Image Formation and Cameras

CSE 455

Linda Shapiro

Projection

http://www.julianbeever.net/pave.htm

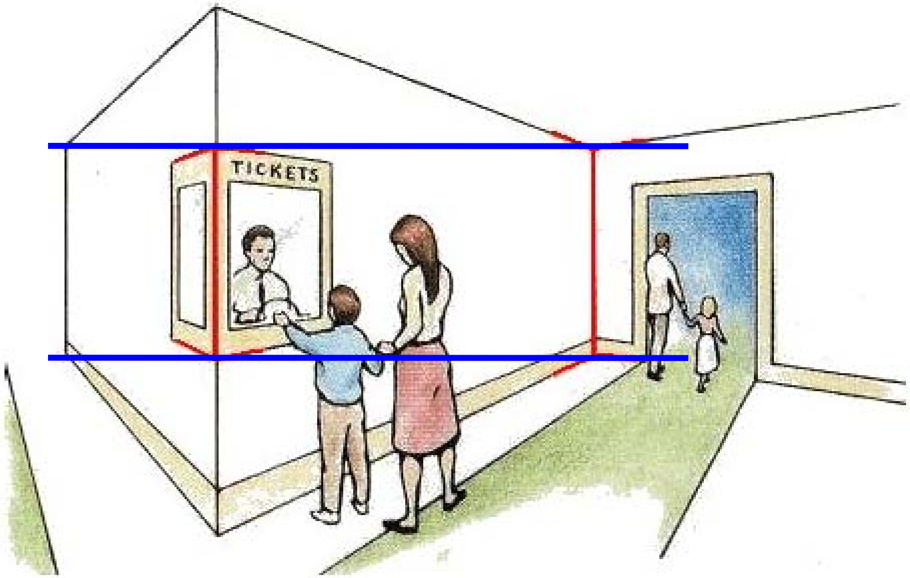
- Do sizes, lengths seem accurate?
- How do you know?

Projection

http://www.julianbeever.net/pave.htm

- What's wrong?
- Why do you think it's wrong?

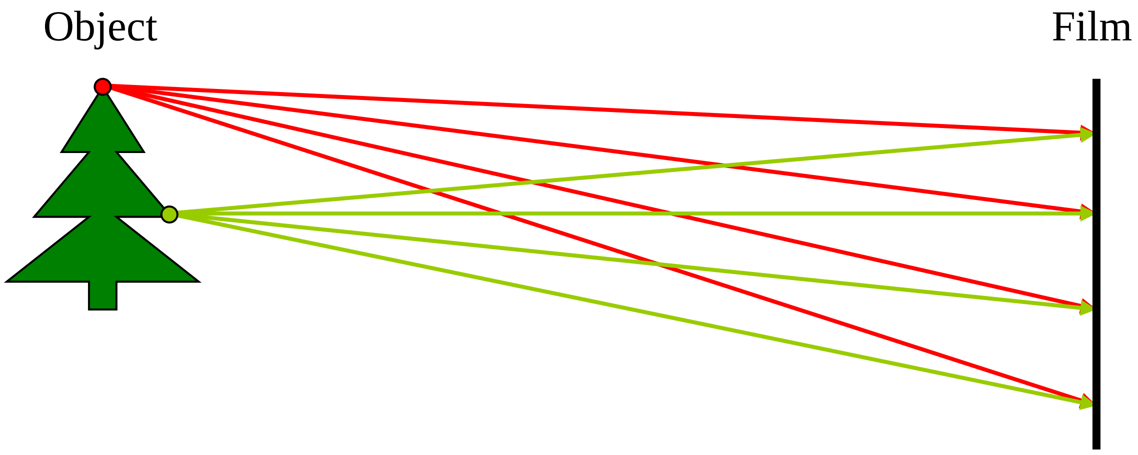
Müller-Lyer Illusion



http://www.michaelbach.de/ot/sze_muelue/index.html

- What do you know about perspective projection?
- Vertical lines?
- Other lines?

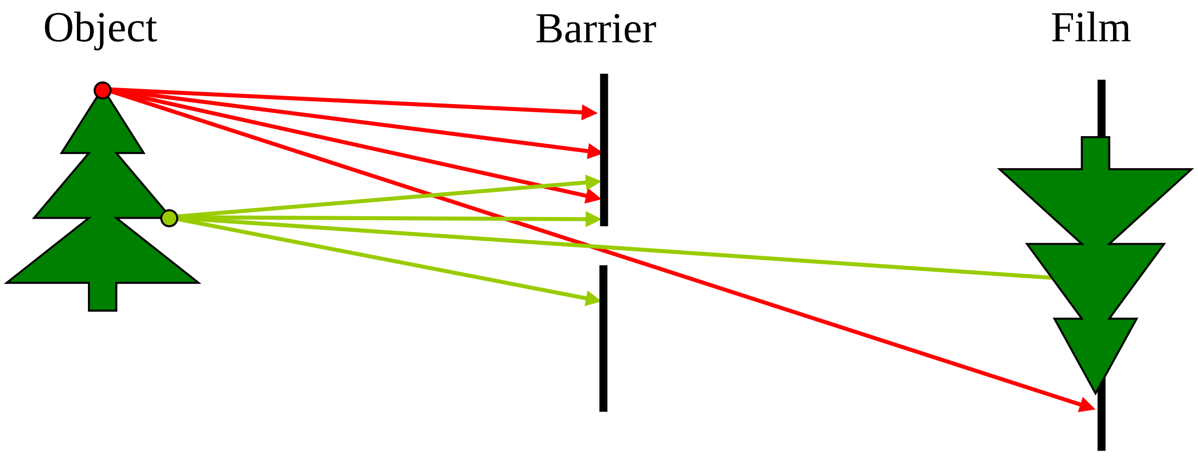
Image formation



Let's design a camera

- Idea 1: put a piece of film in front of an object
- Do we get a reasonable image?

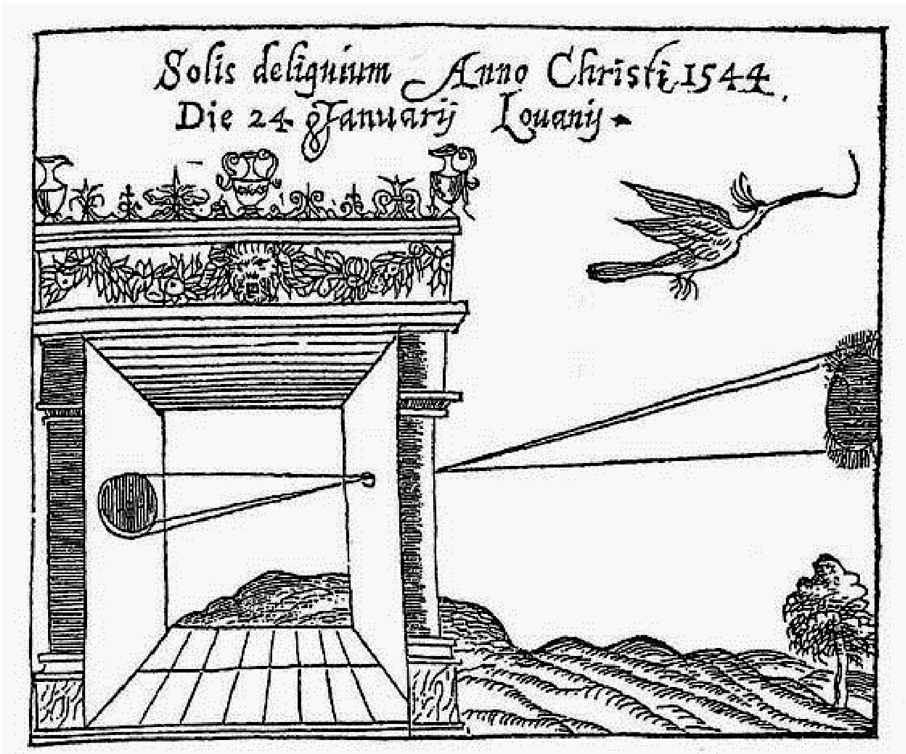
Pinhole camera



Add a barrier to block off most of the rays

- This reduces blurring
- The opening known as the **aperture**
- How does this transform the image?

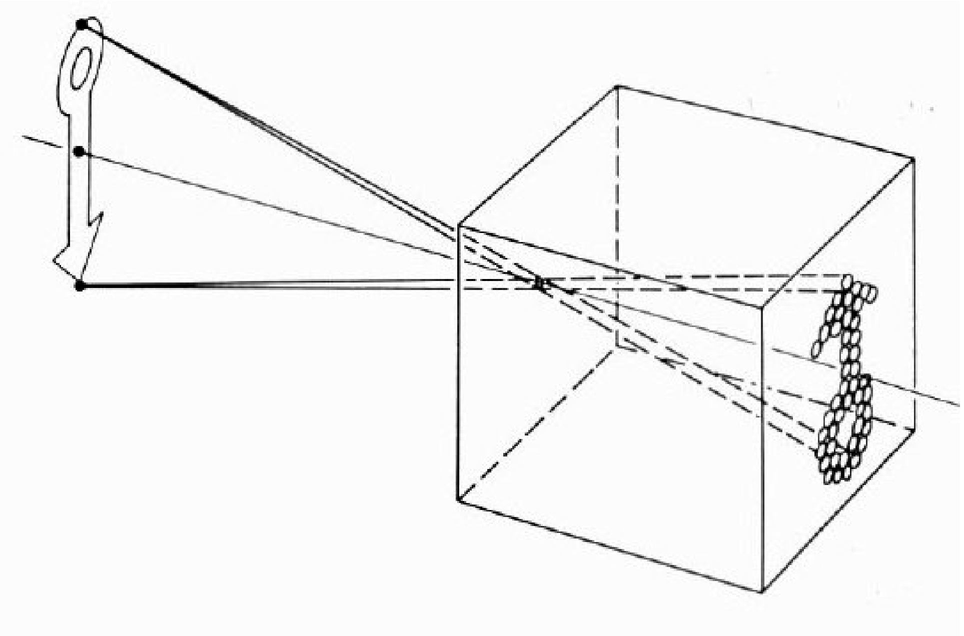
Camera Obscura



Gemma Frisius, 1558

- Basic principle known to Mozi (470-390 BC), Aristotle (384-322 BC)
- Drawing aid for artists: described by Leonardo da Vinci (1452-1519)

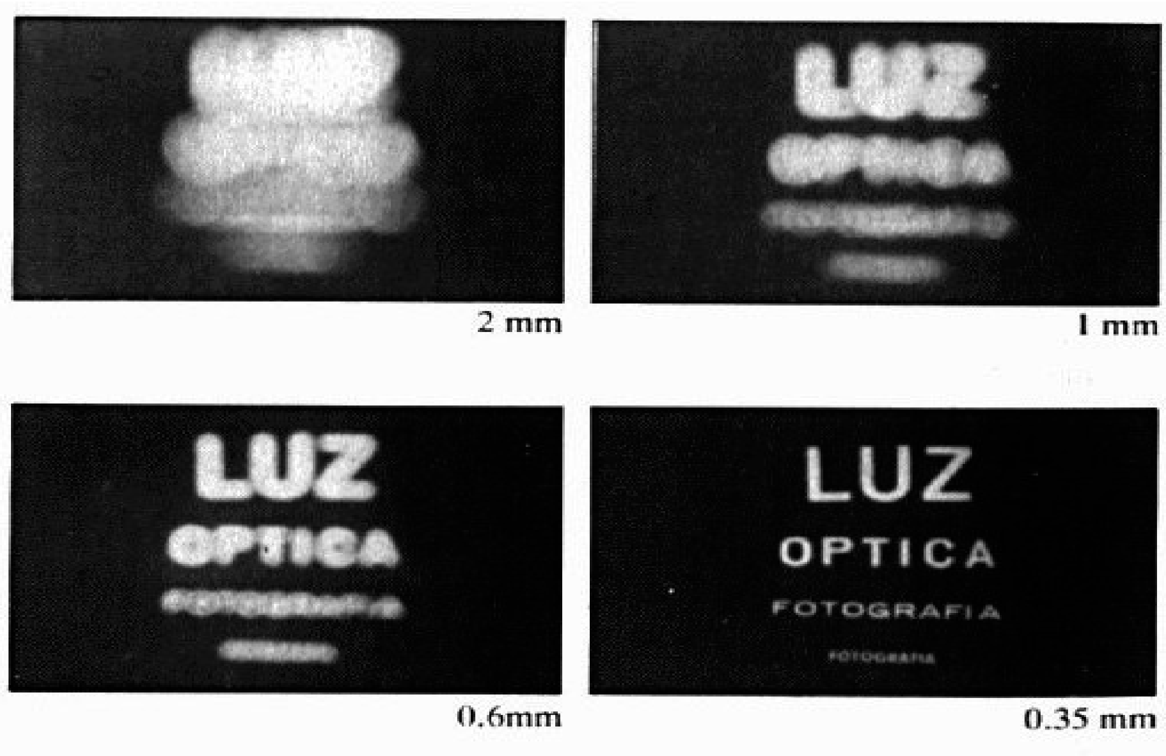
Camera Obscura



The first camera

- How does the aperture size affect the image?

