

SAIR

Spatial AI & Robotics Lab

CSE 473/573-A

L3: COLORING & WARPING

Chen Wang

Spatial AI & Robotics Lab

Department of Computer Science and Engineering



University at Buffalo The State University of New York

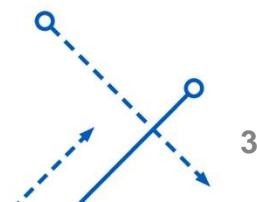


OPTICAL SENSOR

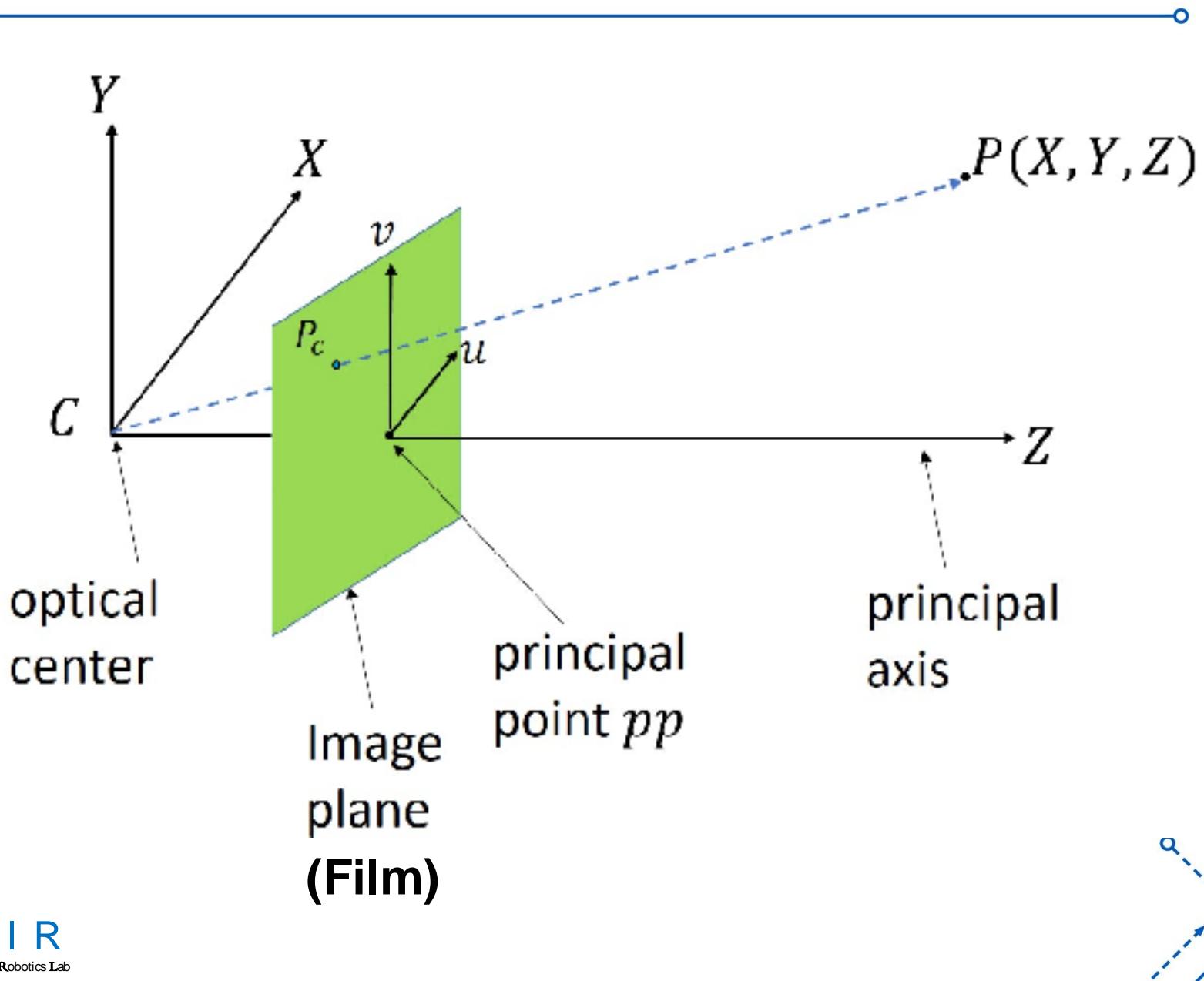


Content

- Optical Sensor
 - Image representation, resolution, sampling & quantization
 - Digital Camera, CCD, CMOS
- Color Space
 - Electromagnetic Spectrum, Bayer pattern
 - RGB, HSV, L*a*b*, YCbCr



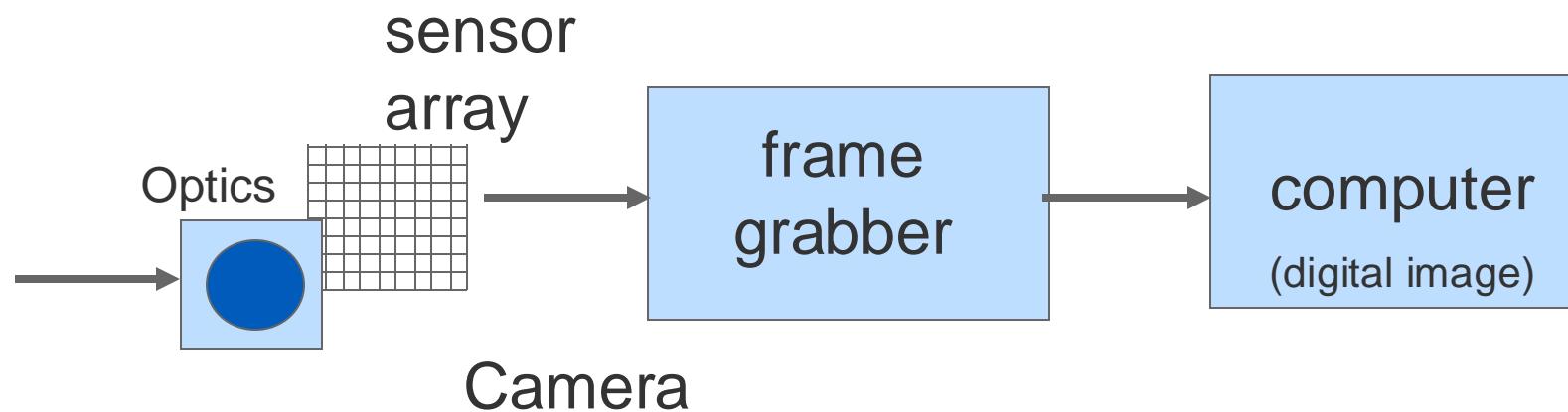
What is the film in a real camera?



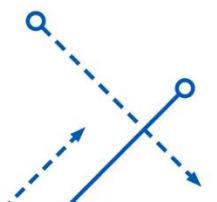
Digital cameras



- Film is a sensor (detector) array
- Each position contains a light sensitive diode that converts **photons** (light energy) to **electrons**.

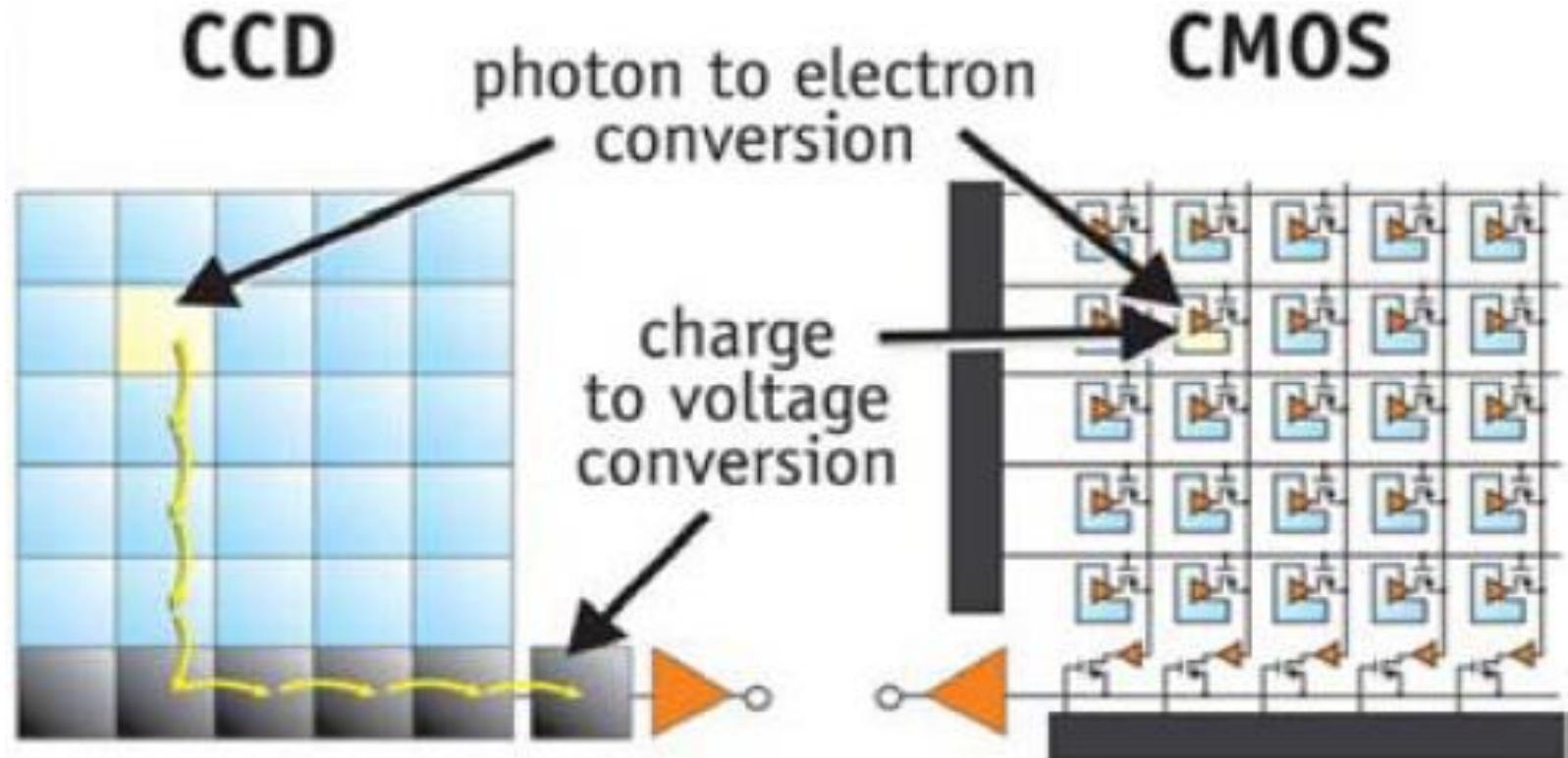


- Basic process:
 - photons hit a detector
 - the detector becomes charged
 - the charge is read out as brightness



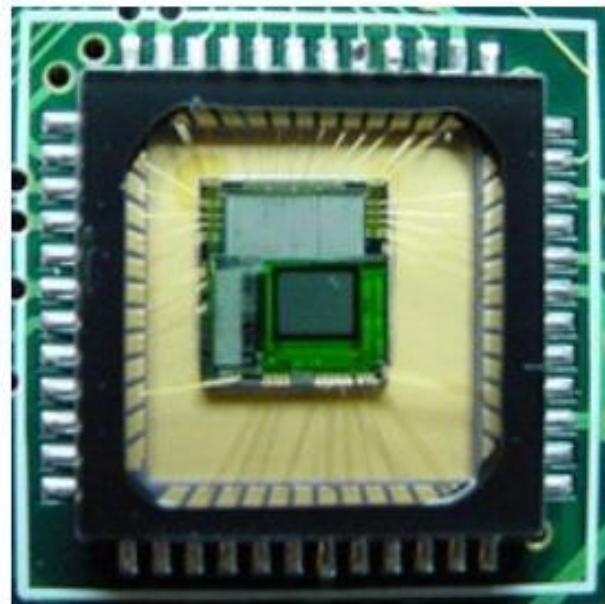
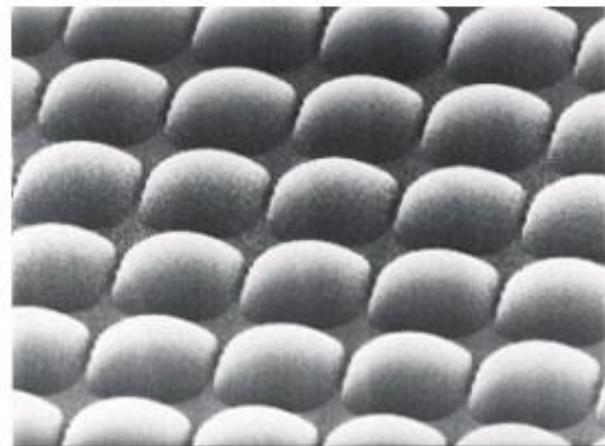
Digital Sensors

- CCD (charge couple devices): moves photogenerated charge from pixel to pixel and convert it to voltage at an output node.
- CMOS (complementary metal oxide semiconductor): convert charge to voltage inside each pixel.

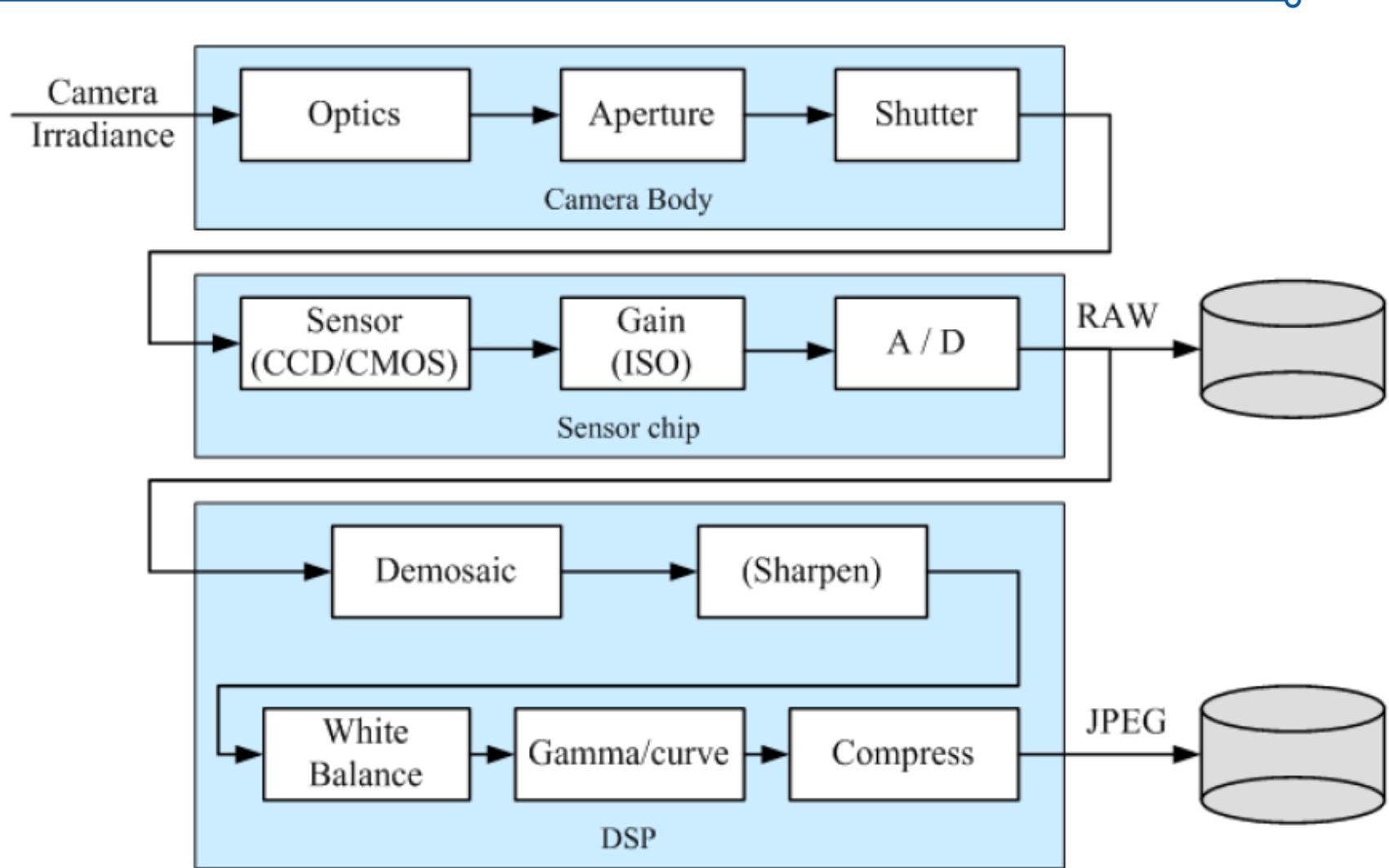


CCD vs CMOS

- Sensor types:
 - CCD (charge-coupled device)
 - most common
 - high sensitivity
 - high power
 - cannot be individually addressed
 - blooming
 - CMOS
 - simple to fabricate (cheap)
 - lower sensitivity, lower power
 - can be individually addressed

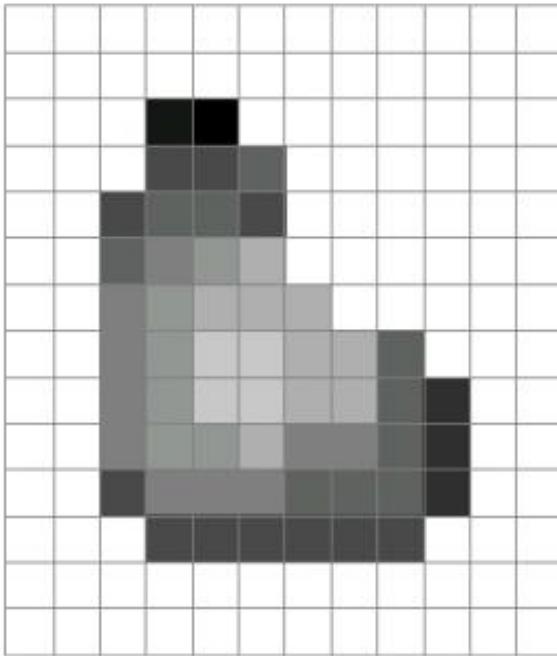
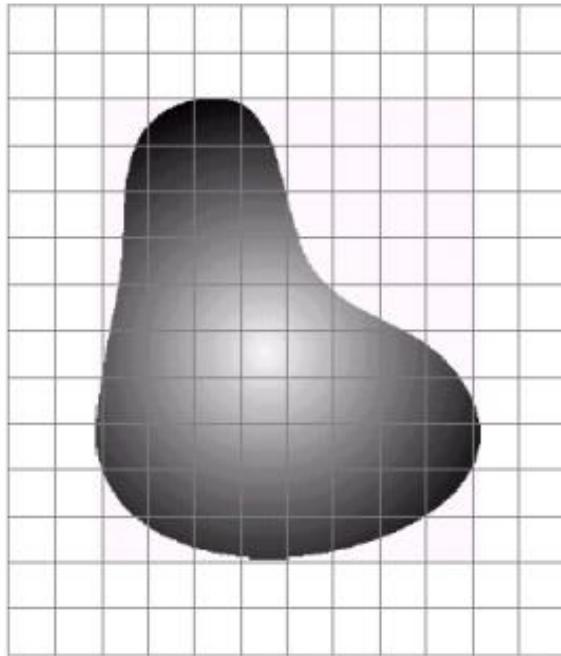


Camera



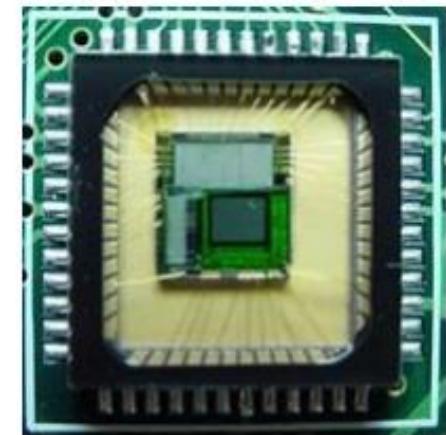
ISO referred to the sensitivity of film

Sampling and Quantization



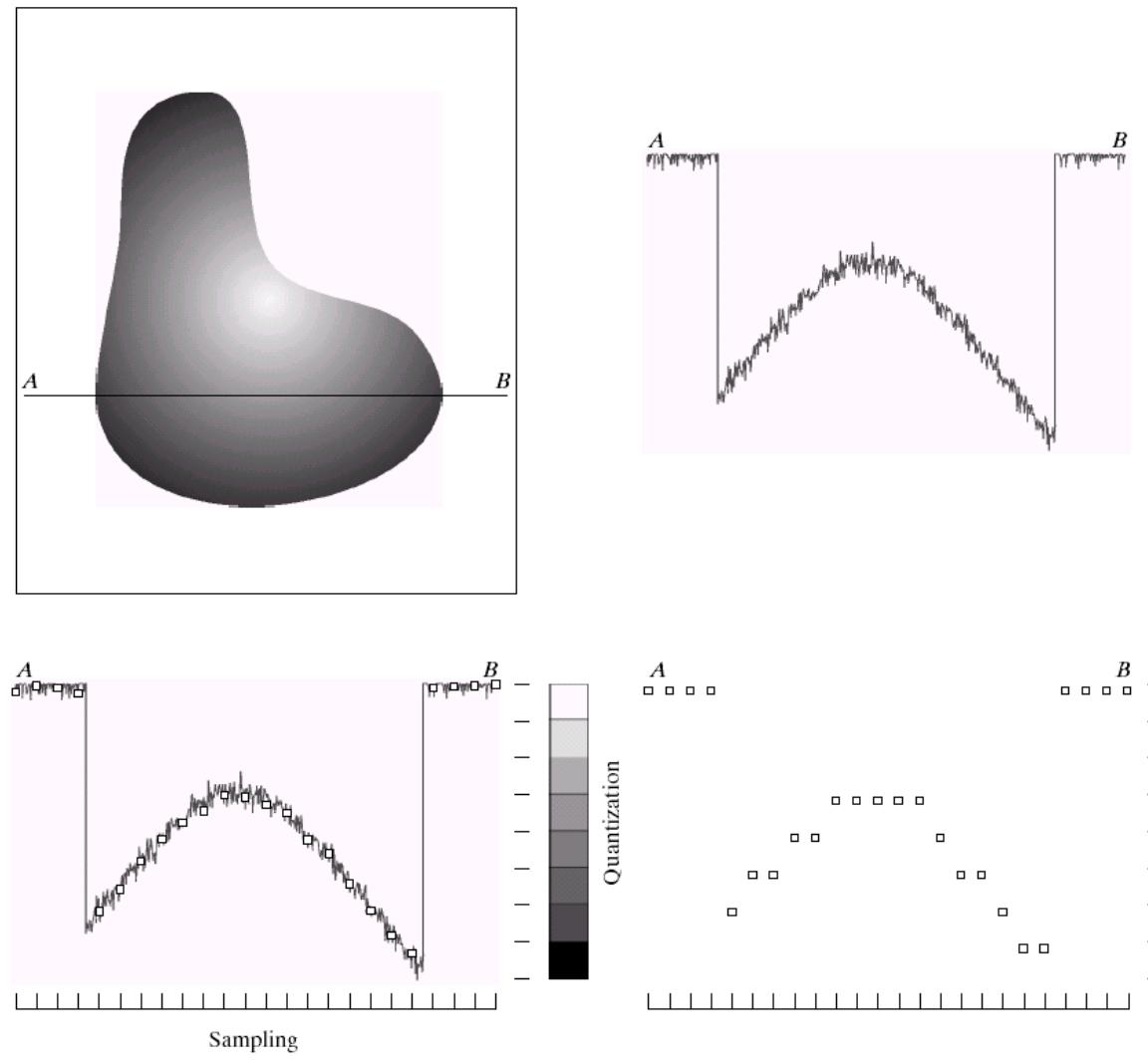
a b

FIGURE 2.17 (a) Continuous image projected onto a sensor array. (b) Result of image sampling and quantization.



