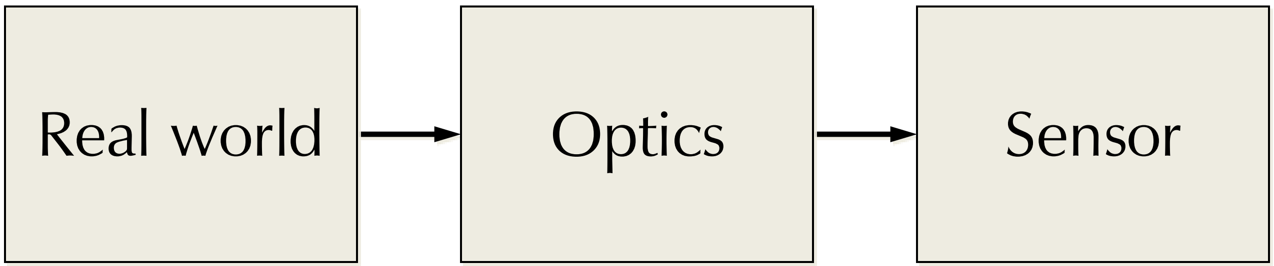
Image Formation and Capture

COS 429: Computer Vision

PRINCETON UNIVERSITY

Figure credits: B. Curless, E. Hecht, W.J. Smith, B.K.P. Horn, A. Theuwissen, and J. Malik

Image Formation and Capture



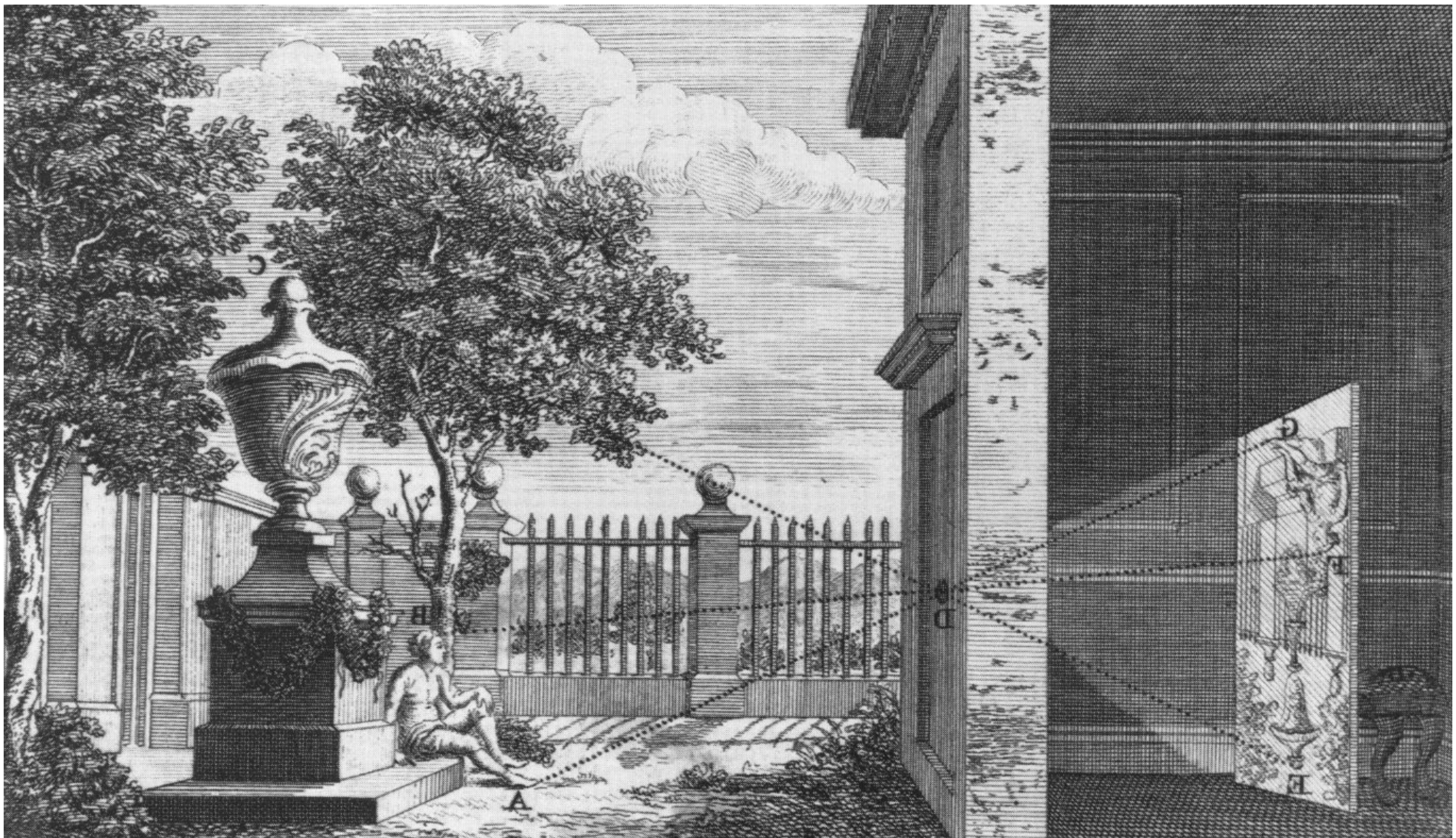
- Devices
- Sources of Error

Optics

- Pinhole camera
- Lenses
- Focus, aperture, distortion

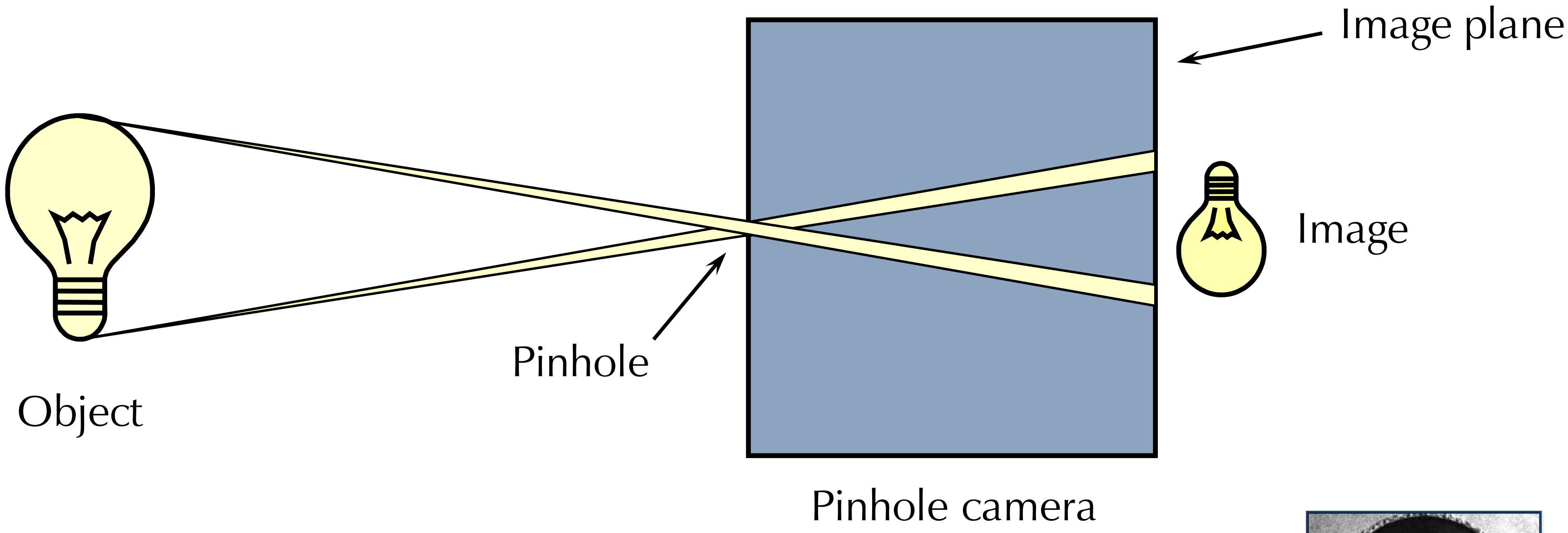
Pinhole Camera

- Camera obscura (“dark room”) – known since antiquity

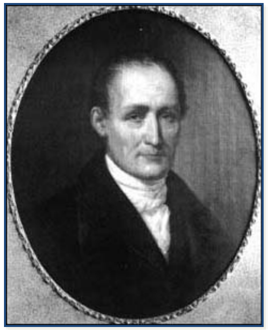


Pinhole Camera

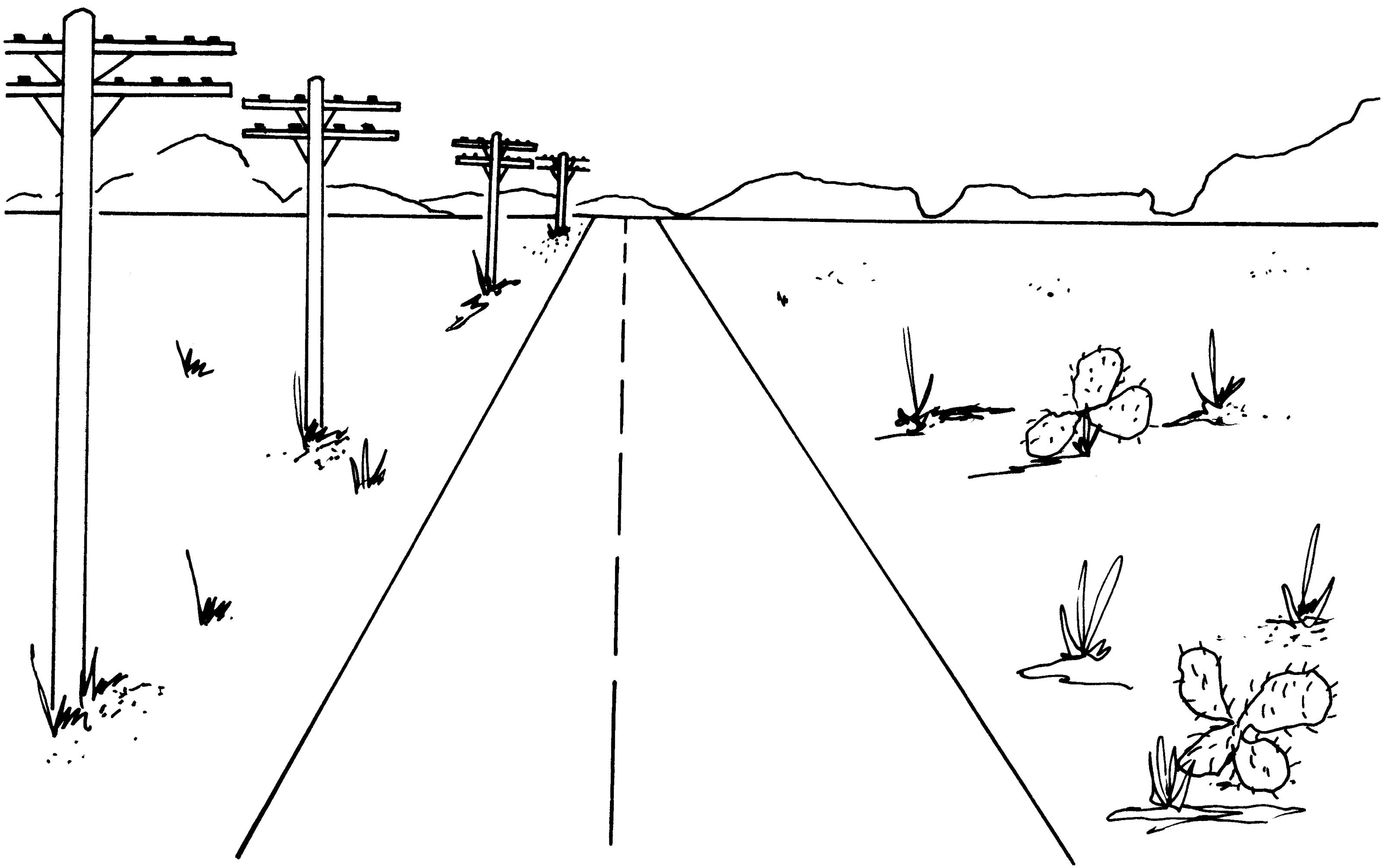
- Each point on image plane illuminated by light from one direction