Shrinking the aperture



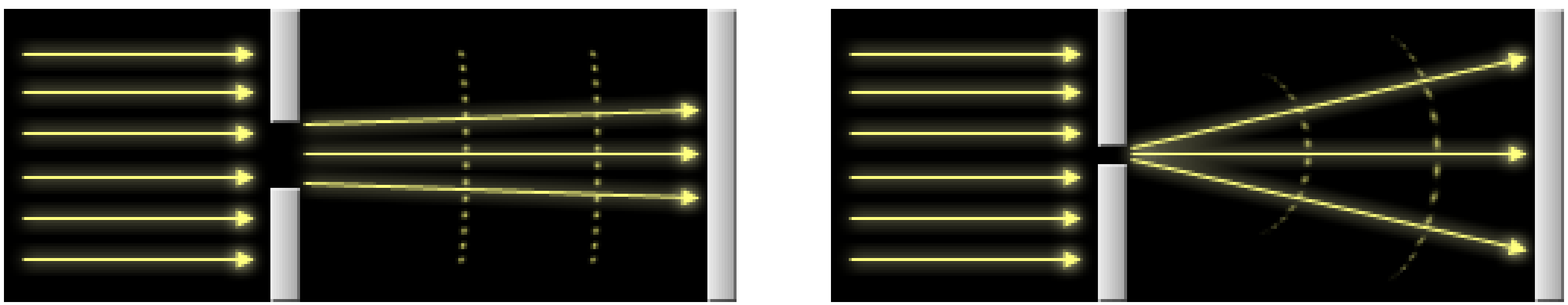
2 mm 1 mm 0.6mm 0.35 mm

Why not make the aperture as small as possible?

- Less light gets through
- *Diffraction* effects...

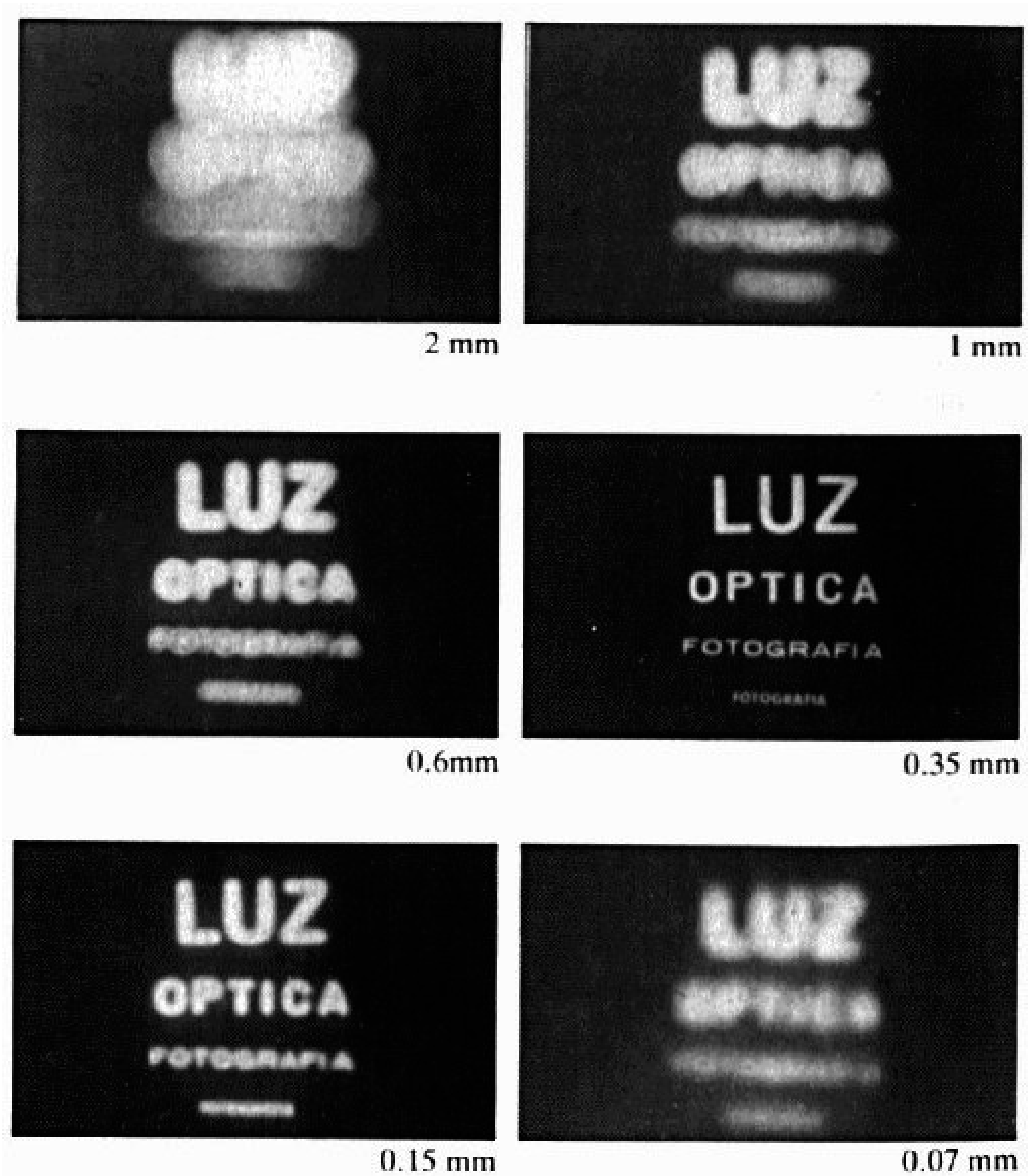
Diffraction

- Light rays passing through a small aperture will begin to diverge and interfere with one another.
- This becomes more significant as the size of the aperture decreases relative to the wavelength of light passing through.



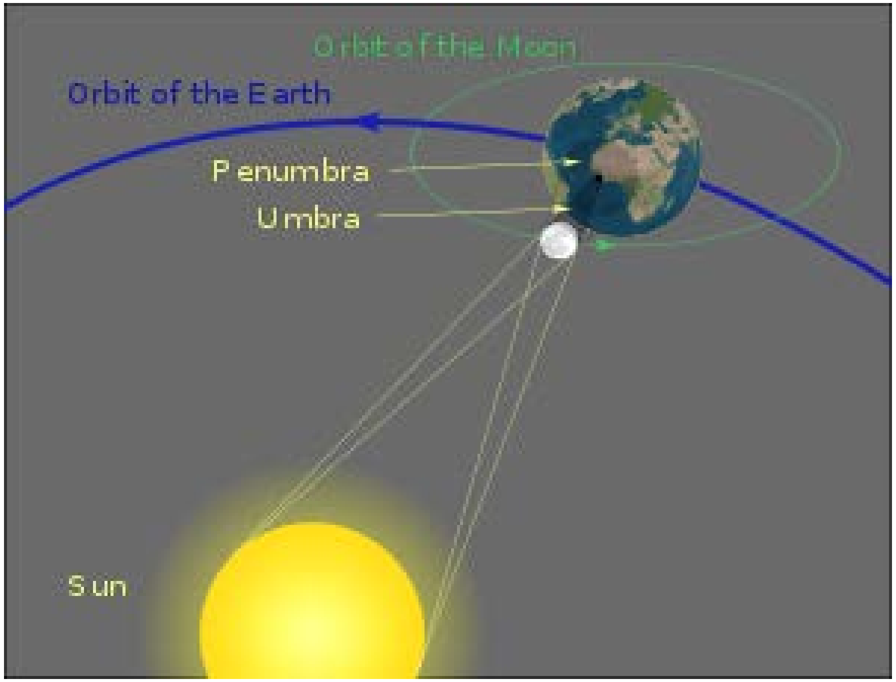
- This effect is normally negligible, since smaller apertures often improve sharpness.
- But at some point, your camera becomes diffraction limited, and the quality goes down.

Shrinking the aperture



Pinhole Cameras: Total Eclipse

- A total eclipse occurs when the moon comes between the earth and the sun, obscuring the sun.