CMOS sensor

Sampling and Quantization: 1D Case



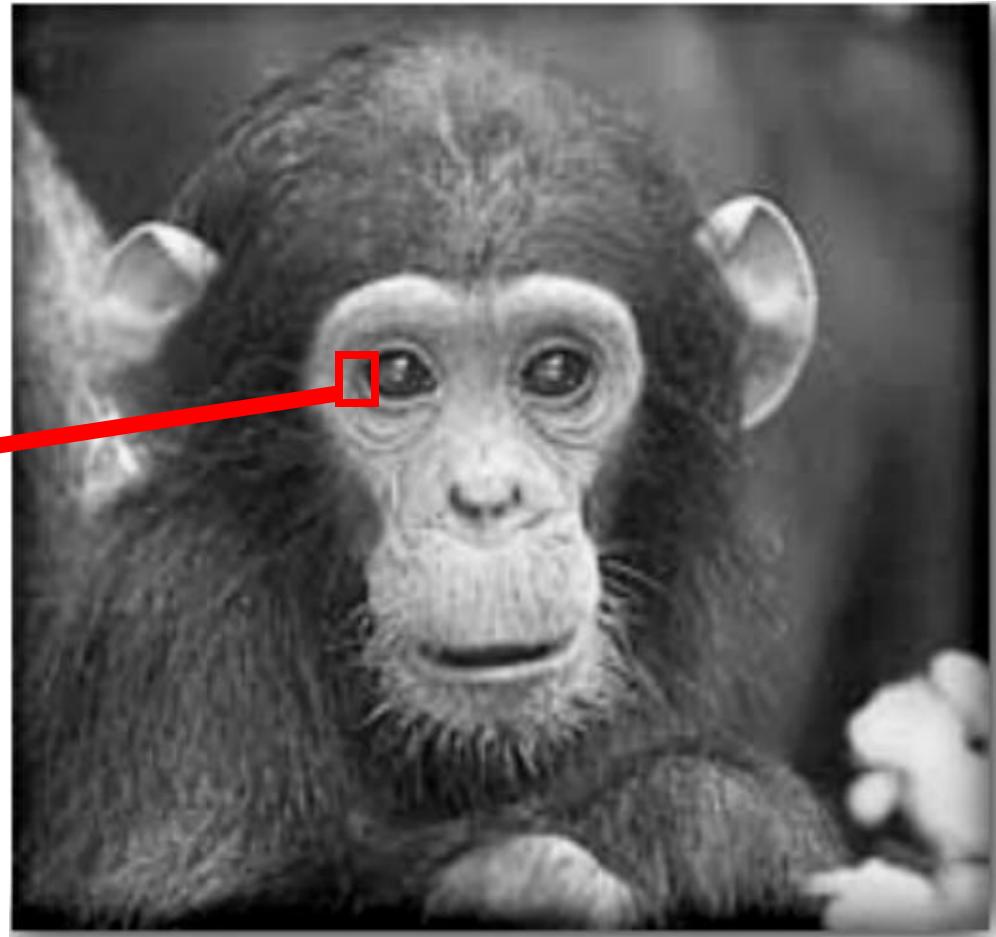
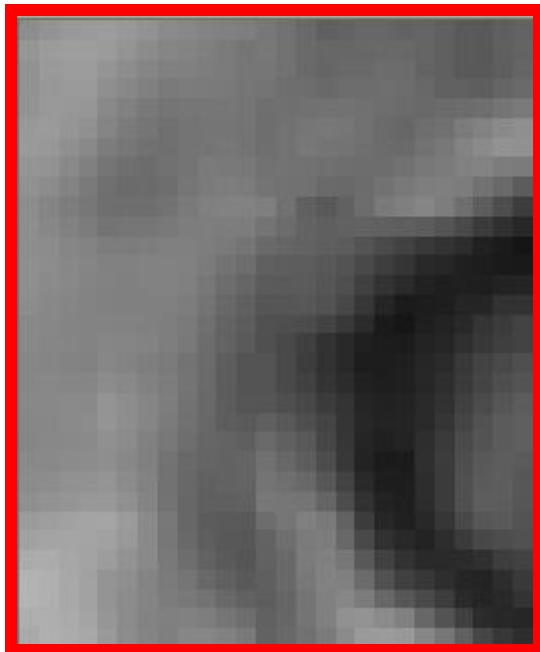
a b

c d

FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from *A* to *B* in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

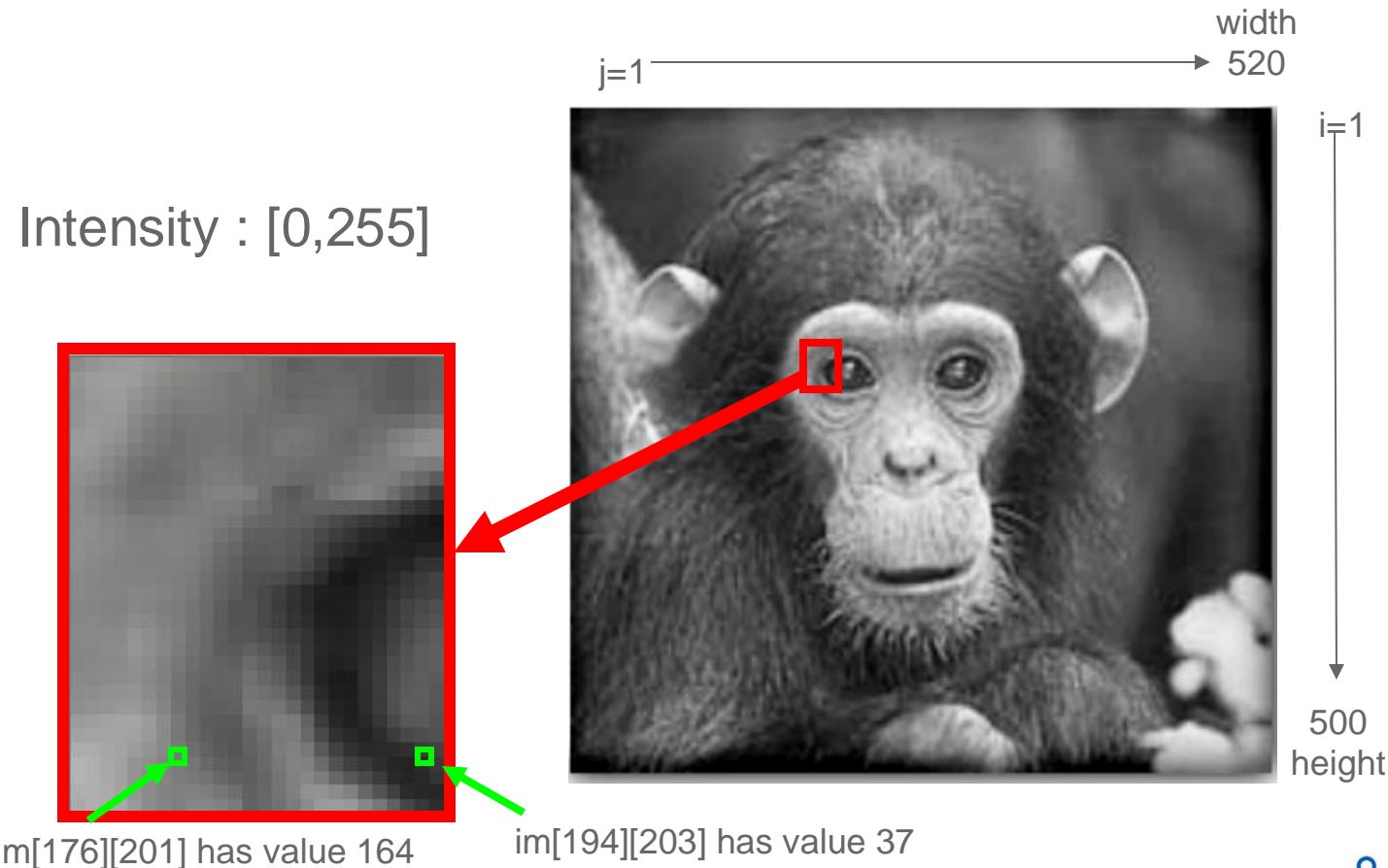
Digital Image

- Think of images as matrices taken from CCD array.



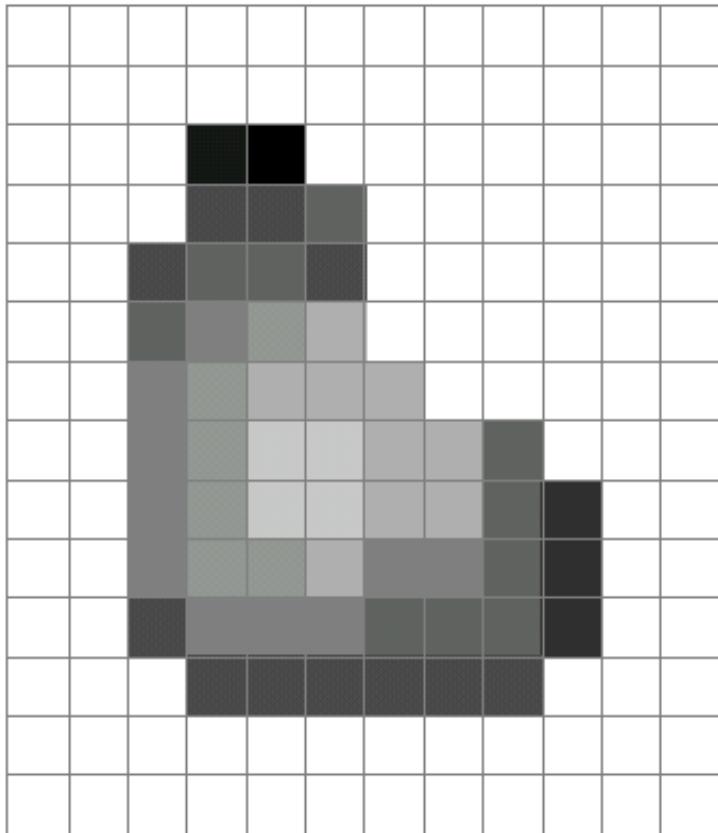
Digital Image

- Think of images as matrices taken from CCD array.



Digital image

- A grid (matrix) of intensity values.
- (common to use one byte per value: 0 = black, 255 = white)

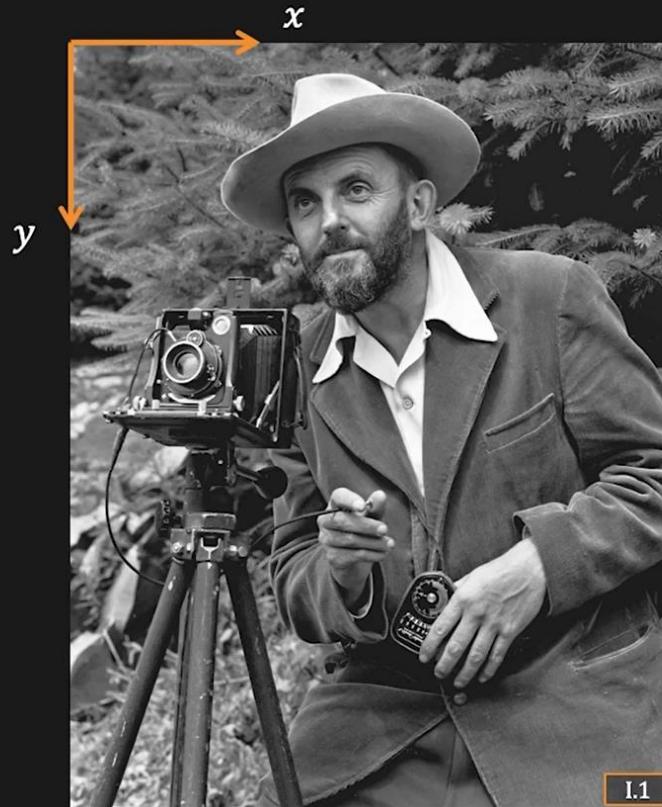


=

255	255	255	255	255	255	255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255	255	255	255	255	255	255
255	255	255	20	0	255	255	255	255	255	255	255	255	255
255	255	255	75	75	75	255	255	255	255	255	255	255	255
255	255	75	95	95	75	255	255	255	255	255	255	255	255
255	255	96	127	145	175	255	255	255	255	255	255	255	255
255	255	127	145	175	175	175	255	255	255	255	255	255	255
255	255	127	145	200	200	175	175	95	255	255	255	255	255
255	255	127	145	200	200	175	175	95	47	255	255	255	255
255	255	127	145	145	175	127	127	95	47	255	255	255	255
255	255	74	127	127	127	95	95	95	47	255	255	255	255
255	255	255	74	74	74	74	74	74	74	255	255	255	255
255	255	255	255	255	255	255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255	255	255	255	255	255	255

Image representation

Image as a Function



$f(x, y)$ is the image intensity at position (x, y)

Image representation

- A (grayscale) image as a **function**, f , from \mathbb{R}^2 to \mathbb{R} :
 - $f(x, y)$ gives the **intensity** at position (x, y) .
 - A **digital image** is a discrete (**sampled, quantized**) version of this function.

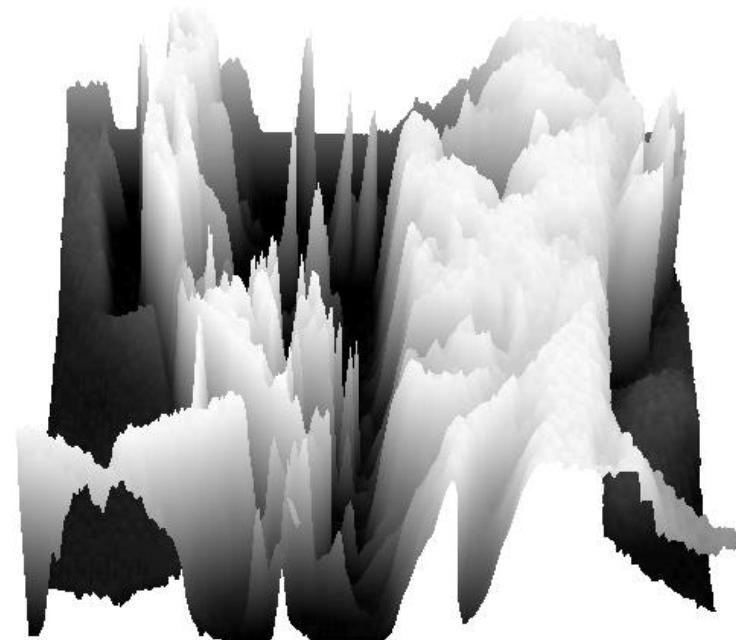
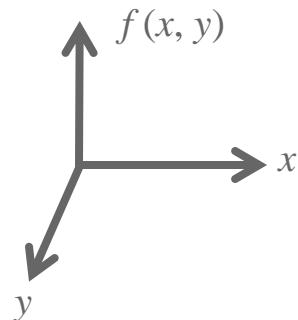


Image Resolution

- Sensor: size of real-world scene element that maps to a single pixel
- Influences what analysis is feasible, affects best representation choice.

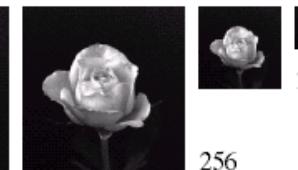
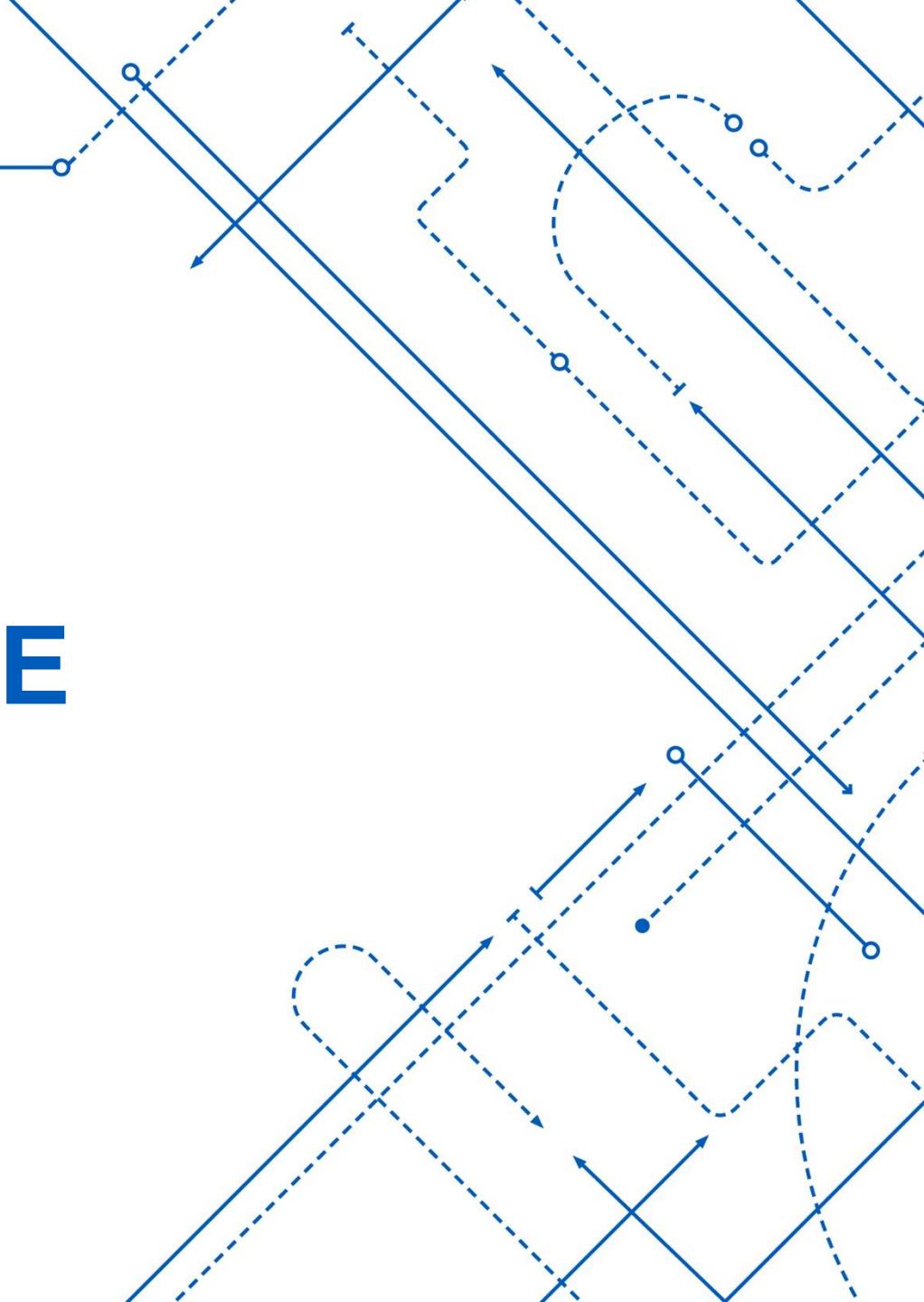


FIGURE 2.19 A 1024×1024 , 8-bit image subsampled down to size 32×32 pixels. The number of allowable gray levels was kept at 256.