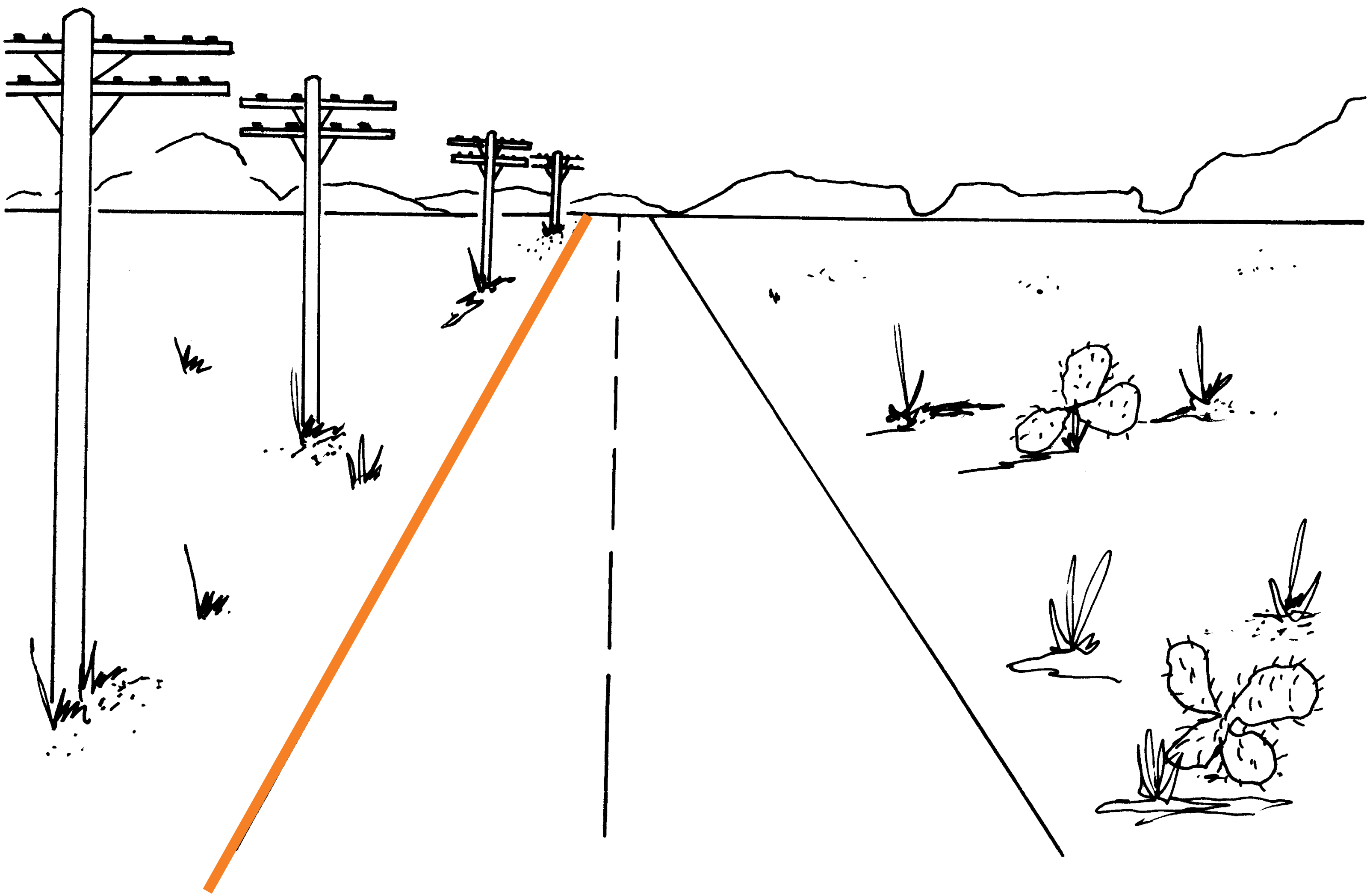
- Joseph Nicéphore Niépce: first recording onto pewter plate coated with bitumen



