

NTNU
Norges teknisk-naturvitenskapelige
universitet

Fakultet for informasjonsteknologi,
matematikk og elektroteknikk

Institutt for datateknikk
og informasjonsvitenskap



**EKSAMEN I FAG SIF8039
GRAFIKK, BILDEBEHANDLING
OG
MENNESKE-MASKINGRENSESNI TT
ONSDAG 15. MAI 2002**

LØSNINGSFORSLAG

OPPGAVE 1 Metaforer vs. vindussystemer (100 poeng)

- a) Når man skal designe et grafisk brukergrensesnitt står man ofte overfor valget mellom å bruke standardelementer fra et vindussystem eller å utvikle en grensesnittmetafor.

Angi fordeler og ulemper med begge valgene.

Fordelen med en god metafor er at det gjør det lettere for brukeren å forstå logikken til grensesnittet, og ofte gir et mer helhetlig design. Stikkordet er *gjenkjenning* ift. den underliggende problestillingen. Det betyr at det er ønskelig å bruke en metafor, dersom en god en finnes. Fordelen med å bruke standardelementer fra et vindussystem er også knyttet til gjenkjenning, men denne gang ift. hvordan en mer konkret opererer grensesnittet. Strukturelle mønster hører også med, f.eks. liste-detalj-view over data, og gjør det lettere å ta en ny applikasjon i bruk. Bruk av standardelementer gjør det også enklere å veksle mellom forskjellig applikasjon og evt. integrere dem i et felles grensesnitt.

- b) Det skal utvikles et system for å editere innholdet i telefonbaserte automatiske svartjenester. Slike tjenester brukes f.eks. av butikker for at kundene skal kunne få informasjon om varetilbud, åpningstider o.l. Kunden som ringer opp tjenester og gjør valg vha. tall-, *-, og #-tastene på telefonen.

Målgruppen for editeringssytemet er butikkpersonale uten spesialkompetanse i data. Anta at systemet tillater å spille inn lydmeldinger, og å angi hvilke tastetrykk på telefonen som tar innringeren automatisk videre til andre meldinger. Du kan anta at informasjonen kan representeres som en trestruktur, dvs. at svartjenesten tilsvareer en hierarkisk meny.

Det er utviklet to forslag til grafisk grensesnitt:

Forslag I: Representer strukturen vha. hierarkiske lister f.eks. Swings JTree.

Forslag II: Representer datastrukturen som et hulespill (rom-metafor), der hvert rom er en lydmelding, og for hvert tastetrykk på telefonen går det en dør videre til et annet rom.

Evaluer fordeler og ulemper med forslag I og II.

Forslag I.

Fordeler: Gjør bruk av eksisterende grensesnittelementer som brukeren vil gjenkjenne.

Det finnes etablerte måter å legge inn nye elementer, editere og slette.

Ulemper: Hierarkiske lister kan virke kompliserte for ikke-tekniske brukere. Det "tekniske" utseendet kan styre resultatene vekk fra mer kreative løsninger.

Forslag II:

Fordeler: Hvis metaforen er riktig utformet kan den oppfattes som intuitiv og minske opplæring.

Ulemper: Hvis metaforen kun realiseres i form av et "first person" spill, så vil brukeren ikke få noe grafisk kart over hele meldingstreet. Dette kan avhjelpes ved nettopp å generere et slikt kart automatisk som en del av programmet.

Det er ikkegitt utifra HCI teori hvilken av de to løsningene som vil være best. Kun gjennom prototyping og brukbarhetstesting med representative brukere kan dette fastslås.

OPPGAVE 2

MMI – Affordance

(100 poeng)

- a) Forklar begrepet *affordance* slik det brukes i læreboka.

Begrepet "affordance" kommer fra Gibsons "Ecological Psychology" og Don Norman. I Normans tolkning betyr det noe slikt som "det ved et objekts form/utseende som indikerer dets bruk/virkemåte".

