IN2070: Digital Bildebehandling

Introduksjon: Avbildning

Marius Aasan

Praktisk Info
Hvem, Hva, Hvor?

- **Pensum:**
  - Digital Image Processing (4th Ed.) - Gonzales & Woods
  - Slides + Tilleggsstoff
- **Programmeringsspråk:**
  - Python
- **Innleveringer og Eksamen:**
  - Obliger: Uke 39-42 + Uke 44-46
  - Skriftlig Eksamen: 05.12.2023 - Silurveien 2
- **Gruppelærer:**
  - Mathias Mellemstuen - mathime@ifi.uio.no - Tor / Fre 10-12
  - Avklaring om hvilken dag kommer.
- **Kontakt og spørsmål?**
  - Send mail eller melding på Mattermost / Discourse

Forelesere



Are Charles Jensen, Førsteemanuensis DSB
arej@ifi.uio.no



Marius Aasan, Stipendiat DSB
mariuaas@ifi.uio.no



Martine Tan, Stipendiat DSB
matan@ifi.uio.no

Praktisk Info
Discourse

Første år med **Discourse forum** i IN2070!

Brukes til:

- Faglig diskusjon.
- Spørsmål og svar.
- Kollokvie / lesegrupper.
- Praktisk info.

- Kan spørre anonymt! Ingen dumme spørsmål!
- **Kursside**: https://astro-discourse.uio.no/c/in2070-23h
- Meld dere gjerne opp så snart som mulig.



Praktisk Info
Plan for Kurset

- **Uke 1**: Avbildning og Optikk
- **Uke 2**: Sampling og Kvantisering
- **Uke 3**: Geometriske Operasjoner
- **Uke 4**: Gråtonemapping
- **Uke 5**: Histogrambaserte Operasjoner
- **Uke 6-7**: Naboskapsoperasjoner
- **Uke 8**: Segmentering
- **Uke 9-10**: Fourier transform
- **Uke 11-12**: Kompresjon og Koding
- **Uke 13**: Morfologi
- **Uke 14**: Fargerom
- **Uke 15**: Repetisjonsforelesning



Praktisk Info

Om obligatoriske innleveringer

- Vi ser av og til forsøk på plagiering og fusk i besvarelsene. **Reaksjonen fra instituttet kan være alvorlig.**
- Les gjerne www.uio.no/studier/admin/obligatoriske-aktiviteter/mn-ifi-oblig.html
- **Hovedregel:**
  - Levér din egen kode på oppgaver som sier "programmer selv"!
  - Ellers, er det OK å bruke annen kode så lenge kilden er angitt.
- Innleveringen **er et individuelt arbeid** og skal være **ditt produkt**.



UNIVERSITETET I OSLO

Introduksjon:
En verden av bilder!

- Ca. 40% av alle nervefibre i hjernen er koblet til retina, og ca. 50% av hjernebarken er relatert til visuelle stimuli. Flere nevroner i hjernen er dedikert til det visuelle sanseapparat enn de andre 4 sansene kombinert!
- Mennesker kan behandle informasjon i bilder opptil 60,000 ganger raskere enn tekst. Et helt bilde kan prosesseres ila. 13 ms.
- Hvert dag produserer menneskeheten mellom 3-5 petabyte med stillbildedata fra smarttelefoner online.
- **Bilder og visuell stimuli står sentralt i menneskelig kommunikasjon og virkelighetsoppfattelse.**



Bildebehandling
Behandling eller Analyse?

- **Bildebehandling**: Prosessering av bilder for å fremheve spesifikk karakteristikk eller representasjon for praktisk bruk eller videre analyse.
- **Bildeanalyse**: Hente ut spesifikk informasjon vedrørende et bilde.
  - ***Klassifikasjon***: Er bilde av en lemur eller en elefant?
  - ***Deteksjon***: Posisjoner til alle personer i bilder.
  - ***Semantisk segmentering***: Punktvis klassifisering av bilder.

Tradisjonelt: Behandling er et nødvendig steg ***før*** analyse.

Idag: Behandling og analyse ***mer parallellisert*** med datadrevne algoritmer.

Videre Studier

Hvilke fag anbefales å ta videre?

Bachelor / Masternivå:

- IN3310/4310 - Deep Learning for Image Analysis
- IN3190/4190 - Digital Signal Processing
- IN3050/4050 - Introduction to Artificial Intelligence and Machine Learning

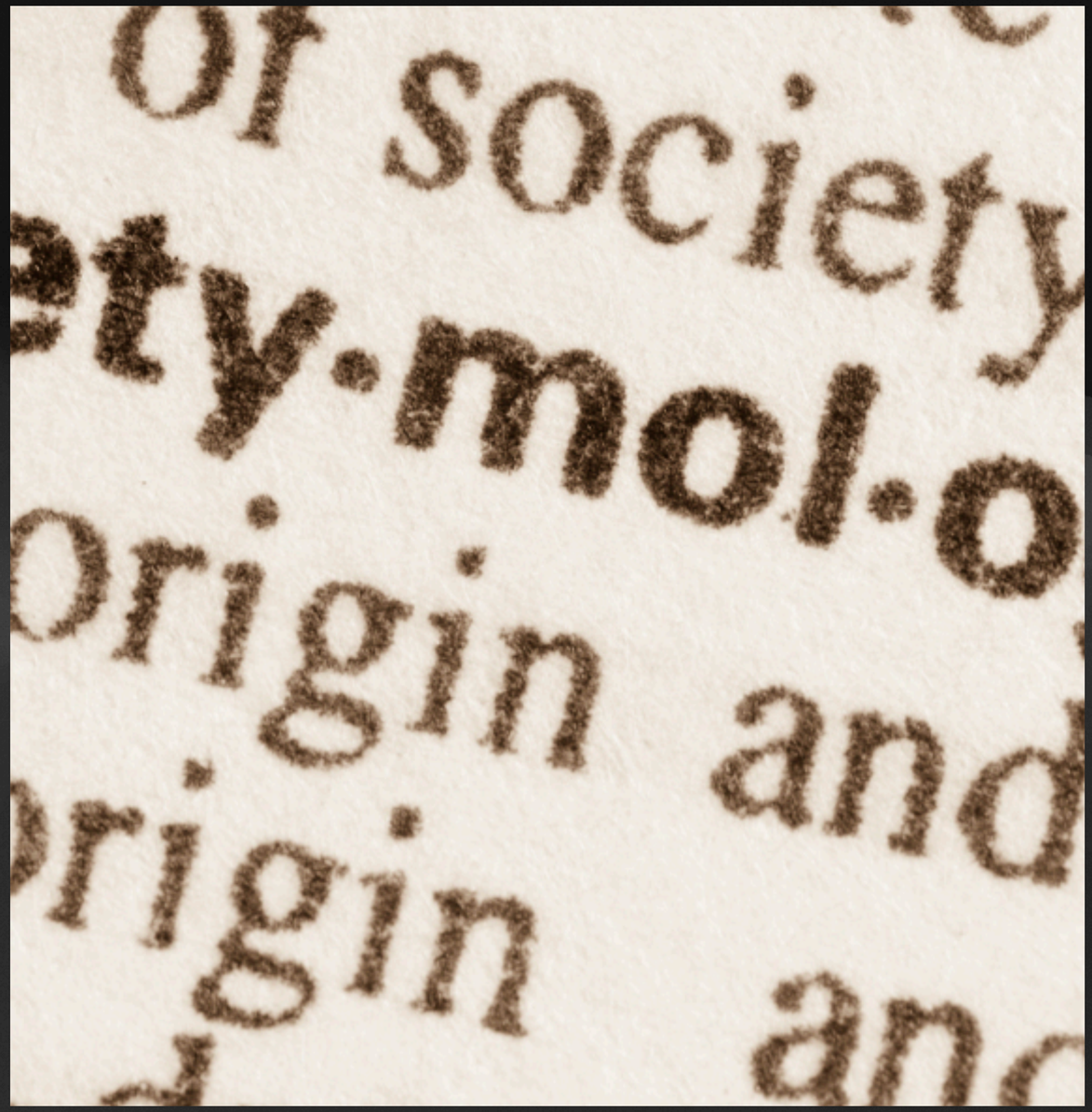
Master / PhD Nivå:

- IN5310/9310 - Advanced Topics in Deep Learning for Image Analysis
- IN5450/9450 - Array Signal Processing
- IN-STK4300/9300 - Statistical Learning Methods in Data Science
- STK4051 - Computational Statistics

Hva menes med “bilde”?

Semantikk og Etymologi

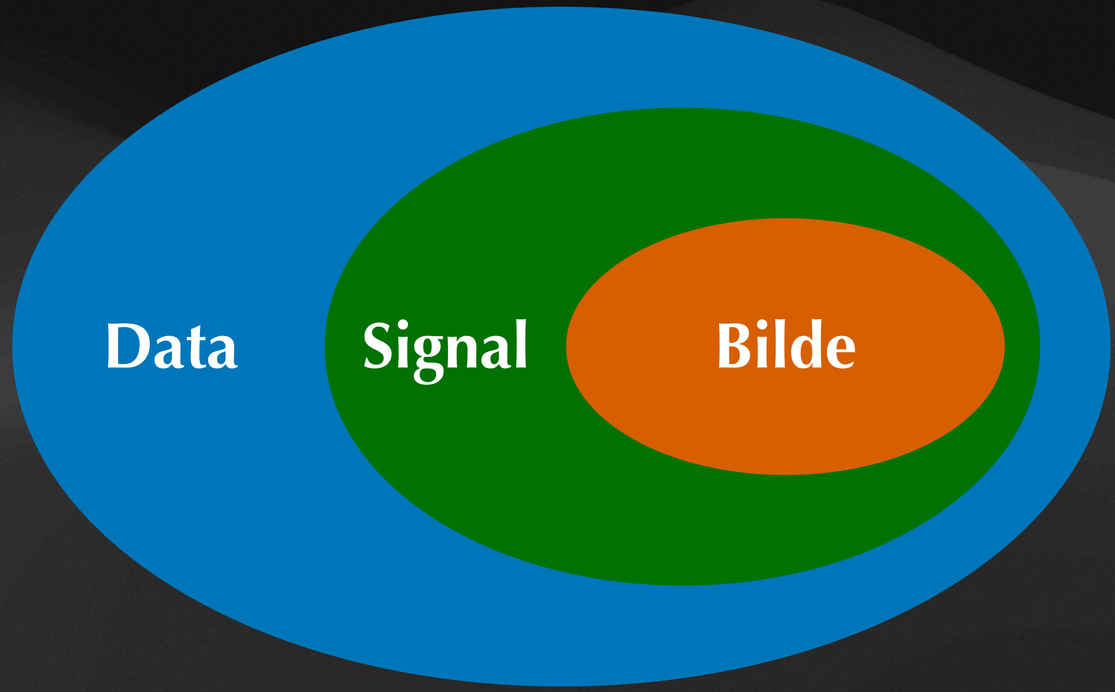
- **Fra latin** *imago*:
  - Av likhet, en kopi, etterlikning eller et utsnitt fra en original.
  - Samme ordstamme som *imitari*: å imitere, etterlikne.
- **Definisjoner**:
  - En fremstilling, avtrykk, eller representasjon. Avbildning.
  - Et ikon, skildring, eller en metafor.
  - Brukes ofte om en todimensjonal representasjon av virkeligheten.



Hva menes med “bilde”?

Data, Signal, eller Bilde?

- **Data**: Representative målinger av observasjoner av interesse.
- **Signaler**: *Homogen data* [1] med en utstrekning over tid og/eller rom.
- **Bilder**: Signaler med eksplisitt utstrekning over rom.



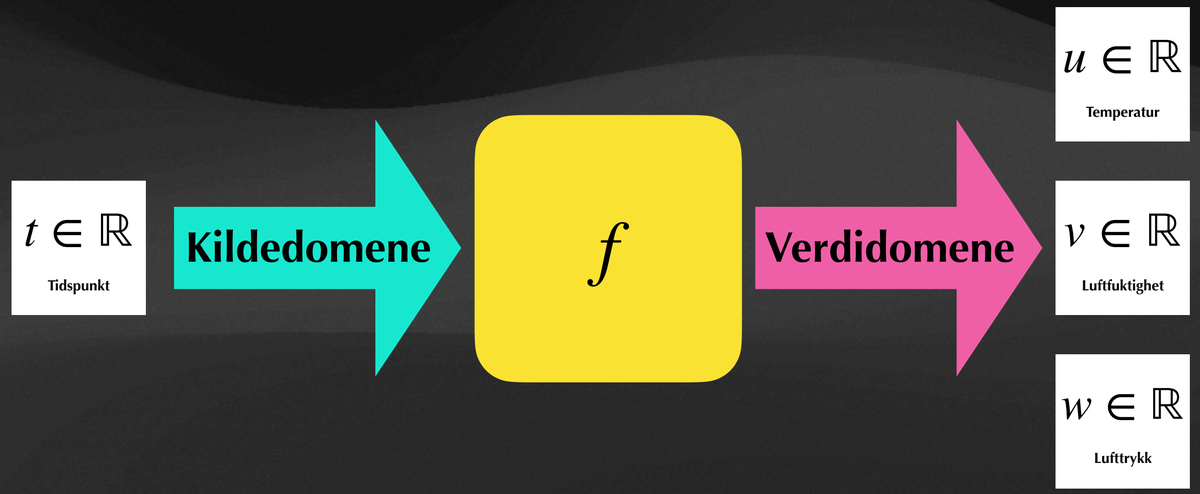
[1] Homogen data: Målinger av samme type observasjoner med uniform representasjon.

