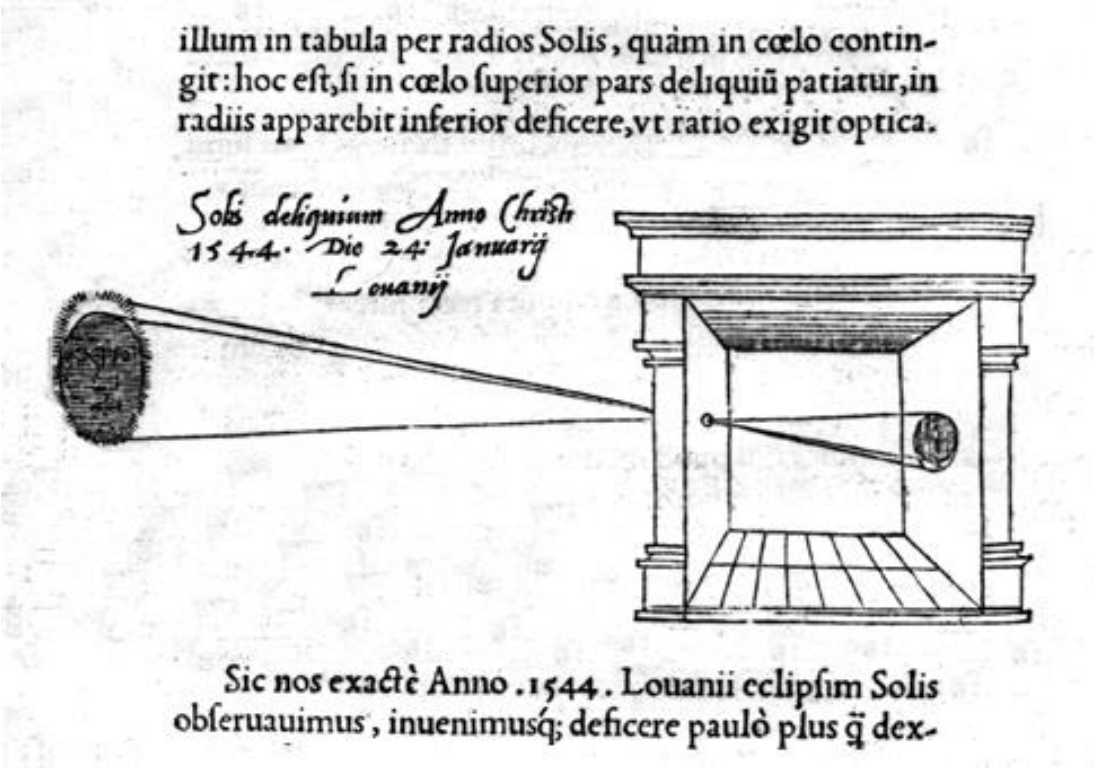
Image formation

Thanks to Peter Corke and Chuck Dyer for the use of some slides

Image Formation

- Vision infers world properties form images.
- How do images depend on these properties?
- Two key elements
  - Geometry
  - Radiometry
  - We consider only simple models of these

Camera Obscura

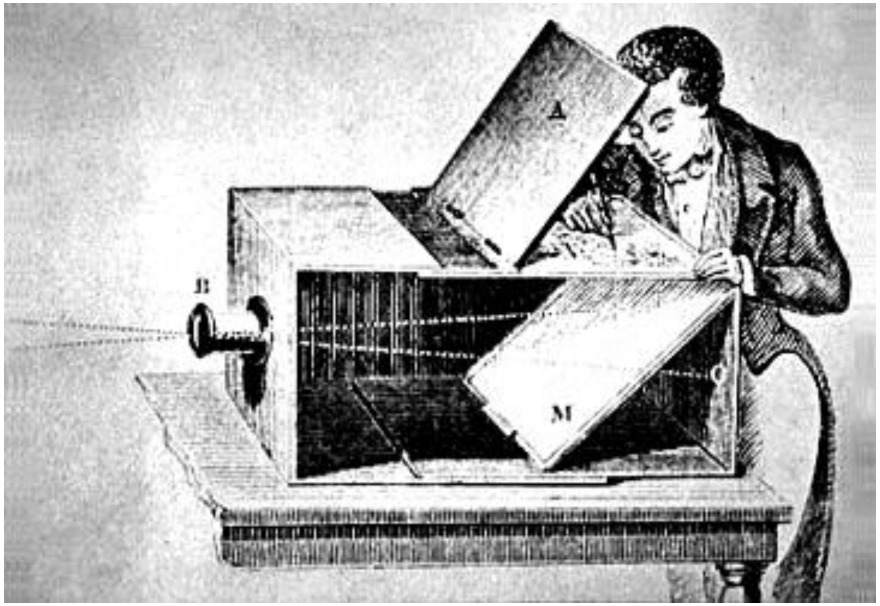

illum in tabula per radios Solis, quàm in cœlo contin-
git: hoc eſt, ſi in cœlo ſuperior pars deliquiũ patiatur, in
radiis apparebit inferior deficere, vt ratio exigit optica.

