Image Formation and Cameras

Computer Vision I

CSE 252A

Lecture 4

Announcements

- HW0 is due on today
- HW1 will be assigned later today.
- Read Chapters 1 & 2 of Forsyth & Ponce
- (Subset of?) Final exam can be use for CSE MS Comprehensive Exam [Pending]

Projective Geometry

- Axioms of Projective Plane
  1. Every two distinct points define a line
  2. Every two distinct lines define a point (intersect at a point)
  3. There exists three points, A,B,C such that C does not lie on the line defined by A and B.
- Different than Euclidean (affine) geometry
- Projective plane is “bigger” than affine plane – includes “line at infinity”

Conversion

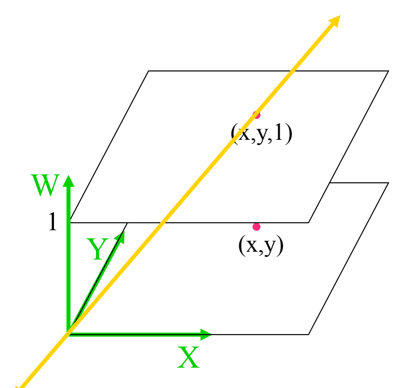
Euclidean → Homogenous → Euclidean

In 2-D

- Euclidean → Homogenous: $(x,y) \rightarrow \lambda(x,y,1) = (\lambda x, \lambda y, \lambda)$
- Homogenous -> Euclidean: $(x,y,w) \rightarrow (x/w, y/w)$

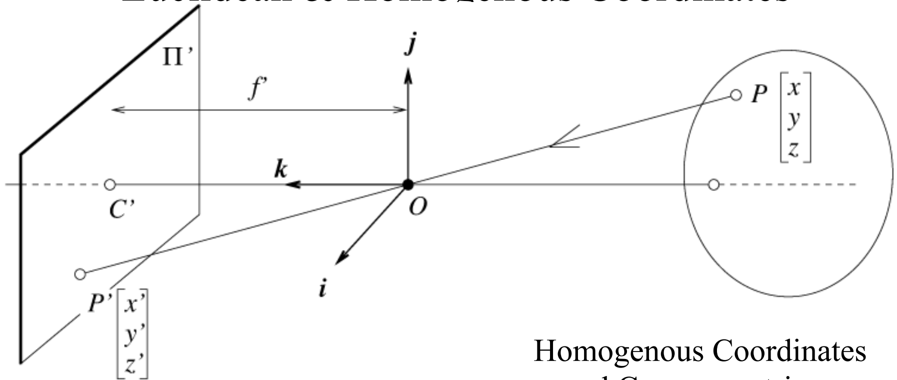
In 3-D

- Euclidean → Homogenous: $(x, y, z) \rightarrow \lambda(x, y, z, 1)$
- Homogenous → Euclidean: $(x, y, z, w) \rightarrow (x/w, y/w, z/w)$



Note: If w=0, then (x,y,w) is a point at infinity

The equation of perspective projection: Euclidean & Homogenous Coordinates



Cartesian coordinates

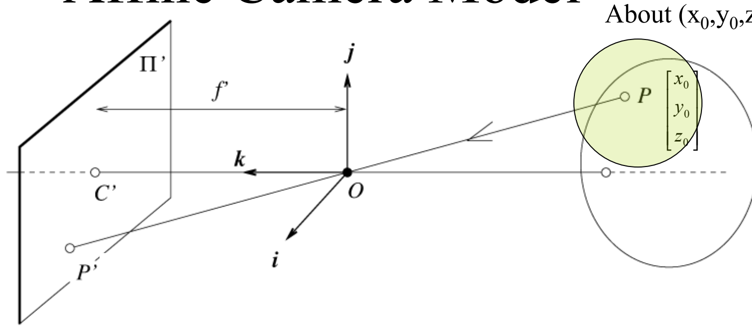
$$(x,y,z) \rightarrow (f\frac{x}{z}, f\frac{y}{z})$$

Homogenous Coordinates and Camera matrix

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

Affine Camera Model

Appropriate in Neighborhood About (x_0,y_0,z_0)