Hva menes med “bilde”?

Bilder som Signaler (1)

Ett **signal** er målinger som en **funksjon** av tid eller rom.

- Du “mater" et funksjonssignal med f.eks. ett tidspunkt, og den “spytter ut” en måling.
- Vi skriver $f : \mathbb{R} \to \mathbb{R}^c$; som vi leser som følger: f er en funksjon fra de reelle tallene ($\mathbb{R}$) til en måling i de reelle tallene med c-dimensjoner ($\mathbb{R}^c$). Med andre ord; f tar oss fra ett **kildedomene** til ett **verdidomene**.
- **Eksempel**: En måling på en værstasjon ved tidspunkt t av temperatur, luftfuktighet, og barometrisk trykk ($c = 3$).

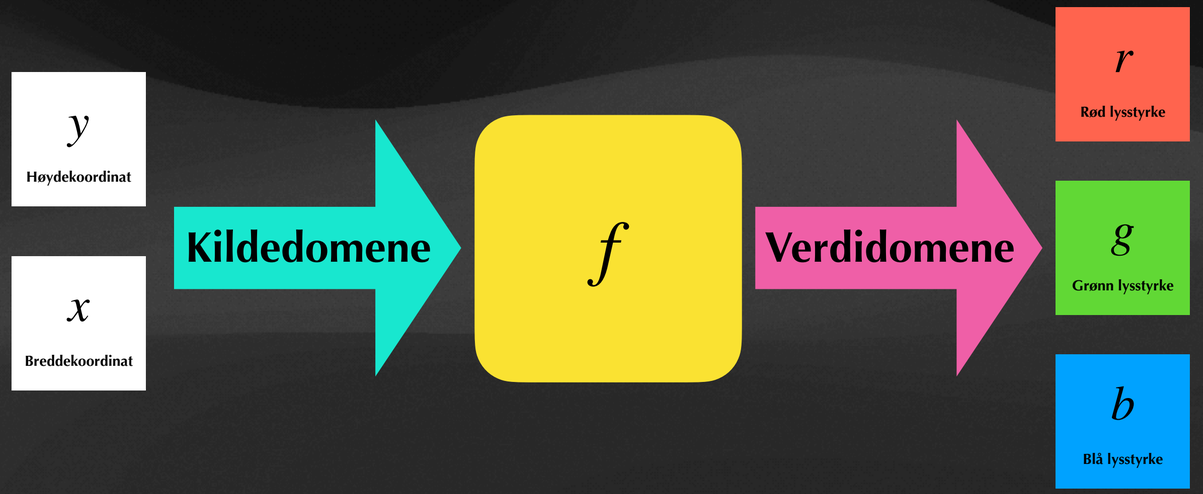


Hva menes med “bilde”?

Bilder som Signaler (2)

Et **bilde fra et kamera** kan derfor sees på som et type signal der hvert **punkt** i bildet har en tilegnet måling av **lys**.

- Med andre ord, et bildesignal $f : \mathbb{R}^d \to \mathbb{R}^c$ har gjerne flere dimensjoner i **kildedomenet**, representert ved romlige koordinater. (f.eks. bruker vi gjerne x, y for todimensjonale stillbilder).
- **Eksempel**: En **dataskjerm** er en funksjonell representasjon av et bildesignal. Hvert punkt (**piksel**) tilegnes en **lysmetning** i tre kanaler; rød, grønn, og blå.



Hva menes med “bilde”?

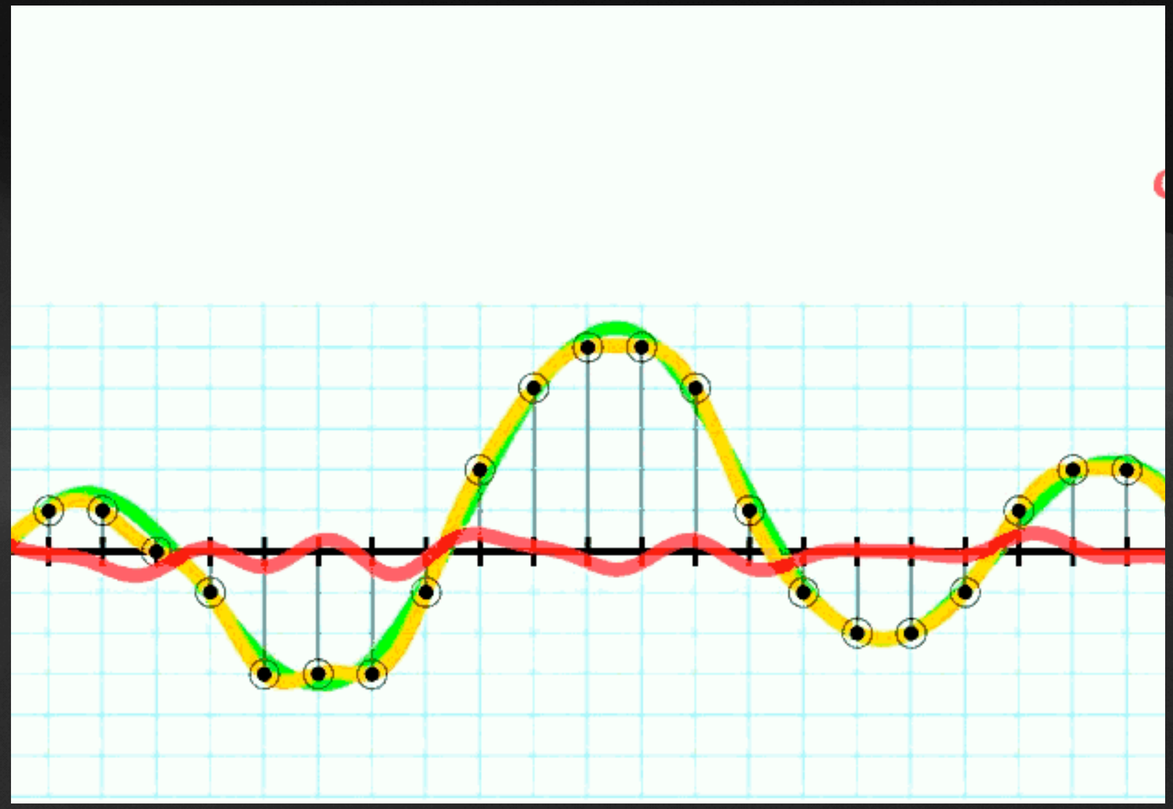
Bilder som Signaler (3)

Bilder og signaler lar seg ikke behandle som **kontinuerlige**, eller **analoge** på datamaskiner.

- **Kildedomene**: endelig antall posisjoner (oppløsning).
- **Verdidomene**: endelig nøyaktighet i målinger (kvantisering).

Behandling av bilder og signaler på datamaskin begrenser oss til de **digitale domener…**

…selv om vi gjerne liker å forestille oss bilder som funksjonelle signaler.



Hva menes med “bilde”?

Kun måling av lys?

Vi knytter ofte ordet “bilde” til synssans og lys, siden vi hovedsaklig observerer virkeligheten via **bilder produsert av vårt synsapparat…**

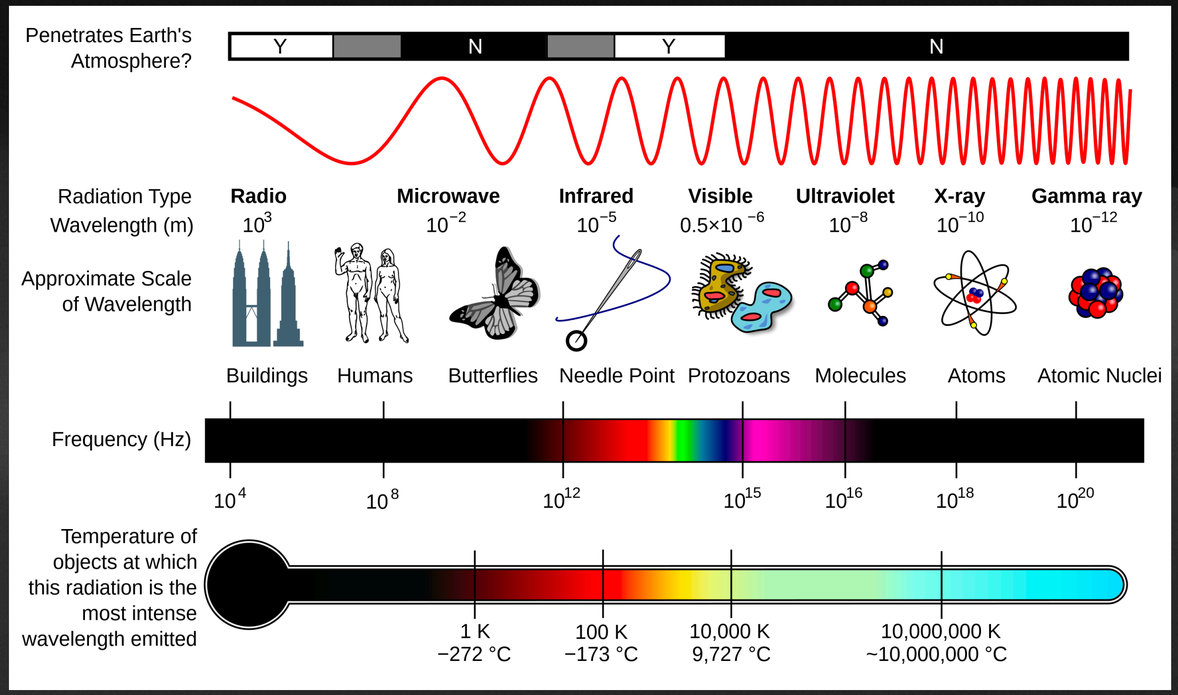
…men bilder kan også dannes gjennom andre medium.

- **Ultralyd og Sonar** (lydbølger)
- **Radar, MRI, Røntgen, PET** (utvidet EM)
- **Massespektrometri** (ionladning)



Lys og EM Stråling
Effekt og Spektrum

- Elektromagnetisk (EM) stråling omfavner alt fra lavfrekvente radiosignaler til lys, røntgen, og radioaktiv stråling (gamma).
- Inndeles i spektrum av bølgelengde (λ) / frekvens (ν) med forholdet: $\lambda = \frac{c}{\nu}$
- Stråling med lavere frekvens har lavere energinivå.

