

UNIVERSITETET I OSLO

Det matematisk-naturvitenskapelige fakultet

Eksamen i: IN3370 — Digital bildebehandling og -analyse

Eksamensdag: Fredag 6. desember 2024

Tid for eksamen: 15.00 – 19.00

Oppgavesettet er på 8 sider.

Vedlegg: Ingen

Tillatte hjelpemidler: Ingen

Kontroller at oppgavesettet er komplett før
du begynner å besvare spørsmålene.

**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.*

**Les gjennom hele oppgavesettet før du begynner å løse
oppgavene.**

*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.*

Merk: Alle svar må begrunnes!

*Gjør rede for bruken av eventuelle teoremer, prinsipper eller forutsetninger
slik at en tredjeperson kan følge dine resonnementer.*

Oppgave 1 Avbildning og sampling

a

Forklar hva som menes med en punktspredningsfunksjon, og hvordan bredden på denne er knyttet opp til et vanlig mål på romlig oppløsning.

Løsningsforslag: En punktspredningsfunksjon beskriver hvordan ett enkelt "punkt" (strukturer med nærmest ingen utstrekning) fremstår etter avbildning, altså i bildeplanet før sampling. Bred slik funksjon tilsvarer et mer utsmørt bilde. Et vanlig mål på romlig oppløsning er hvor nærme to "punkter" kan være i objektplanet (alternativt vinkel mellom dem ved varierende avstand) og hvor man fortsatt kan skille dem fra hverandre i bildet. Jo bredere punktspredningsfunksjon, jo lengre må disse punktene stå fra hverandre for ikke å smøres fullstendig inn i hverandre.

b

Vi flytter to punktlignende strukturer i objektplanet nærmere hverandre helt til de så vidt kan adskilles i avbildningsplanet. Avstanden mellom dem i avbildningsplanet er da $2/3$ mm. Bestem den nødvendige tettheten av samplingspunkter for å digitalisere bilder fra dette avbildningssystemet.

Løsningsforslag: Den minste avstanden mellom repeterende strukturer vil da være $D = \frac{2}{3}$ mm, som tilsvarer (nogenlunde) den minste perioden i bildet. La T_s være samplingsperioden ($1/\text{samlingsfrekvensen}$). Samplingsteoremet krever at $T_s < \frac{1}{2}D = \frac{1}{3}$ mm.

c

Anta at vi sampler nært opptil Nyquist-raten for vårt avbildningssystem. Anta så at vi oppsamler bildet med en faktor 5 gjennom en geometrisk skaleringsoperasjon, hvor baklengsmapping brukes med nesten perfekt interpolasjon. Forklar hvilken effekt denne oppskaleringen har på den høyeste mulige romlige frekvensen ($f_{\max}$) i det digitale bildet. **Løsningsforslag:** Kun pikselmengden i bildet er økt, den høyeste romlige frekvensen $f_{\max}$ endres ikke. Det er ikke lagt til mer informasjon i bildet, kun foretatt interpolasjon.

d

La oss nå anta at det blir gjort en diskret Fourier-transform av de oppskalerte bildene fra c). Beskriv og/eller skisser hva du vet om hvordan frekvensspekteret til bildene vil se ut. **Løsningsforslag:** Etter oppskaleringen vil frekvensspekteret reflektere den opprinnelige informasjonen i bildet, uten å introdusere nye detaljer. Dette betyr at det meste av energien i spekteret fortsatt vil være konsentrert i den sentrale (lavfrekvente) delen. Den høyeste romlige frekvensen ($f_{\max}$) vil nå være plassert kun $1/5$ fra senteret langs hver akse. (Trolig lettere ved bruk av figur.) Magnitudene vil være høyere i det oppskalerte bildet, men dette er ikke det sentrale vi skal frem til.

(Fortsettes på side 3.)

e

Anta nå at vi opererer med et nytt og ukjent avbildningssystem, men hvor vi får tilgang til Fourier-spekteret til dette systemets punktspredningsfunksjon. Forklar hvorfor maksimal frekvens (med magnitude av betydning) på denne vil tilsvare $f_{\max}$ i bildene? **Løsningsforslag:** For et avbildningssystem kan vi modellere bildet som en konvolusjon mellom motivet og systemets punktspredningsfunksjon (PSF). Fourier-transformen av bildet vil dermed være produktet av Fourier-transformen av motivet og PSF. Om denne Fourier-transformen av PSF er (tilnærmet) 0 for frekvenser over $f_{\max}$, vil det så være for bildet også.