Sic nos exactè Anno .1544. Louanii eclipſim Solis
obſeruauimus, inuenimuſq; deficere paulò plus q̃ dex-

"When images of illuminated objects ... penetrate through a small hole into a very dark room ... you will see [on the opposite wall] these objects in their proper form and color, reduced in size ... in a reversed position, owing to the intersection of the rays".

Da Vinci

http://www.acmi.net.au/AIC/CAMERA_OBSCURA.html (Russell Naughton)



- Used to observe eclipses (eg., Bacon, 1214-1294)
- By artists (eg., Vermeer).

Jetty at Margate England, 1898.

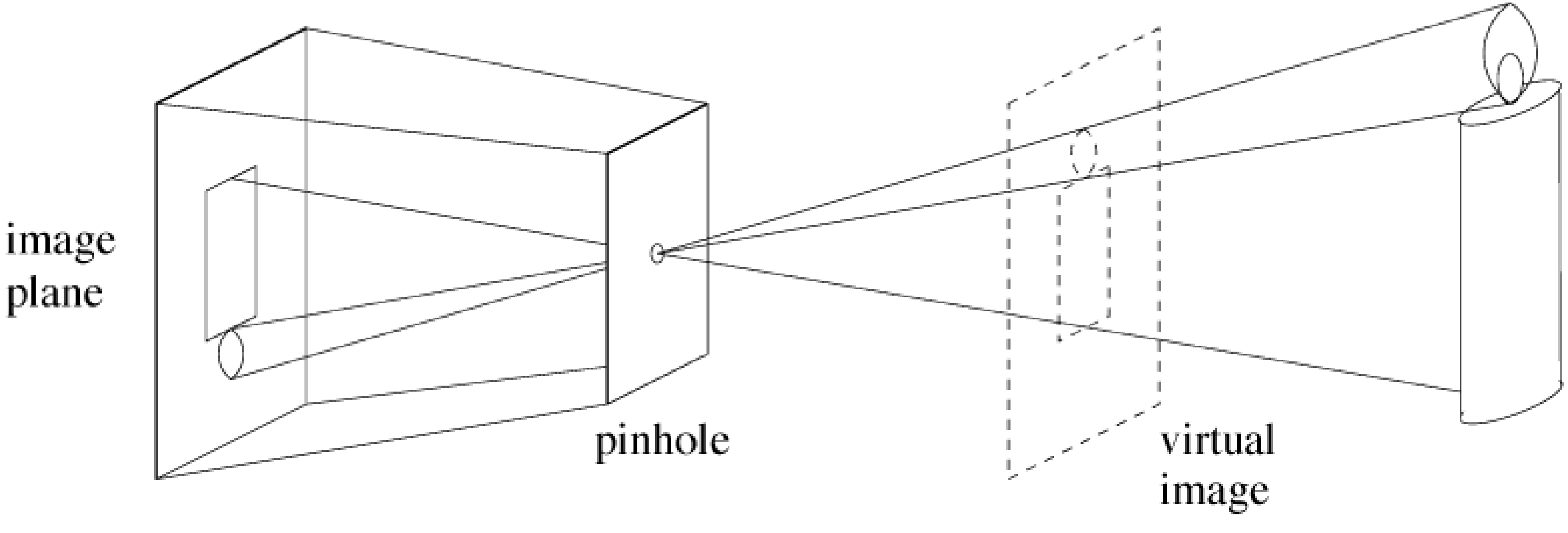
http://brightbytes.com/cosite/collection2.html (Jack and Beverly Wilgus)

