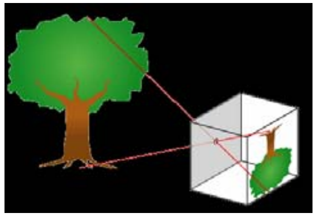
INF 1040
Syn, avbildning og digitale bilder

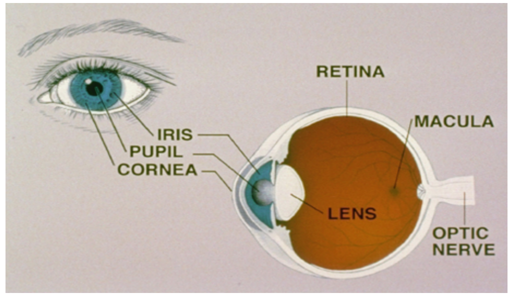
- Temaer i dag :
  1. Synssystemet vårt
  2. Avbildning
  3. Digitalisering av bilder



- Pensumlitteratur: Læreboka, kapittel 12, 13 og 14.

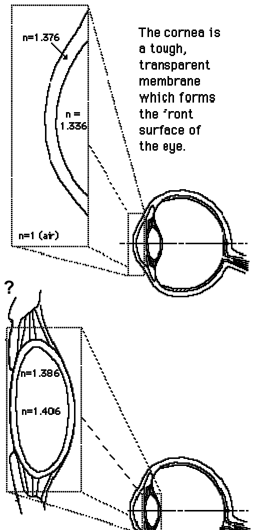
Øyet og synssystemet vårt

- Motivasjon for å kunne noe om dette:
  - Mesteparten av vår sensoriske input kommer via synssansen.
- Fleksibel optikk:
  - Deformerbar linse
- Adaptiv detektor:
  - Variabel oppløsning
  - Logaritmisk respons
  - Pre-prosessering
  - Separate systemer for høylys- og lavlyssyn
- Enorm prosesserings- og lagringskapasitet i hjernen



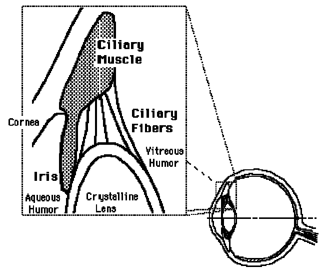
Øyets linsesystem

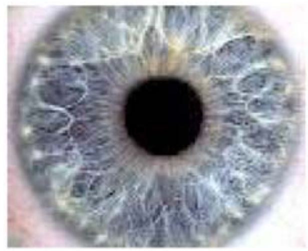
- Øyets linsesystem fokuserer lyset.
- Øyelinsen har en fokallengde, f, som er ca 1.5 cm.
- Linsestyrken angis ofte i "dioptre", d=1/f , der f er gitt i meter.
- Øyelinsen er vanligvis 67 d $(1/1.5 \cdot 10^{-2} \approx 67)$, hvorav hornhinna (*cornea*) står for 45 d.
- Q: Hva betyr det at man korrigerer langsynthet med en +3.0 brille ?
  A: Man bruker en konvergerende linse med f = 1/3 m.
- Q: Hva er fokallengden for en – 4.0 brille?
  A: f = 1/d = -1/4 = - 0.25 m, divergerende linse.
- Øyelinsen er veldig spesiell:
  den kan endre fokallengde.
- Evnen til å skifte fokus raskt (*akkomodasjon*) svekkes med alderen.



Iris og pupillen

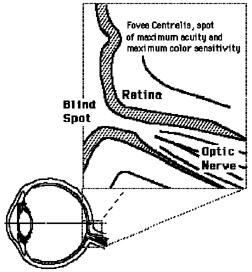
- Iris er den fargede delen av øyet.
- Den fungerer som en blender:
  - Kraftig lys: lukker seg til diameter 2 mm
  - Svakt lys: åpner seg til ca 8 mm
- Digitale bilder av iris kan brukes til verifikasjon ved adgangskontroll
  - mønstret er tilstrekkelig unikt for hver person.
- Pupillen er den svarte åpningen - midt i iris - som slipper lys inn i øyet
  - lyset som går gjennom den blir absorbert i øyet og kommer ikke ut igjen

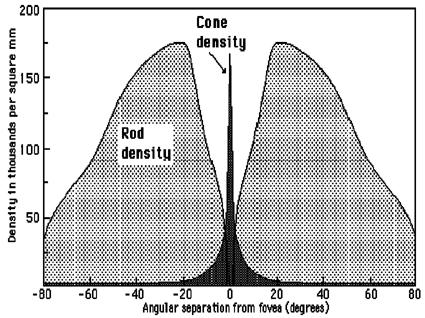




Netthinna (retina)

- Netthinna er det lysfølsomme laget bak i øyet.
- Dekker omtrent 65% av den indre flaten.
- Omtrent 130 millioner detektorer.
- To typer: staver (”rods”) og tapper (”cones”).
- Tappene er konsentrert i *fovea*.
- Stavene er fordelt over resten av netthinna.
- Detektorene vender bort fra lyset!
- Der synsnerven går ut av øyet, har netthinna en blind flekk .





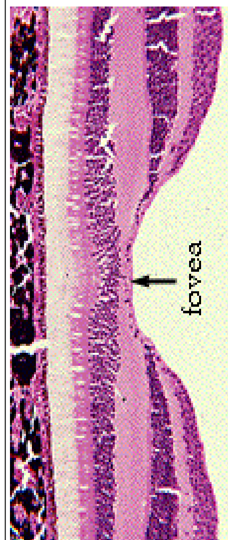
Egenskaper ved synet

- Vi kan se lysintensiteter over ***et intervall på 10 dekader***
  - ”Blendings-intensiteten” er 10^{10} ganger så høy som den svakeste intensitet vi kan oppfatte.
- Vi ser bare et visst antall nivåer samtidig.
  - Den minste gråtone-forskjellen vi kan oppfatte er ca 2%.
  - Vi ser ca 50 forskjellige gråtoner samtidig.
  - Vi kan se mange flere fargernyanser samtidig.
- Når øyet skifter fokus til et annet sted i bildet tilpasser synet seg et annet intensitetsnivå.
  - Vi ser lokale intensitets-forskjeller, både i høylys og lav-lys.