Pinhole cameras everywhere

Sun “shadows” during a solar eclipse

by Henrik von Wendt http://www.flickr.com/photos/hvw/2724969199/

The holes between fingers work like a camera obscura and show the eclipsed sun

Pinhole cameras everywhere

Sun “shadows” during a partial solar eclipse

http://www.flickr.com/photos/73860948@N08/6678331997/

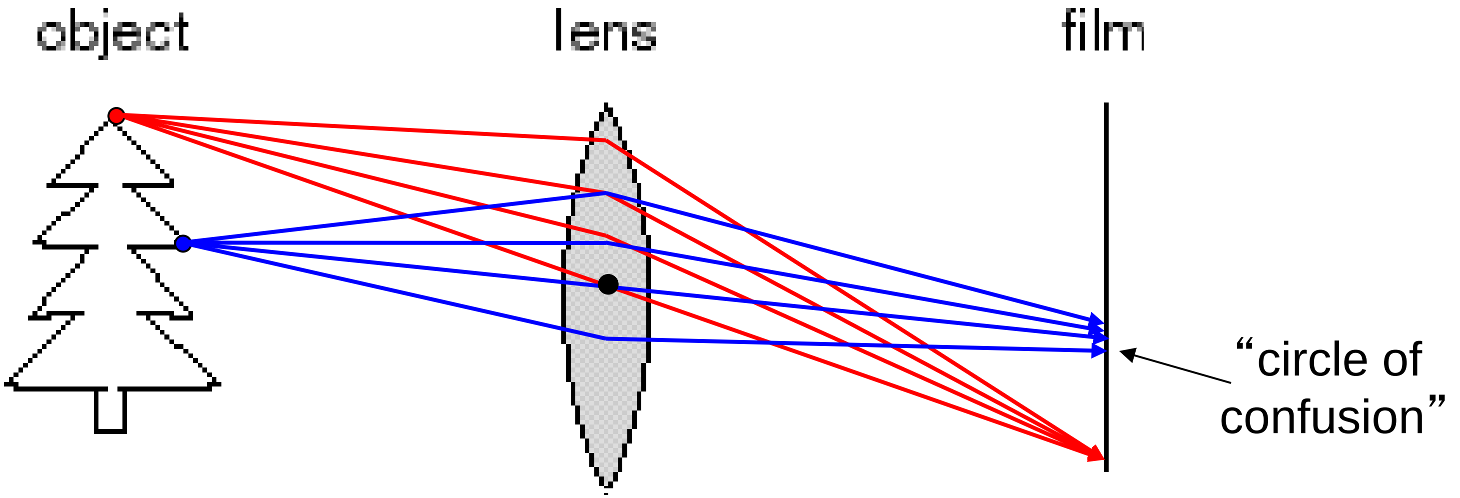
Pinhole cameras everywhere

Tree shadow during a solar eclipse

photo credit: Nils van der Burg

http://www.physicstogo.org/index.cfm

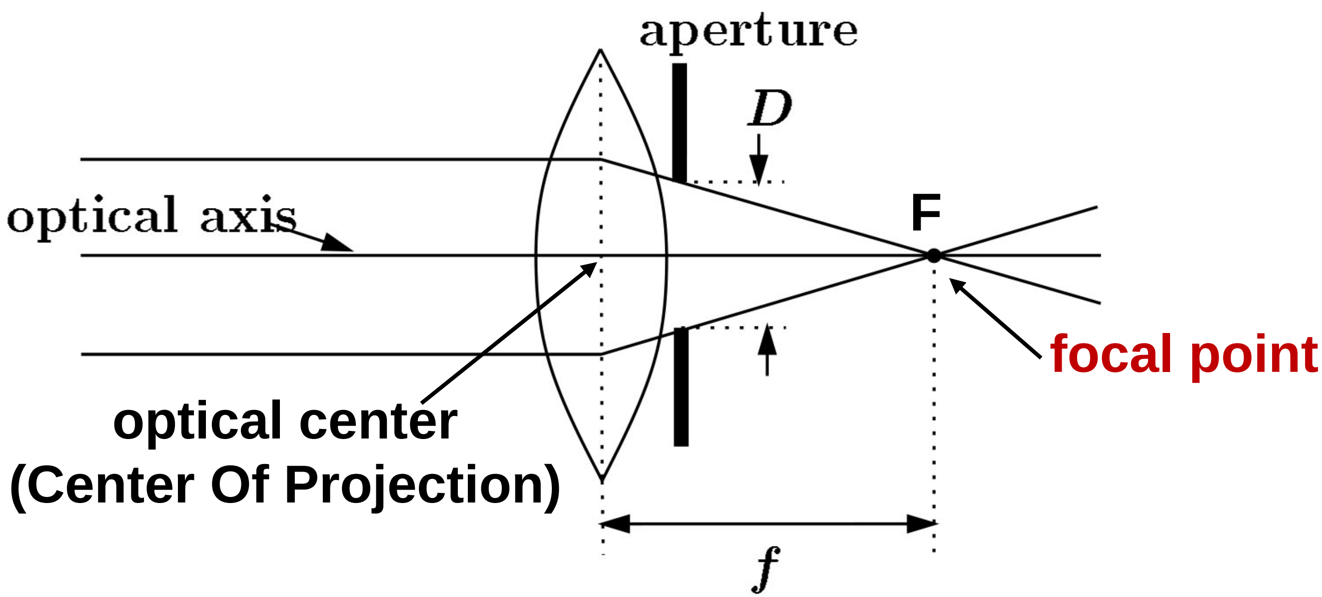
Adding a lens



A lens focuses light onto the film

- There is a specific distance at which objects are "in focus"
  - other points project to a "circle of confusion" in the image
- Changing the shape of the lens changes this distance

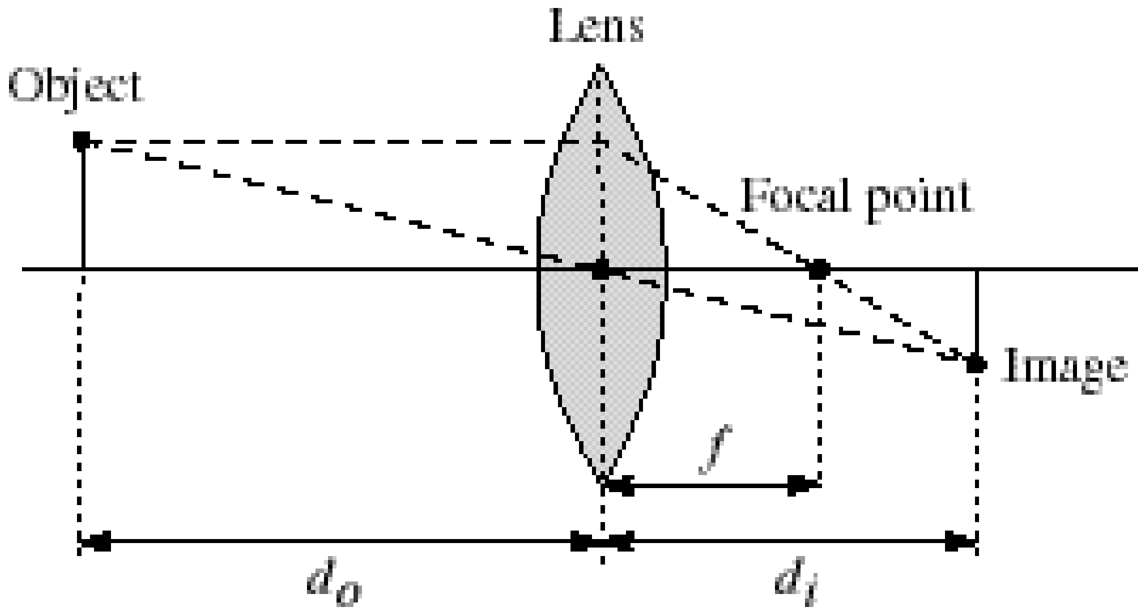
Lenses



A lens focuses parallel rays onto a single focal point

- focal point at a distance *f* beyond the plane of the lens
  - *f* is a function of the shape and index of refraction of the lens
- Aperture of diameter D restricts the range of rays
  - aperture may be on either side of the lens
- Lenses are typically spherical (easier to produce)
- Real cameras use many lenses together (to correct for aberrations)

Thin lenses

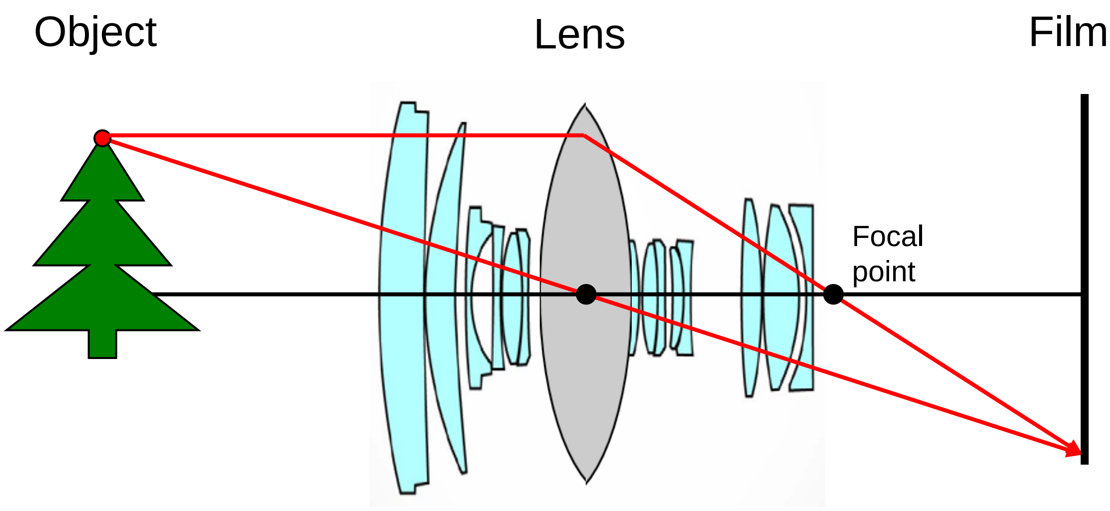


Thin lens equation: $$\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}$$

- Any object point satisfying this equation is in focus

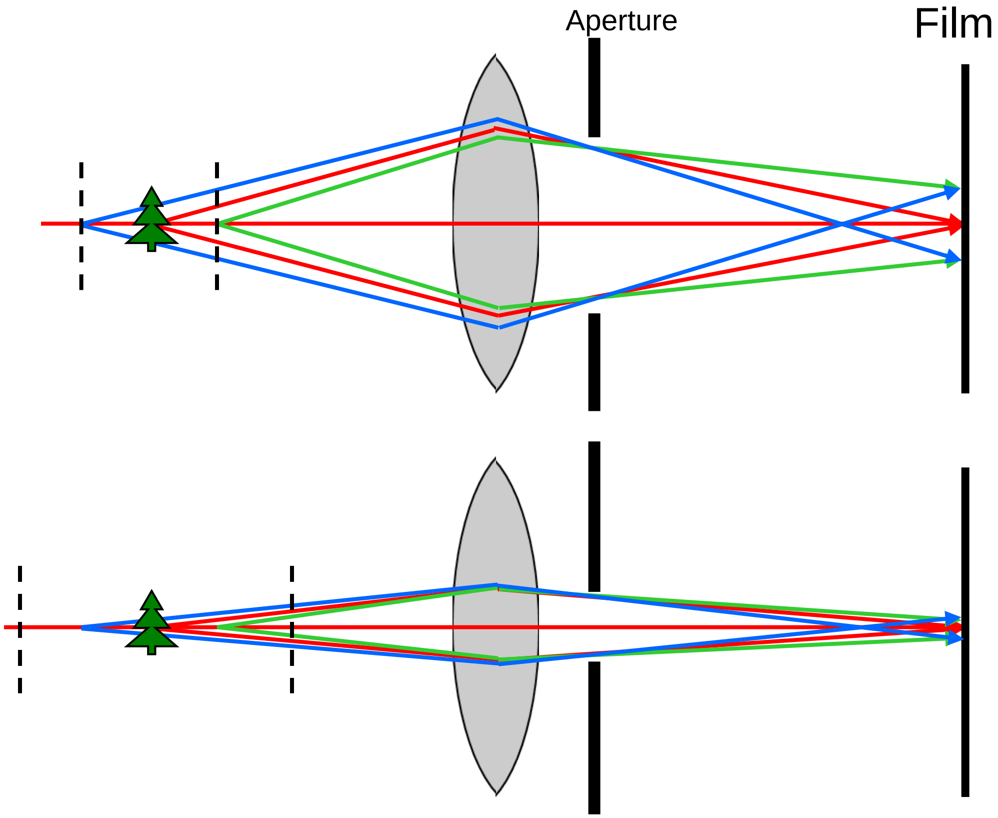
Thin lens assumption

The thin lens assumption assumes the lens has no thickness, but this isn’t true…



By adding more elements to the lens, the distance at which a scene is in focus can be made roughly planar.