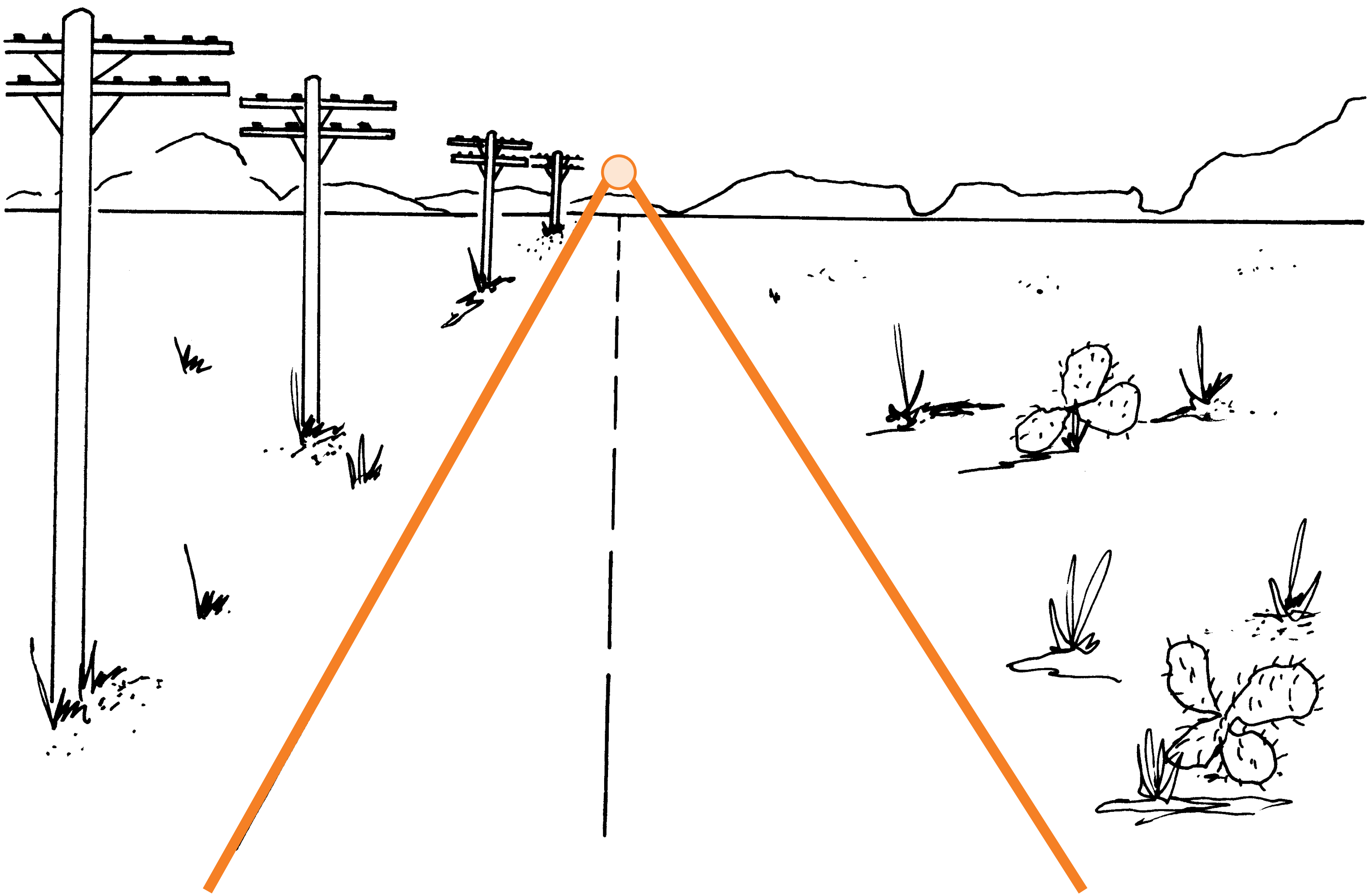
Perspective Projection Phenomena…



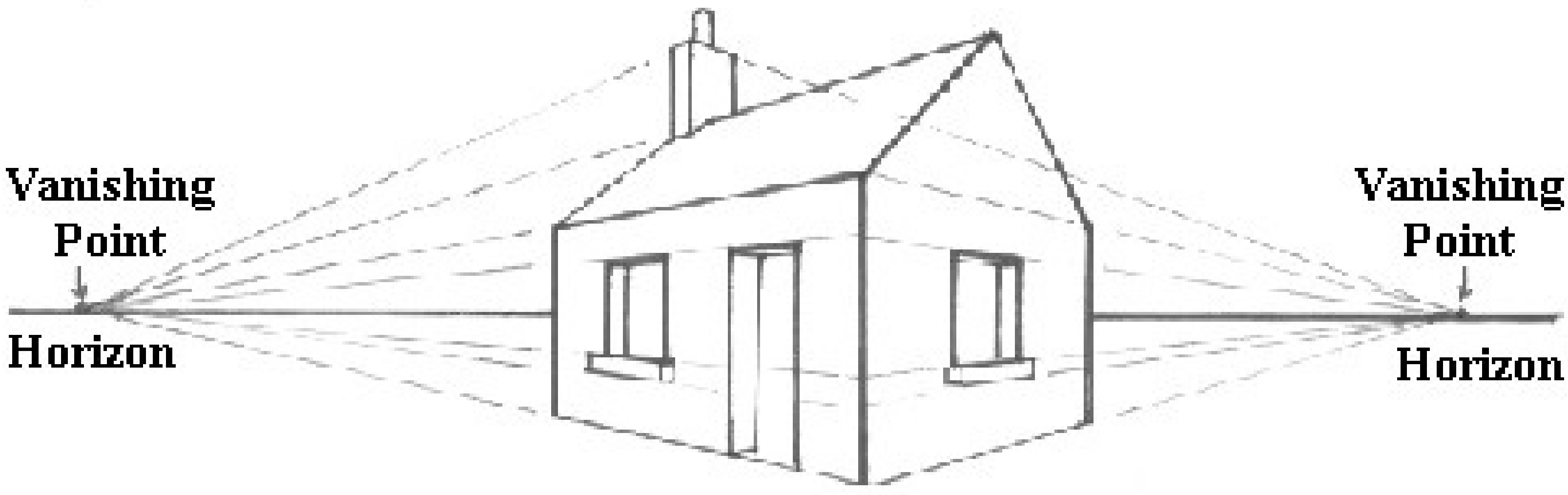
Straight Lines Remain Straight



Parallel Lines Converge at Vanishing Points



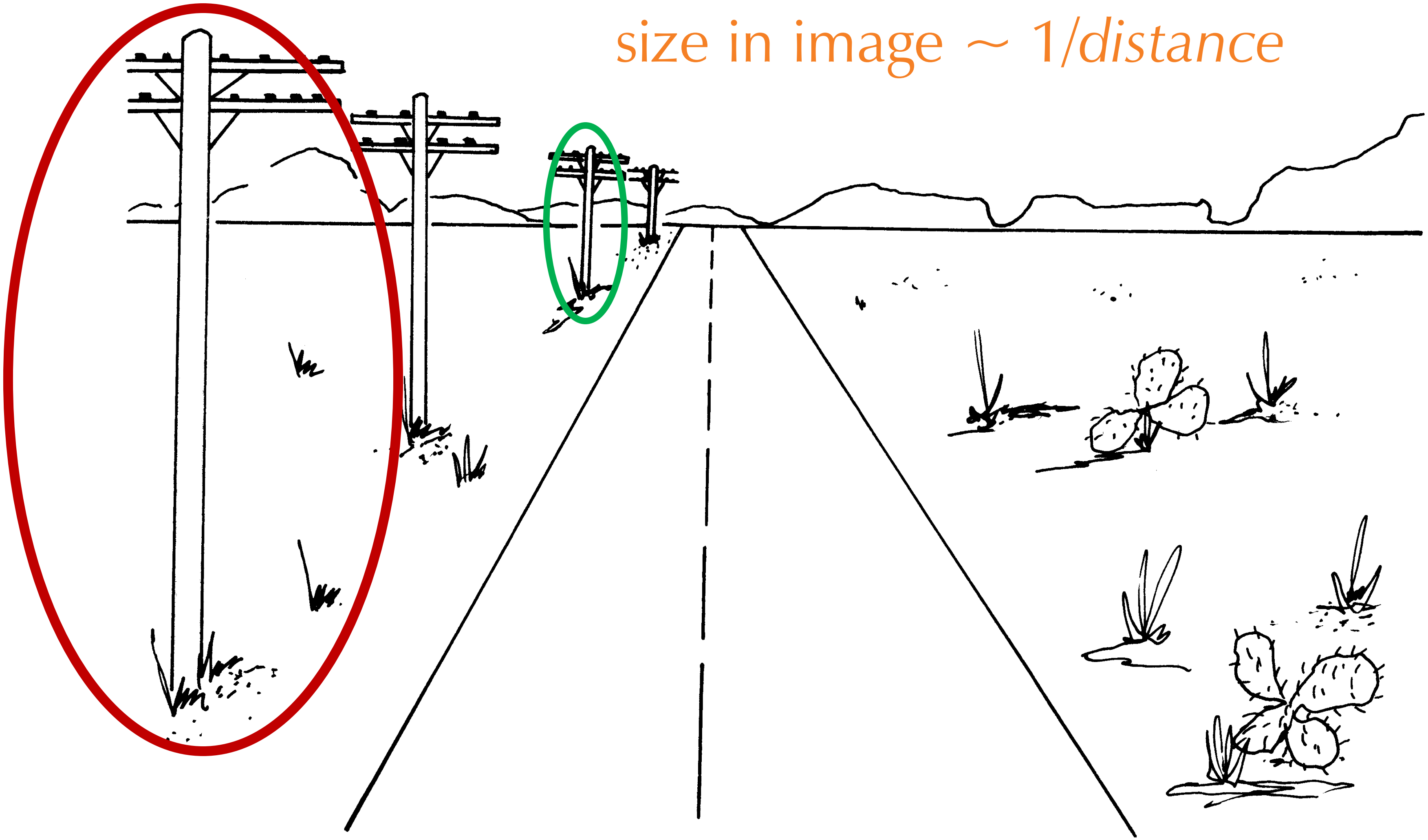
Parallel Lines Converge at Vanishing Points



Each family of parallel lines has its own vanishing point

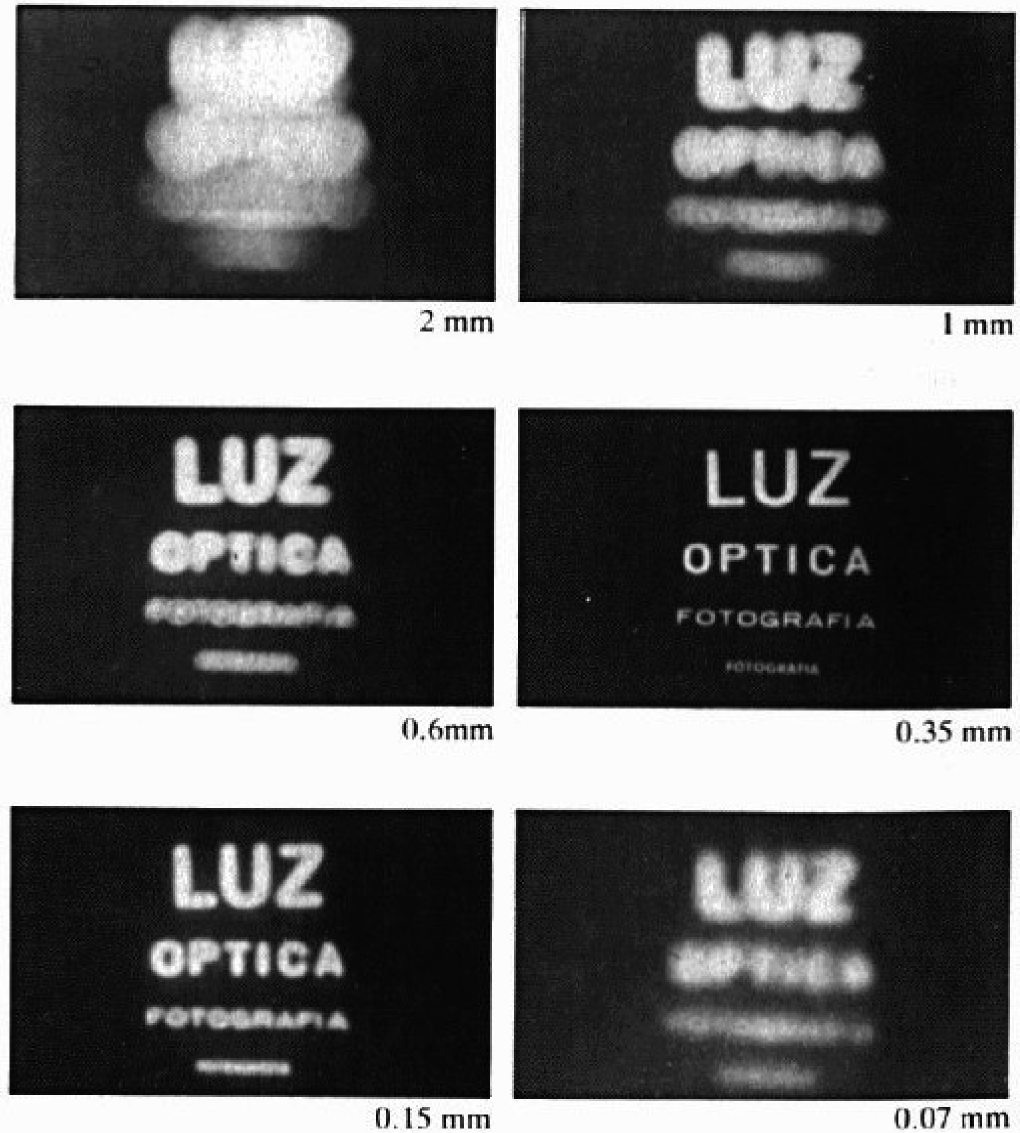
Nearer Objects Appear Bigger

size in image ~ 1/*distance*



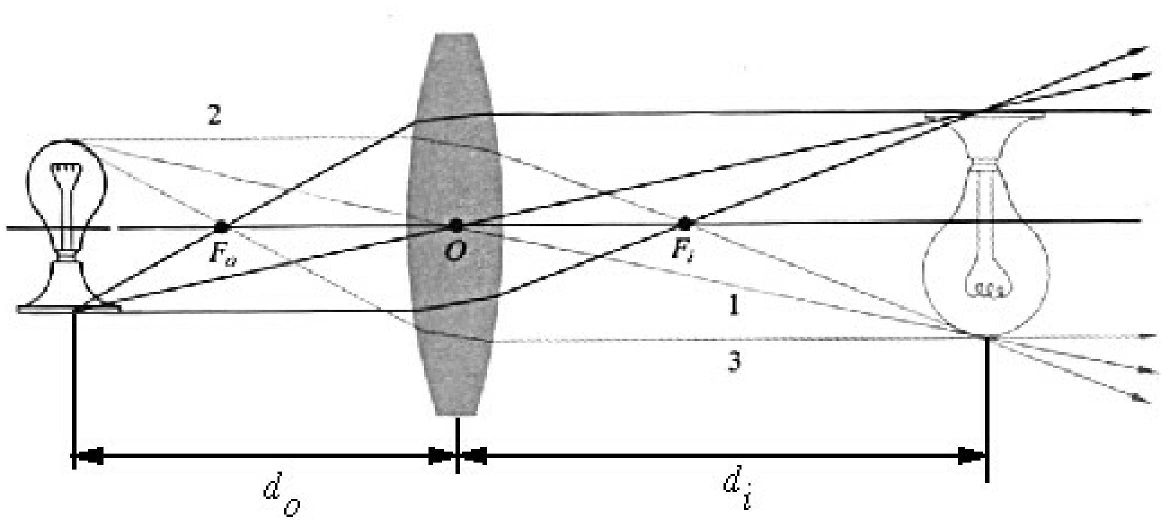
Pinhole Camera Limitations

- Aperture too big:
  blurry image
- Aperture too small:
  requires long exposure or high intensity
- Aperture much too small:
  diffraction through pinhole → blurry image



Lenses

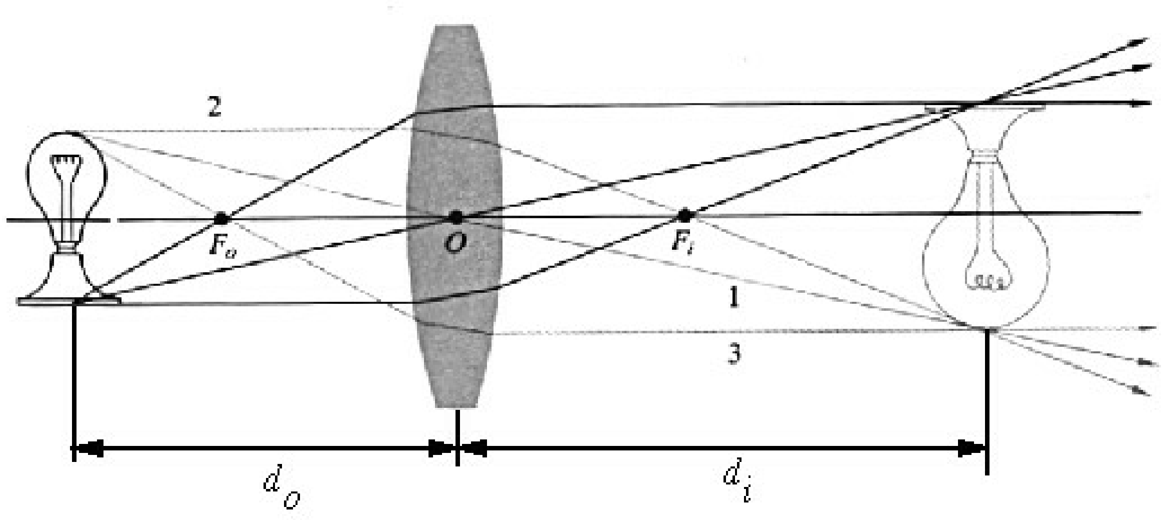
- Focus a bundle of rays from a scene point onto a single point on the imager
- Result: can make clear images with bigger aperture
  - But only one distance “in focus”



Ideal "Thin" Lens Law

- Relationship between focal distance and focal length of lens:

$$1/d_o + 1/d_i = 1/f$$

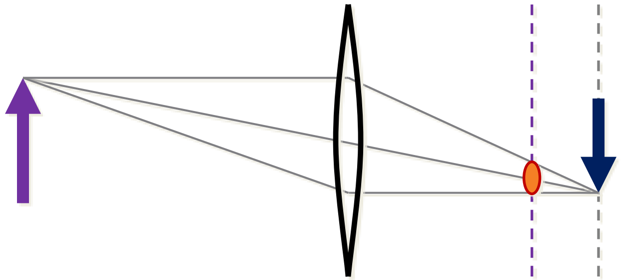


Camera Adjustments

- Focus?
  - Changes d_i
- Iris?
- Zoom?

Focus and Depth of Field

- For a given d_i, "perfect" focus at only one d_o
- In practice, OK for some range of depths
  - Circle of confusion smaller than a pixel



- Better depth of field with smaller apertures
  - Better approximation to pinhole camera
- Also better depth of field with wide-angle lenses

Camera Adjustments

- Focus?
  - Changes d_i
- Iris?
  - Changes aperture
- Zoom?