COLOR SPACE



What is Color?

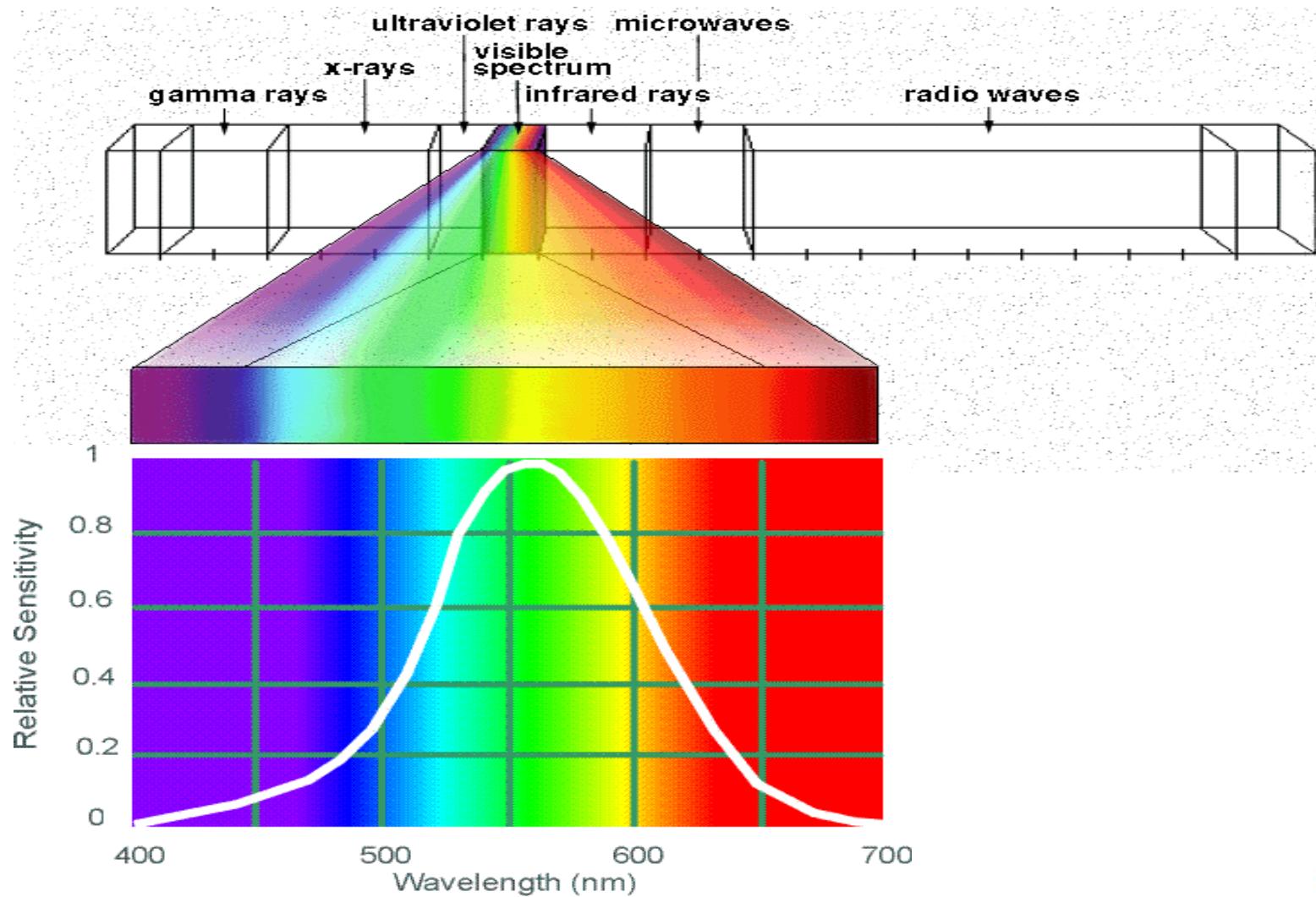
Human Response to different wavelengths

Visible light:



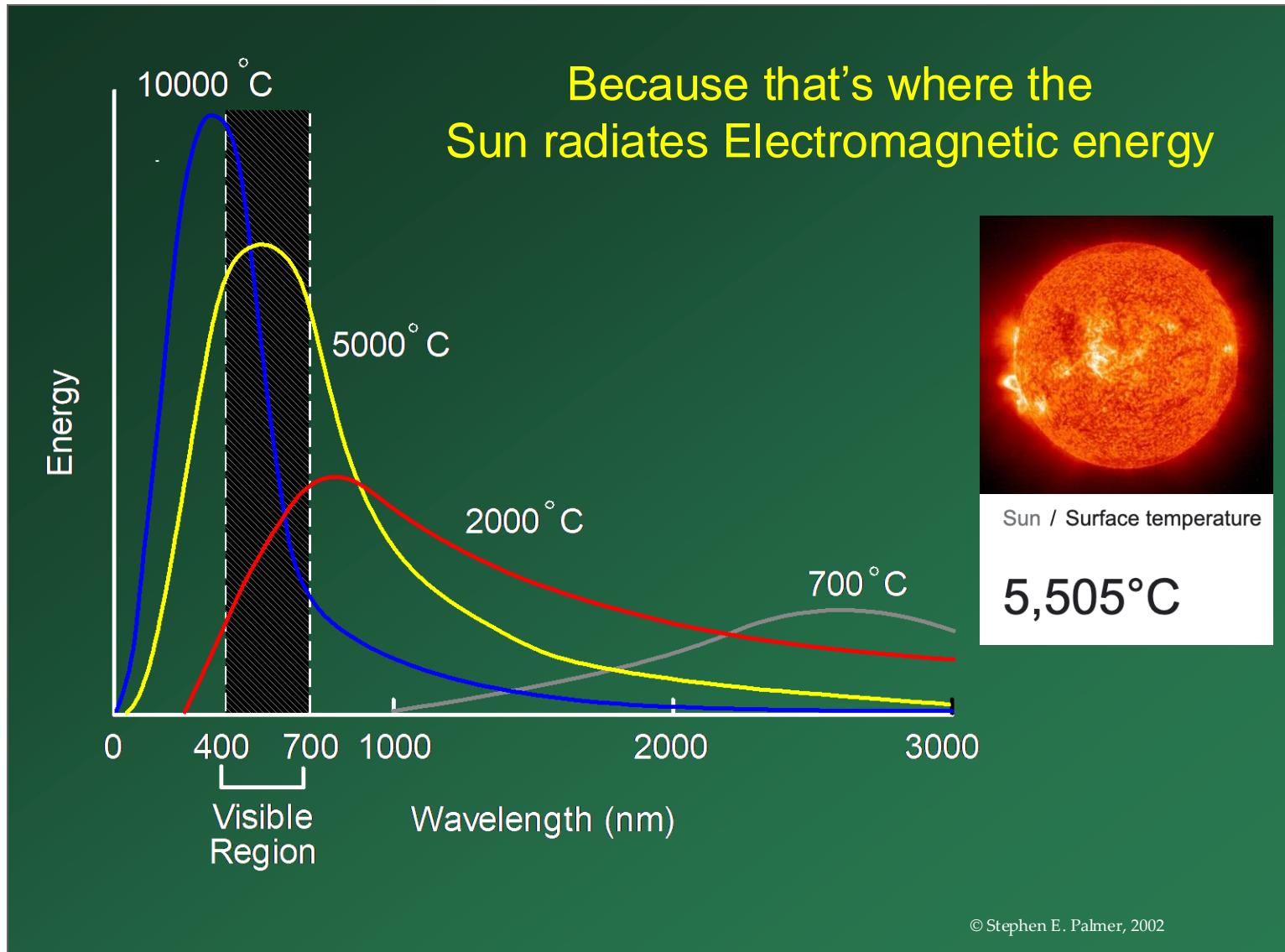
electromagnetic wave

Electromagnetic Spectrum



Human Luminance Sensitivity Function

Why do we see light of these wavelengths



Mixing of color



Human Sensation of nearly all colors can be produced using 3 wavelengths!

$$(\lambda_r, \lambda_g, \lambda_b) = (650, 530, 410) \text{ nm}$$

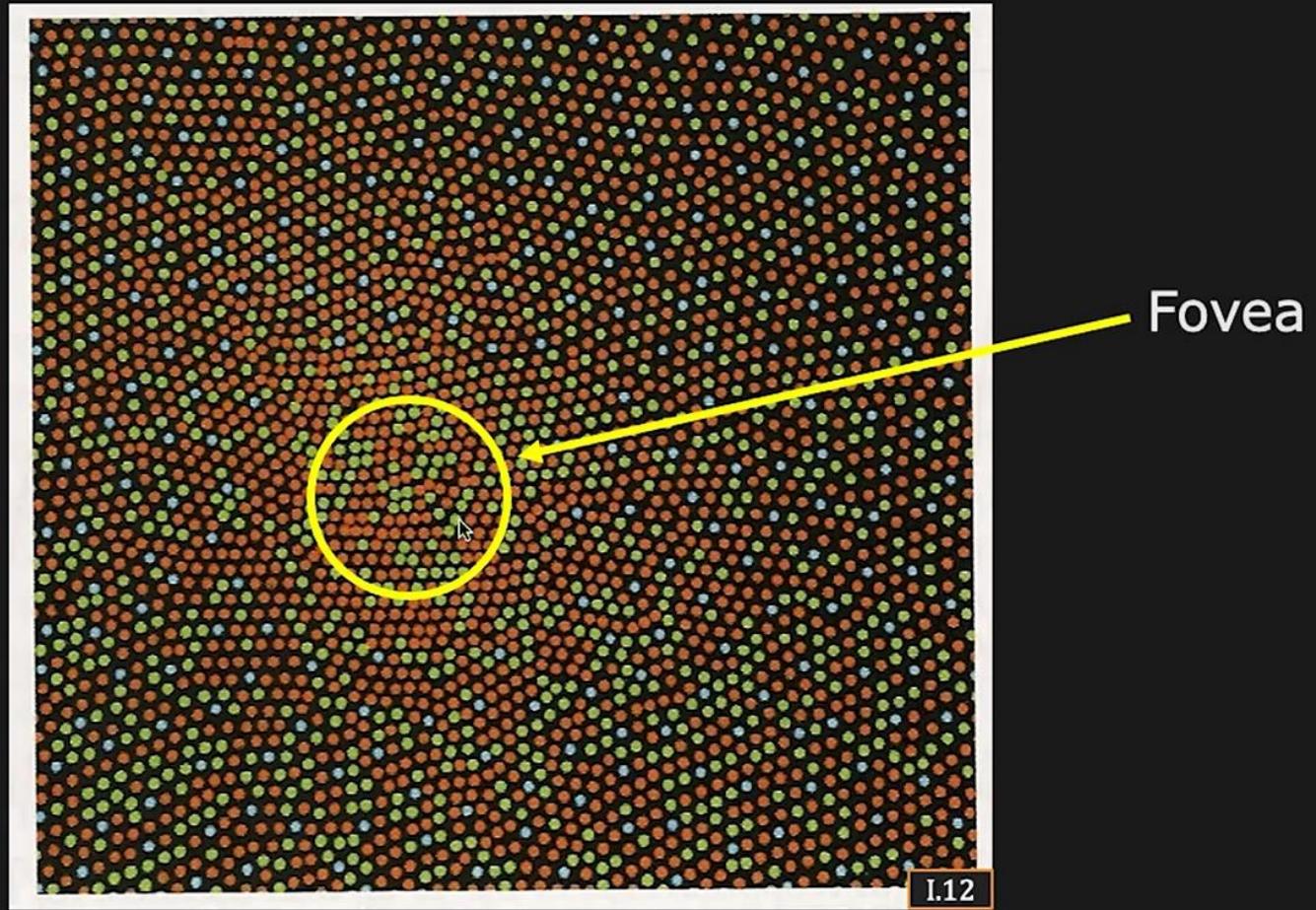
Hence, cameras and displays often use 3 filters:

(red, green, blue)

Young's Experiment on Color Mixture

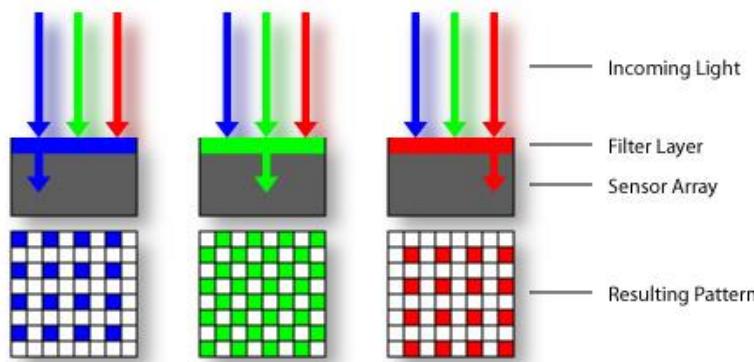
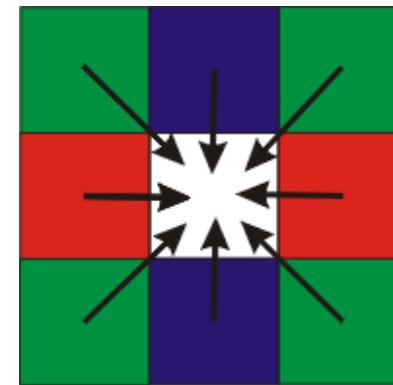
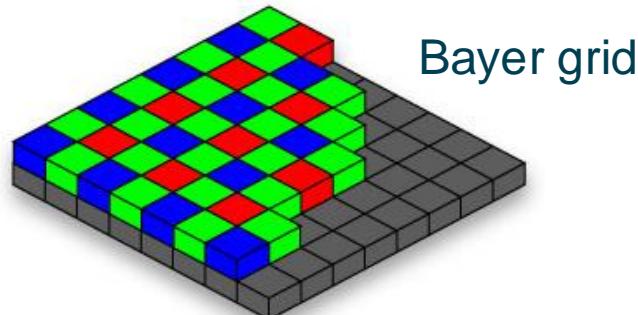
Distribution of Cones in Human Retina

Three types of cones for sensing **red**, **green**, **blue**



Color sensing in digital cameras

- Estimate missing components from neighboring values (demosaicing)



Anatomy of the Active Pixel Sensor Photodiode

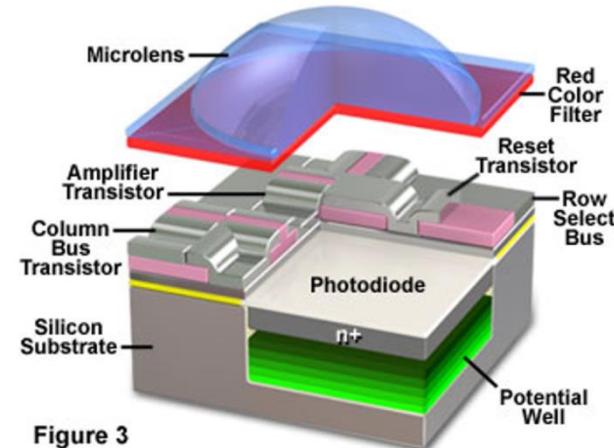
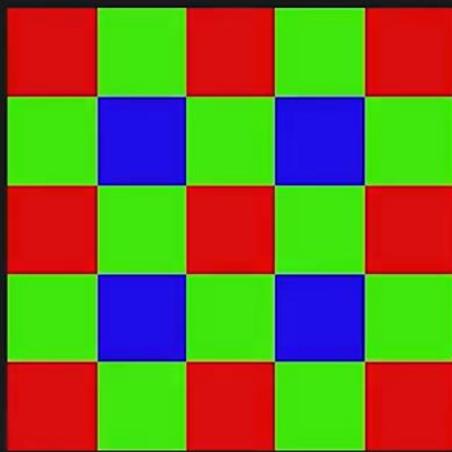
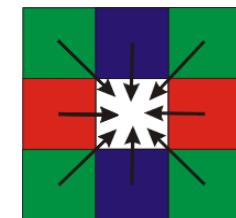
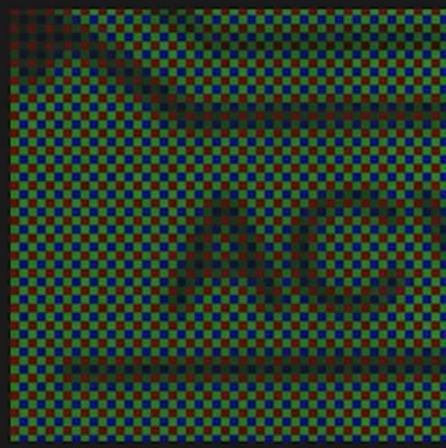


Figure 3

Sensing Color Using Color Mosaic



Bayer Pattern
(Color Filter Mosaic)



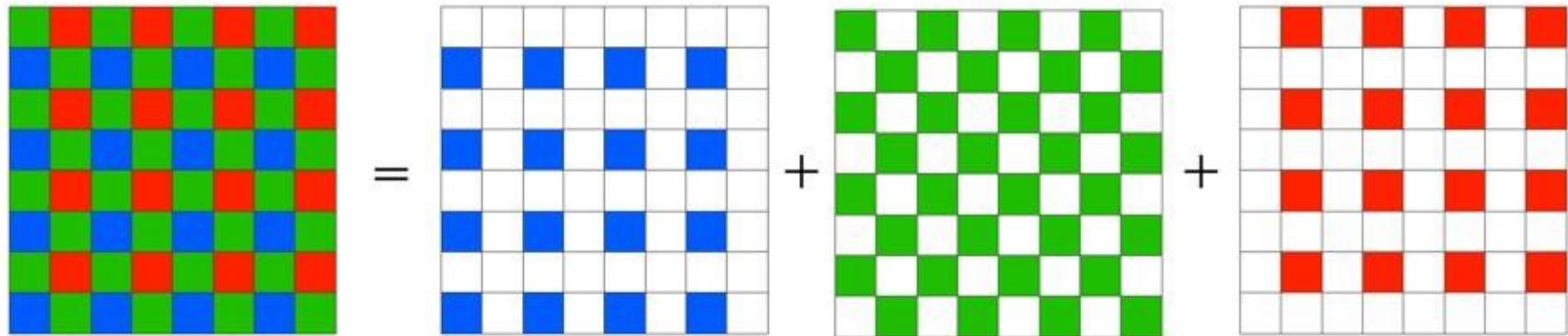
Raw Image



Interpolated Image

Color Filled in by Interpolation (**Demosaicing**)

Bayer pattern



- But why are there more green filters?
- Because human vision is more sensitive to green.
- So, the ratio is 50% green, 25% red, and 25% blue.

