

UNIVERSITETET I OSLO

Det matematisk-naturvitenskapelige fakultet

Eksamen i : INF2310 — Digital bildebehandling

Eksamensdag : Tirsdag 27. mars 2012

Tid for eksamen : 15:00 – 19:00

Løsningsforslaget er på : **9 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 såfall rede for forutsetningene og antagelsene du gjør.
- Det er tilsammen 20 delspørsmål, og det lønner seg å disponere tiden slik at man får besvart alle oppgavene. Hvis du står fast på enkeltoppgaver, gå videre slik at du får gitt et kort svar på alle oppgaver.
- **Alle svar skal begrunnes.** Gjør rede for bruken av eventuelle teoremer, prinsipper eller forutsetninger slik at en tredjeperson kan følge dine resonnerer.

1. Kvantisering

- a) Anta at vi har B biter per piksel i et bilde. Sett opp et uttrykk for hvor mange kvantiseringsnivåer vi vil miste hvis vi dropper de 3 minst signifikante bitene i hvert piksel.

Hvor mange nivåer mister vi hvis utgangspunktet er $B = 9$?

Med B biter per piksel har vi 2^B kvantiseringsnivåer. Etter at vi har fjernet de 3 minst signifikante bitene i hvert piksel har vi 2^{B-3} . Altså får vi $2^B - 2^{B-3}$ færre kvantiseringsnivåer. $B=9$ gir $512 - 64 = 448$.

Vi kan også si at vi mister halvparten av nivåene for hver bit vi tar bort.

Altså mister vi $256 + 128 + 64 = 448$ nivåer.

- b) Anta at vi i utgangspunktet kvantiserer til B biter per piksel. 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 3 biter mindre (fra B til $B-3$)?

Vi dobler kvantiseringsfeilen for hvert bit vi tar bort.

Kvantiseringsfeilen blir altså $2^3 = 8$ ganger så stor.

2. Sampling og geometrisk transform

Anta at vi har et (kontinuerlig) båndbegrenset bilde med en høyeste frekvens $f_{\max} = 1/3 \text{ mm}^{-1}$.

- a) Det som er avbildet er bl.a. noen punktkilder som står så nær hverandre at de så vidt kan skilles fra hverandre i bildet. Gi en nedre grense for hvor tett disse punktkildene står.

Plasser en mengde tenkte punktkilder med avstand P i mellom seg på en rekke. Siden punktene kan skilles får vi da et signal med periode P i bildet.

P kan ikke være mindre enn minste periode i bildet: $P \geq P_{\min} = 1/f_{\max} = 3 \text{ mm}$.

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

Samplingsteoremet krever $f_s > 2f_{\max}$, altså må $f_s > 2 \cdot 1/3 \text{ mm}^{-1} = 2/3 \text{ mm}^{-1}$.

- c) Anta at den geometriske transformen er gitt ved:

$$x' = 0.3x + 30$$

$$y' = 0.3y + 300$$

der x og y er koordinatene i "innbildet", x' og y' er de transformerte koordinatene. La oss videre anta at det benyttes "nærmeste nabo" interpolasjon ved baklengstransformasjon.

Hva vil den effektive samlingsraten være etter en slik transform, og hvilke (uønskede) effekter vil denne geometriske transformen helt sikkert gi opphav til i dette bildet? Begrunn svaret.

Den geometriske transformen vil nedskalere bildet slik at hver retning blir redusert til 30 % av sin opprinnelige størrelse. En slik nedskalering vil påvirke samlingsraten tilsvarende, så den nye, effektive samlingsraten blir redusert til $f_s' = 0.3f_s = 0.2 \text{ mm}^{-1}$. Det transformerte bildet vil derfor ikke kunne representere de høyeste frekvensene som kan være til stede i det originale bildet, og siden punktkildene så vidt kunne skilles i det originale bildet vet vi at det finnes frekvenser i det originale bildet helt opp til Nyquist-raten, noe som gjør at vi helt sikkert vil ende opp med aliasingproblemer i det transformerte bildet.

3. Histogrammer

La oss anta at vi har følgende 5x5 gråtonebilde:

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

- a) Skissér gråtonehistogrammet, h , det normaliserte histogrammet, p , samt det kumulative normaliserte histogrammet, c , til bildet.

[5, 2, 2, 4, 4, 4, 2, 2]

[5/25, 2/25, 2/25, 4/25, 4/25, 4/25, 2/25, 2/25]

[5/25, 7/25, 9/25, 13/25, 17/25, 21/25, 23/25, 1]

- b) Beskriv hvordan du vil utføre histogramutjevning på det oppgitte gråtonebildet.

