UNIVERSITETET I OSLO

Det matematisk-naturvitenskapelige fakultet

Eksamen i :	INF2310 — Digital bildebehandling
Eksamensdag :	Fredag 7. juni 2019
Tid for eksamen :	09:00 – 13:00 (4 timer)
Oppgavesettet er på :	**6 sider**
Vedlegg :	**Ingen**
Tillatte hjelpemidler:	**Godkjent kalkulator**

- Det er **6** 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 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, Fourier, sampling og geometrisk transform

Punktspredningsfunksjonen til et avbildningssystem beskriver hvordan ett punkt i objektplanet (i "virkeligheten") vil fremstå etter avbildningen (altså i bildet vårt). Om vi antar linearitet ved avbildningen så kan vi si at det "sanne" objektet med uendelig detaljnivå blir konvolvert med denne punktspredningsfunksjonen for å danne bildet som senere blir digitalisert/samplet.

a) Hva sier konvolusjonsteoremet?

b) Med utgangspunktet i avbildningssystemet beskrevet i introduksjonen til denne oppgaven, hvordan kan vi ved å studere frekvensspekteret til punktspredningsfunksjonen si noe om den romlige oppløsningen til avbildningssystemet vårt?

c) Hvis vi får opplyst at vi er nødt til å sample (plukke pikselpunkter) med en avstand $T_s < 5$ mm for å unngå aliasing, hva sier det om hvor hovedvekten av våre ikke-ubetydelige frekvenskomponenter er å finne i punktspredningsfunksjonens frekvensspekter?

d) Hvis det blir benyttet en samplingsfrekvens som er 3 ganger Nyquistraten (den minste frekvensen vi må sample med for å unngå aliasing), beskriv og/eller skissér hva du vet om hvordan frekvensspekteret til bildene vil se ut. Du kan gjerne hoppe over dette med at spekteret egentlig er repetitivt, da dette ikke er det sentrale vi skal frem til her.

e) Anta at det etter sampling vil bli utført følgende geometriske operasjon:

$$x' = cx,$$
$$y' = cy,$$

der (x',y') er hvor koordinat (x,y) blir forflyttet, og hvor c er en konstant. La oss videre anta at det som neste steg blir utført en "vanlig" baklengs resampling (dog med «perfekt» interpolasjon). For hvilke c vil vi unngå aliasing om det i utgangspunktet hadde blitt benyttet en samplingsfrekvens på 3 ganger Nyquistraten?

2. En konvolusjons-operator

a) Gitt et 8-bits gråtonebilde f(x,y). Hva inneholder ut-bildet G(x,y) når

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

og $h_0, \ldots, h_3$ er konvolusjonsfiltrene

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

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

c) 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).
Hvilket trinn i Canny-algoritmen kan dette «k-bildet» erstatte,
og hvilke matematiske operasjoner unngår vi med dette?

3. Diskret Fourier-transform

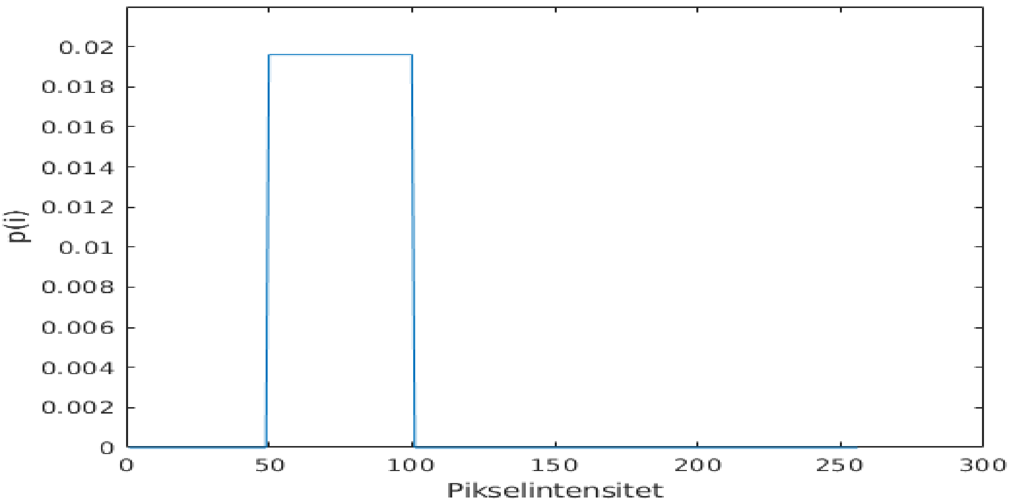
a) Anta at vi har et NxN bilde, f, og at vi har gjort en diskret Fourier transform (i praksis en FFT) av dette bildet og lagret resultatet i NxN-matrisen F, der F(0,0) gir nullfrekvensen («DC komponenten»). Hvis nå alle elementene i F er null bortsett fra F(5,4) og F(N-5, N-4), beskriv hvordan bildet ser ut.

b) Anta samme type matrise F som beskrevet i deloppgave a), men for et generelt bilde f. Gi pseudokode for hvordan man kan hente ut amplituden og fasen til frekvens (u, v), der u er antall repetisjoner vertikalt og v antall repetisjoner horisontalt.

c) Anta så at vi translaterer (forflytter) bildet, f, 10 piksler til høyre og 5 piksler nedover, men på en slik måte at piksler som kommer utenfor bildematrisen kommer inn igjen på motsatt side (det samme som om bildet var repeterende langs begge akser).
Hvordan ville en slik forflytning påvirke amplitudespekteret?