Depth of field



f / 5.6

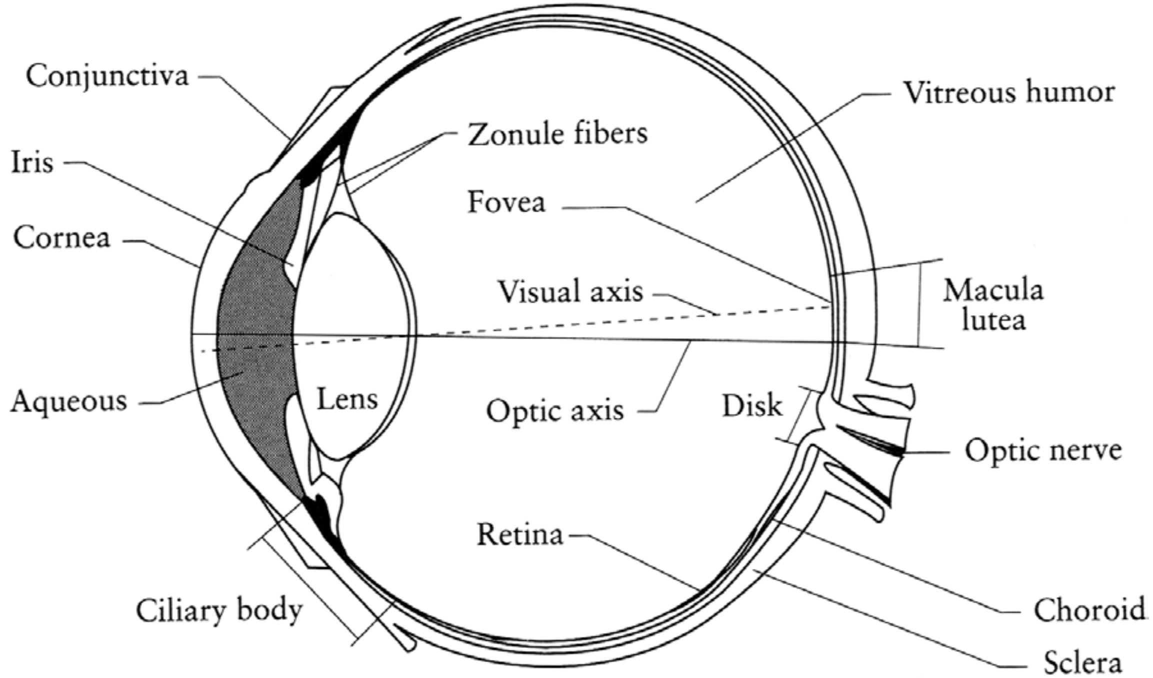
f / 32

Changing the aperture size affects depth of field

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

Flower images from Wikipedia http://en.wikipedia.org/wiki/Depth_of_field

The eye



The human eye is a camera

- **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**
- **How do we refocus?**
  - **Change the shape of the lens**

Digital camera

A digital camera replaces film with a sensor array

- Each cell in the array is a **C**harge **C**oupled **D**evice (CCD)
  - light-sensitive diode that converts photons to electrons
- CMOS is becoming more popular (esp. in cell phones)
  - http://electronics.howstuffworks.com/digital-camera.htm

Issues with digital cameras

Noise
- big difference between consumer vs. SLR-style cameras
- low light is where you most notice noise

Compression
- creates artifacts except in uncompressed formats (tiff, raw)

Color
- color fringing artifacts from Bayer patterns

Blooming
- charge overflowing into neighboring pixels

In-camera processing
- oversharpening can produce halos

Interlaced vs. progressive scan video
- even/odd rows from different exposures

Are more megapixels better?
- requires higher quality lens
- noise issues

Stabilization
- compensate for camera shake (mechanical vs. electronic)

More info online, e.g.,

- http://electronics.howstuffworks.com/digital-camera.htm
- http://www.dpreview.com/

Projection

Mapping from the world (3d) to an image (2d)

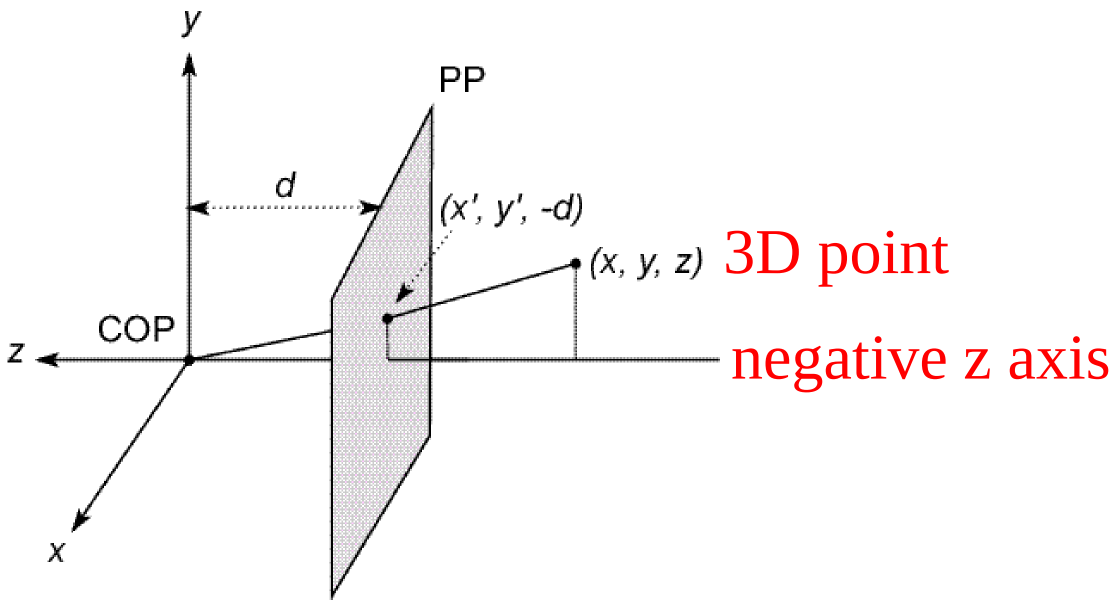
- Can we have a 1-to-1 mapping?
- How many possible mappings are there?

An optical system defines a particular projection. We'll talk about 2:

1. Perspective projection (how we see "normally")

2. Orthographic projection (e.g., telephoto lenses)

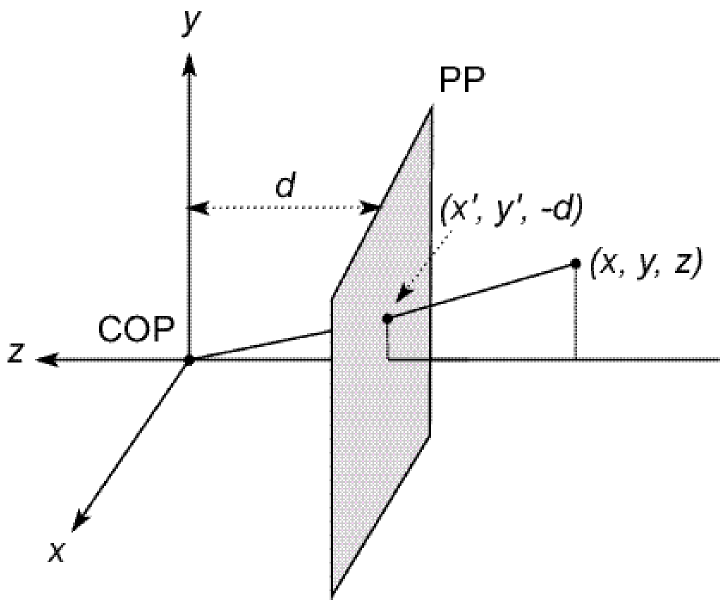
Modeling projection



The coordinate system

- We will use the pin-hole model as an approximation
- Put the optical center (**C**enter **O**f **P**rojection) at the origin
- Put the image plane (**P**rojection **P**lane) *in front* of the COP
- The camera looks down the *negative* z axis
  - we need this if we want right-handed-coordinates

Modeling projection



Projection equations

- Compute intersection with PP of ray from (x,y,z) to COP
- Derived using similar triangles

$$(x, y, z) \rightarrow (-d\frac{x}{z},\ -d\frac{y}{z},\ -d)$$

- We get the projection by throwing out the last coordinate:

$$(x, y, z) \rightarrow (-d\frac{x}{z},\ -d\frac{y}{z})$$

Homogeneous coordinates

Is this a linear transformation?

- no—division by z is nonlinear

Trick: add one more coordinate:

$$(x, y) \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

homogeneous image coordinates

$$(x, y, z) \Rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

homogeneous scene coordinates

Converting *from* homogeneous coordinates

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x/w, y/w)$$

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \Rightarrow (x/w, y/w, z/w)$$

Perspective Projection

Projection is a matrix multiply using homogeneous coordinates:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/d & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ -z/d \end{bmatrix} \Rightarrow (-d\frac{x}{z}, \quad -d\frac{y}{z})$$

divide by third coordinate

projection matrix 3D point 2D point

This is known as **perspective projection**

- The matrix is the **projection matrix**

Perspective Projection Example

1. Object point at (10, 6, 4), d=2

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/d & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/2 & 0 \end{bmatrix}\begin{bmatrix} 10 \\ 6 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 10 & 6 & -2 \end{bmatrix}$$

$$\Rightarrow x' = -5,\ y' = -3$$

2. Object point at (25, 15, 10)

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/d & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/2 & 0 \end{bmatrix}\begin{bmatrix} 25 \\ 15 \\ 10 \\ 1 \end{bmatrix} = \begin{bmatrix} 25 & 15 & -5 \end{bmatrix}$$

$$\Rightarrow x' = -5,\ y' = -3$$

Perspective projection is not 1-to-1!

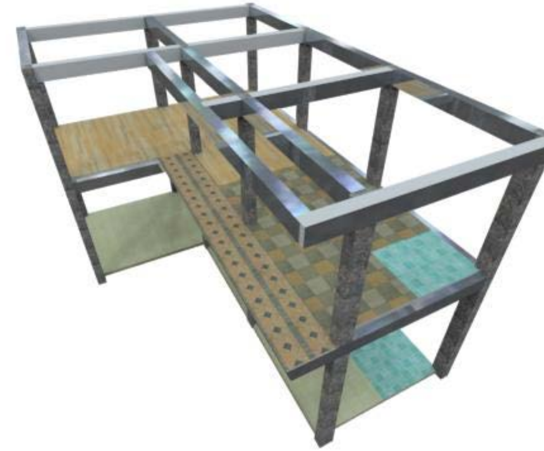
Perspective Projection

How does scaling the projection matrix change the transformation?

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/d & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ -z/d \end{bmatrix} \Rightarrow (-d\frac{x}{z}, \quad -d\frac{y}{z})$$

$$\begin{bmatrix} -d & 0 & 0 & 0 \\ 0 & -d & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} -dx \\ -dy \\ z \end{bmatrix} \Rightarrow (-d\frac{x}{z}, \quad -d\frac{y}{z})$$

Perspective Projection

- What happens to parallel lines?
- What happens to angles?
- What happens to distances?

Perspective Projection

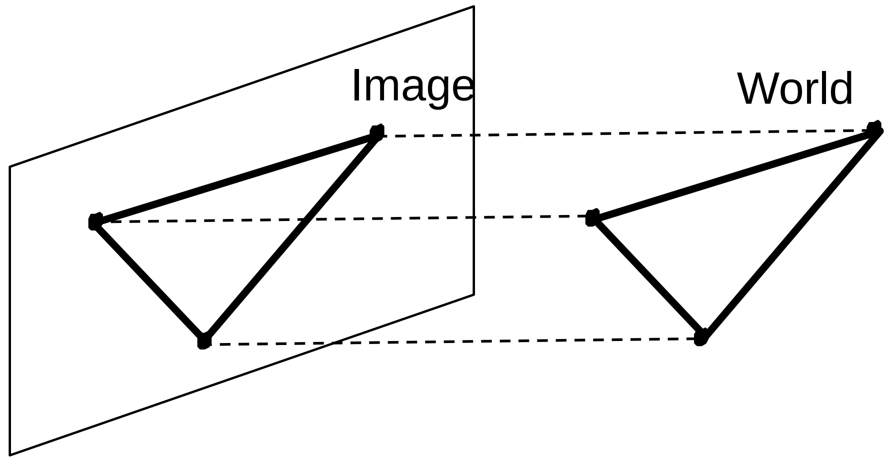
What happens when d➔∞?