Aperture

- Controls amount of light
- Affects depth of field
- Affects distortion (since thin-lens approximation is better near center of lens – stay tuned)

f/1.4 f/5.6 f/16

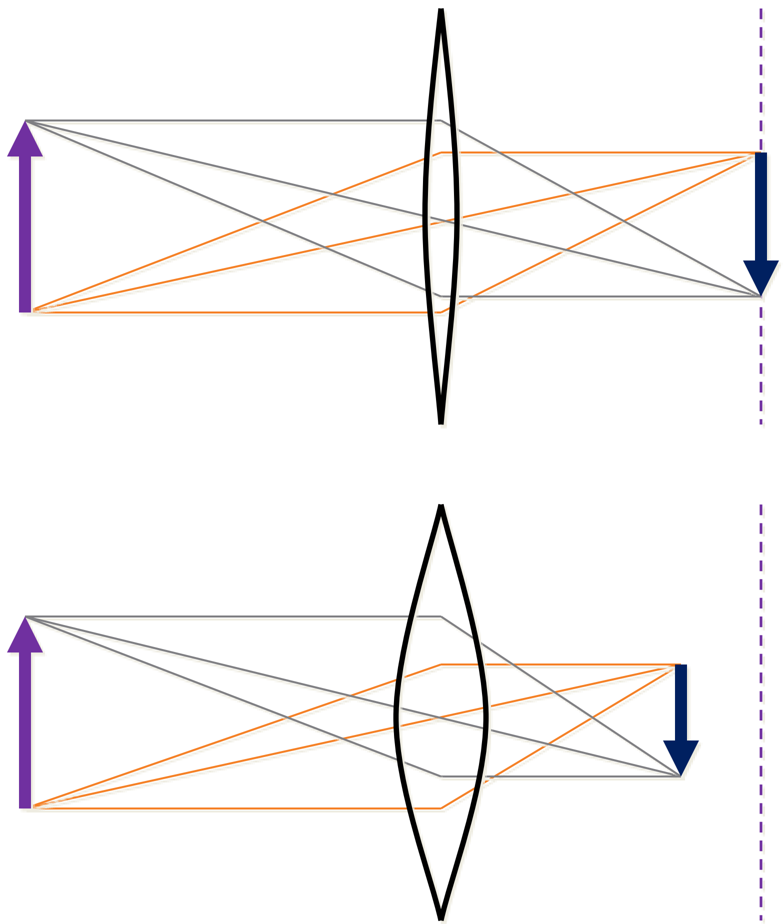
Aperture

- Aperture typically given as "*f*-number"
- What is *f*/4?
  - Aperture *diameter* is ¼ the focal length
- One "*f*-stop" equals change of *f*-number by $\sqrt{2}$
  - Equals change in aperture *area* by factor of 2
  - Equals change in amount of light by factor of 2
  - Example: *f*/2 $\rightarrow$ *f*/2.8 $\rightarrow$ *f*/4 (each one doubles light)

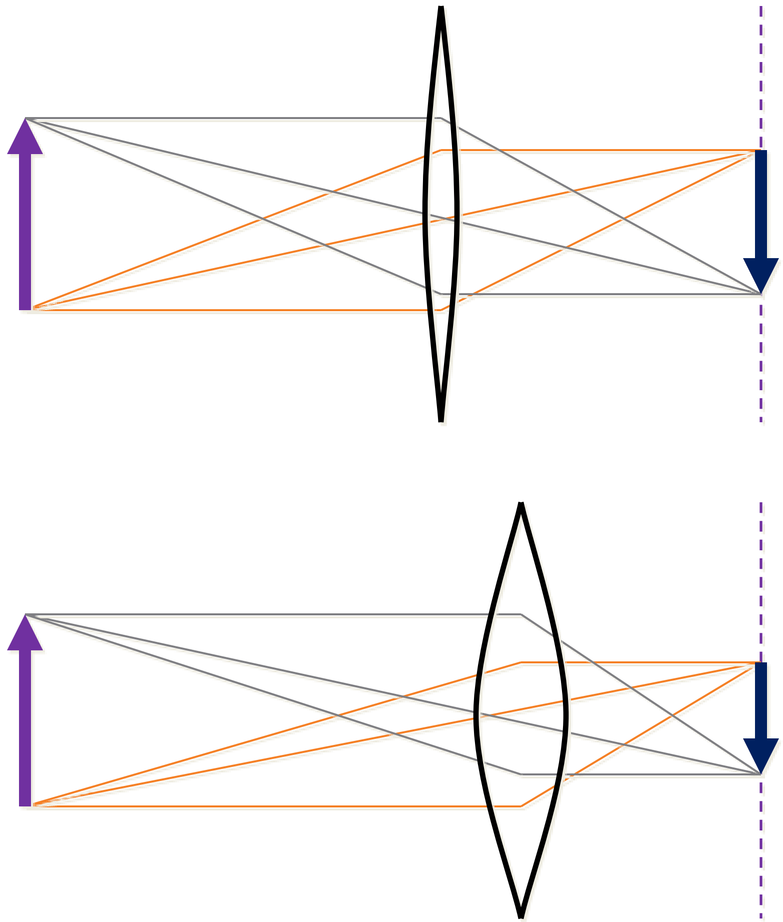
Camera Adjustments

- Focus?
  - Changes d_i
- Iris?
  - Changes aperture
- Zoom?
  - Changes f and sometimes d_i

Zoom Lenses – Varifocal



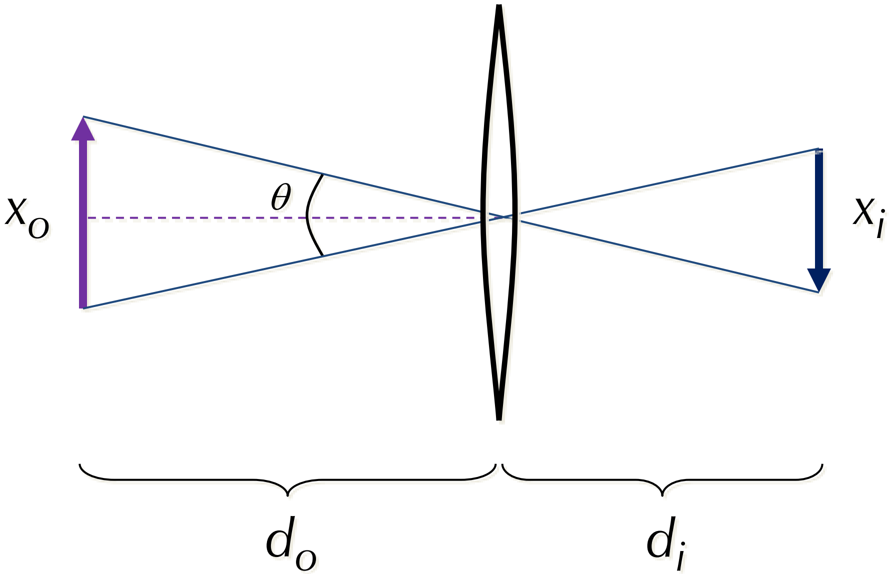
Zoom Lenses – Parfocal



Field of View

- Q: What does field of view of camera depend on?
  - Focal length of lens
  - Size of imager
  - Object distance?

Computing Field of View



$$1/d_o + 1/d_i = 1/f$$

$$\tan \theta/2 = ½\, x_o / d_o$$

$$x_o / d_o = x_i / d_i$$

$$\theta = 2 \tan^{-1} ½\, x_i (1/f - 1/d_o)$$

Since typically $d_o >> f$,

$$\theta \approx 2 \tan^{-1} ½\, x_i / f$$

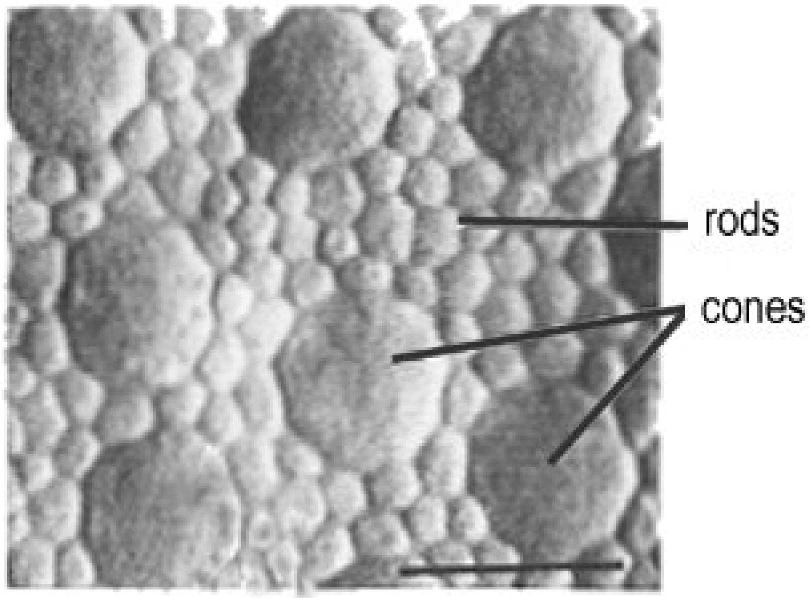
$$\theta \approx x_i / f$$

Photoreceptors

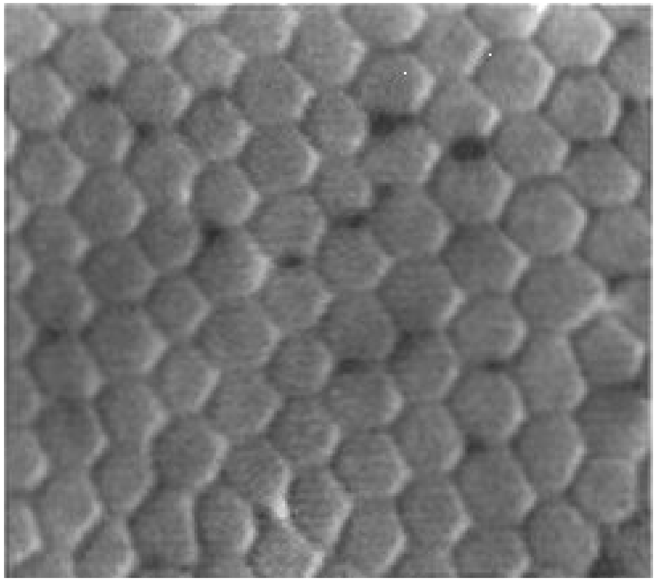
- Human retina
- Vidicon
- CCD and CMOS imagers

Photoreceptors in Human Retina

- Two types of receptors: rods and cones



Rods and cones

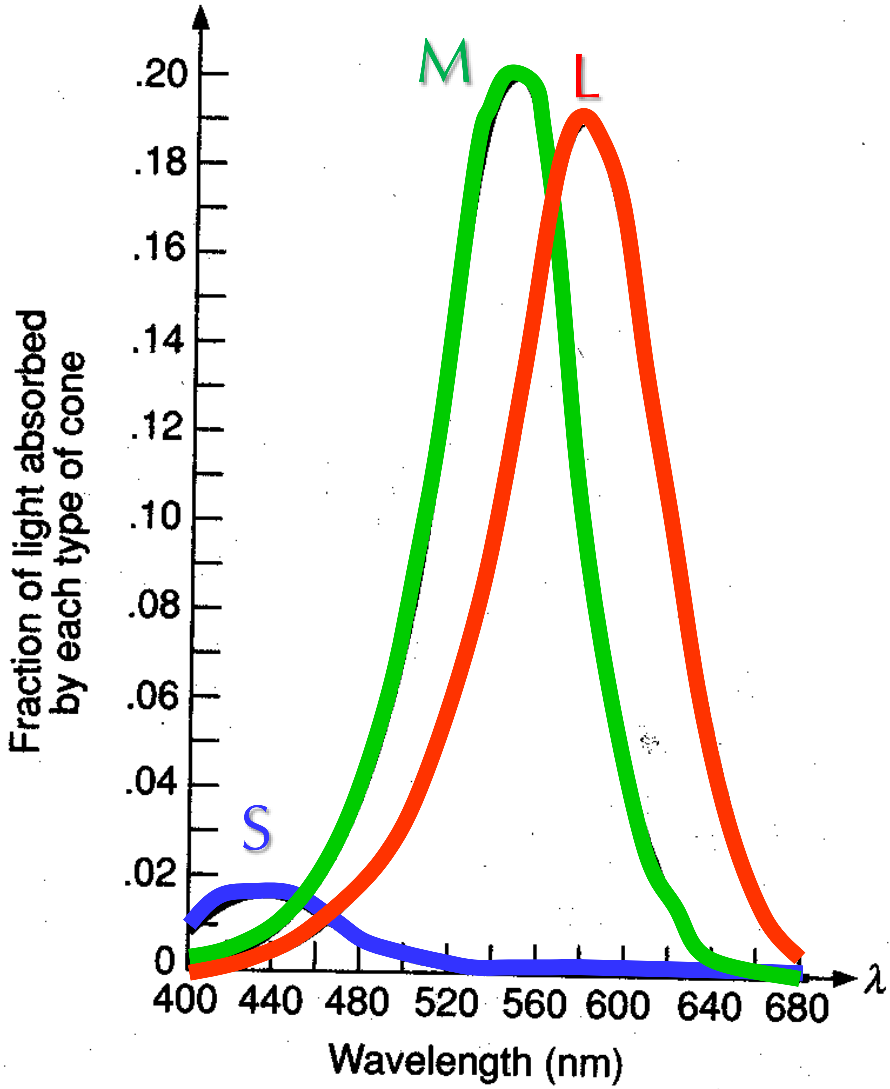


Cones in *fovea*
(central part of retina)

