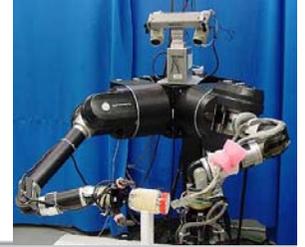


EVC - Computer Vision

Unit 3: Image Acquisition



<http://www.caa.tuwien.ac.at/cvl/teaching/sommersemester/evc>

- Content:
 - Human Eye
 - Image Geometry
 - Lenses
 - Radiometry
 - Resolution/Sampling
 - Image Sensors
 - Cameras
 - Color

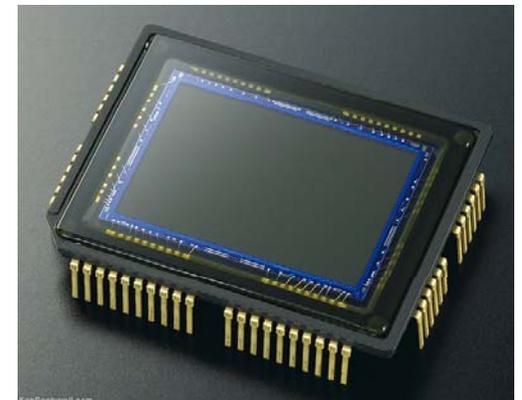
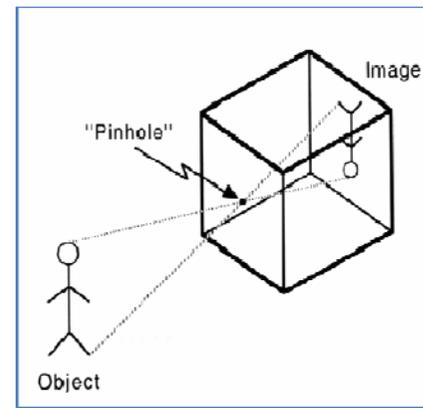
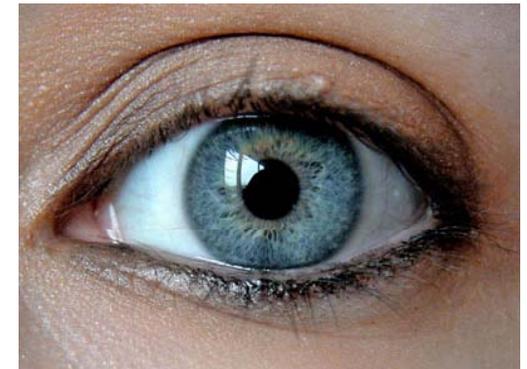
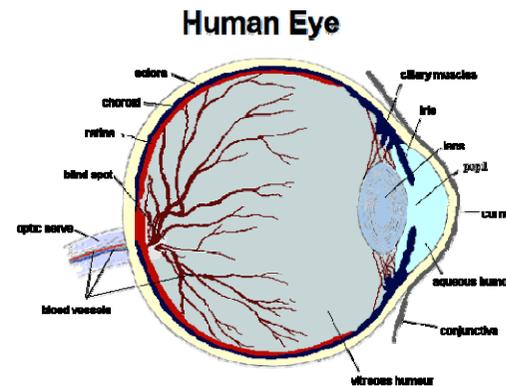
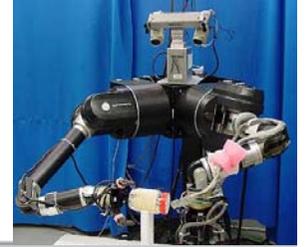
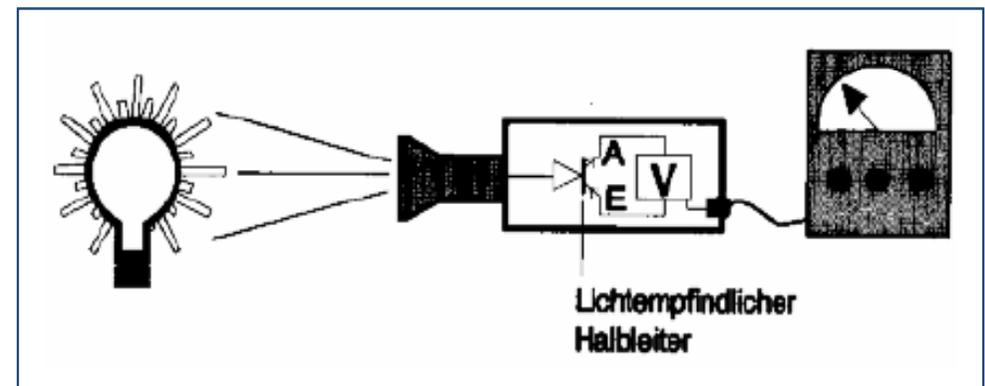
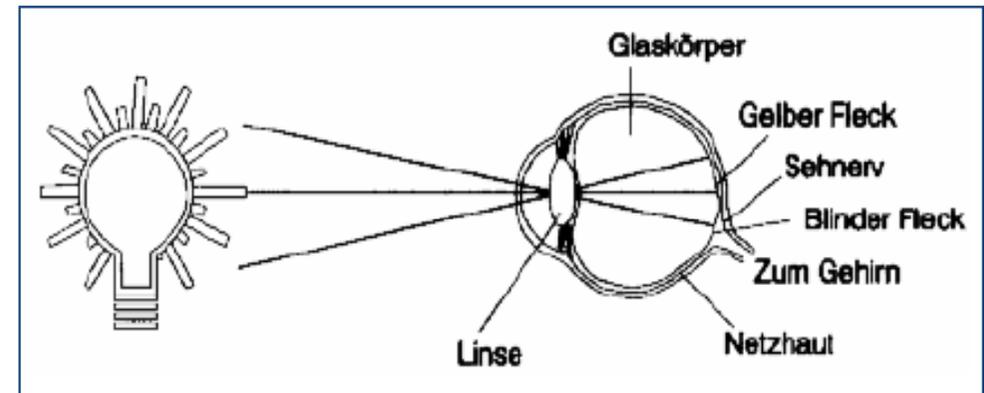


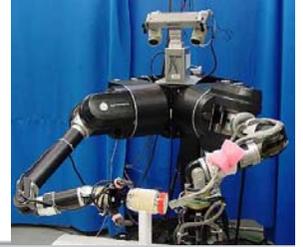
Image Formation



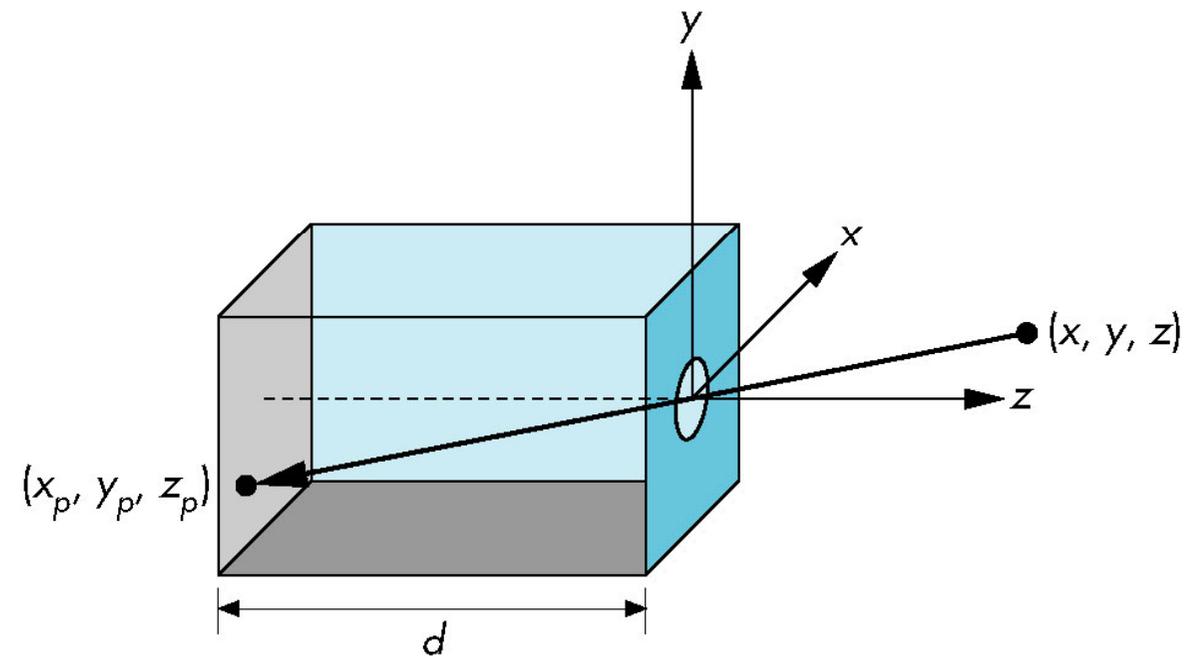
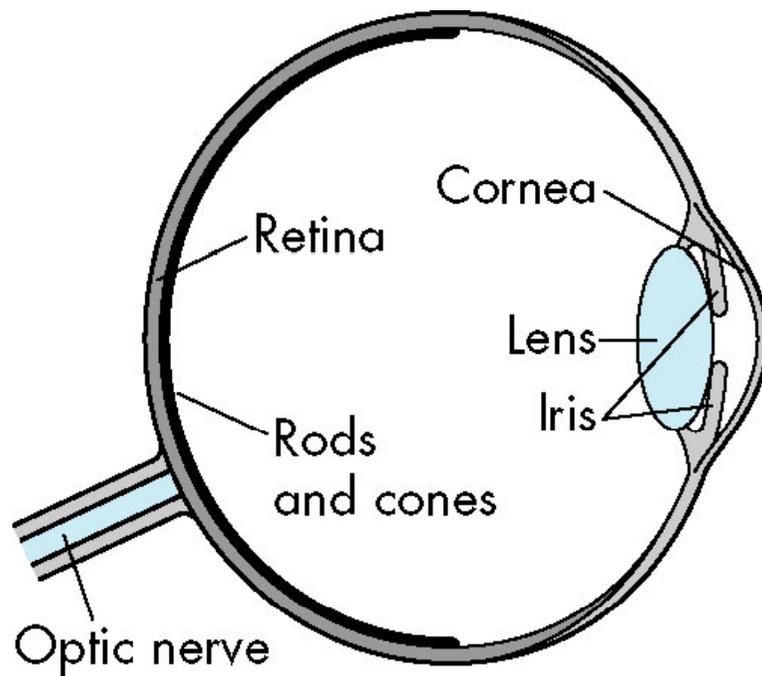
- Input in Human Vision:
 - Eye
- Input in Computer Vision:
 - Image
 - Role model: Human eye
 - Replica: CCD camera
 - Furthermore: Scanner, 3d Scanner,



Human Eye vs. Camera



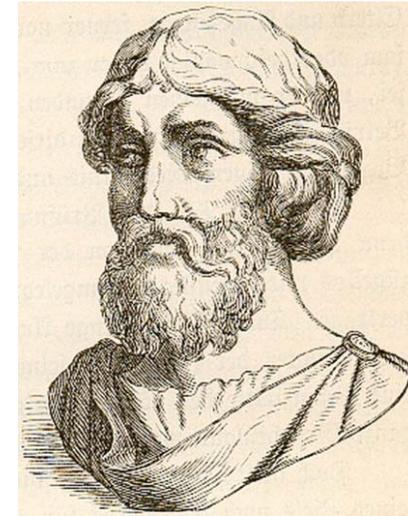
- We make cameras that act “similar” to the human eye



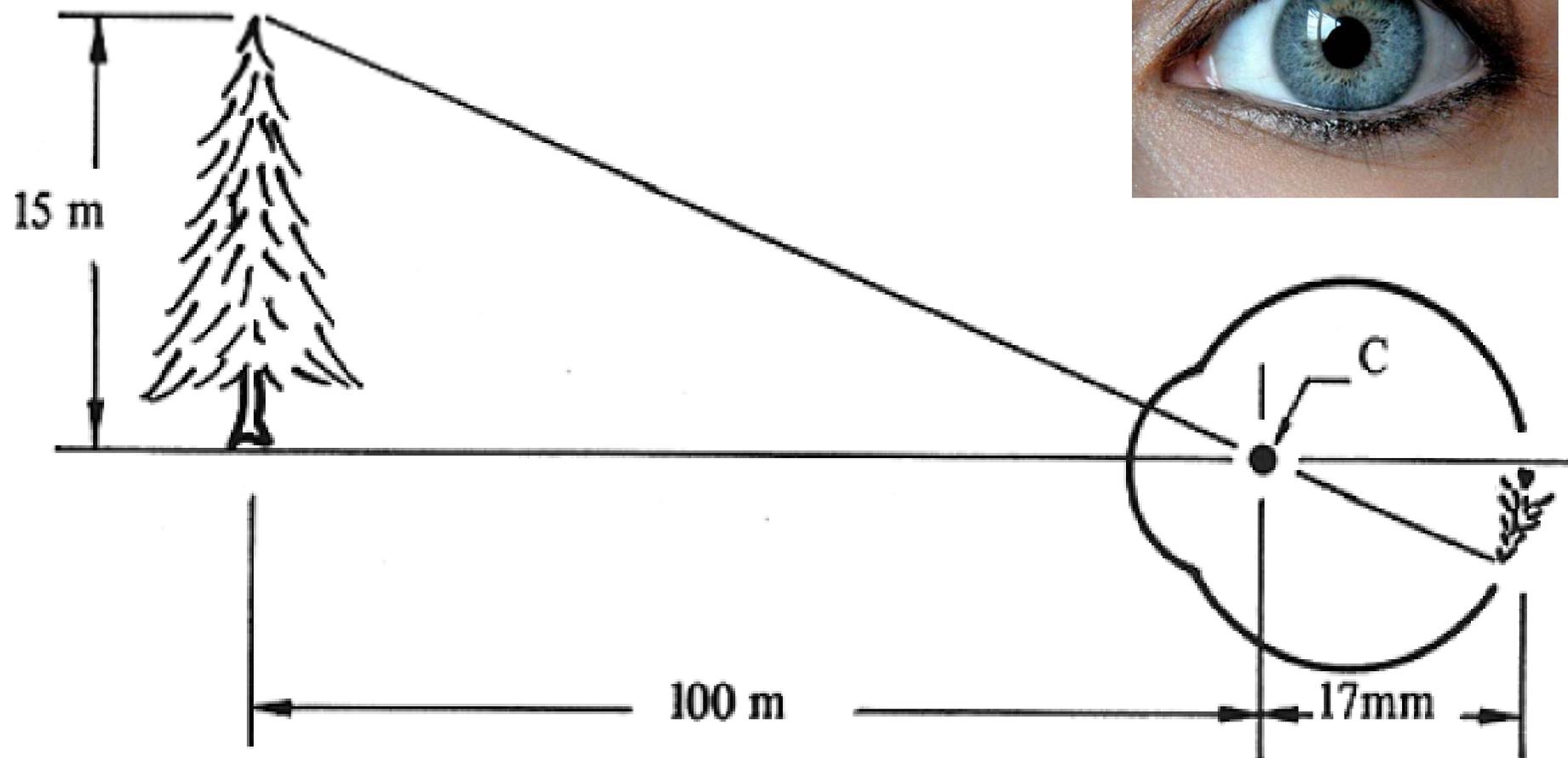
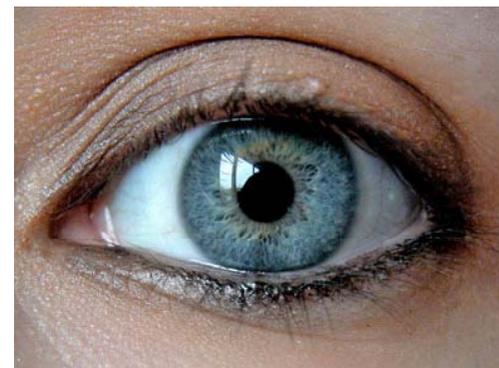
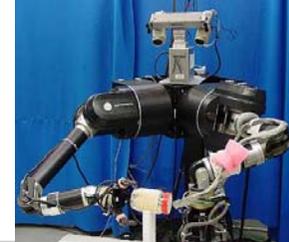
Human Eye - History



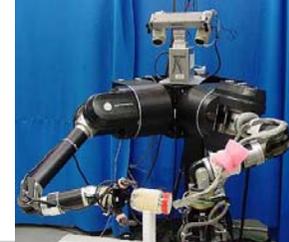
- **Pythagoras** (500 BC): Eye is **sending out rays** – by touching objects the seeing process is initiated (Range Finder Principle)
- **Kepler** (1604 AD): discovers vision process in human eye. On the retina an **upside-down image** of the world is sensed, which is assembled in the **visual center** into a 3d image.



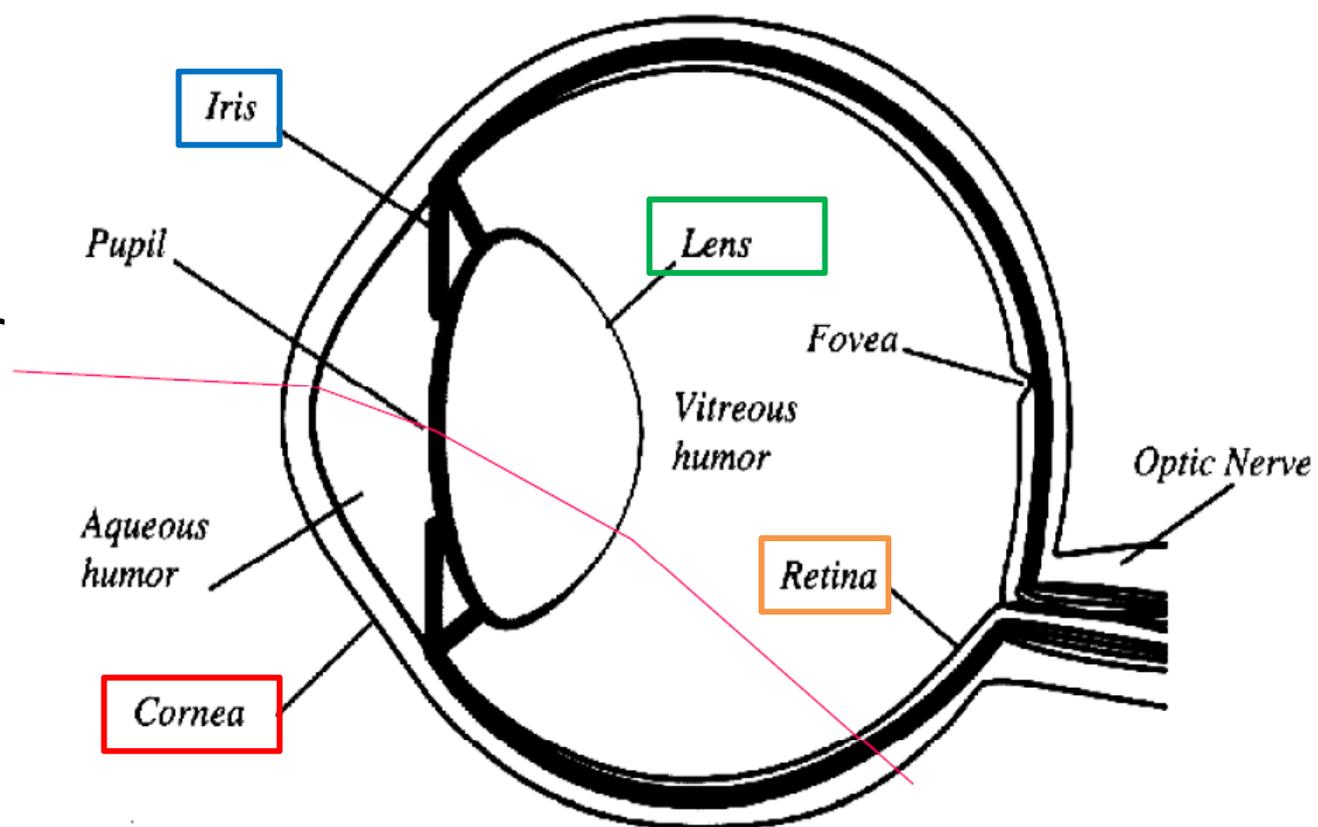
Human Eye



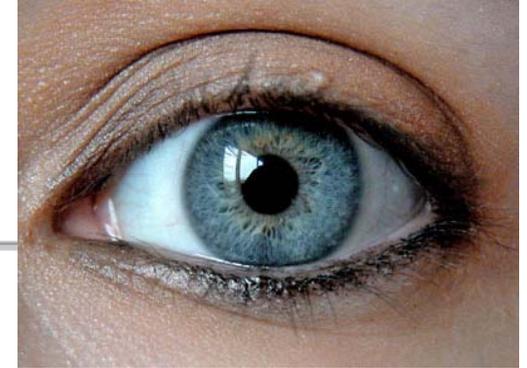
Human Eye - Components



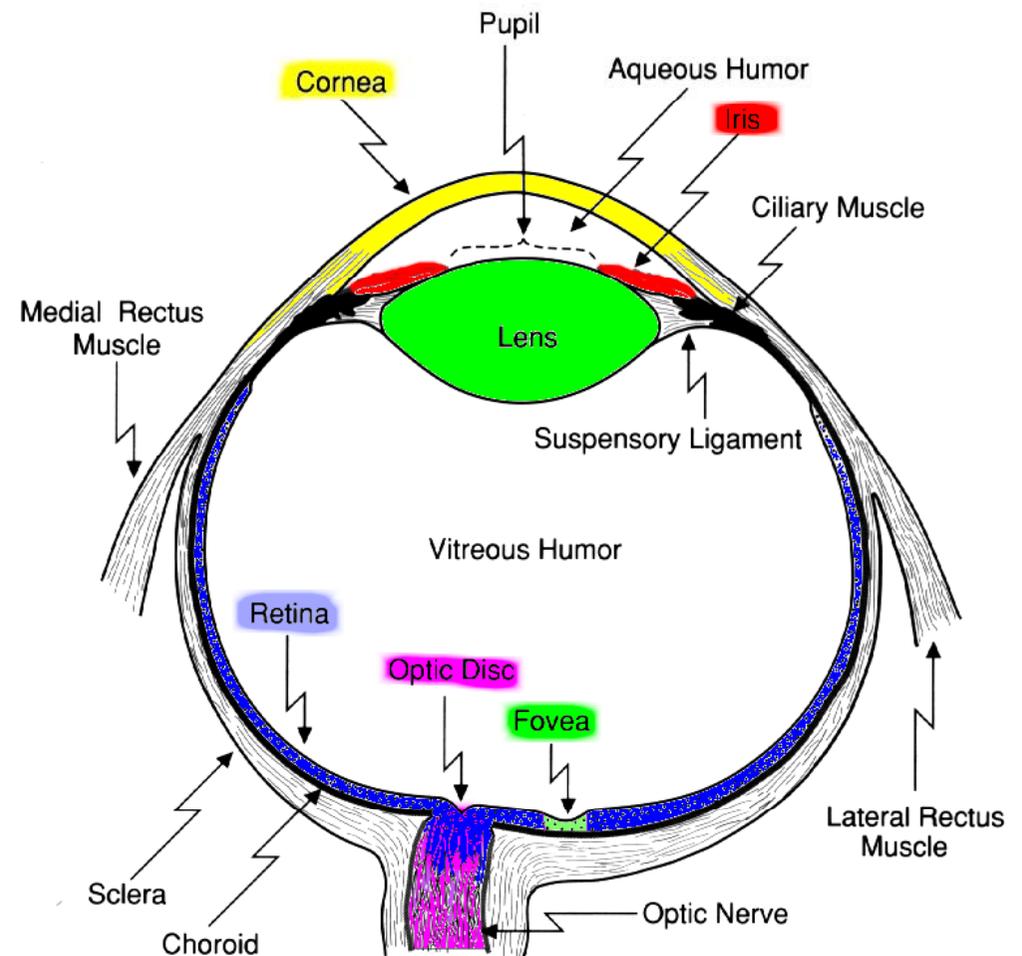
- **Cornea** + **Lens**:
 - Light fraction
- **Iris**:
 - variable aperture
- **Retina**: Image Detector
 - (ca. 100 Mio. Photoreceptors)



Human Eye - Accommodation



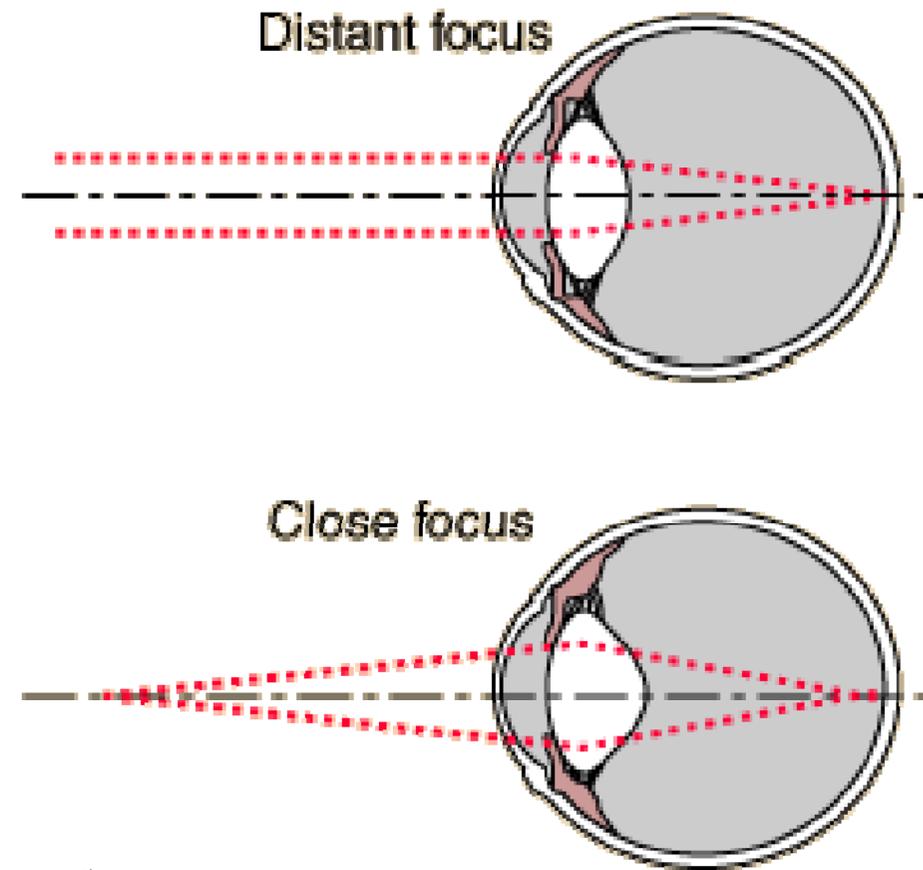
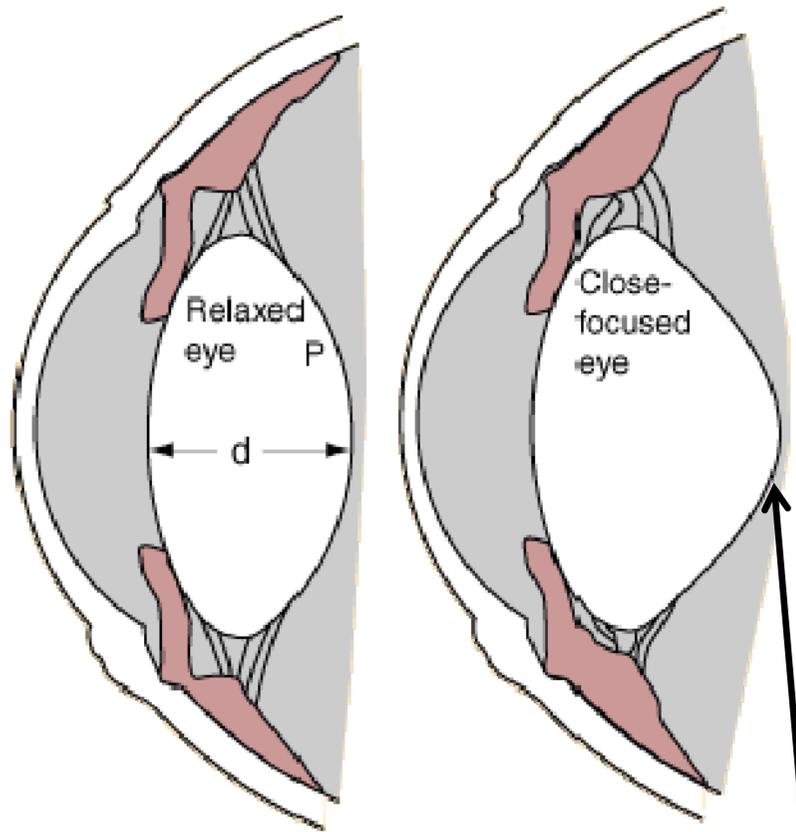
- Is the process by which the vertebrate **eye changes optical power** to maintain a **clear image (focus)** on an object as its distance varies.
- The image of the world is **represented exactly** on the retina. Objects too far forward or too far back to be mapped are **blurred**.