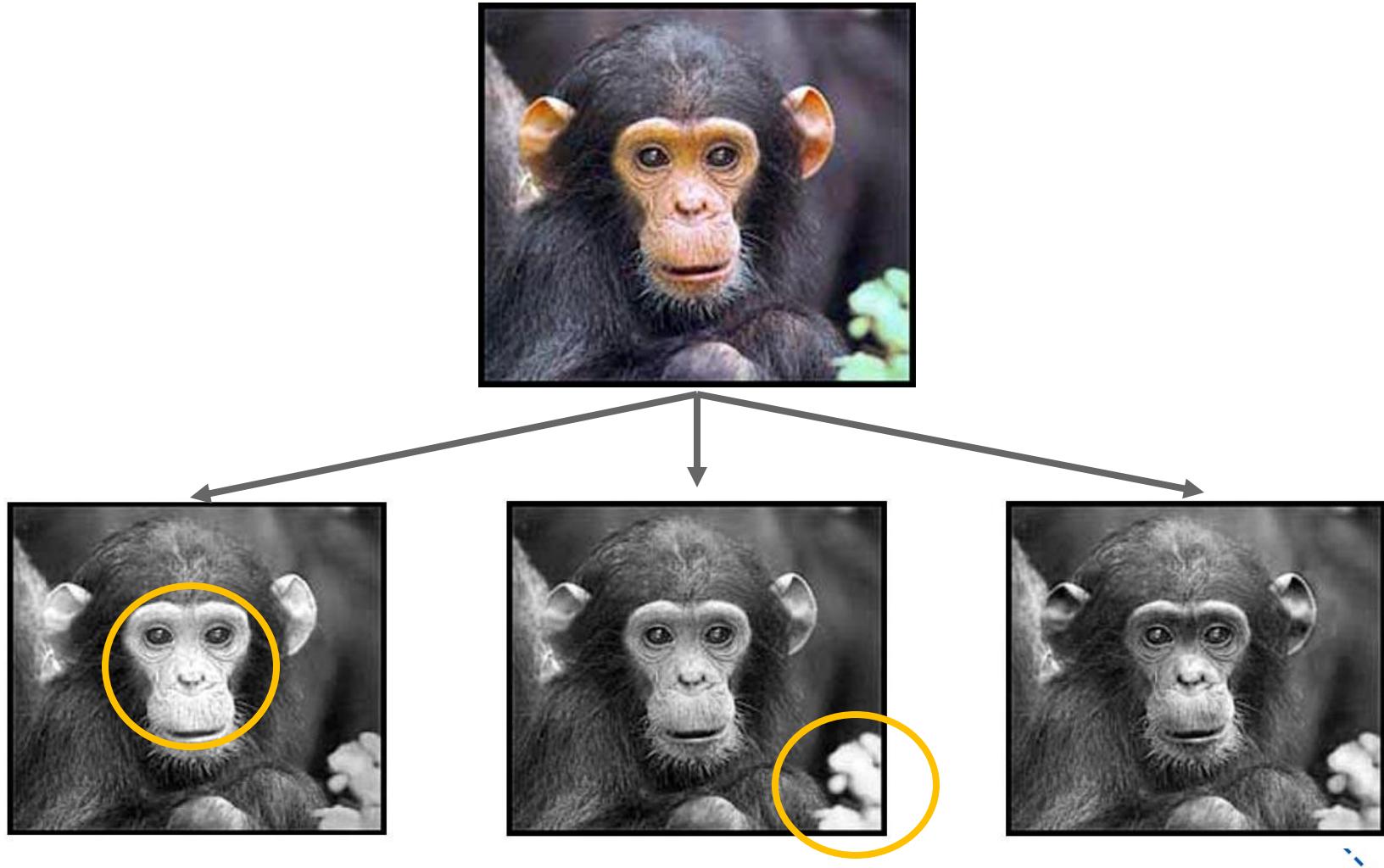
True vs interpolated values of RGB

G	R	G	R
B	G	B	G
G	R	G	R
B	G	B	G

rGb	Rgb	rGb	Rgb
rgB	rGb	rgB	rGb
rGb	Rgb	rGb	Rgb
rgB	rGb	rgB	rGb

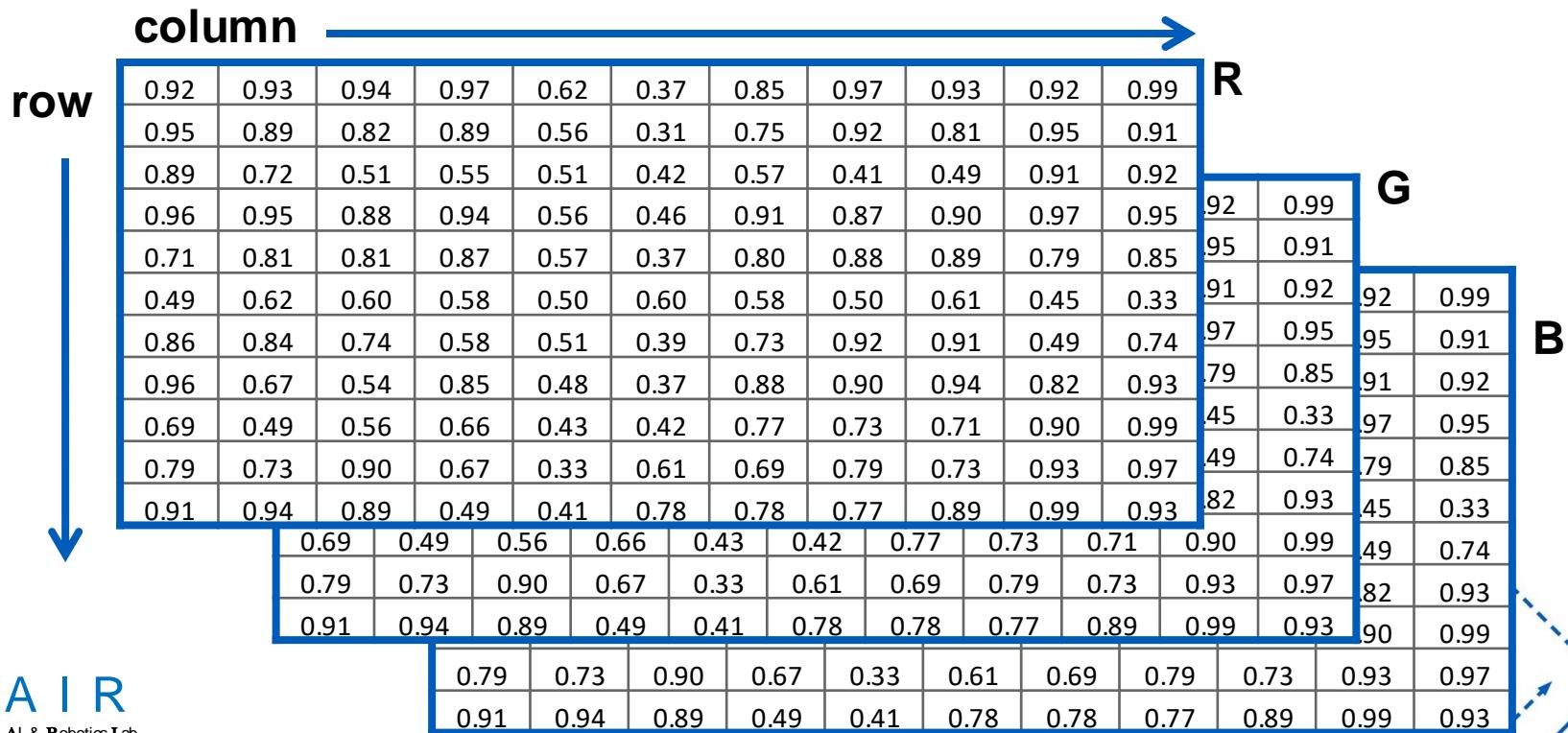
Digital Color images

- RGB color space (3 Channels)



Images in Python (PyTorch)

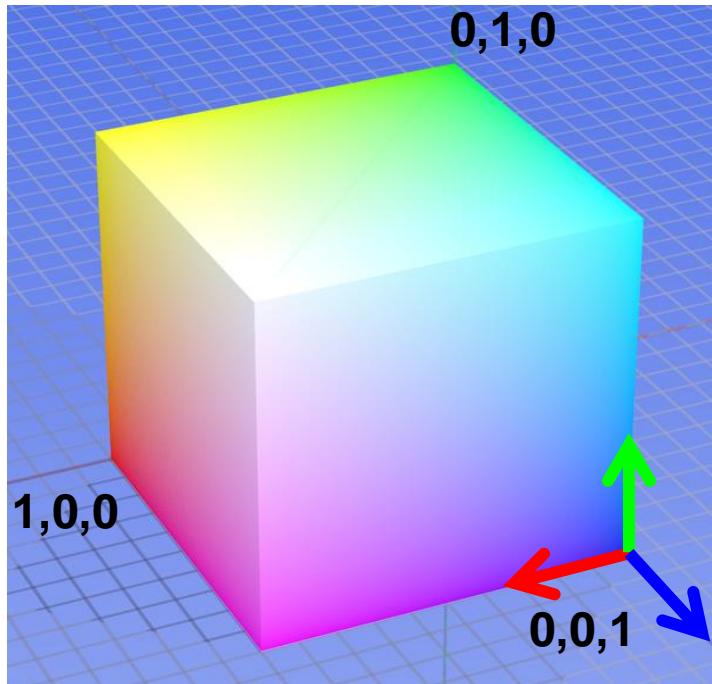
- Images represented as a matrix (Tensor)
- Suppose we have a $N \times M$ RGB image called “I”
 - $I[0,0,0]$ = top-left pixel value in R-channel
 - $I[b, r, c] = c$ pixels to right, r pixels down, in the b^{th} channel
 - $I[B, N-1, M-1]$ = bottom-right pixel in B-channel



Color spaces: RGB

- Drawbacks

- Strongly correlated channels
- Non-perceptual



Default color space



G ($R=0, B=0$)

B ($R=0, G=0$)

Any other color space?

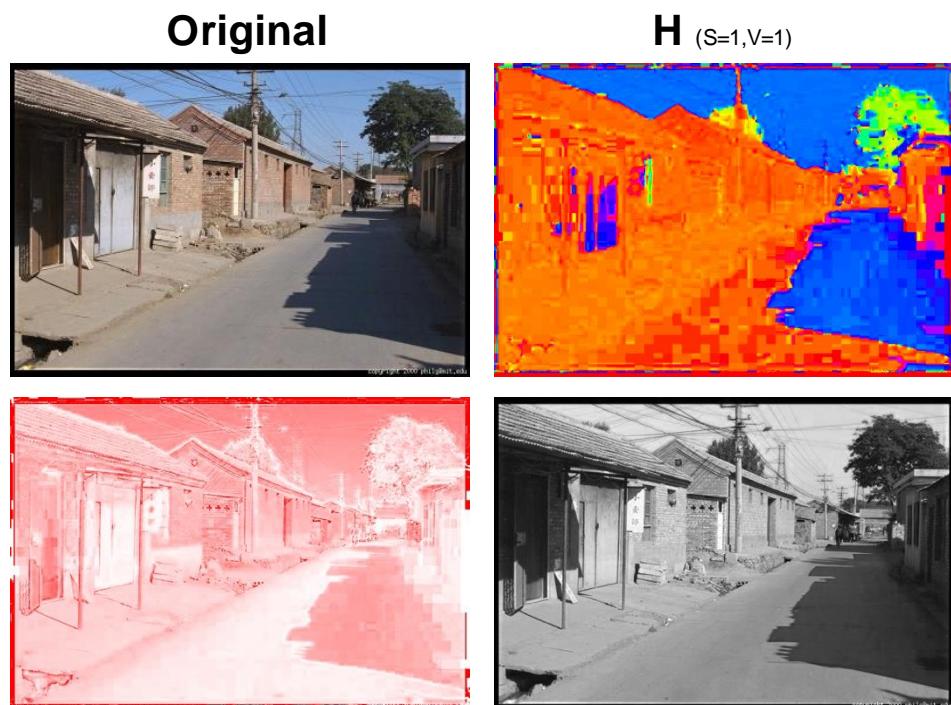
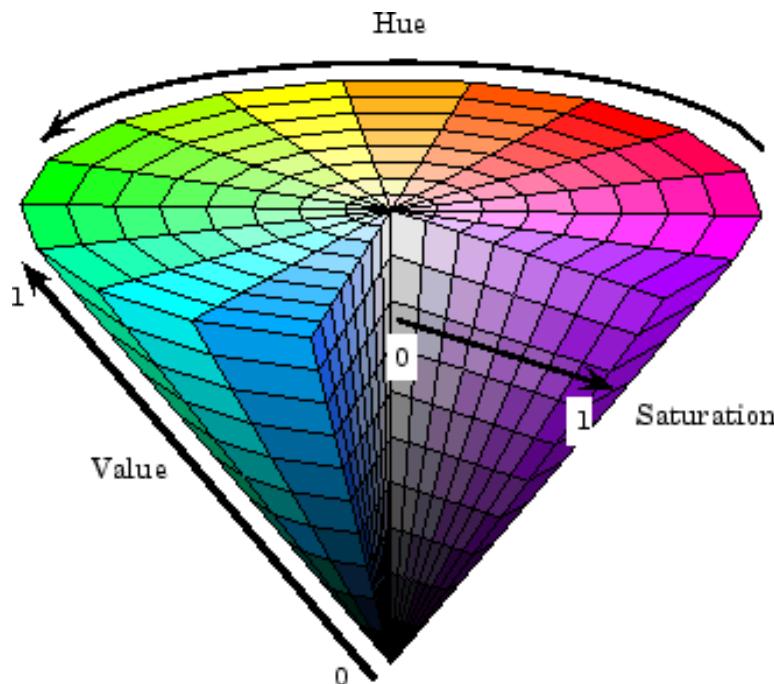


http://en.wikipedia.org/wiki/File:RGB_illumination.jpg

Color spaces: HSV

- **Hue, Saturation, Value (Brightness)**: how colors appear under light.
- **Saturation** (photography): the intensity of a color, expressed as the degree to which it differs from gray.

Intuitive Color Space



S ($H=1, V=1$)

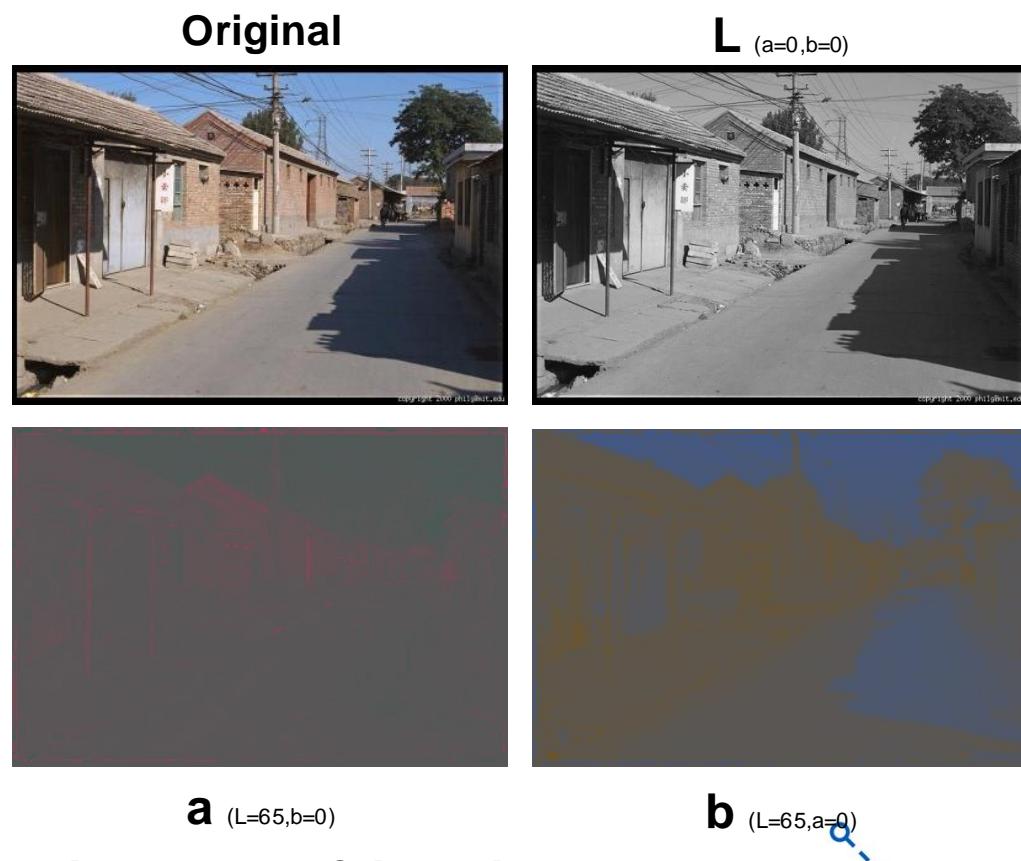
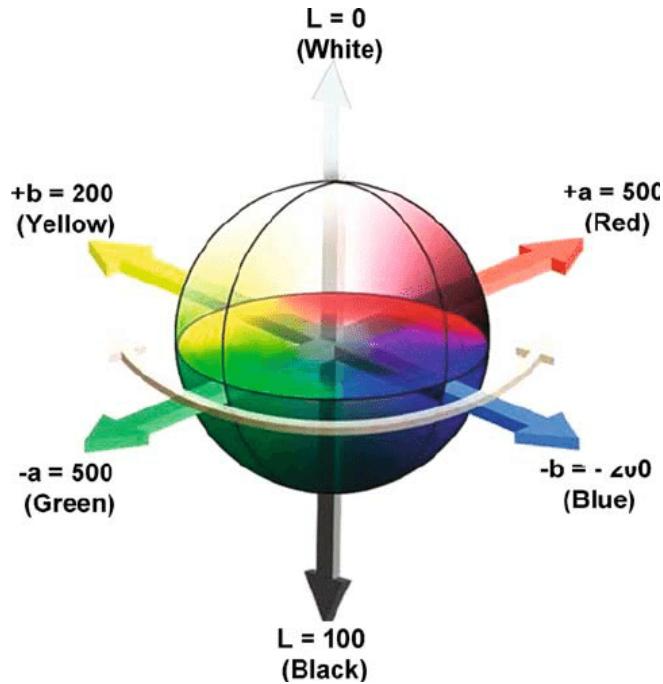
V ($H=1, S=0$)

Friendly for human intuition

Color spaces: L*a*b*

- L is for lightness. It goes from 0 to 100, shows contrast between black and grays.
- a is red (+) to green (-).
- b is yellow (+) to blue (-).

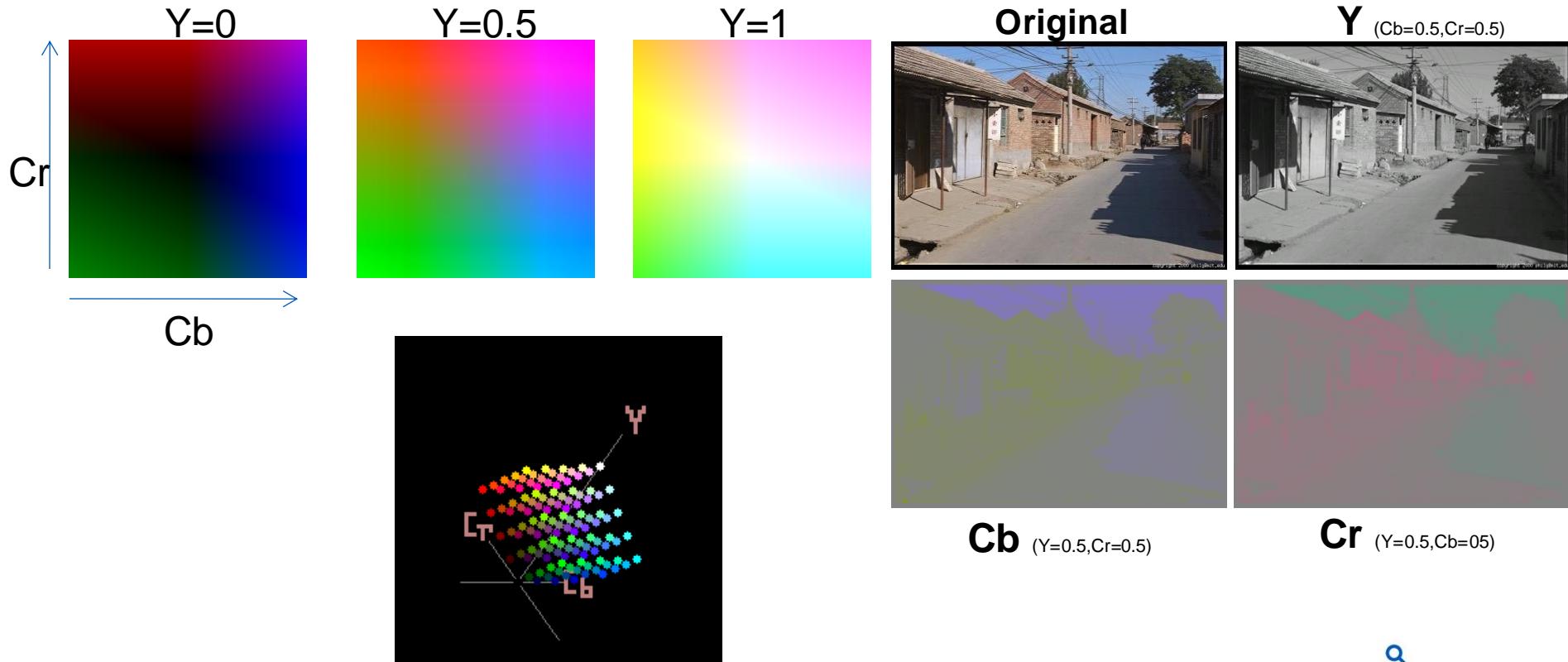
“Perceptually uniform”* color space



If a change of length in any direction X of the color space is perceived by a human as the same change.

Color spaces: YCbCr

- **Y** is for luminance.
- **C_b** is difference between **blue** and a **luminance** component
- **C_r** is difference between **red** and a **luminance** component.



Think and Answer

If you had to choose, would you rather go without luminance or chrominance?