Cameras

- First photograph due to Niepce
- First on record shown in the book - 1822

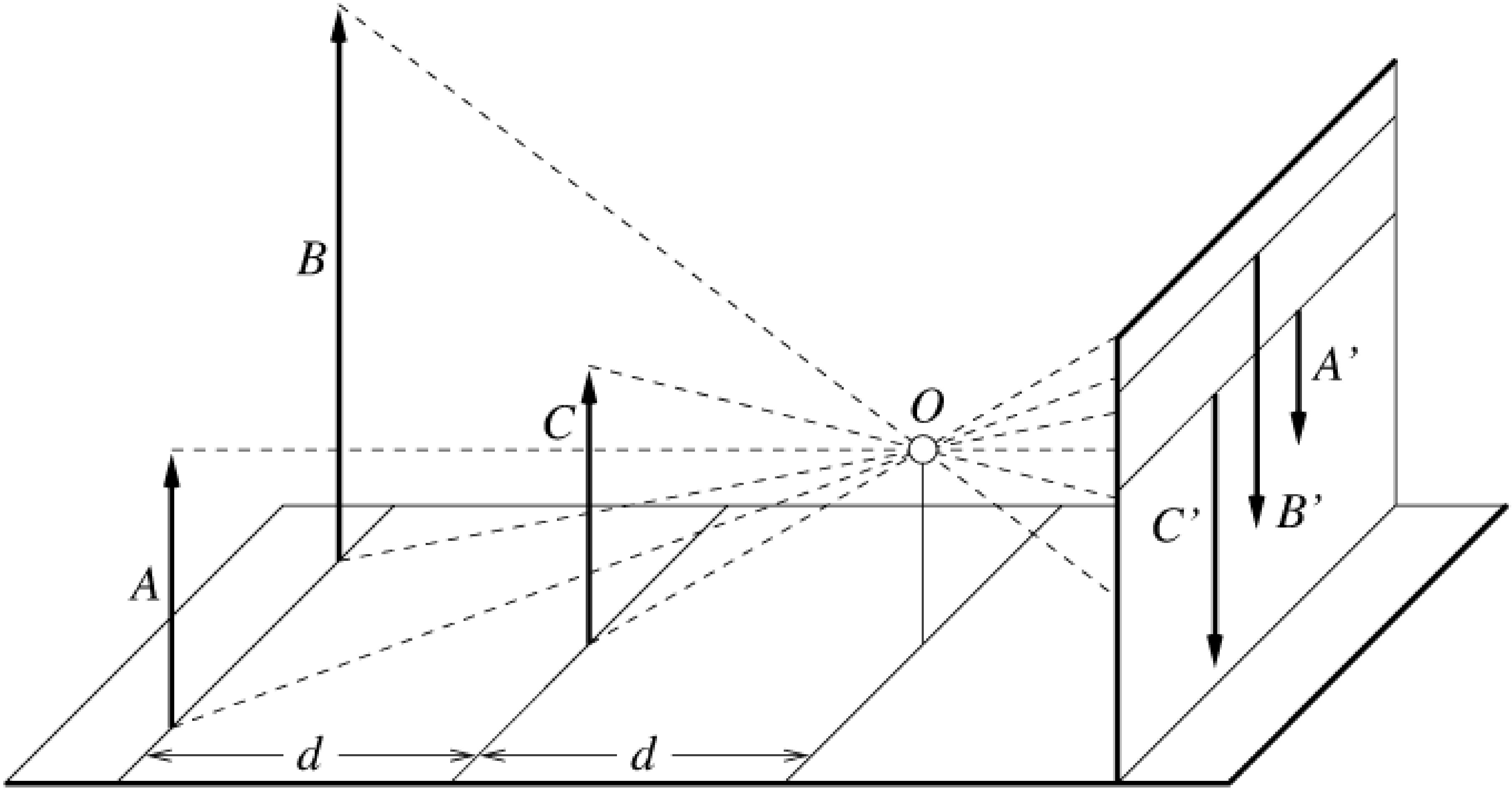
Pinhole cameras

- Abstract camera model - box with a small hole in it
- Pinhole cameras work in practice



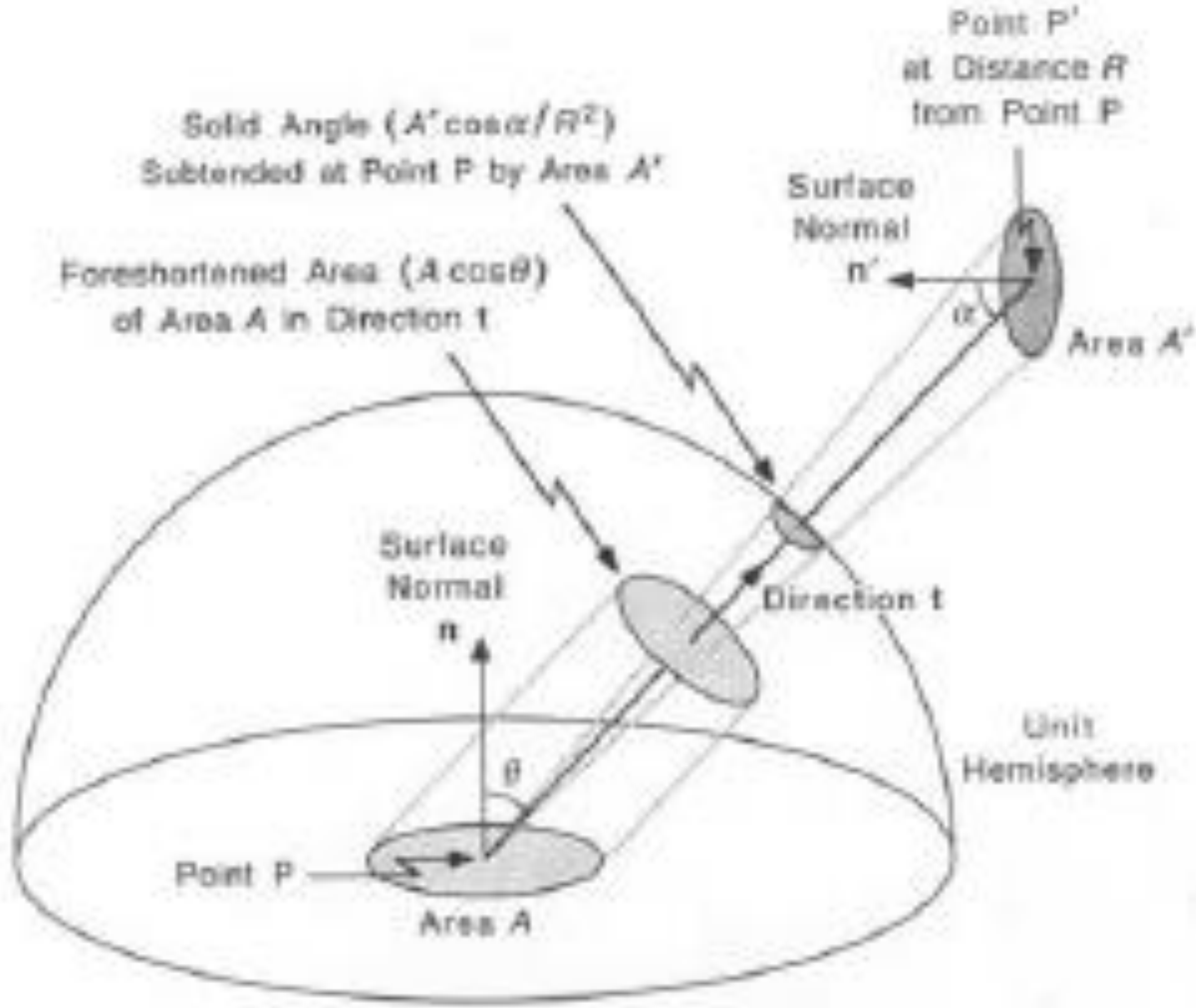
(Forsyth & Ponce)