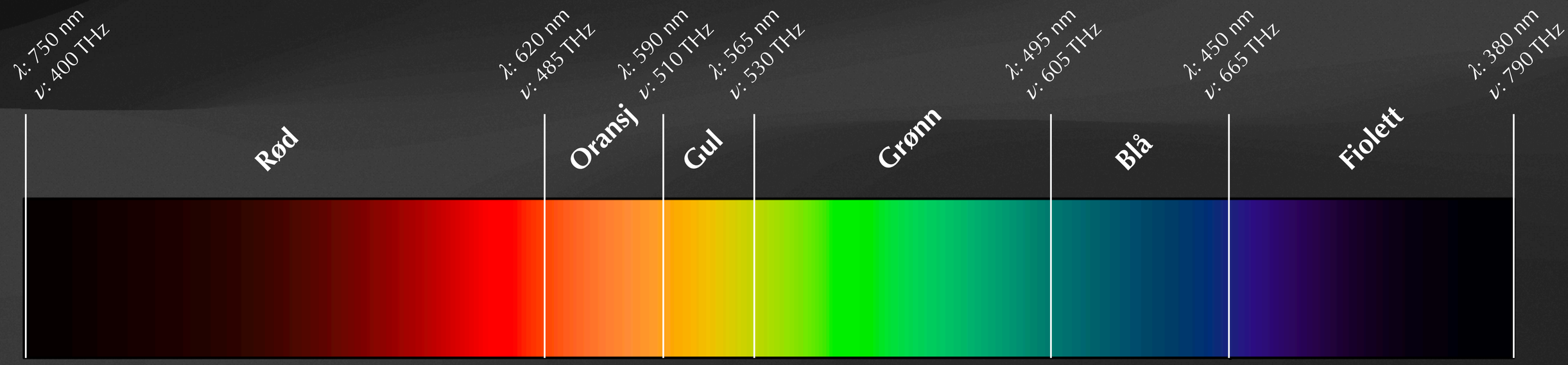


c : Lysets hastighet ~ 300,000,000 m/s

Synlig Lys
Spektrum og Farger

←Infrarød Stråling

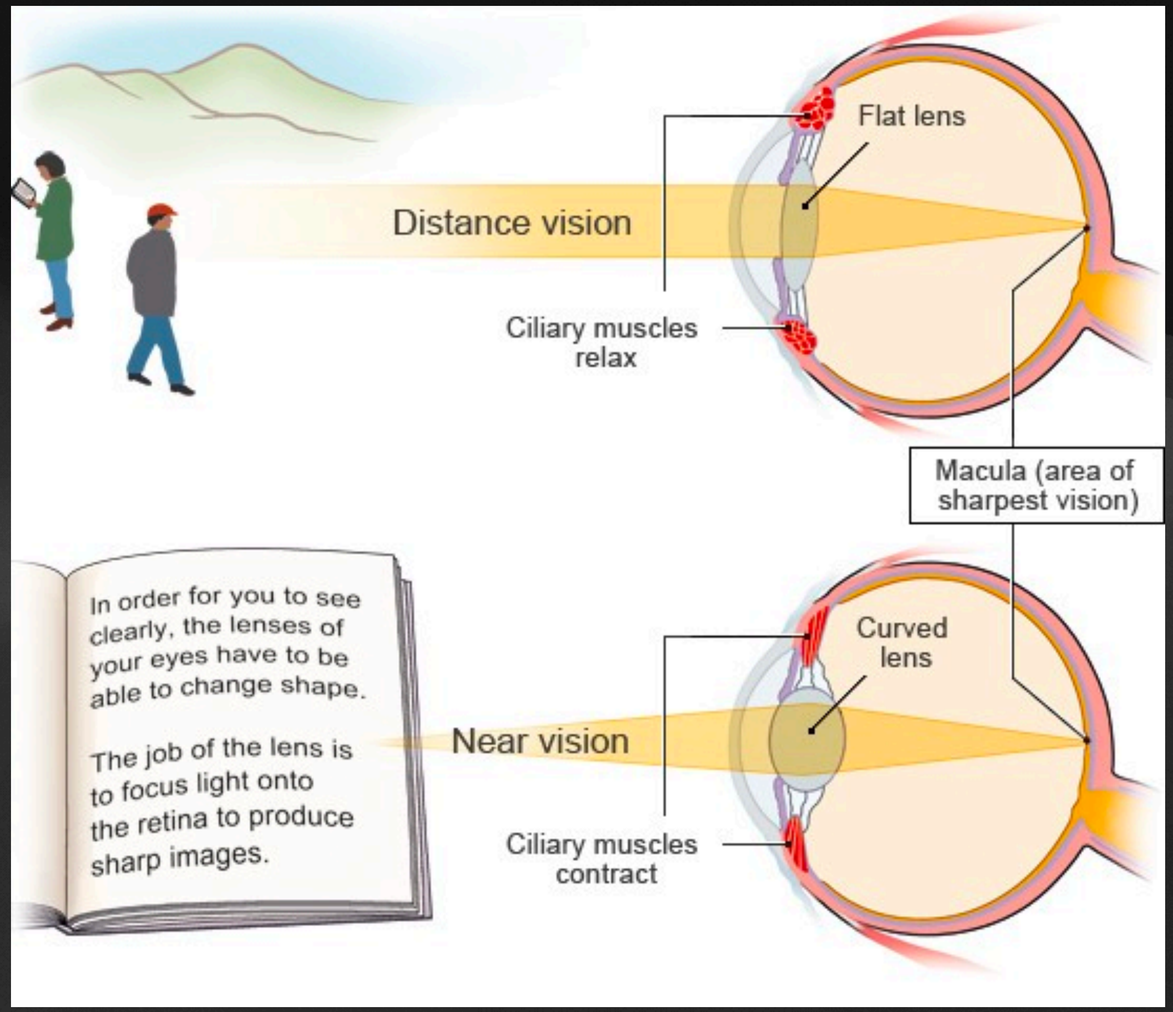
Ultrafiolett Stråling→



Avbildning
Hvordan fange et bilde

Bilder fanges på en overflate eller sensor som er sensitiv for lys. Avbildning skjer ved å kontrollere mengden med lys og dets fokusområde på overflaten.

- **Øyet**:
  - **Deformerbar** linse for fokus på macula.
  - **Iris** kontrollerer lysmengde (blender).
- **Fotokamera**:
  - Mekanisk optisk system med linser av glass.
  - Lukkemekanisme for å slippe igjennom lys.



Avbildning
Menneskeøyet

Øyet er et unikt optisk system:

- Linsen er deformerbar og kan endre brennvidde (fokallengde).
- Adaptiv deteksjon med variabel geometrisk oppløsning.
- Separate systemer for lavlys og høylys. Blenderåpning (pupillen) justeres av iris:
  - 2mm (sterkt lys)
  - 8mm (svakt lys)

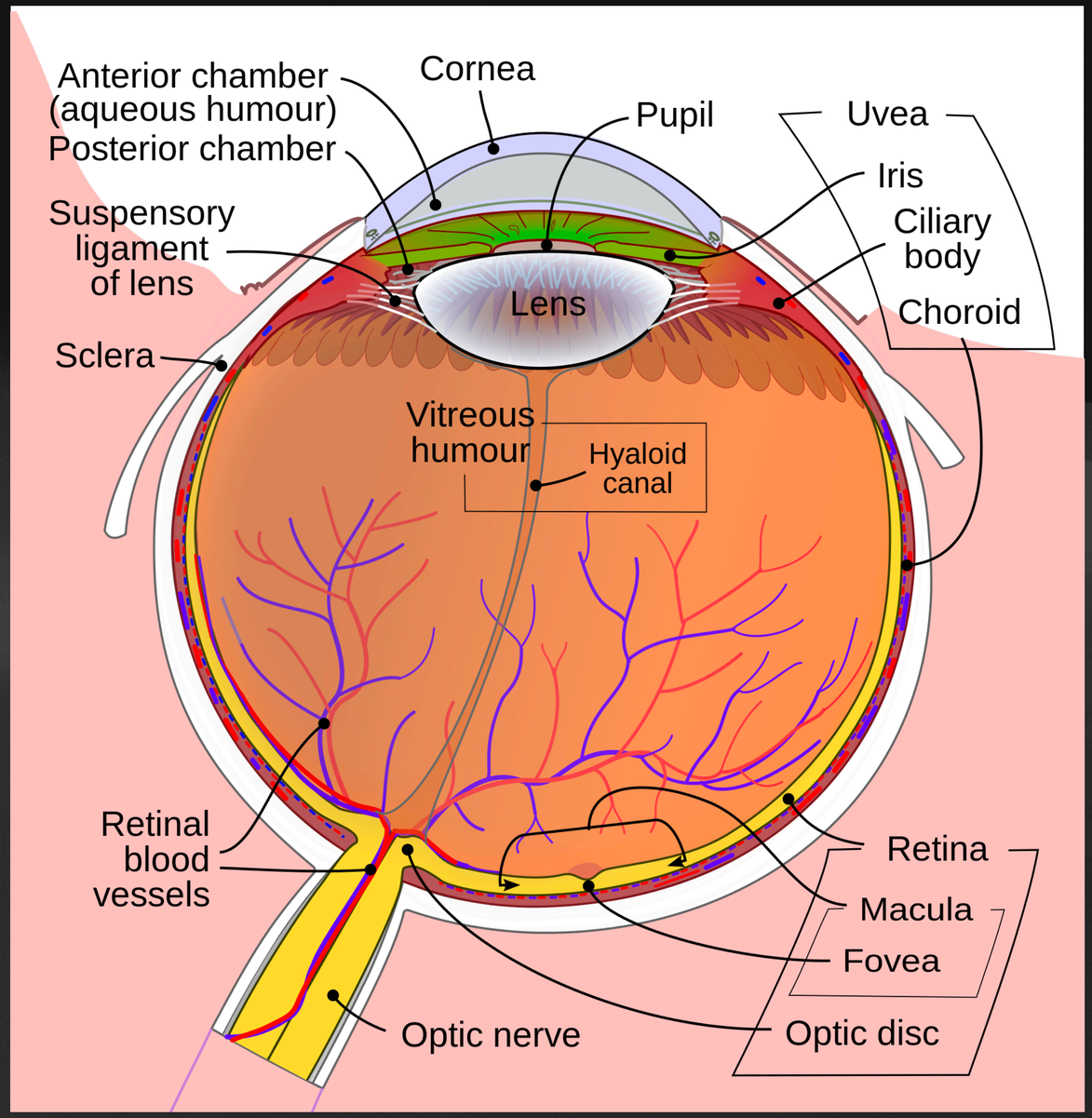


Bilde: Maude Guatteri/EyeEm/Getty Images

Avbildning
Netthinnen

Netthinnen (retina): lysfølsomt lag som dekker 65% av den indre flaten bak øyet, ca. 130 millioner detektorer:

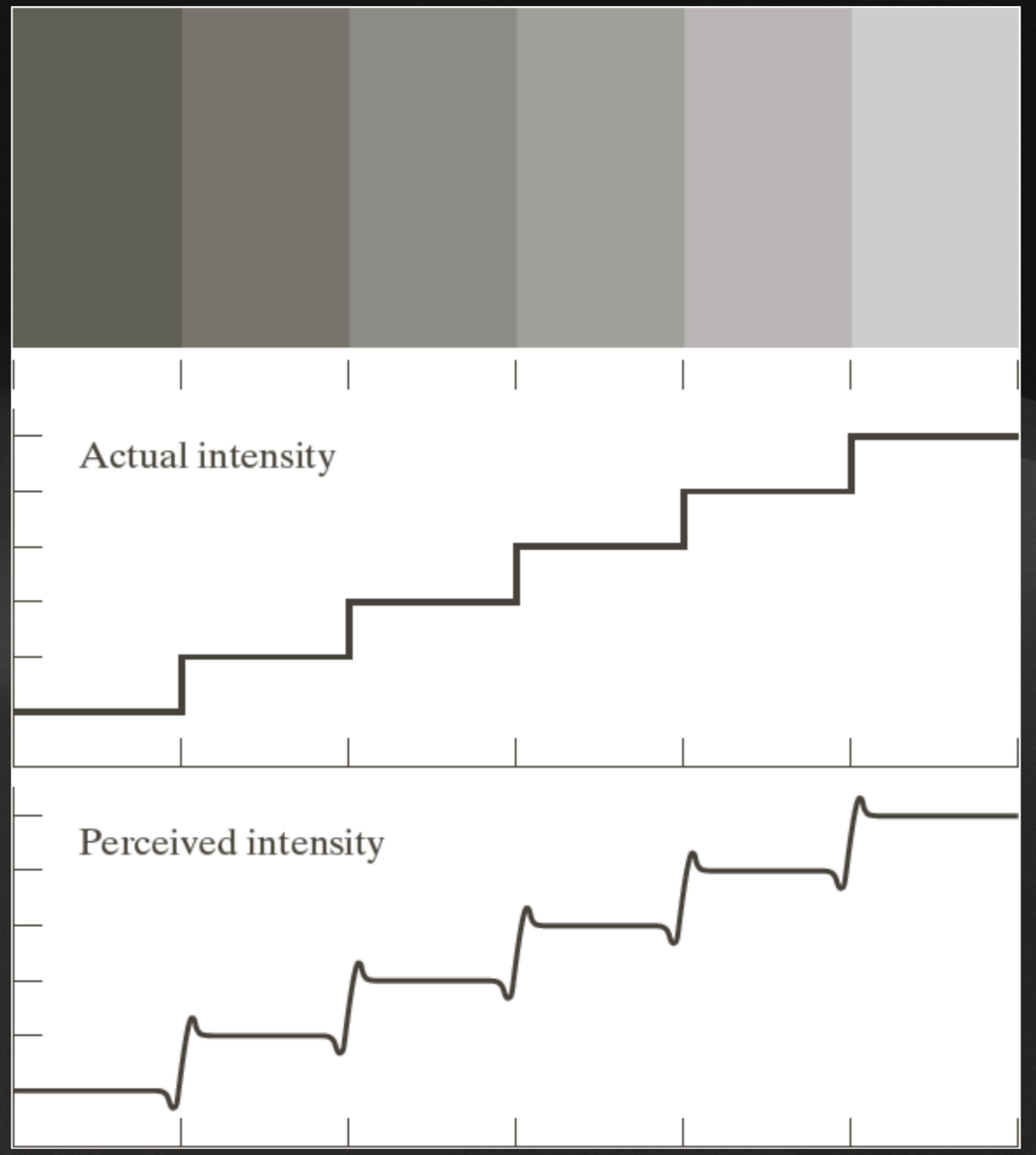
- 120 mill. **staver** (rods)
  - Lavere oppløsning, kun gråtone.
  - Tilpasset lav lysfølsomhet.
- 7 mill. **tapper** (cones)
  - Konsentrert i macula (fovea) med høy oppløsning
  - Tilpasset høy lysintensitet med dedikerte RGB detektorer.