Rods and Cones

- Rods
  - More sensitive in low light: “scotopic” vision
  - More dense near periphery
- Cones
  - Only function with higher light levels: “photopic” vision
  - Densely packed at center of eye: fovea
  - Different types of cones → color vision

Color Perception



Spectral-response functions of the three types of cones (including absorption due to cornea and lens)

FvDFH

Tristimulus Color

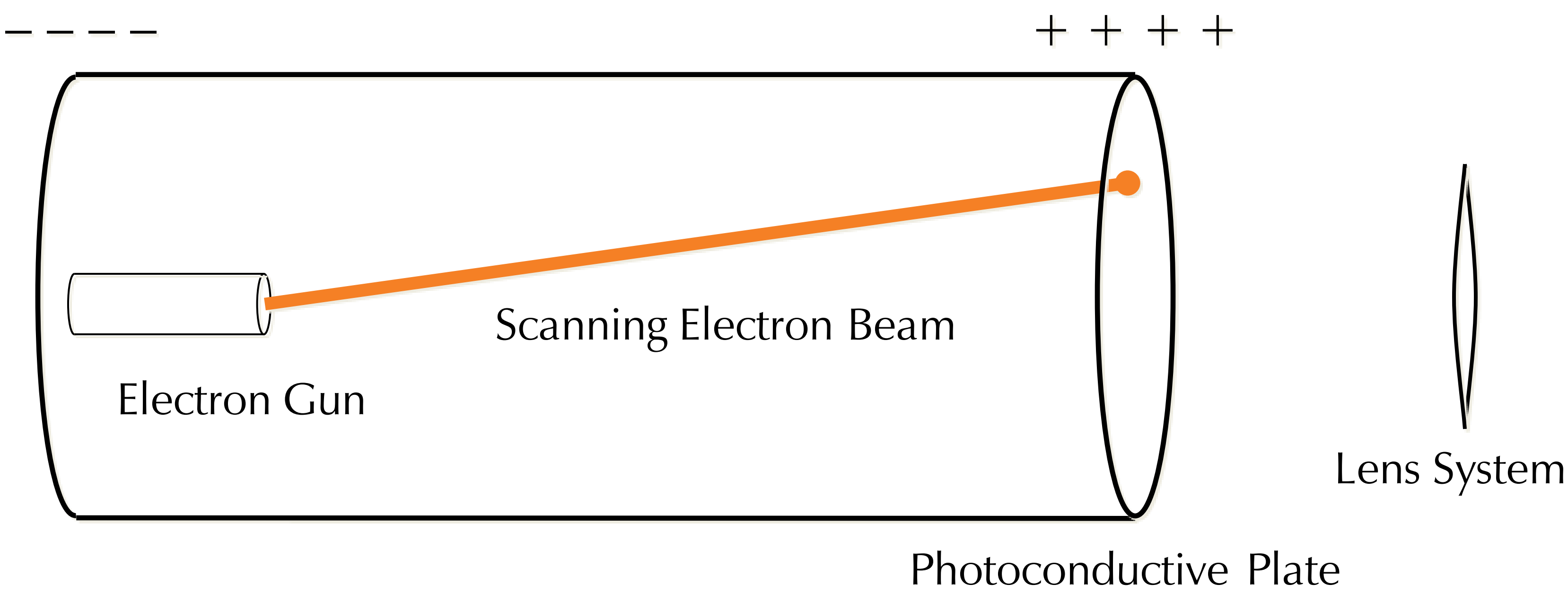
- Any distribution of light can be summarized by its effect on 3 types of cones
- Therefore, human perception of color is a 3-dimensional space
- Metamerism: different spectra, same response
- Color blindness: fewer than 3 types of cones
  - Most commonly L cone = M cone

Electronic Photoreceptors

- Analog technologies:
  - Coated plates
  - Film
- Digital technologies
  - Vidicon
  - CCD
  - CMOS imagers
- Produce regular grid of pixels
  - Measures light power integrated over some time period, over some area on image plane

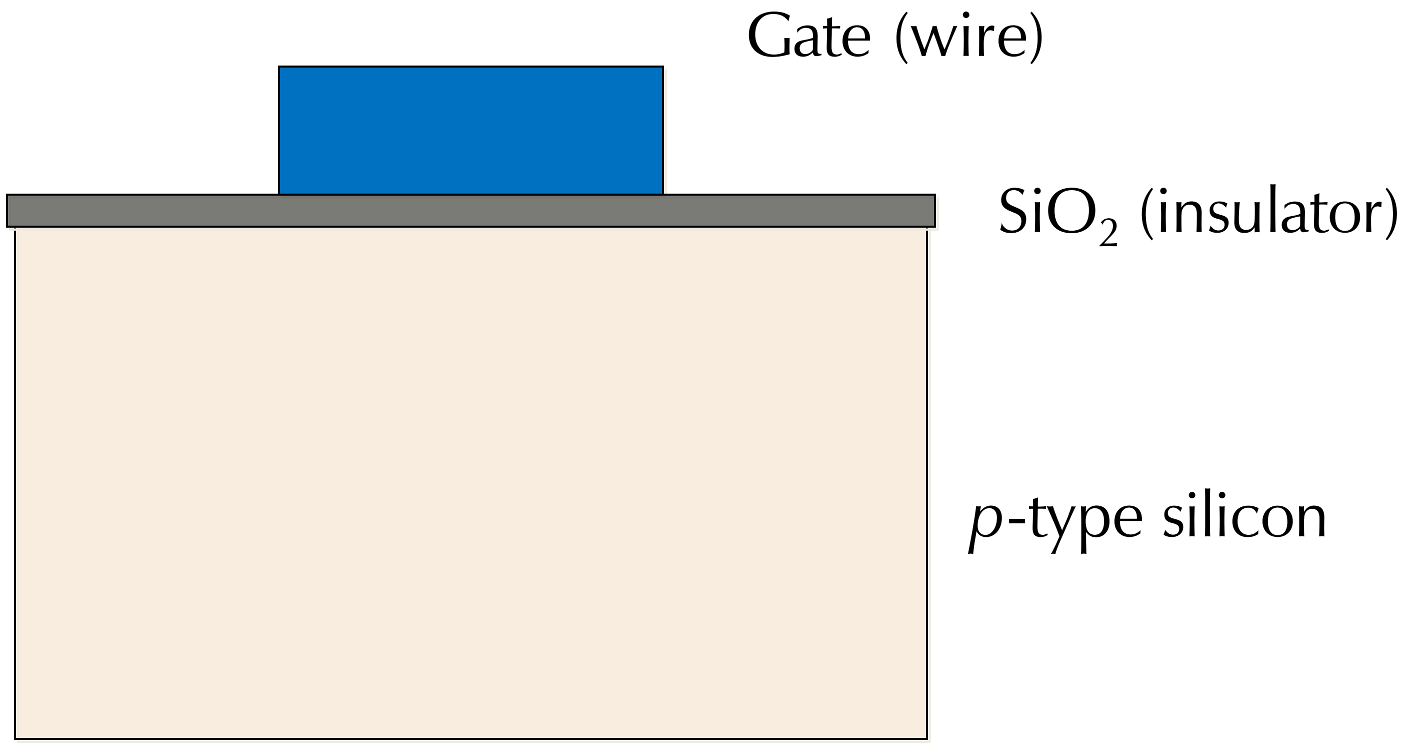
Vidicon

- Best-known in family of "photoconductive video cameras"
- Basically television in reverse