- b) Gi eksempler på *kulturelt betingede affordance* og *universell affordance* som gjelder for alle mennesker på tvers av kulturer.

Kulturell affordance er avhengig av kulturell kontekst. Et eksempel er knappe- og bryter utseende som i vår kultur indikerer "noe man trykker på for å få noe til å skje".

Universell affordance er felles for alle kulturer. Cola flaskens form indikerer at den kan holdes for alle mennesker uansett kultur. Det er verdt å merke seg at dette ikke gjelder for alle arter. F.eks. vil ikke Cola flaskens form ha den samme mening for en delfin (som ikke har hender).

OPPGAVE 3**Gjenkjenning vs. fremkalling****(100 poeng)**

- a) Innen litteraturen om menneskelig hukommelse skilles det mellom recall (fremkalling) og recognition (gjenkjenning).

Hva sier den kognitive psykologien om menneskets evne til henholdsvis recall og recognition?

Vi er klart flinkere til og gjenkjenne enn å fremkalle. F.eks. kan man kjenne igjen tegneserier man leste som barn selv om man ikke på forhånd vil være istand til å ramse opp alle tegneserier man har lest. Andre eksempler er gjenkjenning av ansikt vs. evnen til å tegne/beskrive ansikter.

- b) Det skal utvikles et grafisk brukergrensesnitt for et komplekst system med 144 forskjellige funksjoner. Det har kommet opp flere forslag til løsninger.

Forslag I: Kommandospråk som for UNIX med 144 forskjellige kommandoer.

Forslag II: Hierarkisk pull-down meny system med inntil 4 nivåer av menyer.

Forslag III: En 12x12 matrise av grafiske ikoner med tilhørende tekst som alltid er synlig på skjermen.

Evaluer forslag I, II og III utifra innsikten om recall vs. recognition

Forslag I: De 144 kommandoene må huskes (recall), noe som setter store krav til opplæring og evt. tilhørende brukermanualer og hjelpesystemer.

Forslag II: Med hierarkiske pull-down menyer vil alltid øverste nivå kategorier være synlig (recognition). Dette gir en klar brukbarhetsfordel. Brukeren ser derimot ikke alle muligheter samtidig, og metoden vil derfor kun fungere dersom menystrukturen oppleves logisk for brukeren slik at han kan finne fram top-down . Dette setter krav til riktig kategorisering og klassifisering.

Forslag III: En 12x12 matrise vil gjøre alle valg synlige samtidig (recognition). Utifra teorien om recall/recognition er dette optimalt. I praksis er det begrenset hvor store oversikter en bruker er istand til å lete seg fram i effektivt, men det favnes av andre innsikter i kognitiv psykologi.

OPPGAVE 4**Bildebehandling - Bildeforbedring****(150 poeng)**

For oppgave 1a krever vi bare at det demonstreres noenlunde generell forståelse for å gi toppkarakterer, noen detaljer kan gjerne mangle. Verre hvis det som sies er direkte feil.

Fotografier som er skjemmet av ujevn belysning kan modelleres som pixelvise produkter mellom refleksjons- og belysningsbilder, der sistnevnte antas å ha kun svært lavfrekvent informasjon.

- a) Utled et ikke-lineært homomorft filter som kan forbedre slike fotografier.

Kapt. 4.5 (s.191-194)

- b) Gi alle algoritmer nødvendige for å realisere filteret, inklusive alle transformer.

Algoritme:

$\log()$ av hvert pixel, algo. for 2D DFT: fire løkker og akkumulering av produkt mellom bildet og $\cos()$ og $\sin()$,

kompleks multipl. mellom bildet og filteret $D(u,v)$, inv. 2D DFT, og $\exp()$ av hvert pixel.

OPPGAVE 5 Bildebehandling – Bildeanalyse

(150 poeng)

For oppgave 2b og spesielt 2c krever vi bare at det demonstreres noenlunde generell forståelse for å gi toppkarakterer, noen detaljer kan gjerne mangle. Verre hvis det som sies er direkte feil.