Staver og tapper

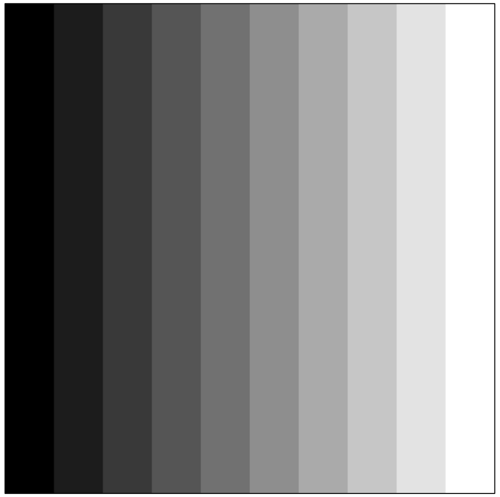
- To klasser av reseptorer:
  - Ca 120 millioner staver (”rods”), spredt over hele netthinna.
    - Flere koblet til hver nerve-ende => lav geometrisk oppløsning
    - Scotopisk (lav-lys) syn: dekker nedre 5 - 6 dekader
    - Gir bare gråtoner (½ time mørke => 10 000 ganger høyere følsomhet)
    - Er ikke følsomme for rødt lys
  - Ca 7 millioner tapper (”cones”), konsentrert i fovea
    - Koblet til hver sin nerve-ende => høy geometrisk oppløsning
    - Fotopisk (høy-lys) syn: dekker øvre 5 - 6 dekader
    - Farge-følsomme: 3 typer (R=700 nm, G=546nm, B=436 nm)

Fovea



- Den gule flekken (*macula*) er ca 3 mm i diameter.
- *Fovea centralis* er ca 0.3 mm i diameter.
  - Overliggende cellelag borte
    - Mer lys til detektorene
  - Bare tapper (høylys, fargesyn)
    - Veldig høy tetthet => høy geometrisk oppløsning.
    - Hver tapp er koblet til en nerve-ende.
- Når vi ser direkte på et objekt, øker oppløsningen, fordi øyet foveerer – flytter bildet til fovea.

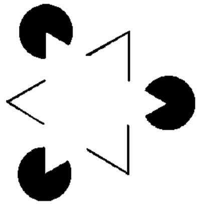
Nevrale prosessorer i netthinna



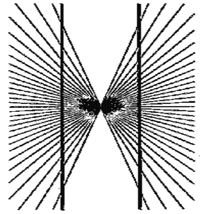
- Netthinna forsterker kanter.
- Stimulering av én del av netthinna undertrykker stimulering av en annen del.
- Dette øker kontrasten ved overgang mellom uniforme regioner i bildet.
- Kalles ”Mach-bånd”

Optiske illusjoner

- Illusoriske konturer

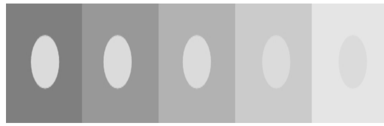


Rette og buete linjer

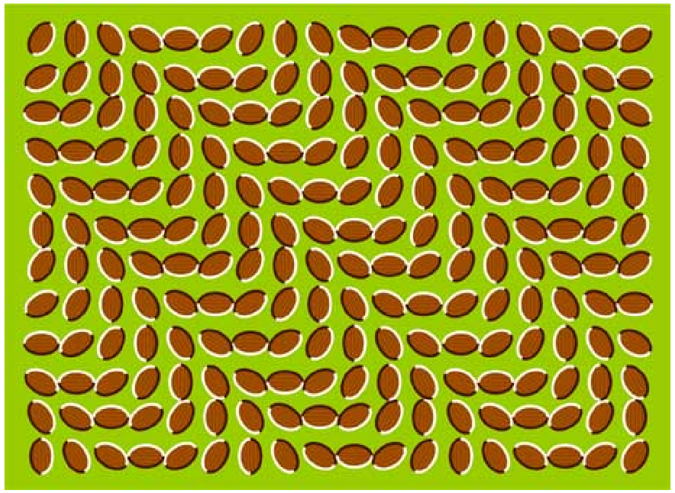


- Multistabile bilder

Simultan kontrast



Optiske illusjoner – ”bevegelse”



Farge-ergonomi

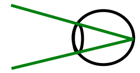
- Øynene stiller krav til arbeidsmiljøet
- Farger gir mer informasjon, hurtig identifikasjon, gir sammenhenger og assosiasjoner, øker motivasjonen.
- Jo flere farger, desto bedre ? Sannheten er som vanlig noe mer nyansert.
- Rødt øker blodtrykk, puls og svette midlertidig. Blått virker motsatt.
- Viktig at ikke overlesset fargebruk forvirrer.
- Få og lett gjenkjennelige farger.
- Samme lysstyrke og fargemetning.
- Stabile farger – minst mulig påvirket av rombelysning.
- Minst mulig farge-stereopsi: Forskjellige farger fokuserer i ulike plan. Resultat: Øye-hjerne arbeider med re-fokusering => hodepine og trøtthet.

"T-Rex and Me"

- I synshjernebarken (visual cortex) sitter flere sett av kant-detektorer, som finner kanter og linjer med forskjellige orienteringer (vinkler) og forskjellige tykkelser.
- Separate sett detektorer for høyre og venstre øye.
- Øyet skanner over objektet, og konsentrerer seg om de interessante, krumme kantene.
- I tillegg har vi flere typer raske øyebevegelser.
- Når øyet beveger seg raskt, er synet "slått av".
- Uten de raske øyebevegelsene, som gir nye reseptorer muligheten til å se bekrefte konturene av objektet, faller synet ut i løpet av sekunder.

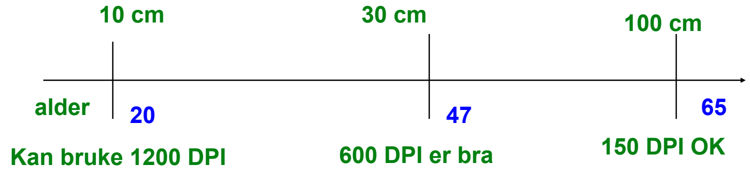
Hvor fine detaljer kan vi se?

- Vi kan se ca. 60 svarte og 60 hvite rader per grad av synsfeltet.
- Et A4-ark holdt ca. 30 cm foran øyet
  - dekker et synsfelt på ca. 50º horisontalt og 40º vertikalt



- 50º horisontalt synsfelt => 50 · 60 = 3 000 vertikale linjer -> 6 000 piksler
- 40º vertikalt synsfelt => 40 · 60 = 2 400 horisontale linjer -> 4 800 piksler
- 6 000 piksler / 11 tommer (A4) → 550 DPI (dots per inch)

Vi kan fokusere skarpt ned til avstander:



Fra bilder til video

- Når vi lukker øynene, tar det litt tid før "etter - bildet" forsvinner, spesielt hvis intensiteten er høy (i deler av) bildet.
- Bildet forsvinner gradvis (eksponensielt) fra retina.
- Det er en kort periode da vi ikke kan ta inn ny informasjon -- selv om øynene er åpne.
- Dette kan vi teste med periodiske lysblink.
- Det viser seg at "critical flicker fusion frequency" øker logaritmisk med økende luminans.
- Den er så lav som 4 s^{-1} for lave lysstyrker, der bare stavene er virksomme.
- Den er ca 60 s^{-1} for kraftig lysstyrke, der bare tappene er virksomme, og konvergerer mot 80 s^{-1}

Luminans-variasjon

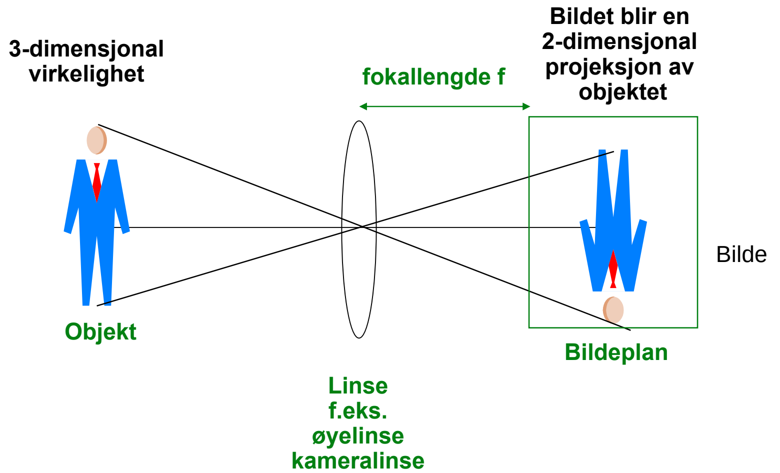
- Når lysblinkene kommer tett nok etter hverandre i tid, vil de oppfattes som et jevnt lys, uten "flicker".
- Den effektive oppfattede luminans vil være gjennomsnittet over belysnings-sykelen, som jo egentlig består av en lys og en mørk del.
- Dette betyr at vi kan variere den effektive luminans ved å holde den fysiske intensiteten til lyskilden konstant, og variere lengden av den mørke delen mellom lysblinkene.