Most information in intensity



Only color shown – constant intensity

Most information in intensity



copyright 2000 philg@mit.edu

Only intensity shown – constant color

Most information in intensity



copyright 2000 philg@mit.edu

Original image

Summary for Important Concepts

- Optical Sensor
 - Image representation, resolution, sampling & quantization
- Coloring
 - RGB, HSV



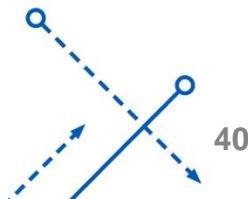
IMAGE FORMATION

Image Warping



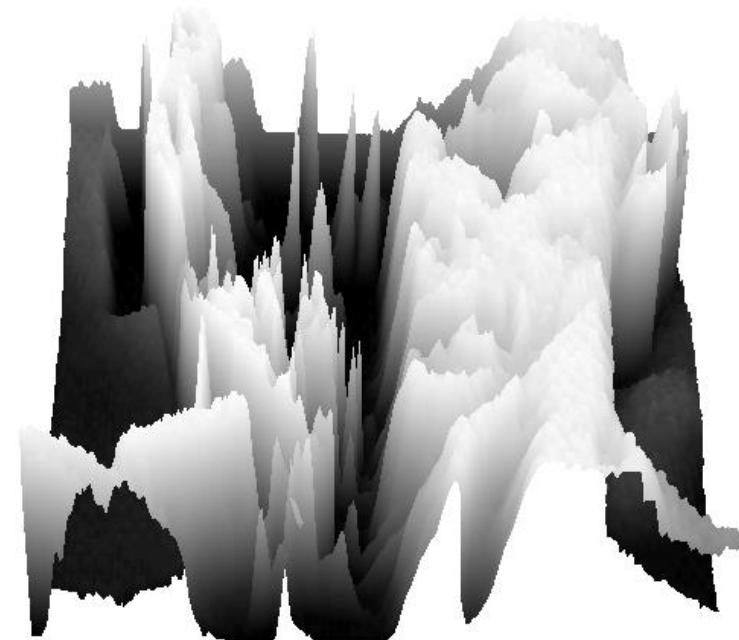
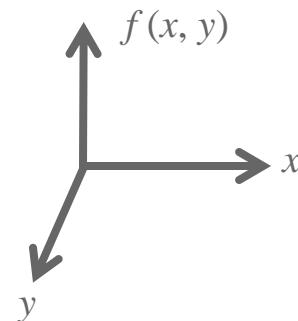
Content

- Image Warping
 - Mirror, rotation, translation, scaling, skewing
 - Euclidean, similarity, affine, projective (homograph)
- Optics
 - Lens, aperture, focal length.
 - Field of view, depth of view.
 - Vignetting, chromatic/spherical aberration
 - Pincushion, barrel distortion



Recap: Image representation

- A (grayscale) image as a **function**, f , from \mathbb{R}^2 to \mathbb{R} :
 - $f(x, y)$ gives the **intensity** at position (x, y) .
 - A **digital image** is a discrete (**sampled, quantized**) version of this function.



Warping (Transformation)

image warping: change **domain** of image

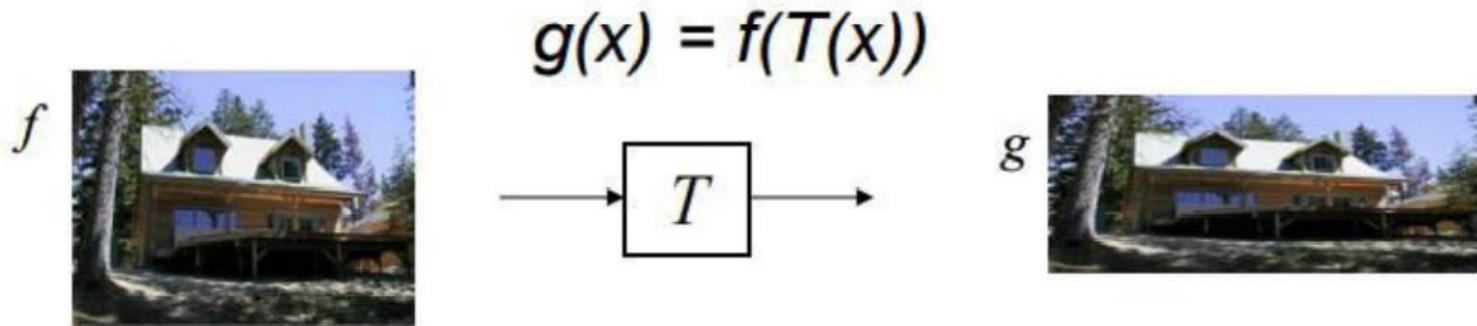
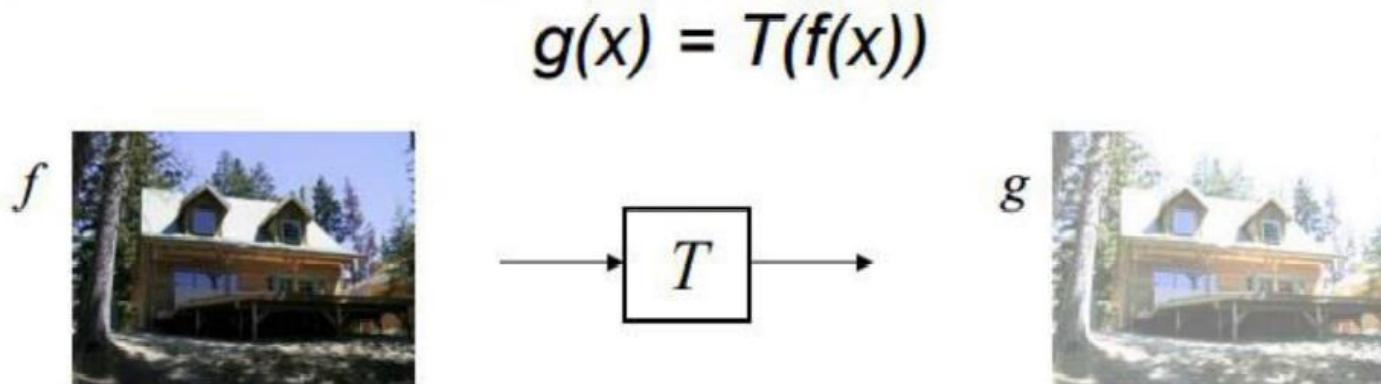


image filtering: change **range** of image (Next Week)



2D Planer Transformation

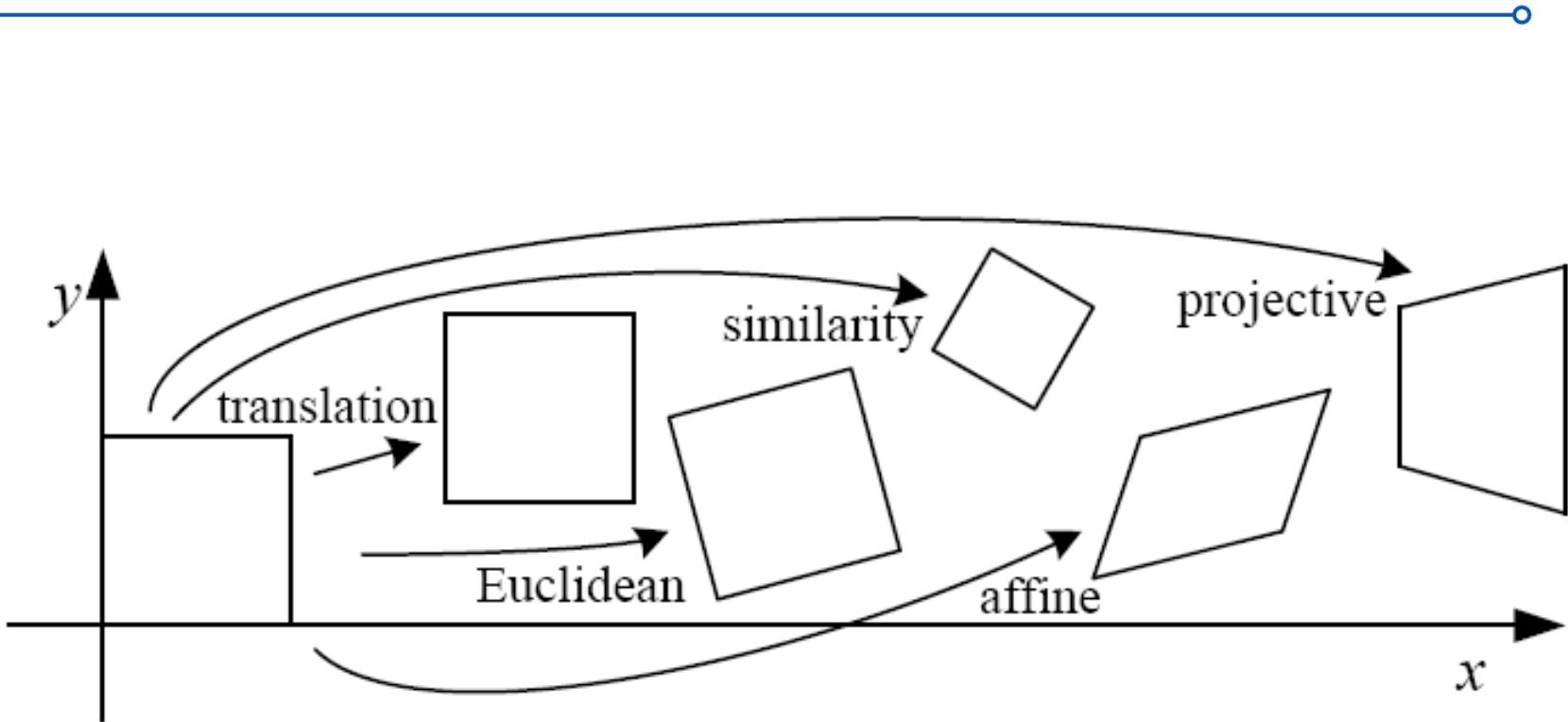


Figure 2.4: Basic set of 2D planar transformations

Image Warping

Global Warping/Transformation



Translation



Rotation



Scaling and Aspect

$$g(x, y) = f(T(x, y))$$



Affine

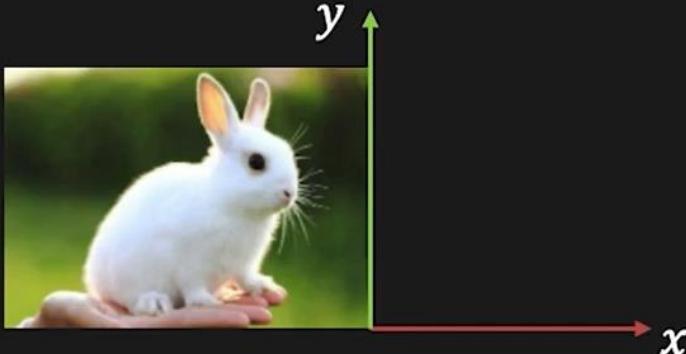


Projective



Barrel

Mirror



Mirror about Y-axis:

$$x_2 = -x_1$$

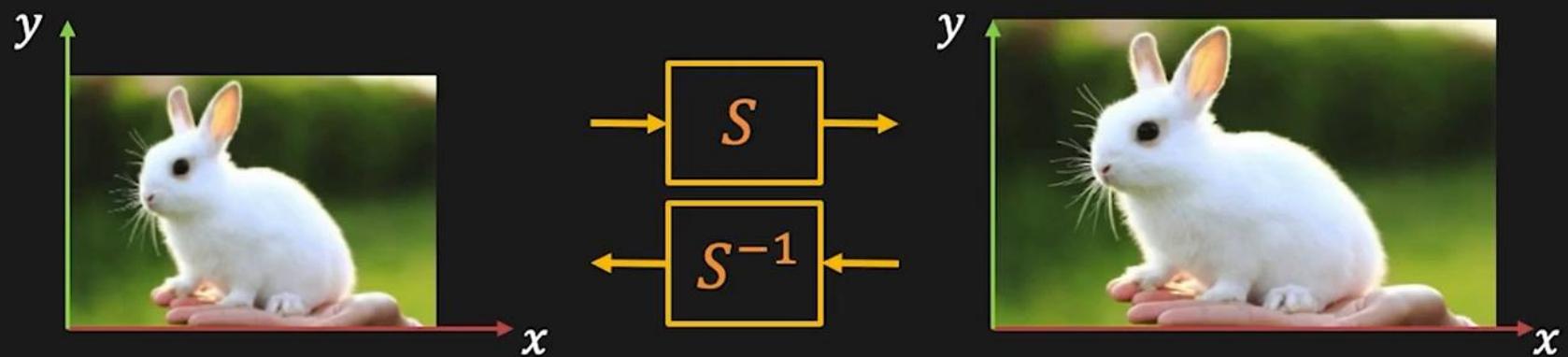
$$y_2 = y_1$$

Mirror about line $y = x$:

$$x_2 = y_1$$

$$y_2 = x_1$$

Scaling (Stretching or Squashing)



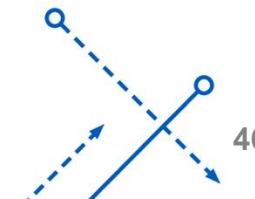
Forward:

$$x_2 = ax_1 \quad y_2 = by_1$$

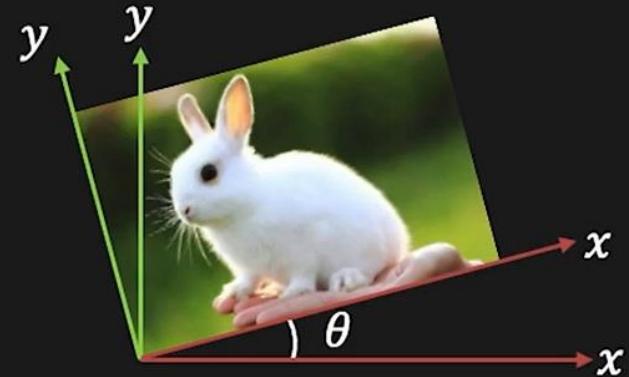
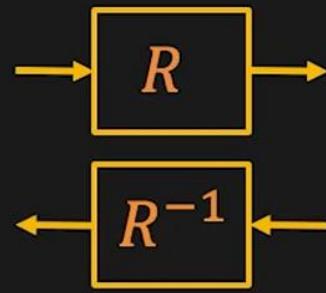
Inverse:

$$x_1 = \frac{1}{a}x_2 \quad y_1 = \frac{1}{b}y_2$$

- Scaling a coordinate means multiplying each component by a scalar
- Uniform scaling means this scalar is the same for all components.



Rotation



Forward:

$$x_2 = x_1 \cos\theta - y_1 \sin\theta$$

$$y_2 = x_1 \sin\theta + y_1 \cos\theta$$

Inverse:

$$x_1 = x_2 \cos\theta + y_2 \sin\theta$$

$$y_1 = -x_2 \sin\theta + y_2 \cos\theta$$

How to derive rotation matrix?

Let $r = |\mathbf{V}|$. Then, we have the relations:

$$v_x = r \cos \alpha$$

$$v'_x = r \cos(\alpha + \theta)$$

$$v_y = r \sin \alpha$$

$$v'_y = r \sin(\alpha + \theta)$$

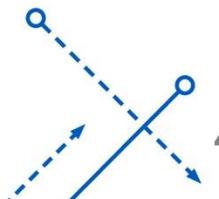
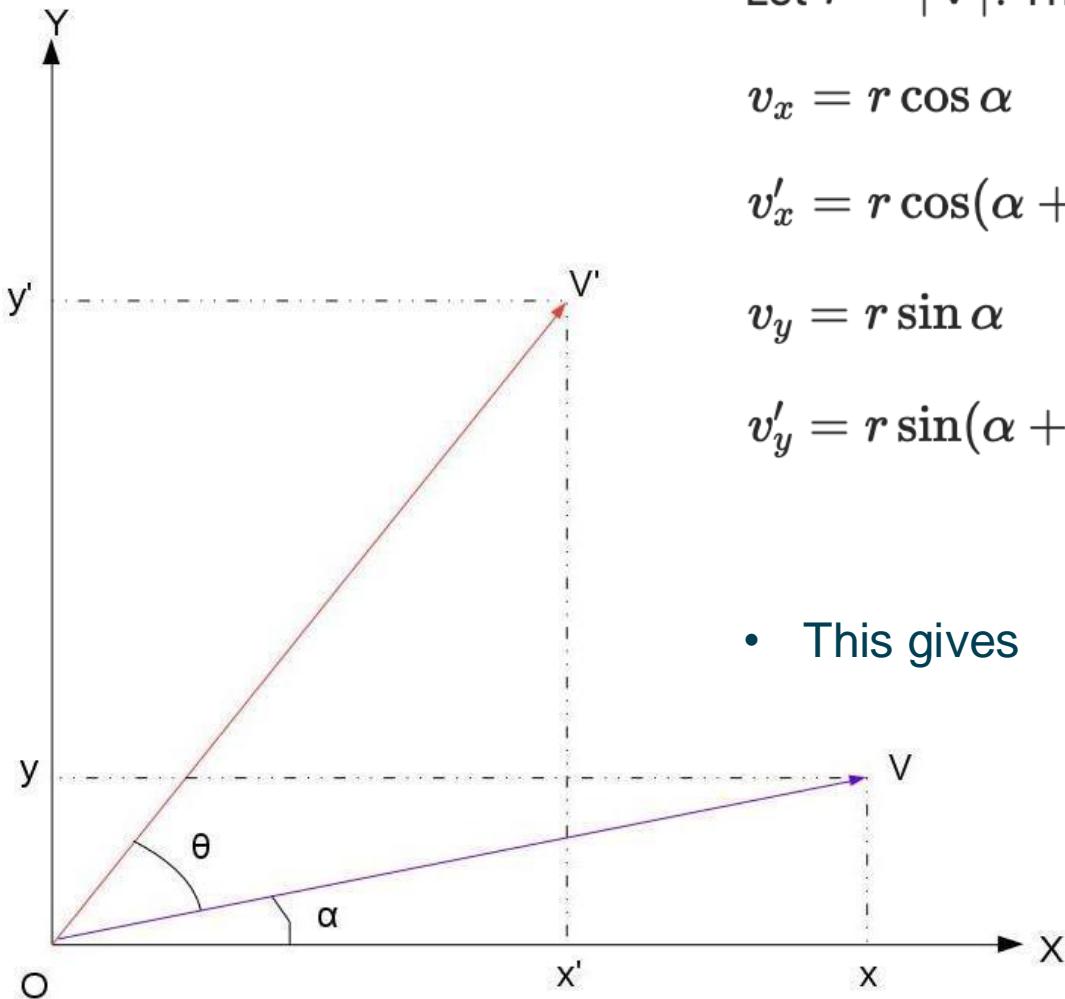


$$v'_x = v_x \cos \theta - v_y \sin \theta$$

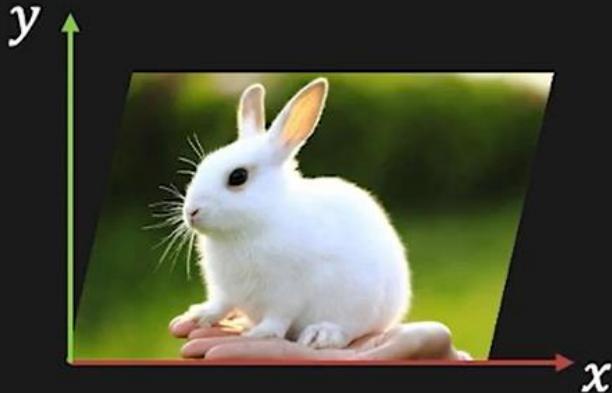
$$v'_y = v_x \sin \theta + v_y \cos \theta$$

- This gives

$$\begin{pmatrix} v'_x \\ v'_y \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} v_x \\ v_y \end{pmatrix}$$



Skew



Horizontal Skew:

$$x_2 = x_1 + m_x y_1$$

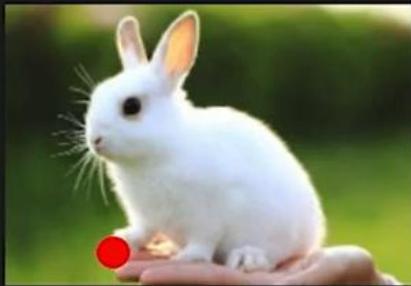
$$y_2 = y_1$$

Vertical Skew:

$$x_2 = x_1$$

$$y_2 = m_y x_1 + y_1$$

Linear Transformations



$$\mathbf{p}_1 = (x_1, y_1)$$

$$\mathbf{p}_2 = (x_2, y_2)$$

T can be represented by a matrix.

$$\mathbf{p}_2 = T\mathbf{p}_1$$

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = T \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

2 x 2 Matrix Transformations

Any transformation of the form:

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

- Origin maps to the origin
- Lines map to lines
- Parallel lines remain parallel
- Closed under composition

Scaling, Rotation, Skew, Translation

$$\begin{bmatrix} \tilde{x}_2 \\ \tilde{y}_2 \\ \tilde{z}_2 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}$$

Scaling

$$\begin{bmatrix} \tilde{x}_2 \\ \tilde{y}_2 \\ \tilde{z}_2 \end{bmatrix} = \begin{bmatrix} 1 & m_x & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}$$

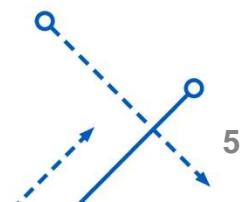
Skew

$$\begin{bmatrix} \tilde{x}_2 \\ \tilde{y}_2 \\ \tilde{z}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}$$

Translation

$$\begin{bmatrix} \tilde{x}_2 \\ \tilde{y}_2 \\ \tilde{z}_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}$$

Rotation



Affine Transformation

Any transformation of the form:

$$\begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} \equiv \begin{bmatrix} \tilde{x}_2 \\ \tilde{y}_2 \\ \tilde{z}_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{x}_1 \\ \tilde{y}_1 \\ \tilde{z}_1 \end{bmatrix}$$