Distant objects are smaller

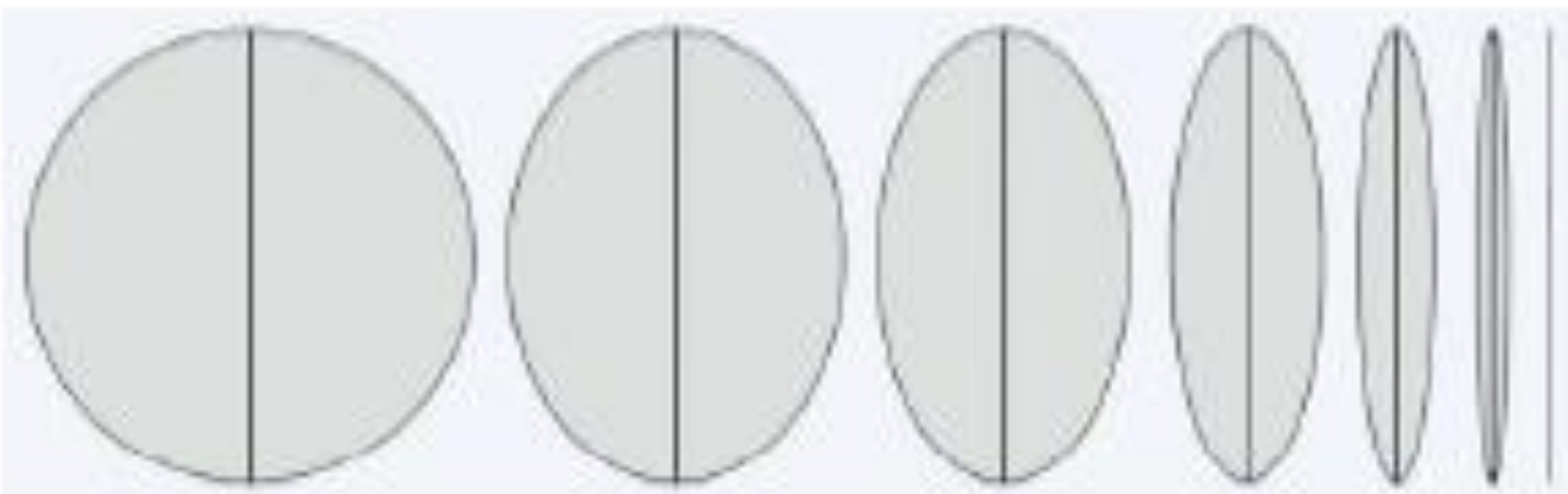


(Forsyth & Ponce)