Accommodation

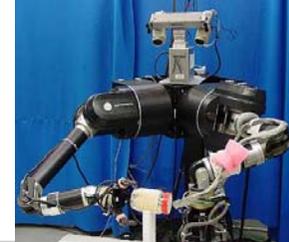


- Changes the focal length of the lens:

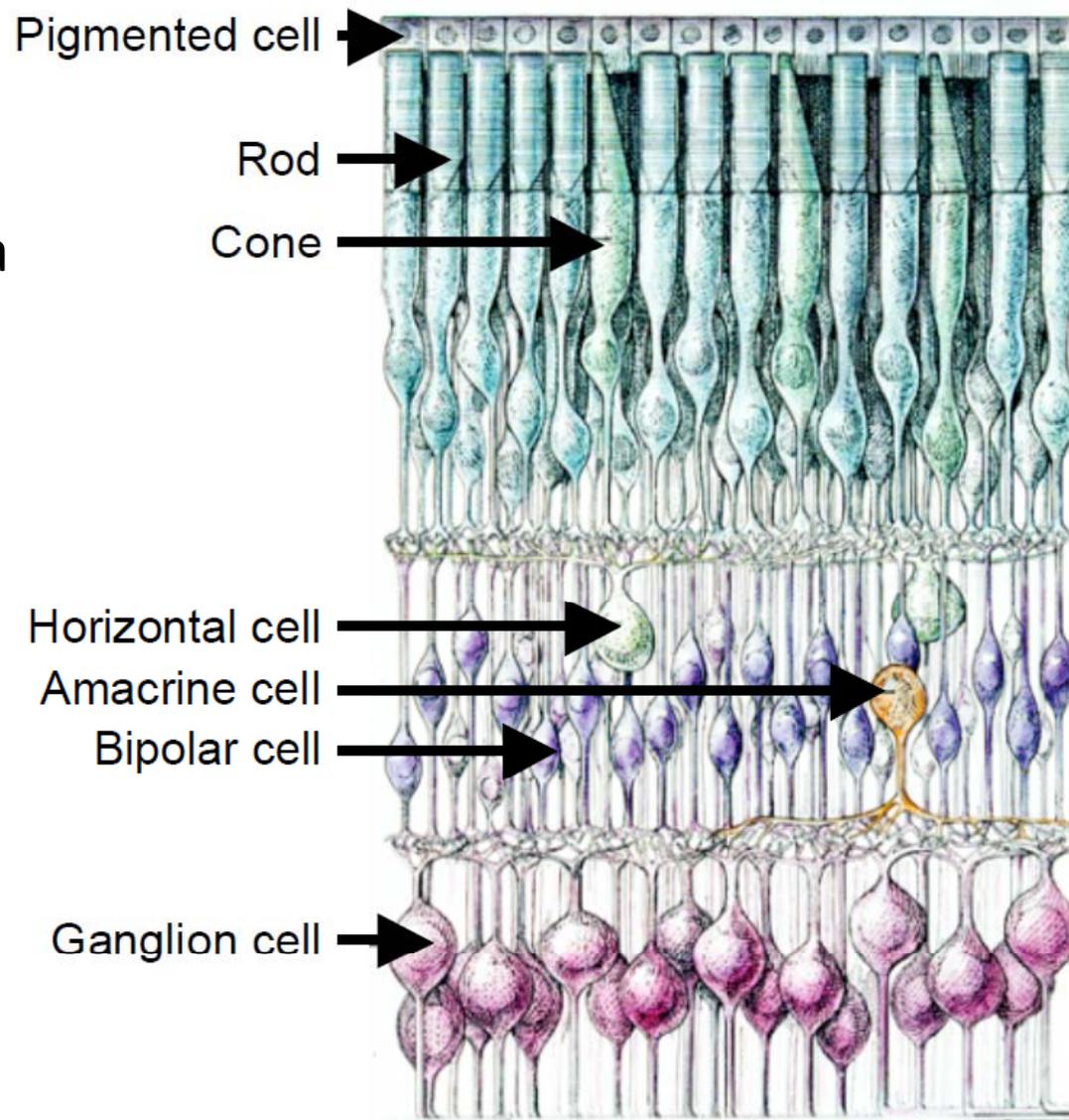


shorter focal length

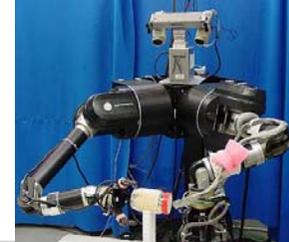
Retina



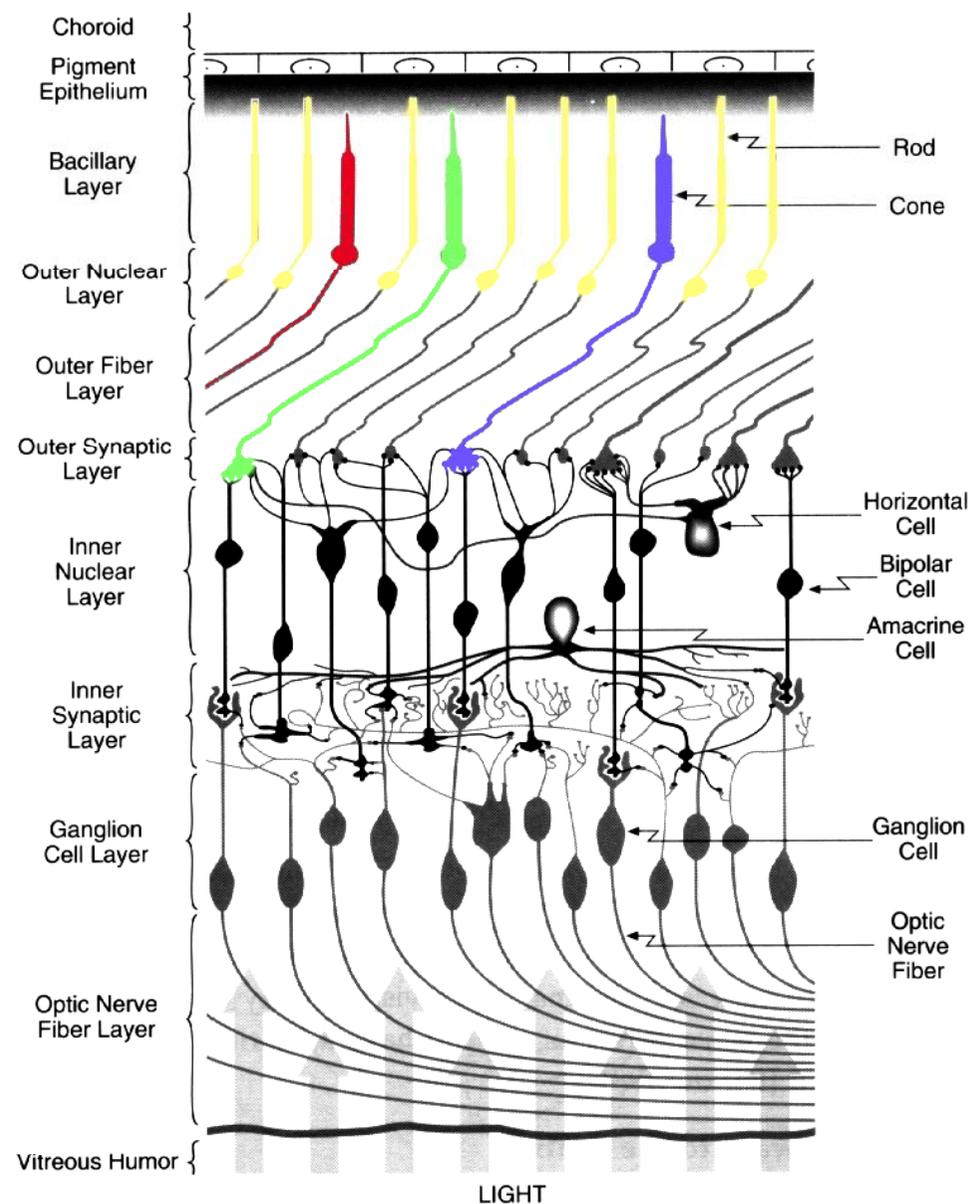
- **Light-sensitive tissue** lining inner surface of the eye.
- Light striking the retina initiates a cascade of **chemical and electrical events** that trigger nerve impulses.
- Retina -> optic nerve -> visual centers
- Fovea:
 - **sharp** central vision
 - **high** concentration of photoreceptors
- **approximately 50%** of the nerve fibers in optic nerve carry information from fovea



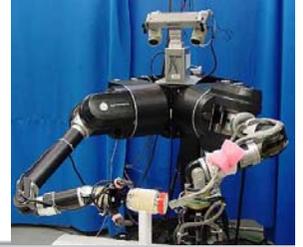
Retina



- Rods: Monochrome
- Cones: Color (RGB)
- Fovea: Cones only
- Number: 6 Mio. Cones
120 Mio. Rods
- But only **1 Mio. nerve fibers** in optic nerve => **intelligent sensor!**

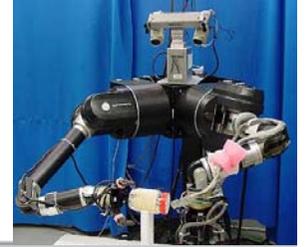


Blind Spot in Eye

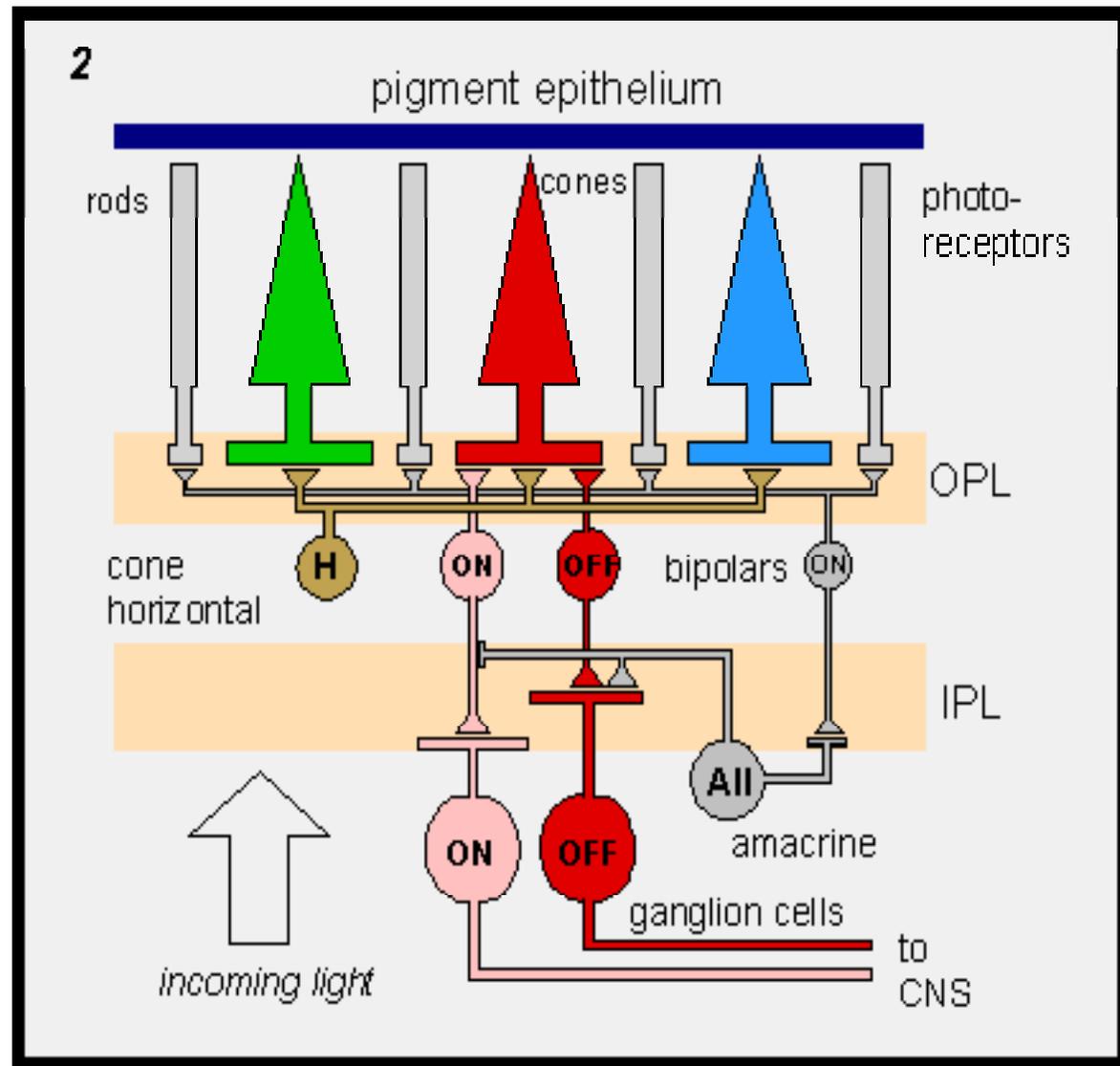


Close your right eye and look directly at the “+”

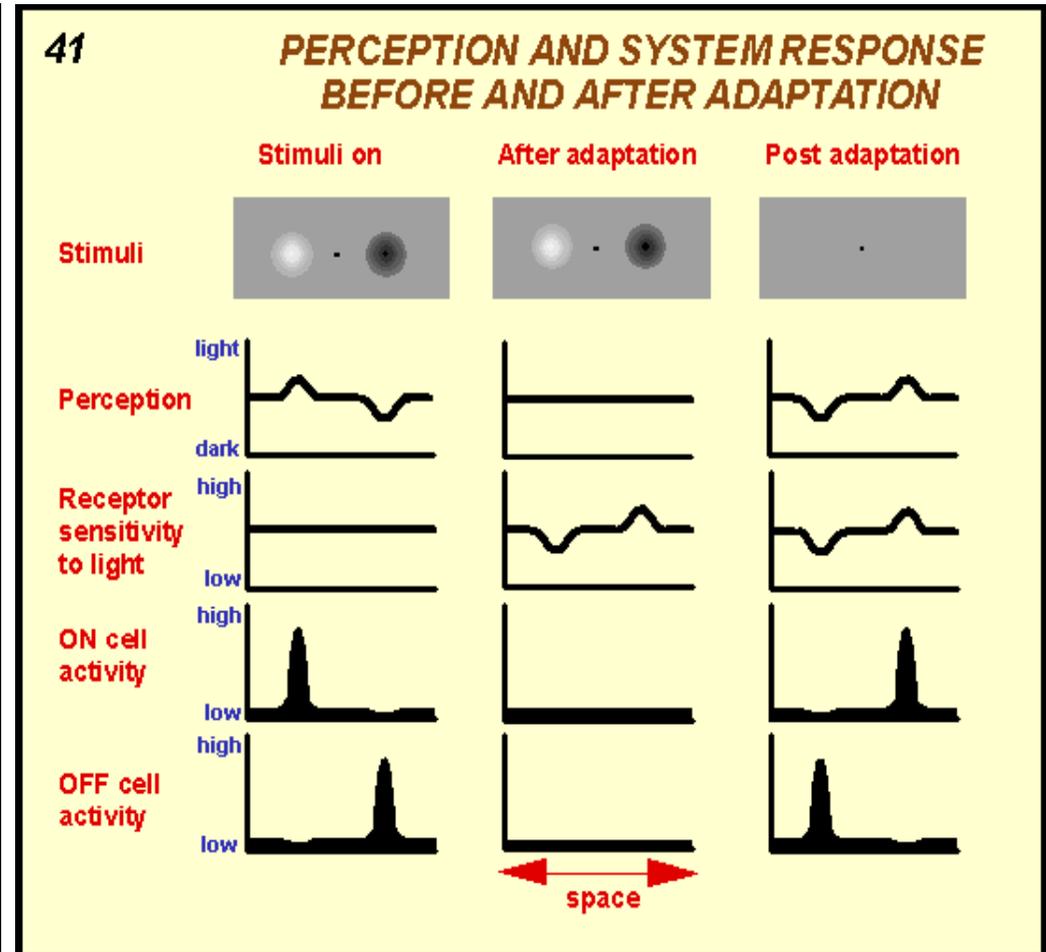
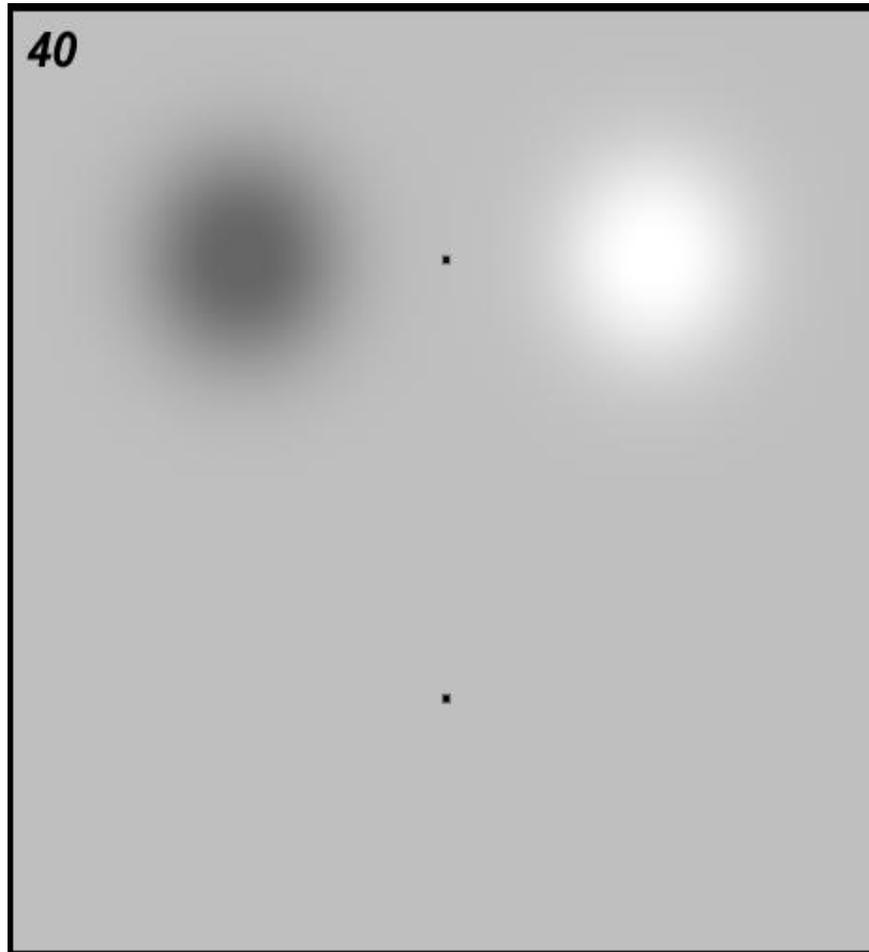
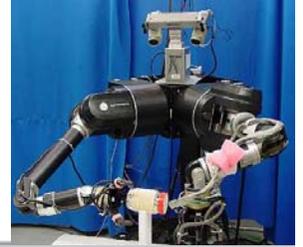
Cells of Retina



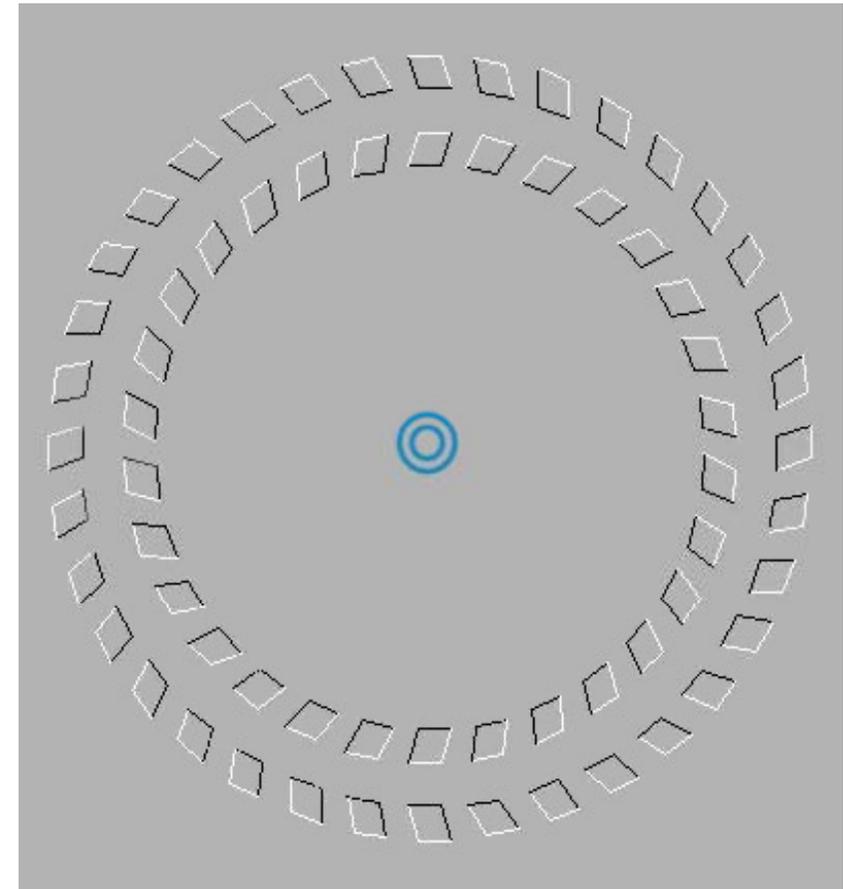
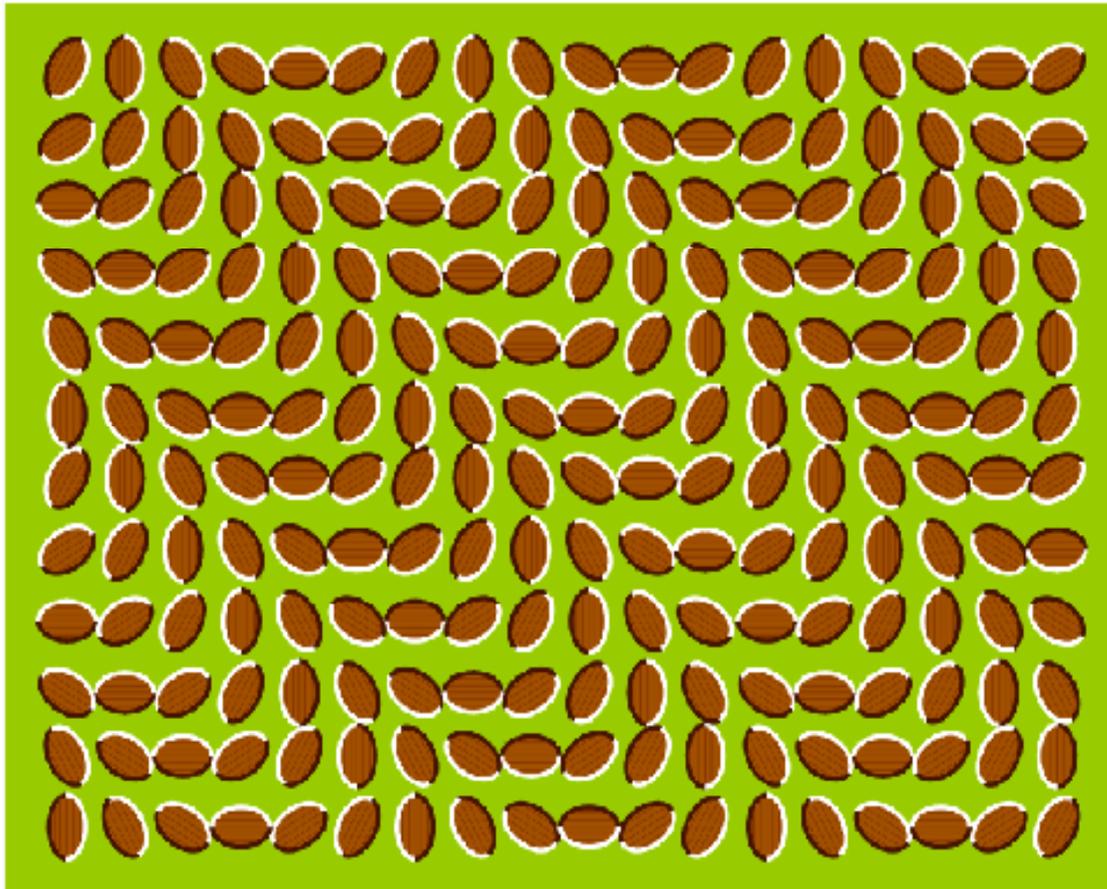
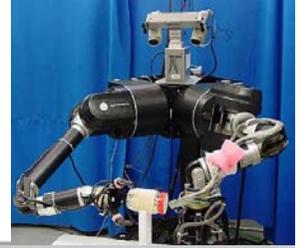
- Rods
- Cones
- Filter cells
 - Horizontal
 - Bipolar
 - Amacrine



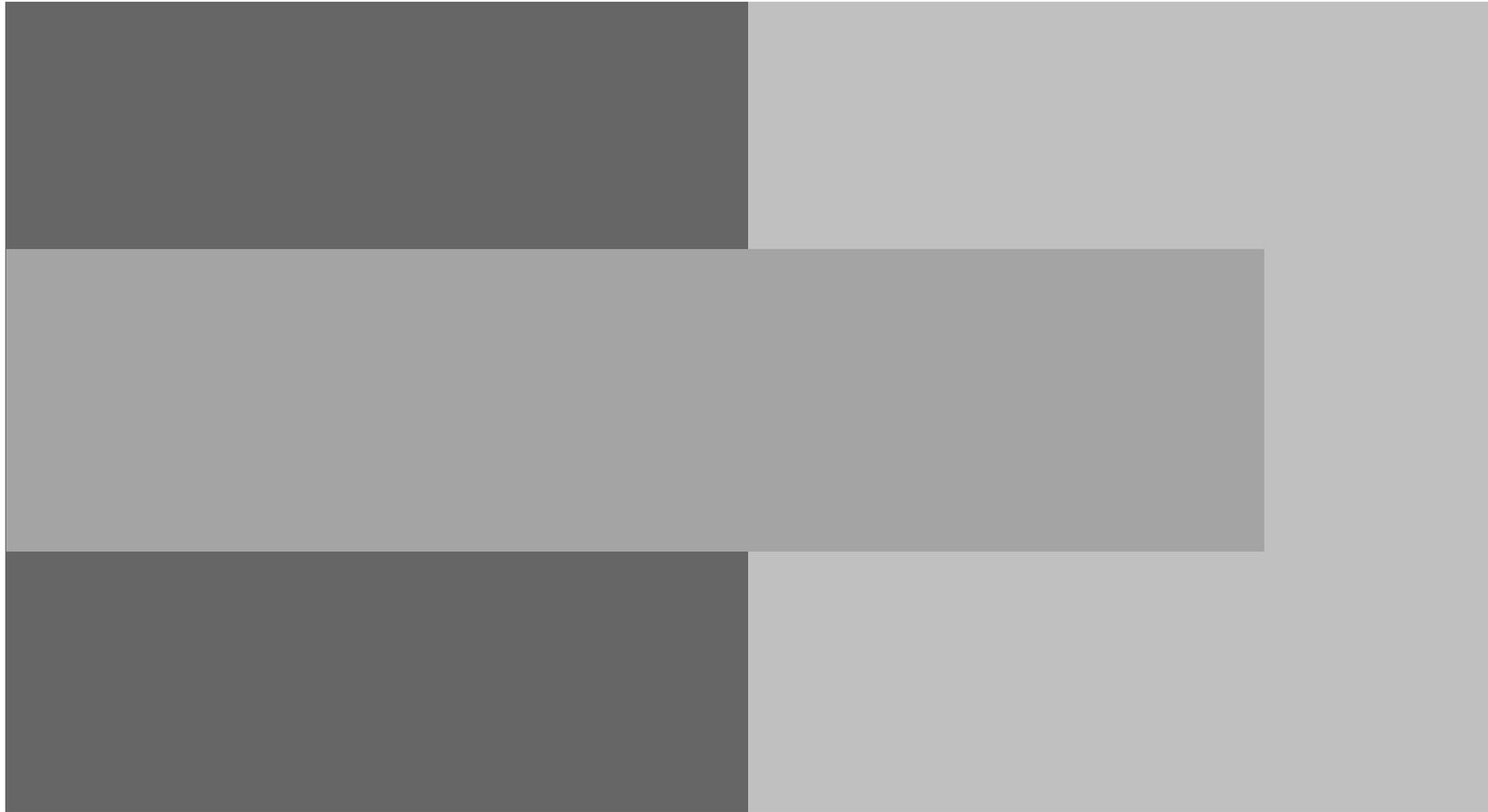
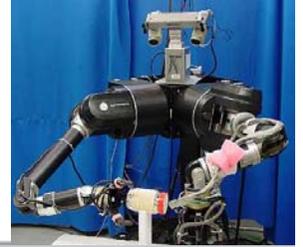
Afterimages