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1/d & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ -z/d \end{bmatrix} \Rightarrow (-d\frac{x}{z}, \quad -d\frac{y}{z})$$

Orthographic projection

Special case of perspective projection

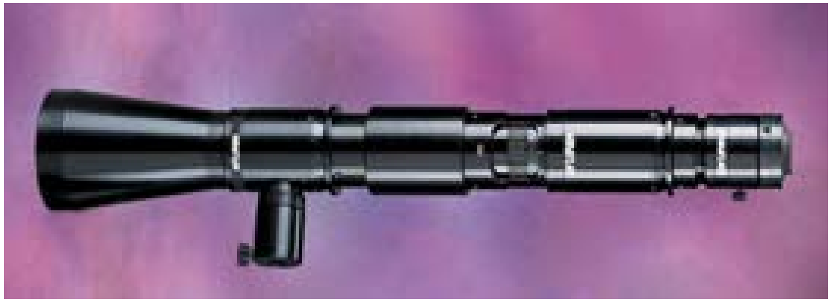
- Distance from the COP to the PP is infinite



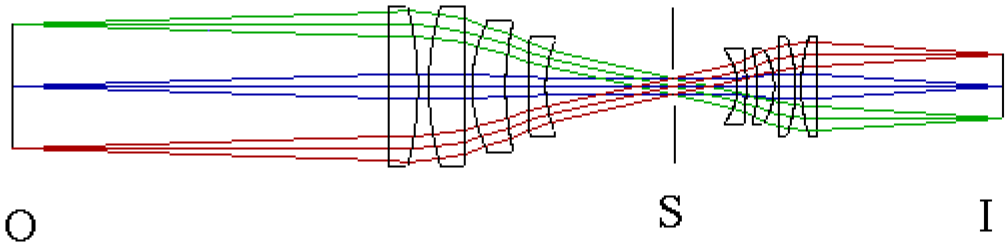
- Good approximation for telephoto optics
- Also called “parallel projection”: (x, y, z) → (x, y)
- What’s the projection matrix?

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \Rightarrow (x, y)$$

Orthographic (“telecentric”) lenses

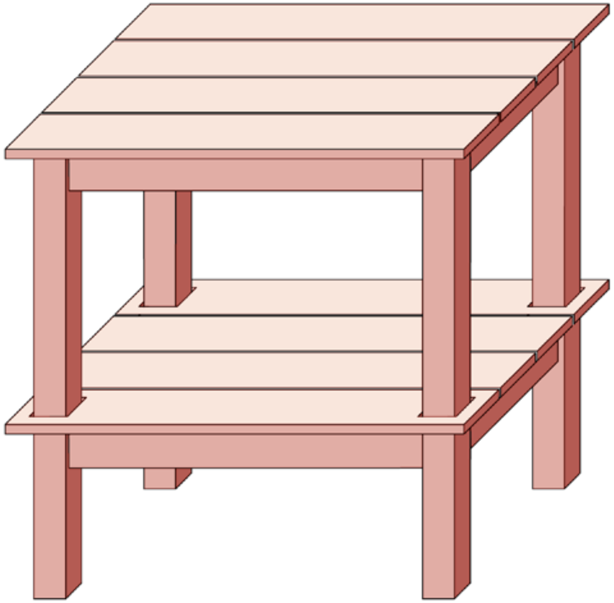


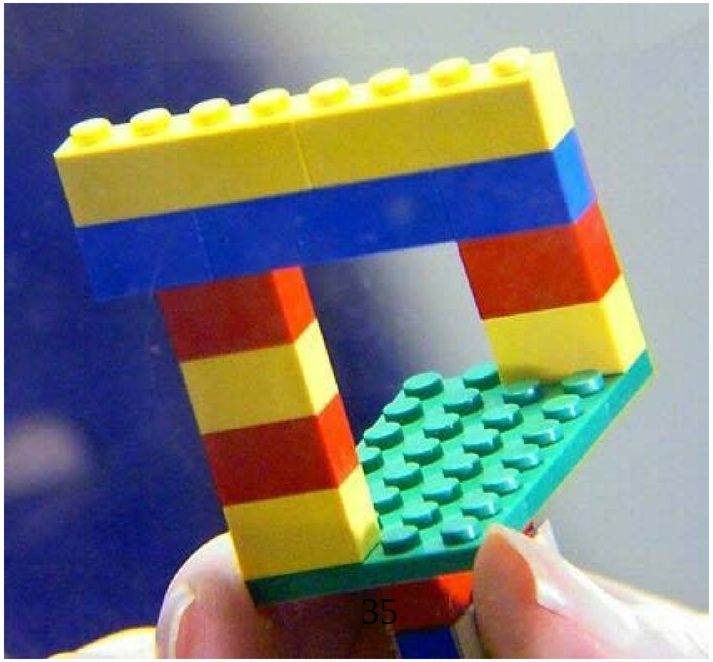
Navitar telecentric zoom lens



http://www.lhup.edu/~dsimanek/3d/telecent.htm

Orthographic Projection





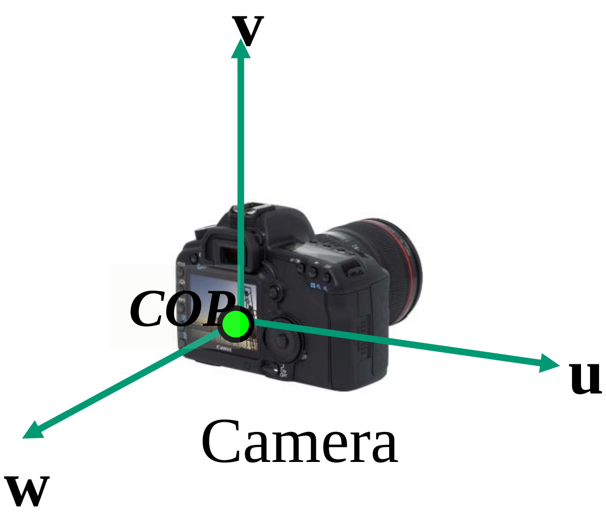
- What happens to parallel lines?
- What happens to angles?
- What happens to distances?

Camera parameters

How many numbers do we need to describe a camera?

- We need to describe its *pose* in the world
- We need to describe its internal *parameters*

A Tale of Two Coordinate Systems



Camera

"The World"

Two important coordinate systems:

1. *World* coordinate system
2. *Camera* coordinate system

Camera parameters

- To project a point (x,y,z) in *world* coordinates into a camera
- First transform (x,y,z) into *camera* coordinates
- Need to know
  - Camera position (in world coordinates)
  - Camera orientation (in world coordinates)
- Then project into the image plane
  - Need to know camera *intrinsics*
- These can all be described with matrices

3D Translation

- 3D translation is just like 2D with one more coordinate

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & tx \\ 0 & 1 & 0 & ty \\ 0 & 0 & 1 & tz \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$= [x+tx, y+ty, z+tz, 1]^T$$

3D Rotation (just the 3 x 3 part shown)

About X axis:
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$$

About Y:
$$\begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}$$

About Z axis:
$$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

General (orthonormal) rotation matrix used in practice:
$$\begin{bmatrix} r11 & r12 & r13 \\ r21 & r22 & r23 \\ r31 & r32 & r33 \end{bmatrix}$$

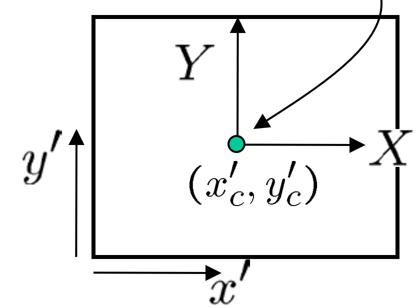
Camera parameters

A camera is described by several parameters

- Translation T of the optical center from the origin of world coords
- Rotation R of the image plane
- focal length f, principle point (x'_c, y'_c), pixel size (s_x, s_y)
- blue parameters are called "extrinsics," red are "intrinsics"

Projection equation

$$\mathbf{x} = \begin{bmatrix} wx \\ wy \\ w \end{bmatrix} = \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \mathbf{\Pi X}$$



- The projection matrix models the cumulative effect of all parameters
- Useful to decompose into a series of operations

identity matrix

$$\Pi = \begin{bmatrix} -fs_x & 0 & x'_c \\ 0 & -fs_y & y'_c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R}_{3x3} & \mathbf{0}_{3x1} \\ \mathbf{0}_{1x3} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{I}_{3x3} & \mathbf{T}_{3x1} \\ \mathbf{0}_{1x3} & 1 \end{bmatrix} \leftarrow [tx, ty, tz]^T$$

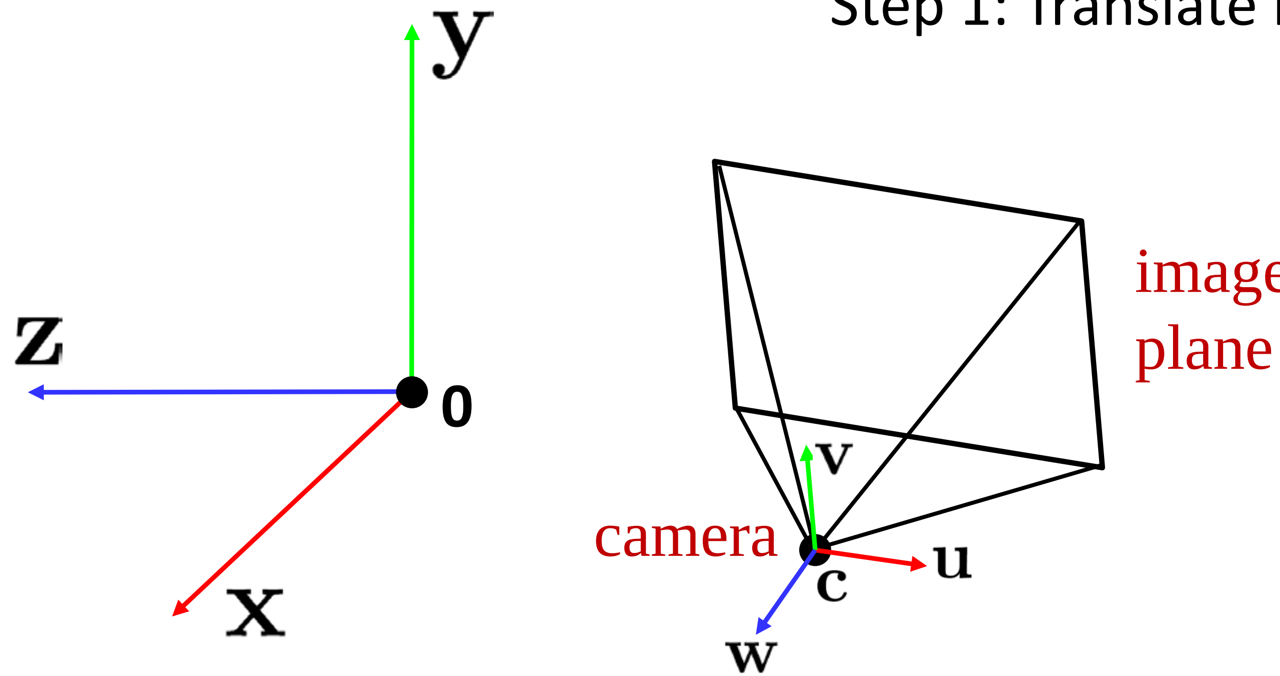
intrinsics projection rotation translation

- The definitions of these parameters are **not** completely standardized
  - especially intrinsics—varies from one book to another

Extrinsics

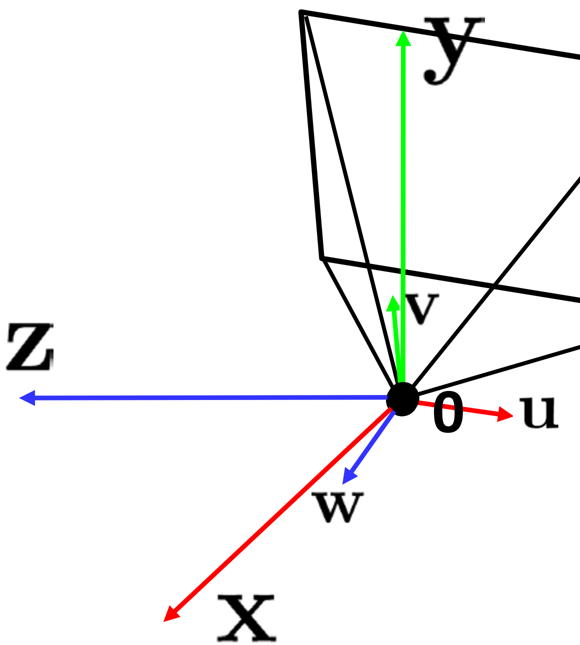
- How do we get the camera to "canonical form"?
  - (Center of projection at the origin, x-axis points right, y-axis points up, z-axis points backwards)

