitude (som består av h_0 og h_2, og erstatter kvadrering-sum-kvadratrot.

d) Anta at vi også tar vare på den verdien av k som gir maksimal absoluttverdi av $g_k(x,y) = f(x,y)*h_k$, (k = 0,1,2,3), for hvert piksel i f(x,y), og lagrer denne i et ut-bilde θ(x,y). Hva inneholder dette ut-bildet, hvilket trinn i Canny-algoritmen kan dette erstatte, og hvilke matematiske operasjoner unngår vi med dette??

Svar: Ut-bildet vil inneholde verdiene [0,1,2,3] – altså 2 bits per piksel, og kan erstatte beregning av gradient-retning ved $\arctg(g_2/g_0)$ med en simpel max-operator, der vi også slipper avrunding av resultatet av arctg til nærmeste 45 grader. Vi kommer altså enklere til tynningen av gradientmagnitude-bildet i eller mot den 8-retningen der gradienten er størst.
Ta gjerne fram koden din fra Oblig1, og implementer disse forenklingene!

TAKK FOR OPPMERKSOMHETEN!