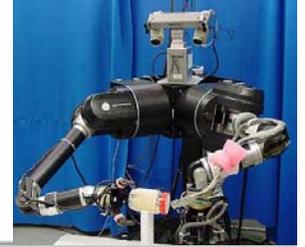
„Movement“ in Static Images



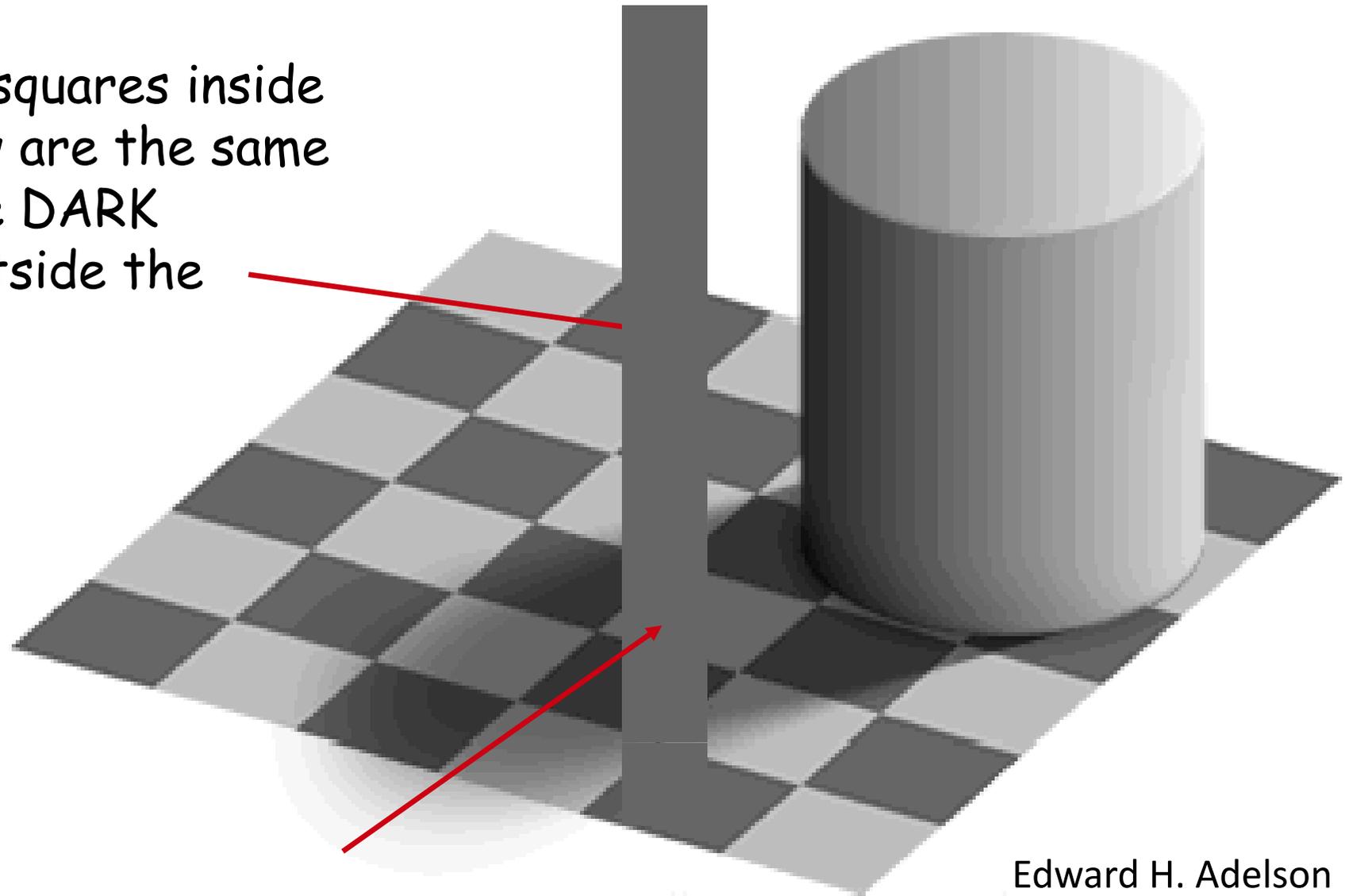
Color Constancy



Color Constancy

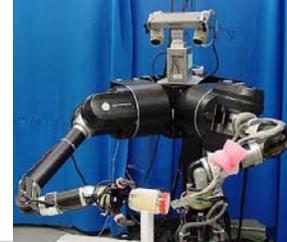


The white squares inside the shadow are the same grey as the DARK squares outside the shadow!

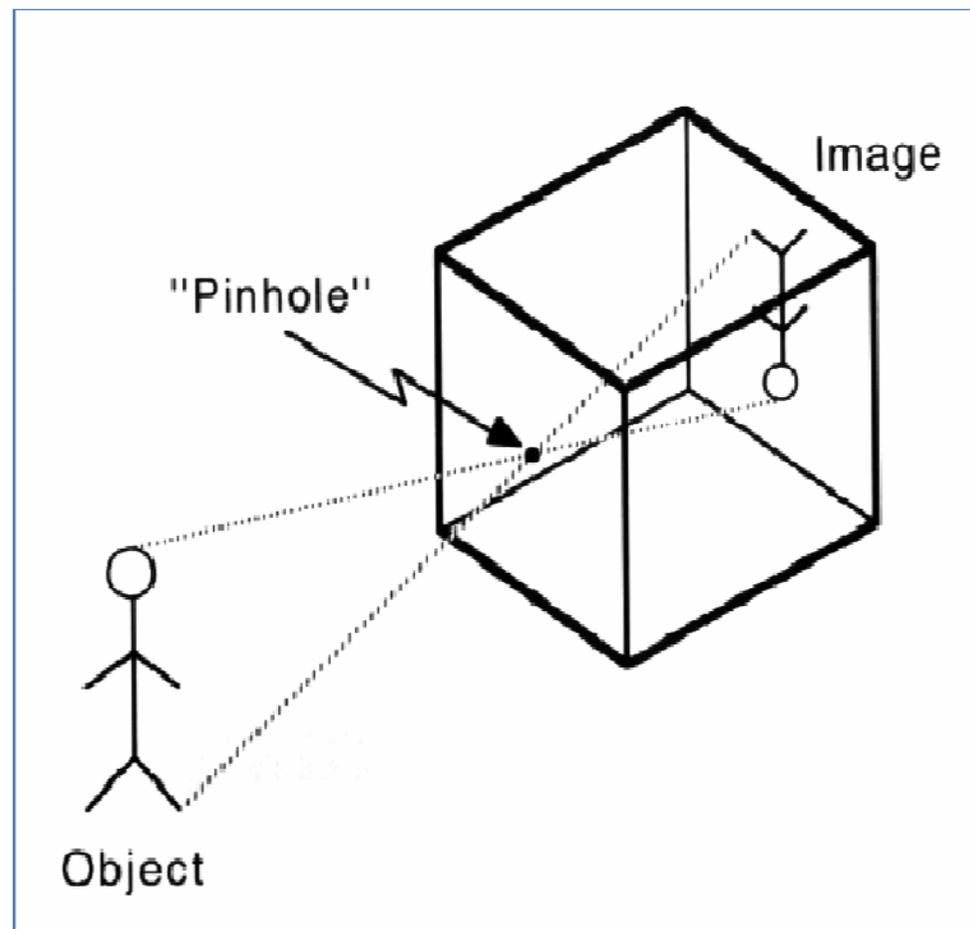


Edward H. Adelson

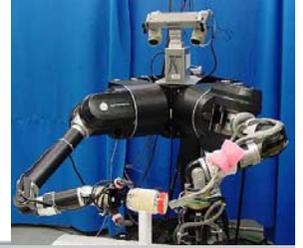
Image Geometry



- Simplest Model: Pinhole camera
 - Has a very small hole (Aperture = ∞), Light is led through the hole and forms an image at the back of the box (upside down and side-inverted)



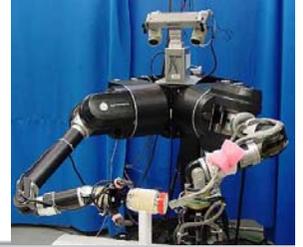
Earliest Surviving Photograph



- First photograph on record, “la table service” by Nicephore Niepce in 1822.



A Brief History of Images



● 1568



● 1837

Still Life, Louis Jaques Mande Daguerre, 1837

A Brief History of Images



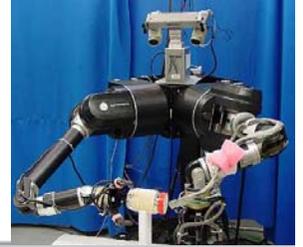
● 1568

● 1840?



Abraham Lincoln?

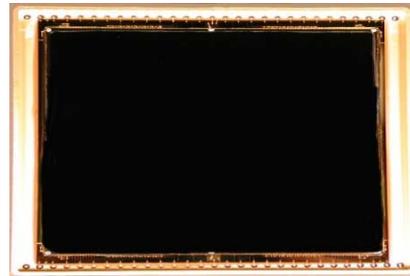
A Brief History of Images



1568

1837

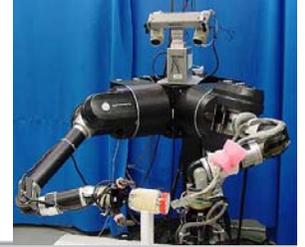
1970



Silicon Image Detector, 1970



A Brief History of Images



Digital Cameras

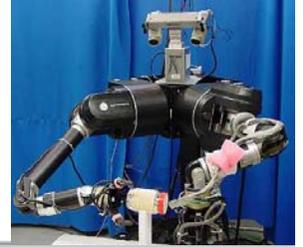
1568

1837

1970

1995

A Brief History of Images



Nikon D3x, 24,5 MPix

● 1568

● 1837

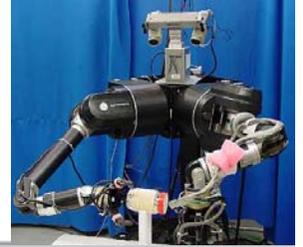
● 1970

● 1995

● 2012

Image Formation

Image Formation



- Images are two-dimensional patterns of brightness values.
- They are formed by the projection of 3D objects.

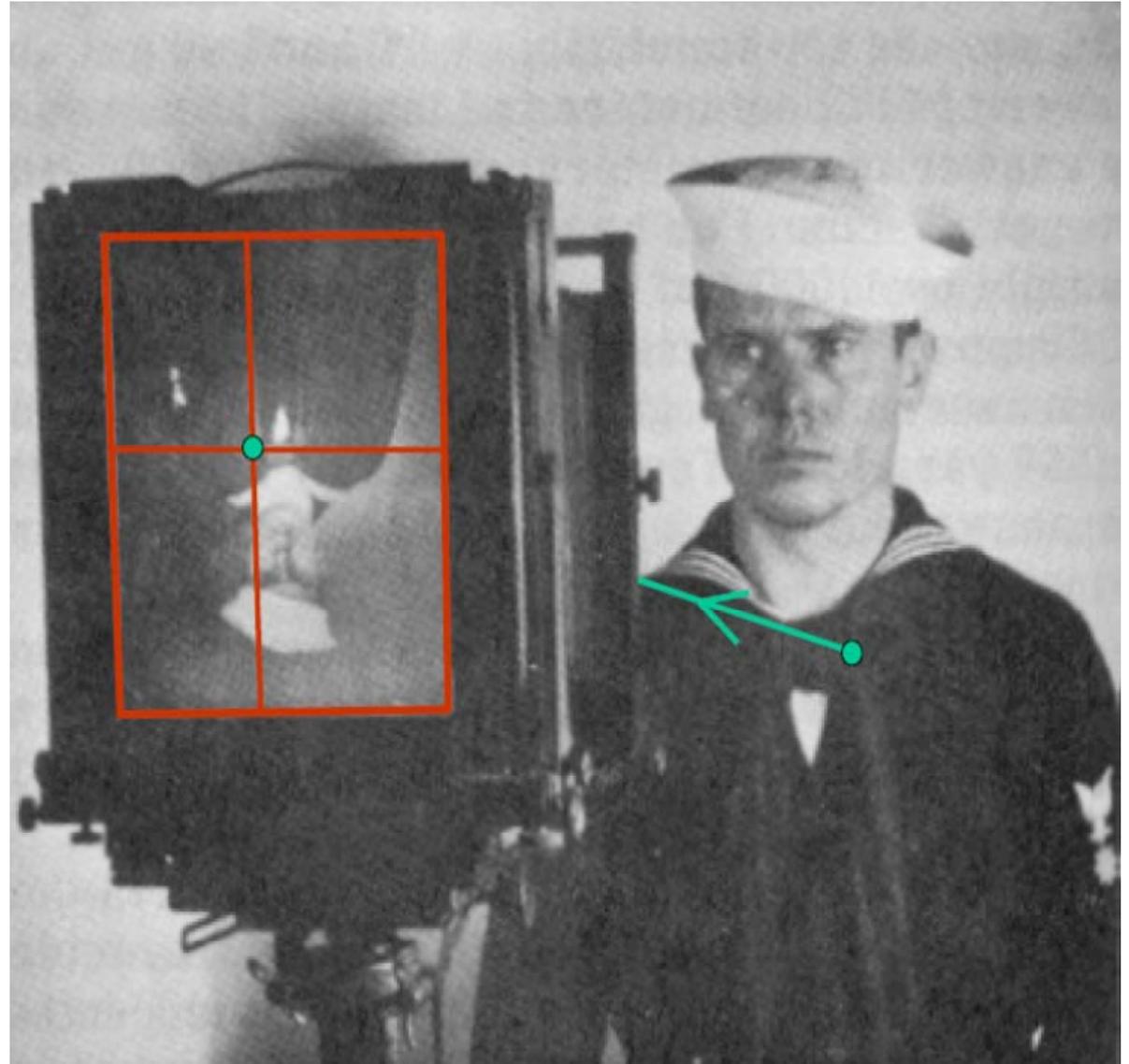
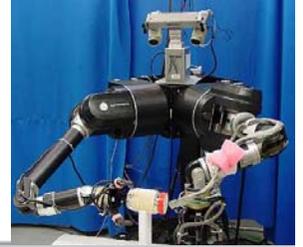
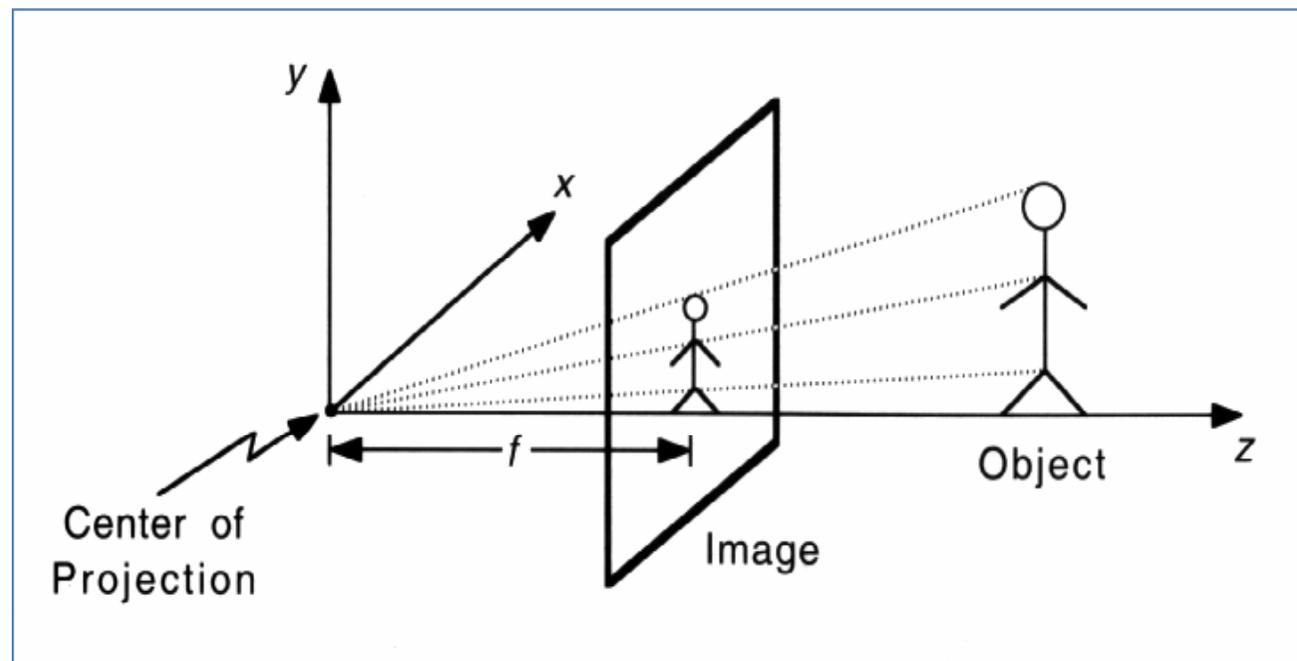


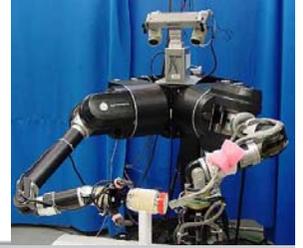
Image Geometry



- Perspective Projection (Central projection)
 - Is the projection of the 3d world onto a 2d plane by rays passing through a common point the center of projection.
 - => models image formation by a pinhole camera

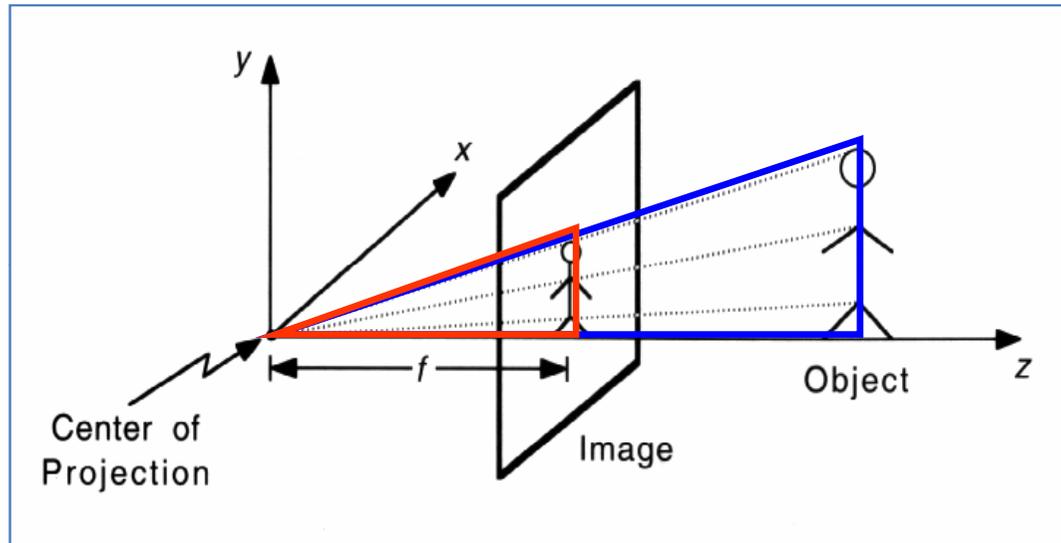


Equations of the perspective projection



$$x = \frac{f}{Z} X$$

$$y = \frac{f}{Z} Y$$



$$\frac{x}{X} = \frac{f}{Z}$$

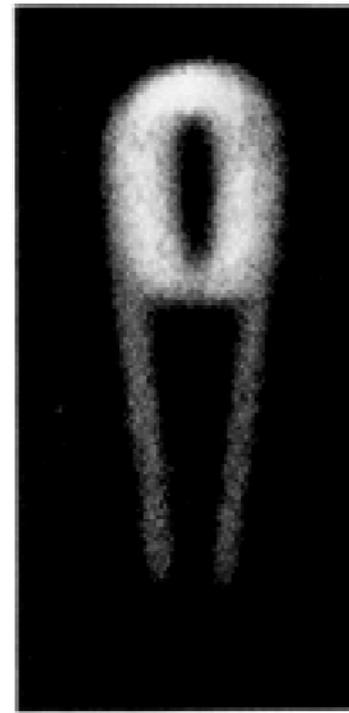
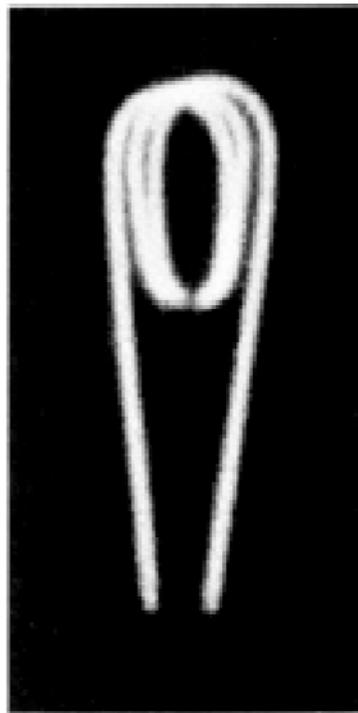
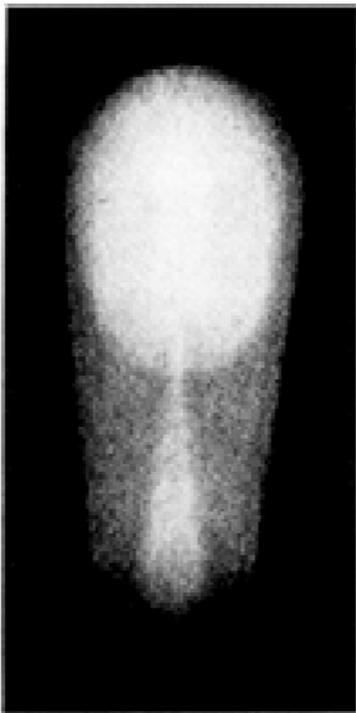
$$\frac{y}{Y} = \frac{f}{Z}$$

- Perspective projection is non-linear !

Recap: Limits of Pinhole Cameras

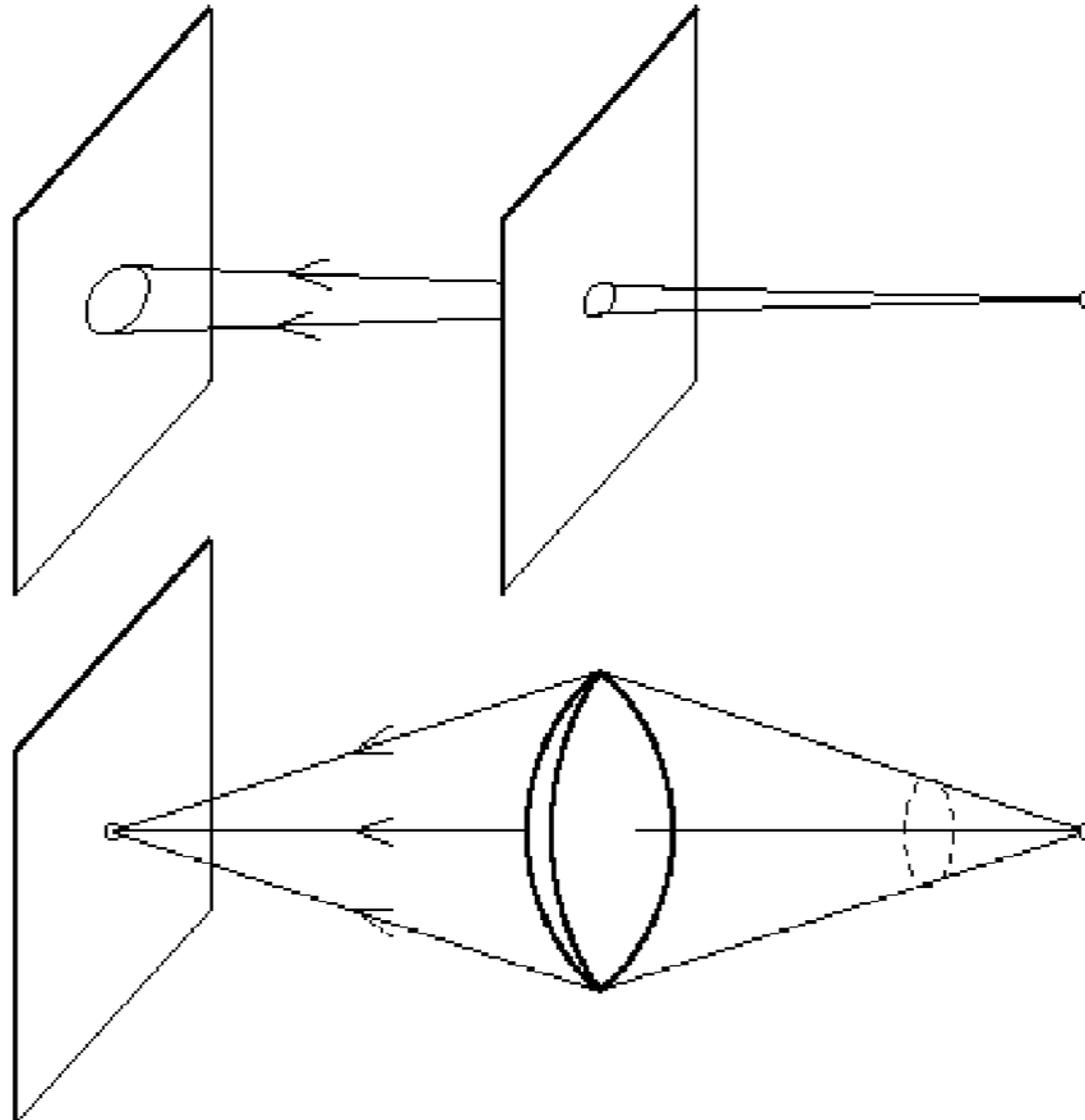
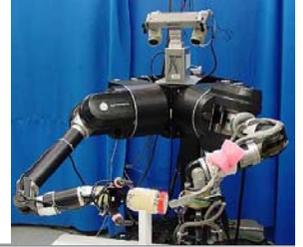


- A picture of a filament taken with a pinhole camera. In the image on the left, the hole was too big (blurring), and in the image on the right, the hole was too small (diffraction).