Finn det normaliserte histogrammet, og fra dette det normaliserte kumulative histogrammet $c[i]$ (se ovenfor). Sett inn verdier i transform-arrayet $T[i] = \text{Round}((G-1)*c[i])$ for $i = 0, 1, \dots, G-1$. Gå deretter gjennom bildet piksel for piksel, og sett $g(x,y) = T[i(x,y)]$.

4. Gråtonetransform

Anta gråtonetransformen $T_1[i] = ai + b$, der a og b er konstanter.

- a) Hvilken effekt har parameteren a på kontrasten og lysheten til bildet for forskjellige verdier av a ?

$a > 1$ øker kontrasten og lysheten, $0 < a < 1$ minsker kontrasten og lysheten.

$a < 0$ vil speilvende histogrammet og gi oss et "negativt" bilde med større eller mindre kontrast, avhengig av tallverdien til a .

- b) Gitt gråtonetransformen $T_2[i] = c \log(i+1)$ for $0 \leq i \leq G-1$. Anta at $0 \leq T_2[i] \leq G-1$.

- Tegn en skisse og forklar hva transformen vil gjøre med kontrasten i henholdsvis de mørke og lyse intensitetsintervallene.
- Hva bør konstanten c være?

Stigningstallet til transformen T_2 er høyt ved lave intensitetsverdier, og lavt ved høye verdier; følgelig vil transformen øke kontrasten i de mørke områdene og minke kontrasten i de lyse områdene.

c er en skaleringsfaktor som skal sikre at vi får ut-verdier $0 \leq T_2[i] \leq G-1$.

Det gjør vi hvis $c = (G-1)/\log(G)$.

5. Lavpass- og medianfiltrering

Bruk 5x5 gråtonebildet fra oppgave 3.

Hvilke pikselverdier får vi i det midtre 3x3 utbildet hvis vi bruker:

a) Et 3x3 kvadratisk middelverdifilter uten skalering?

32	35	36
21	24	35
12	21	32

b) Et 3x3 pluss-formet medianfilter?

Sorterer de 5 pikselverdiene for hver plassering av filtret og får:

[33455], [00456], [05577]

[00234], [03355], [00456]

[01133], [00234], [33455]

I midten av hver sortert tallrekke finner vi medianverdiene:

4	4	5
2	3	4
1	2	4

c) Et 3x3 X-formet medianfilter?

Sorterer de 5 pikselverdiene for hver plassering av filtret og får:

[02446], [35557], [00066]

[13335], [00044], [35557]

[00022], [13335], [02446]

I midten av hver sortert tallrekke finner vi medianverdiene:

4	5	0
3	0	5
0	3	4

NB! Her er det ikke nok å bare gi et svar i form av pikselverdiene i utbildet.

Du må også vise hvordan du finner svarene.

6. En konvolusjonsoperator

En operator T_3 er definert ved

$$T_3 = - \left\{ \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix} * \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix} + \begin{bmatrix} -2 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix} * \begin{bmatrix} -2 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix} \right\}$$

- a) Forklar hva slags operator dette er, og hvordan den skiller seg fra en kjent operator du har sett før.

Dette er en sum av to konvolusjoner.

Den første er en Sobel-operator rotert 45 grader, konvolvert med seg selv, altså den andre deriverte i bi-diagonal retning (oppover mot høyre). Den andre er tilsvarende i diagonal retning. Summen må da være en diskret tilnærming til en rotert Laplace-operator i et 45 grader rotert koordinatsystem, med negativt fortegn:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

(Ligningen er en del av løsningsforslaget)

- b) Beregn resultatet av T_3 . Er denne operatoren separabel? Begrunn svaret.

$$T_3 = - \left\{ \begin{bmatrix} 0 & 0 & 1 & 4 & 4 \\ 0 & -2 & -4 & 2 & 4 \\ 1 & -4 & -12 & -4 & 1 \\ 4 & 2 & -4 & -2 & 0 \\ 4 & 4 & 1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 4 & 4 & 1 & 0 & 0 \\ 4 & 2 & -4 & -2 & 0 \\ 1 & -4 & -12 & -4 & 1 \\ 0 & -2 & -4 & 2 & 4 \\ 0 & 0 & 1 & 4 & 4 \end{bmatrix} \right\} = - \begin{bmatrix} 4 & 4 & 2 & 4 & 4 \\ 4 & 0 & -8 & 0 & 4 \\ 2 & -8 & -24 & -8 & 2 \\ 4 & 0 & -8 & 0 & 4 \\ 4 & 4 & 2 & 4 & 4 \end{bmatrix}$$