- Take perspective projection equation, and perform Taylor series expansion about some point (x_0,y_0,z_0).
- Drop terms that are higher order than linear.
- Resulting expression is affine camera model

Simplified Camera Models

Perspective Projection

↓

Affine Camera Model

↓

Scaled Orthographic Projection

↓

Orthographic Projection

Approximation: Taylor Series Expansion about (X_0, Y_0, Z_0)

Case of affine where $X_0 = Y_0 = 0$

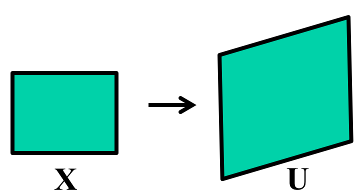
Case of scaled orthograhic where $f/Z_0 = 1$

Projective Transformations

Mappings from a plane to a plane

- 3 x 3 linear transformation of homogenous coordinates
- Points map to points
- Lines map to lines
- If u_3=0, (x_1, x_2, x_3) maps to a point at infinity.

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{21} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$



Mapping from a Plane to a Plane under Perspective is given by a projective transform H

x' = Hx H is a 3x3 matrix,
x, x' are 3x1 vectors of
homogenous coordinates

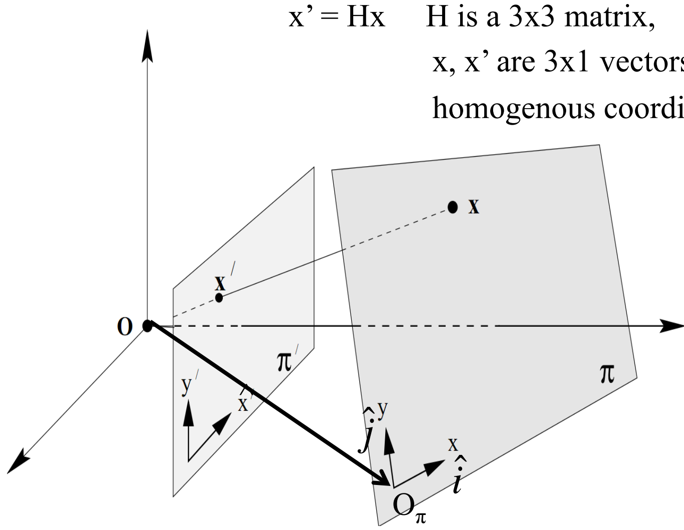
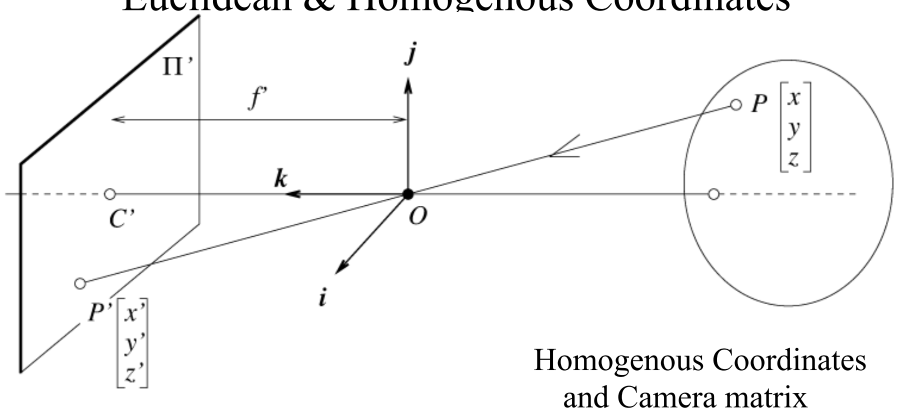


Figure borrowed from Hartley and Zisserman "Multiple View Geometry in computer vision"