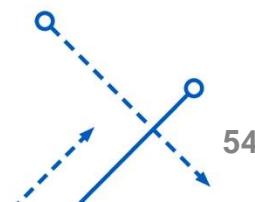
parallelogram

Affine Transformation

Any transformation of the form:

$$\begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} \equiv \begin{bmatrix} \tilde{x}_2 \\ \tilde{y}_2 \\ \tilde{z}_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{x}_1 \\ \tilde{y}_1 \\ \tilde{z}_1 \end{bmatrix}$$

- Origin does not necessarily map to the origin
- Lines map to lines
- Parallel lines remain parallel
- Closed under composition



Projective Transformation

Any transformation of the form:

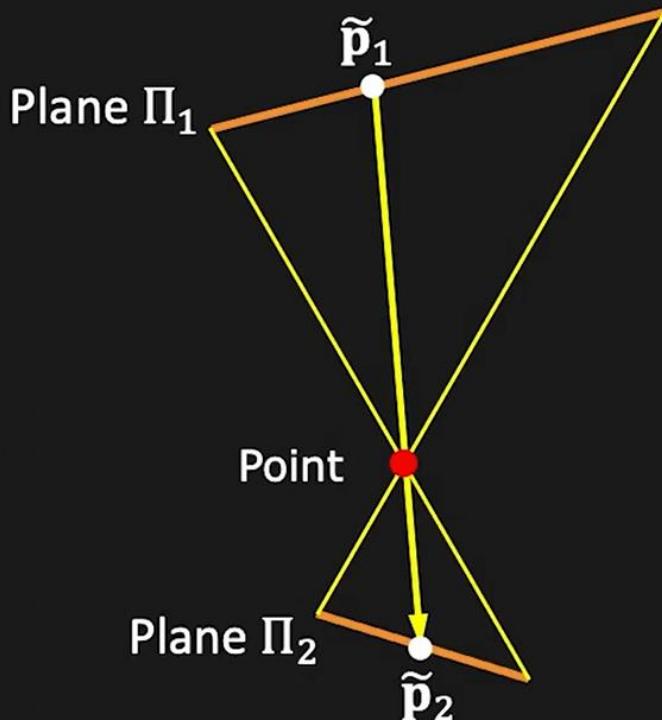
$$\begin{bmatrix} \tilde{x}_2 \\ \tilde{y}_2 \\ \tilde{z}_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} \tilde{x}_1 \\ \tilde{y}_1 \\ \tilde{z}_1 \end{bmatrix} \quad \tilde{\mathbf{p}}_2 = H\tilde{\mathbf{p}}_1$$



Also called Homography, or perspective transform

Projective Transformation

Mapping of one plane to another through a point



$$\tilde{\mathbf{p}}_2 = H\tilde{\mathbf{p}}_1$$

$$\begin{bmatrix} \tilde{x}_2 \\ \tilde{y}_2 \\ \tilde{z}_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} \tilde{x}_1 \\ \tilde{y}_1 \\ \tilde{z}_1 \end{bmatrix}$$

Same as imaging a plane through a pinhole

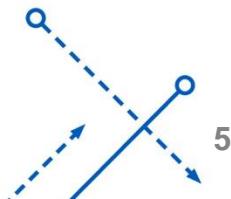
Projective Transformation

Homography can only be defined up to a scale.

$$\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} \tilde{x}_1 \\ \tilde{y}_1 \\ \tilde{z}_1 \end{bmatrix} \equiv \begin{bmatrix} \tilde{x}_2 \\ \tilde{y}_2 \\ \tilde{z}_2 \end{bmatrix} \equiv \textcolor{brown}{k} \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} \tilde{x}_1 \\ \tilde{y}_1 \\ \tilde{z}_1 \end{bmatrix}$$

If we fix scale such that $\sqrt{\sum(h_{ij})^2} = 1$ then **8** free parameters

- Origin does not necessarily map to the origin
- Lines map to lines
- Parallel lines do not necessarily remain parallel
- Closed under composition



2D vs 3D transform

2D

Transformation	Matrix	# DoF	Preserves	Icon
translation	$[\mathbf{I} \mid \mathbf{t}]_{2 \times 3}$	2	orientation	
rigid (Euclidean)	$[\mathbf{R} \mid \mathbf{t}]_{2 \times 3}$	3	lengths	
similarity	$[s\mathbf{R} \mid \mathbf{t}]_{2 \times 3}$	4	angles	
affine	$[\mathbf{A}]_{2 \times 3}$	6	parallelism	
projective	$[\tilde{\mathbf{H}}]_{3 \times 3}$	8	straight lines	

3D

Transformation	Matrix	# DoF	Preserves	Icon
translation	$[\mathbf{I} \mid \mathbf{t}]_{3 \times 4}$	3	orientation	
rigid (Euclidean)	$[\mathbf{R} \mid \mathbf{t}]_{3 \times 4}$	6	lengths	
similarity	$[s\mathbf{R} \mid \mathbf{t}]_{3 \times 4}$	7	angles	
affine	$[\mathbf{A}]_{3 \times 4}$	12	parallelism	
projective	$[\tilde{\mathbf{H}}]_{4 \times 4}$	15	straight lines	



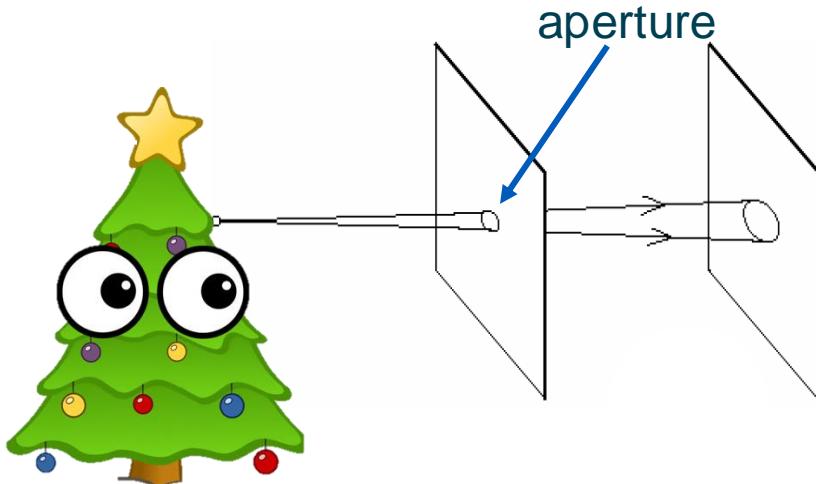
IMAGE FORMATION II

Optics



Previous Model - Pinhole

How does the size of the “Pinhole” affect the image we’d get?



- Make the aperture smaller?
 - How small?
 - Like a pinhole? Infinitely Small?
- What is the problem then?
 - How much light do we need?
- Solution?
 - Use a Lens

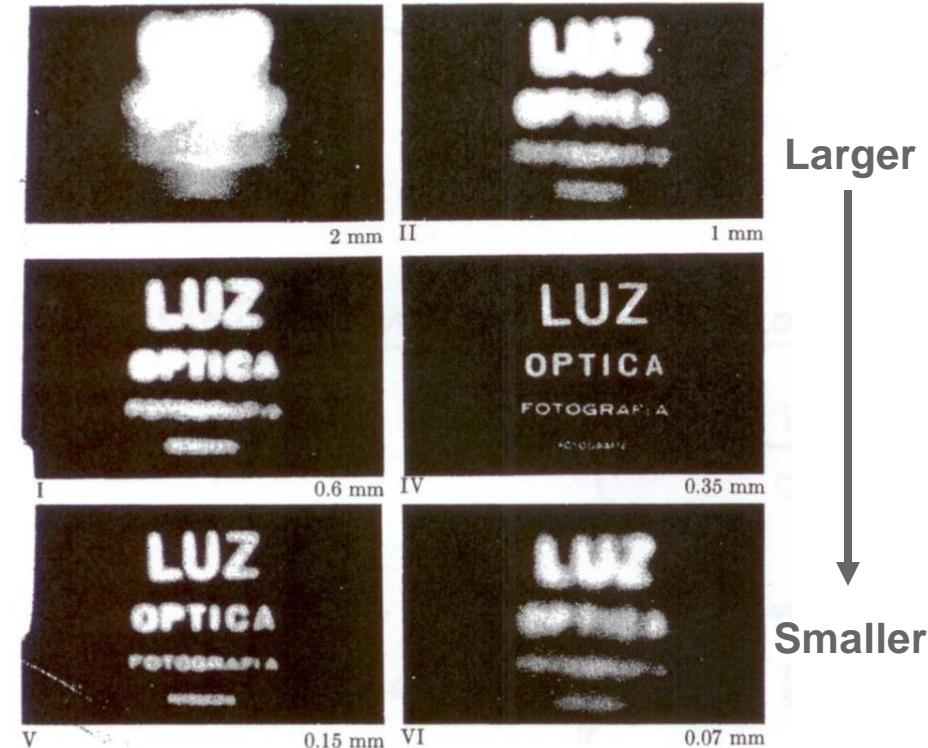
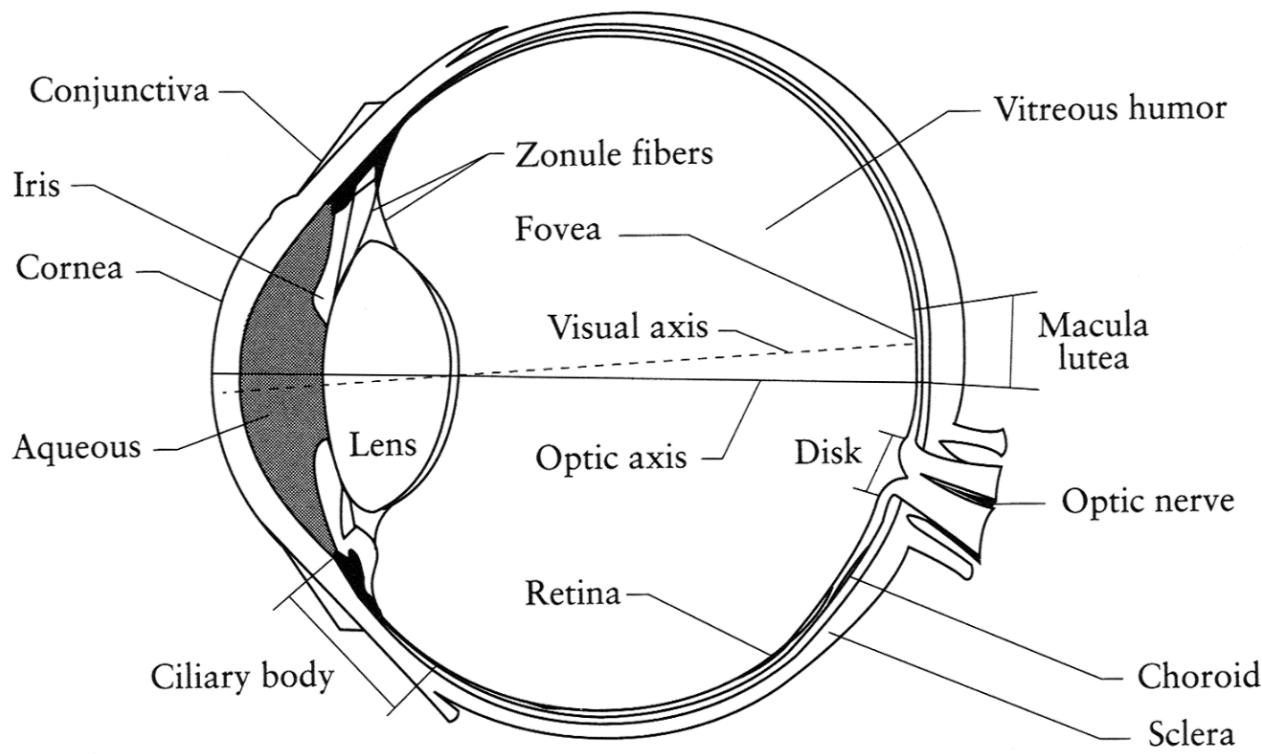


Fig. 5.96 The pinhole camera. Note the variation in image clarity as the hole diameter decreases. [Photos courtesy Dr. N. Joel, UNESCO.]

Human Eye

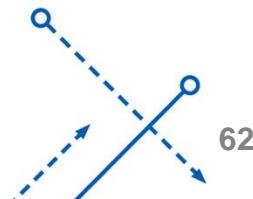
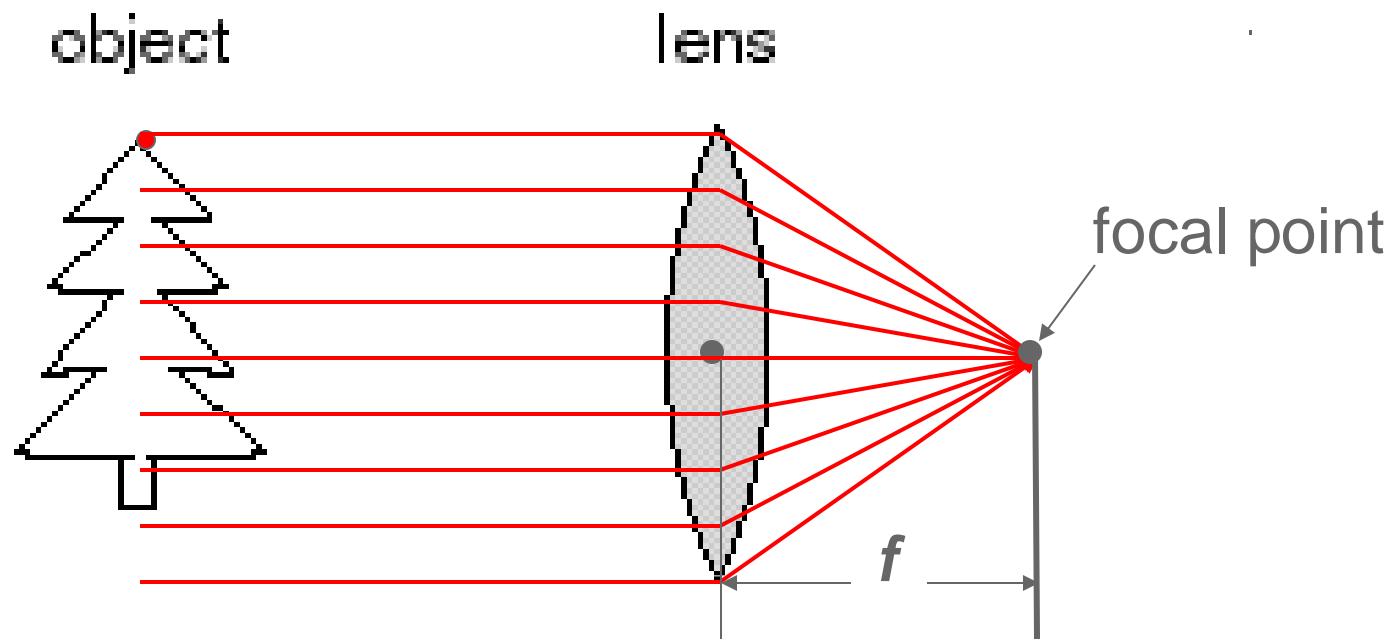
- **Iris:** Colored annulus with radial muscles
- **Pupil:** the hole (aperture) whose size is controlled by the iris.
- What's the “film”?



– photoreceptor cells (rods and cones) in the **retina**

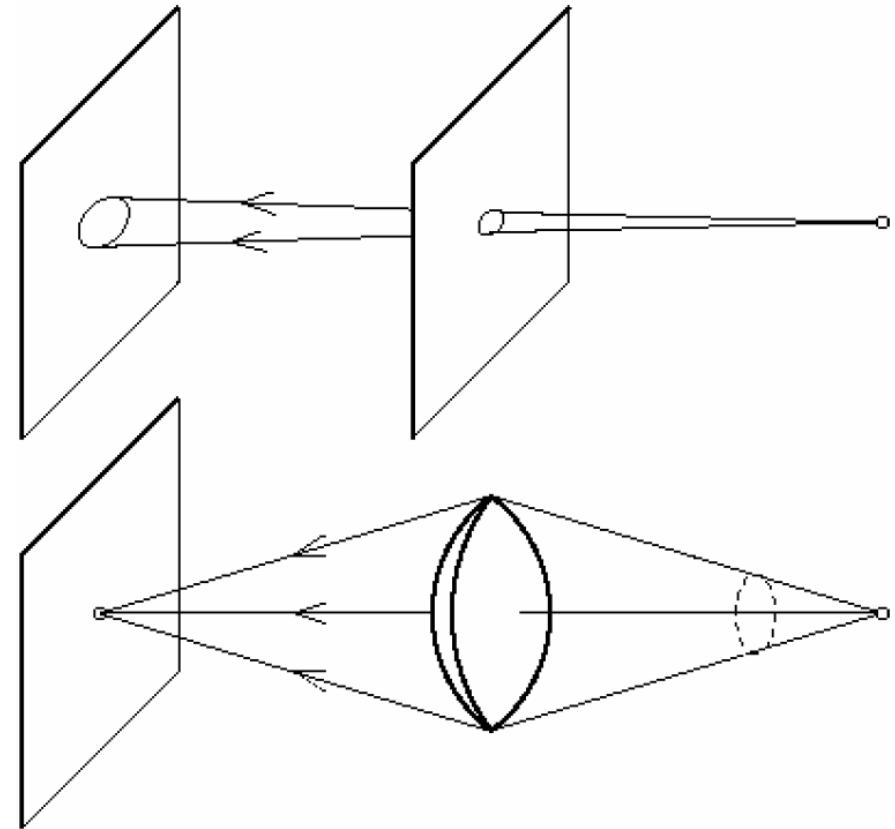
Lens

- A lens focuses parallel rays onto a single focal point
 - Gathers more lights; while keeping focus;
 - Make pinhole perspective projection practical

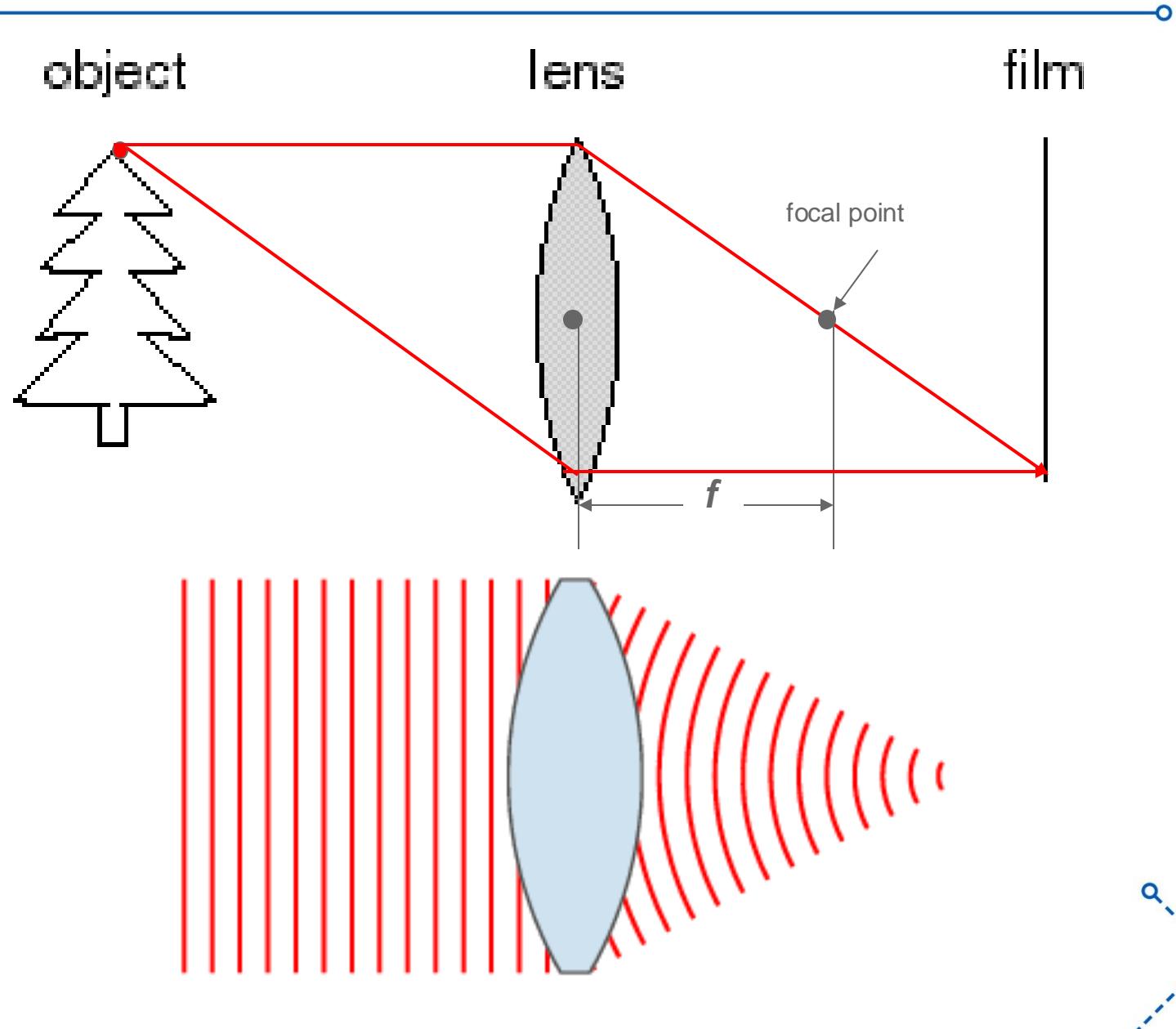


Benefits and challenges of adding lens

- Benefits
 - Light Concentration
 - Change the Focus
 - Depth of Field
 - Field of View
- Problems
 - Vignetting
 - Aberration