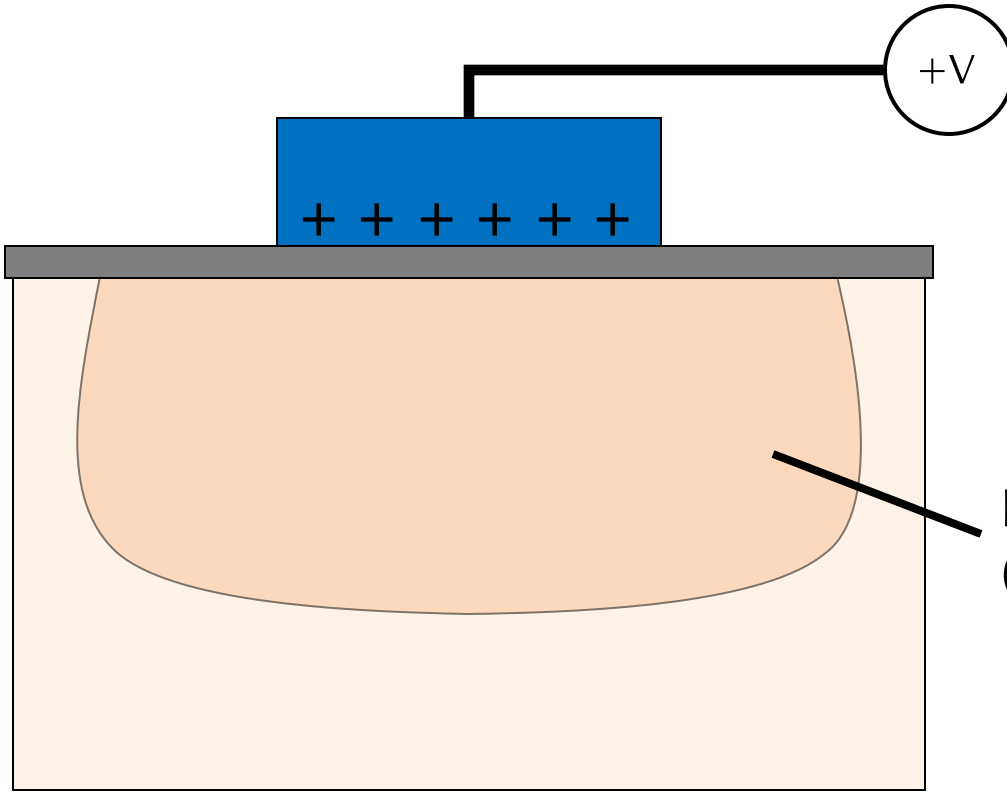
MOS Capacitors

- MOS = Metal Oxide Semiconductor



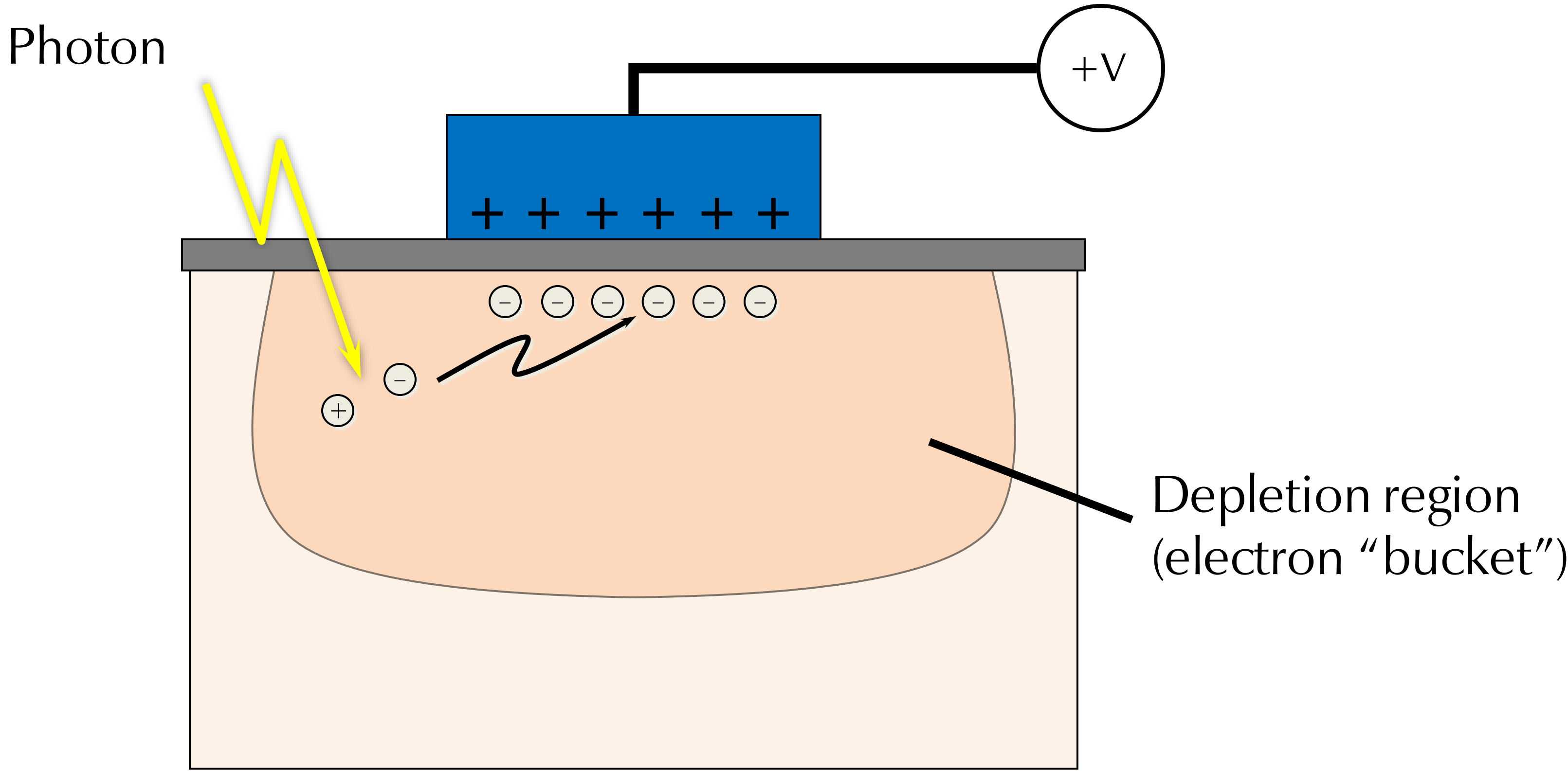
MOS Capacitors

- Voltage applied to gate repels positive “holes” in the semiconductor



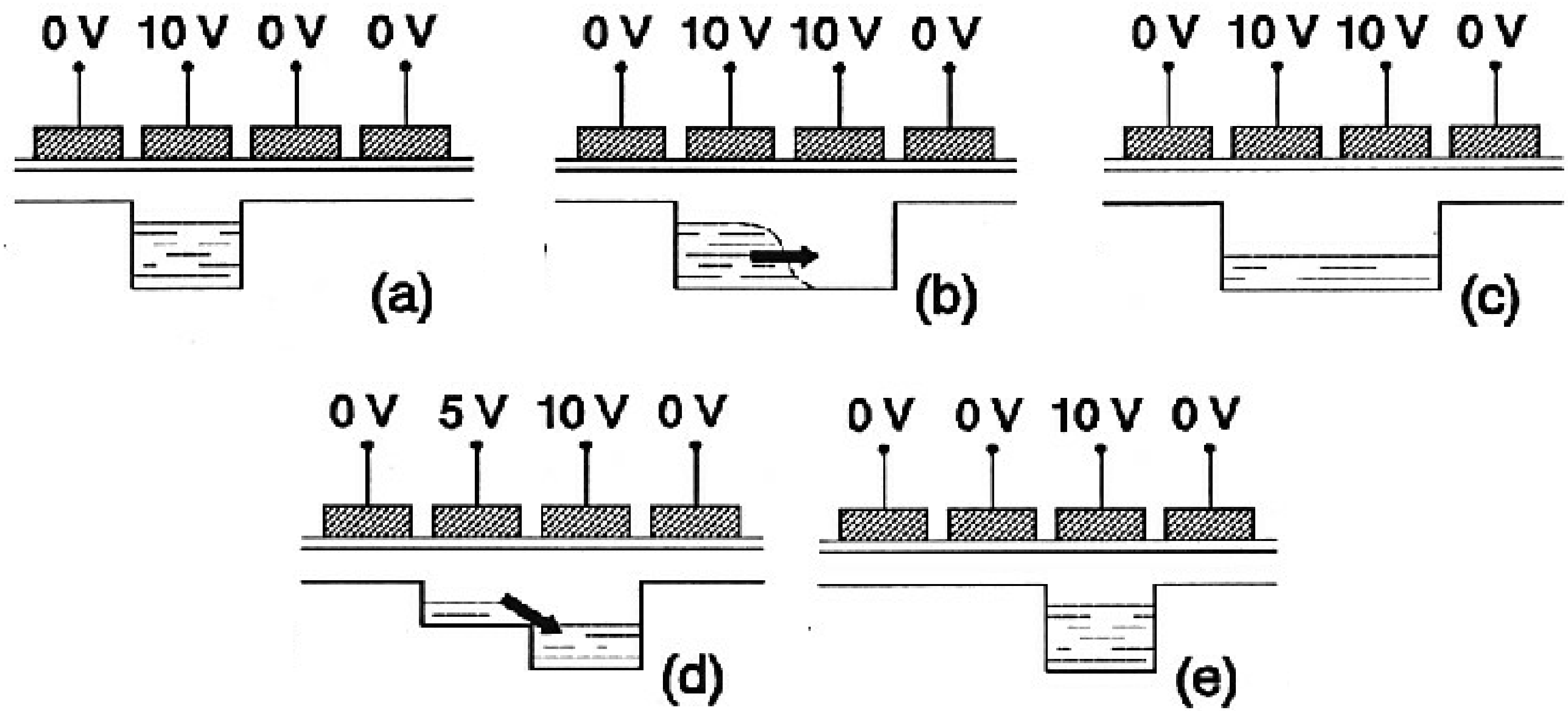
MOS Capacitors

- Photon striking the material creates electron-hole pair



Charge Transfer

- CCDs (Charge-Coupled Devices) move charge from one bucket to another by manipulating voltages

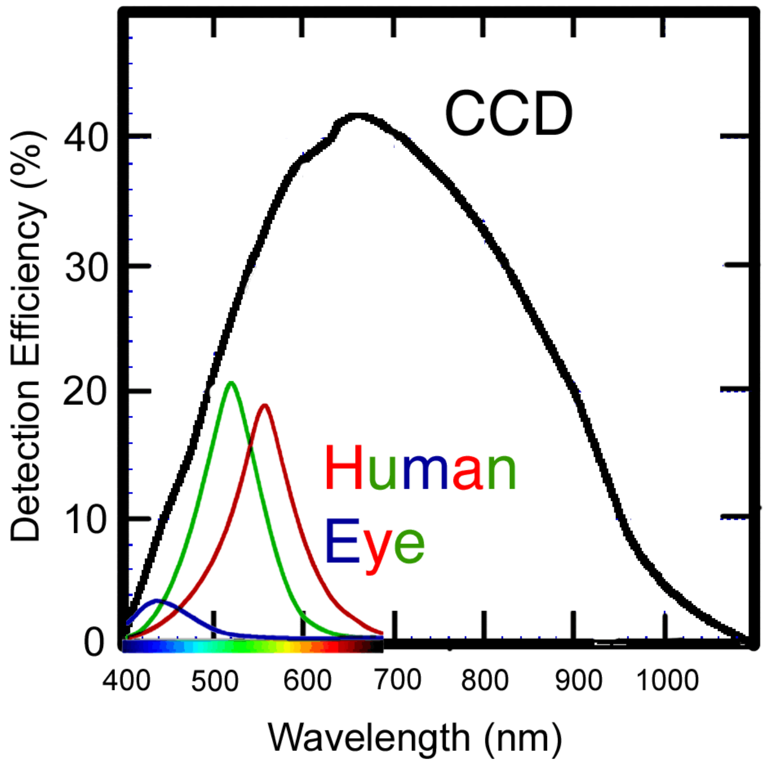


CMOS Imagers

- Recently, can manufacture chips that combine photosensitive elements and processing elements
- Benefits:
  - Partial readout
  - Signal processing
  - Eliminate some supporting chips → low cost

Color Cameras

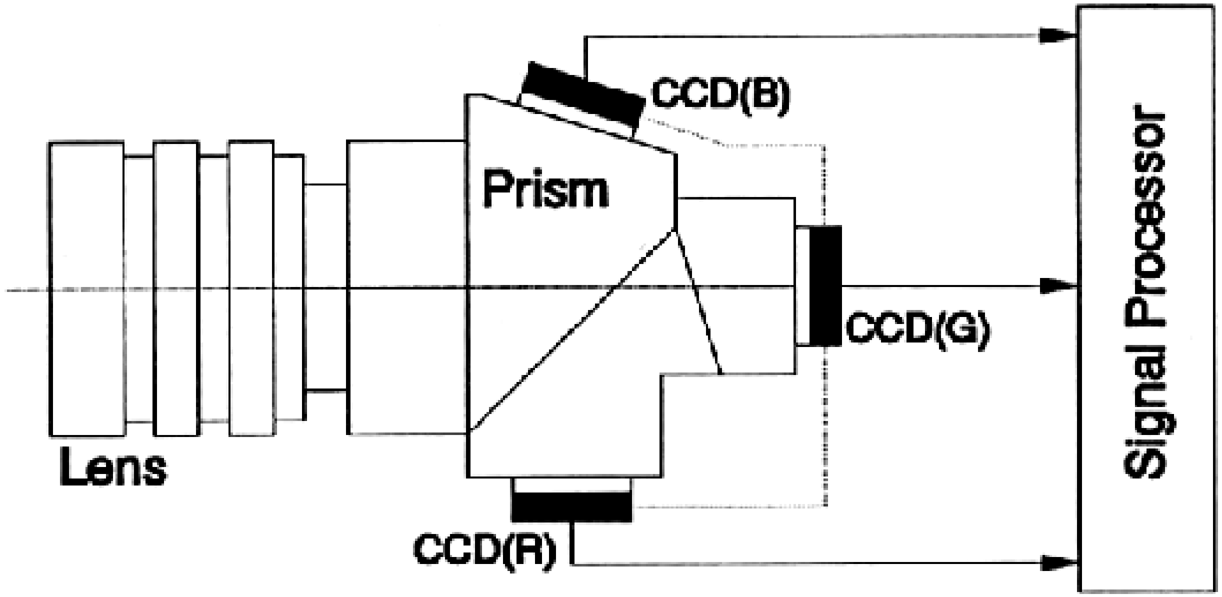
- CCD sensitivity does not match human eye



