

# UNIVERSITETET I OSLO

## *Det matematisk-naturvitenskapelige fakultet*

Eksamen i:

IN2070 — Digital bildebehandling

Eksamensdag:

Torsdag 2. juni 2022

Tid for eksamen:

15:00 – 19:00

Oppgavesettet er på: **6** sider

Vedlegg:

Ingen

Tillatte hjelpemidler: Ingen

- 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 så fall rede for forutsetningene og antagelsene du har gjort.
- Det er tilsammen 26 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 resonnerer.

# 1 Romlig oppløsning, sampling og geometriske operasjoner

Anta et avbildningssystem som avbilder objekter på lang avstand over på et avbildningsplan. I dette avbildningsplanet vil vi senere utføre sampling av bildet.

Blant det du har mulighet til å avbilde er to punkt-lignende objekter (altså objekter med neglisjerbar utstrekning). Ved nitidig utprøving kommer du frem til at det nærmeste man kan sette disse objektene slik at de blir adskillige i bildet er 0.75 m. Anta videre at man ved geometriske betraktninger kan regne seg frem til at dette tilsvarer 4 mm i avbildningsplanet.

- a) Her vil punktspredningsfunksjonen ha en bredde  $\geq 4$  mm. Forklar hvorfor.

Hadde bredden vært mindre, ville man kunne sette objektene nærmere hverandre uten at de ble smurt sammen i bildet. [Studentene må her vise forståelse for hva en psf er.]

- b) Hva kan du si om høyeste mulige romlige frekvens i bildet ( $f_{\max}$ )?

[I kurset fokuserer vi på frekvenser med bidrag "av betydning", og ser altså bort i fra slikt som dette med uendelig båndbredde grunnet endelig utstrekning på bildet.] Minste periode er 4 mm, altså vil høyeste romlige frekvens være  $1/(4 \text{ mm}) = 0.25 \text{ mm}^{-1} = f_{\max}$ .

- c) Hvor tett må vi sample bildet for å unnga aliasing?

Samplingsteoremet krever  $f_s > 2f_{\max}$ . Altså må minste samplingsperiode være mindre enn  $1/f_s = 1/(2f_{\max}) = 2$  mm. [Viktig at de henviser til samplingsteoremet her.]

- d) Det samples nå dobbelt så tett som svaret på forrige deloppgave. Videre skal følgende geometriske transform utføres:

$$\begin{aligned}x' &= c \cdot x, \\y' &= c \cdot y,\end{aligned}$$

der  $x$  og  $y$  er de originale koordinater,  $x'$ ,  $y'$  de transformerte og  $c$  en konstant. Anta en vanlig baklengs transform som del av den geometriske transform-operasjonen.

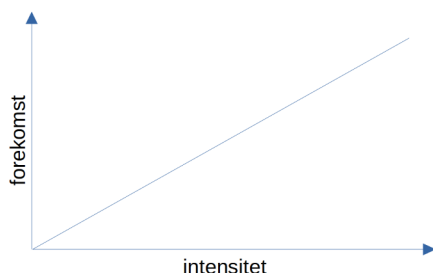
Hvilke krav har vi til  $c$ , og hva risikerer vi om ikke disse kravene oppfylles?

$c > 1/2$ . Om  $c < 1$  vil vi få en nedskalering, og med "vanlig baklengs transform" vil vi få en ny, lavere effektiv samplingsrate:  $c \cdot f_s = c \cdot 4f_{\max}$ . Kravet fra samplingsteoremet er  $c \cdot f_s = c \cdot 4f_{\max} > 2f_{\max}$ . Dette vil ikke tilfredstilles om  $c < 1/2$ , og vi står da i fare for å få aliasing.

- e) La nå  $c=10$ , og la oss anta at det benyttes en nær perfekt interpolasjon under vår geometriske operasjon. Hvilke endringer av bildets høyeste mulige frekvens,  $f_{\max}$ , vil dette gi utslag i?

En oppsampling gir ingen nye ("ekte") detaljer i bildet, og vi vil følgelig ikke endre/øke den romlige oppløsningen.

- f) I samplingsprosessen må også pikslenes intensitet deles opp i intensitetsintervaller, og intensiteter innenfor hver av de lagres med et gitt bit-mønster (tenk heltall). Anta at figuren under viser fordelingen av intensiteter for et gitt bilde før kvantisering.



Anta at du har 4 bit tilgjengelig per piksel. Beskriv og/eller skisser hvor du ville satt slike nevnte intensitetsintervaller.