- a) Gi et eksempel på et mønstergjenkjenningssproblem innen bildeanalyse.

De fleste burde greie å gi en anvendelse som tar input-bilder, og ender med en klassifisering av bilders objekter.

- b) Beskriv i grove trekk hvordan man kan gå frem for å løse problemet med et feed-forward nevralt nettverk med overvåket læring.

Kapt. 12.2, se Multilayer feed-forward neural networks (s.719-722) og eksempel 12.6 s. 726-729

- c) Utled back-propagation læringregelen for output-laget i et slik nettverk.

Kapt. 12.2, se Training by back-propagation (s.722-725)

OPPGAVE 6

Grafikk - Avbildningstransformasjoner

(150 poeng)

a) Forklar **kort** og **konsist** følgende begreper:

- Parallellprojeksjon

Parallellprojeksjon er projeksjonsmetoder der projeksjonsstrålene er parallelle.

- Perspektivisk projeksjon

Perspektivisk projeksjon er projeksjonsmetoder der projeksjonsstrålene går mot et projeksjonssenter.

- Ortografisk projeksjon

En ortografisk projeksjon er en parallellprojeksjon der projeksjonsstrålene står normalt på projeksjonsplanet

- Aksonometrisk projeksjon

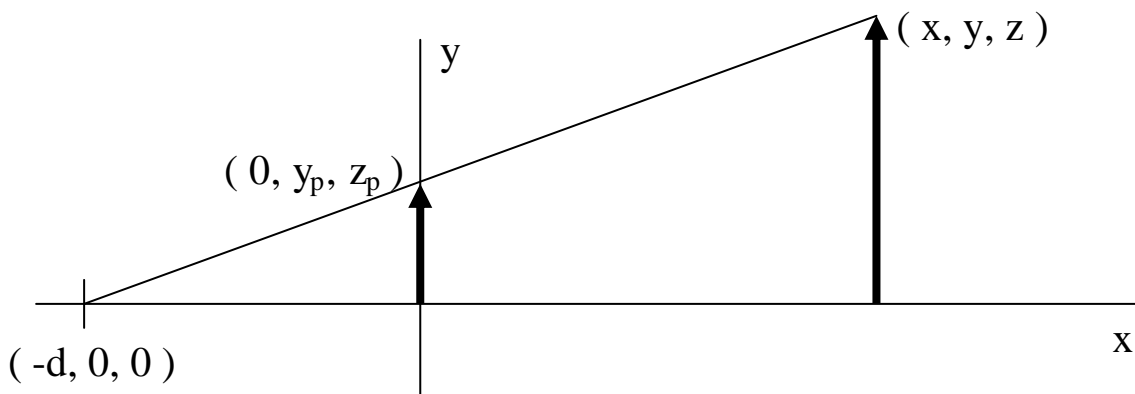
En aksonometrisk projeksjon er en ortografisk projeksjon der projeksjonsplanet ikke er parallelt med noe av hovedkoordinatsystemets koordinatplan (projeksjonsplanet står skjevt i forhold til objektet)

- Isometrisk projeksjon

En isometrisk projeksjon er en aksonometrisk projeksjon der projeksjonsplanet danner samme vinkel med alle koordinataksene (skjærer aksene i samme avstand fra origo)

- Forsvinningspunkt

Forsvinningspunkt opptrer i perspektiviske projeksjoner og er punkter der sett av innbyrdes parallelle linjer i objektet som ikke er parallelle med projeksjonsplanet, skjærer hverandre i projeksjonen. Forsvinningspunkt representerer det som er uendelig langt borte.

b) Utled avbildningsmatrisen for perspektivisk projeksjon når bildet skal være i planet $x = 0$ og projeksjonssenteret skal være punktet $(-d, 0, 0)$ med $d > 0$.