Step 1: Translate by -**c**



Extrinsics

- How do we get the camera to "canonical form"?
  - (Center of projection at the origin, x-axis points right, y-axis points up, z-axis points backwards)



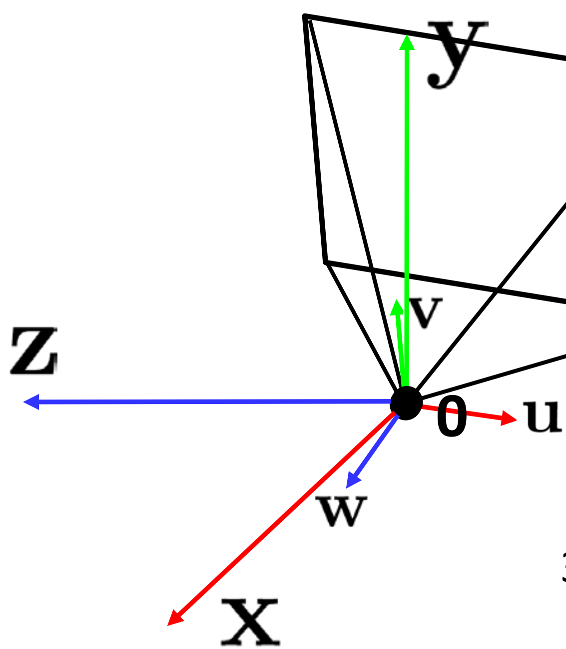
Step 1: Translate by -**c**

How do we represent translation as a matrix multiplication?

$$\mathbf{T} = \begin{bmatrix} \mathbf{I}_{3\times3} & -\mathbf{c} \\ 0 \quad 0 \quad 0 & 1 \end{bmatrix}$$

Extrinsics

- How do we get the camera to “canonical form”?
  - (Center of projection at the origin, x-axis points right, y-axis points up, z-axis points backwards)



Step 1: Translate by -**c**

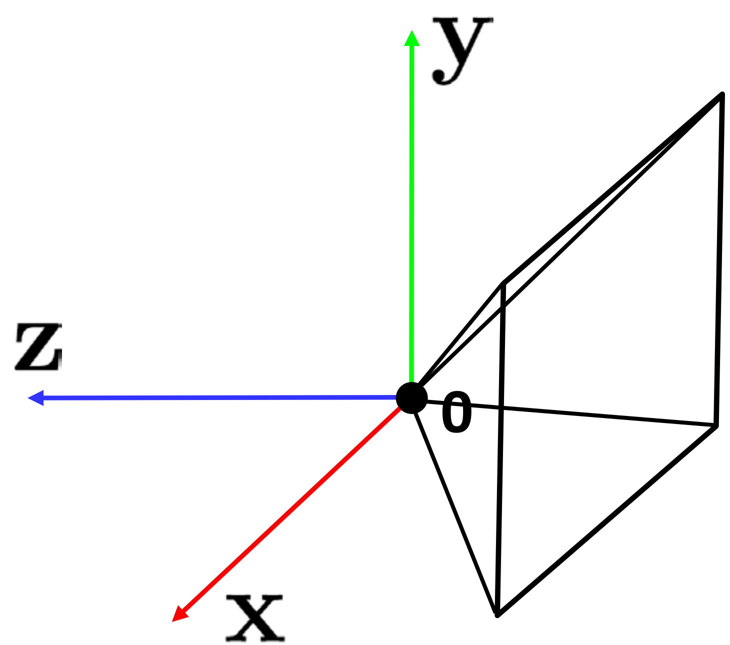
Step 2: Rotate by **R**

$$\mathbf{R} = \begin{bmatrix} \mathbf{u}^T \\ \mathbf{v}^T \\ \mathbf{w}^T \end{bmatrix}$$

3x3 rotation matrix

Extrinsics

- How do we get the camera to "canonical form"?
  - (Center of projection at the origin, x-axis points right, y-axis points up, z-axis points backwards)



Step 1: Translate by -**c**

Step 2: Rotate by **R**

$$\mathbf{R} = \begin{bmatrix} \mathbf{u}^T \\ \mathbf{v}^T \\ \mathbf{w}^T \end{bmatrix}$$

Perspective projection

$$\underbrace{\begin{bmatrix} -f & 0 & 0 \\ 0 & -f & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{K}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

(intrinsics) (converts from 3D rays in camera coordinate system to pixel coordinates)

in general, $\mathbf{K} = \begin{bmatrix} -f & s & c_x \\ 0 & -\alpha f & c_y \\ 0 & 0 & 1 \end{bmatrix}$

f is the focal length of the camera

α : **aspect ratio** (1 unless pixels are not square)

s : **skew** (0 unless pixels are shaped like rhombi/parallelograms)

(c_x, c_y) : **principal point** ((0,0) unless optical axis doesn't intersect projection plane at origin)

Focal length

- Can think of as "zoom"

24mm

50mm

200mm

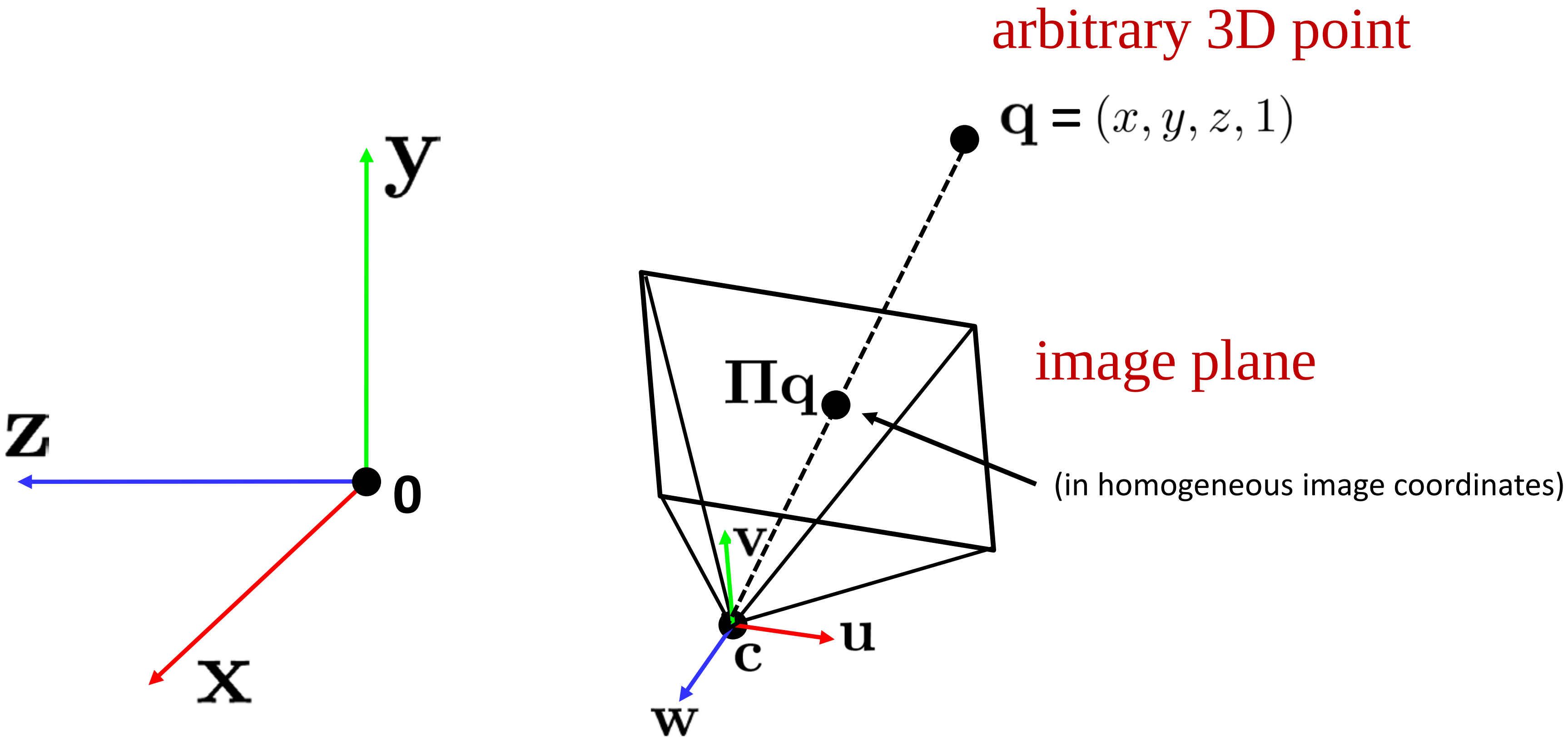
800mm

- Related to *field of view*

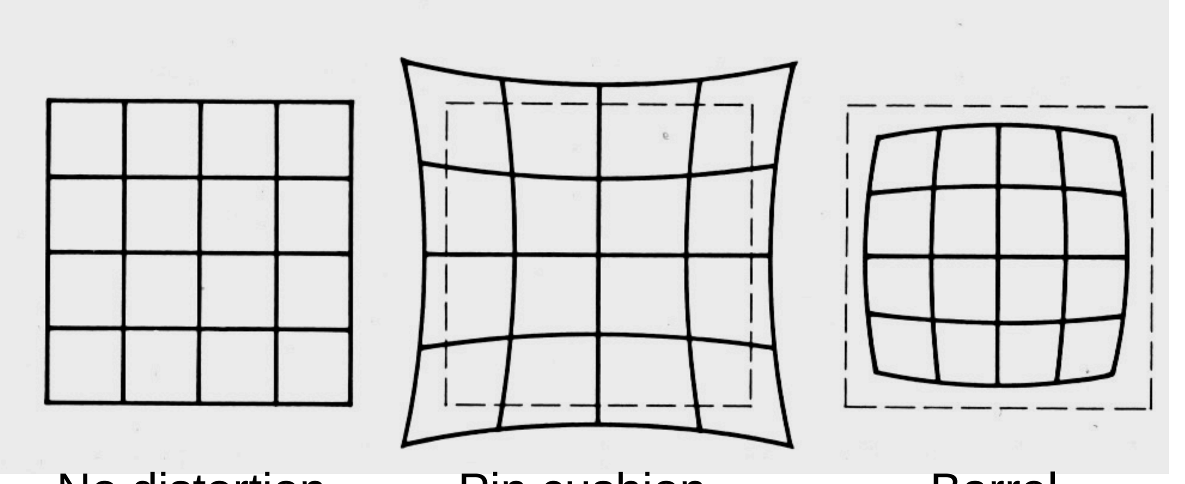
Projection matrix

$$\mathbf{\Pi} = \underbrace{\mathbf{K}}_{\text{intrinsics}} \underbrace{\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\text{projection}} \underbrace{\begin{bmatrix} & \mathbf{R} & & \begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\text{rotation}} \underbrace{\begin{bmatrix} & \mathbf{I}_{3\times3} & & -\mathbf{c} \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\text{translation}}}$$