Avbildning

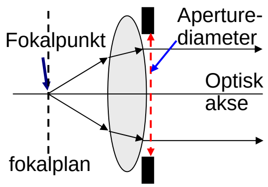
- Avbildning er en prosess som produserer et bilde av en del av våre omgivelser.
  - Visualisering er ikke avbildning.
  - Et bilde, men ikke avbildning:

- Vi konsentrerer oss om digitale bilder.
- Motivasjon for å kunne noe om dette:
  - *Hvor små må detektorene i bildeplanet være for å få med seg alle detaljene ?*
    - Da må vi finne hvor store de minste synlige detaljene i bildet er, når vi kjenner linsens diameter og lysets bølgelengde.
  - *Hvor mange detektorer trenger vi for å dekke hele bildet ?*
    - Da må vi finne ut hvor stort bildet av objektet blir, gitt at vi kjenner fokallengden til linsen, avstanden til objektet og objektets størrelse.

Kamera og optikk



"The lensmakers equation"

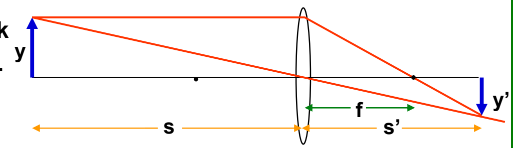


- En *konveks* linse vil samle lysstråler som er parallelle med *optisk akse* i *fokalpunktet*.
- Avstanden fra fokalpunktet til midten av linsen er *fokallengden*.
- Bildet av et fjernt objekt dannes i *fokalplanet.*
- Linsen kan blendes ned til en *aperture-diameter ≤ fysisk diameter*.
- For de interesserte: Fokallengden avhenger av krumningsradiene:
  - "Lensmakers equation" gir linsestyrken i *dioptre*, invers fokallengde:
  - n = brytningsindeksen for glasset i linsen,
  - R_1 og R_2 er krumningsradiene til linseflatene på forsiden og baksiden av linsen.

$$\boxed{d = \frac{1}{f} = (n-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)}$$

  - En veldig krummet linse har kort fokallengde, en flatere linse har lang fokallengde.

Objekt-bilde relasjonen (avsnitt 13.2.1)



- Lysstråler parallelle med optisk akse går gjennom fokalpunktet.
- Lysstråler gjennom linsens sentrum avbøyes ikke.
- En trekant til venstre og en til høyre for linsen gir oss

$$\frac{y}{s} = \frac{y'}{s'} \Rightarrow \frac{y'}{y} = \frac{s'}{s}$$

- To trekanter med felles toppunkt i fokalpunktet til høyre for linsen gir

$$\frac{y}{f} = \frac{y'}{s'-f} \Rightarrow \frac{y'}{y} = \frac{s'-f}{f}$$

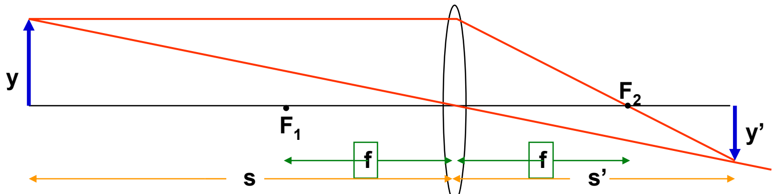
- To uttrykk for y'/y settes lik hverandre og gir

$$\frac{s'}{s} = \frac{s'-f}{f}$$

- Rydder vi litt og får "*objekt-bilde relasjonen*":

$$\boxed{\frac{1}{s} + \frac{1}{s'} = \frac{1}{f}}$$

Objekt-bilde relasjonen (avsnitt 13.2.1)



- Objekt-bilde relasjon: $\boxed{1/s + 1/s' = 1/f}$

  Mer praktisk: $$\boxed{s' = \frac{s\,f}{(s-f)}}$$

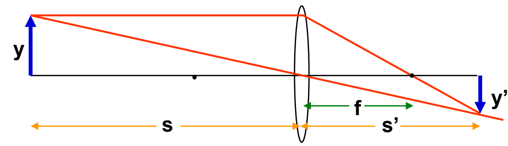
- Forstørrelse: m = y'/y = s'/s
- Hvor stort blir bildet? y' = ys'/s.
  - Setter vi inn det uttrykket vi har for s' ovenfor, finner vi et nyttig uttrykk for størrelsen på bildet i fokalplanet:

$$\boxed{y' = \frac{y\,f}{(s-f)}}$$

Hvor stort blir bildet av Månen ?

- Hvor stort blir bildet av månen med f = 50 mm?
- Månen har en diameter på 3 476 km, og avstanden fra jorda til månen er 384 405 km.
- s = 384 405 km, f = 50 mm, y = 3 476 km i figuren vår

$$\boxed{y' = \frac{y\,f}{(s-f)}}$$



- Da blir
  y' = yf / (s - f) = 3 476 km · 50 mm/ (384 405 km - 50 mm) = <u>0.45 mm.</u>
- Dette bildet fyller bare 0.2 promille av arealet på en 24 · 36 mm film!

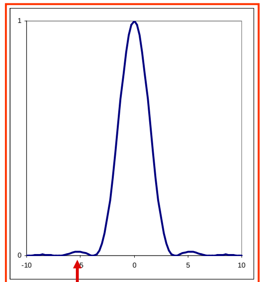
Synsfelt og perspektiv

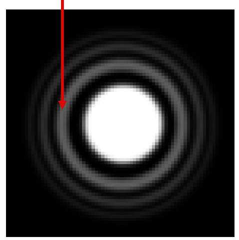
- For et gitt bildeutsnitt vil fokallengden bestemme hvor stort synsfelt vi får.
- Hvis bildeutsnittet i fokalplanet er 24 · 36 mm:
  - f = 28 mm gir et vidvinklet synsfelt: 75º
  - f = 300 mm zoomer inn synsfeltet til bare 8º.

- Fokallengden kan forvrenge perspektivet.
  - Kort brennvidde gir stor nese i *en face* portrett.
  - Telelinser komprimerer dybden i bildet.
- ”Normalobjektiver” gir samme perspektiv som øyet
  - f = 50 mm gir 47º synsfelt på 24 ·36 mm film
  - En liten detektor-brikke i et digitalkamera kan gi normalt perspektiv med en liten linse med kort brennvidde, men oppløsningen vil bli dårligere.