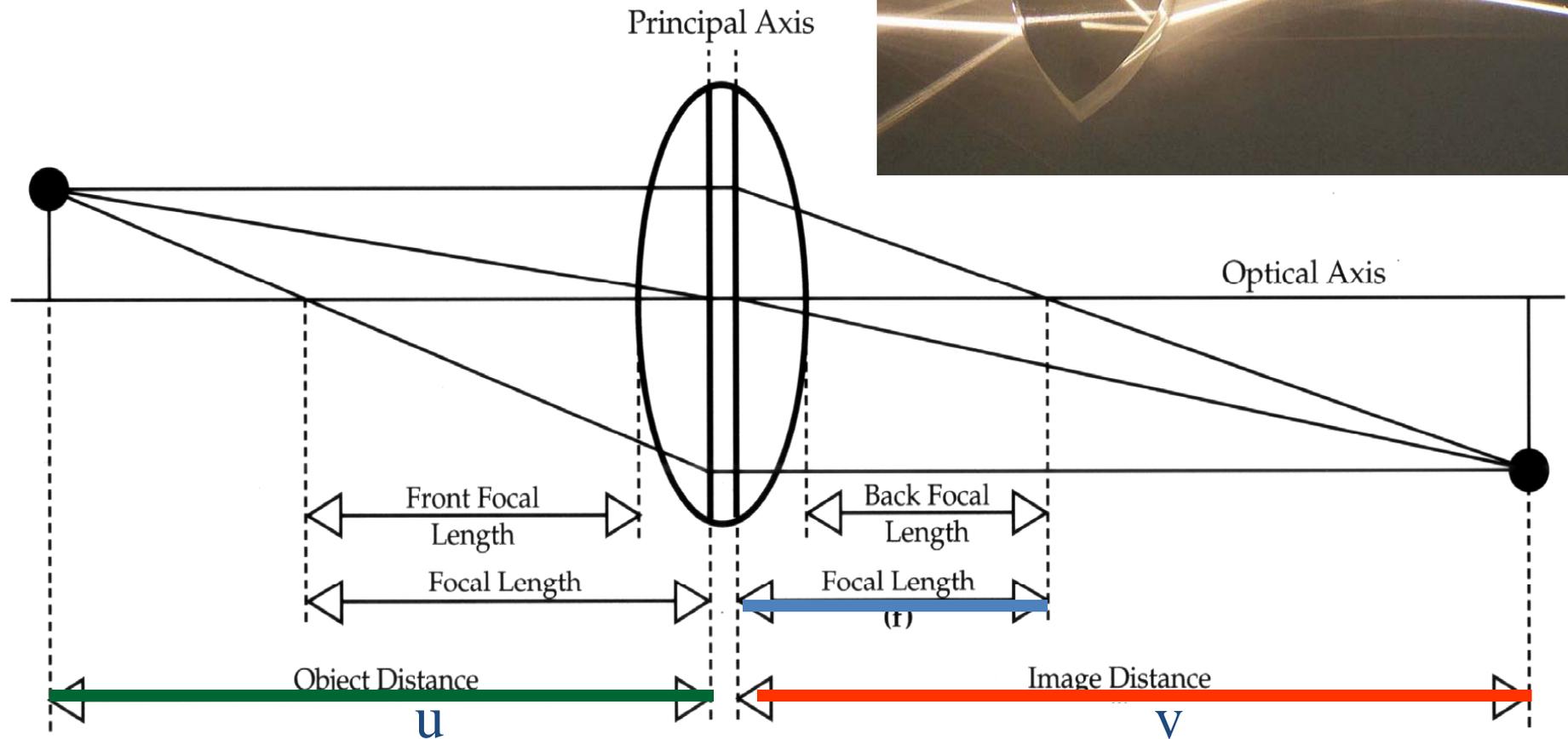
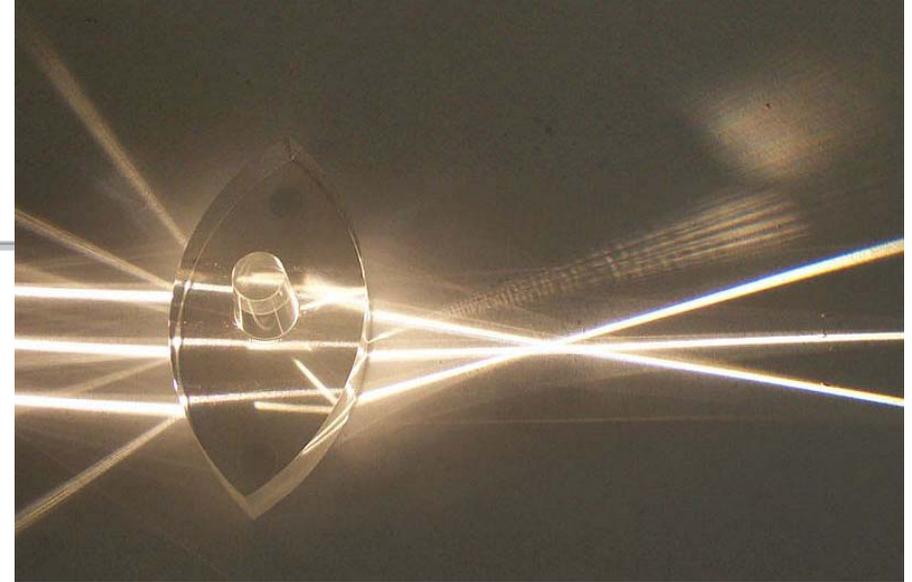


Ruechardt, 1958

Cameras with Lenses

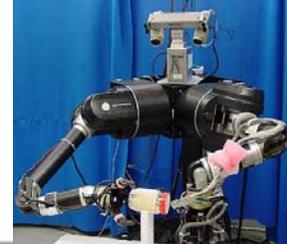


Simple Lens Parameters



FOCAL LENGTH = Distance from focus point to principal axis.

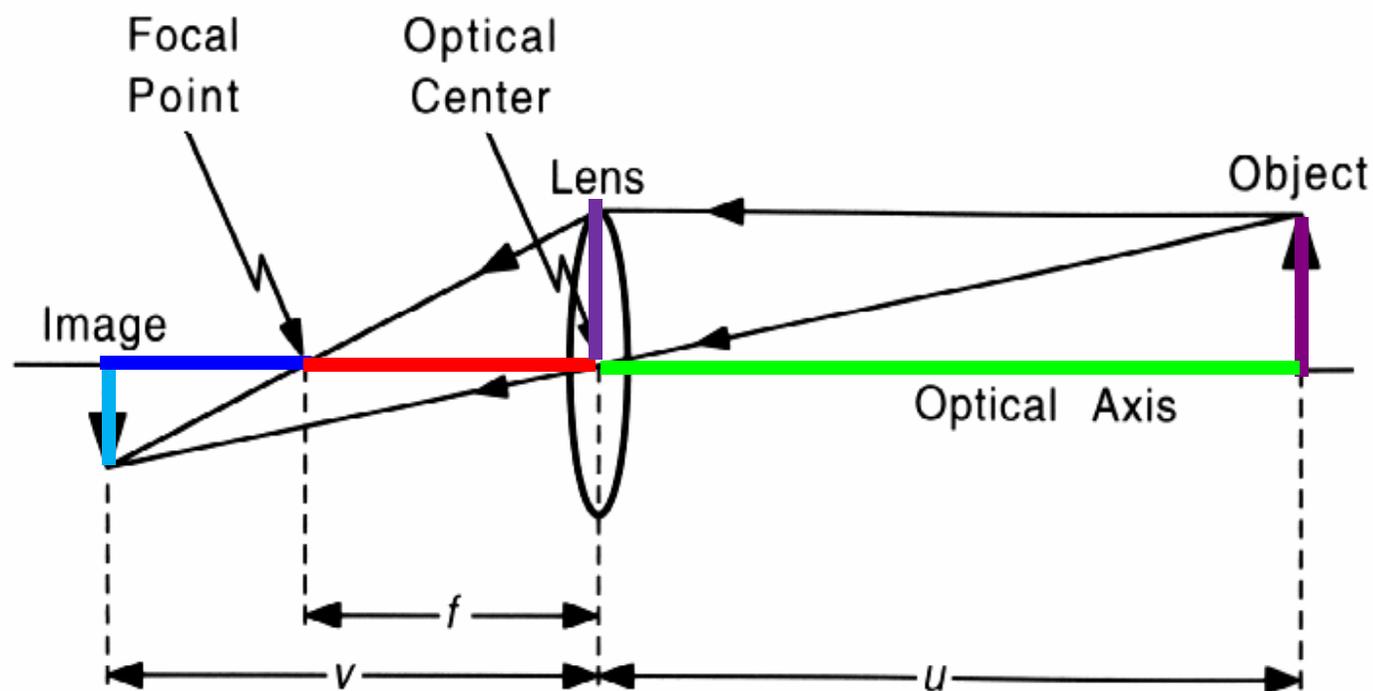
Lenses



- Pin has no lens => small Aperture => few light
 - „thin" lenses: small Aperture but much light
- Thin lens law:

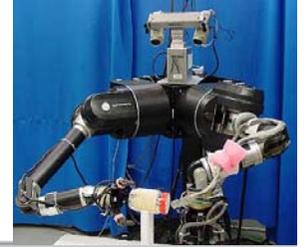
$$\frac{|y_0|}{|y_i|} = \frac{u}{v}$$

$$\frac{|y_0|}{|y_i|} = \frac{f}{v - f}$$

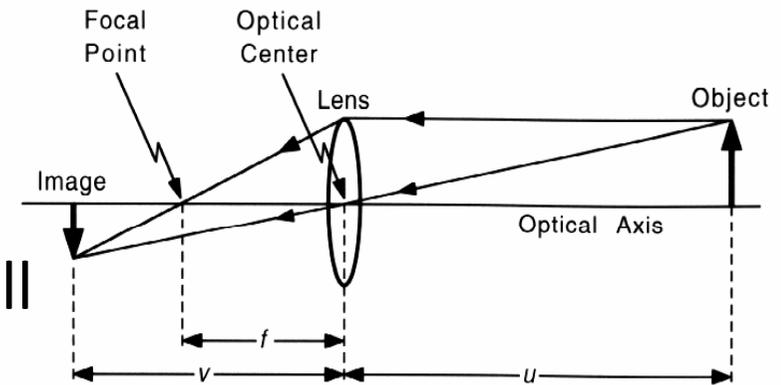


Thin-Lens Equation: $\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$

Lenses



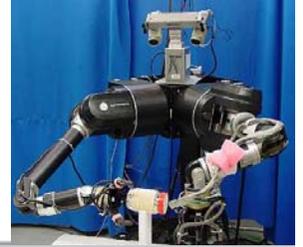
- f : focal length = distance of the point on the optical axis where all rays emerging from infinity meet to the lens plane (= all rays are parallel to the optical axis)
- if $u = \infty$ then $v = f$
- Rays going through the optical center of the lens are not diffracted
- Field of view: area that is recorded by a camera:
 - The bigger f the smaller the area that is imaged
 - Wide-angle - small f ; Zoom - large f



Thin-Lens Equation: $\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$

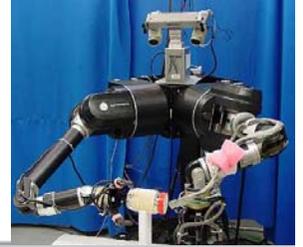
$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$$

Depth of Field



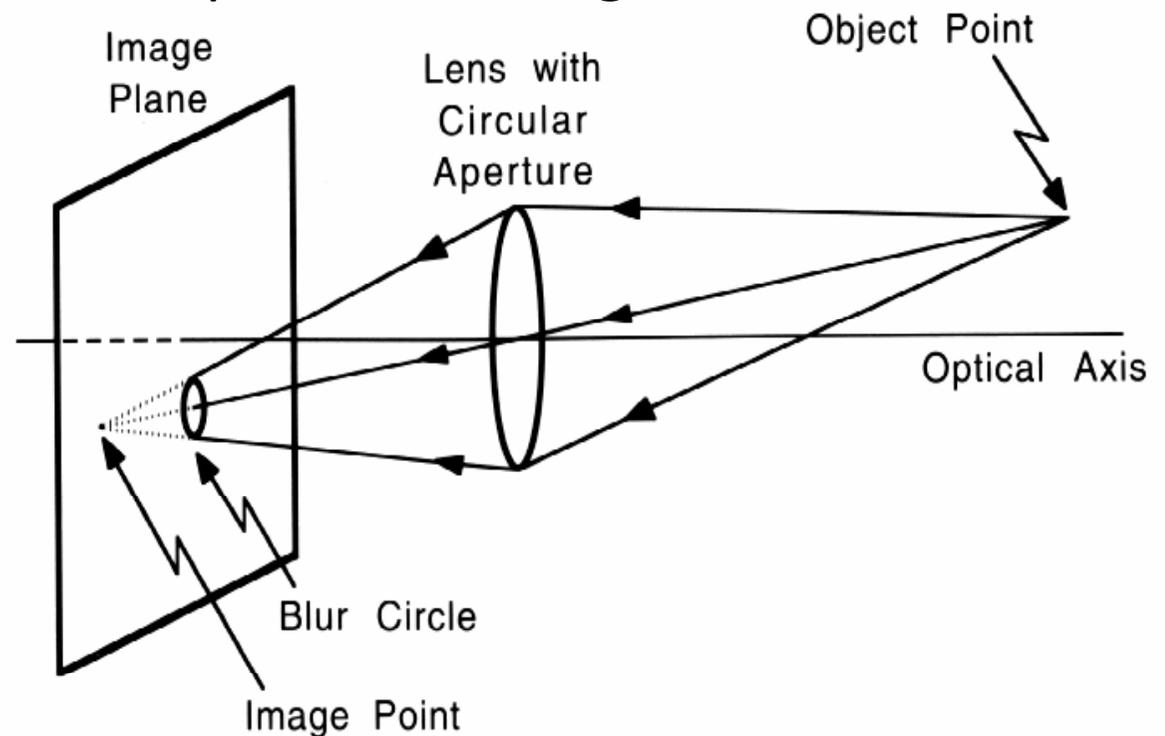
Same F/stop setting was used on all three lenses.
Note the difference in depth of field.

Depth of Field



- Only objects in a certain distance are imaged sharply at the image plane, all other distances are blurred because of blur circles.
- The bigger the aperture, the bigger the blur circles
- The smaller the aperture, the sharper is the image

- The bigger the depth of field the darker the image
- Large Aperture = small depth of field



Depth of Field

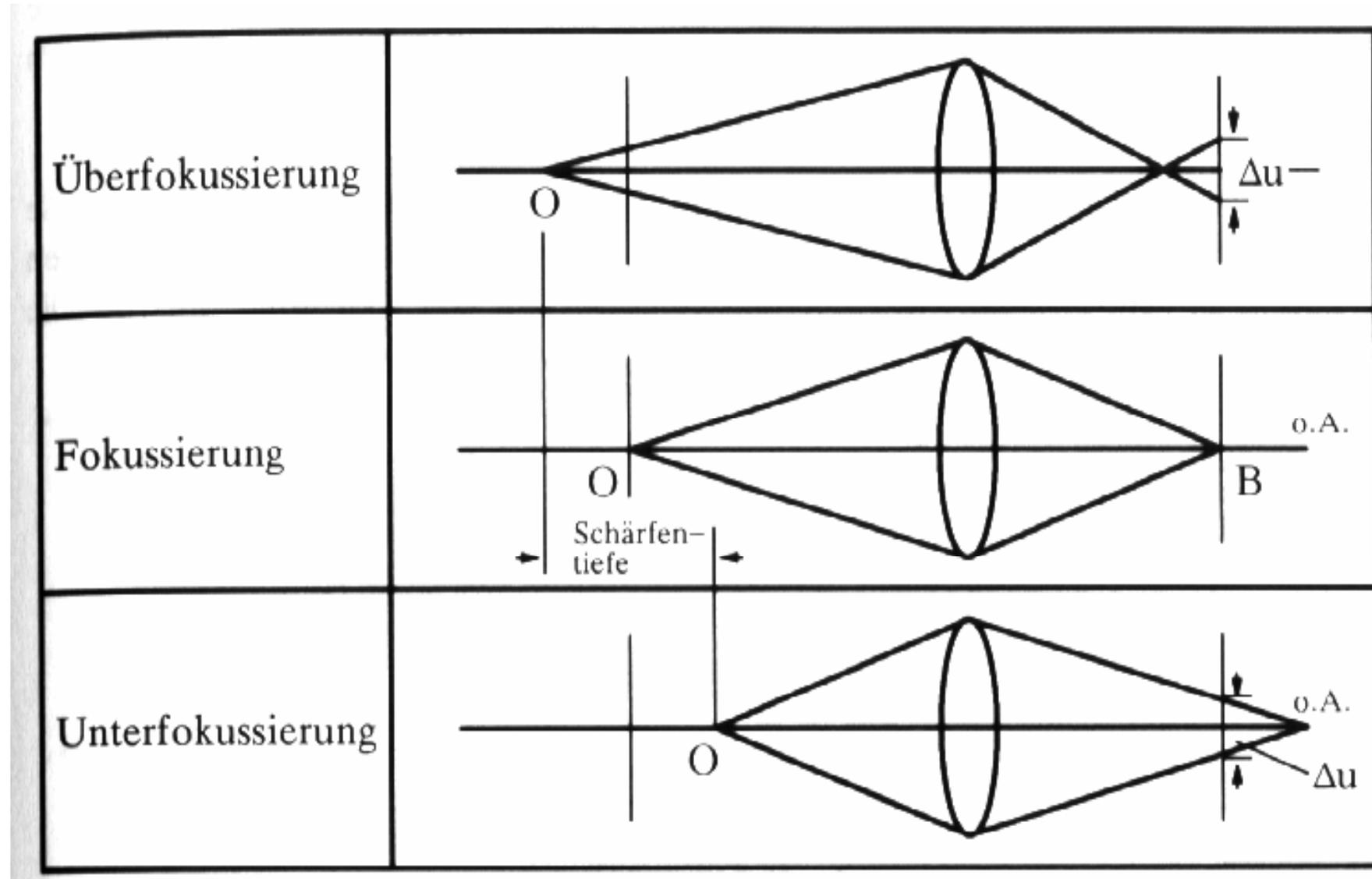
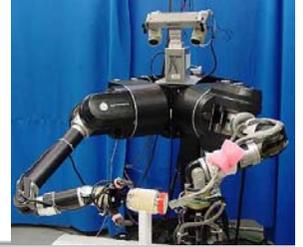
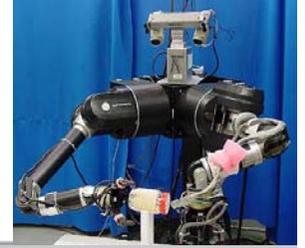


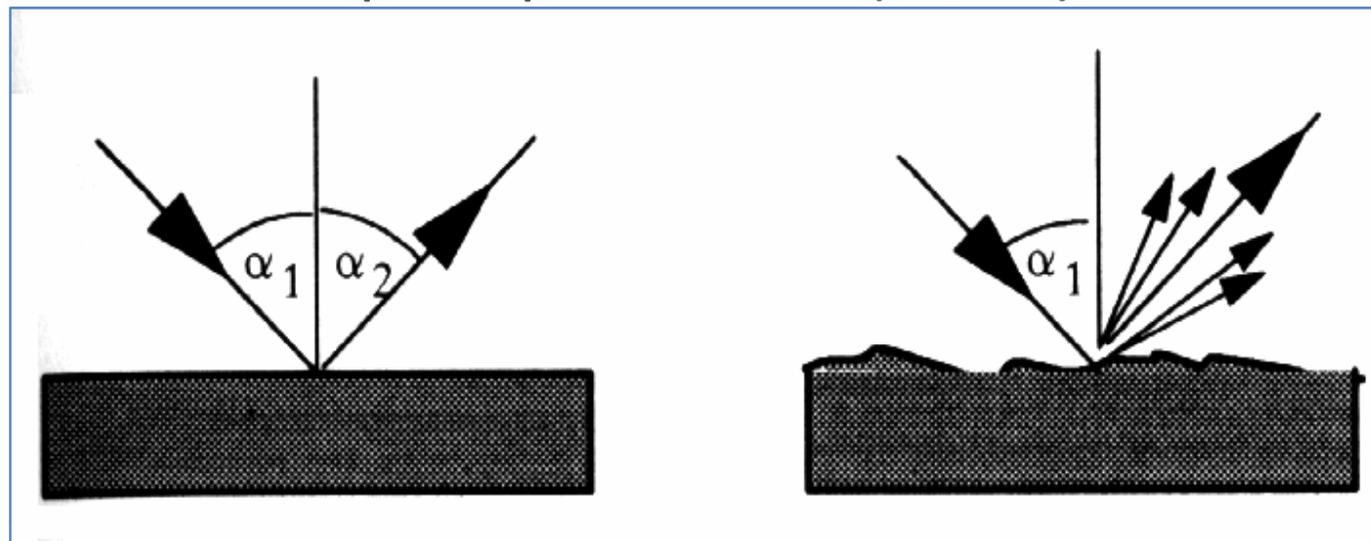
Image Generation

Radiometry



The radiometric relation between the world and its projection is formed by:

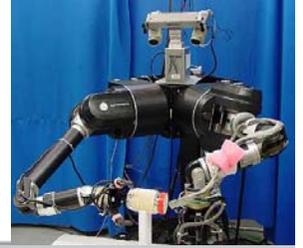
- Amount of light that is reflected by a surface point = **Radiance**
- Amount of light that is projected from this point onto the image = **Irradiance**
- measured in watts per square meter (W/m^2),



Smooth Surface

Rough Surface

Radiometric Resolution



- Number of digital values (“gray levels”) that a sensor can use to express variability of signal (“brightness”) within the data
- Determines the information content of the image
- The more digital values, the more detail can be expressed
- Determined by the number of bits of within which the digital information is encoded

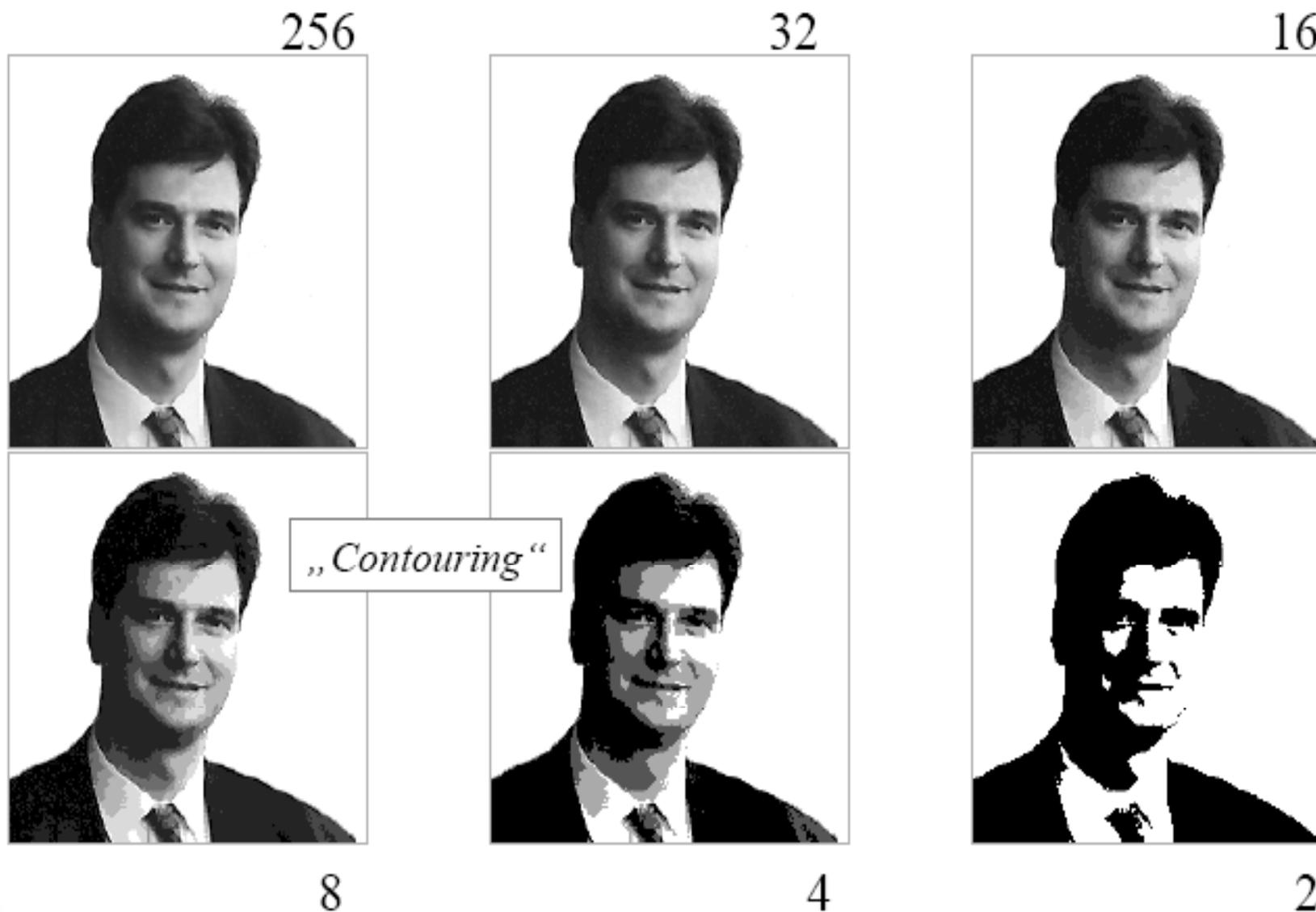
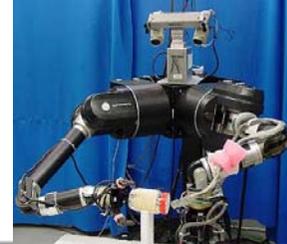
$$2^1 = 2 \text{ levels (0,1)}$$

$$2^2 = 4 \text{ levels (0,1,2,3)}$$

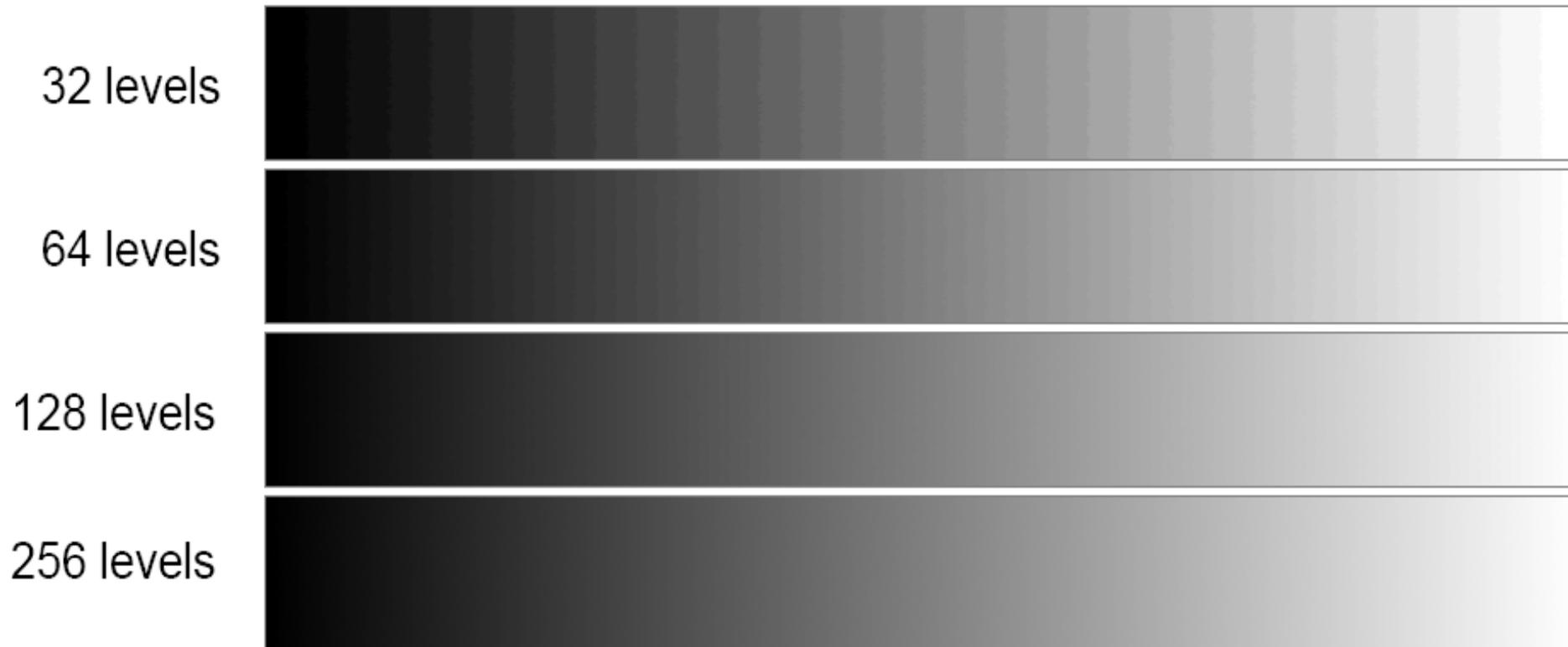
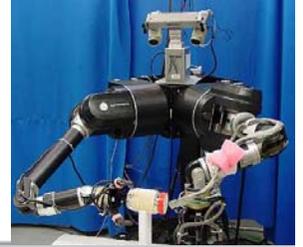
$$2^8 = 256 \text{ levels (0-255)}$$