Oppgave 2 Gråtonemapping

Vi er gitt følgende lineære (affine) gråtonemapping: $T[i] = ai + b$, der $T[i]$ er hva pikselintensiteten i mappes til, og a og b er koeffisienter.

a

Beskriv effekten a og b har på resultatbildets normaliserte gråtonehistogram. **Løsningsforslag:** $|a| > 1$ vil strekke histogrammet med utgangspunkt i $i = 0$ (og gjøre det lavere, da summen forblir 1), $|a| < 1$ vil motsatt "sammenklemme" histogrammet (og gjøre det høyere), $a < 0$ vil i tillegg speilvende histogrammet. b vil forskyve histogrammet. Eventuelle klipping-effekter kommer i tillegg.

b

Beskriv hvordan vi kan benytte en slik transform til å gi en serie bilder noenlunde den samme lysheten og kontrasten. **Løsningsforslag:** Vi har benyttet gjennomsnittsverdien som mål på lyshet og variansen som mål på kontrast. Et ønsket gjennomsnitt (μ_t) og en ønsket varians (σ_t^2) kan bestemmes. Deretter kan man for hvert bilde beregne nåværende gjennomsnitt (μ) og varians (σ^2), og ut ifra det finne koeffisientene a og b i gråtonemappingen, som så igjen anvendes på bildet. Mer presist settes $a = \sigma_t/\sigma$, $b = \mu_t - a\mu$.

c

Beskriv eller skisser en $T[i]$ (ikke nødvendigvis affin) som kan benyttes til å fremheve mørke strukturer i et gråtonebilde på bekostning av de lyse. Input-bildene har 8 bit, men vi ønsker å kunne lagre bildet i etterkant med 10 bit. **Løsningsforslag:** Noe med en form à la logaritmisk skalering. Må dog skaleres slik at $T[\max \text{ input}] < 2^{10}$, altså $T[255] \approx 2^{10} - 1$.

Oppgave 3 Geometriske transformer

Vi har jobbet med affine, geometriske transformasjoner som kan uttrykkes slik at et koordinat (x, y) blir transformert til (x', y') ved:

$$x' = a_1 \cdot x + a_2 \cdot y + a_0$$

(Fortsettes på side 4.)

$$y' = b_1 \cdot x + b_2 \cdot y + b_0$$

Koeffisientene $a_0, a_1, a_2, b_0, b_1, b_2$ bestemmer altså transformasjonen.

a

Hvilke vanlige bildeoperasjoner kan man oppnå ved å benytte slike transformer? **Løsningsforslag:** Translasjon, rotasjon, skalering (inkludert refleksjon), skjæring, og alle kombinasjoner av disse.

b

Vis hvordan en hel sekvens med slike affine transformer kan uttrykkes som én enkelt affin transform. **Løsningsforslag:** Hver av disse transformene kan ses på som en matrisemultiplikasjon med en 1-utvidet posisjonsvektor. Matrisemultiplikasjon er assosiativ, og følgelig kan man slå sammen alle disse matrisene før den siste matrise-posisjonsvektormultiplikasjonen. (Vi har i kurset jobbet nesten eksklusivt med en slik matrisevariant.)

c

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

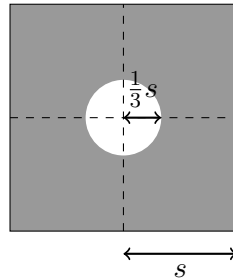
Ved (praktisk) bruk av geometriske transformer ønsker man som oftes igjen å ha sampler/pikslar på et rektangulært grid. En vanlig måte her er å benytte en baklengsmapping. Forklar hva som menes med baklengsmapping, samt redegjør for fordelene med denne kontra resampling etter en forlengsmapping. **Løsningsforslag:** Ved en baklengsmapping traverseres ut-pikselposisjonene, og hver av disse kjøres igjennom en invers variant av transformen for å finne dets koordinater i inn-bildet. Å finne de nærmeste (romlige) samplene for interpolasjon er her trivielt (samplene er allerede romlig sortert (i matrisen) og heltallsindeksert).

Oppgave 4 Fourier-transform

a

Vi har designet følgende filter i frekvensdomenet:

(Fortsettes på side 5.)