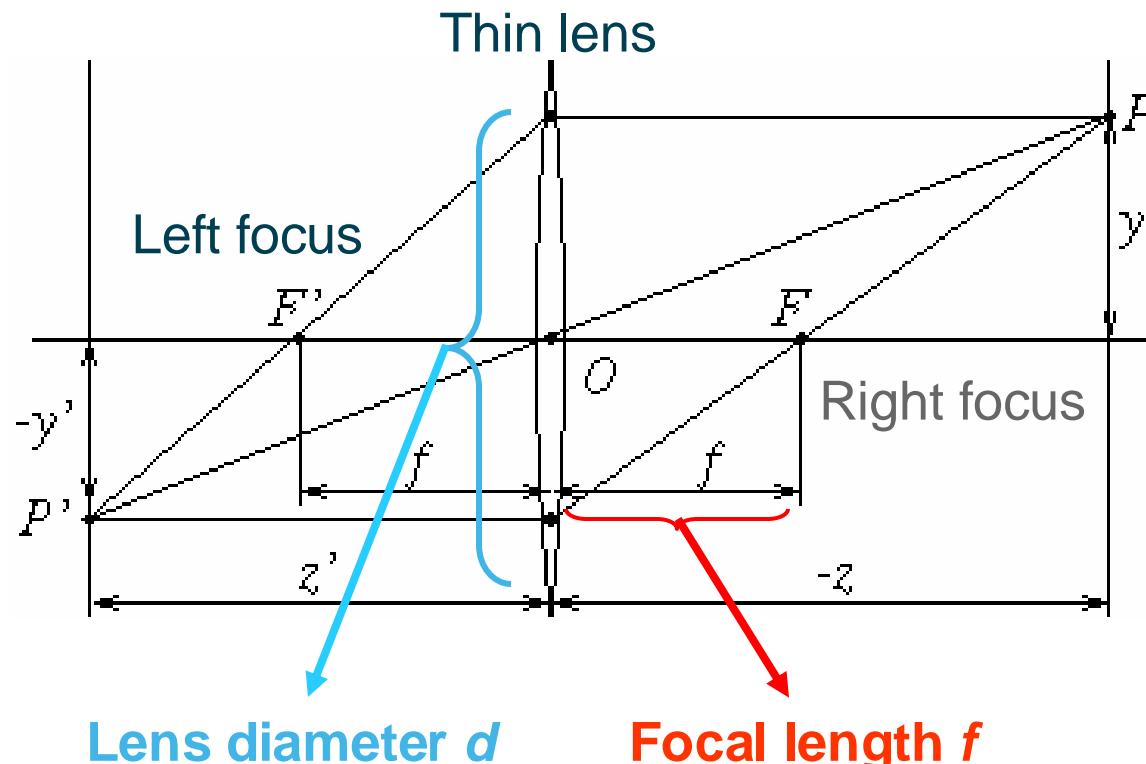


Light Concentration



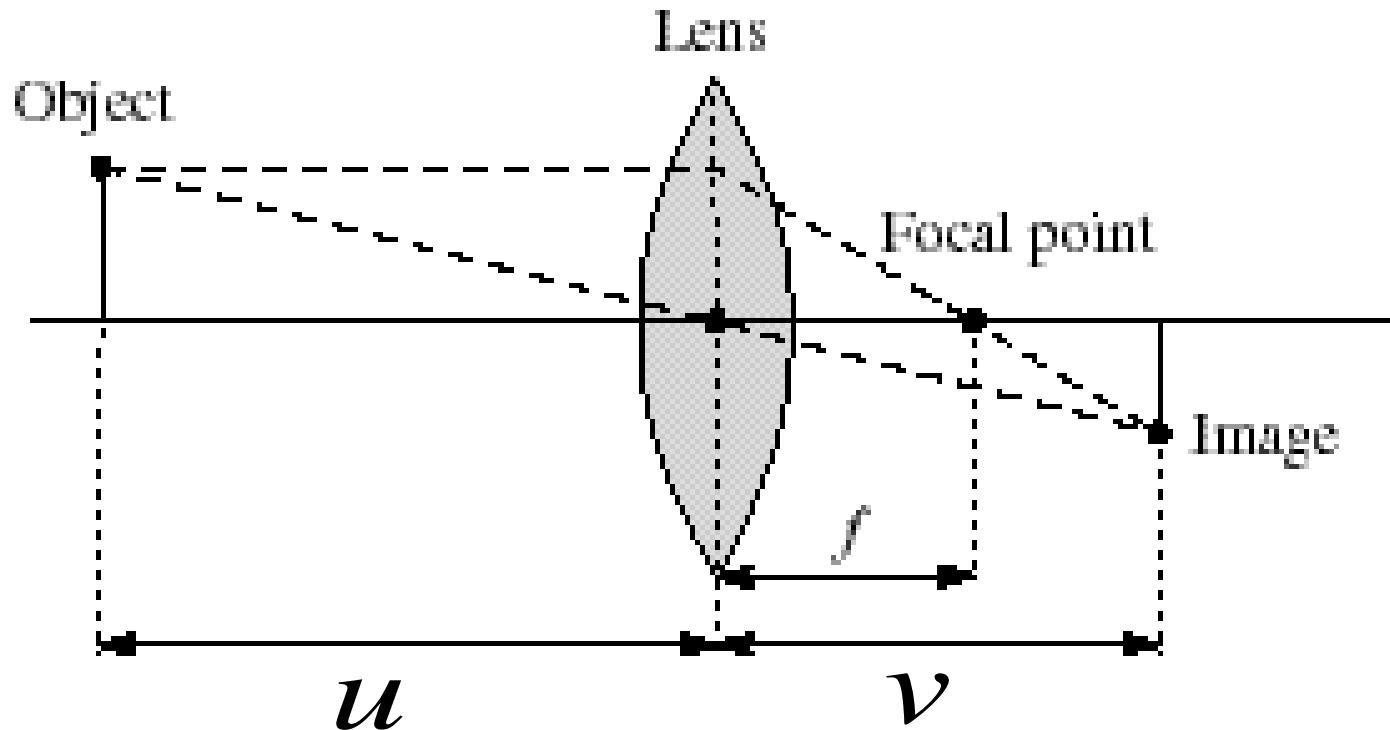
Thin lens

- Rays entering parallel on one side go through focus on other, and vice versa.
- In ideal case, all rays from P imaged at P'.



Thin Lens

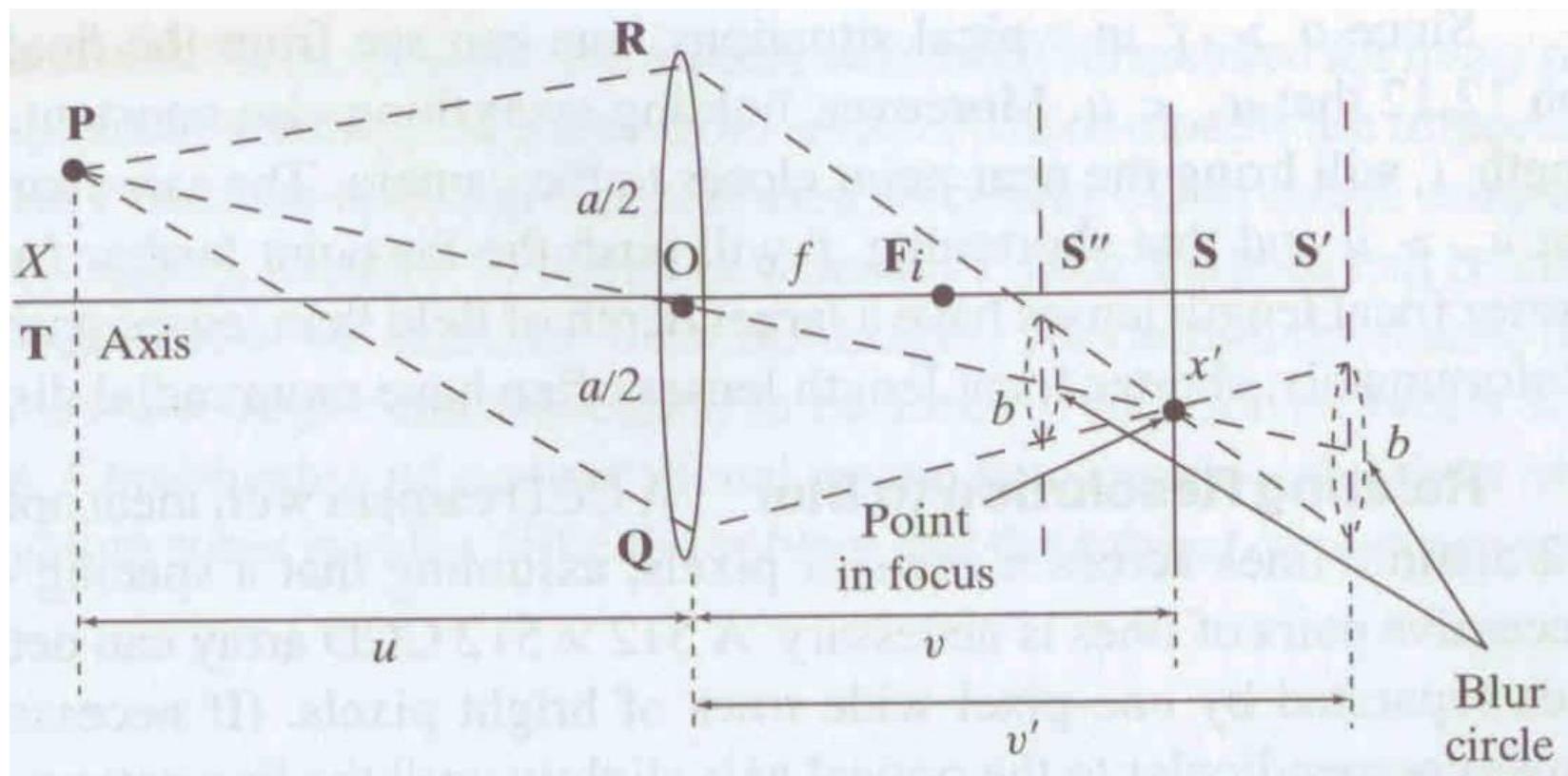
- Scene points at distinct depths come in focus at different image planes.



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

Depth of field

- Distance between image planes where blur is tolerable
- (Real camera lens systems have greater depth of field)

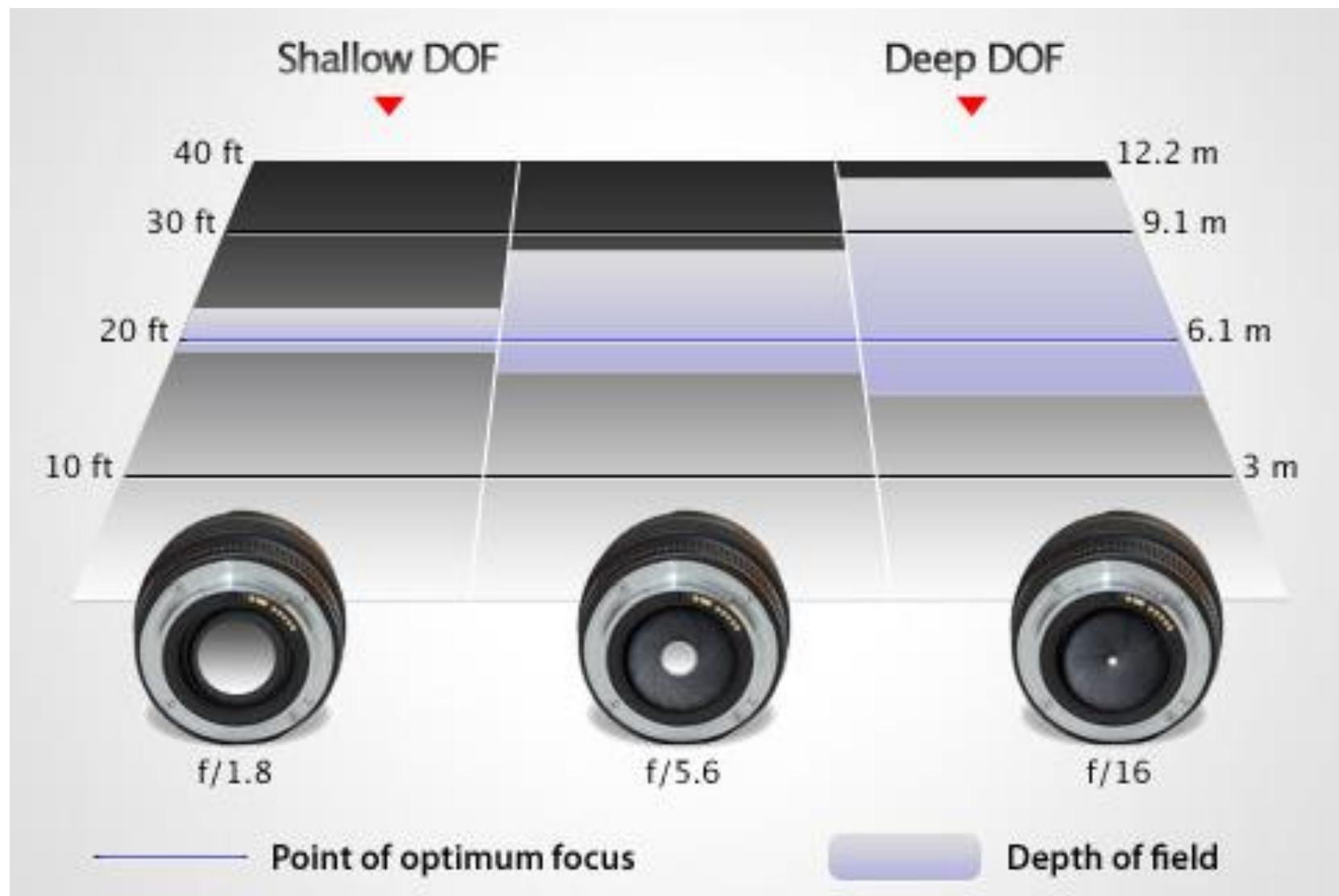


Focus and depth of field



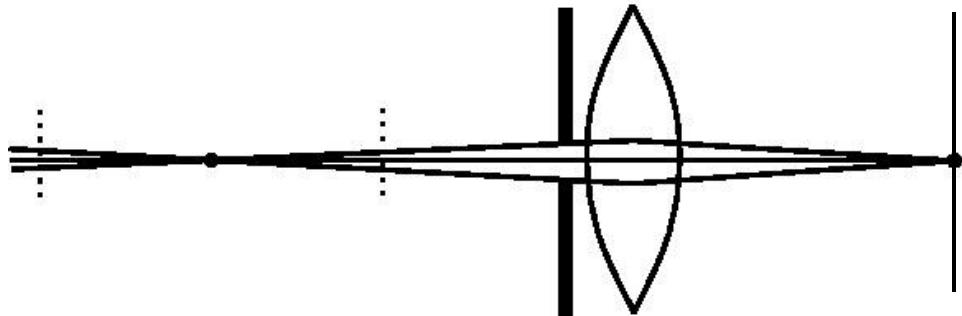
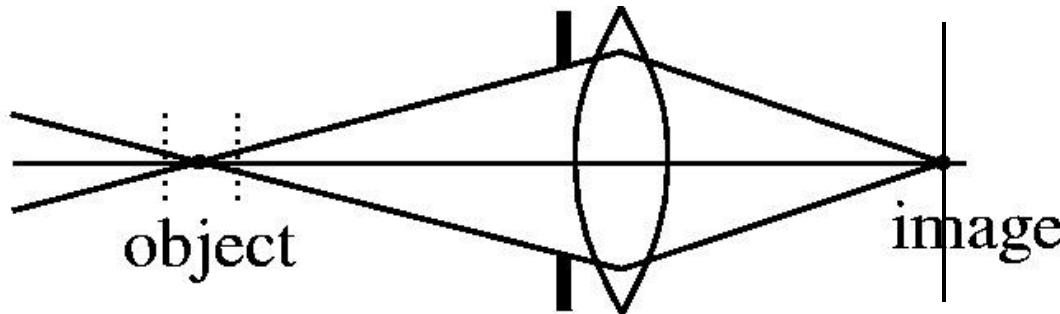
68

Depth of field



Aperture affects the depth of field

- A smaller aperture increases the range in which the object is approximately in focus



Aperture settings



f/22 - small aperture
Deep Depth of Field



f/2.8 - large aperture
Shallow Depth of Field



f/2



f/2.8



f/4



f/5.6



f/8



f/11

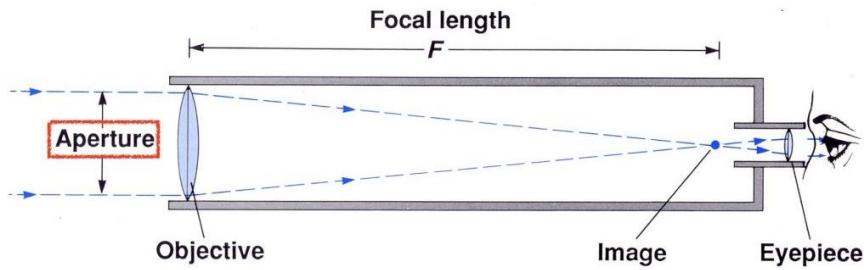


f/16



f/22

Brightness vs Aperture



- **F-number**

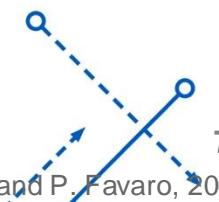
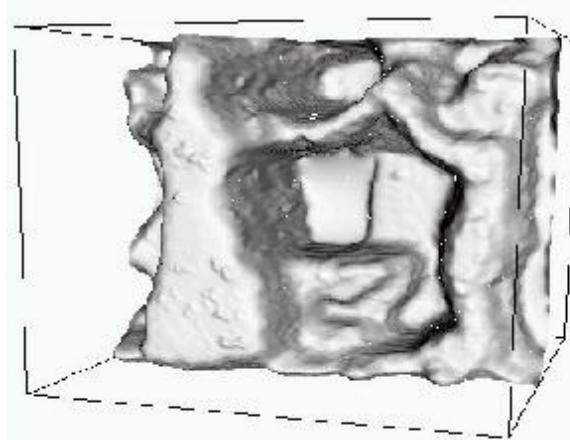
- Ratio of focal length to the diameter of the lens.

$$\bullet \text{ F-number} = \frac{f}{d}$$

- Greater f-number projects darker images.
- Brightness of the projected image (illuminance):
 - relative to the brightness of the scene (luminance)
 - decreases with the **square** of the f-number.
- Doubling the f-number decreases:
 - the relative brightness by a factor of four.
 - To maintain the same photographic exposure
 - The exposure time would need to be four times as long.