The equation of perspective projection: Euclidean & Homogenous Coordinates



Homogenous Coordinates and Camera matrix

Cartesian coordinates

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

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

3D point coordinates on plane $\boldsymbol{\pi}$

$$\mathbf{x} = x\hat{i} + y\hat{j} + O_\pi$$

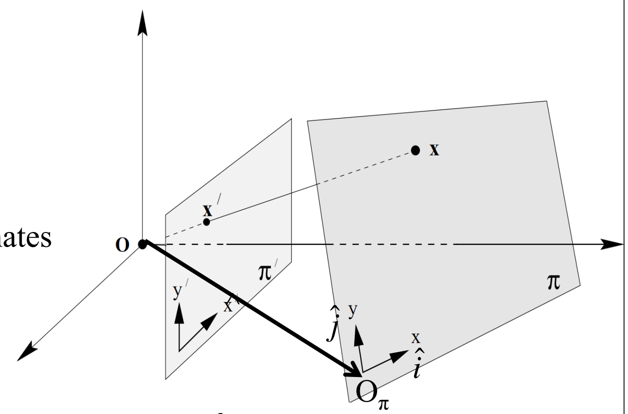
As a matrix equation

$$\mathbf{x} = \begin{bmatrix} \hat{i} & \hat{j} & O_\pi \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

In homogenous coordinates

$$\mathbf{x} = \begin{bmatrix} \hat{i} & \hat{j} & O_\pi \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

4x3



And taking perspective projection onto π'

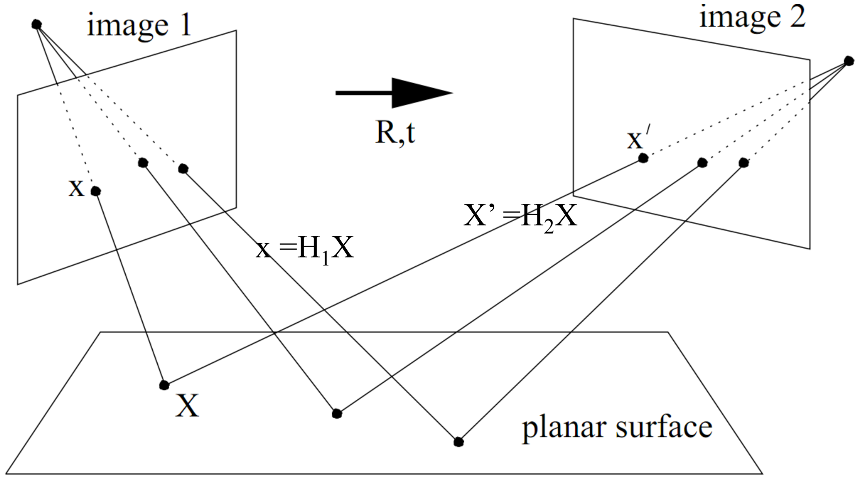
$$\begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \begin{bmatrix} \hat{i} & \hat{j} & O_\pi \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = H \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \longrightarrow \mathbf{x}' = H\mathbf{x}$$

Application of Homography: OCRs, scan,…

$$\mathbf{x}' = \mathbf{H}\,\mathbf{x}$$

Homography: Can be estimated from four points.

Planar Homography



$$x' = H_2X = H_2(H_1^{-1}x) = (H_2H_1^{-1})x$$

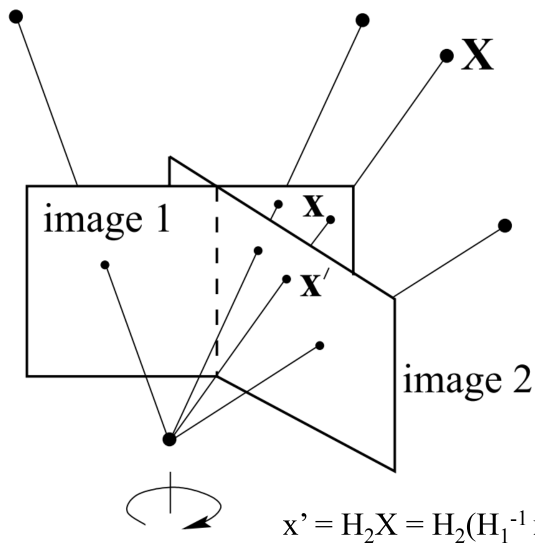
Application: Two photos of a white board

Figure borrowed from Hartley and Zisserman "Multiple View Geometry in computer vision"

What about the mapping from this slide to your eye?