$$2^{12} = 4096 \text{ levels (0-4095)}$$

Different numbers of Gray Levels

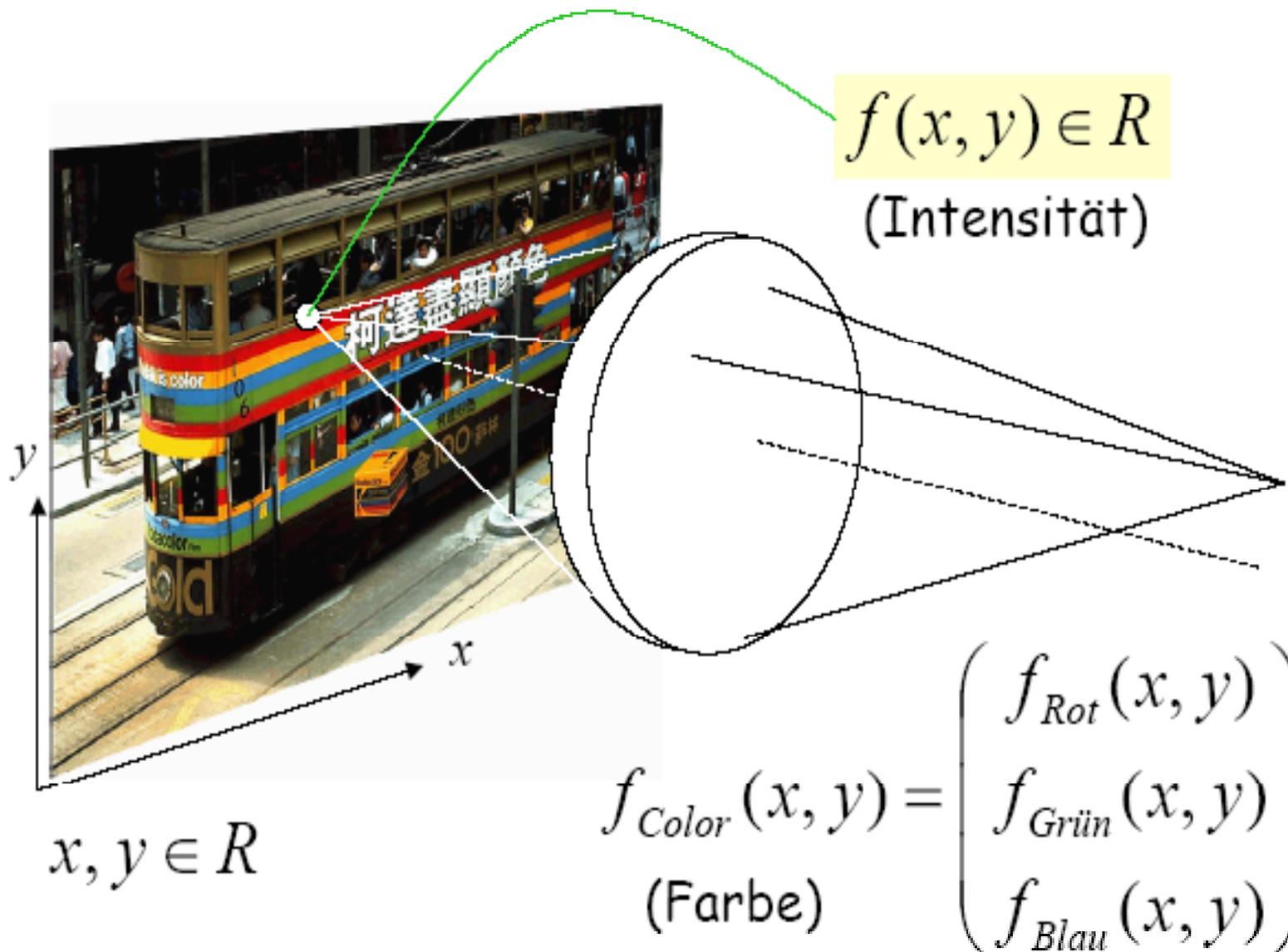
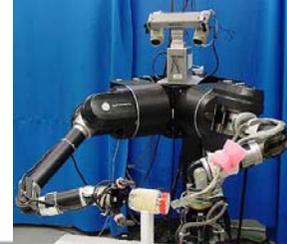


How many gray levels are required?

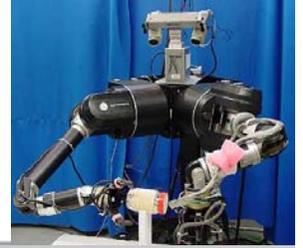


- Contouring is most visible for a ramp
- Digital images typically are quantized to 256 gray levels.

Continuous Image Function



Transition to a Digital Image



$f(x, y):$

$$R \times R \rightarrow R$$



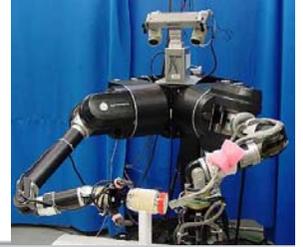
50	23	7	9	
19	8	4		
6	10			



$g(u, v):$

$$Z \times Z \rightarrow Z$$

Transition to a Digital Image - 1

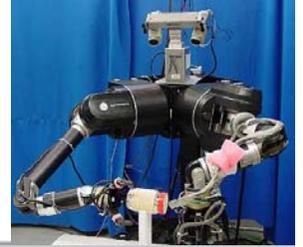


1. Räumliche Abtastung
(Sampling)

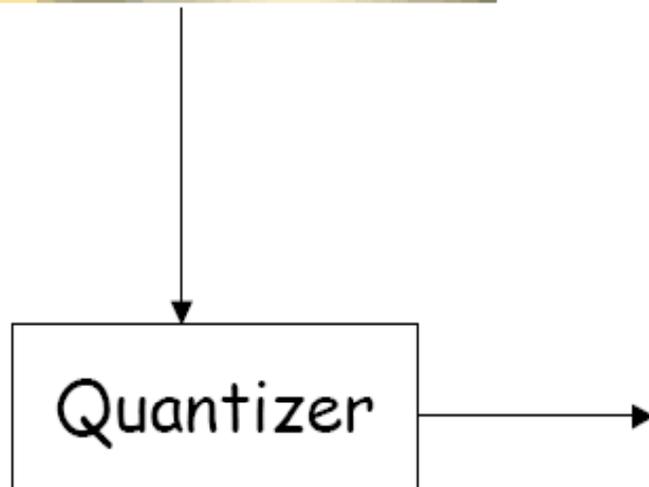
Sampler



Transition to a Digital Image - 2



2. Diskretisierung
der Bildwerte
(Quantisierung)



50	23	7	9	
19	8	4		
6	10			

Spatial Sampling

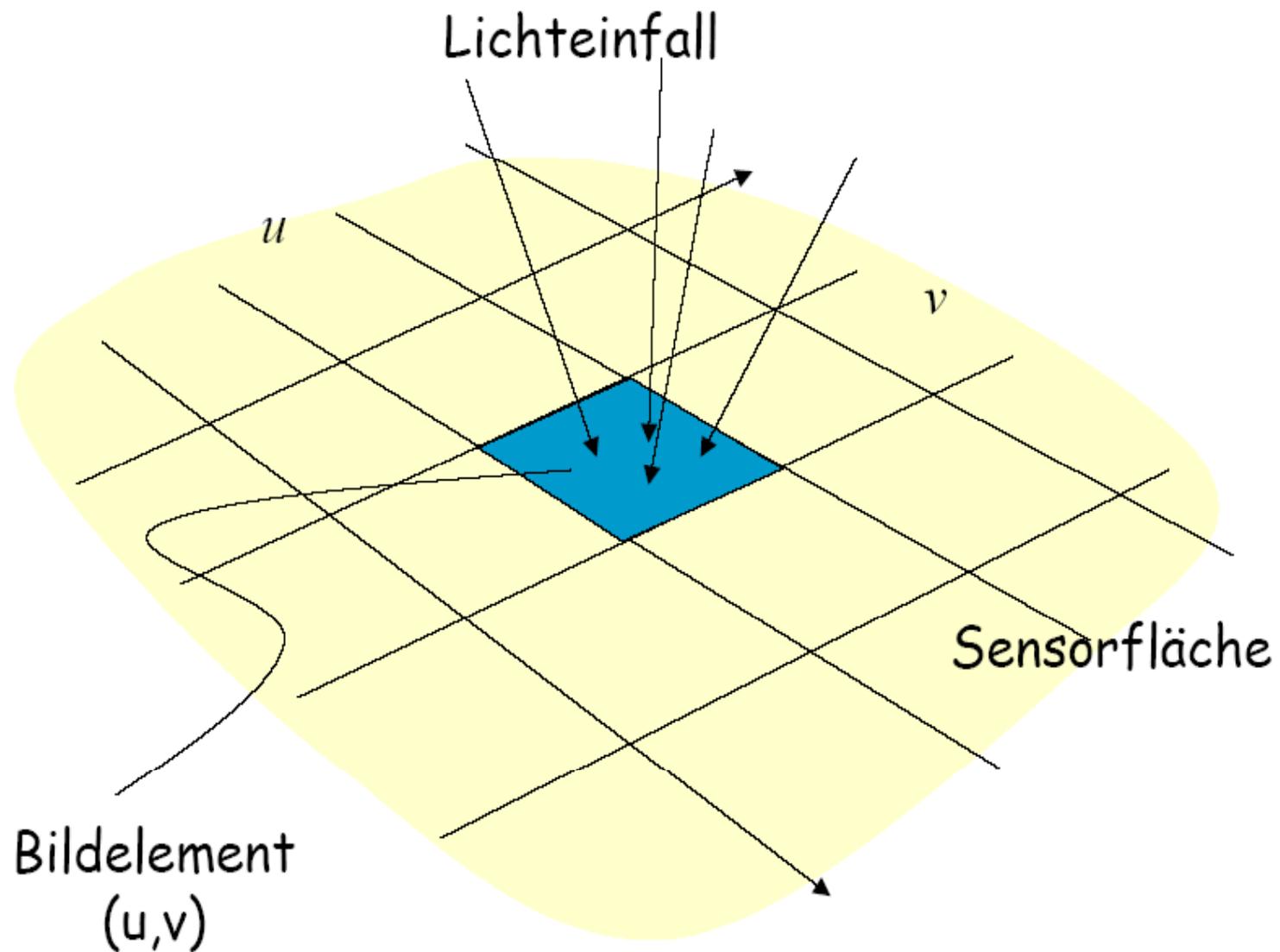
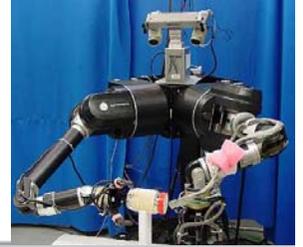
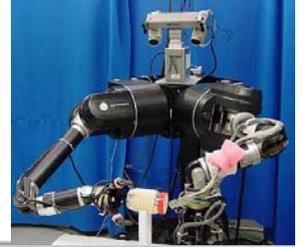


Image Size and Resolution



- These images were produced by simply picking every n-th sample horizontally and vertically and replicating that value $n \times n$ times:



200x200



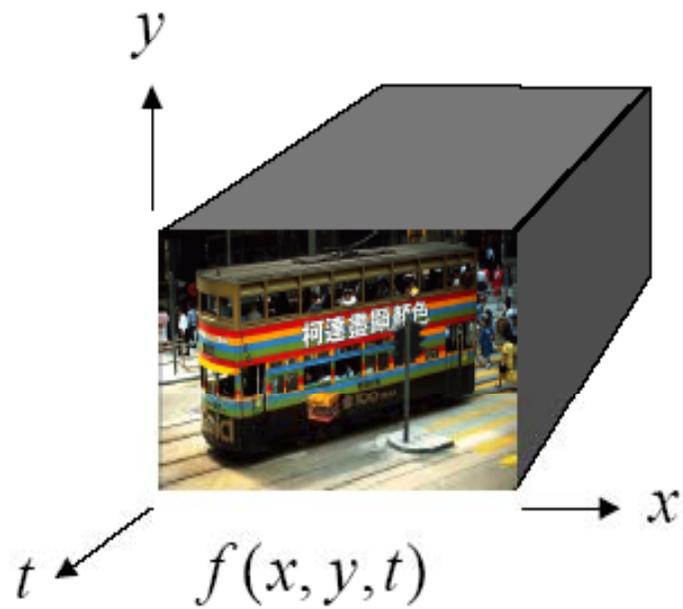
100x100



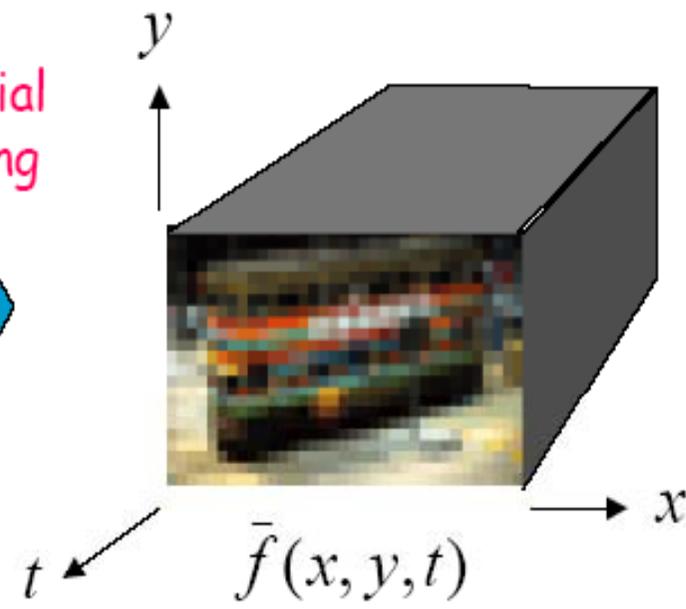
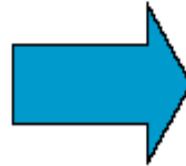
50x50



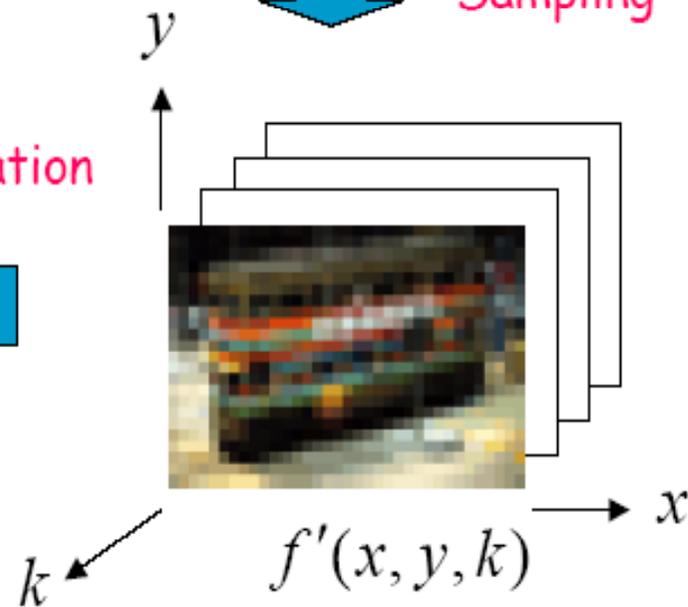
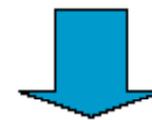
25x25



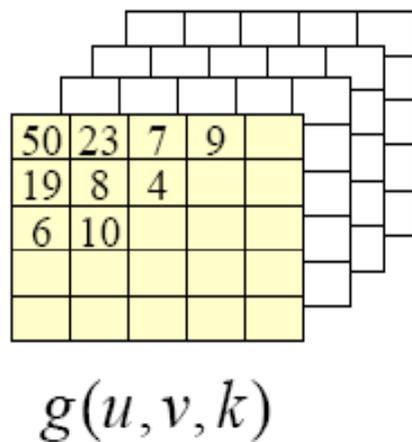
1. Spatial Sampling



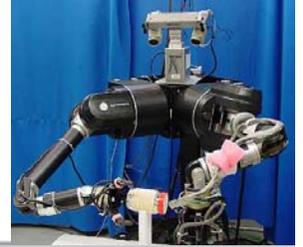
2. Temporal Sampling



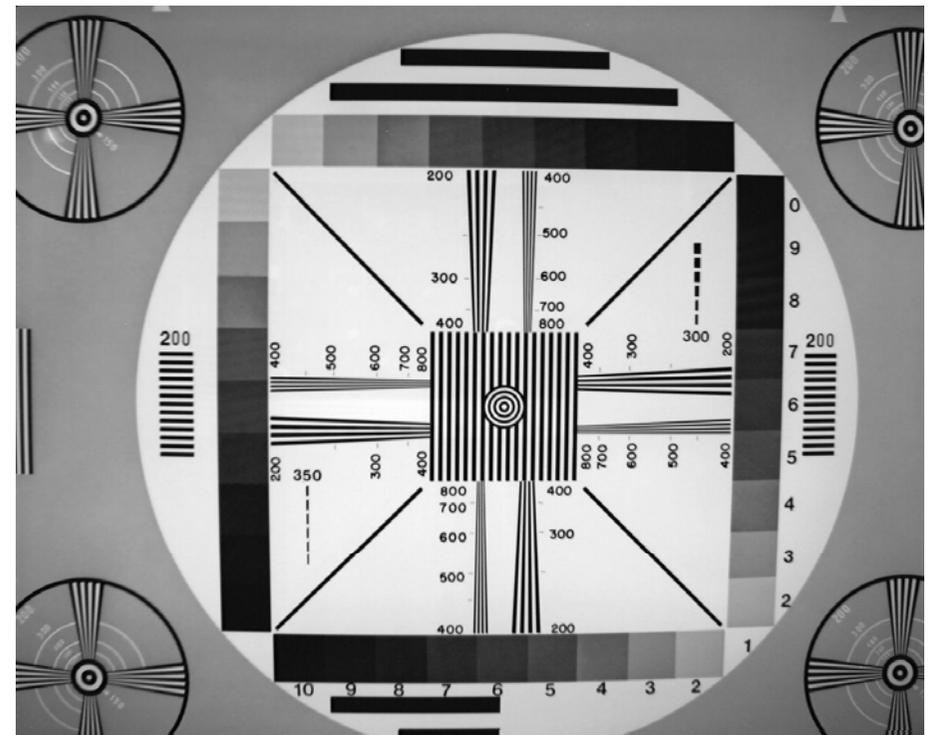
3. Quantization



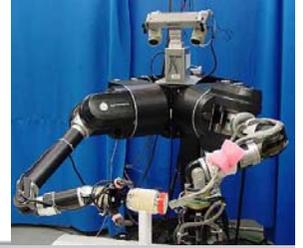
Spatial “Resolution” of a Sensor System



- **Spatial density** = Number of sensor elements (horizontal/vertical)
- **Optical resolution** = Quality of the optical system
- **Spatial resolution** = Relation pixel – object size (ppi)
- **Effective resolution:** Spatial density + Optical resolution (spatial resolution)



Signal generation



Transformation of the optical image into an „electrical“:

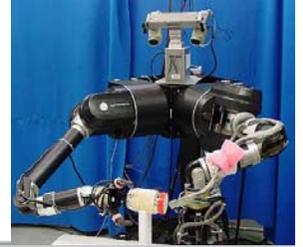
- Dependent on the wavelength:

$$E(x, y) = \iint Irrad(x, y, t, \lambda) s(\lambda) \tau(t - t_0) d\lambda dt$$

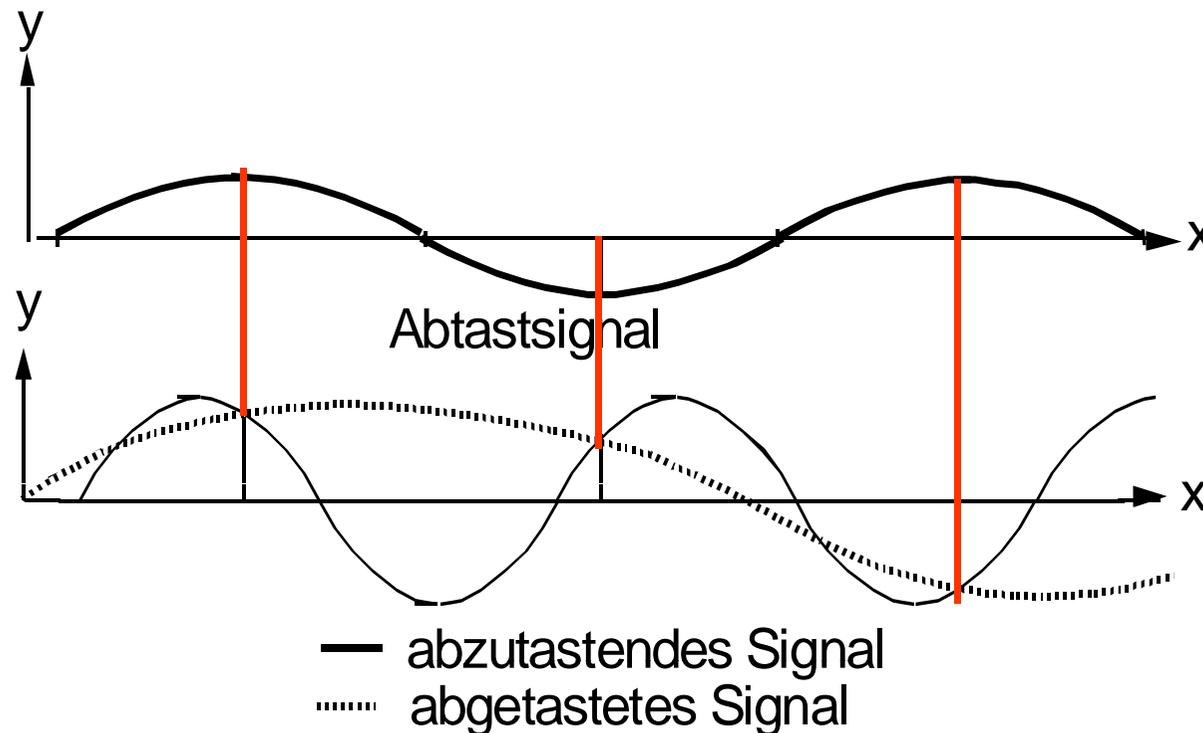
$s(\lambda)$ = Spectral sensitivity of the sensor

$\tau(t)$ = timeframe of the acquisition at $t = 0$

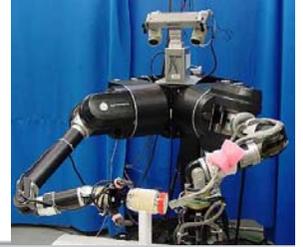
Sampling Theorem



Shannon Theorem: Exact reconstruction of a continuous-time baseband signal from its samples is possible if the signal is bandlimited and the **sampling frequency** is **greater than twice the signal bandwidth**.



Sampling Theorem

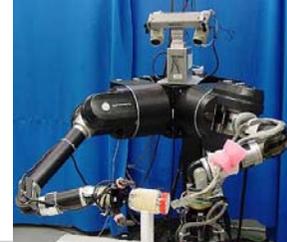


- Is always valid!
 - Space
 - Greylevel / Color
 - Time

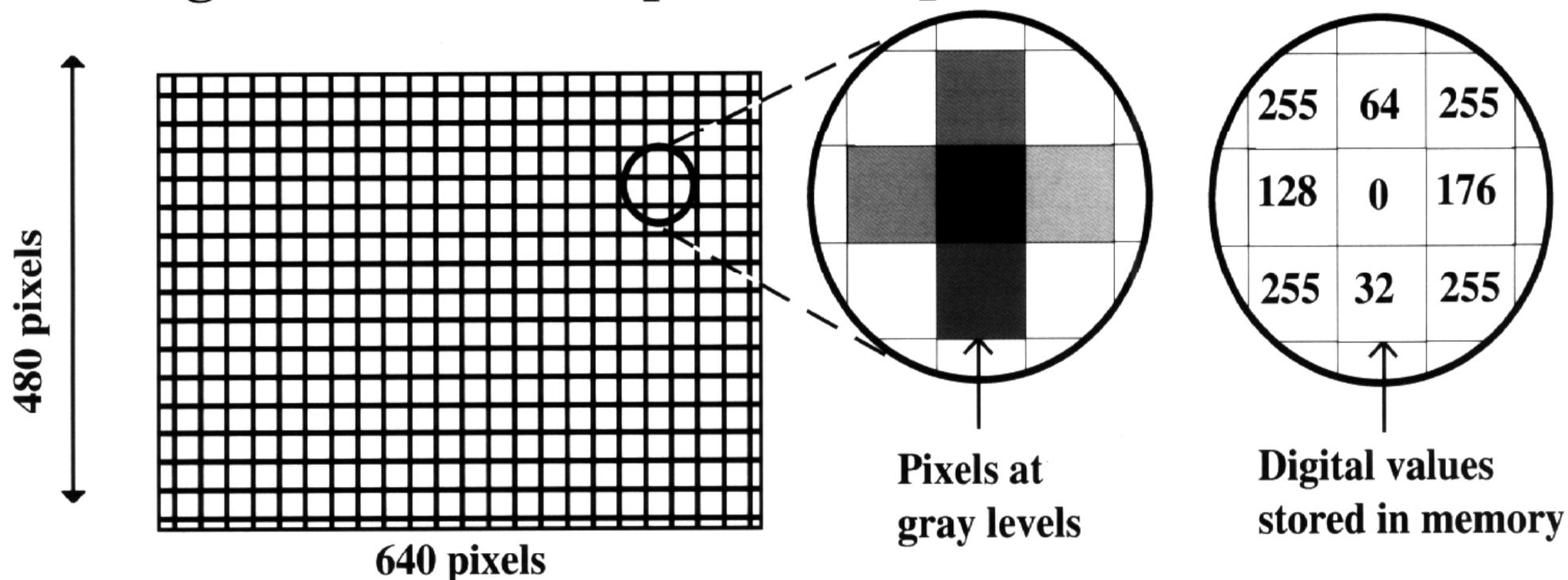


Sensors

Image Representation



Light intensities on photo-receptors



Cameras



- Mainly used class: **CCD Cameras**
- CCD = **C**harge **C**oupled **D**evice
- **MOS Transistors**, that charge themselves while light is projected