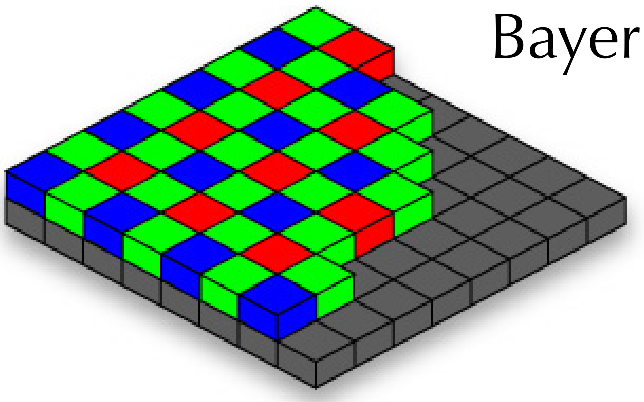
- Use band-pass color filters to adapt…

3-Chip Color Cameras

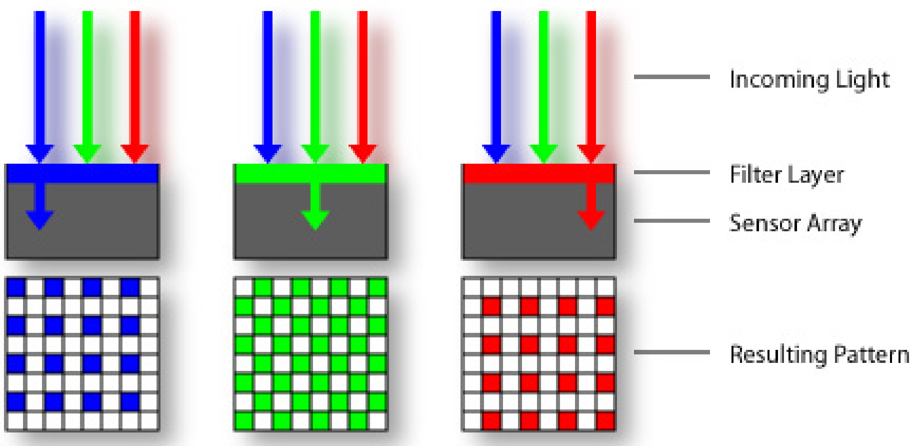
- Use prisms and filters to split image across 3 sensors
- Expensive, hard to align



1-Chip Color Cameras

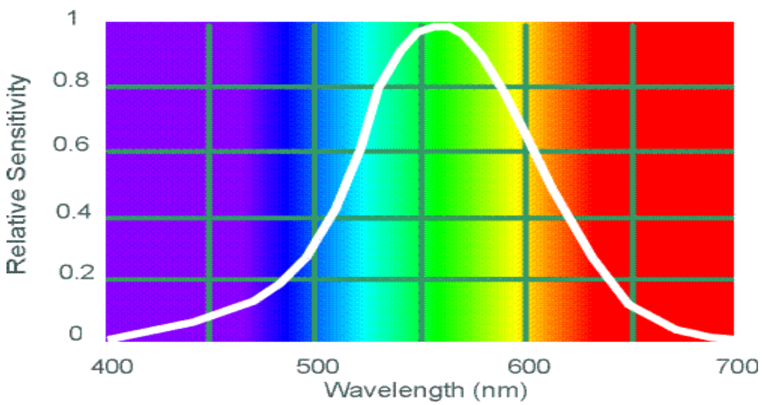


Bayer grid



Estimate missing components from neighboring values (demosaicing)

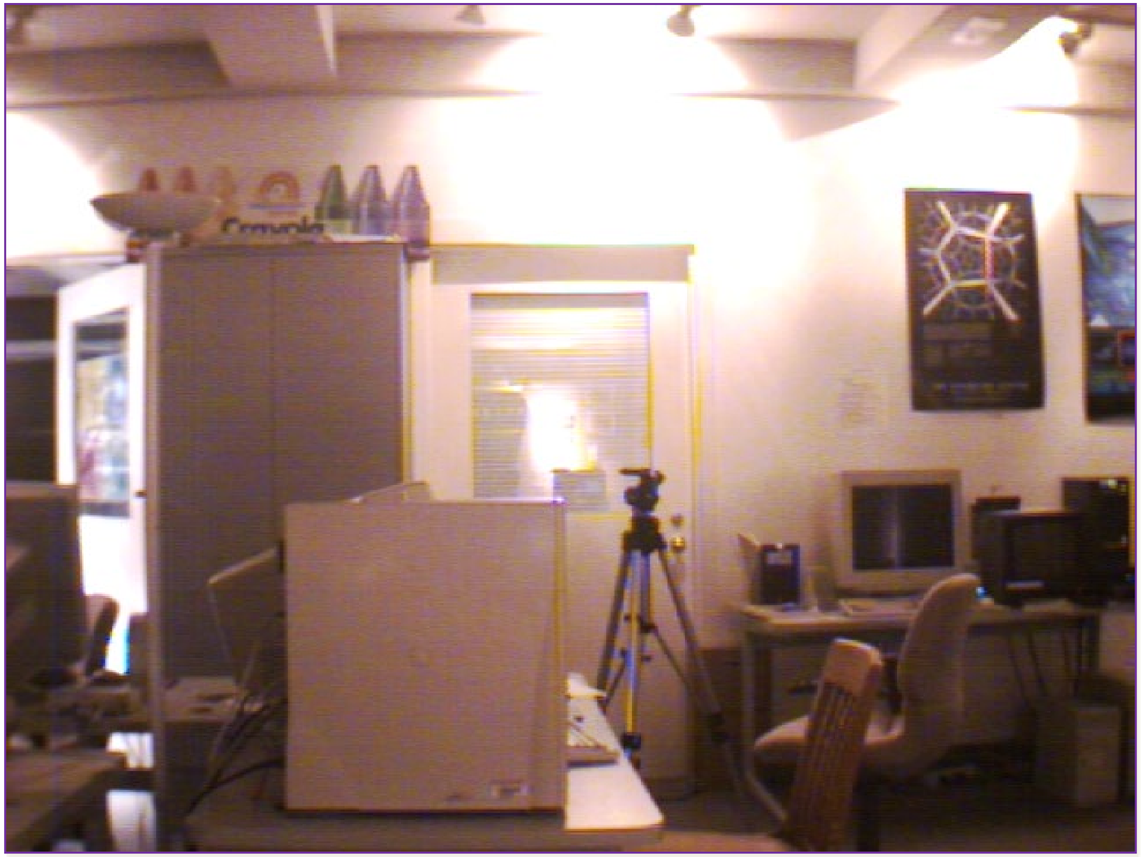
Why more green?



Human Luminance Sensitivity Function

Errors in Digital Images

- What are some sources of error in this image?



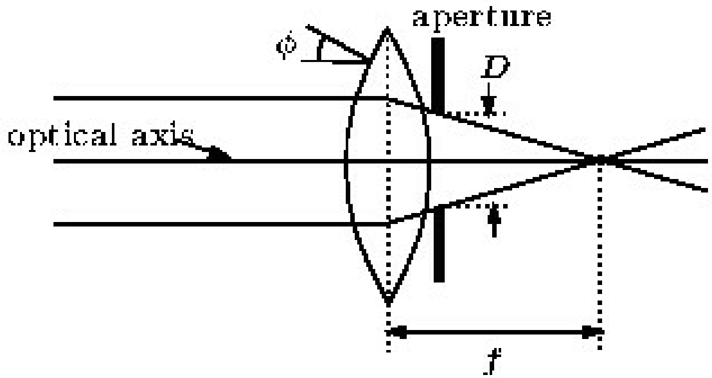
Sources of Error

- Geometric (focus, distortion)
- Color (1-chip artifacts, chromatic aberration)
- Radiometric (cosine falloff, vignetting)
- Bright areas (flare, bloom, clamping)
- Signal processing (gamma, compression)
- Noise

Monochromatic Aberrations

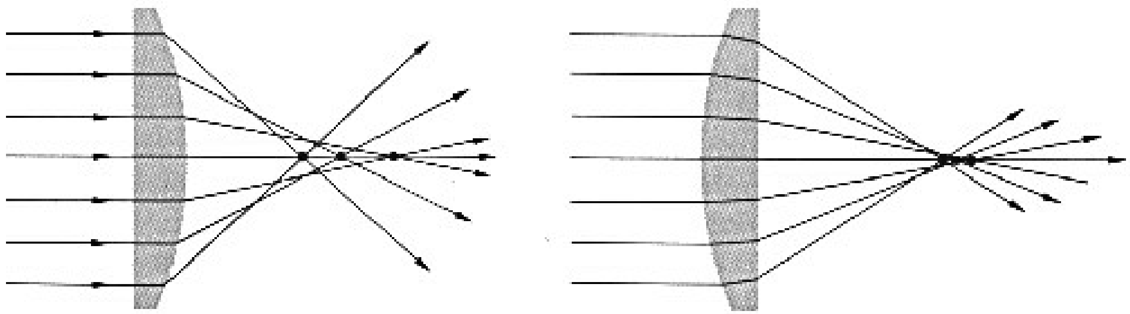
- Real lenses do not follow thin lens approximation because surfaces are spherical (manufacturing constraints)
- Result: thin-lens approximation only valid iff

$$\sin \varphi \approx \varphi$$



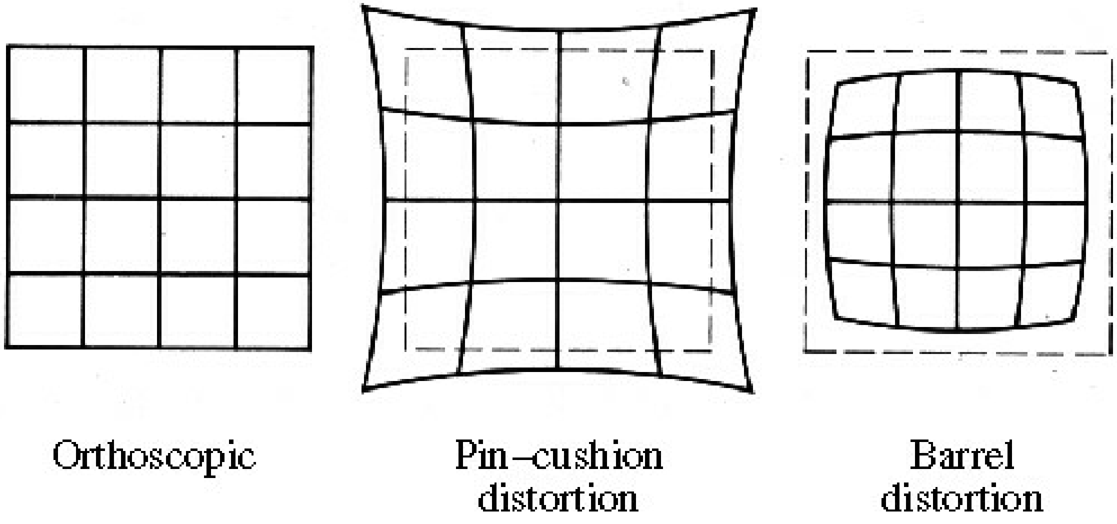
Spherical Aberration

- Results in blurring of image, focus shifts when aperture is stopped down
- Can vary with the way lenses are oriented

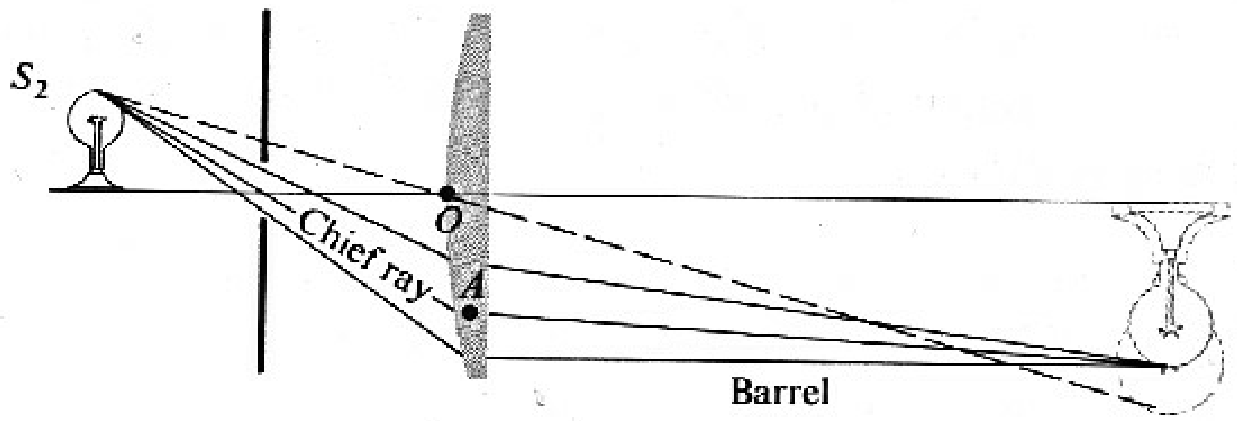


Distortion

- Pincushion or barrel radial distortion
  - Straight lines in the world no longer straight in image