Projection matrix



Distortion



No distortion Pin cushion Barrel

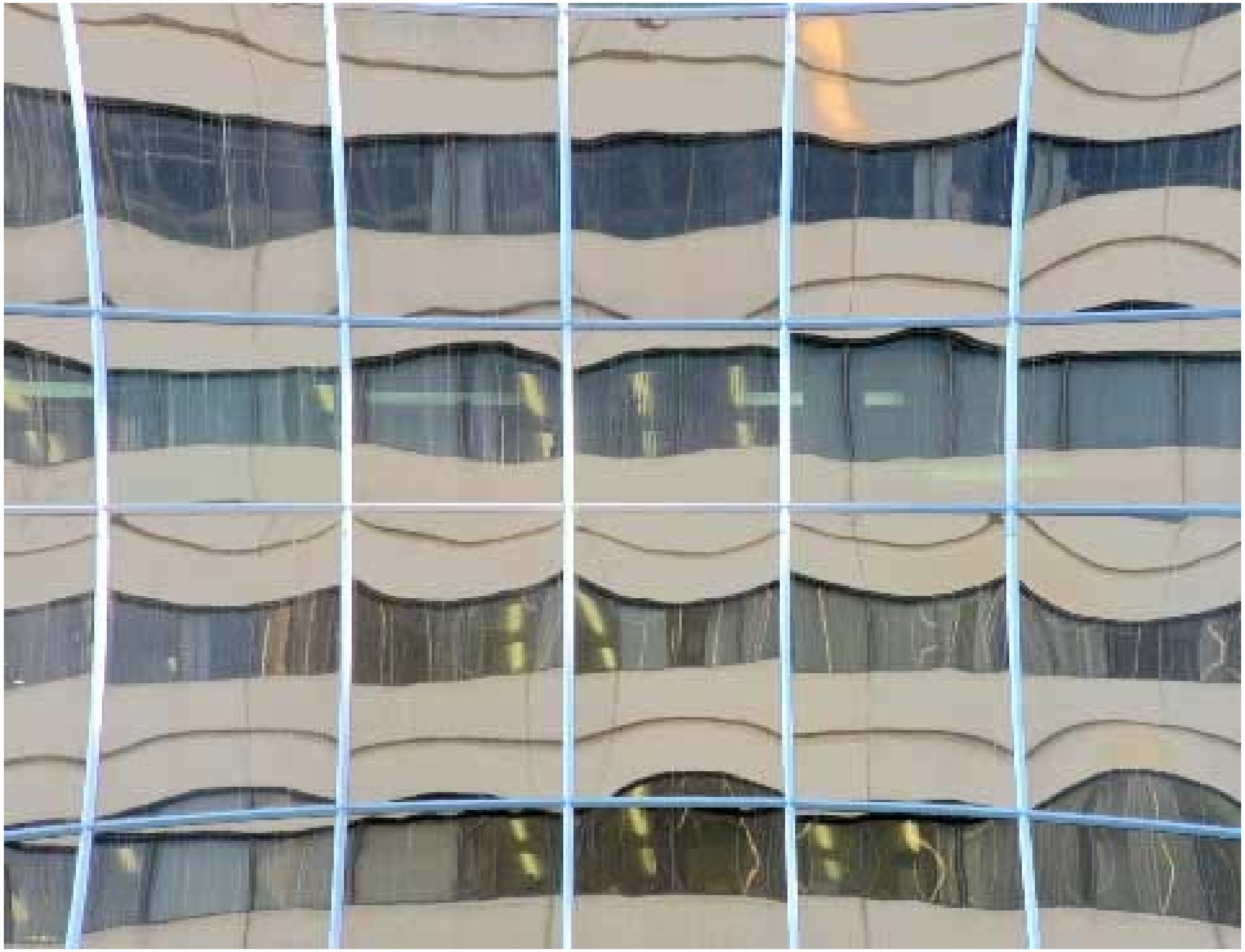
Radial distortion of the image

- Caused by imperfect lenses
- Deviations are most noticeable for rays that pass through the edge of the lens

Correcting radial distortion

from Helmut Dersch

http://blog.photoshopcreative.co.uk/general/fix-barrel-distortion/



Many other types of projection exist...

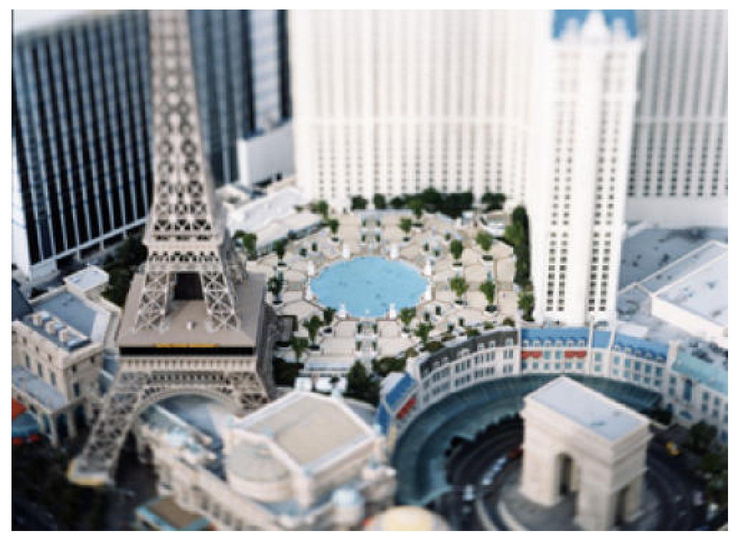
360 degree field of view…

Basic approach

- Take a photo of a parabolic mirror with an orthographic lens (Nayar)
  - http://www.cs.columbia.edu/CAVE/projects/cat_cam_360/gallery1/index.html
- Or buy one a lens from a variety of omnicam manufacturers…
  - See http://www.cis.upenn.edu/~kostas/omni.html

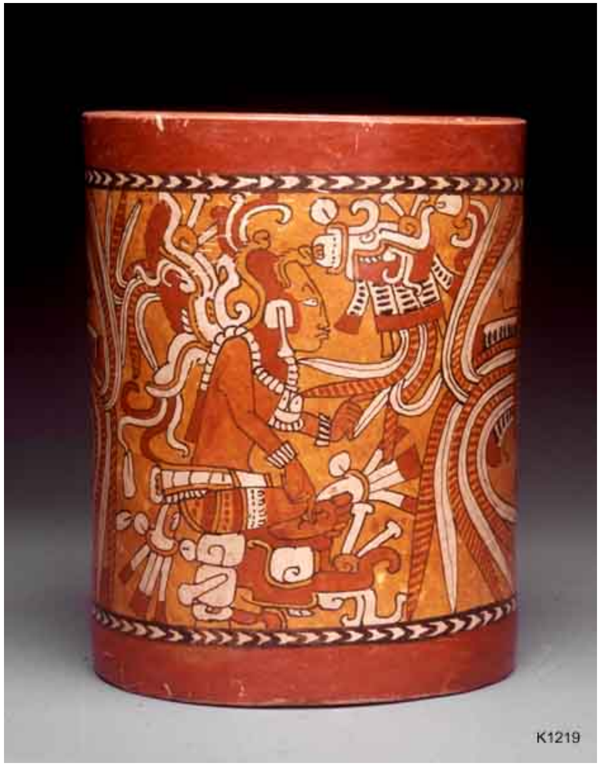
Tilt-shift

http://www.northlight-images.co.uk/article_pages/tilt_and_shift_ts-e.html



Tilt-shift images from Olivo Barbieri
and Photoshop imitations

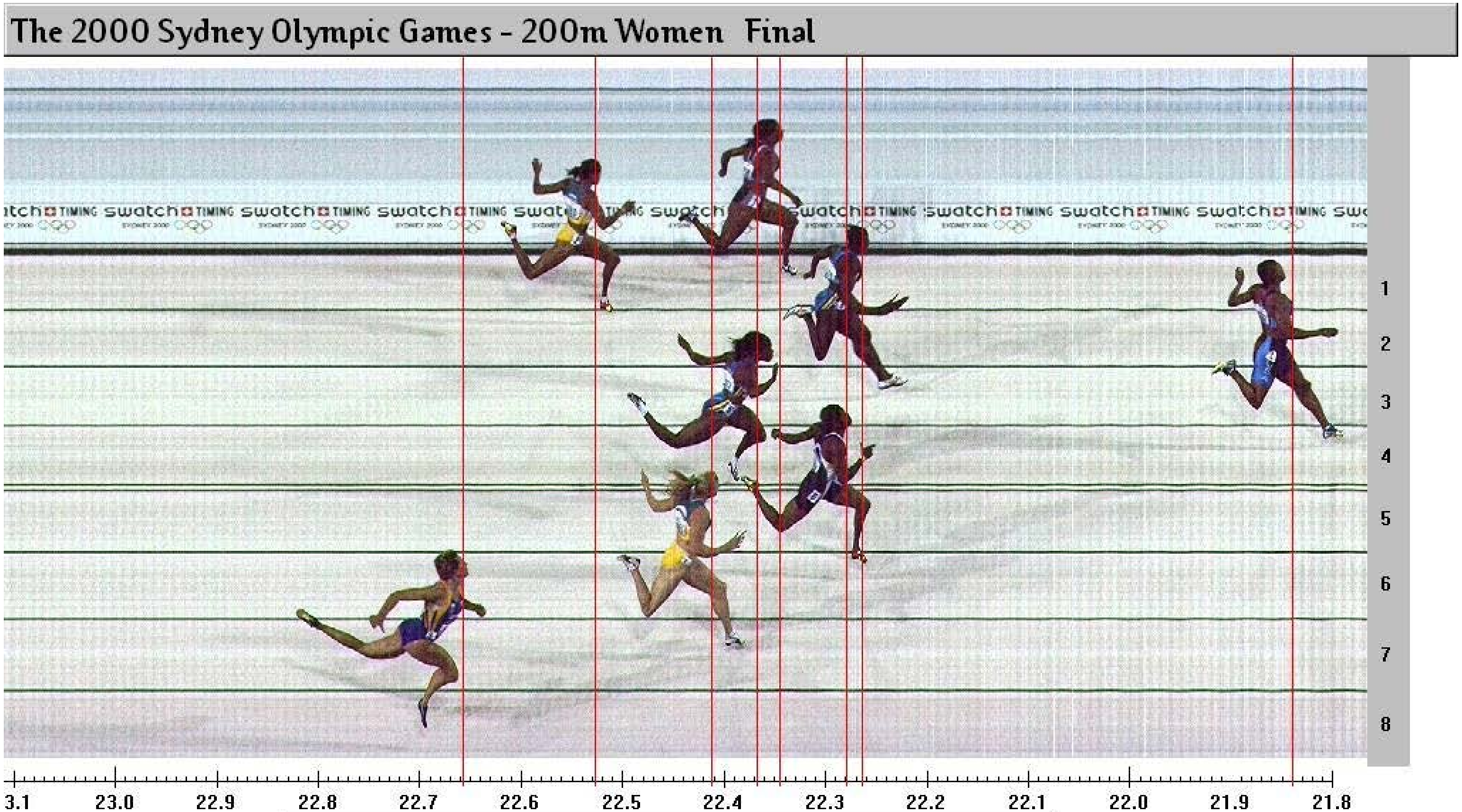
Rotating sensor (or object)



Rollout Photographs © Justin Kerr
http://research.famsi.org/kerrmaya.html

Also known as "cyclographs", "peripheral images"

Photofinish



Results Wind : + 0.7 m/s

Rank	La	Bib	Nu		Time	R_time
1.	4	3357	Jones Marion	USA	21.84	0.174
2.	3	1174	Davis-Thompson Pauline	BAH	22.27	0.185
3.	6	3058	Jayasinghe Susanthika	SRI	22.28	0.207
4.	1	2291	McDonald Beverly	JAM	22.35	0.151
5.	5	1178	Ferguson Debbie	BAH	22.37	0.196
6.	7	1111	Gainsford-Taylo Melinda	AUS	22.42	0.178
7.	2	1110	Freeman Cathy	AUS	22.53	0.235
8.	8	3239	Pintusevych Zhanna	UKR	22.66	0.190