Antall intervaller tilgjengelig:  $2^4=16$ . Et vanlig ønske er å minimere kvantiseringsfeil. Da ville man laget mindre intervaller ved høyere forekomst; altså, her gradvis smalere intervaller mot høyre i figuren.

## 2 Konvolusjon

- a) I denne deloppgaven skal vi utføre konvolusjon der vi håndterer bilderandproblemet med nullutvidelse. Konvolver bildet  $f$  med filteret  $h$  slik at resultatet blir like stort som bildet  $f$ . Se figur under for hvordan  $f$  og  $h$  ser ut i denne oppgaven. Vi lar filterets origo være i filterets senter, som er uthevet i en tykk kant i figuren under.

| Bildet $f$ |   |   |   |
|------------|---|---|---|
| 1          | 2 | 3 | 1 |
| 0          | 3 | 1 | 2 |
| 2          | 1 | 0 | 1 |
| 1          | 3 | 1 | 3 |

| Filteret $h$ |   |   |
|--------------|---|---|
| 1            | 0 | 0 |
| 0            | 2 | 0 |
| 0            | 0 | 2 |

Resultater etter konvolusjonen er:

5 5 8 2  
 1 8 7 10  
 7 3 9 4  
 2 10 4 6

- b) Gitt bildet som er vist under, hvilket filter du ville ha brukt til å redusere/fjerne så mye av støyen som mulig? Forklar hvordan du kom fram til akkurat det filteret.



Median-filtrering er best egnet til salt-og-pepper-støy

- c) Beskriv hva korrelasjonsoperatoren gjør, og hva en slik operator kan brukes til. Hva er sammenhengen mellom korrelasjon og konvolusjon, slik vi har definert begrepene i dette kurset?

Korrelasjon kan brukes som «template-matching», altså finne mønstre i bildet. Den eneste forskjellen mellom korrelasjon og konvolusjon, er at i korrelasjon så roteres ikke filteret 180 grader før vi anvender det på bildet.

### 3 Morfologi

- a) Vi har fått et bilde  $f$  og et  $3 \times 3$  pluss-formet strukturelement  $S$ , se  $f$  og  $S$  under. Utfør en erosjon på bildet  $f$  med strukturelementet  $S$ . Her er forgrunnen, representert med 1-ere, grå og bakgrunnen, representert med 0-ere, hvit. Bildet er  $13 \times 12$  stort.

Bildet

|   |   |   |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Strukturelementet

|   |   |   |
|---|---|---|
| 0 | 1 | 0 |
| 1 | 1 | 1 |
| 0 | 1 | 0 |

Bruk dette eksempelet til å forklare hvordan erosjon og dilasjon påvirker konvekse og konkave hjørner til forgrunnsobjekter.

Her kan studenten se at etter erosjonen, vil konkave hjørner bli avrundet mens konvekse hjørner blir ikke det.

## Bildet

|   |   |   |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Fra forelesningen har vi sett at erosjon påvirker hjørner på motsatt måte enn det dilasjon gjør, altså når vi vet at erosjon avrunder konvekse hjørner og bevarer konkave hjørner, vil dilasjon bevare konvekse hjørner og avrunde konkave hjørner.

- b) Anta at vi har et strukturelement  $S$  og et bilde  $f$ . Hvordan kan vi regne ut en dilasjon av bildet  $f$  med strukturelementet  $S$  ved å kun bruke erosjon?

Kan bruke at  $f$  erodert med  $S$  er det samme som komplementet av  $f^c$  erodert med  $S$  rotert 180 grader.

- c) Forklar hva vi ønsker å oppnå etter en:

- Åpning
- Lukking
- Hit – or – miss transformasjon

Ved åpning ønsker vi å bl.a. fjerne små strukturer som lenker sammen forgrunnsobjekter, samt fjerne støy.

Ved lukking ønsker vi å bl.a. «tette» igjen hull som ikke nødvendigvis representerer objektet i forgrunnen.

Ved hit-or-miss-transformasjonen ønsker vi å finne spesifikke mønstre slik at vi kan enten fjerne disse eller bruke informasjonen til analyse av bildet.

## 4 Diskret Fourier-transform

- a) Anta at vi har et  $N \times N$  gråtonebilde,  $f$ , og at vi har gjort en diskret Fourier-transform (i praksis en FFT) av dette bildet og lagret resultatet i  $N \times N$ -matrisen  $F$ , der  $F(0,0)$  gir nullfrekvensen. Hvis nå alle elementene i  $F$  er null bortsett fra  $F(4,3)$  og  $F(N-4, N-3)$ , beskriv hvordan bildet ser ut.