(Ligningen er en del av løsningsforslaget)

Denne operatoren er ikke separabel fordi den ikke kan uttrykkes som en konvolusjon av to 1D-filtre. (I en ordinær Laplace-operator er hver av delene i summen separable (dette er vist i foil 32, forelesning 7, 2012), men her er ikke dette tilfelle, siden Sobel-operatoren er rotert.)

- c) Hva er et LoG-filter, og hva brukes det til, og hvor i resultatbildene finner vi det vi søker?

Laplacian of Gaussian (LoG) er en konvolusjon av Laplace-operatoren (se oppgaven ovenfor) med en 2D Gauss-funksjon $G(\sigma)$. (LoG anvendt på et bilde $f(x,y)$ gir samme resultat som om vi hadde lavpassfiltrert $f(x,y)$ med $G(\sigma)$ og så konvolvert resultatet med Laplace-operatoren.) Filtret brukes til kantdeteksjon i bilder, og siden Laplaceoperatoren beregner den andre-deriverte, er det nullgjennomgangene i resultatbildet som gir oss posisjonen til kantene. (Fint om man nevner at LoG er en mer støyrobust estimering enn Laplace-operatoren.)

7. Todimensjonal diskret Fourier-transform (2D DFT)

- a) Gi en kort beskrivelse av hva 2D DFT er. Er representasjonen unik for hvert bilde?

Tiltenkt beskrivelse: 2D DFT er et (ortogonalt) representasjonsskifte (eventuelt basisskifte) (med hensyn på en basis som er) basert på sinuser og cosinuser med forskjellige frekvenser.

Alternativ, fullgod beskrivelse: 2D DFT er et representasjonsskifte.

Representasjonen er en mengde komplekse tall som kan (eventuelt normalt) representeres ved fasene og amplitudene. Amplitudene beskriver hvilke frekvenser bildet inneholder.

Representasjonen er unik for hvert bilde (forskjellige bilder vil alltid ha forskjellig 2D DFT).

- b) Beregn $F(2,1)$ der F er 2D DFT-en av følgende 4×4 -utsnitt av gråtonebildet fra oppgave 3:

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

Du kan få bruk for følgende matriser:

1	0	-1	0
-1	0	1	0
1	0	-1	0
-1	0	1	0

Cosinus-bildet av størrelse 4×4 og frekvens $(2,1)$

0	-1	0	1
0	1	0	-1
0	-1	0	1
0	1	0	-1

Sinus-bildet av størrelse 4×4 og frekvens $(2,1)$

(alle frekvenskomponentene i denne deloppgaven og videre er nullindeksert)

Summen av punktproduktet av 4×4 -bildet og cosinus-bildet er 4. Summen av punktproduktet av 4×4 -bildet og sinus-bildet er 4. $F(2,1) = 4 + 4i$

- c) Finn koeffisienten til 2D DFT-en av 4×4 -utsnittet over (deloppgave 7b) for en annen frekvens, altså en frekvens (u,v) der $0 \leq u,v \leq 3$ og $(u,v) \neq (2,1)$. Hint: Bruk egenskap(er) ved 2D DFT.

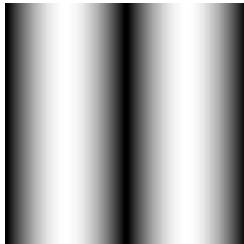
Alternativ 1: $F(0,0) = \langle \text{sum av alle gråtoneintensitetene i bildet} \rangle = 52$

Alternativ 2: 2D DFT er konjugert symmetrisk, så $F(-2,-1) = 4 - 4i$, og også periodisk med periode M og N henholdsvis vertikalt og horisontalt, så $4 - 4i = F(-2,-1) = F(-2+4,-1+4) = F(2,3)$. Dermed er $F(2,3) = 4 - 4i$

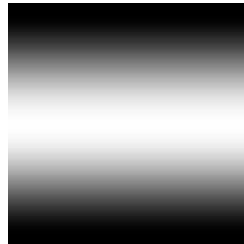
Alternativ 3: Finne 4×4 cosinus- og sinus-bildet for en annen frekvens og beregne 2D DFT-koeffisienten til den frekvensen på samme måte som i deloppgave 7b. Forslag: Cosinus-bildet av størrelse 4×4 og frekvens $(1,2)$ er den transponerte av cosinus-bildet i deloppgave 7b. Tilsvarende er sinus-bildet av størrelse 4×4 og frekvens $(1,2)$ den transponerte av sinus-bildet i deloppgave 7b. Ved å summere hvert av punktproduktene finner vi at $F(1,2) = 4 - 4i$.