Tilted Objects are Foreshortened

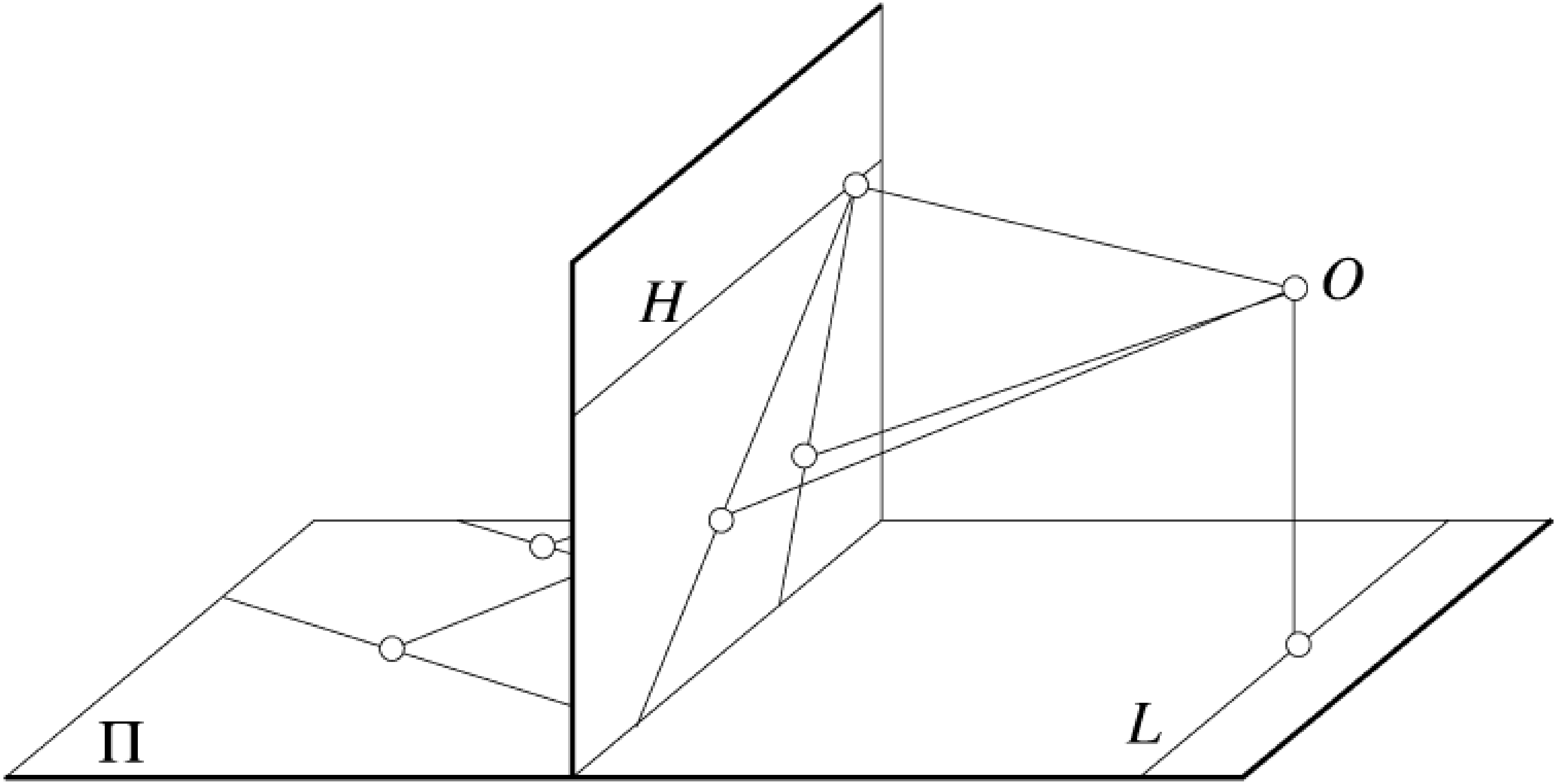


Tilted Objects are Foreshortened



Parallel lines meet

Common to draw image plane *in front* of the focal point.
Moving the image plane merely scales the image.



(Forsyth & Ponce)