If your eye were an image plane?

Planar Homography: Pure Rotation



$$x' = H_2X = H_2(H_1^{-1}x) = (H_2H_1^{-1})x$$

Application: Panoramas

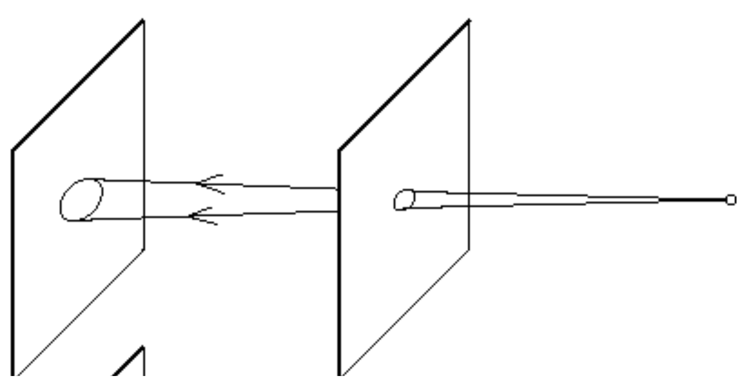
 Figure borrowed from Hartley and Zisserman "Multiple View Geometry in computer vision"

Application: Panoramas and image stitching

All images are warped to central image

Lenses

Beyond the pinhole Camera
Getting more light – Bigger Aperture

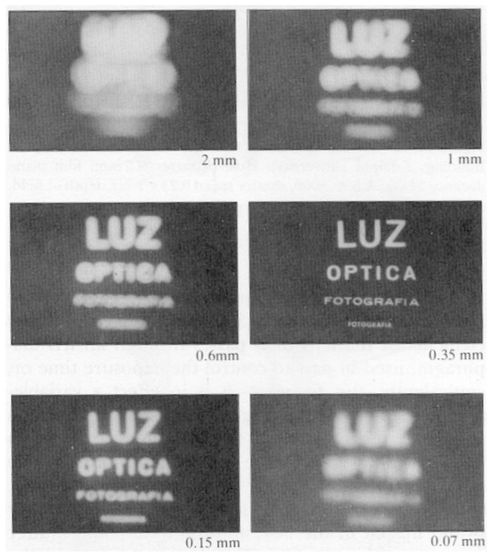


Pinhole Camera Images with Variable Aperture

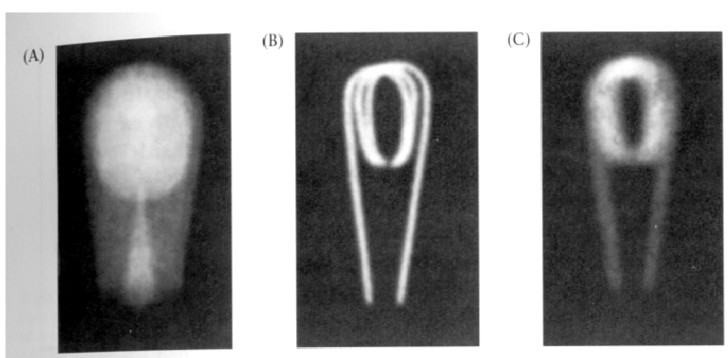
2 mm **1mm**

.6 mm **.35 mm**

.15 mm **.07 mm**

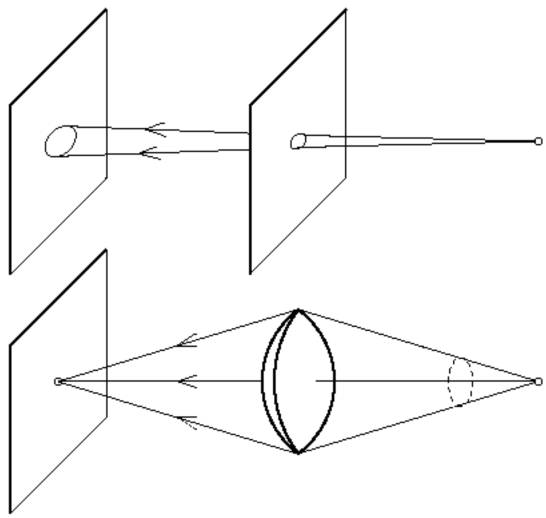


Limits for pinhole cameras

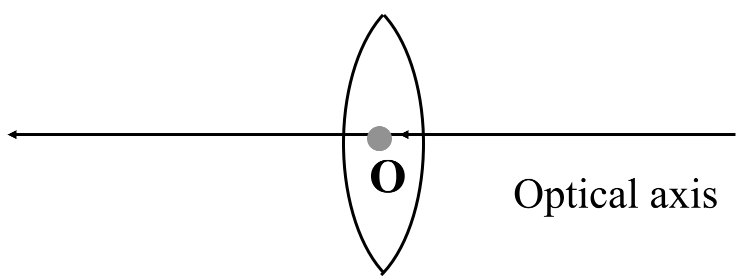