I figuren er nullfrekvensen midt i figuren, og grått er 0 og hvitt 1. Hvordan vil du beskrive filteret? (Lavpass, høypass, båndpass, notch-filter, ...). Hvilke uønskede effekter kan fremkomme i bildet etter filtrering? **Løsningsforslag:** Lavpassfilter som reduserer f_{max} til $1/3$ av opprinnelig. De skarpe overgangene i dette naive/ideelle filteret kan gi opphav til ringing-effekter.

b

Vi sier ofte at man må huske på å lage symmetriske (rundt nullfrekvensen) filtre i frekvensdomenet (som ble gjort for filteret i oppgave a)). Hva skjer i praksis om vi ikke gjør det? **Løsningsforslag:** Enhver reel filterkjerne vil ha (konjugert) symmetri i frekvensdomenet. Filtring med ikke-symmetriske filtre i frekvensdomenet vil tilsvare en konvolusjon med komplekse verdier, og altså gi oss komplekse pikselverdier.

c

Er det noen konvolusjonsfiltre vi kan lage i (det diskrete) frekvensdomenet som ikke kan lages i billedomenet, eller omvendt? Begrunn. **Løsningsforslag:** Nei. Konvolusjonsteoremet sier at en (for oss sirkel-)konvolusjon i billedomenet tilsvarer multiplikasjon i frekvensdomenet, og omvendt. Det er ingen føringer/begrensninger.

d

I hvilke tilfeller kan det være en fordel å utføre en DFT av både bildet og filterkjernen for så å gjøre filtreringen der før resultatbildet igjen blir transformert til billedomenet med en invers DFT? **Løsningsforslag:** I den obligatoriske oppgaven demonstrerte de at store filtre (store filterkjerne) kan implementeres raskere som beskrevet i oppgaven.

Oppgave 5 Filtring i billedomenet

Gitt følgende utsnitt av et 3-bits gråtonebilde med pikselverdier. I svarene må du vise litt utregning eller begrunnelse, ikke bare gi svaret.

1	3	7
2	5	7
1	7	3

(Fortsettes på side 6.)

a

Hvilken verdi vil senterpikselet her få hvis vi bruker:

- Et 3×3 uniformt lavpassfilter?
- Et 3×3 kvadratisk medianfilter?
- Et 3×3 plussformet medianfilter?

Løsningsforslag:

- $(1+3+7+2+5+7+1+7+3)/9 = 36/9=4$.
- Sorterer: 1,1,2,3,3,5,7,7. I midten finner vi medianen, $M = 3$.
- Sorterer og finner 2,3,5,7,7. I midten finner vi medianen $M = 5$.

Oppgave 6 Segmentering ved terskling

Anta at du har et digitalt gråtonebilde av størrelse 8×8 med $G = 8$ gråtoner. Om bildet vet vi følgende:

- Det normaliserte *kumulative* histogrammet til bakgrunns pikslene stiger lineært fra $1/8$ ved $g = 0$ til 1 ved $g = 7$.
- Det normaliserte *kumulative* histogrammet til forgrunns pikslene stiger lineært fra 0 ved $g = 1$ til 1 ved $g = 3$.
- Forgrunns pikslene har a priori sannsynlighet $1/4$.

a

Beskriv eller tegn en skisse av forgrunns- og bakgrunns histogrammene, og gi en kort forklaring på hvordan du kommer fram til dem.

Løsningsforslag: Siden det kumulative histogrammet til forgrunns pikslene stiger lineært må det være like mange piksler av hver gråtone. Det er 2 gråtoner i forgrunnen ($3-1=2$), og tilsammen $8 \times 8/4 = 16$ forgrunns piksler, altså 8 piksler av hver av de to gråtonene.

Det kumulative histogrammet til bakgrunns pikslene stiger også lineært, så det må være like mange piksler av hver gråtone. Det er 8 gråtoner i bakgrunnen, og tilsammen $8 \times 8 \times \frac{3}{4}$ bakgrunns piksler, altså 6 piksler av hver gråtone.

b

Beskriv eller tegn en skisse av de normaliserte forgrunns- og bakgrunns histogrammene skalert med a priori sannsynlighet, og gi en kort forklaring på hvordan du kommer fram til dem.

Løsningsforslag: Det det etterspørres tilsvarende de to histogrammene fra forrige deloppgave (h_f og h_b) dividert med antall piksler. OK med referanse til "intuisjon" her, men: $h_f = (N * P_f) * p_f$, og $h_b = (N * P_b) * p_b$. Dette gir $P_f * p_f + P_b * p_b = (h_f + h_b)/N$.

(Fortsettes på side 7.)

c

Beskriv eller indiker på skissen fra b) hvor tersklene(e) må settes for å få færrest mulig feilklassifiserte piksler, og gi et uttrykk for hvordan du vil terskle (vær presis med terskelverdi og ulikhetstegn), og angi hvor stor del av for- og bakgrunns pikslene som nå blir feilsegmentert.