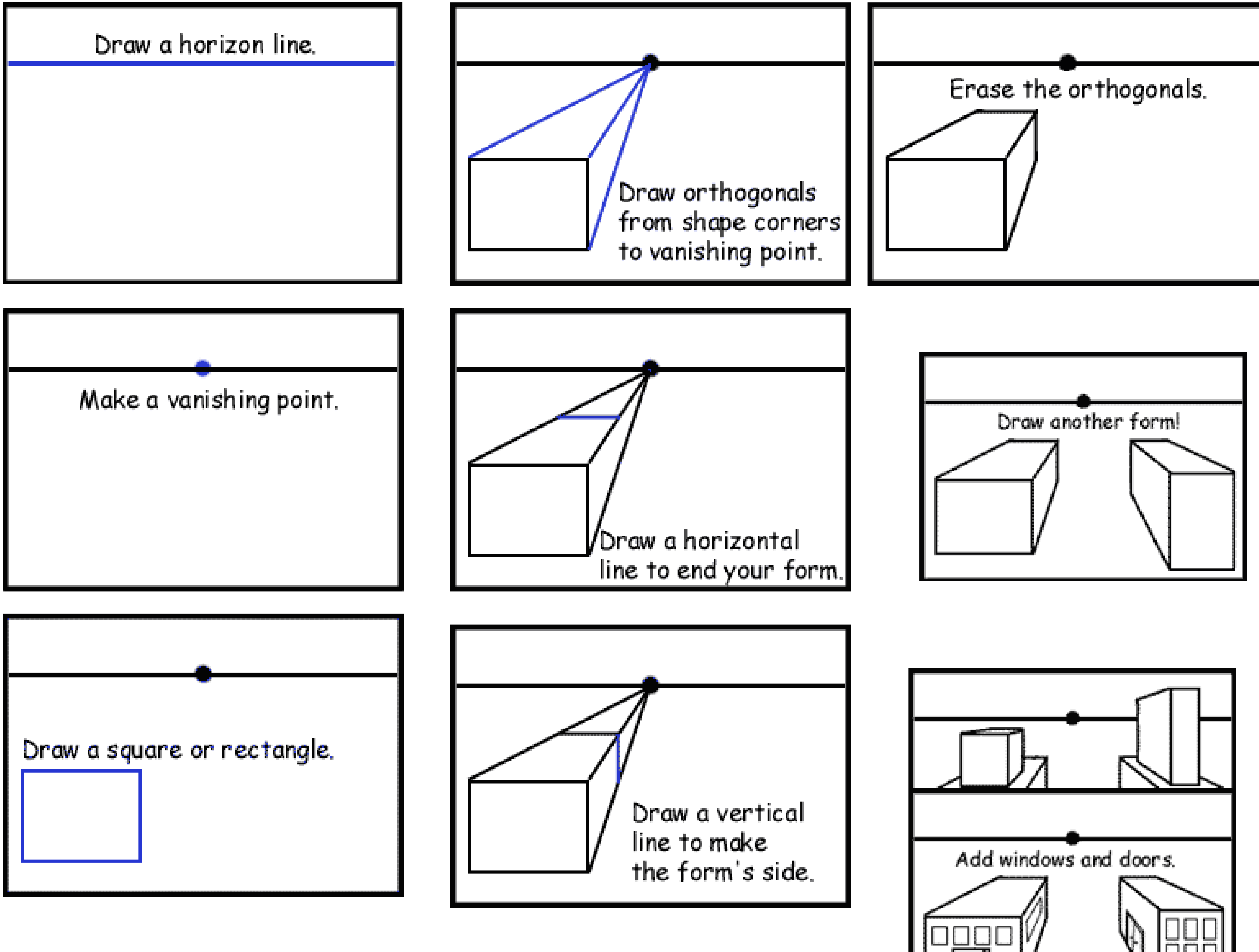
Vanishing points

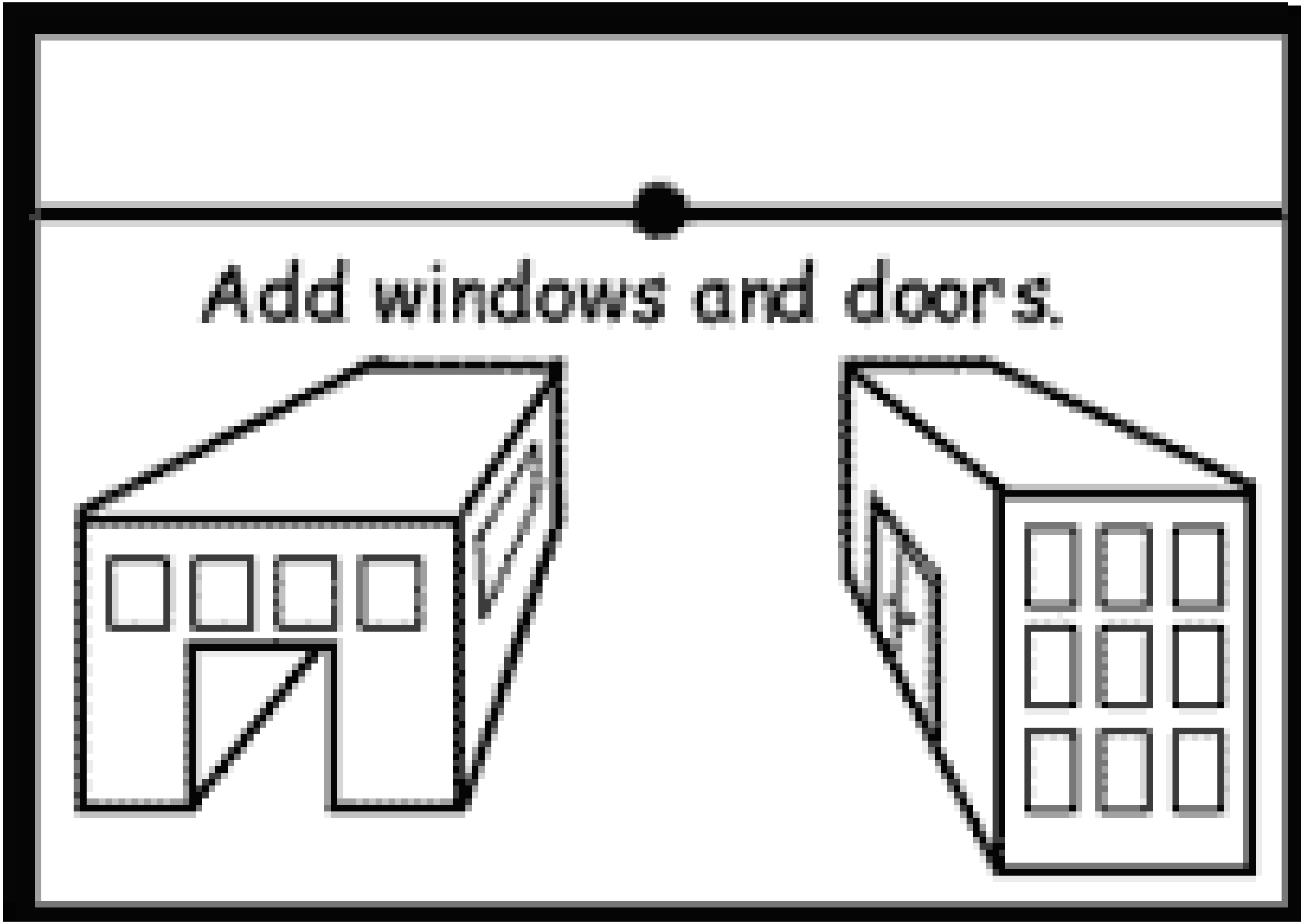
- Each set of parallel lines meets at a different point
  - The *vanishing point* for this direction
- Sets of parallel lines on the same plane lead to *collinear* vanishing points.
  - The line is called the *horizon* for that plane