Ved betraktning av likedannede trekanter får vi:

$$\frac{y_p}{d} = \frac{y}{x+d}$$

$$y_p = \frac{y}{\frac{x}{d} + 1}$$

Tilsvarende vil vi få:

$$z_p = \frac{z}{\frac{x}{d} + 1}$$

Med homogene koordinater gir dette:

$$\begin{bmatrix} x' \\ y' \\ z' \\ w \end{bmatrix} = M_{\text{perspektiv}} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ y \\ z \\ \frac{x}{d} + 1 \end{bmatrix} \quad \underline{\underline{M_{\text{perspektiv}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{1}{d} & 0 & 0 & 1 \end{bmatrix}}}$$

Kontroll:

$$x_p = \frac{0}{w} = 0 \quad y_p = \frac{y'}{w} = \frac{y}{\frac{x}{d} + 1} \quad z_p = \frac{z'}{w} = \frac{z}{\frac{x}{d} + 1}$$

Dette viser at matrisen $M_{\text{perspektiv}}$ er den søkte.

- c) Hvordan kan du enkelt komme frem til avbildningsmatrisen for parallellprojeksjon i det samme planet når projeksjonsretningen er langs x-aksen. Skriv opp matrisen.

En kan se parallellprojeksjon som et spesialtilfelle av perspektivprojeksjon der projeksjonssenteret ligger uendelig langt borte. Dette svarer til at vi lar d bli uendelig stor:

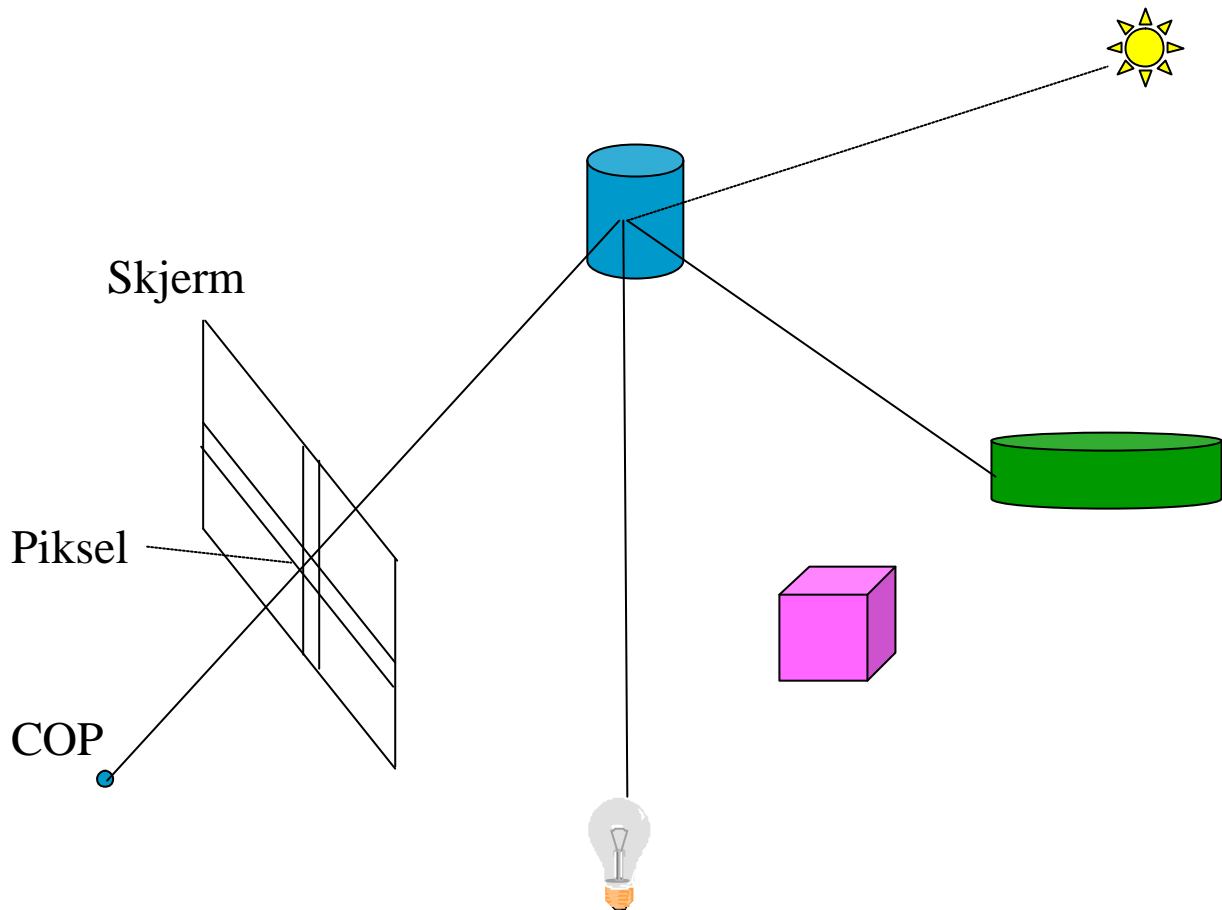
$$M_{\text{parallell}} = \lim_{d \rightarrow \infty} (M_{\text{perspektiv}}) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