2.18 DIFFRACTION LIMITS THE QUALITY OF PINHOLE OPTICS. These three images of a bulb filament were made using pinholes with decreasing size. (A) When the pinhole is relatively large, the image rays are not properly converged, and the image is blurred. (B) Reducing the size of the pinhole improves the focus. (C) Reducing the size of the pinhole further worsens the focus, due to diffraction. From Ruechardt, 1958.

The reason for lenses

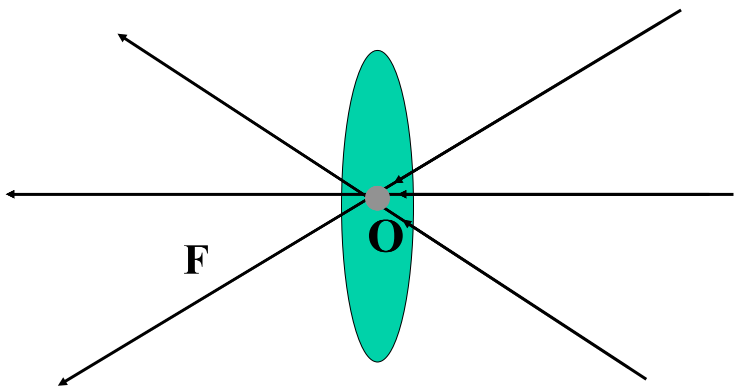


Thin Lens



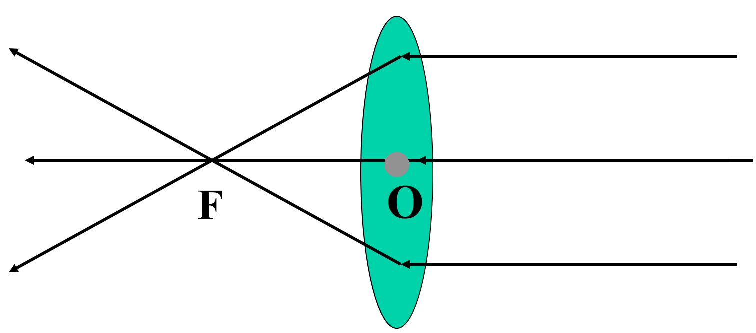
- Rotationally symmetric about optical axis.
- Spherical interfaces.

Thin Lens: Center



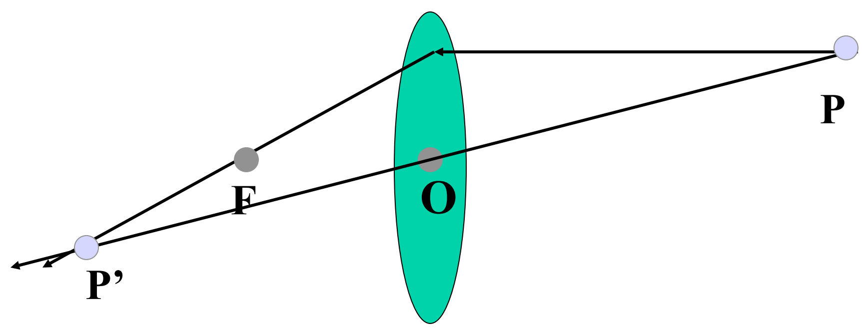
- All rays that enter lens along line pointing at **O** emerge in same direction.