Properties of Projection

- Points project to points
- Lines project to lines
- Planes project to the whole image or a half image
- Angles are not preserved
- Degenerate cases
  - Line through focal point projects to a point.
  - Plane through focal point projects to line
  - Plane perpendicular to image plane projects to part of the image (with horizon).

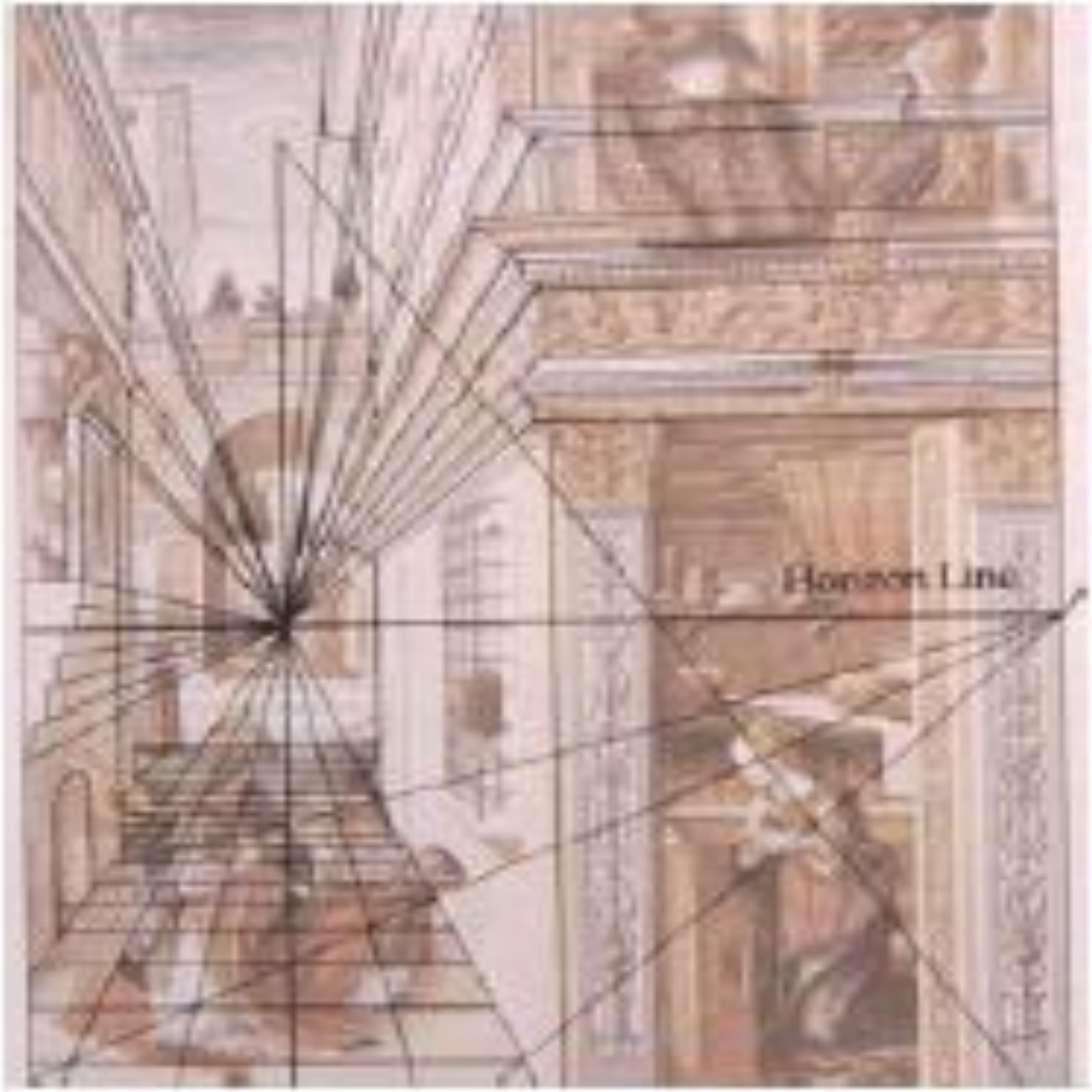
Take out paper and pencil


Draw a horizon line.
Make a vanishing point.
Draw a square or rectangle.
Draw orthogonals from shape corners to vanishing point.
Draw a horizontal line to end your form.