Vi har et "bølgebilde" med fire bølgetopper/repetisjoner vertikalt, og 3 horisontalt. Vi vet intet om fasen.

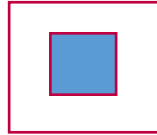
- b) Anta så at vi translaterer (forflytter) bildet,  $f$ , 5 piksler til høyre og 10 piksler nedover, men på en slik måte at piksler som kommer utenfor bildematriksen kommer inn igjen på motsatt side (det samme som om bildet var repeterende langs begge akser).

Hvordan ville en slik forflytning påvirke amplitudespekteret?

Ved shift-teoremet vil kun fasen endres, og amplitudespekteret vil altså forbli det samme. [Studentene trenger ikke å referere til shift-teoremet om argumentasjonen ellers er tilfredstillende.]

- c) Anta nå at vi har fått tildelt et nytt bilde, også dette med størrelse  $N \times N$ . Bildet har blitt samplet med en samplingsfrekvens som er dobbelt av hva samplingsteoremet krever. Lag en skisse av amplitudespekteret hvor frekvenser som det er mulig har ikke-ubetydelig amplitude er markert.

Se figur hvor området med ikke-neglisjerbare frekvenser kan opptre er markert i blått. Antar null-frekvensen i midten av illustrasjonen. Ekstrema har frekvens  $N/2$  (om vi mener antall hele repetisjoner som frekvens), og ekstrema for markert område  $N/4$ . [Egentlig repeteres spekteret i alle retninger.]



- d) I lys av diskrete bilder som vi jobber med; hva sier konvolusjonsteoremet?

Sirkelkonvolusjon i billededomenet tilsvarer multiplikasjon i frekvensdomenet. Og omvendt.

- e) Forklar hvorfor amplitudespekteret til et avbildningssystemets punktspredningsfunksjon kan gi oss informasjon om avbildningssystemets høyeste mulige frekvens,  $f_{\max}$ ?

Om vi antar et lineært avbildningssystem (som vi hele tiden har gjort i kurset), så vil bildet som fremkommer etter avbildningen være "det sanne" (med "uendelig" oppløsning) bildet konvolvert med punktspredningsfunksjonen. I følge konvolusjonsteoremet vil altså frekvensinnholdet til resultatbildet være begrenset av frekvensinnholdet til punktspredningsfunksjonen.

## 5 Segmentering

Anta at vi har et bilde som består av noen objekter av interesse, som fremstår som lysere, på en bakgrunn, som fremstår som mørkere, i bildet. Vi ønsker så å segmentere bildet i forgrunn og bakgrunn kun basert på pikselintensitetene.

Anta også at man har histogrammet for forgrunnen og bakgrunnen hver for seg, henholdsvis  $h_1$  og  $h_2$ . Histogrammet for hele bildet er  $h = h_1 + h_2$ .

- a) Gi en generell regel for om en intensitet  $i$  bør settes til forgrunn eller bakgrunn, basert på de oppgitte histogrammene. Hvilke antagelser legger du til grunn?

Om målet er å minimere antall feilklassifiserte piksler: Velg forgrunn om  $h_1[i] > h_2[i]$ , for intensitet  $i$ . Hipp som happ for  $i$  om  $h_1[i] = h_2[i]$ .

- b) For et gitt, generelt bilde,  $f$ , er fremgangsmåten for å hente ut  $h[i]$  noe vi som har gjennomført dette kurset kjenner veldig godt. Men hvilken tillegginformasjon må vi tross alt ha for å kunne hente ut  $h_1[i]$  og  $h_2[i]$ ?

Vi må ha en oversikt over hvilke piksler som \_faktisk\_ er forgrunn og bakgrunn. Altså, vi må ha fasiten/løsningen på hva vi prøver å løse med vår segmentering for å hente ut  $h_1$  og  $h_2$  i det generelle tilfellet.

- c) Hvilke antagelser om fordelingen til  $h_1$  og  $h_2$  ligger til grunn for Otsus metode (og Ridler og Calvards metode)?

Strengt tatt normalfordelte, med lik varians og med lik a-priori sannsynlighet.

## 6 Koding

a) Gitt følgende melding:

AAABBBCCDE

Lag kodeboken til symbolene i meldingen ved Huffman-koding og vis hvordan du går fram. Regn ut det gjennomsnittlige bitforbruket, samt kompresjonsraten i forhold til å benytte et fast antall bits per symbol. I vårt tilfelle, antar vi at ett symbol opprinnelig bruker 8 bit. Vi ser bort i fra lagring av selve kodeboken.