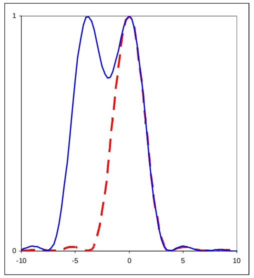
Punktspredningsprofil (PSF)

- En linse avbilder ikke en punktkilde (for eksempel en stjerne) som et lite punkt.
- På grunn av diffraksjon vil en sirkulær linse avbilde en punktkilde som en lys flekk med mørke og lyse ringer rundt, der intensiteten til ringene avtar ganske raskt utover.
- Hver aperture har en”punktspredningsprofil” (PSF) eller diffraksjonsprofil.
  - PSF for en gitt aperture kan beregnes ved hjelp av enkle ligninger.





Rayleigh-kriteriet

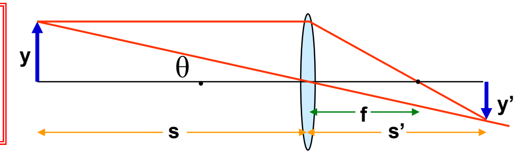


- To lys-punkter kan akkurat adskilles i bildet hvis de ligger slik at sentrum i det ene diffraksjonsmønstret faller sammen med den første mørke ringen i det andre.
- Linse med diameter D, bølgelengde λ.
- La maksimum til den ene falle i første minimum til PSF for den andre.
  - Vinkelen mellom dem er da gitt ved

    $\sin \theta = 1.22 \lambda / D$ radianer.
  - Dette er "Rayleigh-kriteriet".
  - *Vi kan ikke se detaljer mindre enn dette.*

Den minste detalj i et bilde ...

- Fokallengde: f = 35 mm.
- F-tall: f/D = 3 => D = f/3.
- Avstand til objektet: s = 3.5 meter.
- Bølgelengde: λ = 500 nm.



- Vinkeloppløsningen, sett fra linsens sentrum, er gitt ved Rayleigh-kriteriet:

  $\sin(\theta) = 1.22 \lambda / D = 1.22 \cdot 500 \cdot 10^{-9} \cdot 3 / 35 \cdot 10^{-3} = \underline{5.23 \cdot 10^{-3}}$.

- Q: Hva er avstanden y mellom to akkurat adskillbare punkter på objektet ?
- A: Gitt ved: $\tan(\theta) = (y/s)$.

  For små vinkler er $\sin(\theta) = \tan(\theta) = \theta$, når θ er gitt i radianer.

  $y = 3.5 \cdot 5.23 \cdot 10^{-3}$ m $= 1.83 \cdot 10^{-4}$ m $\approx$ 0.2 mm.

- Q: Hva er den tilsvarende avstanden y' i fokalplanet?
- A: $y' = y \cdot f / (s - f)$

  $y' = 0.2 \cdot 35 / (3500 – 35)$ mm $\approx$ 0.002 mm = 2 µm.

De minste detaljer i et bilde …

- … tatt med et kompakt digitalkamera ?
  - Kortere "brennvidde", f = 5.8 mm, og f/D = 3.1 .
- Q: Hva er nå den minste avstand y mellom to punkter på objektet som kan adskilles?
- A: Nå er linsediameteren D = f / 3.1 = 5.8 mm / 3.1 = 1.87 · 10-3 meter .
  - Rayleigh-kriteriet: $\sin(\theta) = 1.22 \lambda / D = 1.22 \cdot 500 \cdot 10^{-9} / 1.87 \cdot 10^{-3}$
  - Men vi må fortsatt ha at $\tan(\theta) = (y / 3.5)$, og $\sin(\theta) = \tan(\theta) = \theta$.
  - Da får vi at $y = 3.5 \cdot 1.22 \cdot 500 \cdot 10^{-9} / 1.87 \cdot 10^{-2} \approx$ 1.1 mm.
  - Det mer moderne kameraet gir altså 5.5 ganger dårligere oppløsning.
- Q: Hva er den tilsvarende avstand y' i bildeplanet?
- A: Vi husker y' = yf / (s - f), altså: y' = 1.1 · 5.8 / (3500 – 5.8) mm ≈ 2 µm.
  - Kravene til fysisk samplingstetthet i bildeplanet er altså omtrent de samme.

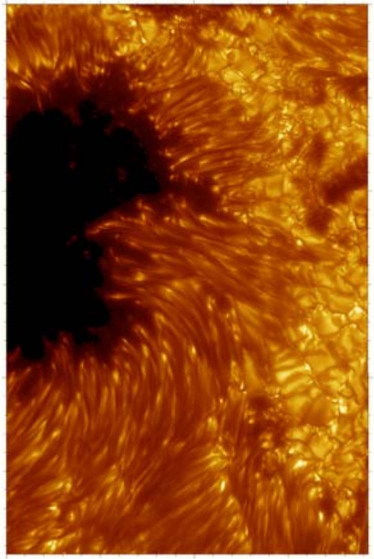
De minste detaljer i et bilde …

… i et bilde tatt med et mobil-kamera ?

- Anta f = 4 mm og f / D ≈ 4. Da er D = f / 4 = 1.
- På 3.5 meters avstand vil den minste synlige detaljen være
  - $y = 3.5 \cdot 1.22 \cdot 500 \cdot 10^{-9} / 1.0 \cdot 10^{-3}$ m $\approx$ 2.1 mm.
  - En faktor 2 dårligere oppløsning enn det kompakte digitalkameraet.
- I bildeplanet er den minste detaljen gitt ved
  - y' = yf / (s - f) = 2.1 · 4 /(3500 - 4) ≈ 2.4 µm.
- Synsfelt er 47 · 34 grader, med 640 · 480 detektorer i fokalplanet.
  - På 3.5 meters avstand vil én detektor dekke 0.45 · 0.45 mm.
  - Mer enn nok detektorer til å utnytte oppløsningen.
  - 2.3 ganger oversampling (2.1 / 0.45 /2 ≈ 2.3).