Løsningsforslag: Vi ser at det er flere forgrunns piksler enn bakgrunns piksler for $g=2,3$, ellers omvendt. Uttrykket for tersklingen blir da:

$$g(x, y) = \begin{cases} 0, & \text{hvis } f(x, y) \leq T_1, \\ 1, & \text{hvis } T_1 < f(x, y) \leq T_2, \\ 0, & \text{hvis } f(x, y) > T_2. \end{cases}$$

Bakgrunns pikslene fra og med $g=2$ til og med $g=3$ blir feilsegmentert. Ingen forgrunns piksler blir feilsegmentert. (Siden bakgrunns pikslene utgjør $3/4$ av bildet, blir den totale feilsegmenteringen $3/16$).

Oppgave 7 Huffman-koding

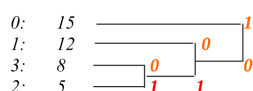
Vi har følgende 2-bits gråtonebilde med 8×5 piksler.

1	1	2	0	3	3	1	0
0	0	2	3	1	0	0	0
1	1	3	0	3	3	1	0
0	1	2	3	1	0	0	3
1	2	2	0	1	0	1	0

a

Vis hvordan du vil gå fram for å Huffman-kode gråtonebildet.

Løsningsforslag: Vi har følgende forekomster, sortert. Slår sammen 2 og 3, så dette med 1, som vist i grafen under:



*Merk: vi ikke har laget frekvenstabell – bare brukt histogrammet.
Huffman-kodene blir da (her er det flere mulige riktige løsninger!):*

0: 1
1: 00
3: 010
2: 011

b

Beregn gjennomsnittlig antall bits pr. piksel etter koding. Du kan se bort fra å måtte lagre selve kodeboken. Vil det være betydelig gevinst å lagre bildet med Huffman-koder?

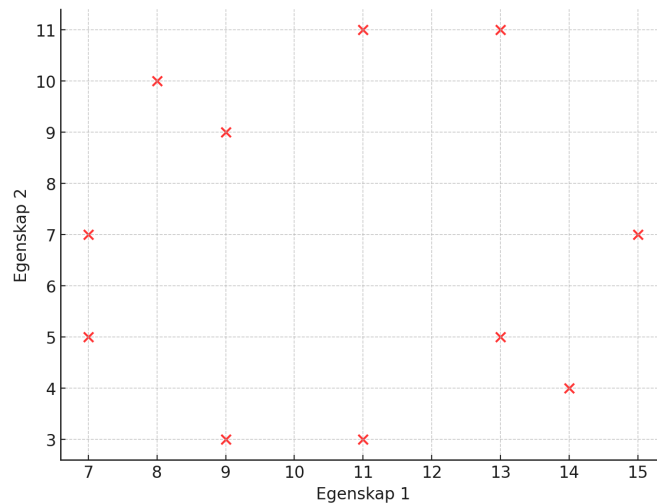
Løsningsforslag: Gjennomsnittlig antall bits pr. piksel etter koding blir: $((8+5)*3 + 12*2 + 15)/40 = (39+24+15)/40 = 78/40 (=1.95)$ bits per piksel. Minimal gevinst, da $1.95 \approx 2$.

(Fortsettes på side 8.)

Oppgave 8 Klyngeanalyse og segmentering

Anta at vi har følgende 2-dimensjonale egenskapsvektorer (i egenskapsrommet, ikke billedomenet):

$(15,7), (11,11), (13,11), (8,10), (9,9), (7,7), (7,5), (13,5), (14,4), (9,3), (11,3)$. Egenskapsvektorene er visualisert i scatterplottet under.



Vi ønsker å utføre en K-means-klustering av disse punktene med $K = 3$. Anta at vi begynner med $(7,5)$, $(9,9)$ og $(11,3)$ som klustersentre.

a

Forklar kort hvordan denne algoritmen fungerer. Hvilke punkter ville blitt satt til de forskjellige klustrene etter første iterasjon? Vil algoritmen konvergere allerede etter første iterasjon i dette tilfellet?

Løsningsforslag: 1) Tildeling: Hvert datapunkt tildeles det klustersenteret det er nærmest, målt ved (typisk) euklidsk avstand. 2) Oppdatering av klustersentre: For hvert kluster beregnes det nye senteret som gjennomsnittet av punktene tildelt clusteret. Iterer til konvergens. Etter første iterasjon, vil klustersentrene bli $(7,6)$, $(10.25, 10.25)$, $(12.4, 4.4)$. Tildelingen av punkt $(9,3)$ vil endres, så algoritmen har ikke konverget.

b

Beskriv kort hvordan SLIC-superpikselalgoritmen utnytter romlig informasjon i segmenteringen.

Løsningsforslag: Egenskapsvektoren utvides med pikslenes koordinater. Og startpunktene for hvert superpiksel plasseres på et regulært grid i bildet.

SLUTT