Vi har følgende hyppighet pr. symbol:

A : 3

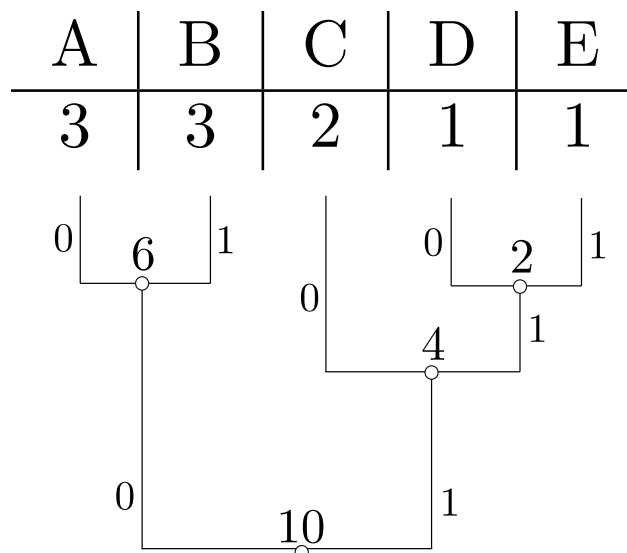
B : 3

C : 2

D : 1

E : 1

Et mulig Huffman-tre, er



Som gir følgende kodebok:

A : 00

B : 01

C : 10

D : 110

E : 111

Fra dette, får vi at gjennomsnittlig bitlengde  $b$  er lik:  $b = 22 / 10$ . Siden det opprinnelige bitforbruket pr symbol er 8 bit, er kompresjonsraten  $8/b = 80 / 22$ .

b) I en setting som beskrevet over, er entropien er en nedre grense for forventet antall bits per symbol. La  $p(x)$  være sannsynligheten for at symbol  $x$  opptrer. Skriv uttrykket for entropi,  $H$ .

$$H = -\sum_x p(x) \log_2 \{p(x)\}$$

- c) Hva vet du om fordelingen,  $p(x)$ , om forventet antall bits per symbol er lik entropien?

Denne grensen nåes kun hvis  $p(x) = 2^{-k}$  for  $k \in \mathbb{N}$  for alle unike  $x$  i meldingen. Forventet antall bits per symbol er  $L = \sum_x p(x)l(x)$ , der  $l(x)$  er kodelengden for  $x$ . For å få  $L=H$ , må  $\log_2\{p(x)\}$  være heltall, og følgelig må  $p(x) = 2^{-k}$  for  $k \in \mathbb{N}$  for alle unike  $x$  i meldingen.

- d) Hvordan kan det ha seg at aritmetisk koding kan gi et gjennomsnittlig bitforbruk som er lavere enn  $H$ ?

Denne algoritmen begrenser seg ikke til å kode symbol for symbol.

## 7 Farger og fargerom

- a) I emnet har vi sett på bl.a. tre fargerom; RGB, CMY og HSI. Beskriv kort hvordan hvert av disse fargerommene representerer farger og i hvilke sammenhenger hver av dem kan være gunstige.

RGB beskriver hvor mye rødt (R), grønt (G) og blått (B) det er i fargen. RGB brukes mest til å genere og fremvise farger digitalt, som f.eks. i skjermer fordi rødt, grønt og blå er primærfarger til lys.

CMY beskriver hvor mye cyan (C), magenta (M) og gul (Y) det er i fargen. CMY brukes ofte til print, da cyan, magenta og gul er primærfarger når vi jobber med pigmenter.

HSI beskriver hvilken dominerende bølgelengde (hue, H), metning (saturation, S) og intensitet (I) fargen har. Ved HSI, kan vi separere kromatisitet fra intensitet og anvende bildebehandlingsoperasjoner på I-komponenten uten å endre på bildets kromatisitet.

- b) Anta vi legger på støy i RGB-kanalene til et fargebilde. Med «støy» så mener vi at vi adderer til små, tilfeldige verdier på hver kanal slik at de opprinnelige verdiene får en *liten* endring. Anta at vi deretter konverterer det støyfylte RGB-bildet til HSI. Forklar hvorfor I-komponenten ikke er like støyfylt som H - og S – komponentene.

I-komponenten vil ikke bli like støyfylt, da den er definert som gjennomsnittet av RGB komponentene og derav vil støyen bli «glattet» ut.