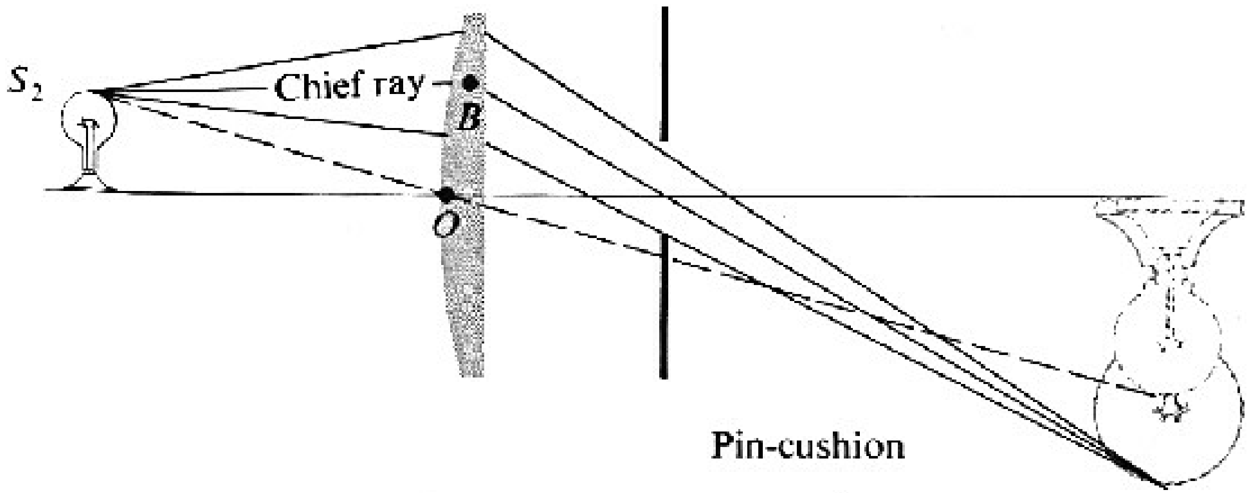


Distortion



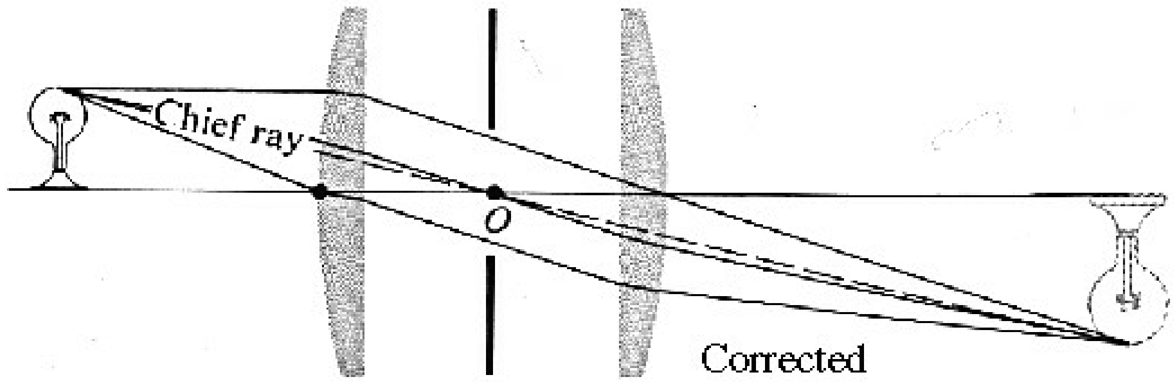
- Varies with placement of aperture

Distortion



- Varies with placement of aperture

Distortion



- Varies with placement of aperture

First-Order Radial Distortion

- Goal: mathematical formula for distortion
- If small, can be approximated by "first-order" formula (like Taylor series expansion):

$$
\begin{aligned}
r' &= r\,(1 + \kappa r^2) \\
r &= \text{ideal distance to center of image} \\
r' &= \text{distorted distance to center of image}
\end{aligned}
$$

- Higher-order models are possible

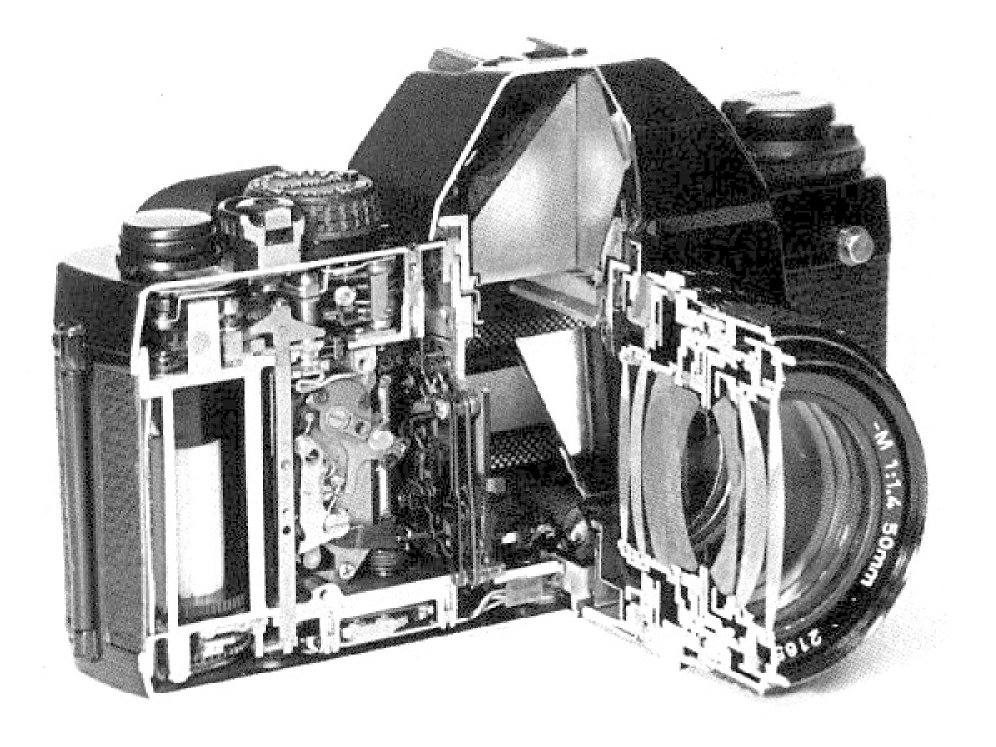
Chromatic Aberration

- Due to dispersion in glass (focal length varies with the wavelength of light)
- Result: color fringes
- Worst at edges of image

- Correct by building lens systems with multiple kinds of glass

Correcting for Aberrations

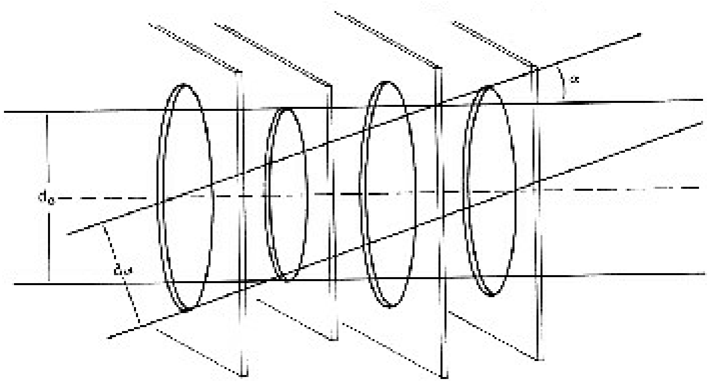
- High-quality compound lenses use multiple lens elements to cancel out distortion and aberration



- Often 5-10 elements, more for zooms

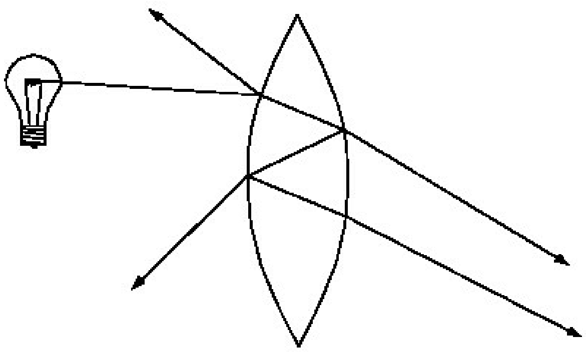
Other Limitations of Lenses

- Optical vignetting:
  less power per unit area
  for light at an oblique angle
  - Approximate falloff $\sim \cos^4 \varphi$
  - Result: darkening of edges
  - Also mechanical vignetting due to multiple apertures



Other Limitations of Lenses

- Flare: light reflecting (often multiple times) from glass-air interface
  - Results in ghost images or haziness
  - Worse in multi-lens systems
  - Ameliorated by optical coatings (thin-film interference)



- Bloom: overflow of charge in CCD buckets
  - Spills to adjacent buckets
  - Streaks (usually vertical) next to bright areas
  - Some cameras have "anti-bloom" circuitry

Flare and Bloom

Tanaka

Dynamic Range

- Most common cameras have 8-bit (per color channel) dynamic range
  - Can be nonlinear: more than 255:1 intensity range
- Too bright: clamp to maximum
- Too dim: clamp to 0
- Specialty cameras with higher dynamic range (usually 10-, 12-, and 16-bit)

High Dynamic Range (HDR) from Ordinary Cameras

- Take pictures of same scene with different shutter speeds
- Identify regions clamped to 0 or 255
- Average other pixels, scaled by 1 / shutter speed
- Can extend dynamic range, but limitations of optics and imager (noise, flare, bloom) still apply

Gamma

- Vidicon tube naturally has signal that varies with light intensity according to a power law:

$$\text{Signal} = E^{\gamma}, \quad \gamma \approx {}^{1}/_{2.5}$$

- CRT (televisions) naturally obey a power law with gamma ≈ 2.3–2.5
- Result: video signal standard has gamma of 1/2.5
- CCDs and CMOS linear, but gamma ≈ 2.2 almost always applied

Consequences for Vision

- Output of most *camera systems* is not linear
- Know what it is! (Sometimes system automagically applies "gamma correction")
- Necessary to correct raw pixel values for:
  - Reflectance measurements
  - Shape from shading
  - Photometric stereo
  - Recognition under variable lighting

Consequences for Vision

- What about e.g. edge detection?
  - Often want "perceptually significant" edges
  - Standard nonlinear signal close to (inverse of) human response
  - Using nonlinear signal often the "right thing"

Noise

- Thermal noise: in all electronics
  - Noise at all frequencies
  - Proportional to temperature
  - Special cooled cameras available for low noise
- Shot noise: discrete photons / electrons
  - Shows up at extremely low intensities
  - CCDs / CMOS can have high efficiency – approaching 1 electron per photon

Noise

- 1/*f* noise – inversely proportional to frequency
  - Amount depends on quality, manufacturing techniques
  - Can be dominant source of noise
- All of the above apply for imager and amplifier

Filtering Noise

- Most common method – simple blur
  - e.g., convolution with Gaussian
- Adaptive filters to prevent bleed across intensity edges
- Other filters for specialized situations
  - e.g., "despeckling" (median filters) for dead pixels

- Next time!

David Macaulay
Great Moments in Architecture
Plate XV: Locating the Vanishing Point (June 8, 1874)