OPPGAVE 7

Grafikk – Strålesporingsmodellen

(150 poeng)

- a) Forklar prinsippene for strålesporingsmodellen (ray tracing). Bruk figurer og ikke skriv mer tekst enn høyst nødvendig.



Strålesporingsmodellen sender en stråle fra projeksjonssenteret (øyepunktet) gjennom hvert piksel i avbildningsflaten. I det punktet der strålen treffer ett av scenens objekter, benyttes en lokal belyningsmodell til å beregne fargen. Stråler skytes mot lyskildene for å bestemme hvilke som er synlige og dermed kan bidra til belysningen av punktet. Dersom objektets overflate har evnen til blank (specular) refleksjon, reflekteres en stråle videre inn i scenen. I likhet med en eventuell brutt stråle dersom objektet er gjennomsiktig eller gjennomskinnelig, bidrar denne strålen til fargen i punktet. Bidraget som de reflekterte og brutte strålene gir fra sine treffpunkt, beregnes ved hjelp av samme lokale belyningsmodell. Forskjellige kriterier brukes til å avslutte videre inntrenging av sekundærstråler i scenen.

- b) Navngi og beskriv med matematiske uttrykk en lokal belyningsmodell som kan brukes til beregning av farge i et gitt punkt når strålesporingsmodellen anvendes.

Phongs refleksjonsmodell brukes som lokal belyningsmodell i Whitteds strålesporingsmodell.

$$I_l = \sum_i \left\{ \frac{1}{a + bd_i + cd_i^2} \left[k_{ld} L_{ild} (l_i \cdot n) + k_{ls} L_{ils} (r_i \cdot v)^a \right] + k_{la} L_{ila} \right\} + k_{la} L_{la}$$

L er belysningen i punktet. Det summeres over lyskildene. Indeksene står for:

- l fargekomponent (R, G og B)
- d diffus spredning
- s blank (specular) refleksjon
- a bakgrunnslys (ambient)

k er refleksjonskoeffisienter.

d_i er avstanden til lyskilde i , og det tas hensyn til demping av belysningen med avstanden ved å bruke faktoren:

$$\frac{1}{a + bd_i + cd_i^2}$$

Leddene:

$$k_{ld} L_{ild} (l_i \cdot n)$$

gir den delen refleksjonen som skyldes diffus spredning. l er retningen til lyskilden mens n er flatenormalen i punktet. Leddet viser at belysningen pr. flateenhet avtar med økende innfallsvinkel og at spredningen er uavhengig av retningen til projeksjonssenteret (øyet).

Leddene:

$$k_{ls} L_{ils} (r_i \cdot v)^a$$

gir den blanke (specular) refleksjonen fra punktet. r er retningen til den speilede strålen med v er retningen til projeksjonssenteret (øyet). a er et blankhetsmål som er uendelig stor for ideelle speil og mindre for virkelige flater. Virkningen er en mer eller mindre utflytende speiling av lyskilden.