Bilde fra: www.wikipedia.org

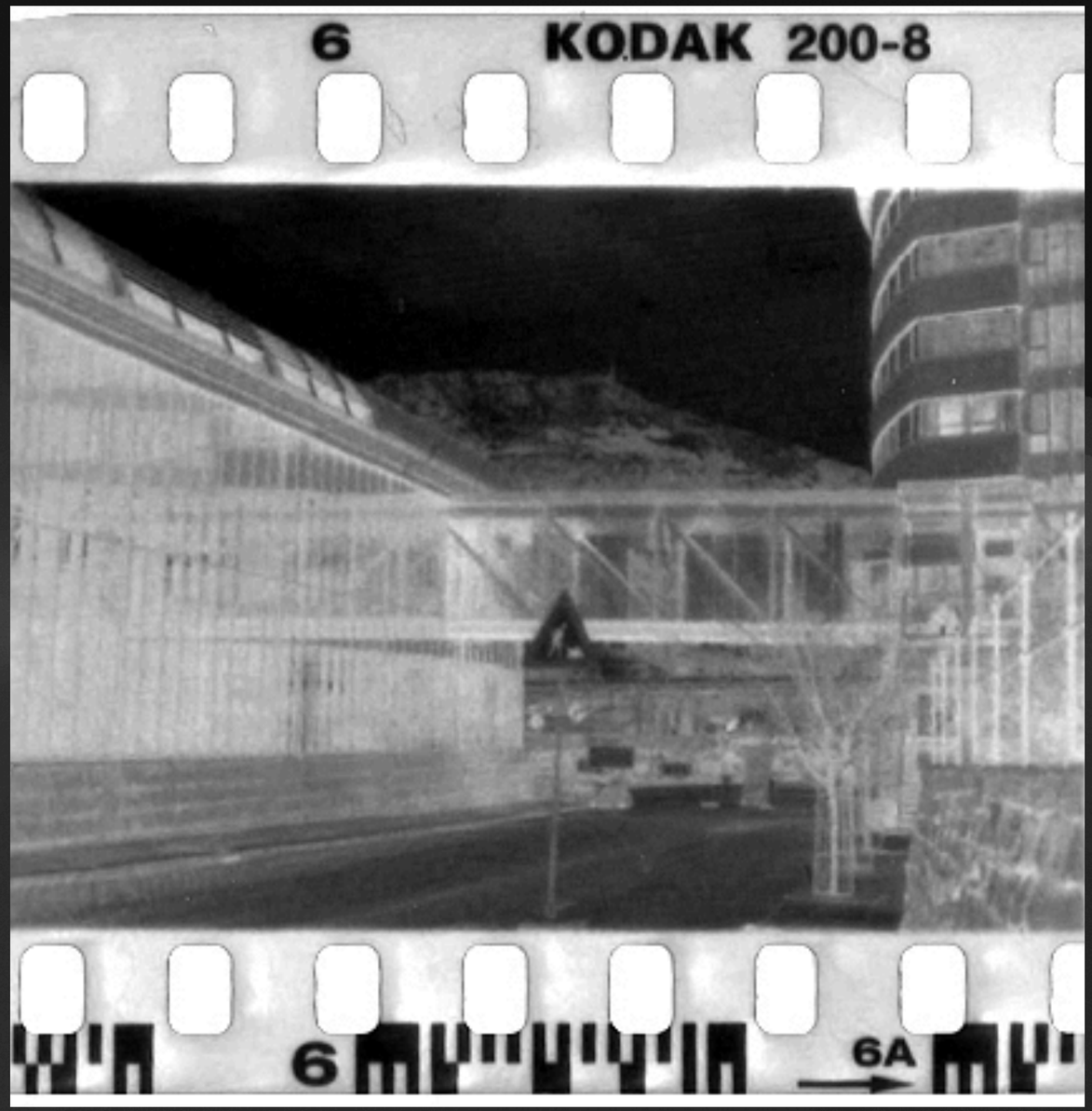
Avbildning
Gråtoner og Mach Bånd

- Når øyet skifter fokus til annet bakgrunnsintensitetsnivå, tilpasser øyet seg dette og ser fint lokale intensitets-forskjeller.
- Vi oppfatter kun en viss andel lysnivåer samtidig, den minste gråtone-forskjellen vi merker er 2%. Vi oppfatter ca. 50 gråtoner.
- **Mach bånd**: Nevrale prosesser i netthinnen forsterker kontrast i overgang mellom uniforme regioner.



Kameratеknologi
Analoge Signaler og Film

- Før digitale sensorer ble normen i kameratеknologi, ble lys fanget med ***kjemiske*** *prosesser*.
- ***Fotografisk film*** er tynn celluloid dekket med en emulsjon av ***sølvsalter***, som er følsomt for lys.
- Avbildningen er ***negativ***, fremkalt ved videre kjemisk prosessering.
- Den endelige avbildningen blir gjort ved en endelig fremkalling; negativet projiseres på fotosensitivt papir.



Bilde fra: www.wikipedia.org

Kameratеknologi

Bartlane vs. Belin

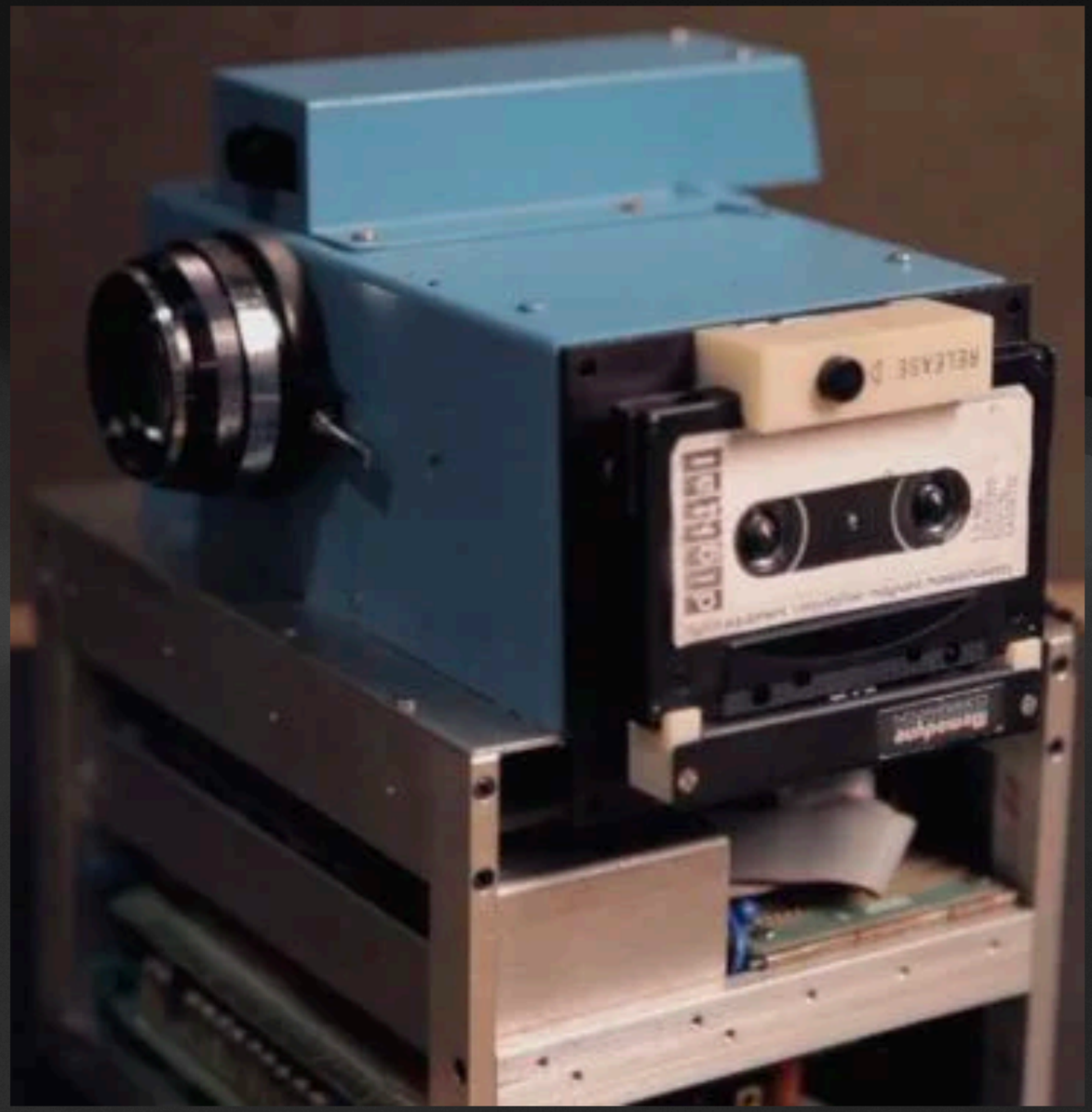
- De tidligste digitale representasjonene for bilder var laget for å sende bilder via telegrafi, og ble brukt hovedsaklig for illustrasjoner og fotografier brukt av nyhetsbyrå og aviser.
- **1929**: 5 gråtoner via Bartlane's hullkort.
- **1936**: Associated Press (AP) sender det første telefotografiet med Belin's analoge lydbølger. Tjenesten fortsetter frem til 2004.
- Bartlane's system var banebrytende innen digital representasjon, men var tregere og hadde oftere feil en Belin's telefoto.



Bilde fra: www.wikipedia.org

Kamerateknologi
Fra Film til Sensorer

- I 1975 utviklet Steven Sasson (hvis mor var norsk) verdens første heldigitale kamera under Kodak.
- Det første kameraet benyttet en 100×100 piksel CCD sensor og fanget bilder i svart hvitt.
- Bildet ble først fanget i digitalt minne, og så overført til kassettbånd som holdt 30 bilder.
- Kodak var verdens største leverandør av fotografisk film og en av verdens største selskap.
- 2003 var det første året digitalkameraer solgte flere eksemplarer enn filmkameraet. Kodak gikk konkurs i 2012.

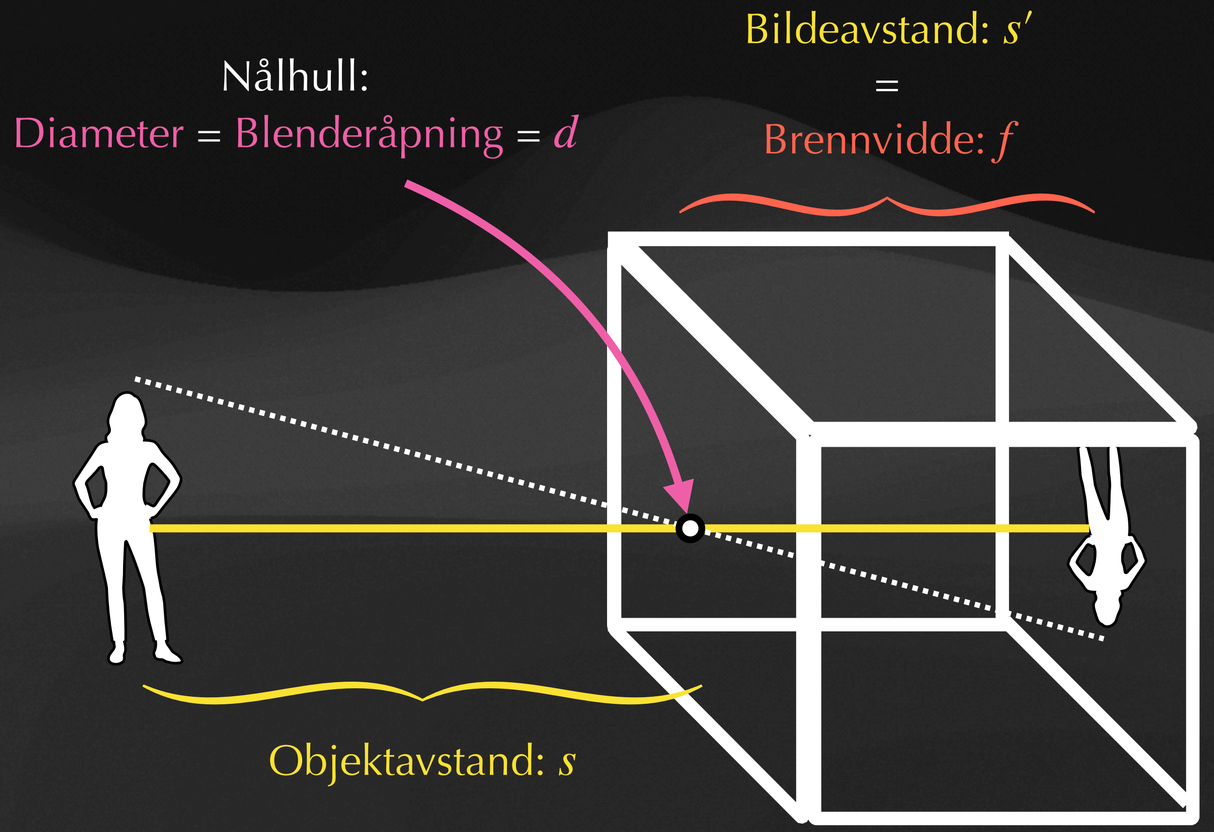


Bilde fra: www.wikipedia.org


Intermission
You look like you could use a snack!
Mr. Pops
We'll meet you in the lobby.

Optikk
Camera Obscura

- **Fra Latin**: *Mørkt kammer*.
- Først beskrevet av den Arabiske fysikeren **Ibn al-Haytham** ca. 1000 AD.
- Enkleste optiske system - ingen linse, kun en minimal blenderåpning d.
- Bildeavstand = Brennvidde
  $s' = f$
- **Optimal blenderåpning uten linse**:
  $d = \sqrt{2 f \lambda}$ med bølgelendge λ.
- Lite lys krever lang eksponering.

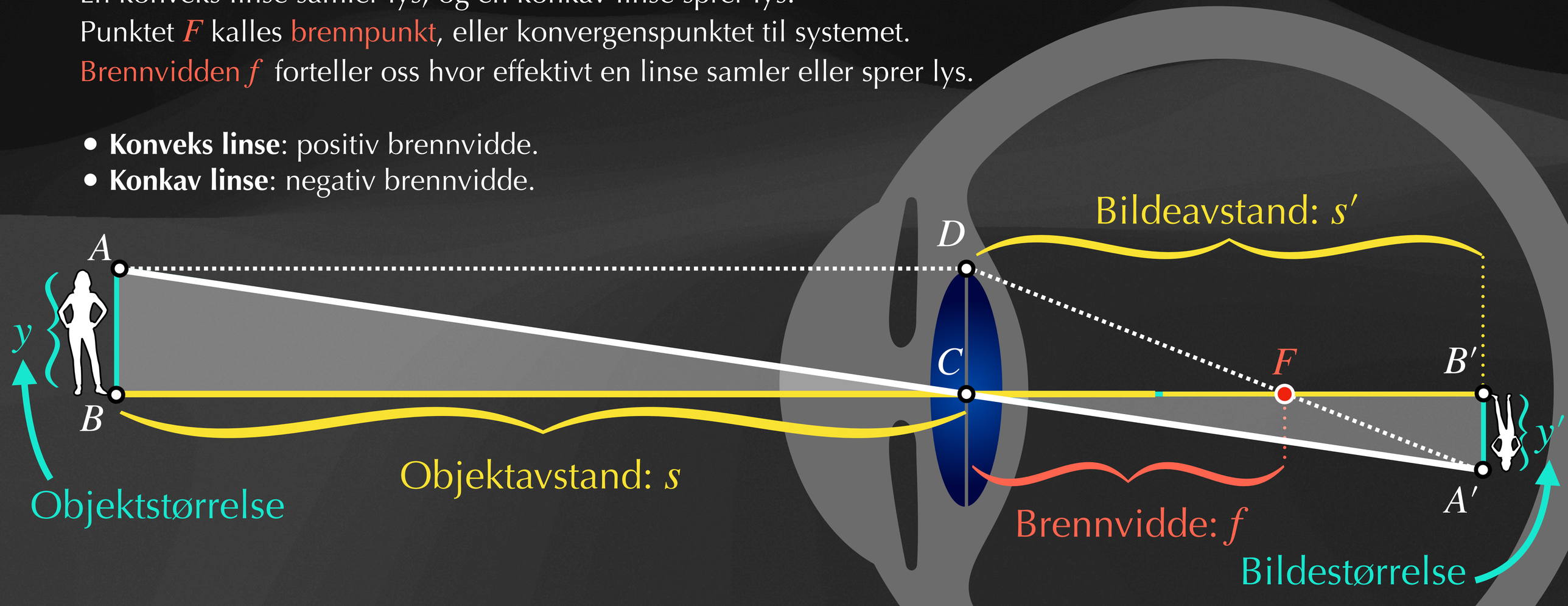


Optikk

Hvordan linser fanger lys

Figuren viser hvordan et objekt avbildes med en konveks linse i et optisk system.
En konveks linse samler lys, og en konkav linse sprer lys.
Punktet F kalles brennpunkt, eller konvergenspunktet til systemet.
Brennvidden f forteller oss hvor effektivt en linse samler eller sprer lys.