8. Filtrering i frekvensdomenet

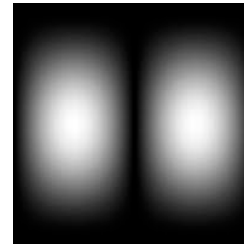
- a) Bildene under viser Fourier-spektrene til tre (nullutvidede) konvolusjonsfiltre, der det siste (nullutvidede) konvolusjonsfilteret er sirkelkonvolusjonen av de to andre. For fremvisning er hvert Fourier-spekter forskjøvet slik at DC, (0,0)-frekvensen, ligger midt i bildet. I hvert bilde er 0 visualisert som svart og maksimalverdien i Fourier-spekteret visualisert som hvitt.
- Angi og begrunn ut fra hvert Fourier-spekter hva slags type filter det er.
 - Hvordan og hvorfor kan vi beregne Fourier-spekter 3 ut fra de to andre Fourier-spektrene?



Fourier-spekter 1



Fourier-spekter 2



Fourier-spekter 3

Fourier-spekter 1 representerer et horisontalt høypassfilter/kantdeteksjonsfilter. Merk at lave horisontale frekvenser dempes/fjernes (spesielt er DC 0), mens høye horisontale frekvenser (stort sett) bevares. (Konvolusjonsfilteret før nullutvidelse var $[1 \ 0 \ -1]$.)

Fourier-spekter 2 representerer et (gaussisk) lavpassfilter. Merk at lave vertikale frekvenser bevares (spesielt er DC lik maksimalverdien i Fourier-spekteret), mens høye vertikale frekvenser dempes/fjernes. (Konvolusjonsfilteret før nullutvidelse var $[1 \ 2 \ 1]^T$.)

Fourier-spekter 3 representerer et støyrobust horisontalt høypassfilter/kantdeteksjonsfilter. Merk at frekvensene som (stort sett) bevares er både høye horisontalt og lave vertikalt, mens alle andre frekvenser dempes/fjernes, mer presist dempes/fjernes alle frekvensene som enten er lave horisontalt (spesielt er DC 0) eller høye vertikalt. (Konvolusjonsfilteret før nullutvidelse var det horisontale Sobel-filteret.)

Ved konvolusjonsteoremet vet vi at sirkelkonvolusjonen av de to første konvolusjonsfiltrene (i billedomenet) har en 2D DFT som er lik det punktvis produktet av konvolusjonsfiltrenes 2D DFT-er. Siden et Fourier-spekter inneholder magnitudene til de komplekse tallene i en 2D DFT og magnituden til produktet av to komplekse tall er produktet av magnitudene til hver av faktorene, så er Fourier-spekteret til konvolusjonsresultatet lik punktproduktet av konvolusjonsfiltrenes Fourier-spektre.

- b) Når vi skal utføre en filtrering i frekvensdomenet, støter vi noen ganger på et problem som kalles wraparound-feil. Hva er dette problemet, når forekommer det og hvordan kan vi forhindre det?

Wraparound-feil er en feil som forekommer når vi filtrerer i frekvensdomenet, men ikke ønsker å behandle randproblemet ved sirkulær indeksering. Det vil forårsake at filtreringsresultatet i piksler nær en bildekant vil være påvirket av intensitetsverdiene til piksler på motsatt bildekant, uten at dette er ønskelig. (At denne feilen forekommer når vi filtrerer i frekvensdomenet er en konsekvens av at vi implisitt må anta at bildet som skal filtreres er periodisk ettersom vi benytter dets Fourier-representasjon.) Problemet vil alltid forekomme når vi filtrerer i frekvensdomenet og sirkulær indeksering ikke er en ønskelig måte å behandle randproblemet på, og er mest synlig dersom avviket som gjøres ved å bruke sirkulær indeksering (fremfor den ønskelige behandlingen av randproblemet) er stort. Vi kan forhindre problemet ved å utvide bildet på ønskelig måte før 2D DFT-en beregnes og brukes til filtrering.

TAKK FOR OPPMERKSOMHETEN!