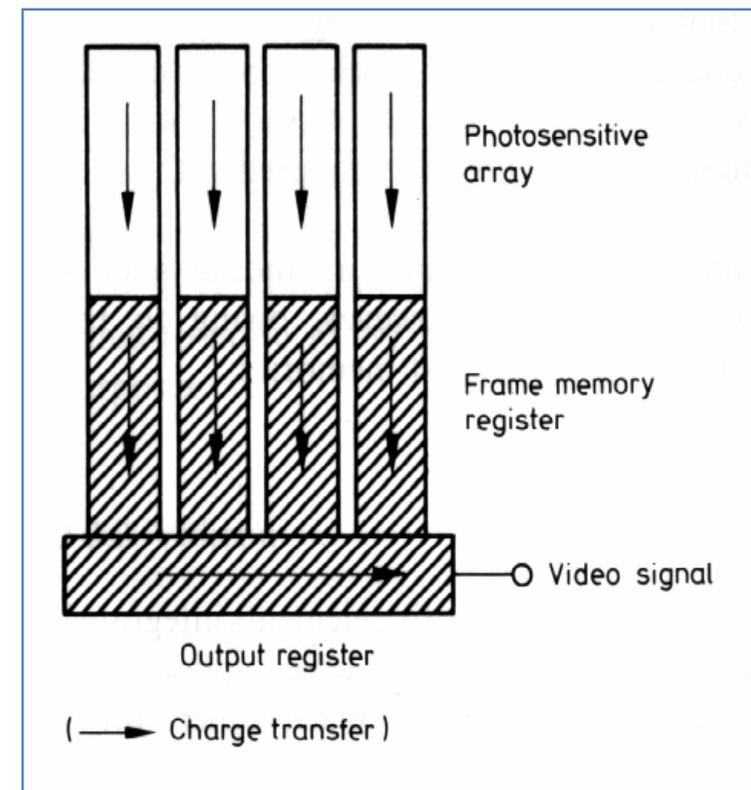
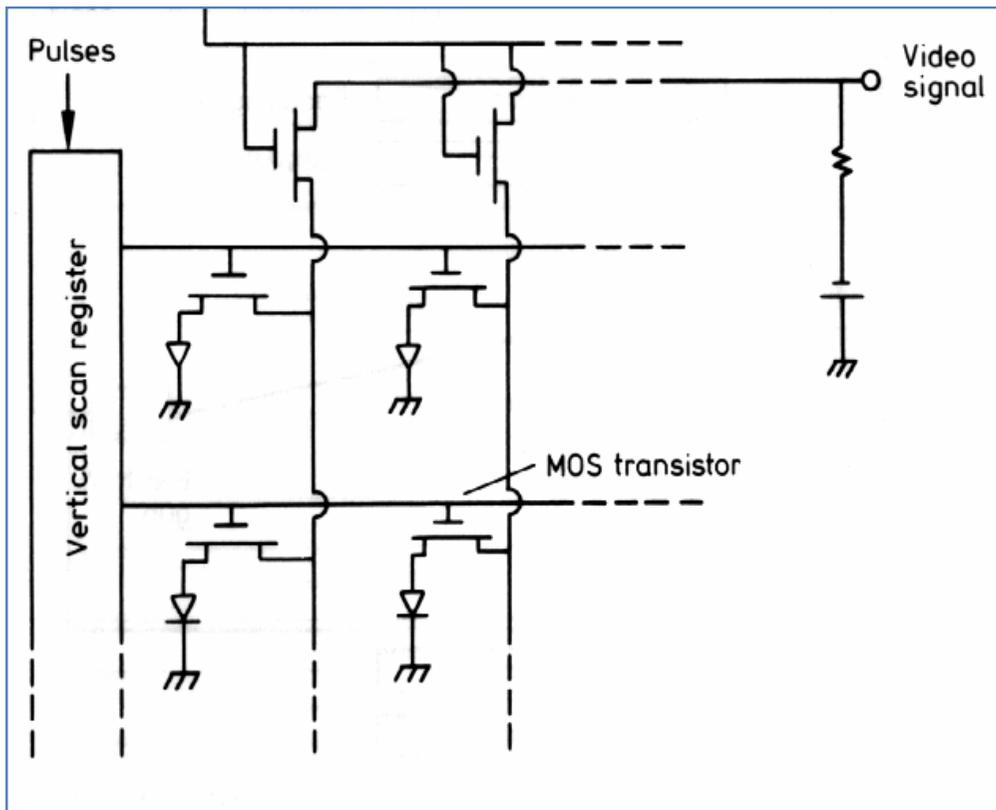
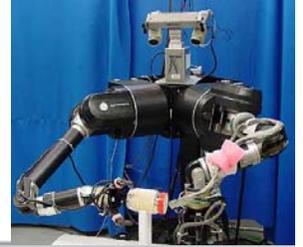


Image Scanning



- Path from the transistor of the camera to the A/D converter
- Image is read out line by line

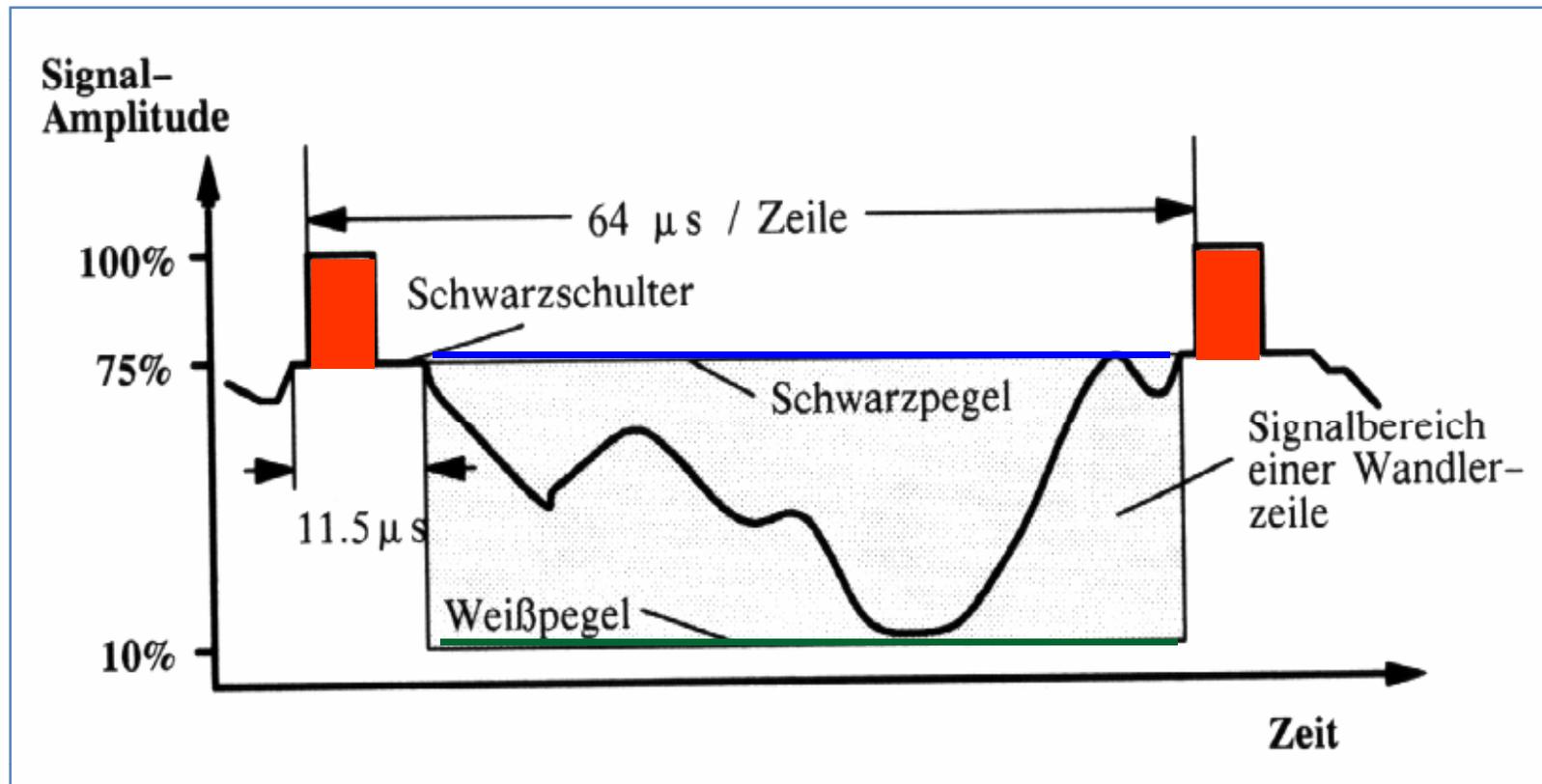


Image Scanning

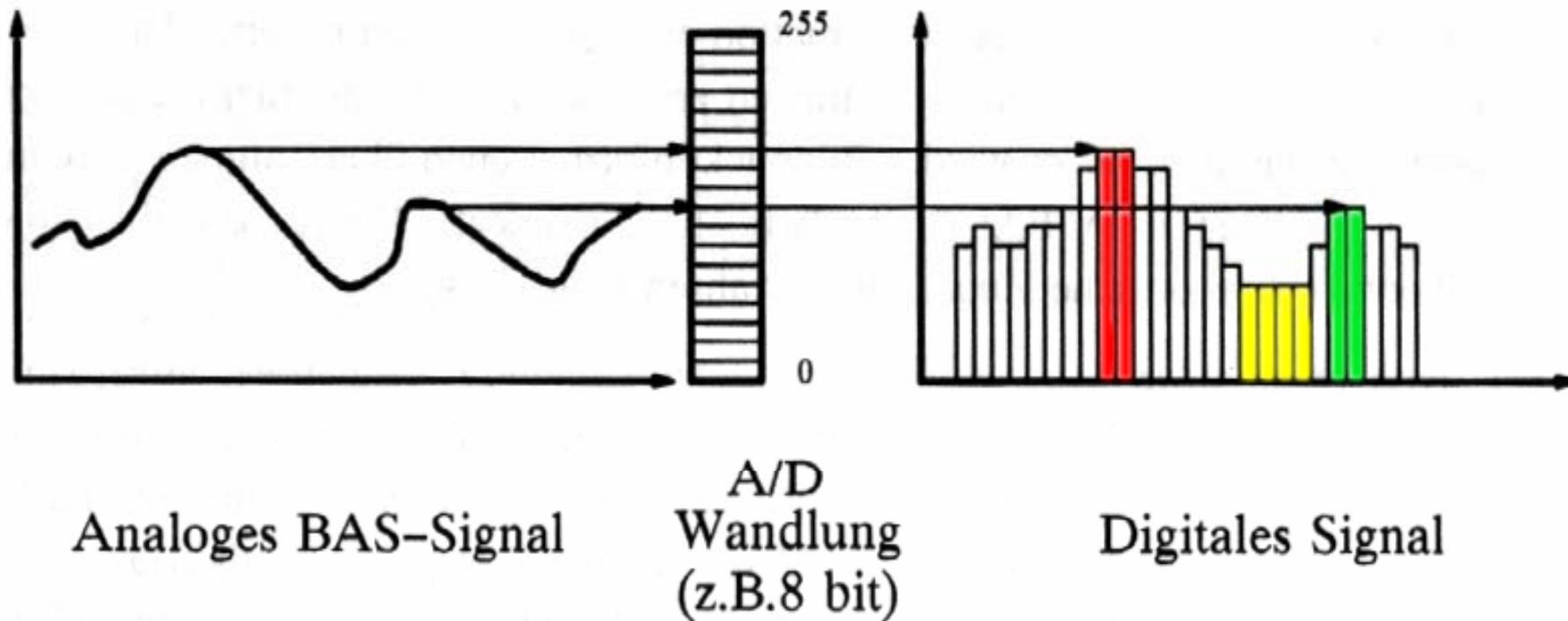
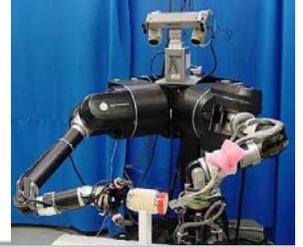
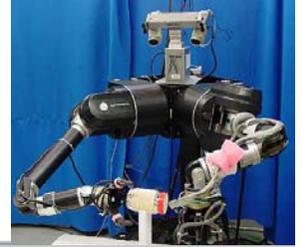
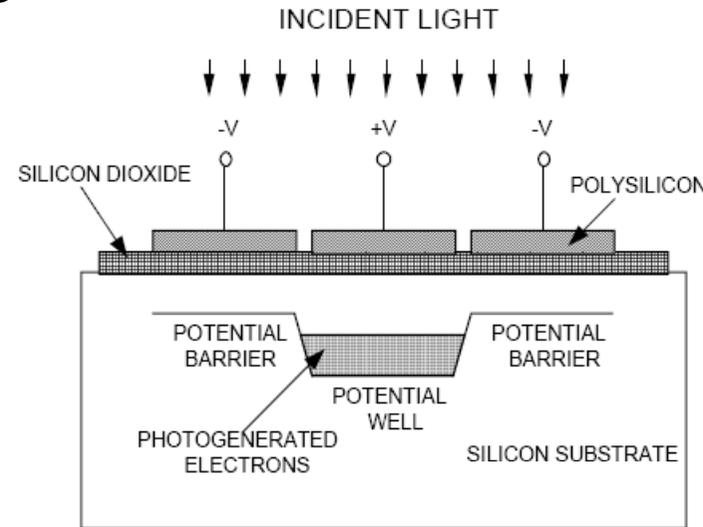
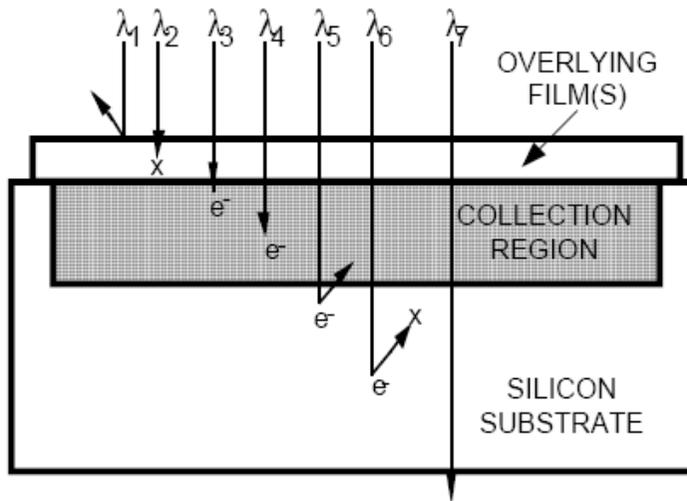


Image Sensors

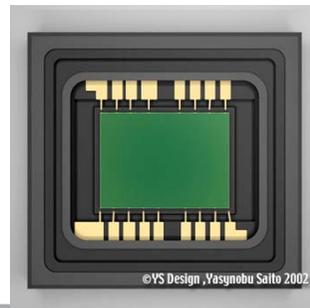


- Convert light into electric charge



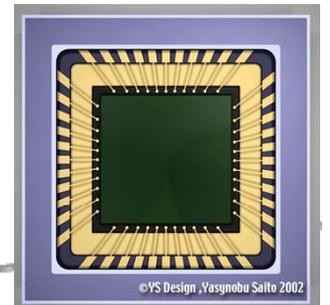
- CCD (charge coupled device)

Higher dynamic range
High uniformity
Lower noise

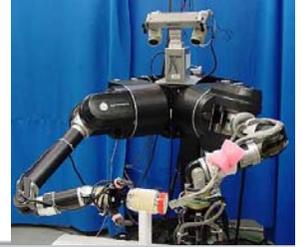


- CMOS (complementary metal Oxide semiconductor)

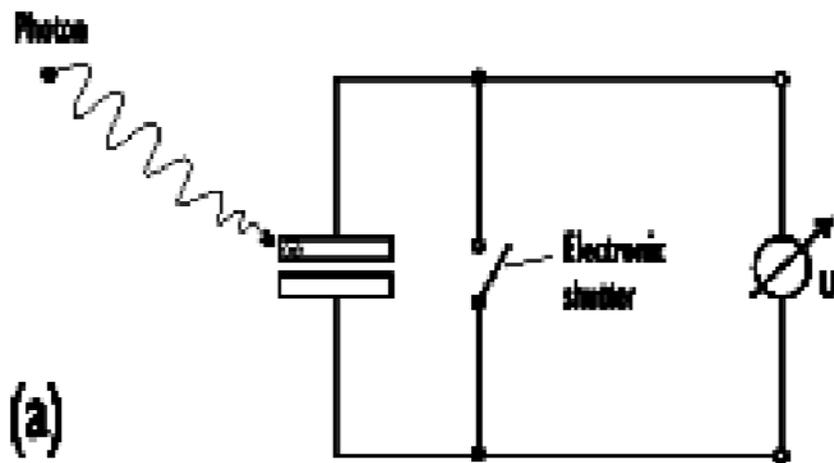
Lower voltage
Higher speed
Lower system complexity



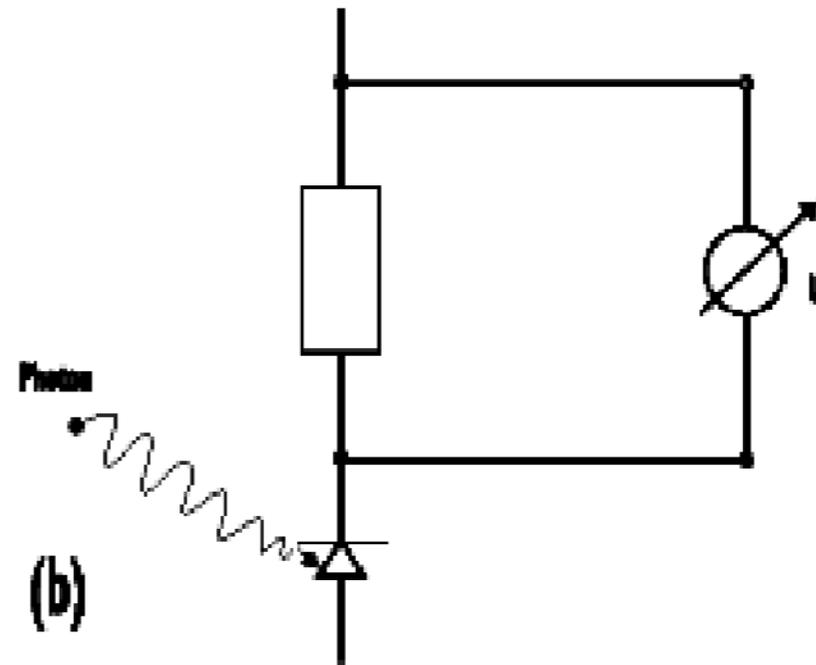
Charging Models



- CCD Cameras work integrative
- CMOS Cameras non-linear

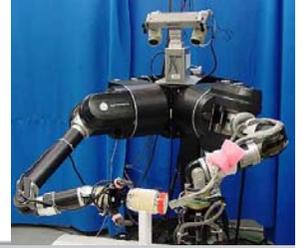


Integrative/linear



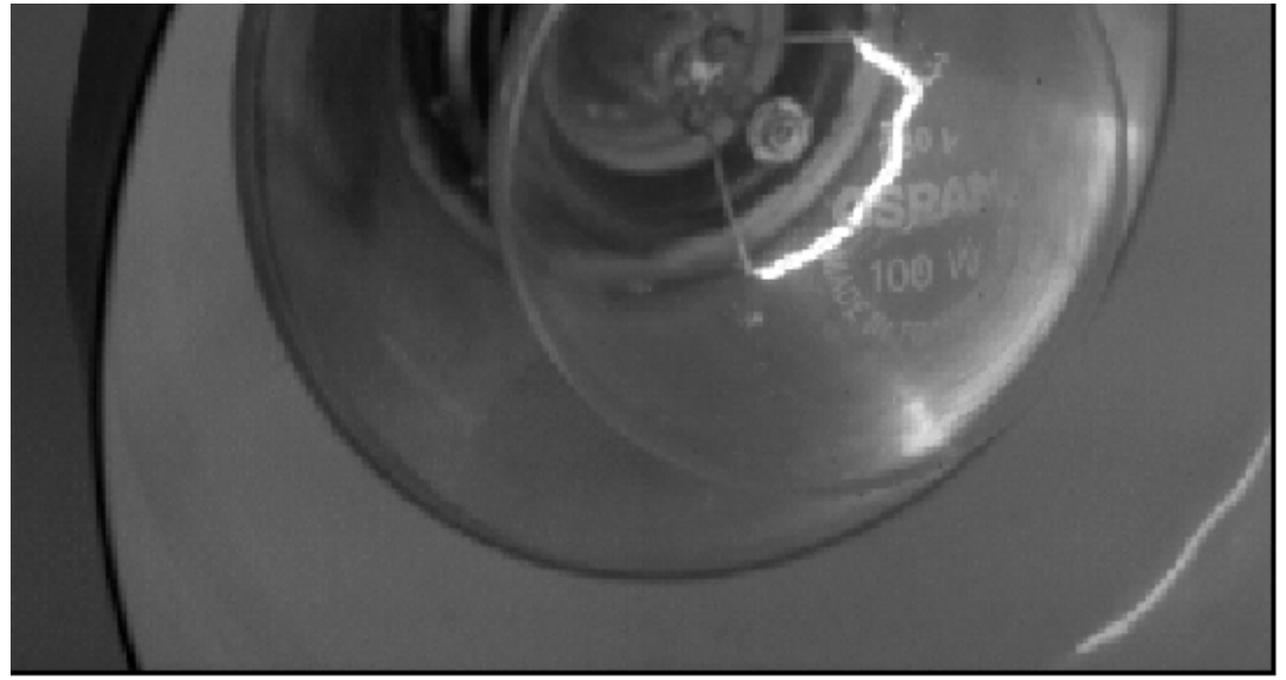
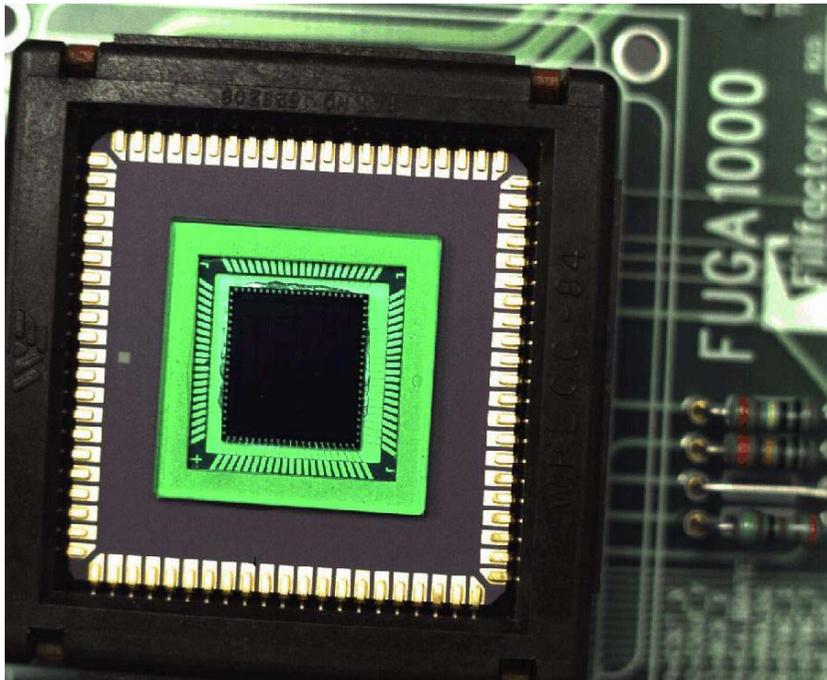
Non-linear

Example

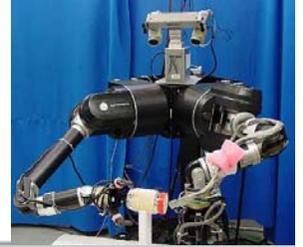


Fuga 1000 Sensor

Image: log.Char.!!



Pros and Cons



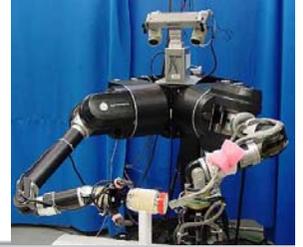
CCD

- Needs extra circuitry to convert to digital signal
- **High dynamic** range of lighting
- **Less noise** due to less on-chip circuitry
- Integrative

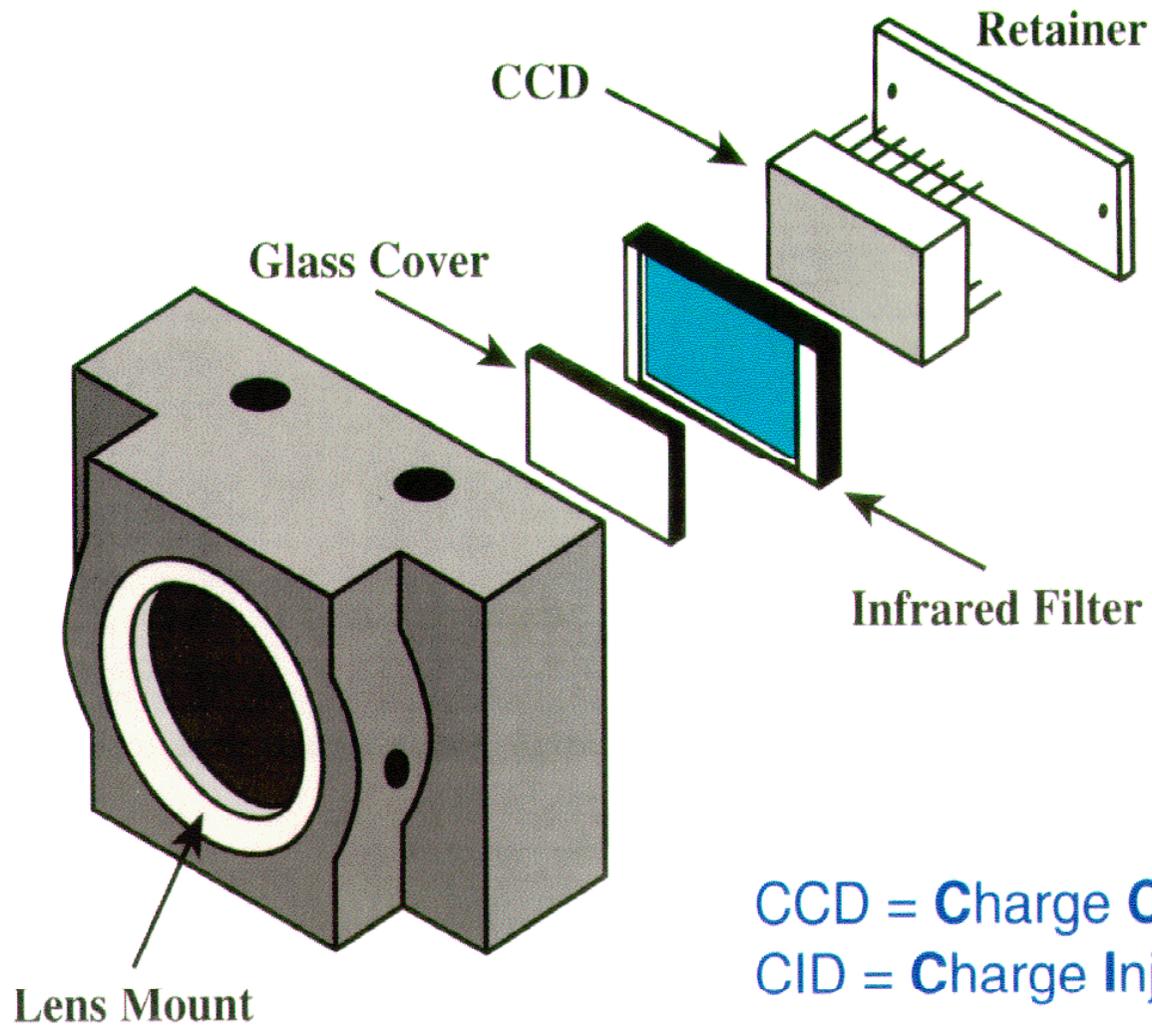
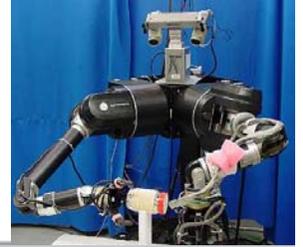
CMOS

- Higher cost to develop
- On-chip analog-to-digital conversion
- Lower complexity on the sensor leading to **faster image capture**
- Reduced power consumption
- Non Linear

Types of Cameras

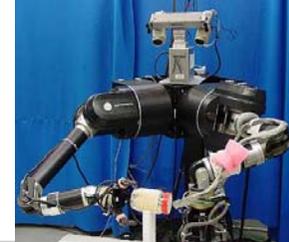


Camera Principles

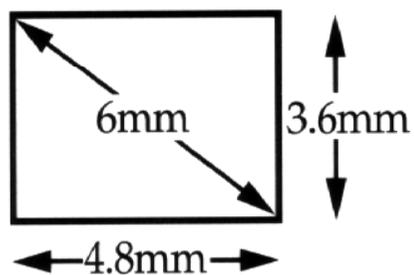


CCD = **C**harge **C**oupled **D**evice
CID = **C**harge **I**njection **D**evice