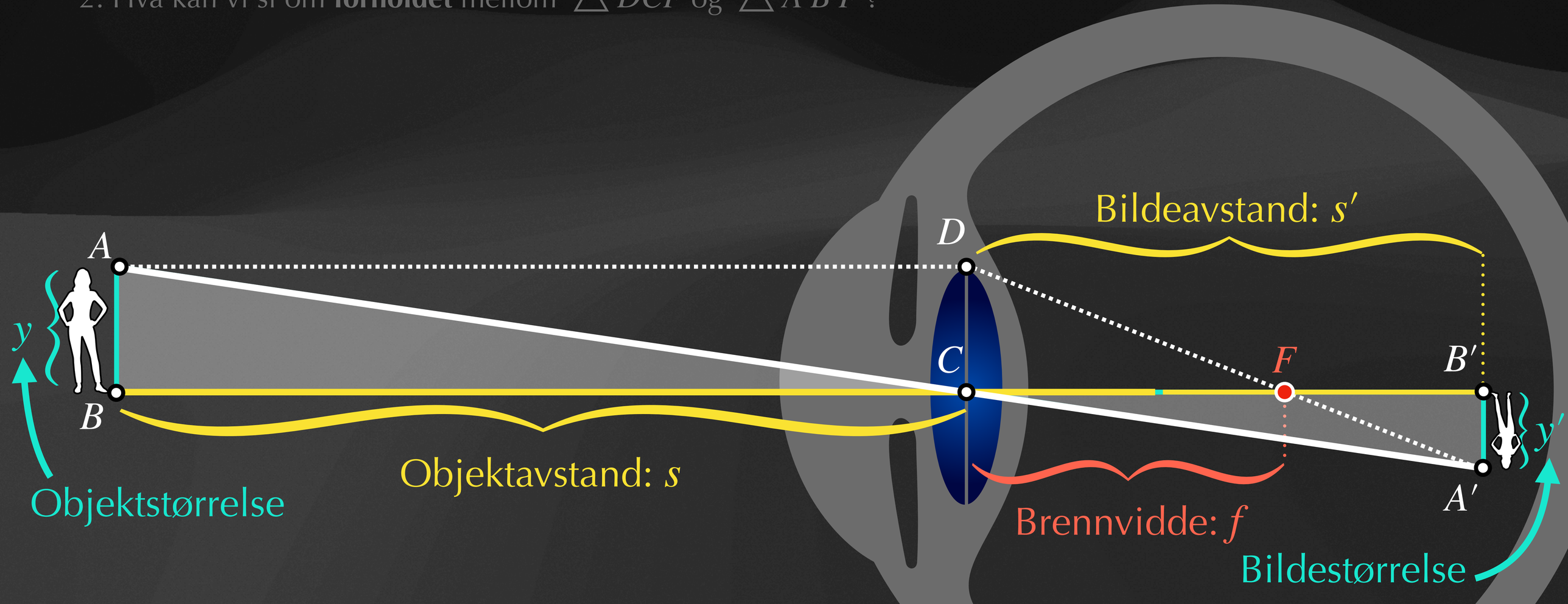
- **Konveks linse**: positiv brennvidde.
- **Konkav linse**: negativ brennvidde.



Optikk

Objekt-bilde relasjonen

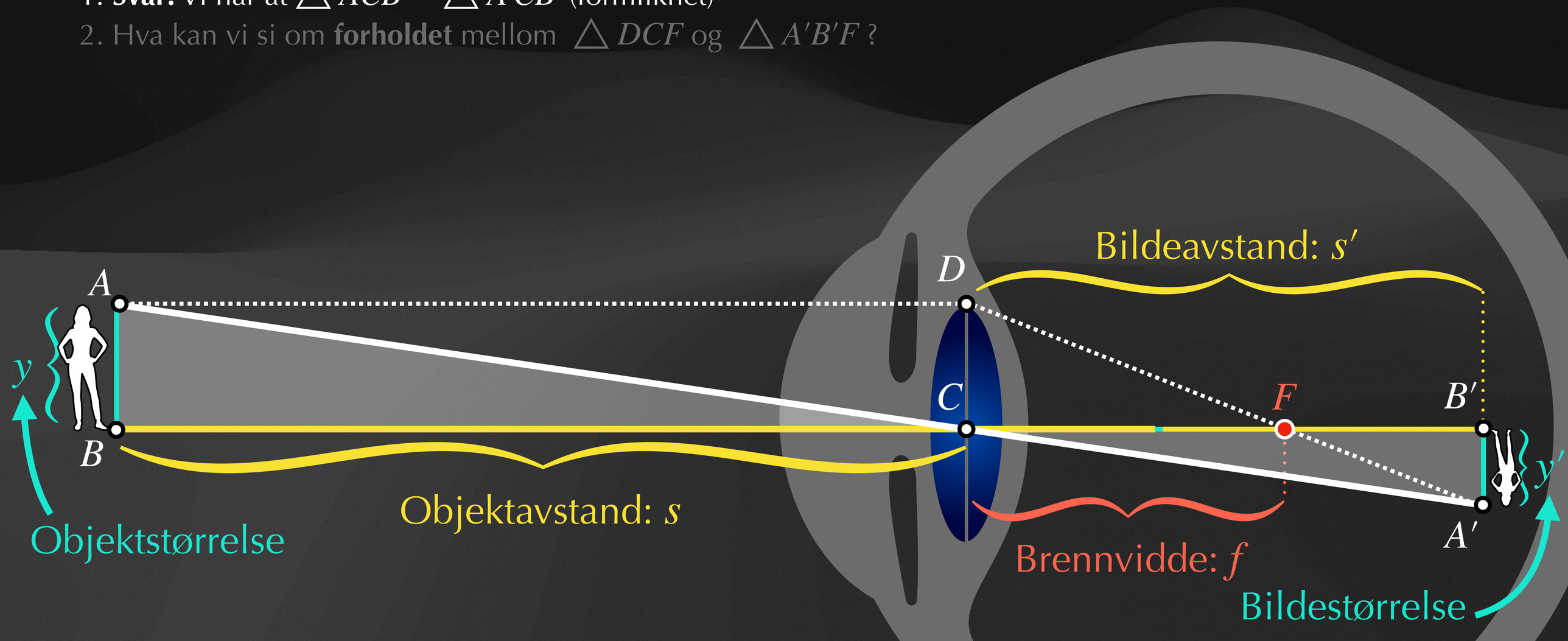
1. Hva kan vi si om **forholdet** mellom $\triangle ABC$ og $\triangle A'B'C$?
2. Hva kan vi si om **forholdet** mellom $\triangle DCF$ og $\triangle A'B'F$?



Optikk

Objekt-bilde relasjonen

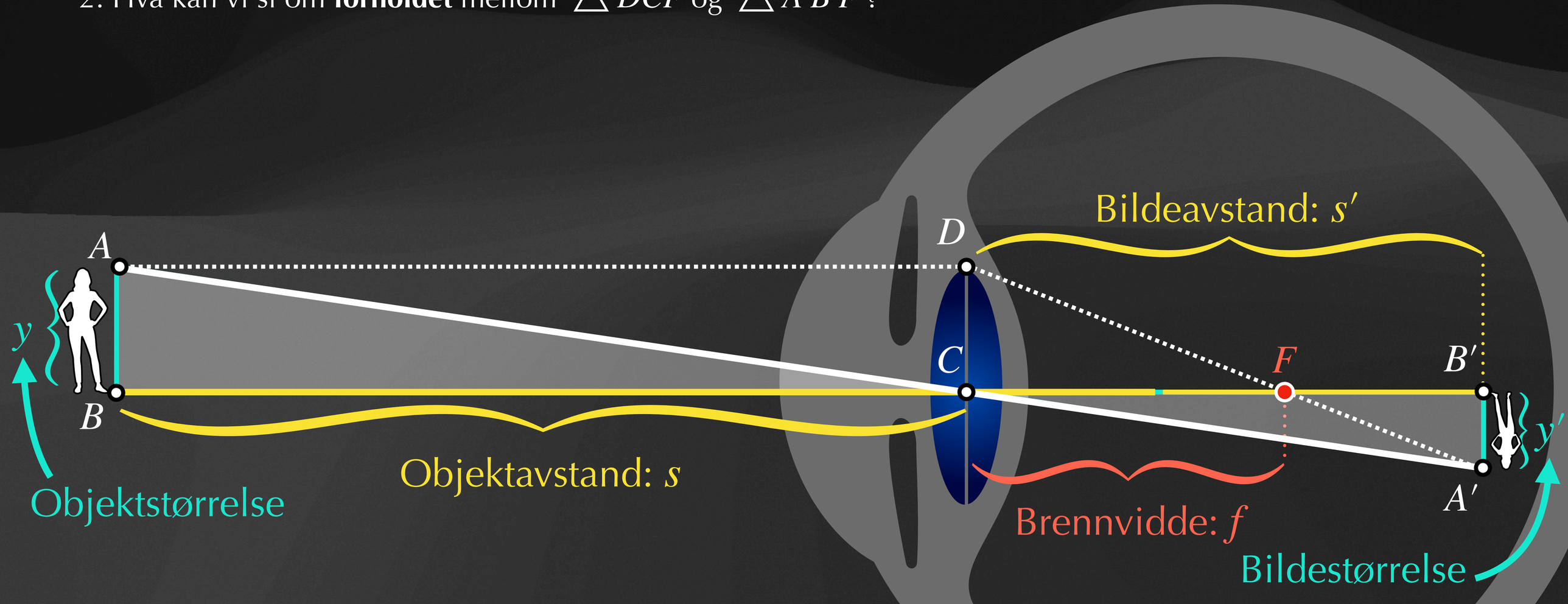
1. **Svar:** Vi har at $\triangle ACB \sim \triangle A'CB'$ (formlikhet)
2. Hva kan vi si om **forholdet** mellom $\triangle DCF$ og $\triangle A'B'F$?



Optikk

Objekt-bilde relasjonen

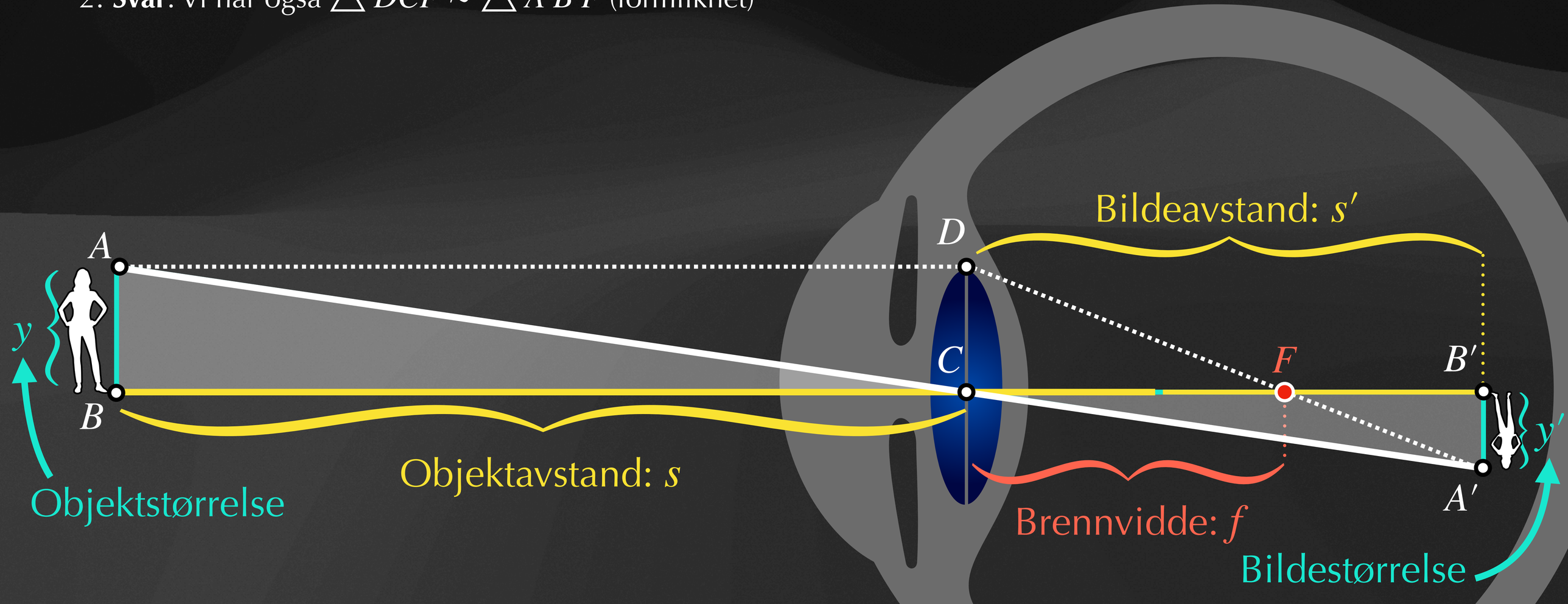
1. **Svar:** Vi har at $\triangle ACB \sim \triangle A'CB'$ (formlikhet)
2. Hva kan vi si om **forholdet** mellom $\triangle DCF$ og $\triangle A'B'F$?