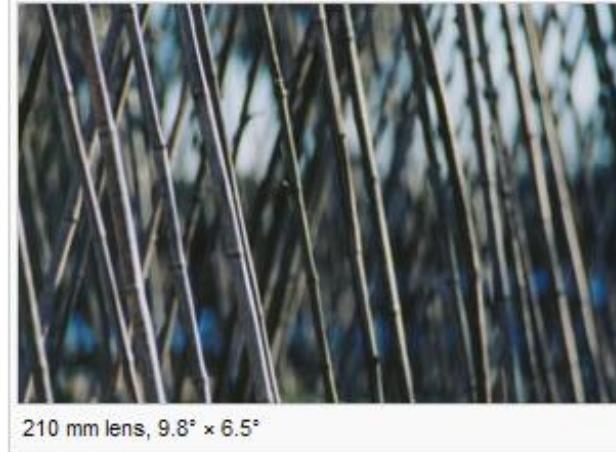
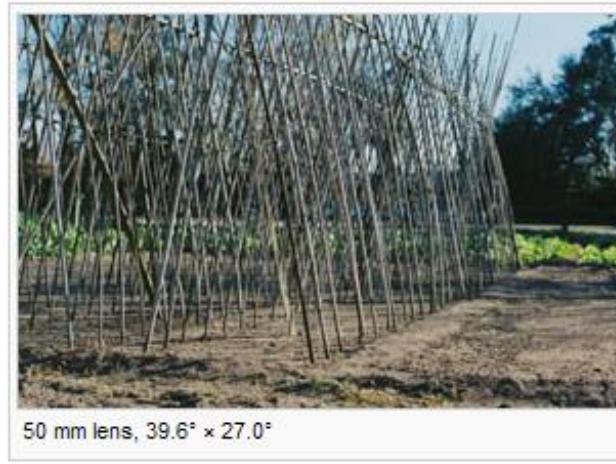
Depth from focus

- Same point of view, different camera parameters
- 3d shape / depth estimates



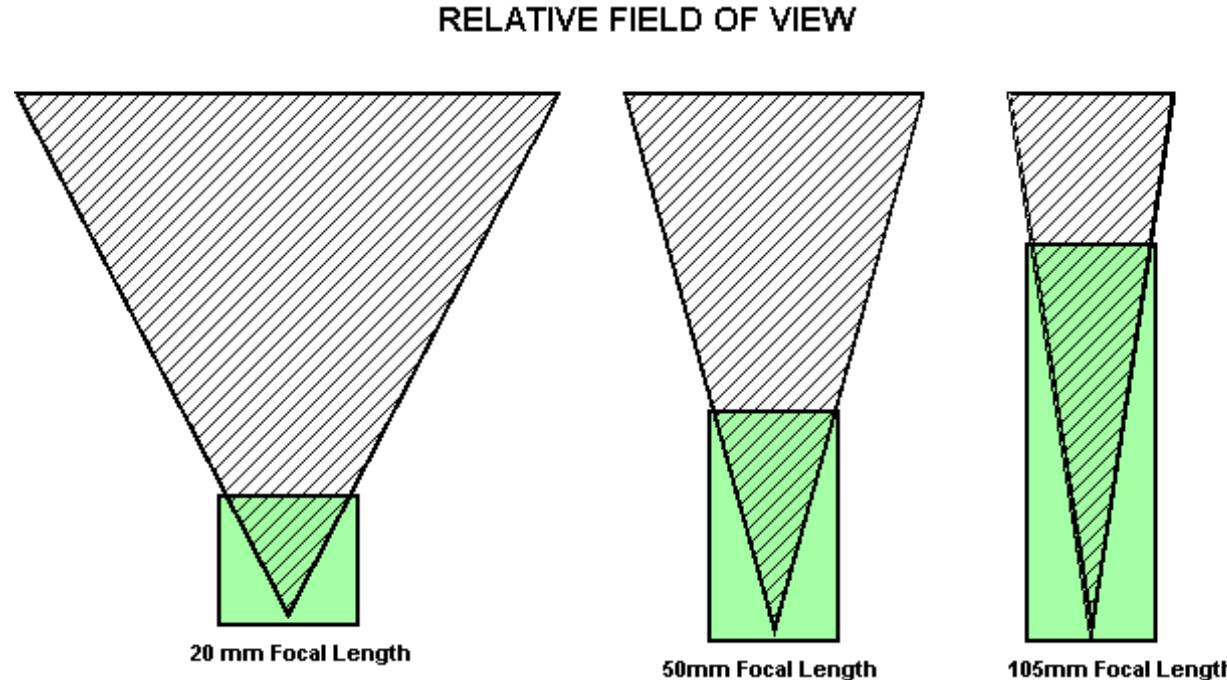
Field of view

- Angular measure of portion of 3D space seen by the camera



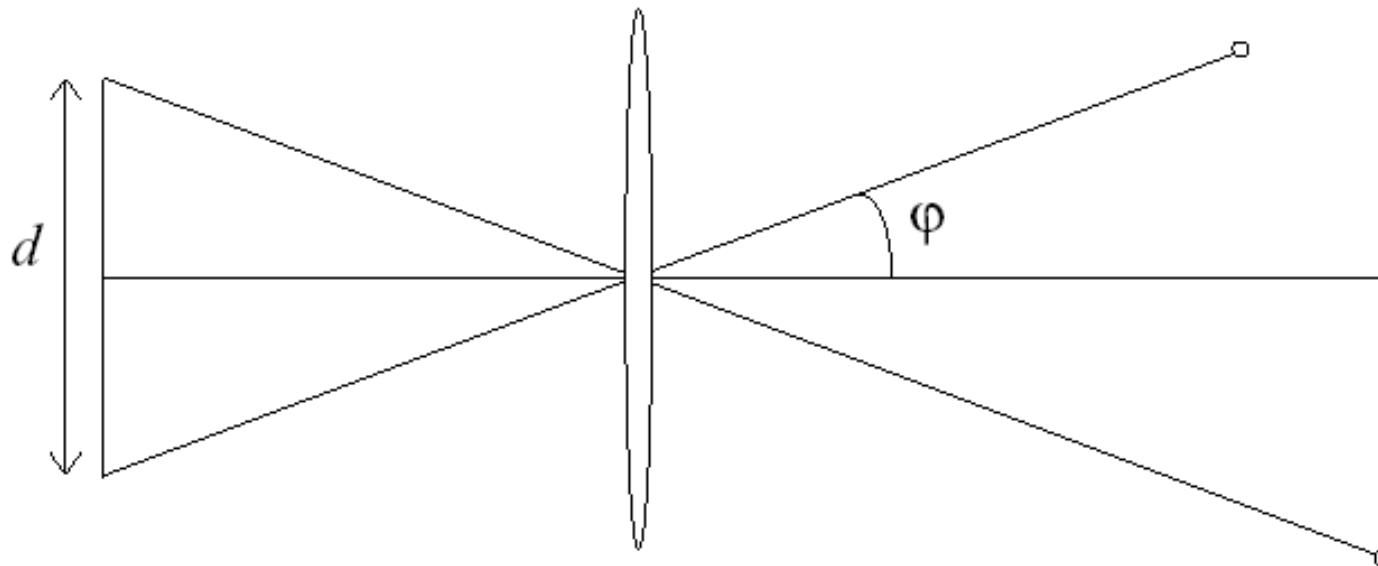
Field of view depends on focal length

- As F gets smaller, image becomes **wider angle**.
 - More world points project onto the image
- As F gets larger, image becomes **more telescopic**.
 - Less world points project onto the image



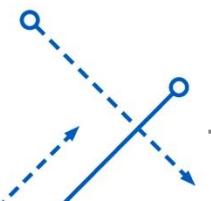
Field of view depends on camera retina

- Smaller FOV = larger Focal Length



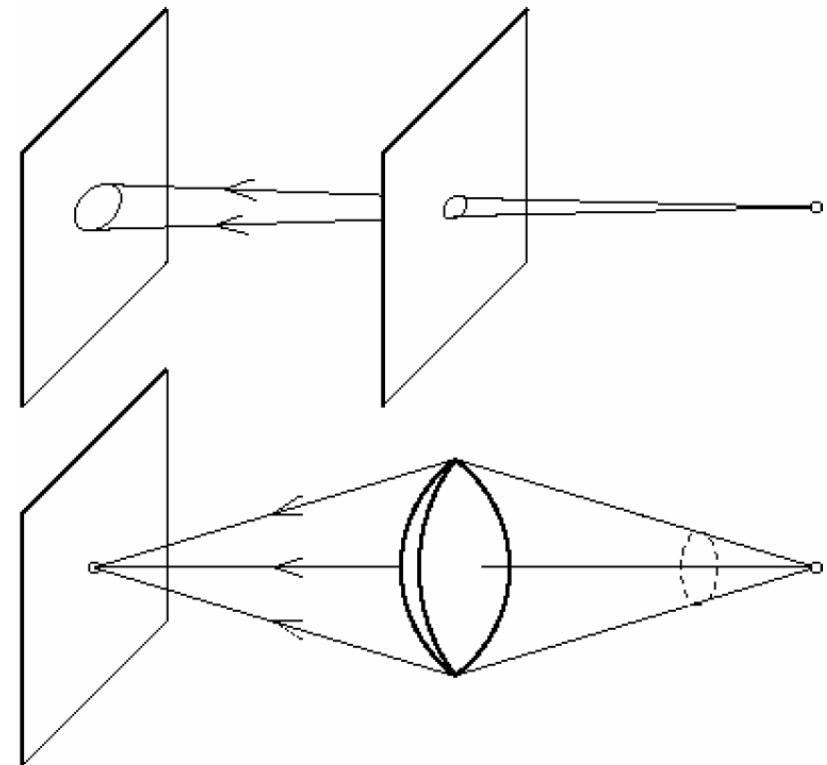
Size of field of view governed by size of the camera retina:

$$\varphi = \tan^{-1}\left(\frac{d}{2f}\right)$$



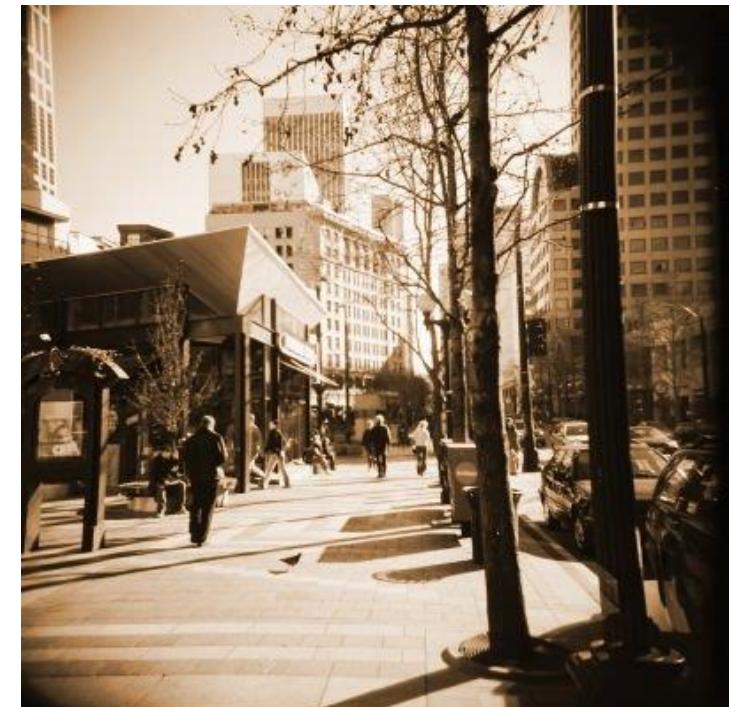
Benefits and challenges of adding lens

- Benefits
 - Light Concentration
 - Change the Focus
 - Depth of Field
 - Field of View
- Problems
 - Vignetting
 - Aberration



Vignetting

- Lens vignetting is a reduction in brightness or saturation on the periphery of an image.



Vignetting

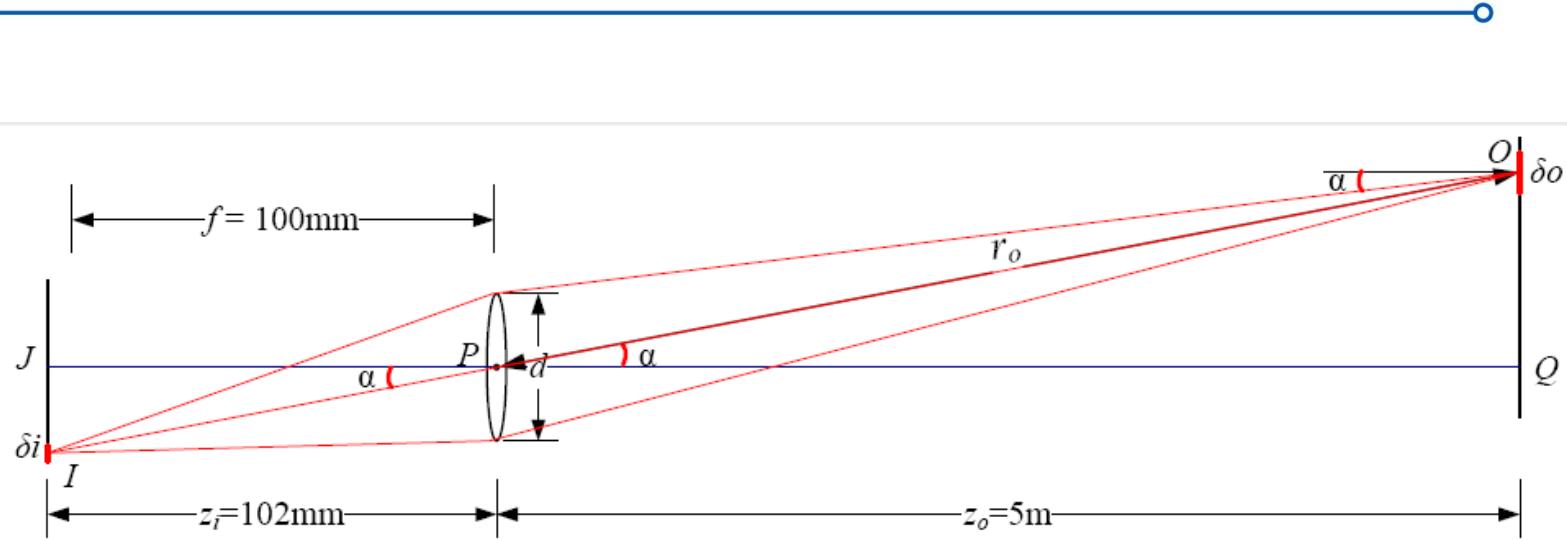
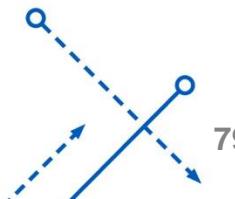


Figure 2.23: The amount of light hitting a pixel of surface area δi depends on the square of the ratio of the aperture diameter d to the focal length f , as well as the fourth power of the off-axis angle α cosine, $\cos^4 \alpha$.

$$l \propto \left(\frac{d}{f}\right)^2 \cos^4 \alpha$$



Chromatic aberration

- Transverse chromatic aberration
 - a blur and a rainbow edge in areas of contrast.

low quality lens

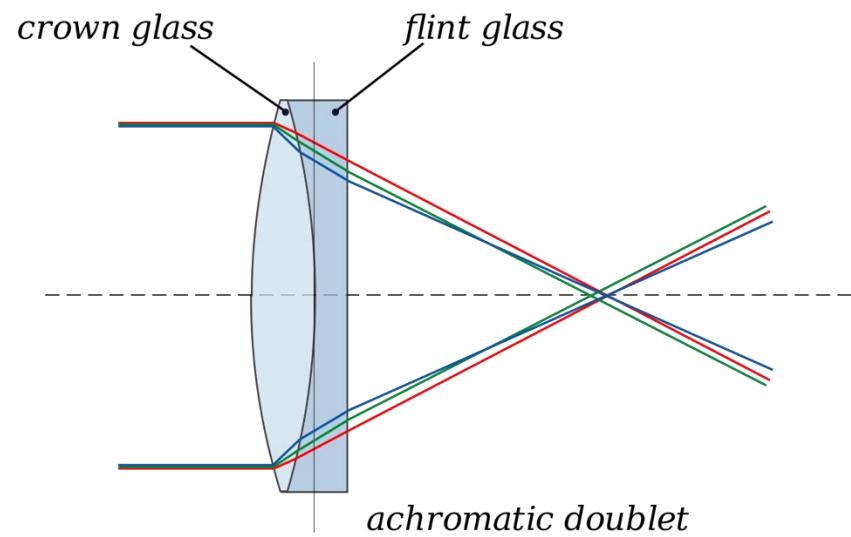
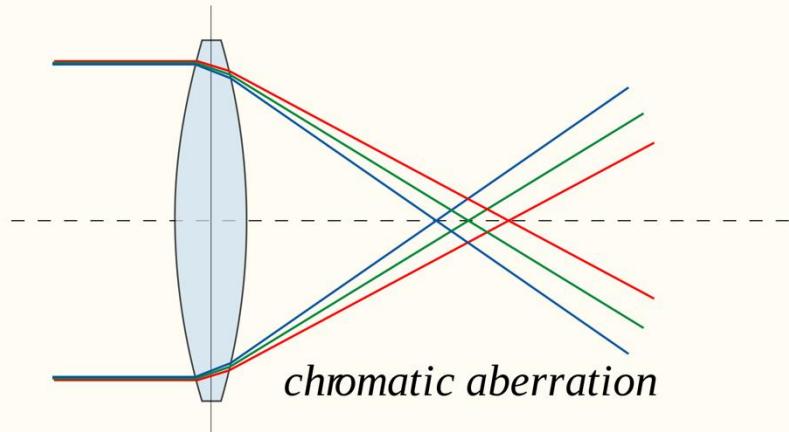


high quality lens

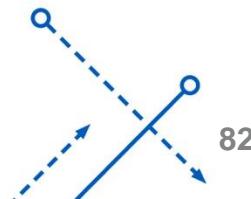
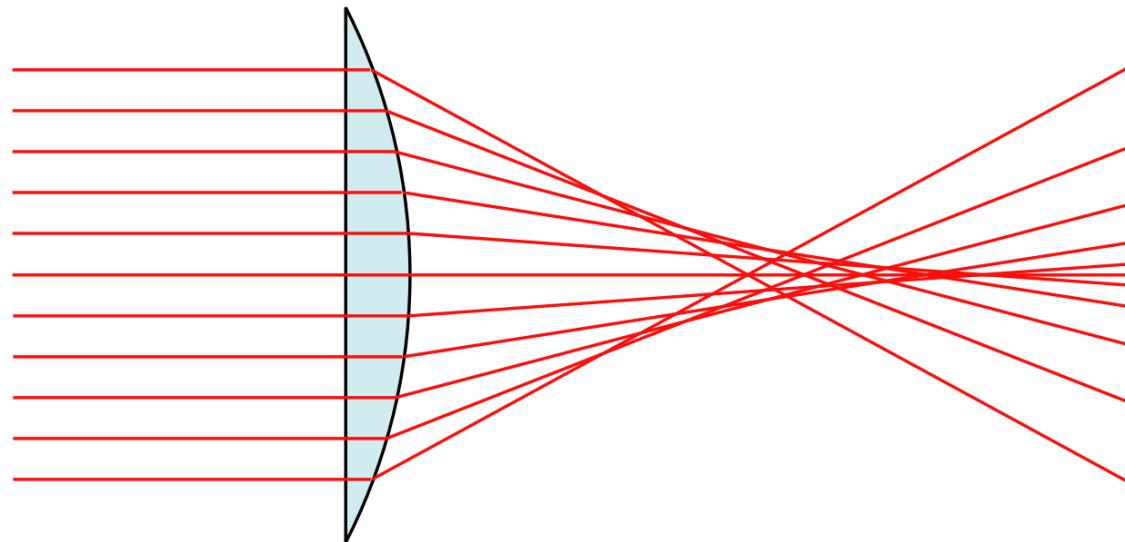
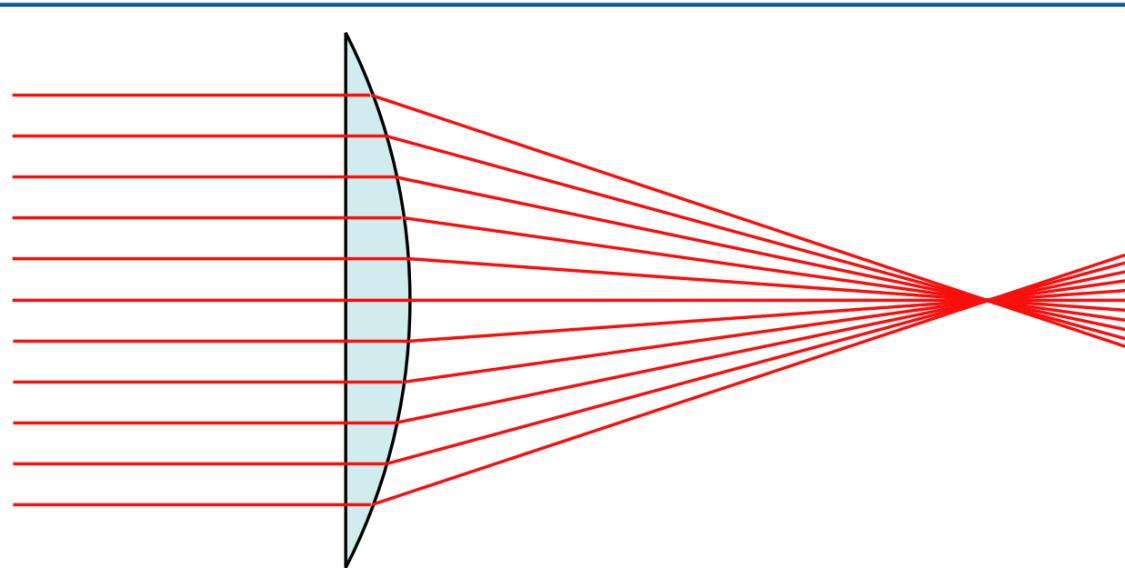


Chromatic aberration

- Dispersion of light

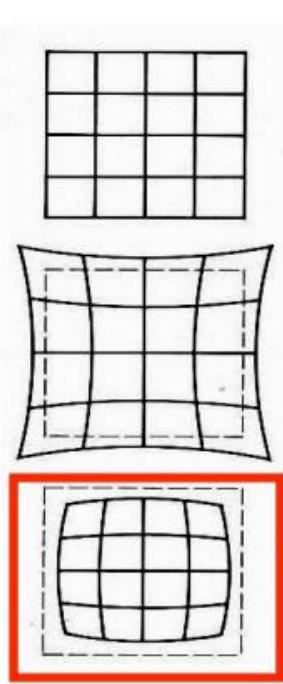


Spherical aberration



Other Distortions

- Imperfect Lenses
 - Deviations are most noticeable for rays passing through the edge of lens



No distortion

Pin cushion

Barrel