Thin Lens: Focus



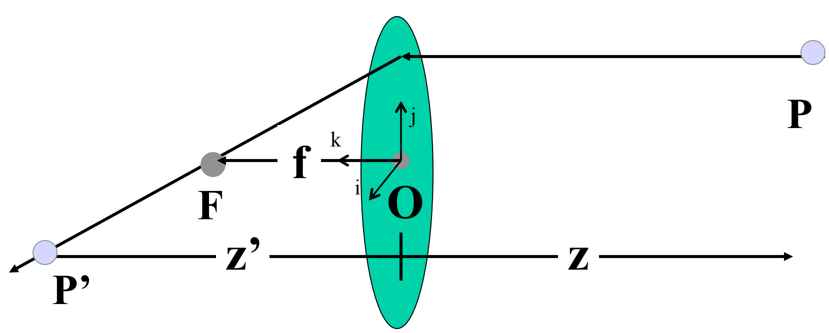
Parallel lines pass through the Focal Point, F

Thin Lens: Image of Point



All rays passing through lens and starting at **P** converge upon **P'**

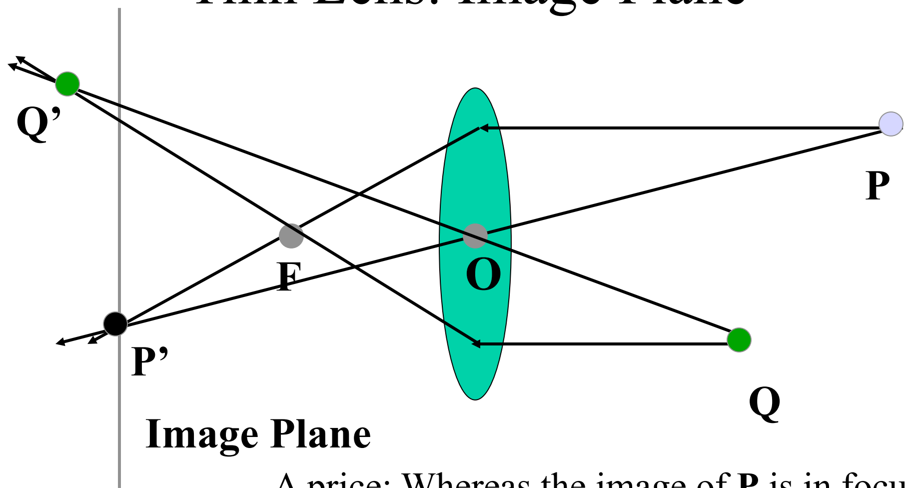
Thin Lens: Image of Point



$$\frac{1}{z'} - \frac{1}{z} = \frac{1}{f}$$

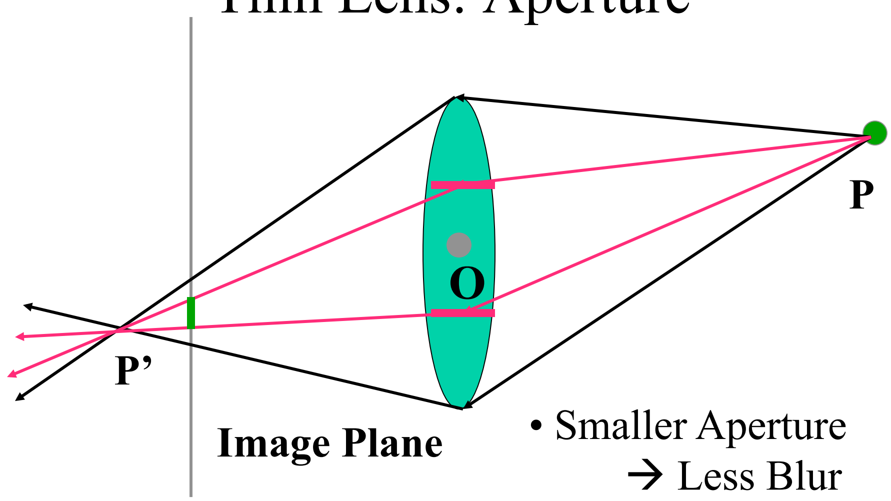
Where z, z', and f are the z-coordinates of P, P, and F. i.e., z is negative.