Optikk

Objekt-bilde relasjonen

1. **Svar:** Vi har at $\triangle ACB \sim \triangle A'CB'$ (formlikhet)
2. **Svar:** Vi har også $\triangle DCF \sim \triangle A'B'F$ (formlikhet)



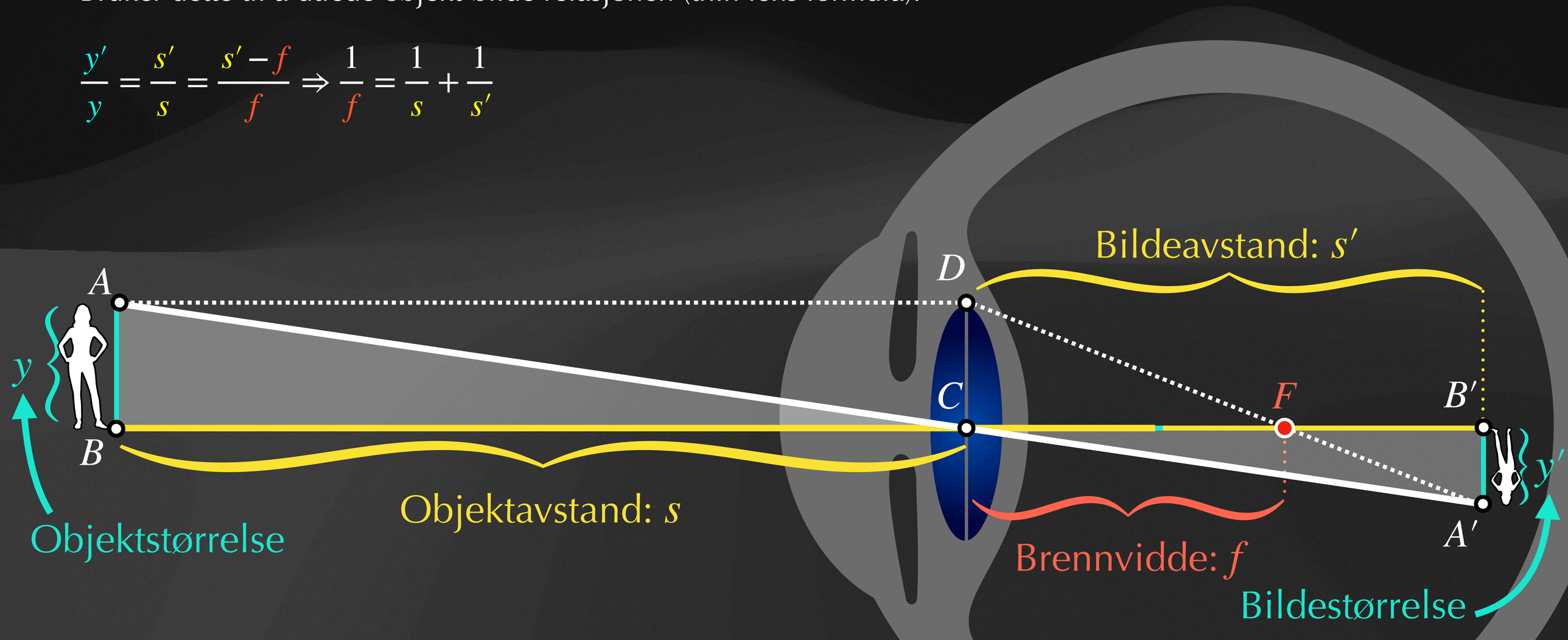
Optikk

Objekt-bilde relasjonen

$$\triangle ACB \sim \triangle A'CB'$$
$$\triangle DCF \sim \triangle A'B'F$$

Bruker dette til å utlede objekt-bilde relasjonen (thin lens formula):

$$\frac{y'}{y} = \frac{s'}{s} = \frac{s'-f}{f} \Rightarrow \frac{1}{f} = \frac{1}{s} + \frac{1}{s'}$$



Optikk

$$\frac{y'}{y} = \frac{s'}{s} = \frac{s'-f}{f} \Rightarrow \frac{1}{f} = \frac{1}{s} + \frac{1}{s'}$$

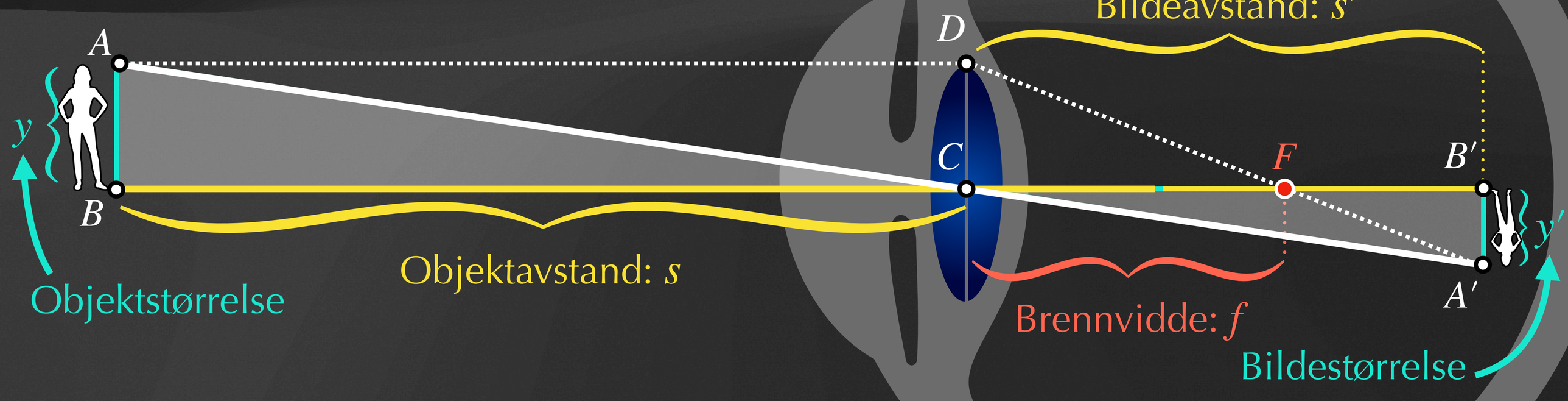
$$\triangle ACB \sim \triangle A'CB'$$
$$\triangle DCF \sim \triangle A'B'F$$

Utregning av bildestørrelse y' gitt f, y, s

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

$$\frac{1}{f} = \frac{1}{s} + \frac{1}{s'} \Rightarrow s' = \frac{sf}{s-f}$$

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



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

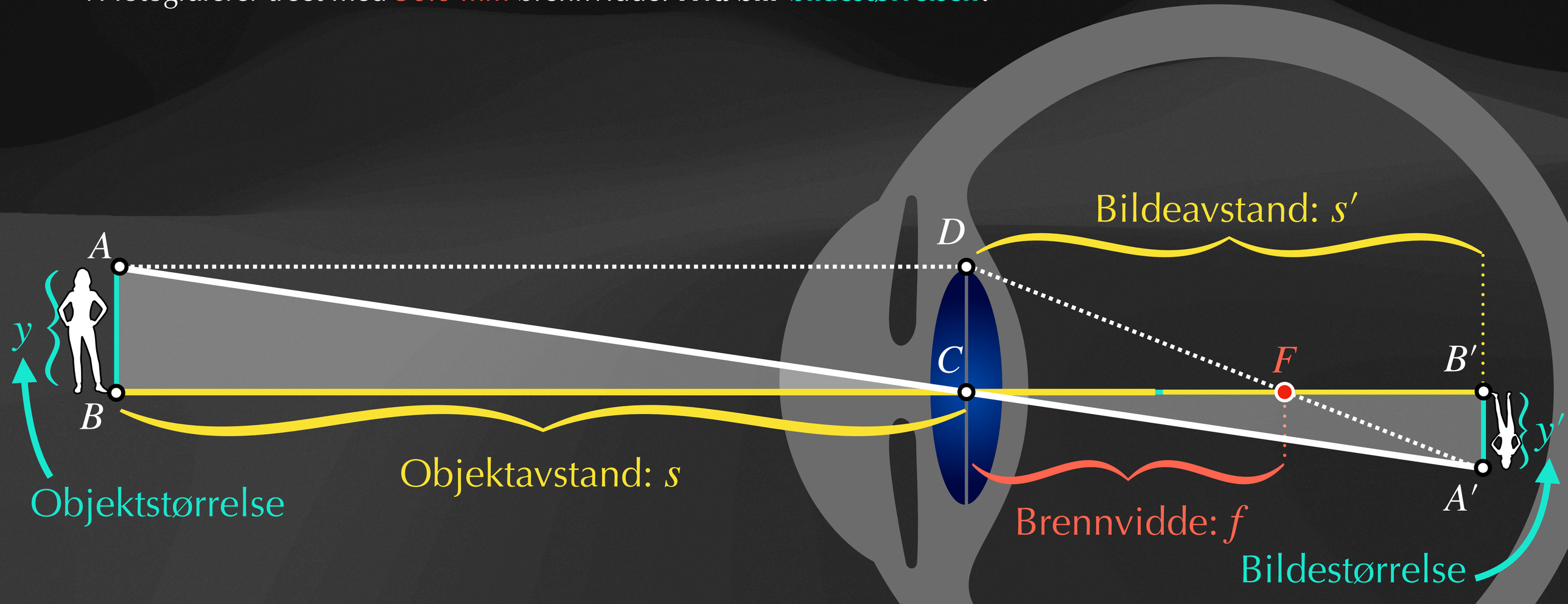
$$\frac{y'}{y} = \frac{s'}{s} = \frac{s' - f}{f} \Rightarrow \frac{1}{f} = \frac{1}{s} + \frac{1}{s'}$$

$$\triangle ACB \sim \triangle A'CB'$$
$$\triangle DCF \sim \triangle A'B'F$$

Optikk

Eksempel: utregning av bildestørrelse

I figuren under, anta at personen er 168.1 cm høy og objektavstanden er 2.452 m.
Vi fotograferer treet med 50.0 mm brennvidde. **Hva blir bildestørrelsen?**



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

$$\frac{y'}{y} = \frac{s'}{s} = \frac{s'-f}{f} \Rightarrow \frac{1}{f} = \frac{1}{s} + \frac{1}{s'}$$

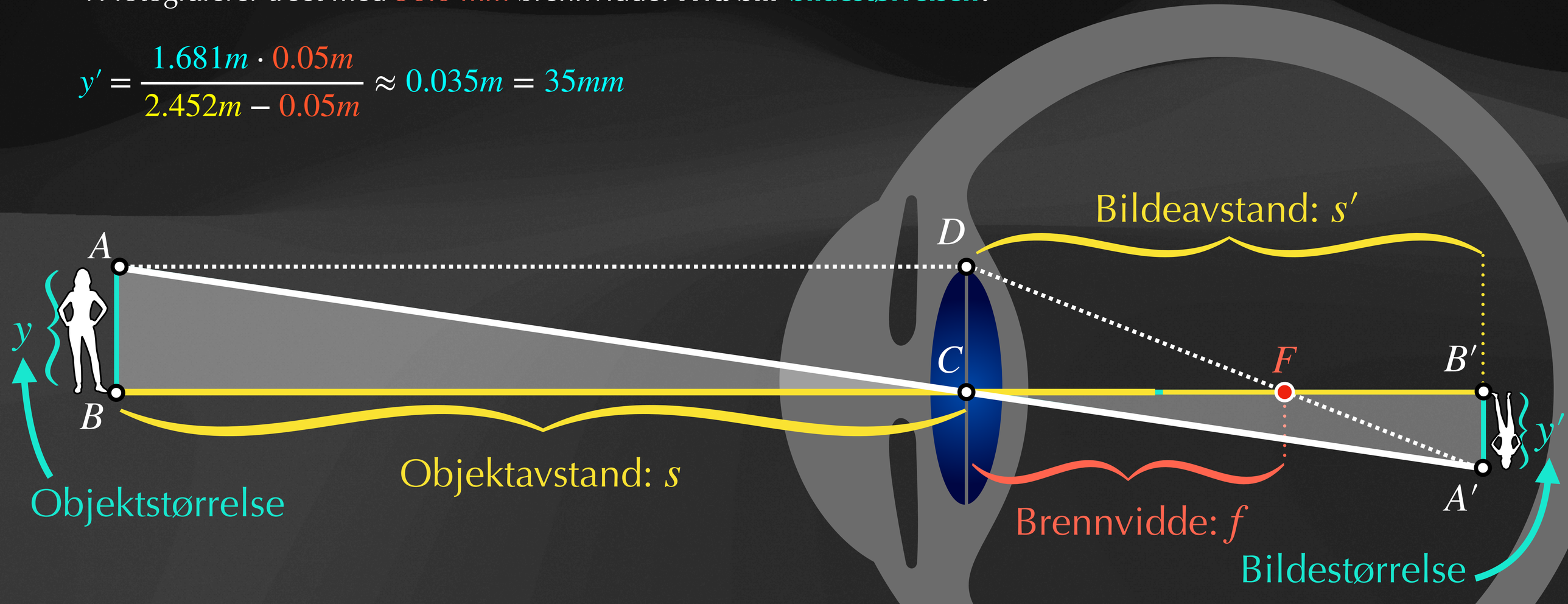
$$\triangle ACB \sim \triangle A'CB'$$
$$\triangle DCF \sim \triangle A'B'F$$