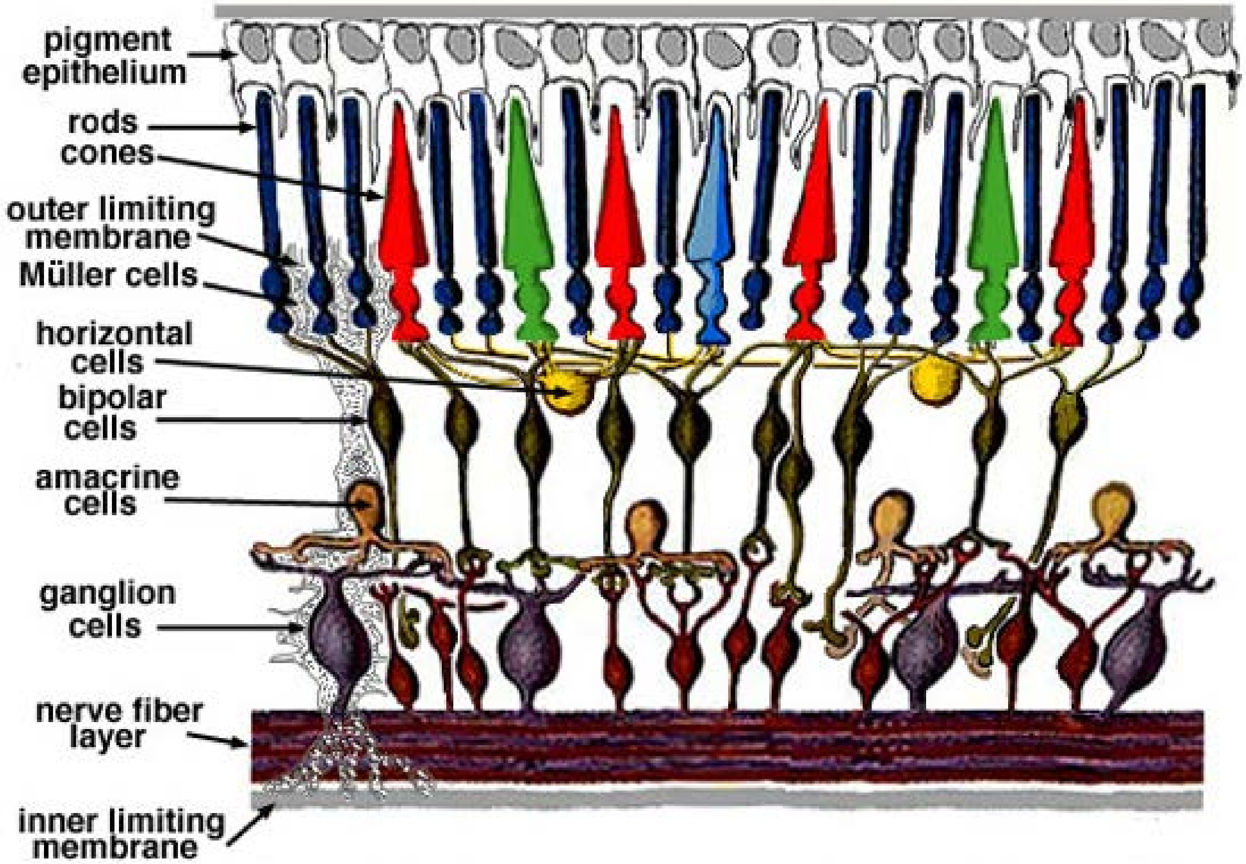
Start: 28. 9.2000 19:57:19.033 @414
Print: 28. 9.2000 20:00:54 @417

Scan'O'Vision Color
Race ID: W200FI00

TM © SOCOG 1996

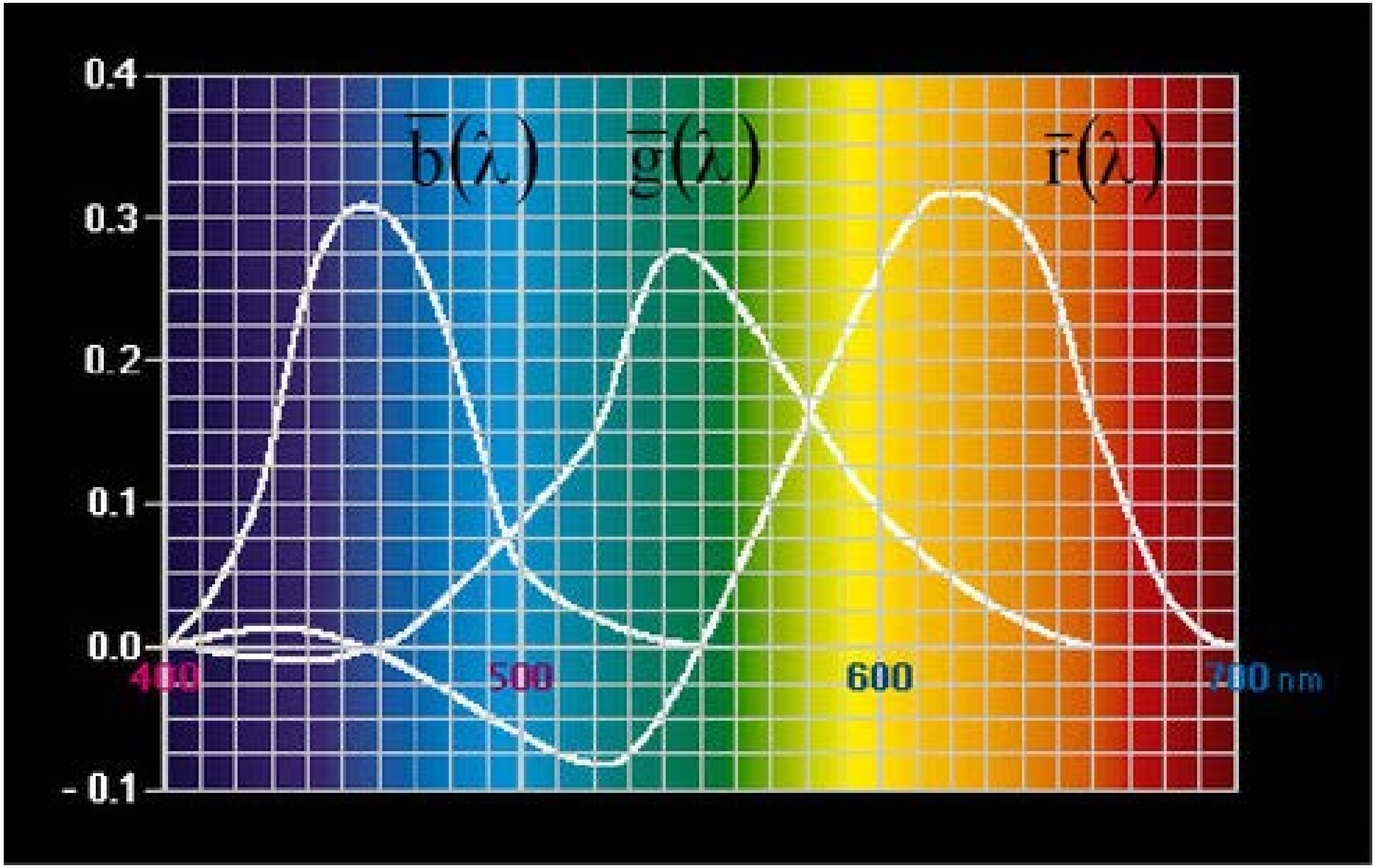
swatch TIMING

Human eye


pigment epithelium
rods
cones
outer limiting membrane
Müller cells
horizontal cells
bipolar cells
amacrine cells
ganglion cells
nerve fiber layer
inner limiting membrane

Colors

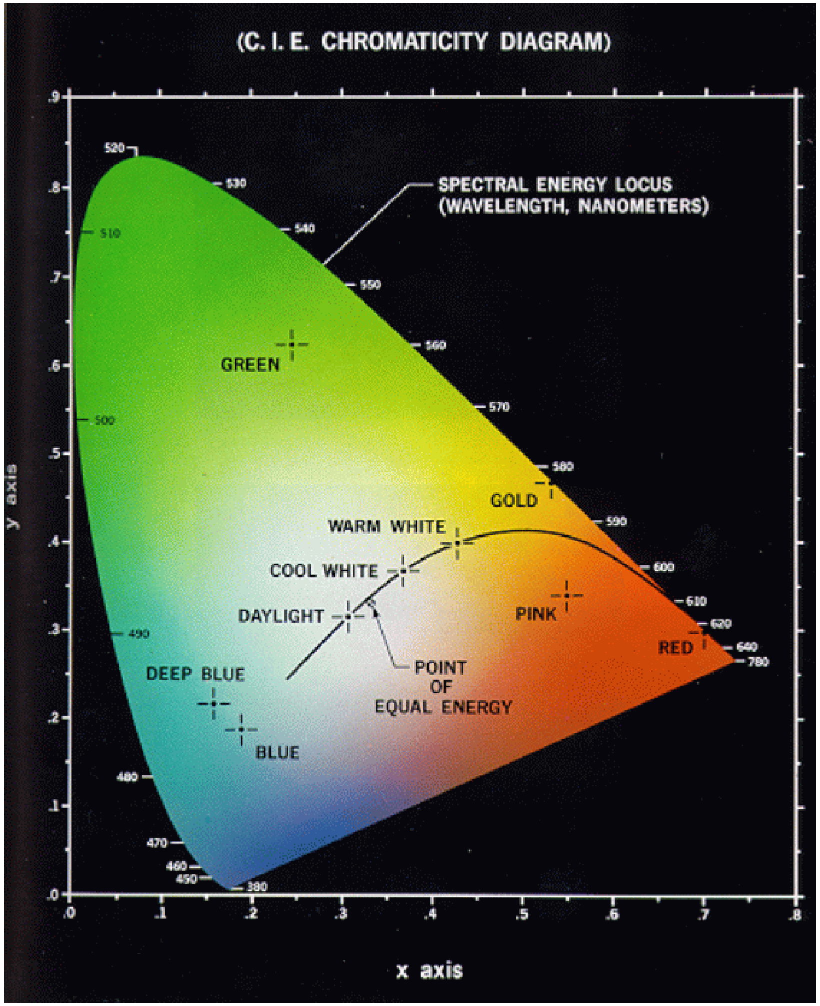
What colors do humans see?



RGB tristimulus values, 1931 RGB CIE

Colors

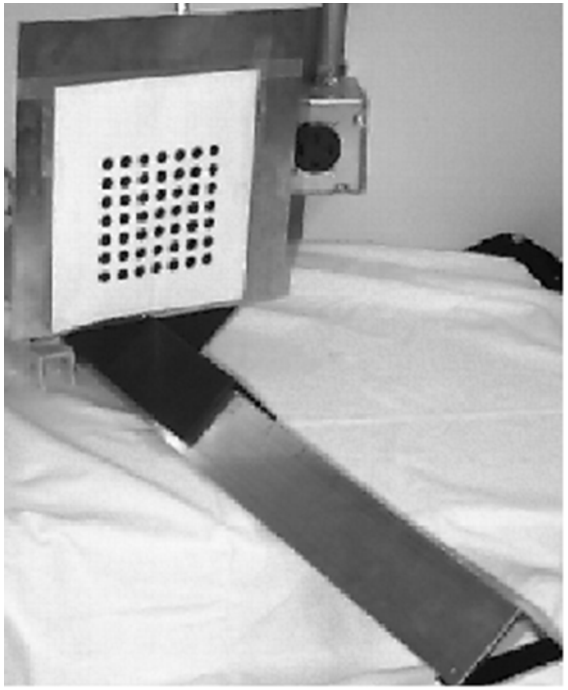
Plot of all visible colors (Hue and saturation):



Where does all this lead?

- We need it to understand stereo
- And 3D reconstruction
- It also leads into camera calibration, which is usually done in factory settings to solve for the camera parameters before performing an industrial task.
- The extrinsic parameters must be determined.
- Some of the intrinsic are given, some are solved for, some are improved.

Camera Calibration



The idea is to snap images at different depths and get a lot of 2D-3D point correspondences.

x1, y1, z1, u1, v1
x2, y2, z1, u2, v2
.
.
xn, yn, zn, un, vn

Then solve a system of equations to get camera parameters.