Eksempel på ekstremt høy oppløsning

- Et teleskop med 1 meter diameter har tatt noen av de aller beste bildene av Sola.
- For $\lambda = 600 \cdot 10^{-9}$ meter og D = 1 meter gir <u>Rayleigh-kriteriet</u> $\theta = 4.2 \cdot 10^{-5}$ grader
- Med en avstanden Jord - Sol $\approx 150 \cdot 10^6$ km blir teoretisk oppløsning ca 100 km.
- Eksempel: del av en solflekk.
- Merke langs bildekanten for hver 1000 km.
- Teleskopet benytter et ”aktivt speil”: bølgefronten til lyset analyseres på vei inn i teleskopet, og små stempler på baksiden av speilet korrigerer dets fasong i nær sann tid, slik at det kompenseres for luft-uroen.
- *© Swedish 1-meter Solar Telescope (SST), Royal Swedish Academy of Sciences.*



Andre sensorer enn øyet

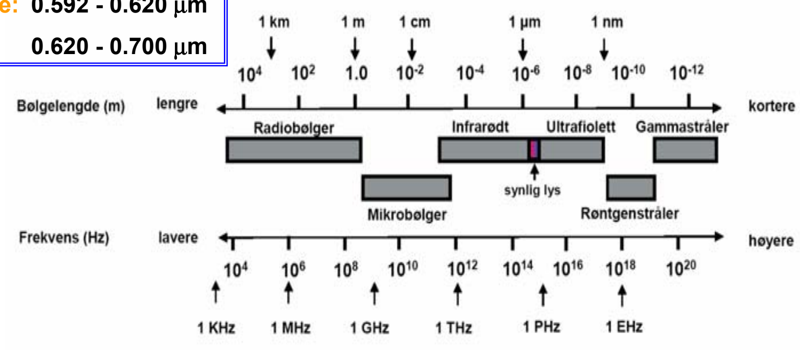
- Aktive og passive sensorer: ”belyse og se” eller bare ”se”.
- Optisk satellittbilde: Landsat
- Radarbilde fra satellitt: SAR
- Infrarødt satellittbilde
- Medisinsk ultralyd
- Røntgen og CT
- NMR – magnetisk resonnans
- Sonar, seismikk – lyd
- Mikroskopi
- Laser avstand scanner

Det elektromagnetiske spekteret

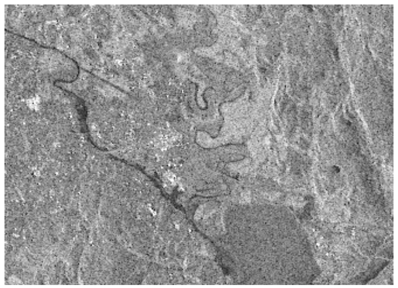
Fiolett:	0.400 - 0.446 µm
Blå:	0.446 - 0.500 µm
Grønn:	0.500 - 0.578 µm
Gul:	0.578 - 0.592 µm
Oransje:	0.592 - 0.620 µm
Rød:	0.620 - 0.700 µm

<u>Sammenheng mellom bølgelengde og frekvens:</u>

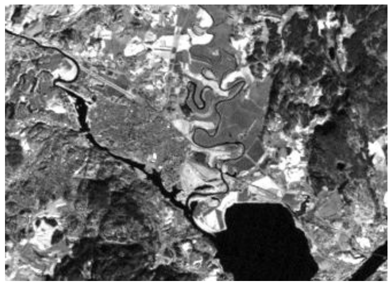
$c = f\lambda$ (bølgeligningen)
c = lysets hastighet ($3 \cdot 10^8$ m/s)
λ = bølgelengde (m)
f = frekvens (svingninger pr. sekund, Hz)



Eksempel: radarbilde vs. optisk

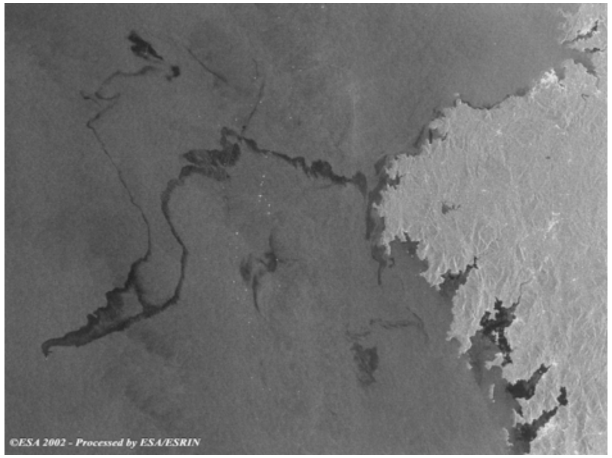


Bilde fra ERS-1 SAR-satellitten
Radaren viser røffheten på overflater



Landsat-bilde fra samme område

Eksempel: radarbilde av oljesøl fra skipsforlis



- Radar detekterer overflatens ”røffhet”.
- Olje demper vindbølger på vann.
- Radarbilder fra satellitt og fly kan overvåke oljesøl
  - Fra skipsforlis
  - Utslipp fra skip
  - Utslipp fra oljerigger
- Eksempel:
  - M/S Prestige, 2002.

  © ESA

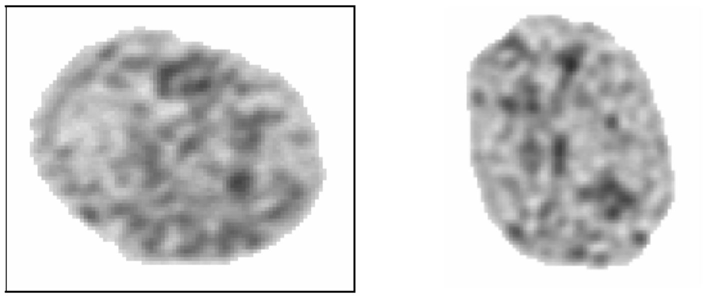
Satellittbilder med lav og høy oppløsning

- Lavoppløsningsbilder gir oversikt, f.eks innen meteorologi.

- Høyoppløselige bilder nyttige til kartlegging, arealplanlegging, spionasje, …

Medisinsk mikroskopi

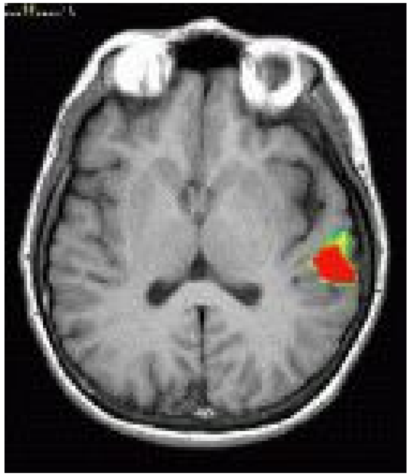
- Eksemplet viser mikroskopi-bilder av cellekjerner fra kreftsvulst i eggstokkene (ovarie) for en pasient med god prognose (venstre) og en pasient med dårlig prognose (høyre).
- Visuelt kan man ikke se forskjell, men med matematisk analyse av teksturen kan man klassifisere dem riktig.



CT og MR

- Plasser røntgenkilden foran en pasient, og en detektor-matrise bak.
  - Roter dette oppsettet sakte rundt pasienten – ta flere bilder.
  - Etterpå kan vi regne oss fram til absorpsjonen i hvert punkt i pasienten.
  - Vi får altså et 3-D røntgenbilde.
  - Kan ta snitt-bilder av dette volumet i de plan og de retninger vi ønsker.
  - Dette kalles” Computed Tomography”, forkortet CT.
- Magnetisk resonans avbildning (MR) avbilder protonene i kroppen.
  - Dette gjøres ved å eksitere hydrogen-atomene,
  - Registrerer hvordan atomene de-eksiteres.
  - Tiden dette tar er avhengig av vevstype og av sykdomstilstand.

Funksjonell MR



- Funksjonell MR (f MRI) avbilder de delene av vevet hvor oksygenforbruket er høyt eller endres mens man gjør opptaket.
- Kan f.eks avbilde de delene av hjernen som er i aktivitet når man utfører en spesiell oppgave.
- Eksempel:
  axialt snitt fra lyttende person (Miami Childrens Hospital)
- Tilsvarende verbal aktivitet ligger nær dette området i hjernen (ved venstre tinning).