Optikk

Eksempel: utregning av bildestørrelse

I figuren under, anta at personen er 168.1 cm høy og objektavstanden er 2.452 m.
Vi fotograferer treet med 50.0 mm brennvidde. **Hva blir bildestørrelsen?**

$$y' = \frac{1.681m \cdot 0.05m}{2.452m - 0.05m} \approx 0.035m = 35mm$$

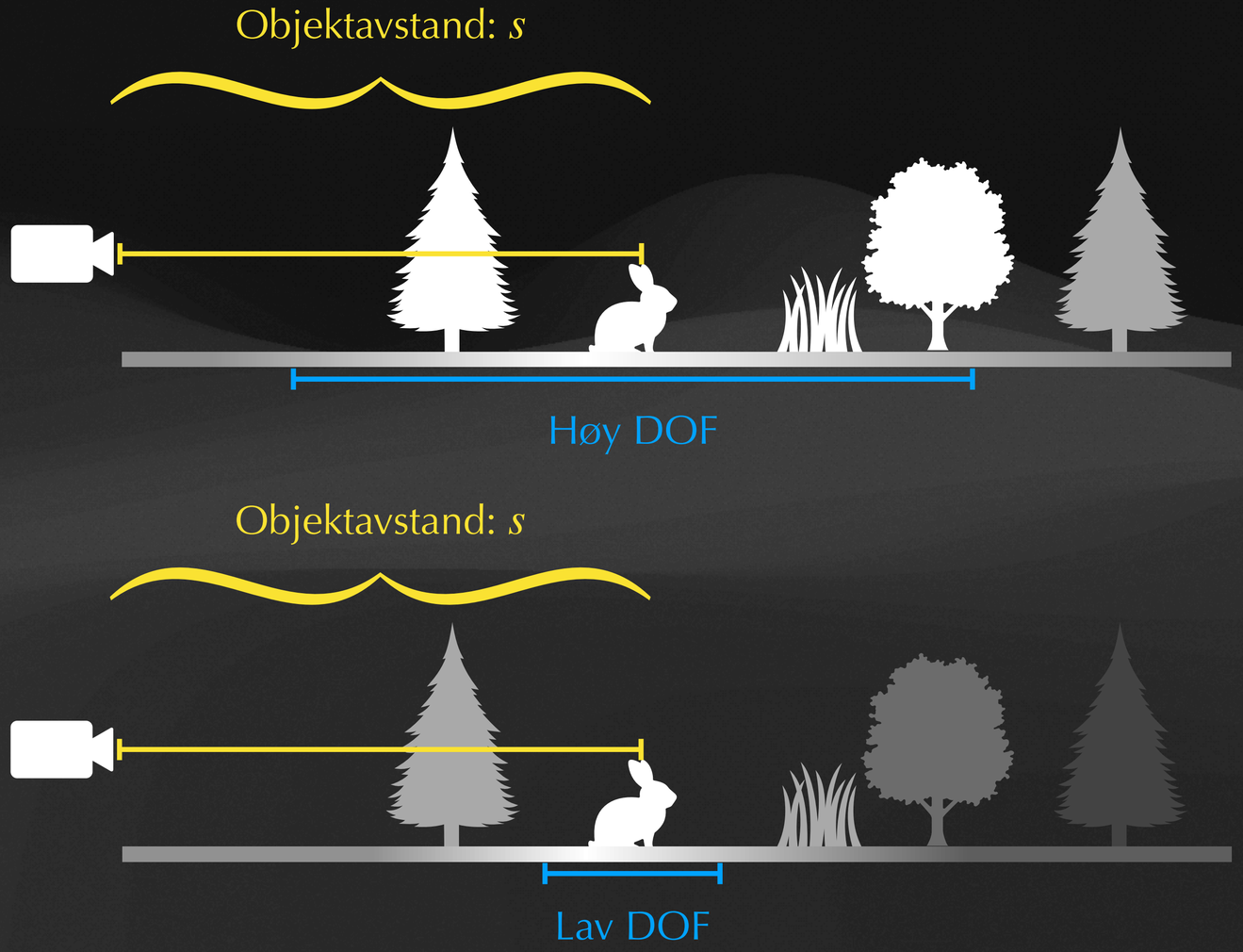


Hvilket bilde ser “finest” ut?

Bilde fra: www.wikipedia.org

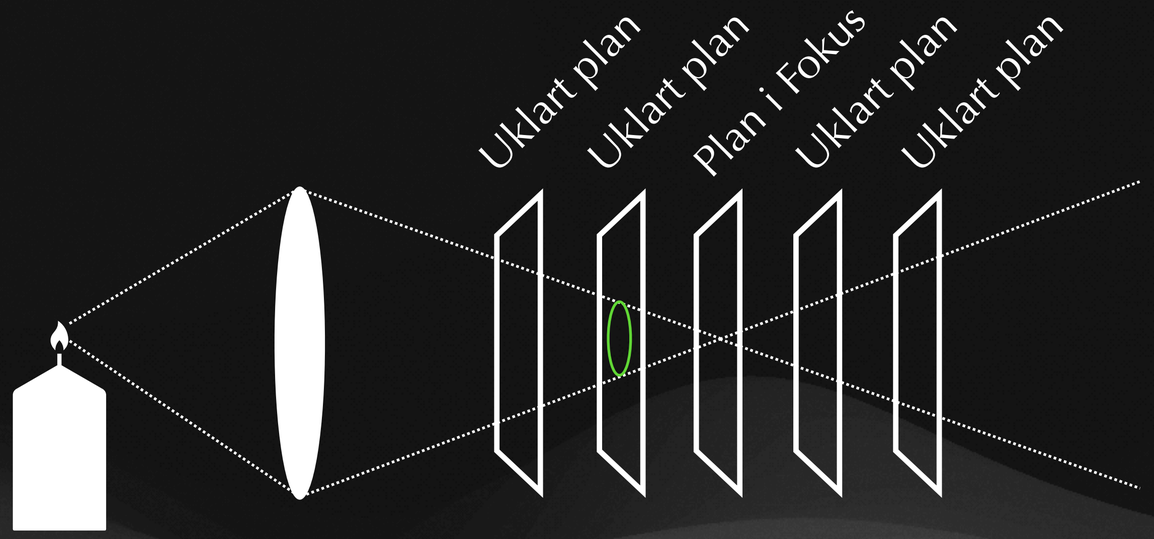
Dybdeskarphet
Depth of Field (DOF)

- Dybdeskarphet er avstanden foran og bak det fokuserte objektet som syntes å være i fokus.
- DOF $\rightarrow \infty$ gir skarpt bilde av både forgrunn og bakgrunn.
- DOF $\rightarrow 0$ gir bløtere fokus på forgrunn og bakgrunn, dermed økt fokus på objektet.

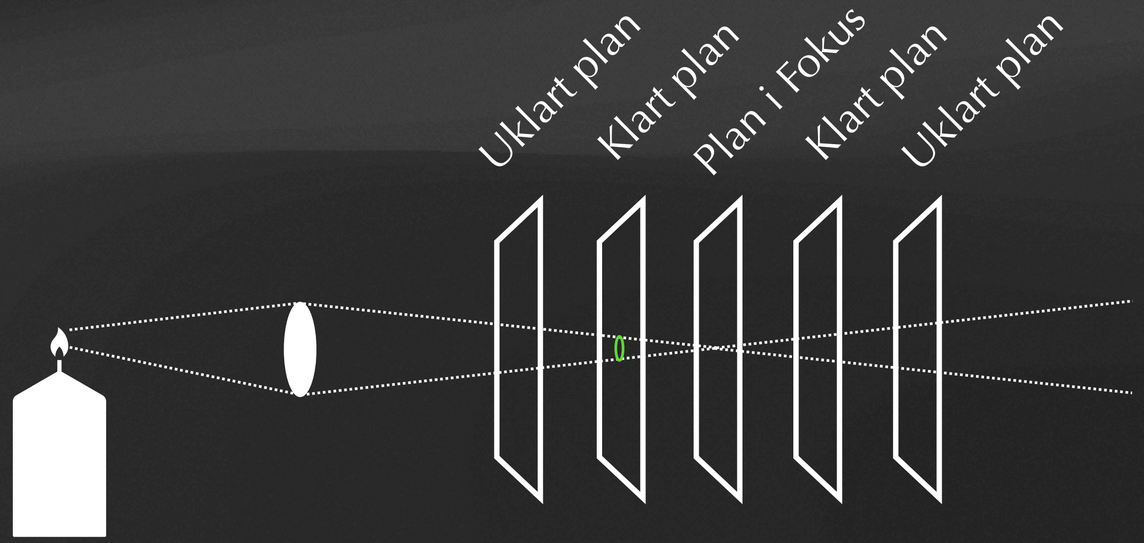


Dybdeskarphet
F-nummer og Blenderåpning

- **Diameter på blenderåpning**: d
- **Brennvidde / fokalstørrelse**: f
- **F-nummer**: f / d
- $d \to \infty \Rightarrow f/d \to 0$:
  - Mer lys, kortere eksponering.
  - Lav dybdeskarphet.
- $d \to 0 \Rightarrow f/d \to \infty$:
  - Mindre lys, lengre eksponering.
  - Høyere dybdeskarphet.

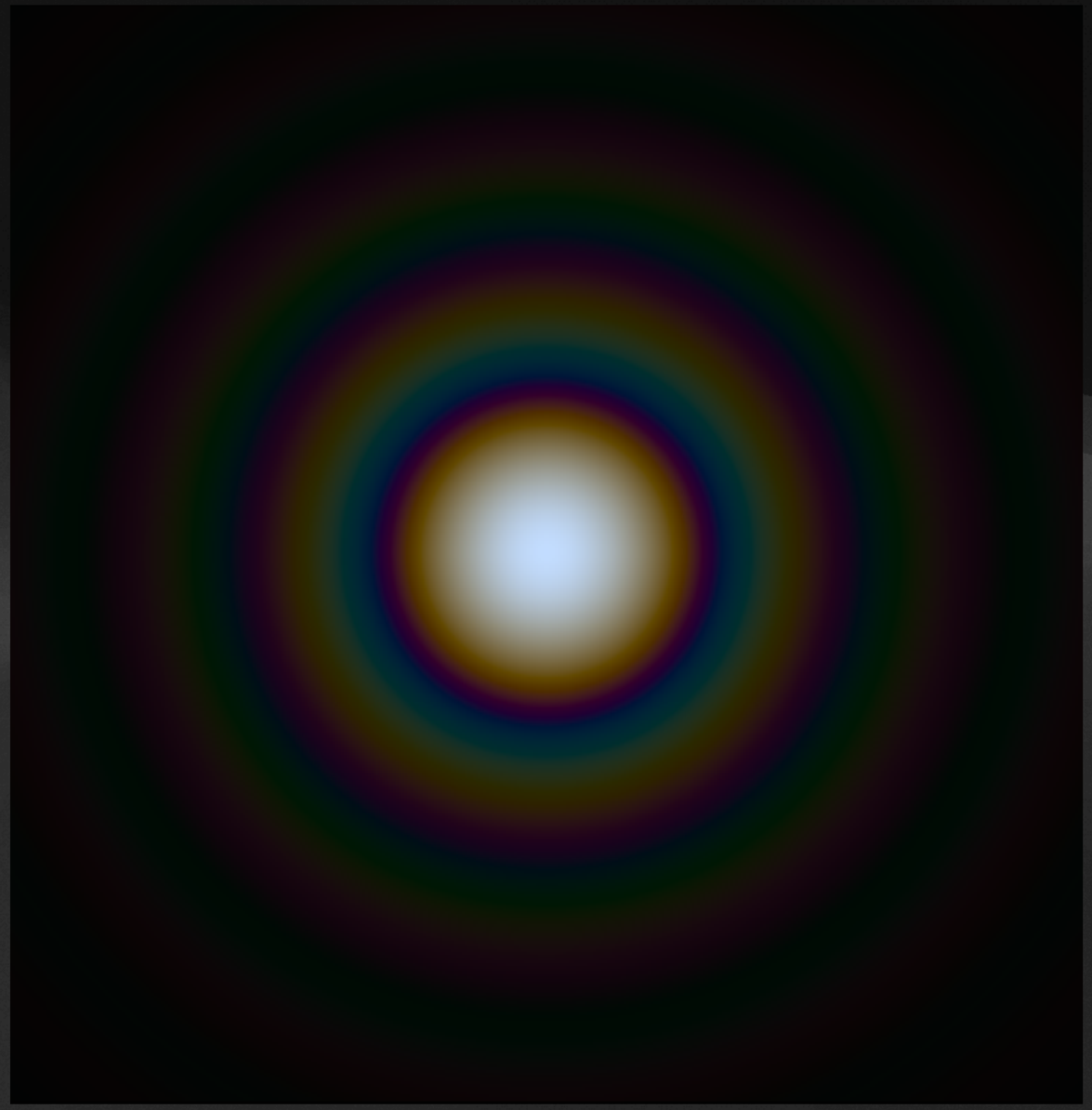


Uskarphetssirkel (Circle of Confusion): C



Vinkeloppløsning
Punktspredningsfunksjon

- Skildringen av lys som stråler er en forenkling av virkeligheten.
- Lys oppfører seg også som **bølger**, med interferens og diffraksjon.
- En sirkulær linse vil avbilde en punktkilde som en lys flekk med ringer rundt (**Airy disk**).
- Vi kan enkelt beregne systemets punktspredningsfunksjon (PSF) for en gitt blenderåpning.