Thin Lens: Image Plane



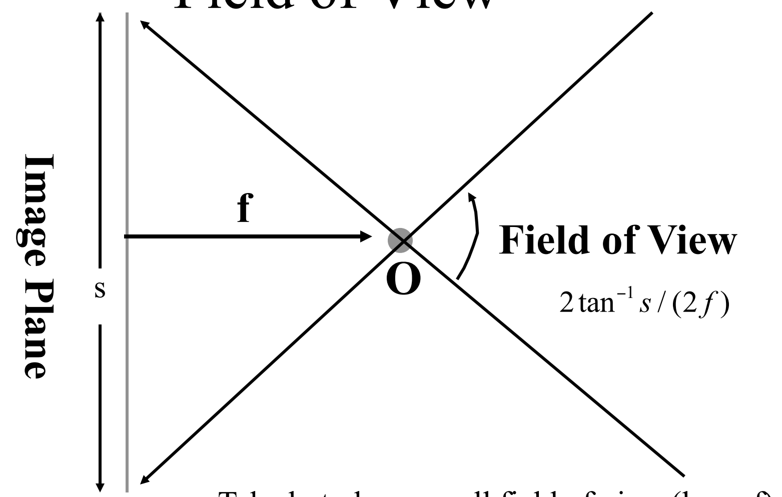
A price: Whereas the image of **P** is in focus, the image of **Q** isn't.

Thin Lens: Aperture



- Smaller Aperture → Less Blur
- Pinhole → No Blur

Field of View



- Telephoto lens: small field of view (large f)
- Wide angle lens: large field of view(small f)
- Zoom lens: variable f using compound lens

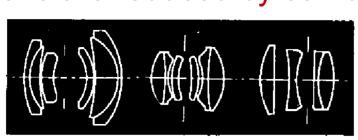
Deviations from the lens model

Deviations from this ideal are *aberrations*

Two types

1. Geometrical
   - distortion
   - spherical aberration
   - astigmatism
   - coma
2. Chromatic

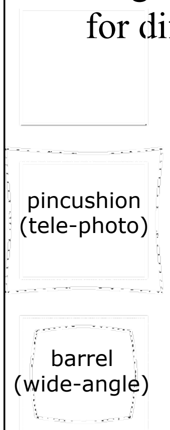
Aberrations are reduced by combining lenses

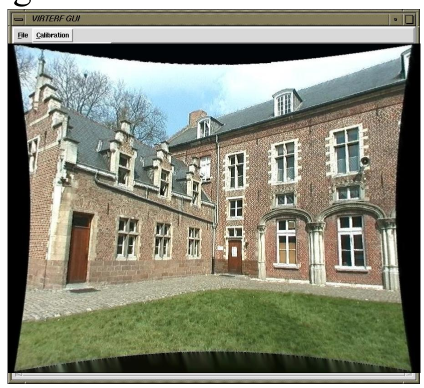


Compound lenses

Distortion

magnification/focal length different for different angles of inclination



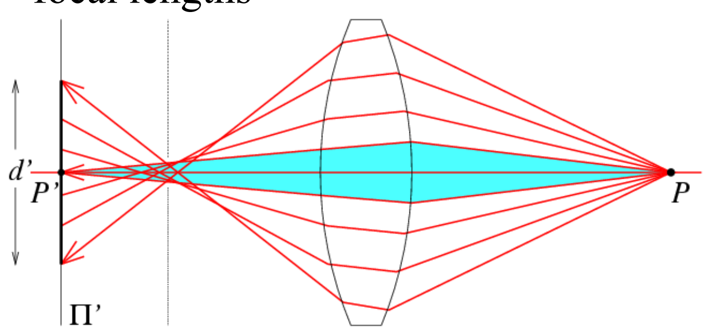


Can be corrected! (if parameters are know)