Flerdimensjonale bilder

- Et 2-D bilde er en projeksjon av et 3-D objekt.
  - For å gjenskape objektet i 3-D må vi ha flere 2-D projeksjoner.
  - Vi må løse "korrespondanseproblemet",
    hvilke punkter i bildene svarer til samme punkt i virkeligheten.
  - Dette gjør hjernen vår. Kombinerer høyre og venstre bilde - *stereo-syn*.
- Laser avstandsmåler gir 2-D avbildning av den tredje dimensjonen.
- CT og MR gir 3-D bilder av organer inne i kroppen vår.
- Mikroskopi
  - konfokal mikroskopi belyser og avbilder på flere dyp i vevet,
  - serielle tynne snitt av cellepreparater gir 3D bilder med meget høy oppløsning.
- En tidssekvens av 2-D bilder kan sees på som et 3-D datasett.
- En tidssekvens av 3-D bilder kan betraktes som et 4-D bilde.
- Avbildning på flere bølgelengder for hvert tidspunkt gir et 5-D bilde.

Digitalisering i tre enkle steg

- Et kontinuerlig bilde er en reell funksjon f(x,y)
  av to (eller flere) reelle og kontinuerlige variable.

1. Vi legger på et rutenett, og beregner gjennomsnittsverdien av f(x,y) i hver rute.
   - Dette er samplingen.
     - Rutenettet er vanligvis kvadrater.
     - Vi har nå en reell funksjon f'(x',y'), der heltallene x' og y' gir nummereringen av rutene.
2. Vi skalerer f'(x',y') slik at den passer innenfor det tall-området vi skal bruke som gråtoneskala.
3. Vi kvantiserer verdiene av f'(x',y') til nærmeste heltall i gråtoneskalaen.
   - Vi har nå funnet f''(x',y'), som er en heltalls funksjon av to heltalls variable.

Viktige begreper

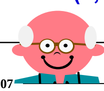
- Et digitalt bilde er en funksjon av to (eller flere) heltallsvariable f(x,y) (x og y er heltall)
- Et 2-dimensjonalt digitalt bilde er en 2-dimensjonal array/matrise. Dette kalles raster-representasjon.
- Hvert element i matrisen kalles et piksel, og angis ved koordinater x og y.
- Tallverdien til hvert piksel angir intensiteten til pikslet.
- Lagres pikselverdiene i matriser, trenger vi ikke ta vare på koordinatene.

x

y

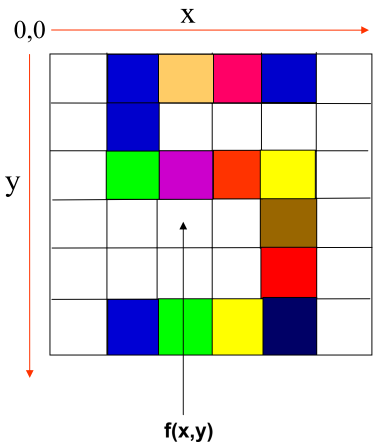
1	5	7	3	6	4
3	7	7	6	4	3
2	4	4	3	4	4
3	3	1	1	1	1
3	3	3	9	9	9
5	6	5	5	6	6

Merk: origo (0,0) er ofte oppe i venstre hjørne i bildet.
Det første pikslet kan ha indekser (0,0) eller (1,1)



Representasjon av digitale bilder

- f(x,y) er bildeverdien i piksel (x,y).
- f(x,y) er et tall som forteller noe om intensiteten (lysstyrken) målt i punktet (x,y).
- f(x,y) kan også være en vektor, f.eks. (r,g,b) for et fargebilde.
- f(x,y) må representeres med en gitt datatype (integer, byte, float etc.).



Eksempel - Pikseltyper

0	1	1	1	0
0	1	0	0	0
0	1	1	1	0
0	0	0	1	0
0	1	1	1	0

Bit

0	25	23	13	0
0	14	0	0	0
0	14	21	31	0
0	0	0	21	0
0	41	14	51	0

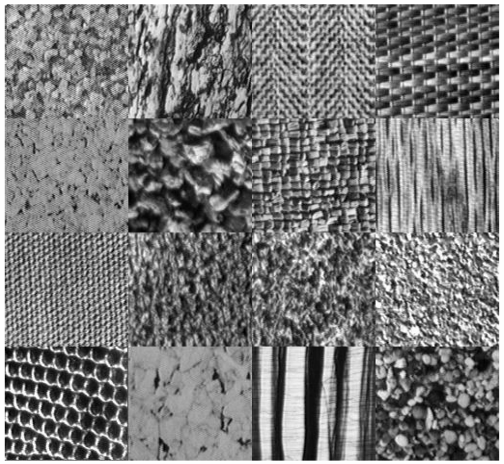
Byte

0.1	14.5	12.3	51.3	0.0
0.2	41.6	0.0	0.2	0.3
0.3	18.7	33.1	44.1	0.8
0.1	0.1	0.1	1.2	0.4
0	55.5	76.4	24.5	0.4

Real

Naturlige frekvenser i bilder

- Naturlige bilder: har et "uendelig" sett av frekvenser.
- Fourier-teori: et bilde kan beskrives ved en <u>endelig</u> sum av sinus og cosinus-signaler med ulike frekvenser.
- Digitale bilder består av et endelig sett frekvenser.

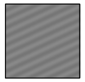


Sampling

- Bildet varierer i intensitet med ulike frekvenser.
- Antar bildet inneholder av et endelig antall frekvenser.
- En høy frekvens beskriver et mønster som varierer hurtig i bildet, lav frekvens noe som varierer langsomt.
- Nyquist-kriteriet: Samplingsraten må minst være dobbelt så stor som den høyeste frekvensen i bildet.
  - Dette betyr: Vi må sample minst to ganger pr. periode av det objektet som varierer hurtigst i det kontinuerlige bildet.
- Det digitale bildet må altså ha minst to piksler pr periode for den minste periodisk struktur i det kontinuerlige bildet.
- I tillegg må vi ha god nok optikk til å avbilde slike strukturer !

Kommentar om sampling av bilder

- Når et kamera tar bilde av et objekt, vil hvert piksel i bildet inneholde lys målt fra hele det området som pikselet dekker.
- Eksempel: la oss si at 1 piksel dekker det området som er vist til høyre, og at dette lille området inneholder noe fin-struktur:

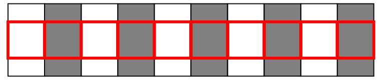


- Dette representeres etter samplingen ved gjennomsnittlig lysstyrke i området:

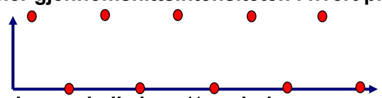
- Vi har målt en middelverdi over et areal, og gjengir det likedan.
- Dermed er all struktur mindre enn pikselstørrelsen blitt visket ut.
- Dette er forskjellig fra sampling av lyd, der samplet lages ved at man "fryser" lydstyrken på et gitt tidspunkt.