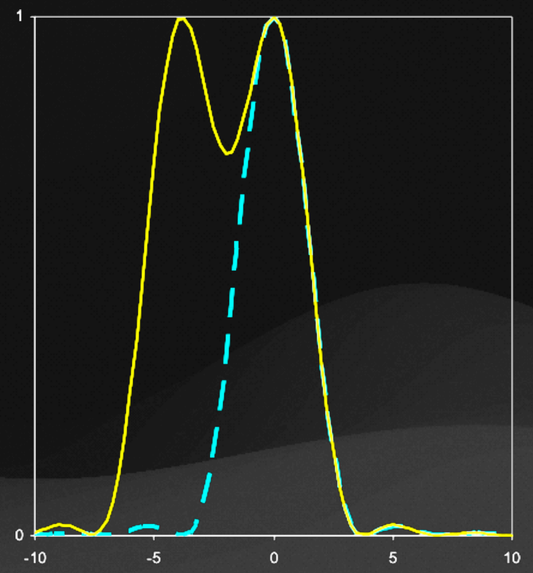


Bilde fra: www.wikipedia.org

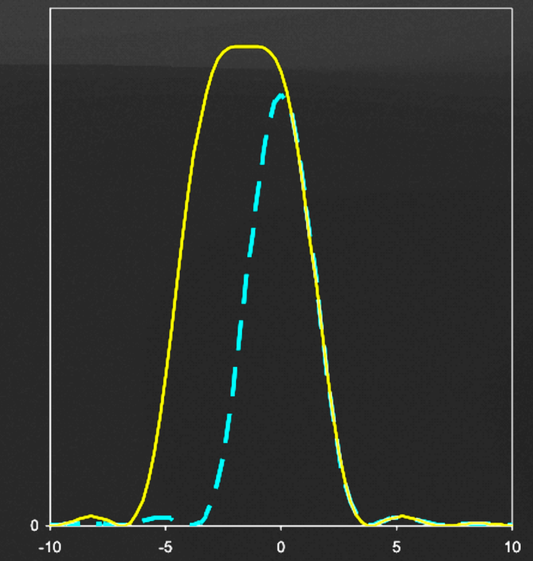
Vinkeloppløsning
Adskillelseskriterier

- Anta linsediameter D og bølgelengde λ.
- To punkter kan akkurat adskilles hvis vinkelen mellom dem θ er gitt ved $\sin\theta = 1.22\lambda/D$ (Rayleigh-kriteriet)
- Dip-en i profilen forsvinner helt ved $\sin\theta = 0.952\lambda/D$ (Sparrow-kriteriet)

Rayleigh



Sparrow



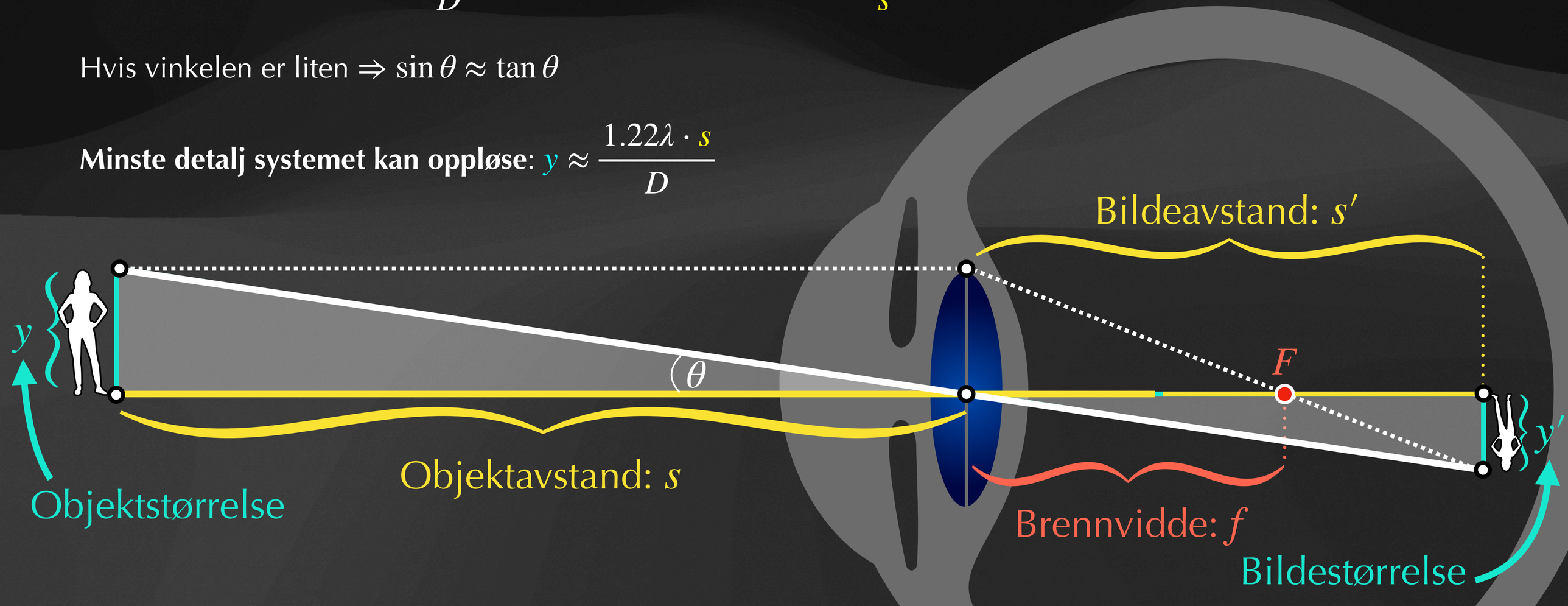
Vinkeloppøsning

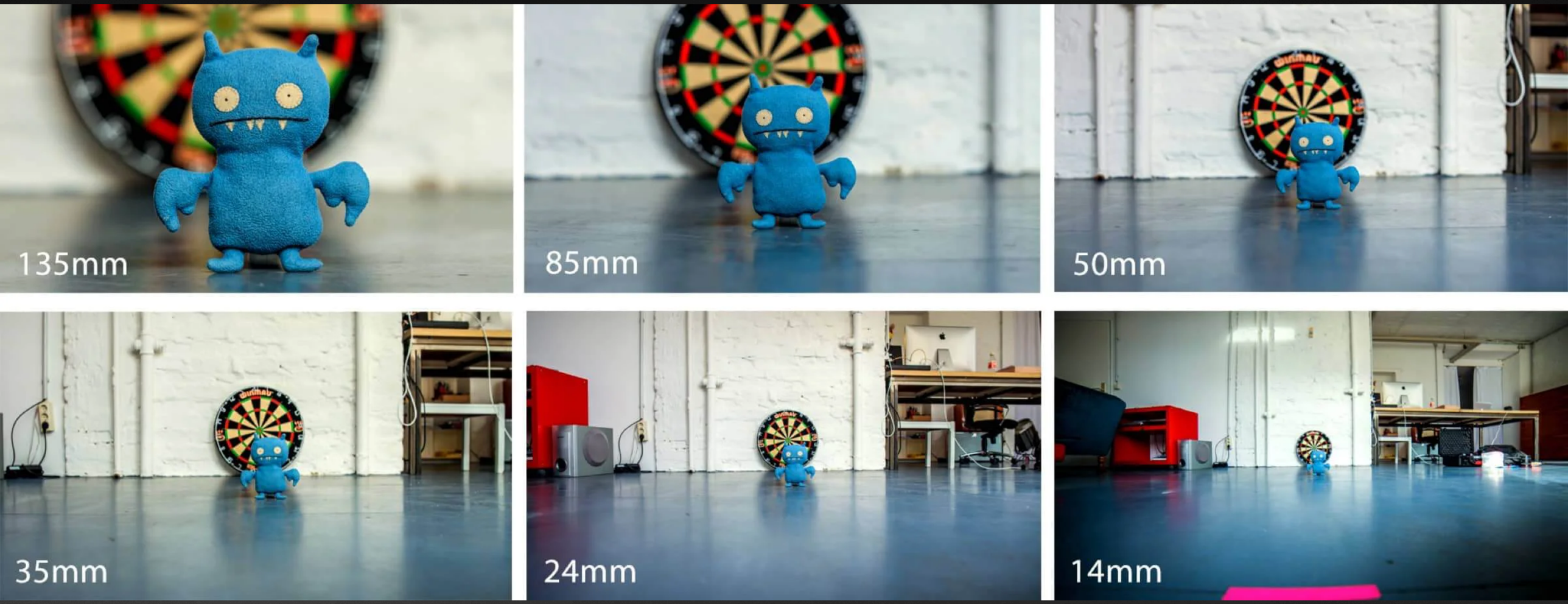
Hvor små detaljer kan en linse oppløse?

Rayleigh kriteriet: $\sin\theta = \dfrac{1.22\lambda}{D}$. Tangens til vinkelen: $\tan\theta = \dfrac{y}{s}$.

Hvis vinkelen er liten $\Rightarrow \sin\theta \approx \tan\theta$

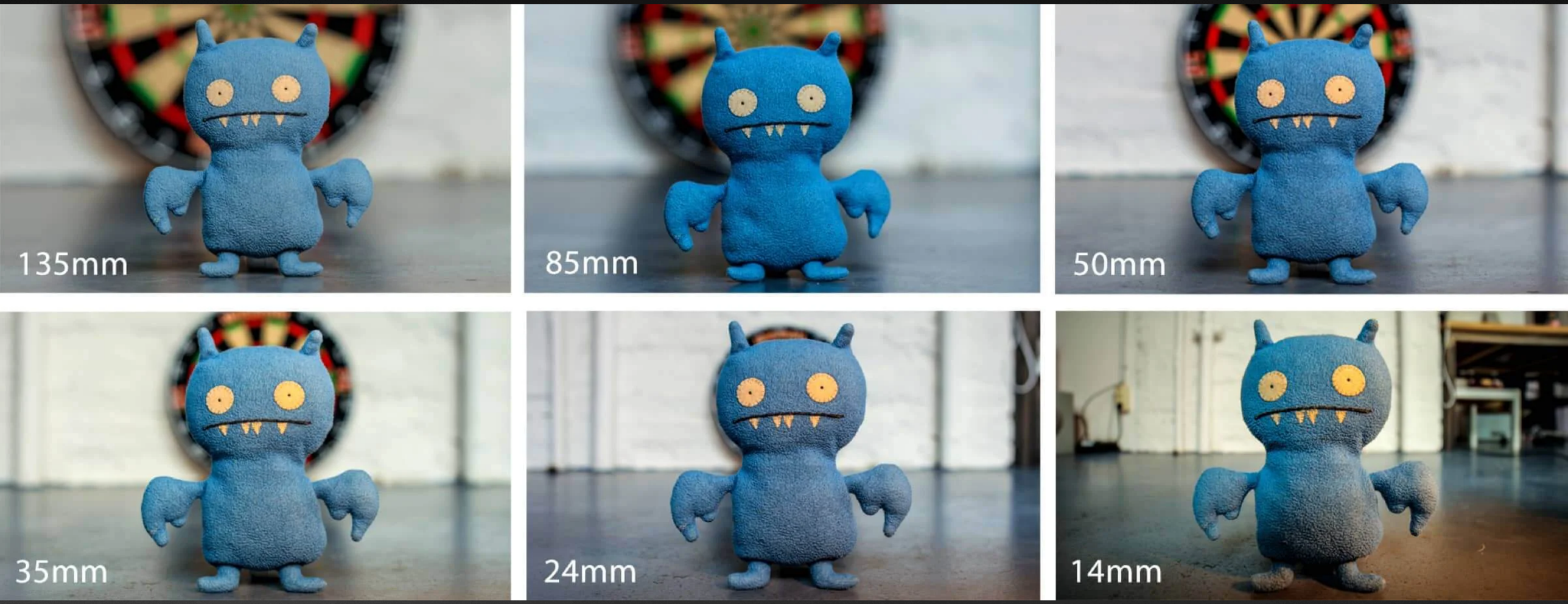
Minste detalj systemet kan oppløse: $y \approx \dfrac{1.22\lambda \cdot s}{D}$





Brennvidde / Fokalstørrelse: Effekt på utsnitt

Bilde fra: www.studiobinder.com



Brennvidde / Fokalstørrelse: Effekt på perspektiv

Bilde fra: www.studiobinder.com

Bildebehandling i Python

Lynrask introduksjon

Hovedmoduler: `numpy, scipy, PIL, matplotlib`

Andre: `skimage, opencv, torchvision`

Åpne et bilde med PIL og Numpy:

```
import numpy as np

import matplotlib.pyplot as plt

from PIL import Image

img = np.array(Image.open(“path/to/file.jpg”)) # Shape: [channels, height, width]

plt.imshow(img)

plt.show()
```