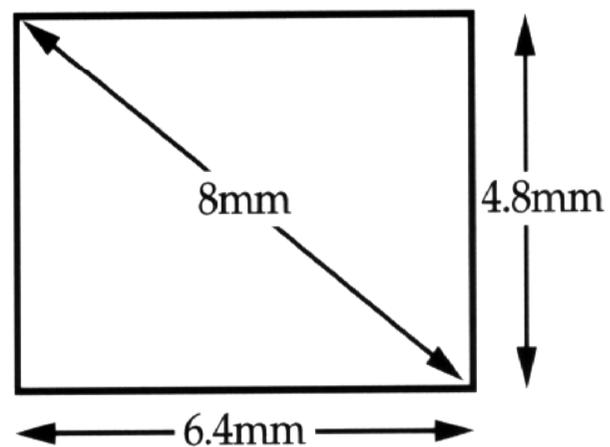
Common CCD Sizes



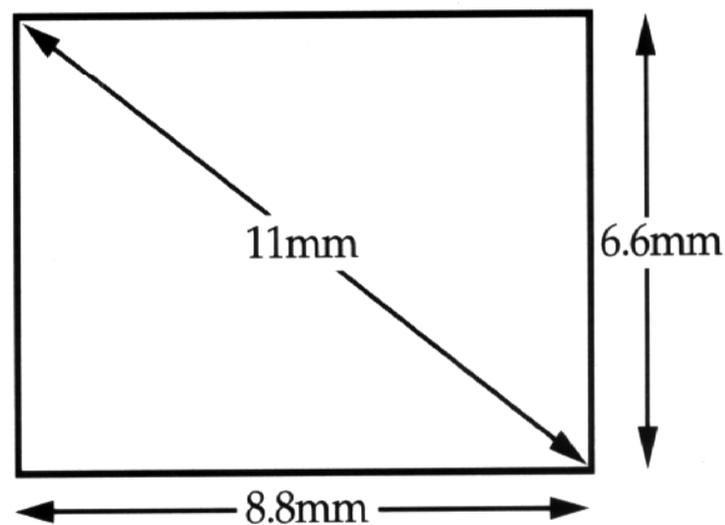
1/3" Format



1/2" Format

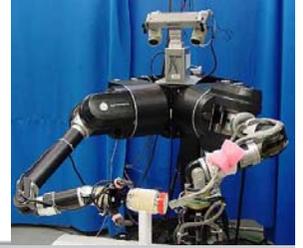


2/3" Format



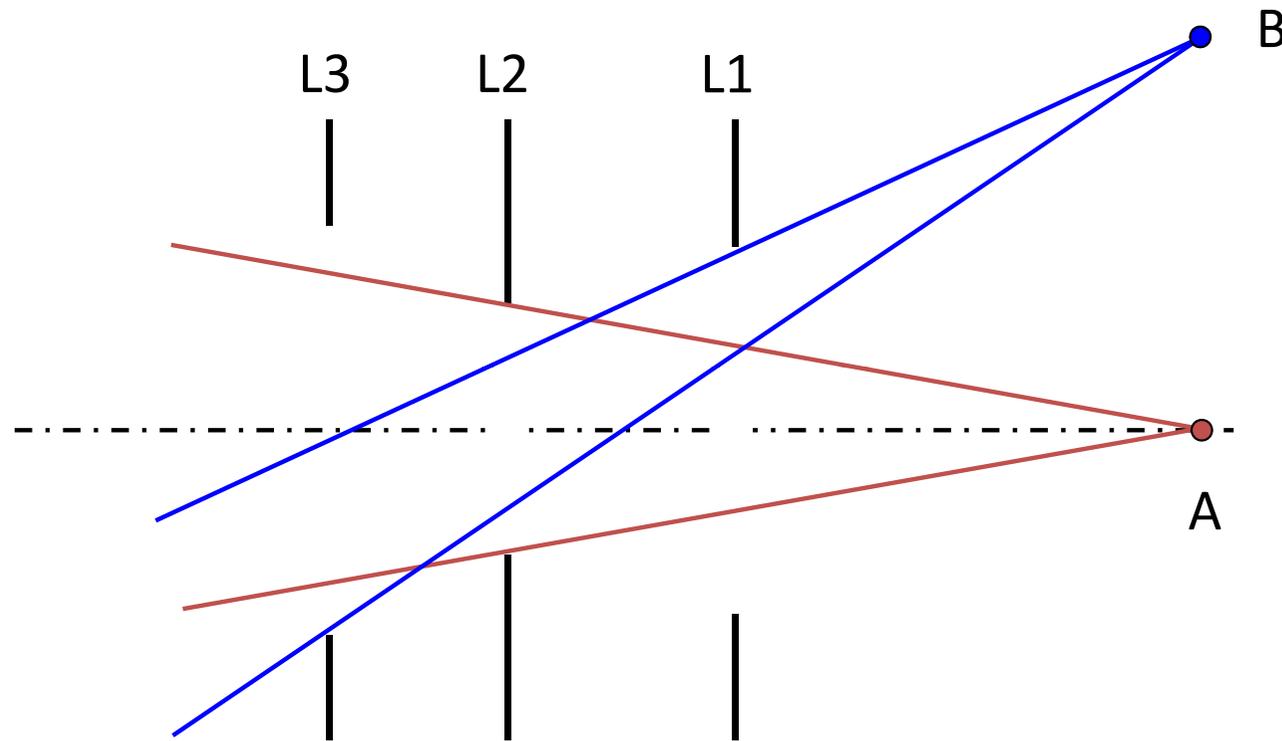
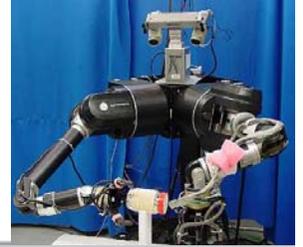
Problems with Real Cameras

Lens Glare



- Stray interreflections of light within the optical lens system.
- Happens when very bright sources are present in the scene.

Vignetting



- More light passes through lens L3 for scene point A than scene point B
- Results in spatially non-uniform brightness (in the periphery of the image)

Vignetting

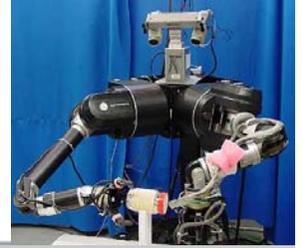
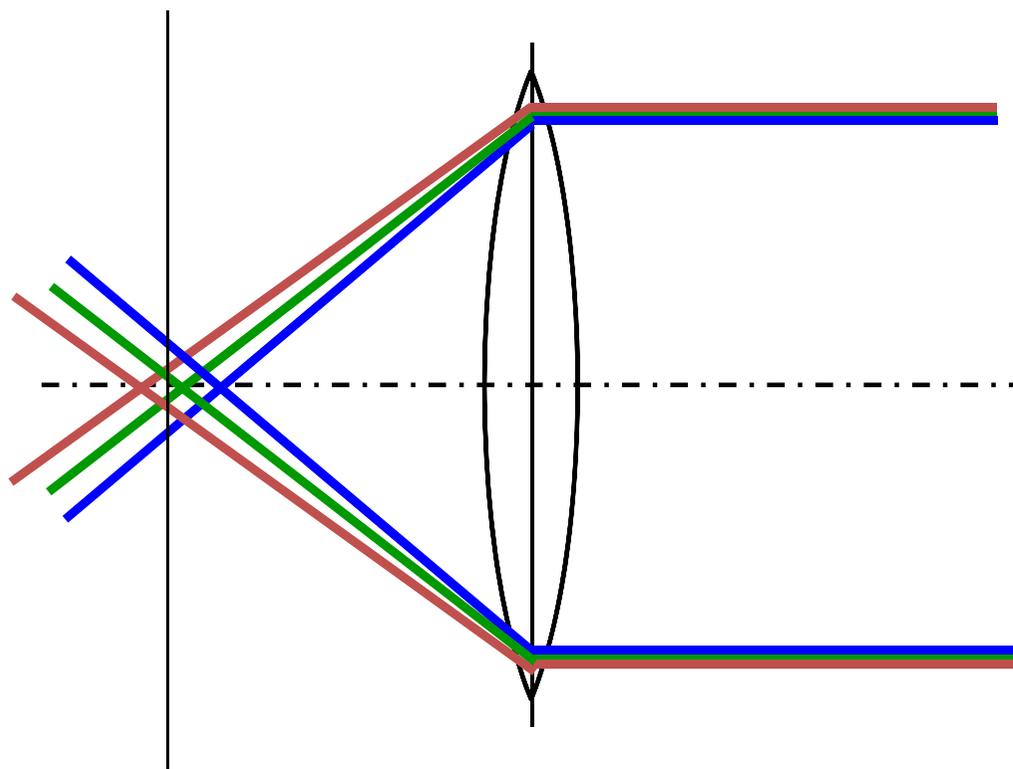
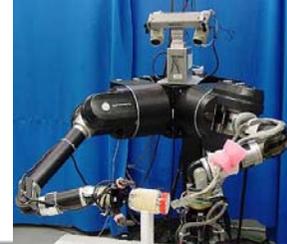
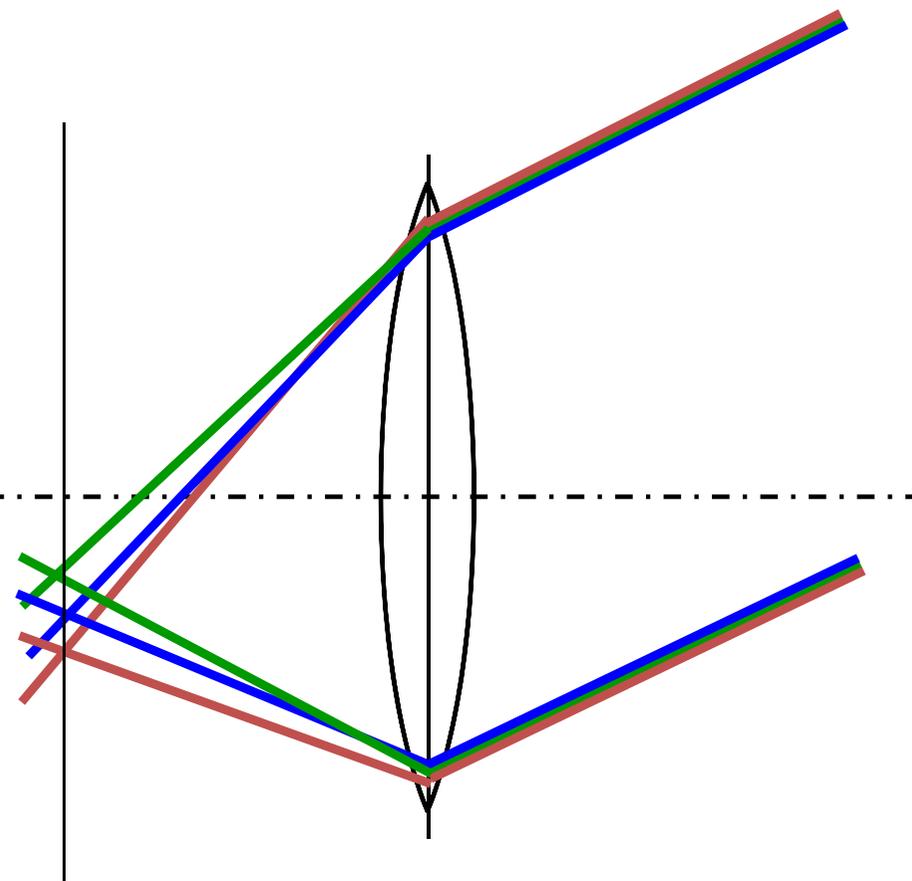


photo by Robert Johnes

Chromatic Aberration

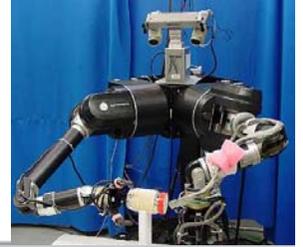


longitudinal chromatic aberration
(axial)

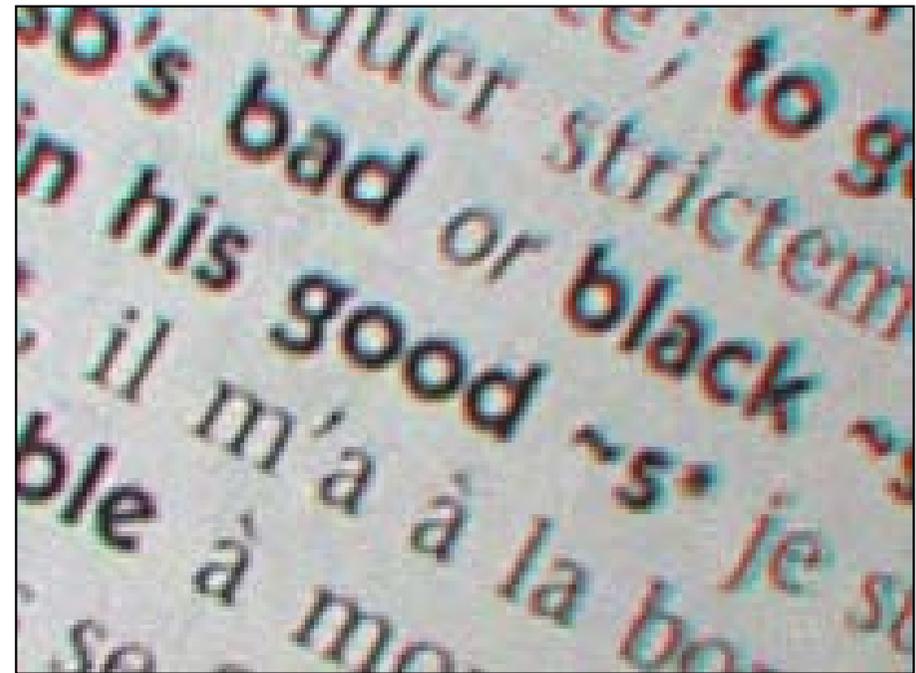


transverse chromatic aberration
(lateral)

Chromatic Aberrations

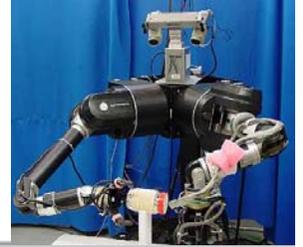


longitudinal chromatic aberration
(axial)

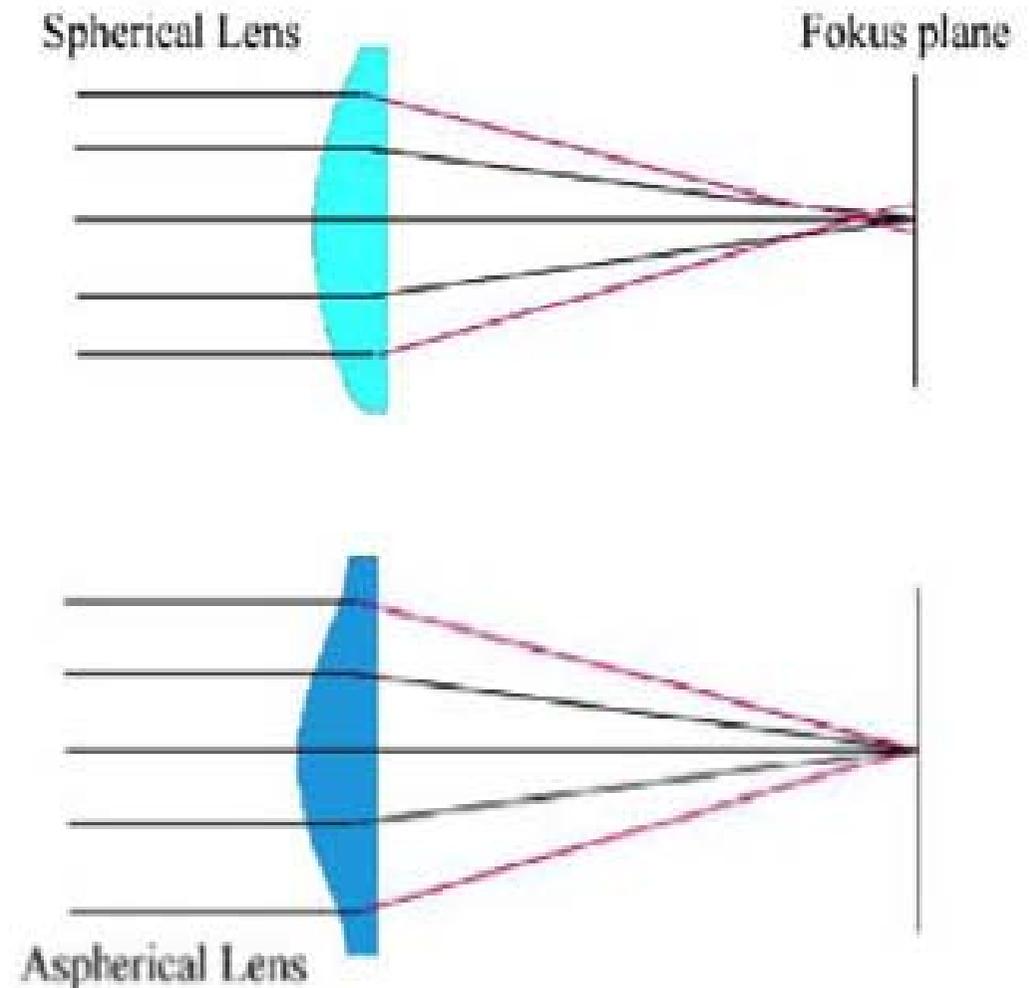


transverse chromatic aberration
(lateral)

Spherical Aberration



- Effect: sharp image superimposed by a blurred one
- Caused by spherical lens surfaces (manufacturing)
- Parallel rays are focused in one point only if they are close to the optical axis
- Can be avoided by using aspherical lenses with parabolic surfaces



Geometric Lens Distortions

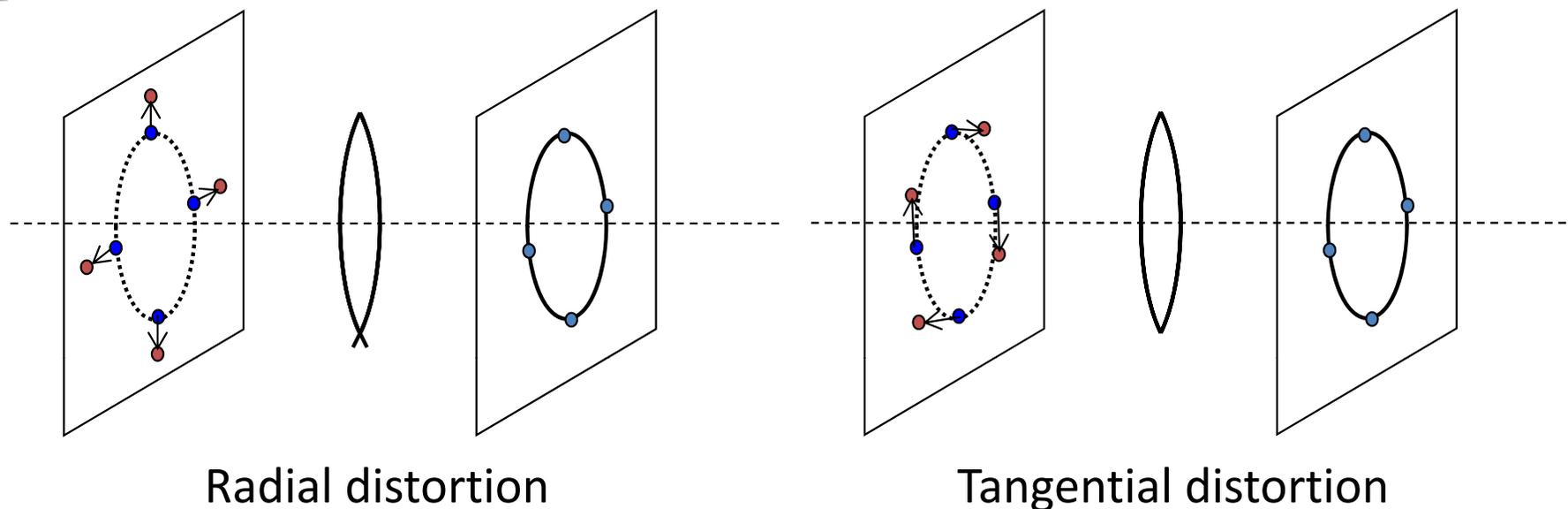
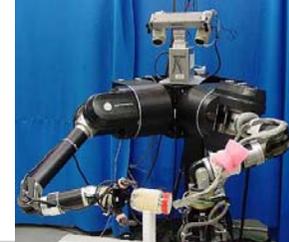
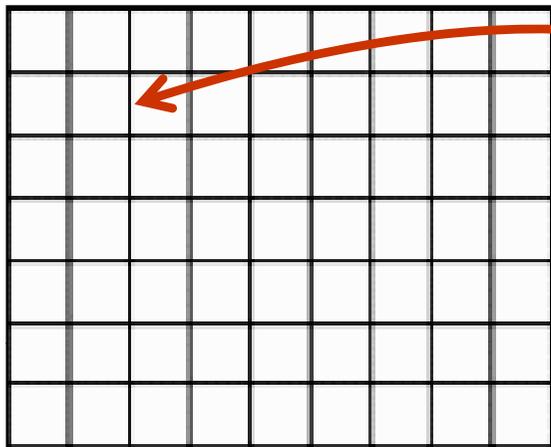
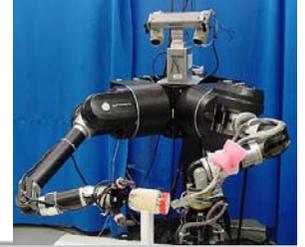


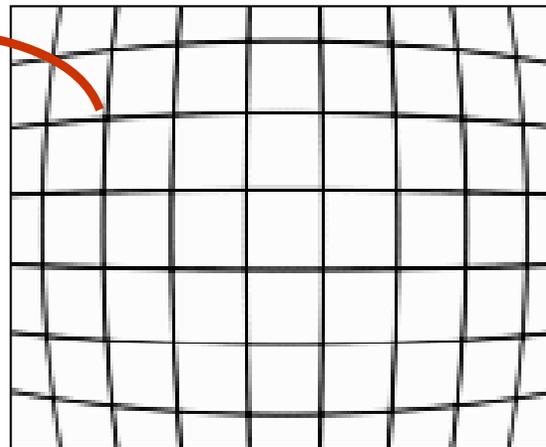
Photo by Helmut Dersch

Both due to lens imperfection

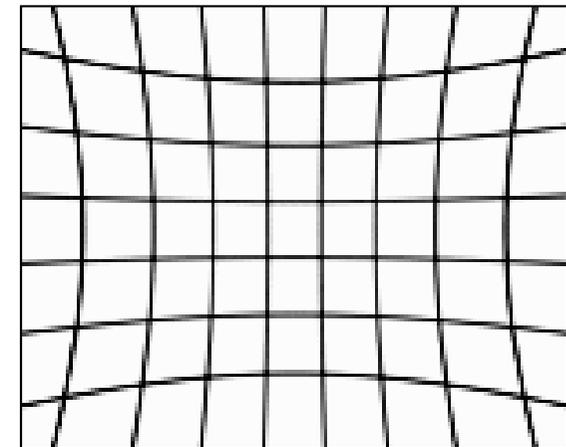
Radial Lens Distortions



No Distortion



Barrel Distortion



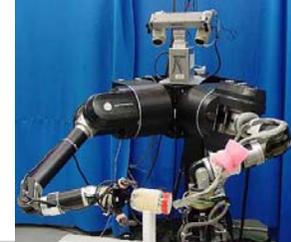
Pincushion Distortion

- Radial distance from Image Center:

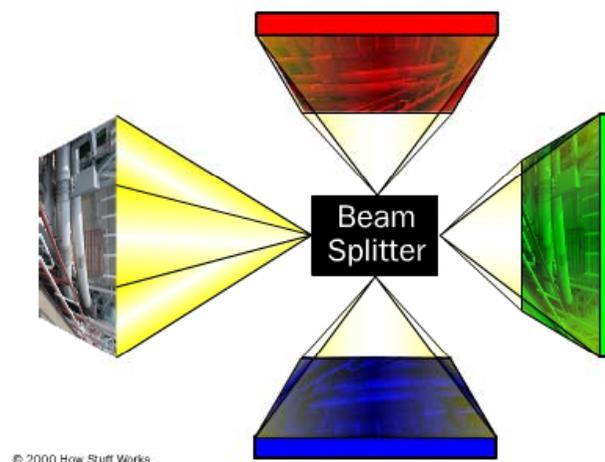
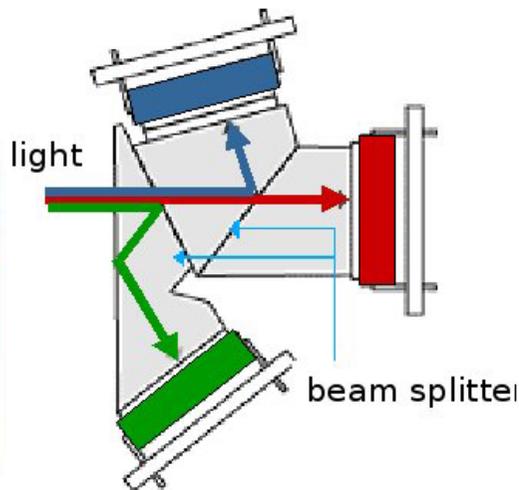
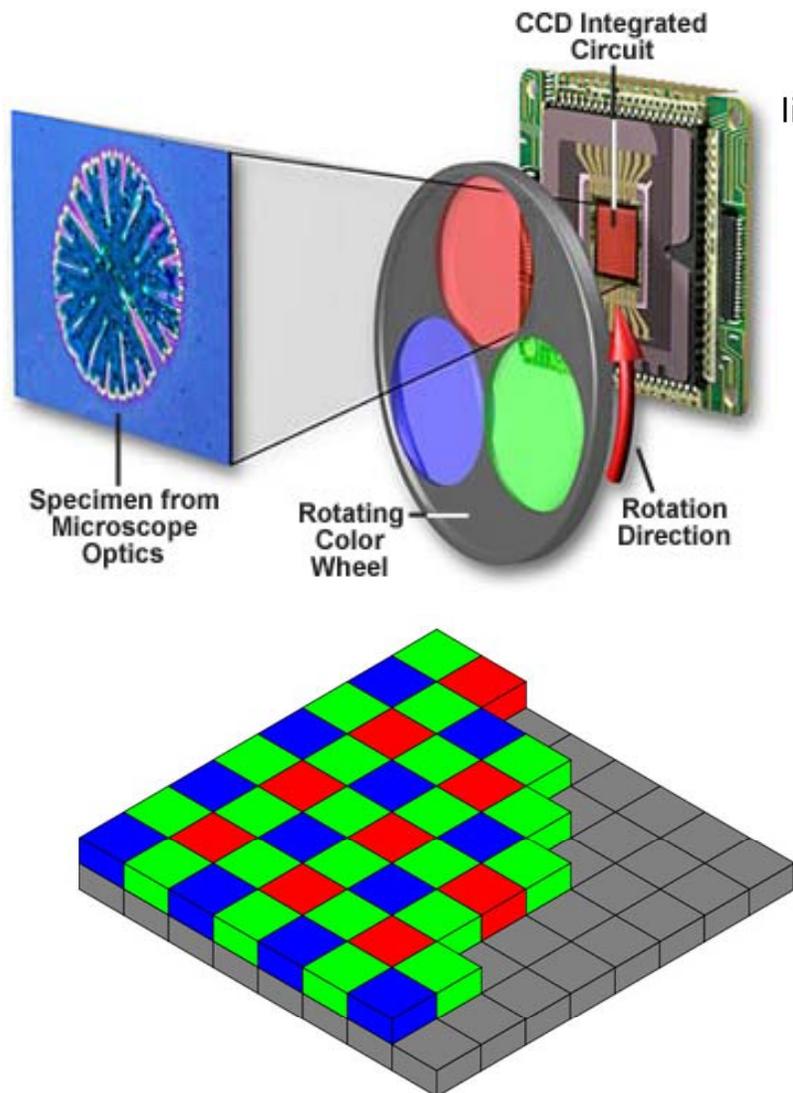
$$r_u = r_d + k_1 r_d^3$$