Nyquist … på et stakitt …

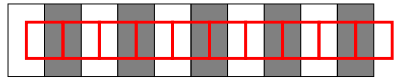
- Ta et analogt bilde av et stakitt med 10 cm sprosser og 10 cm mellomrom.
  - Anta at det er to piksler per periode …



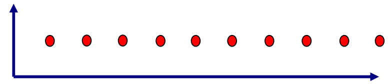
  - Vi finner gjennomsnittsintensiteten i hvert piksel …



  - Så forskyver vi pikslene ¼ periode …



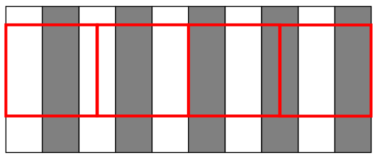
  - Og ser at pikselverdiene blir helt annerledes …



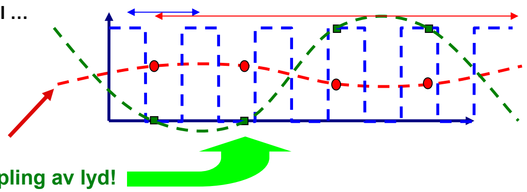
- Vi må ha <u>MINST</u> to piksler per periode!

Et undersamplet Nyquist-stakitt …

- Ta et analogt bilde av et stakitt med 5 sprosser og 5 mellomrom per meter.
  - Vi har en periodisk struktur
    - f = 5 svingninger per meter
    - Periode T = 20 cm
  - Anta at pikslene svarer til 25 · 25 cm, dvs f_s = 4 sampler per meter.



  - Finner gjennomsnitt i hvert piksel …
    - Amplituden reduseres …
    - Vi får aliasing, fa = $|f_s - f|$ = |4 - 5| = 1
    - Perioden $T_a = 1/f_a$ = 1 meter



- Dette er annerledes enn ved sampling av lyd!
  - Vi får redusert amplitude i forhold til lydsampling
  - Vi får samme aliasingfrekvens
  - Vi kan få faseforskyvning

Valg av rutenett/gridstørrelser

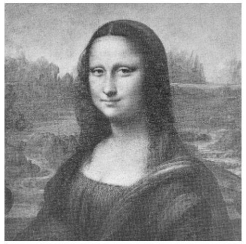
- Gridstørrelsen er ofte gitt av sensorer:
  - Video: f.eks. 640 x 480 (kolonner · linjer)
  - Digitalt kamera: f.eks. 1024 · 768 (0.7 megapiksel)
    - Bedre kamera: 1760 · 1168 (2.1 megapiksel)
    - Enda bedre: 2272 · 1704 (4.0 megapiksel)
- Gridstørrelse og fokallengde bestemmer den romlige oppløsningen, gitt som en vinkel.
- Hvis vi kjenner avstand fra kamera til objektet som avbildes, kan vi angi romlig oppløsning som størrelsen av hvert piksel på objektet, for eksempel i meter.

Romlig oppløsning avhenger av anvendelsen

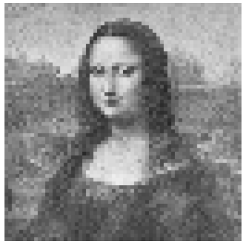
- Et kamera på mobiltelefon skal sende bildet over en GSM-linje og har relativt dårlig oppløsning.
- Bilder på Internett skal også overføres raskt og ligger ofte lagret med dårlig oppløsning.
- Skal vi skanne inn et håndtegnet kart, må vi ha god nok oppløsning til at tekst og linjer kommer korrekt med.
- Skal vi ta bilder av jorden fra satellitt, kan vi enten dekke
  - et stort område med liten oppløsning (1km)
  - et mindre område med moderat oppløsning (10m)
  - et enda mindre område med veldig høy oppløsning (10 cm)

Eksempler - romlig oppløsning


512 · 512 piksler


256 · 256 piksler


128 · 128 piksler


64 · 64 piksler


32 · 32 piksler


16 · 16 piksler

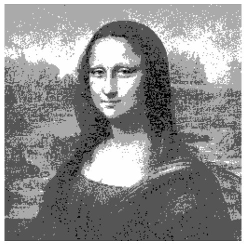
Eksempler - antall biter pr. piksel


8 biter


6 biter

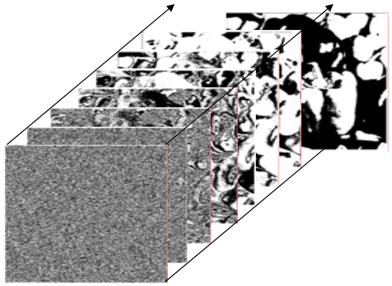

4 biter


3 biter


2 biter


1 bit

Bitplan i et gråtonebilde



- Et 8 bits bilde har 8 bitplan.
  - LSB forrest i figuren - bare tilfeldig støy.
  - MSB bakerst i figuren.
- Pikselverdien i (x,y) er gitt ved:

$f(x,y) = b_0(x,y) \cdot 2^0 + \ldots + b^7(x,y) \cdot 2^7 = \sum b_i(x,y) \cdot 2^i$

  - Se representasjon av tall, kapittel 7.
- Hvis bit 1- 6 i (x,y) = 1:
  - f(x,y) = 1+2+4+8+16+32+64 = 127.
  - MSB = 0 (svart) => f(x,y) Є [0,127].
  - MSB = 1 (hvit) => f(x,y) Є [128, 255].
  - MSB-planet er en terskling av f(x,y).

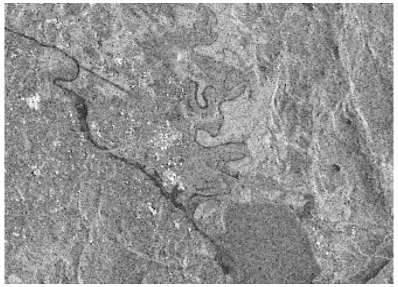
Kvantisering og datatyper

- Hvert piksel lagres vha. *n* bit
- Pikselet kan da inneholde verdier fra 0 til 2^n-1
- 8 bits bilder: 256 ulike verdier (0-255)
- 16 bits bilder: 65536 verdier
  - unsigned: fra -32 768 til 32 767
  - signed: fra 0 til 65 535
- 32 bits integer
- 32 bits float
- Merk: Display og videre bildeanalyse av det kvantiserte bildet stiller ulike krav til presisjon.

Eksempel: plassbehov

Radarbilde fra ERS-satellitten:

- Dekker 100 · 100 km
- Pikselstørrelse 20 · 20m
- 5000 · 5000 piksler



Utsnitt av ERS1 SAR bilde

8-biter per piksel:	25 MB
16 biter per piksel:	50 MB
32 biter per piksel:	100 MB

Bilder tar mye plass

- En side i en lærebok er på 50 linjer a 80 tegn.
- Hvor mange bok-sider svarer et 4 Mpix fargebilde til ?
  $(4 \cdot 10^6 \cdot 3 \cdot 8) / (50 \cdot 80 \cdot 8)$ bok-sider = $3 \cdot 10^3$ bok-sider
- Hvor stort må et 24 bits fargebilde være for at utsagnet ”et bilde sier mer enn 10 000 ord” skal være riktig ?
  10 000 ord · 6 tegn/ord · 8 bits/tegn / 24 bits/piksel = 20 000 piksler, altså et kvadratisk bilde med 141 piksler hver vei.
- For et gråtonebilde: 10 000 · 6 · 8 / 8 = 60 000 piksler, altså et kvadratisk bilde med 245 piksler langs aksene.