Spherical aberration

Rays parallel to the axis do not converge

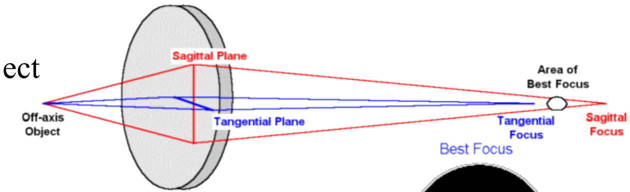
Outer portions of the lens yield smaller focal lengths

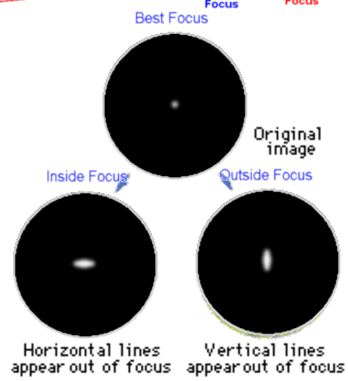


Astigmatism

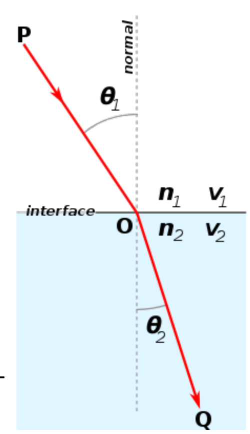
An optical system with astigmatism is one where rays that propagate in two perpendicular planes have different foci. If an optical system with astigmatism is used to form an image of a cross, the vertical and horizontal lines will be in sharp focus at two different distances.

object





Chromatic aberration



Snells Law

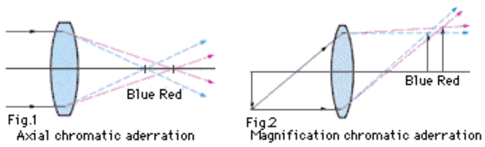
n_i is the index of refraction

λ_i is the wavelength of light in the medium

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{n_2}{n_1} = \frac{\lambda_1}{\lambda_2}$$

Chromatic aberration

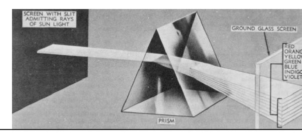
rays of different wavelengths focused in different planes



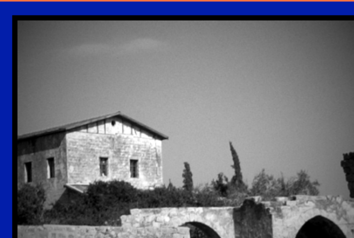
The image is blurred and appears colored at the fringe.

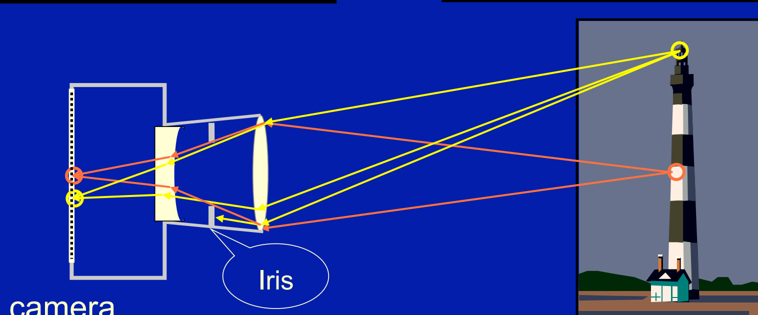
cannot be removed completely

sometimes *achromatization* is achieved for more than 2 wavelengths



Vignetting: Spatial Non-Uniformity





Litvinov & Schechner, *radiometric nonidealities*

Vignetting


- Only part of the light reaches the sensor
- Periphery of the image is dimmer