Color in Cameras

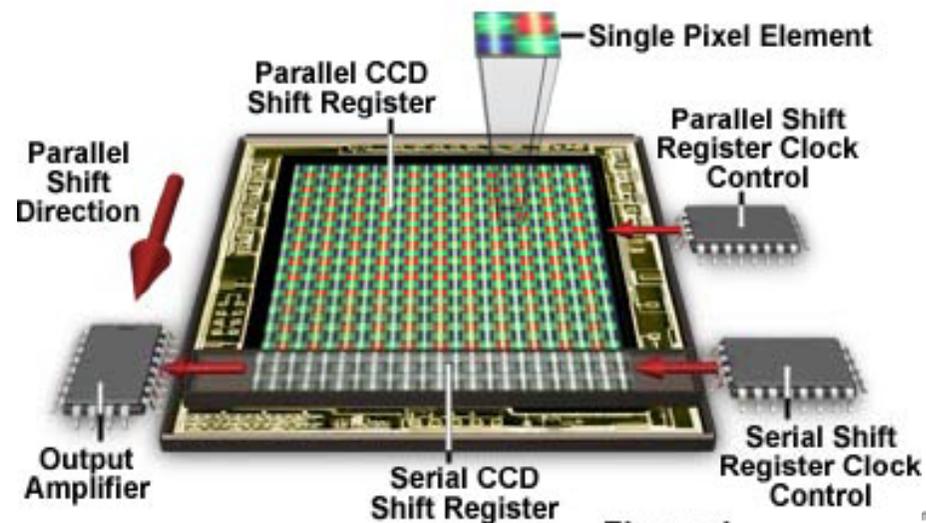
How CCDs Record Color



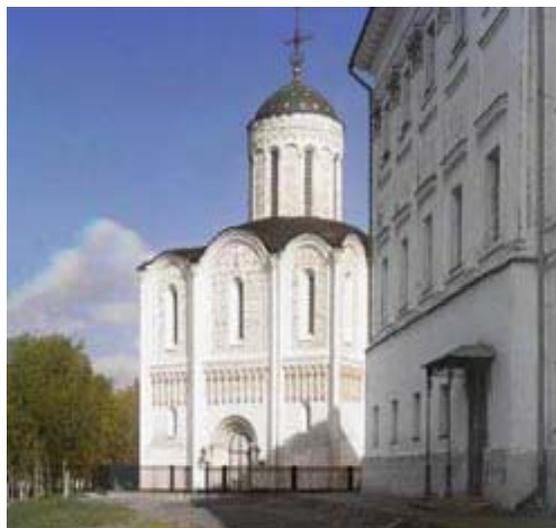
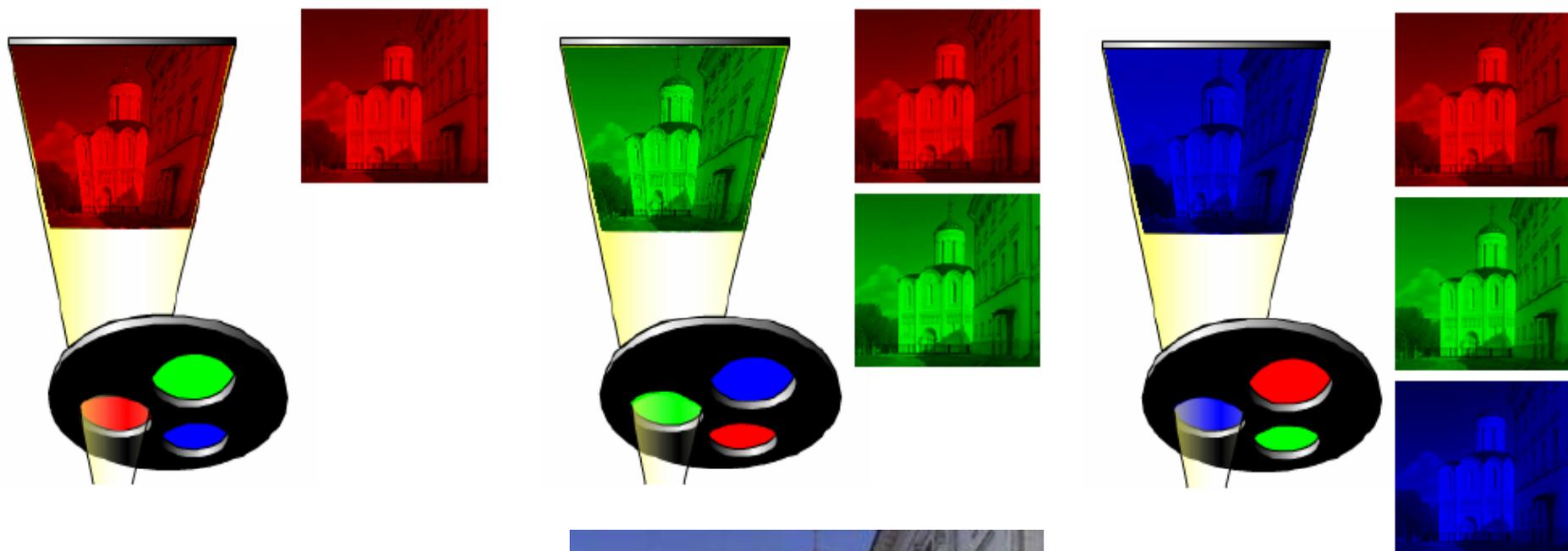
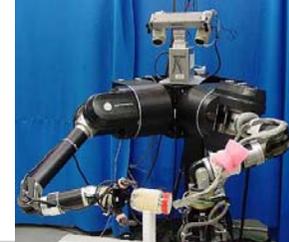
Sequential Color Three-Pass CCD Imaging System



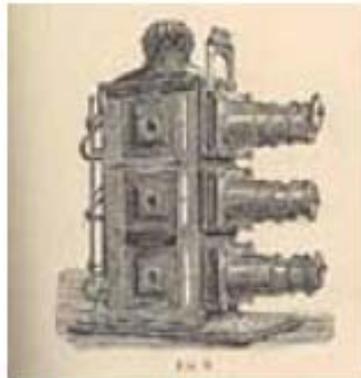
Full-Frame CCD Architecture



Field Sequential



Prokudin-Gorskii (early 1900's)

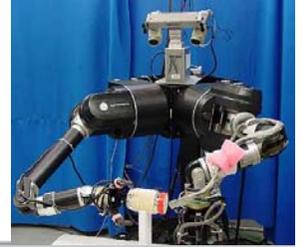


Lantern projector

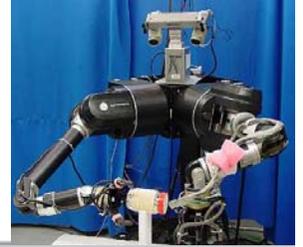


<http://www.loc.gov/exhibits/empire/>

Prokudin-Gorskii (early 1900's)



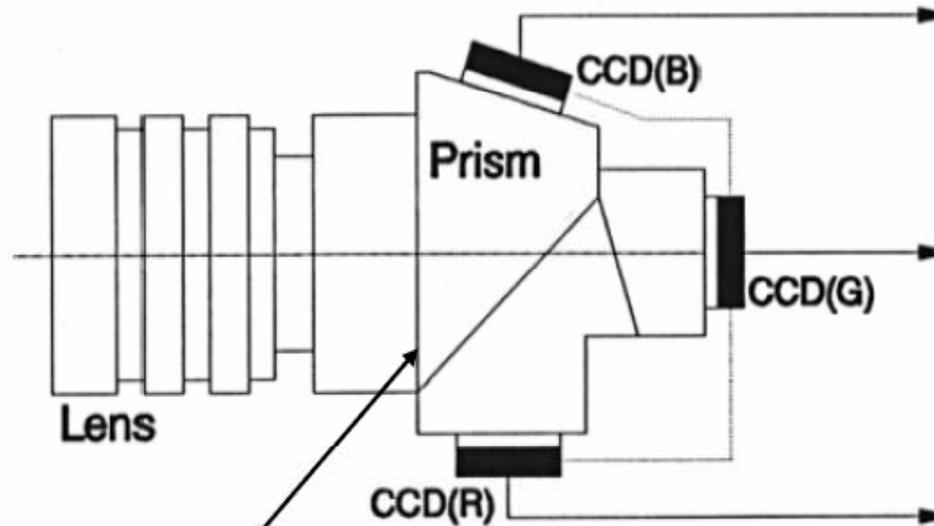
3-Chip Camera



- Accurate color per pixel
- Expensive
- 1/3 light per chip
- No DSLR cameras



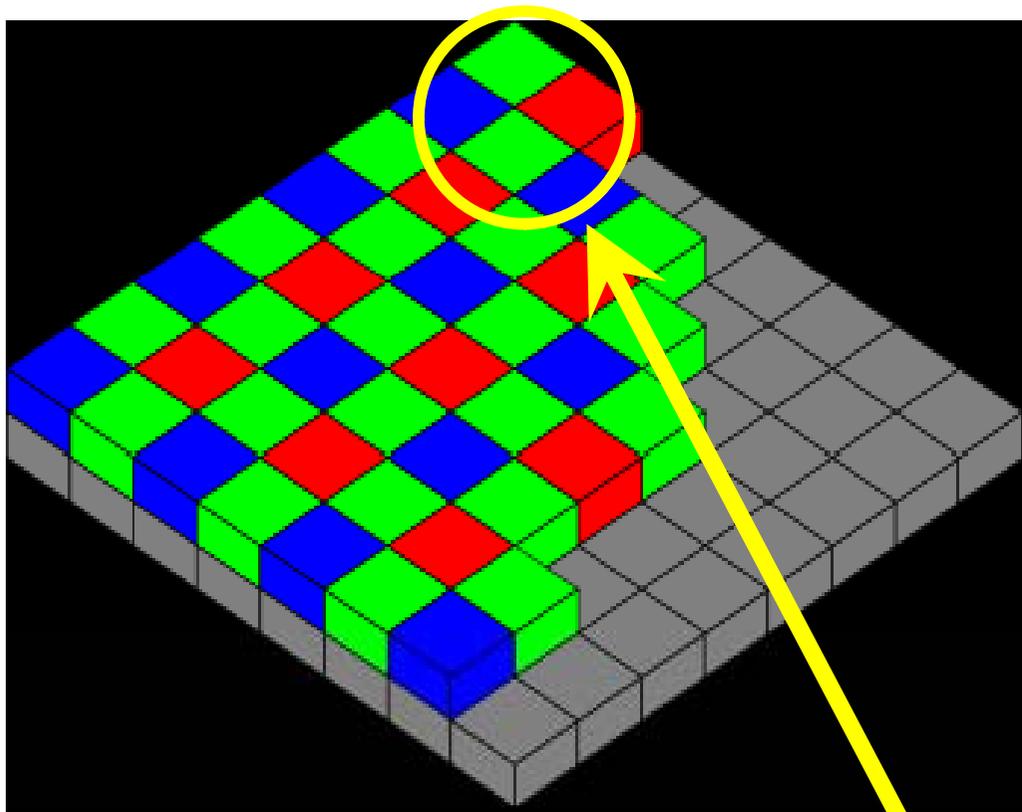
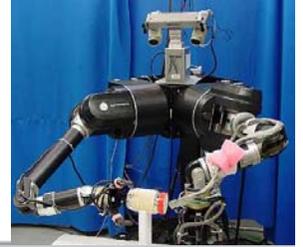
Sony DXC-D55PL



wavelength
dependent

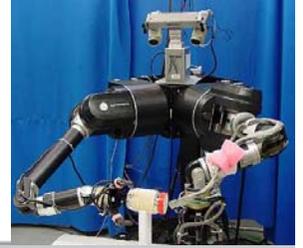


How CCDs Record Color

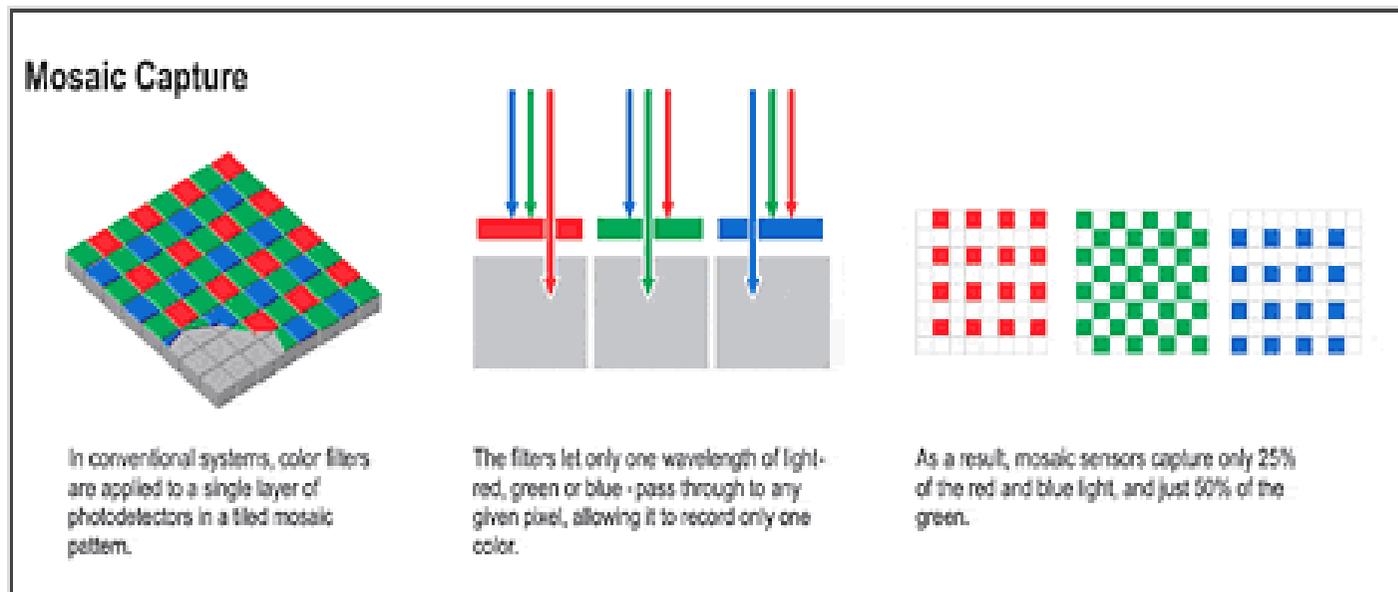


- Each CCD cell in CCD array produces *single value independent of color*.
- To make color images, CCD cells are organized in *groups of four cells* and color filters are placed on top of the group to allow red, blue and green light to hit *one of the four cells*.

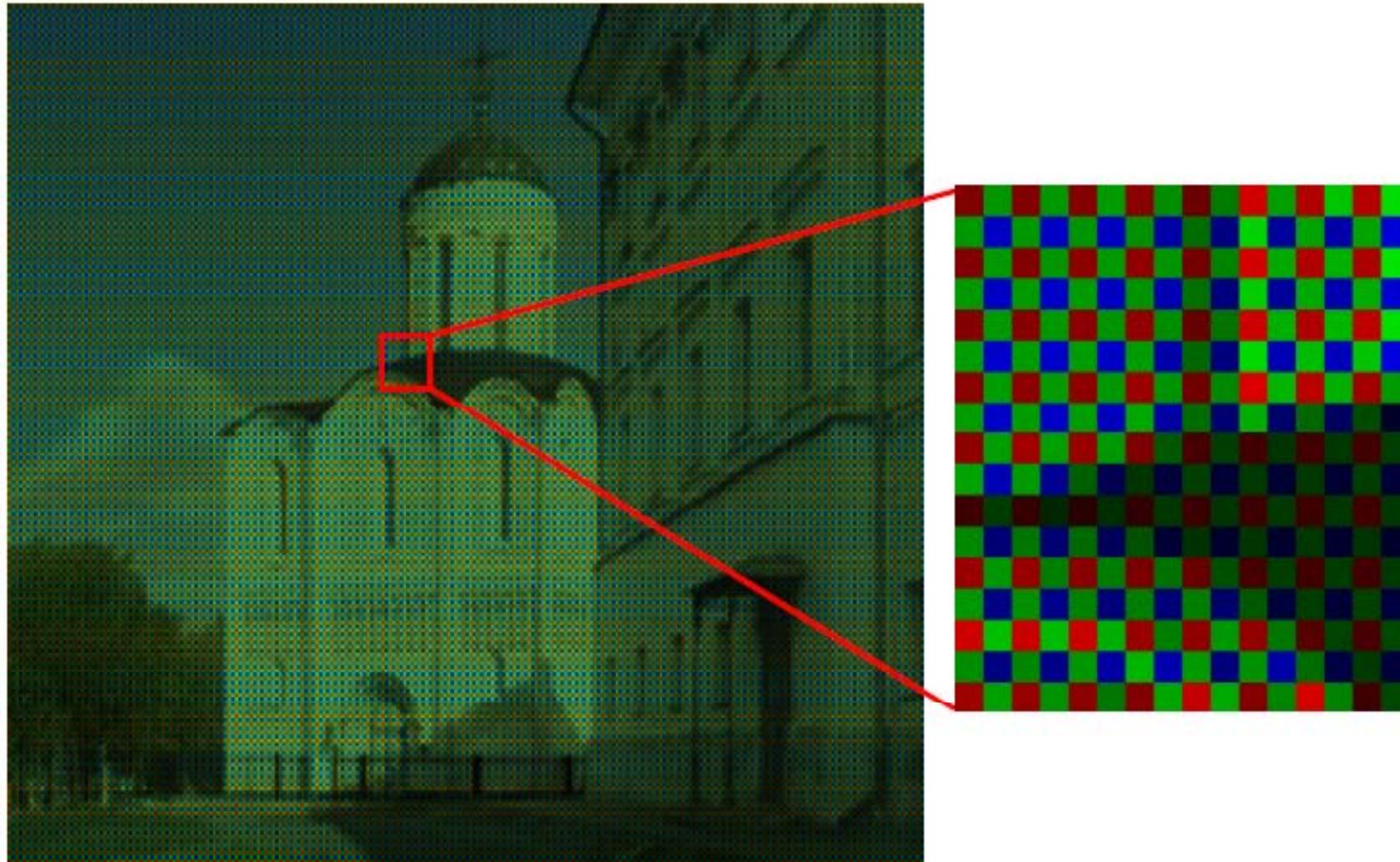
Bayer Filter



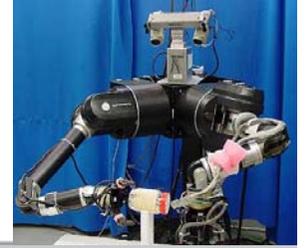
- **Specific arrangement** of a color filter array on a photo sensor due to B. E. Bayer.
- Color pattern has **50% green** elements, 25% red and 25% blue, also referred to as RGBG or GRGB.
- Human eye has greater resolving power with green light.
- **Demosaicing algorithms** convert from Bayer color pattern to RGB by interpolating neighboring values.



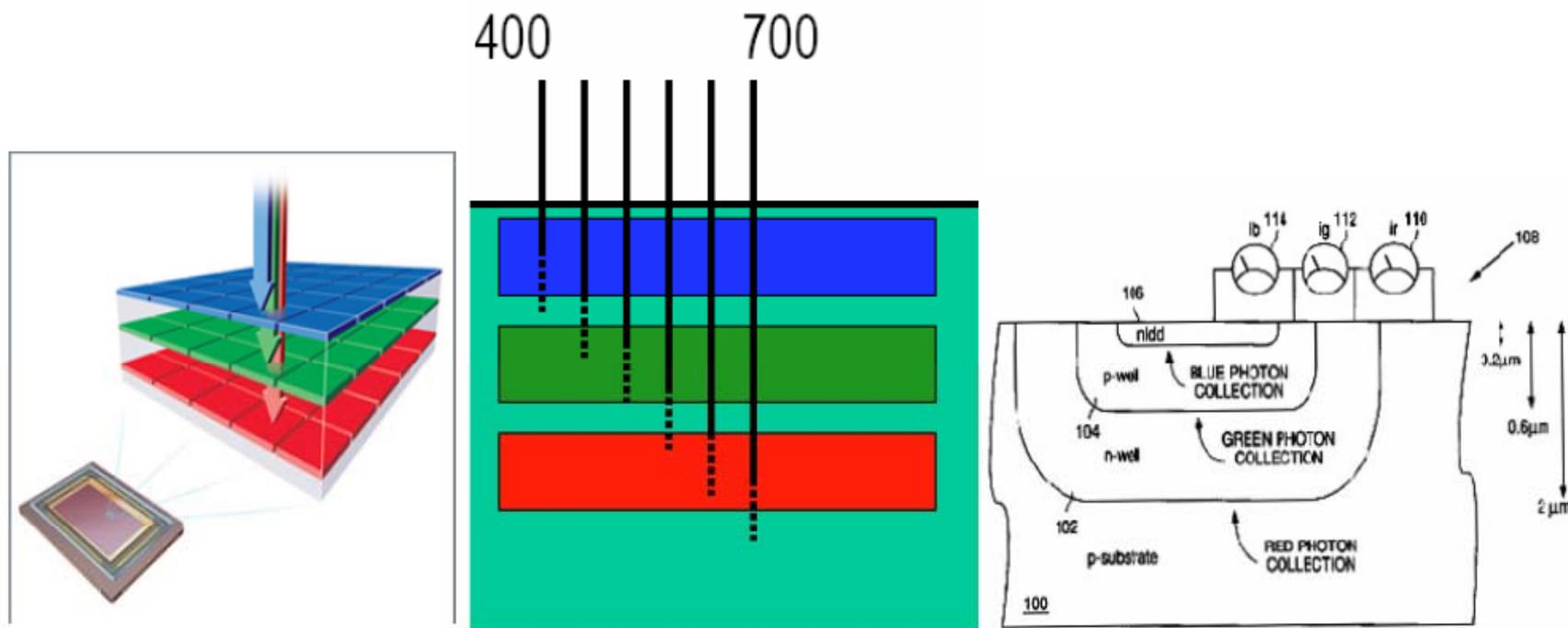
Bayer Filter



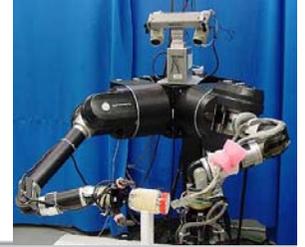
Foveon X3 sensor



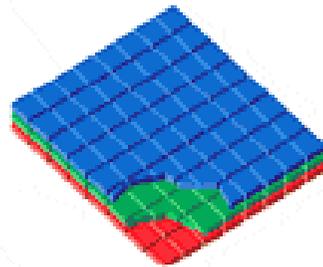
- Light penetrates to different depths for different wavelengths
- **Multilayer CMOS** sensor gets 3 different spectral sensitivities



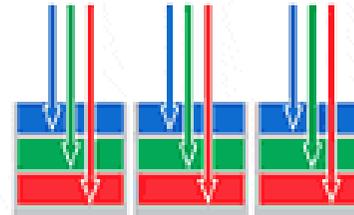
Foveon X3



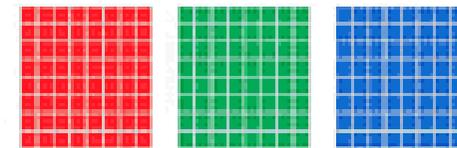
Foveon® X3™ Capture



A Foveon® X3™ image sensor features three separate layers of photodiodes embedded in silicon.



Since silicon absorbs different colors of light at different depths, each layer captures a different color. Stacked together, they create full-color pixels.



As a result, only Foveon X3 image sensors capture red, green and blue light at every pixel location.



red



green

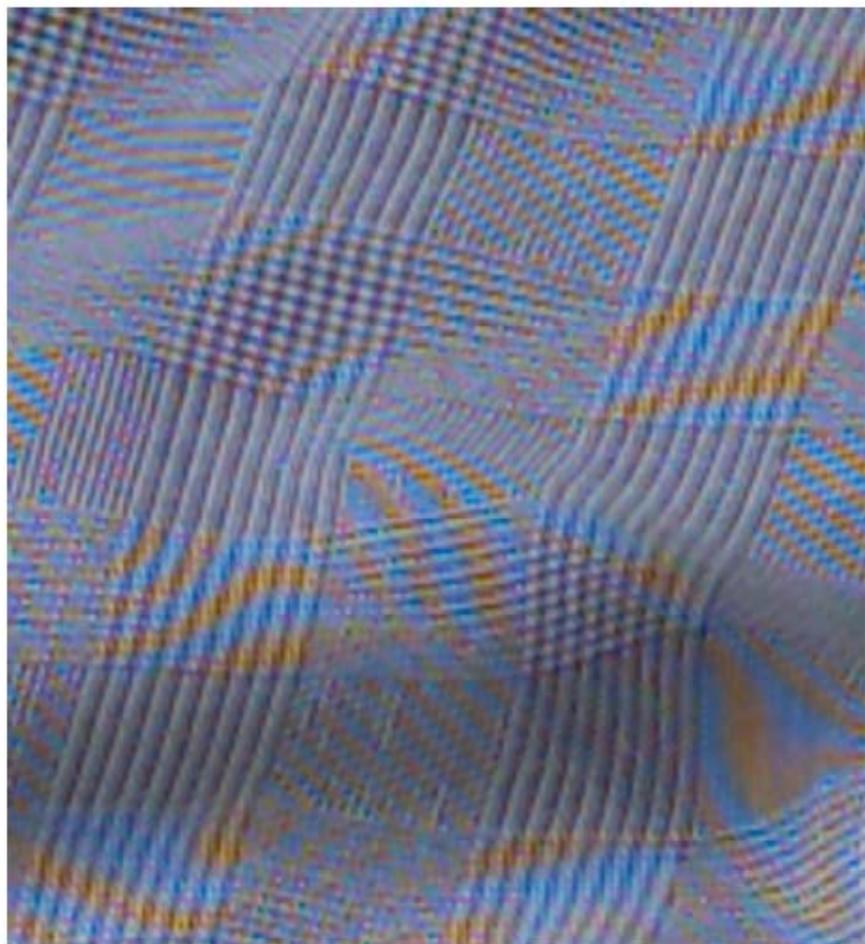
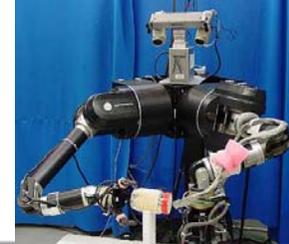


blue

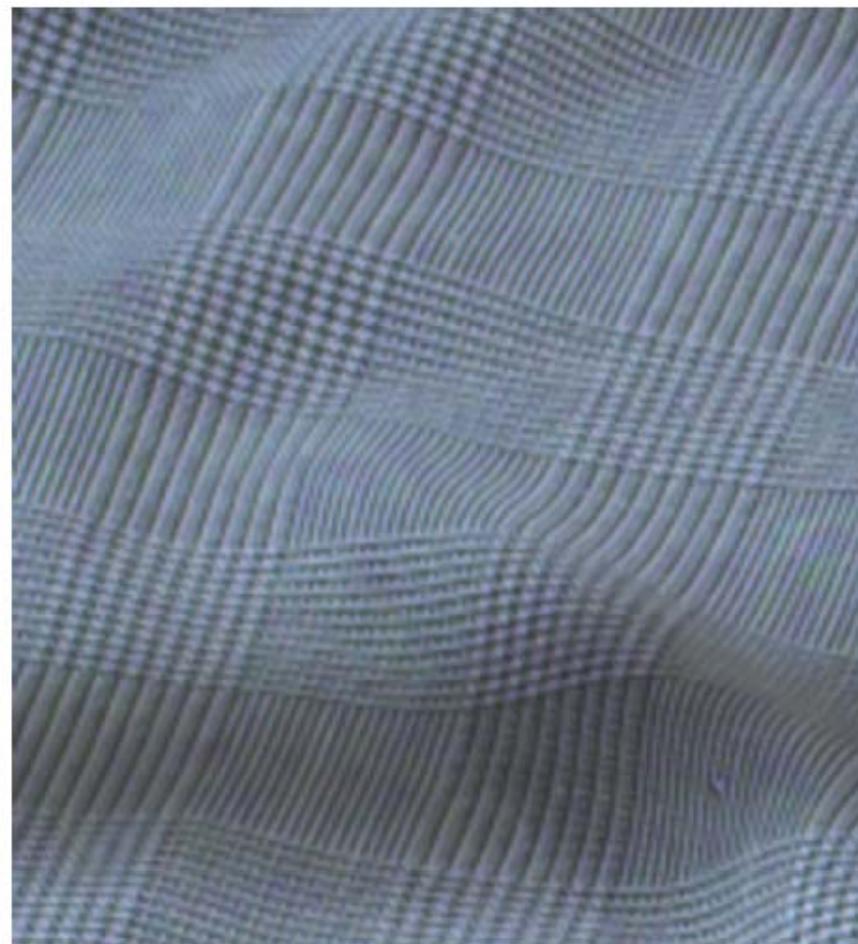


output

Foveon X3 sensor

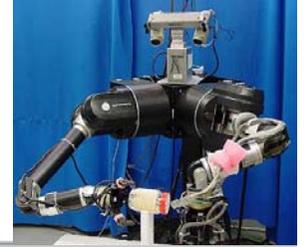


Bayer CFA



X3 sensor

Cameras with X3



Sigma SD10, SD9



Polaroid X530