To representasjoner for bilder

- Det er to fundamentalt forskjellige måter å representere et bilde på:

1. Lagre alle pikselverdiene (gråtoneverdi, eller fargekomponenter)
2. Lagre en parametrisert beskrivelse av bildets innhold

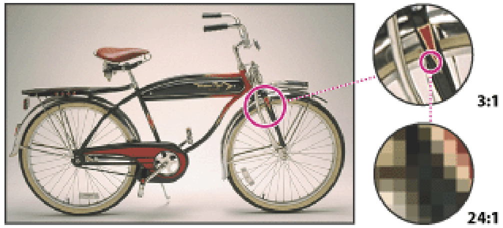
- Den siste metoden krever at bildet deles opp i objekter, og at hvert objekt beskrives ved en rekke parametre.
- Dette forutsetter

1. Enten at bildet er forholdsvis enkelt (skisser, tegninger, CAD, kart, …)
2. Eller at det brukes veldig mange parametre for å generere noe som ligner på et naturlig bilde (”virtual reality”).

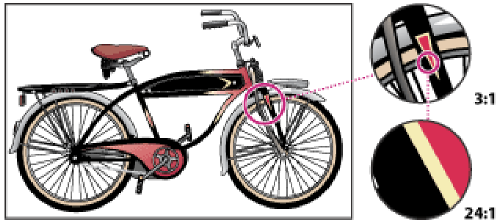
- I det siste tilfellet er det naturlig å la objekter ha overflate-egenskaper (varierende farge, refleksjonsegenskaper, tekstur, etc.).

Raster- eller vektor-bilde

- Rasterbilder er enkle å lage
- Tåler ikke mye forstørrelse
- Krever mye lagringsplass

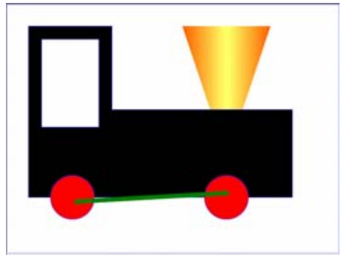


- Vektorbilder må konstrueres
- Krever god programvare
- Tåler mye forstørrelse
- Krever mindre lagringsplass



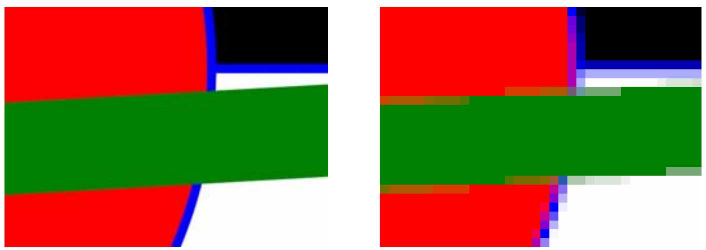
Scalable Vector Graphics (SVG) – XML vektorgrafikk

- SVG er et språk for å beskrive 2-D grafikk i XML.
- Egner seg best for regulære, konstruerte former
- Et objekt kan ligge foran / bak et annet objekt.
- Naturtro scener kan oppnås ved hjelp av ”støy” og fraktaler.
- Grafiske objekter kan grupperes og transformeres.
- Transformasjoner kan nøstes,
  man kan ha transparente masker,
  filtre og templater.



Vektor og raster

- Vi kan forstørre deler av et vektorbilde uten at det går ut over bildekvaliteten.
- For rasterbilder er dette begrenset av pikselstørrelsen.
- Merk at vektorbilder også må konverteres til rasterbilder før framvisning.



Animasjon

- Når objektene er beskrevet som vektor-grafikk, kan man flytte på øye-punktet, og få fram et nytt bilde av objektene.
- Hvis øyet beveger seg langs en parametrisert bane,
  vil man få en sekvens av resultat-bilder.
- Man kan legge på fade-in og fade-out effekter.
- Objekter kan endre størrelse, form og farge underveis.
- Objekter kan bevege seg i forhold til hverandre, og rotere.
- Netto-effekten er en animasjon.
- Er animasjonen tilstrekkelig kompleks,
  kan ”bilde”-sekvensen se ganske virkelighetsnær ut,
  uten at det er lagret noen bilder.

Litt om digitale kart

- Digitale kart ligger lagret på vektorform:
  - Regioner ligger som flater
  - Veier ligger som linjer
  - Bygninger o.l. ligger som punkter med en viss størrelse.
  - Stedsnavn ligger som tekst
  - Høydekurver ligger som linjer
  - Vann ligger som konturer
- Kartene ligger med mange lag, f.eks. for M711-serien:
  - Stedsnavn, bygg og anlegg, høydepunkt, bygg og anlegg, veier, markslagsgrenser, samferdsel, høydedata, kommunegrenser, fylkesgrenser, vannkontur, elver, etc.

Dette bør du ha fått med deg i dag …

- Hornhinna og øyelinsen danner et bilde på netthinna.
- Øyelinsen kan forandre form slik at den kan fokusere på objekter på ulike avstander uten å endre avstanden mellom linsen og bildet.
- I netthinna er det ca 130 millioner lysfølsomme detektorer.
  - Ca 120 millioner staver er svært lysfølsomme, men ser ikke farger.
  - 6-7 millioner tapper gir oss et fargesyn med høy geometrisk oppløsning.
- I synshjernebarken sitter mengder av kantdetektorer som finner kanter og linjer i bildet
  - Separate detektorer for hvert øye, for ulike orienteringer og tykkelser.
- Vi kan se ca 60 svarte linjer på hvit bunn per grad i synsfeltet.
- Vi kan bare oppfatte ca 60 lysblink per sekund

Dette bør du ha fått med deg i dag …

- Et 2-dimensjonalt digitalt bilde er en 2-dimensjonal heltalls matrise (array).
- Hvert tall i matrisen svarer til et lite (oftest kvadratisk) område i det opprinnelige analoge bildet.
- Hvert tall i matrisen kalles et piksel (”picture element”).
- I gråtone-bilder angir piksel-verdien hvor lyst / mørkt det lille området i det opprinnelige bildet er.
- I fargebilder er piksel-verdien en vektor (3 heltall).
- Hvert piksel har implisitte romlige koordinater (x,y), gitt ved posisjonen i matrisen.

Dette bør du også ha fått med deg …

- Det digitale bildet må ha minst to piksler pr periode for den minste periodisk struktur i det kontinuerlige bildet.
- Vi må ha god nok optikk til å avbilde slike strukturer ! Vinkeloppløsningen er gitt ved $\sin \theta = 1.22 \cdot \lambda / D$.
- Vi velger en kvantisering som tilfredsstiller de krav som display og/eller senere bildebehandling og analyse stiller.
- Digitale bilder krever generelt mye lagringsplass.
- Vektorrepresentasjon kan være et alternativ til rasterbilder.