4. Histogramutjevning og histogramtilpasning

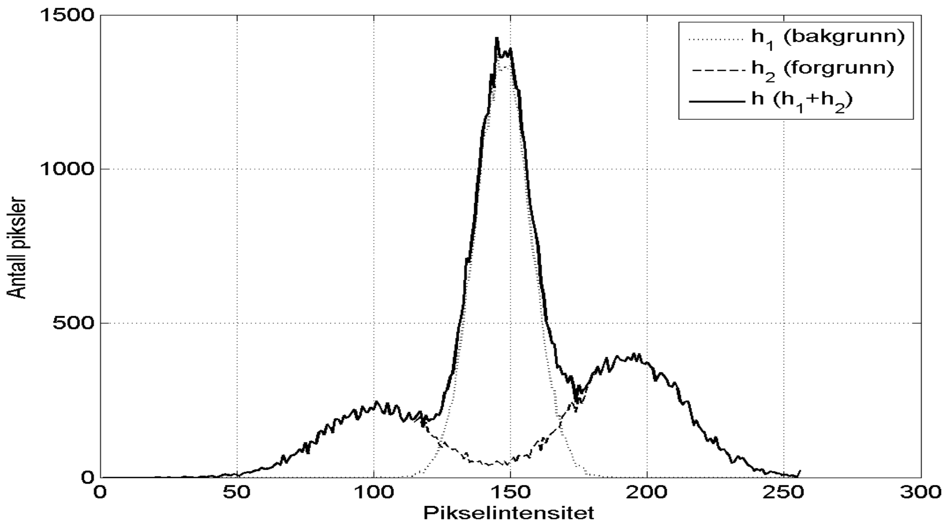
Ved histogramutjevning av gråtonebilder benyttes en global gråtonetransform i et forsøk på å gi et resultatbilde med mest mulig flatt gråtonehistogram.
Slike globale gråtonetransformer er gitt ved en funksjon T[i], for pikselintensiteter i.
Ved histogramtilpasning mener vi her at vi åpner opp for å benytte slike gråtone-transformer i et forsøk på gi resultatbildet et histogram som ligner på et hvilket som helst ønsket resultathistogram.



a) La oss anta at vi har et bilde med et normalisert gråtonehistogram som vist i figuren over. Bildet har heltallige gråtoneverdier fra 0 til 255.
Beskriv og skisser gråtonetransformen, T, du ville benyttet for å utføre en histogramutjevning på dette bildet.

b) Anta at image er en bildematrise med h som gråtonehistogram, og at vi har følgende (selvforklarende) funksjoner tilgjengelig:
T = getHistogramEqTransform(h),
T_{inv} = getInverseTransform(T),
im_{res} = transformImage(image, T).
Skisser (pseudokode) hvordan du kan benytte disse funksjonene til å utføre en histogramtilpasning til et ønsket histogram $h_{ønsket}$.

5. 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 over, hvilke pikselintensitetsintervaller ville du klassifisert til bakgrunn og hvilke intervaller ville du klassifisert til forgrunn? (Som alltid, forklar og begrunn svaret ditt.)

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

c) 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 Ridler og Calvards algoritme (jfr. k-means) for automatisk terskling trolig gi et tilfredstillende resultat på dette bildet?

d) Om det samlede objekt-areal vi vil segmentere ut enten er veldig stort eller veldig lite i forhold til den totale bildeflaten vil vi fort få utfordringer ved automatisk terskling a la Ridler og Calvards metode. Om vi får markert et sett med piksler som befinner seg nær randen av objektene vi er interessert i, hvordan kan vi benytte denne informasjonen til å finne bedre terskelverdier?

e) Ved bl.a. Otsus metode for automatisk terskling kan vi få ut et kvantitativt mål, for eksempel et tall mellom 0 og 1, på hvor godt vår tilpassede modell skiller våre klasser, for eksempel mellom forgrunn og bakgrunn. Beskriv hvorfor og hvordan dette kan komme til nytte ved lokal, adaptiv terskling.

6. Kompresjon og koding

Gitt et 8 x 8 piksels utsnitt av et 3 bits gråtonebilde med pikselverdier:

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

a) Finn bildets gråtonehistogram, normaliserte histogram og kumulative normaliserte histogram.

b) Beskriv trinnene i en Huffman-koding av dette utsnittet, finn kodeboken, og finn det gjennomsnittlige antall bits per piksel etter koding, og kompresjonsraten.

c) Hva er forholdet mellom det gjennomsnittlige antall bits per piksel etter Huffman-koding og entropien til bildeutsnittet, når vi ser bort fra størrelsen på kodeboken? Forklar og begrunn!

d) Forklar hvordan du vil gjøre en histogram-utjevning av det gitte gråtonebildet til et ut-bilde med færre gråtoneverdier, og hvordan dette generelt vil påvirke Huffman-kodingen av bildet.

e) Beskriv stegene i en «lossy» JPEG-kompresjon av bare dette utsnittet.

f) Hvis et større bilde ble delt opp i flere slike 8 x 8 utsnitt, hvilke artefakter ville da kunne oppstå etter kompresjon/dekompresjon, og hvorfor?

Lykke til!