Draw a vertical line to make the form's side.
Erase the orthogonals.
Draw another form!
Add windows and doors.



http://www.sanford-artedventures.com/create/tech_1pt_perspective.html

Carlo Crivelli (1486) *The Annunciation, with St. Emidius*



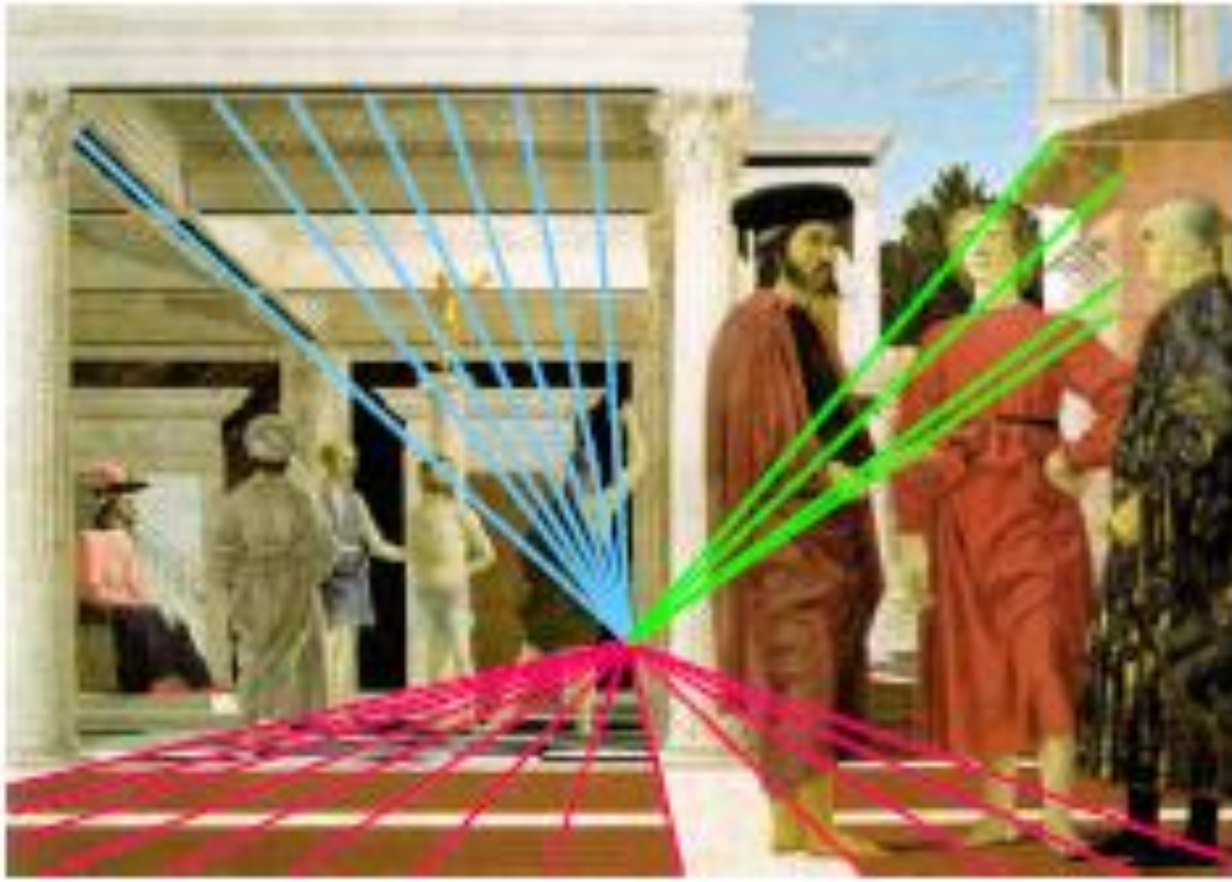
Perspective analysis of Crivelli's *Annunciation*

Masaccio's "Trinity" (c. 1425-8)

- The oldest existing example of linear perspective in Western art
- Use of "snapped" rope lines in plaster
- Vanishing point **below** orthogonals implies looking **up** at vaulted ceiling

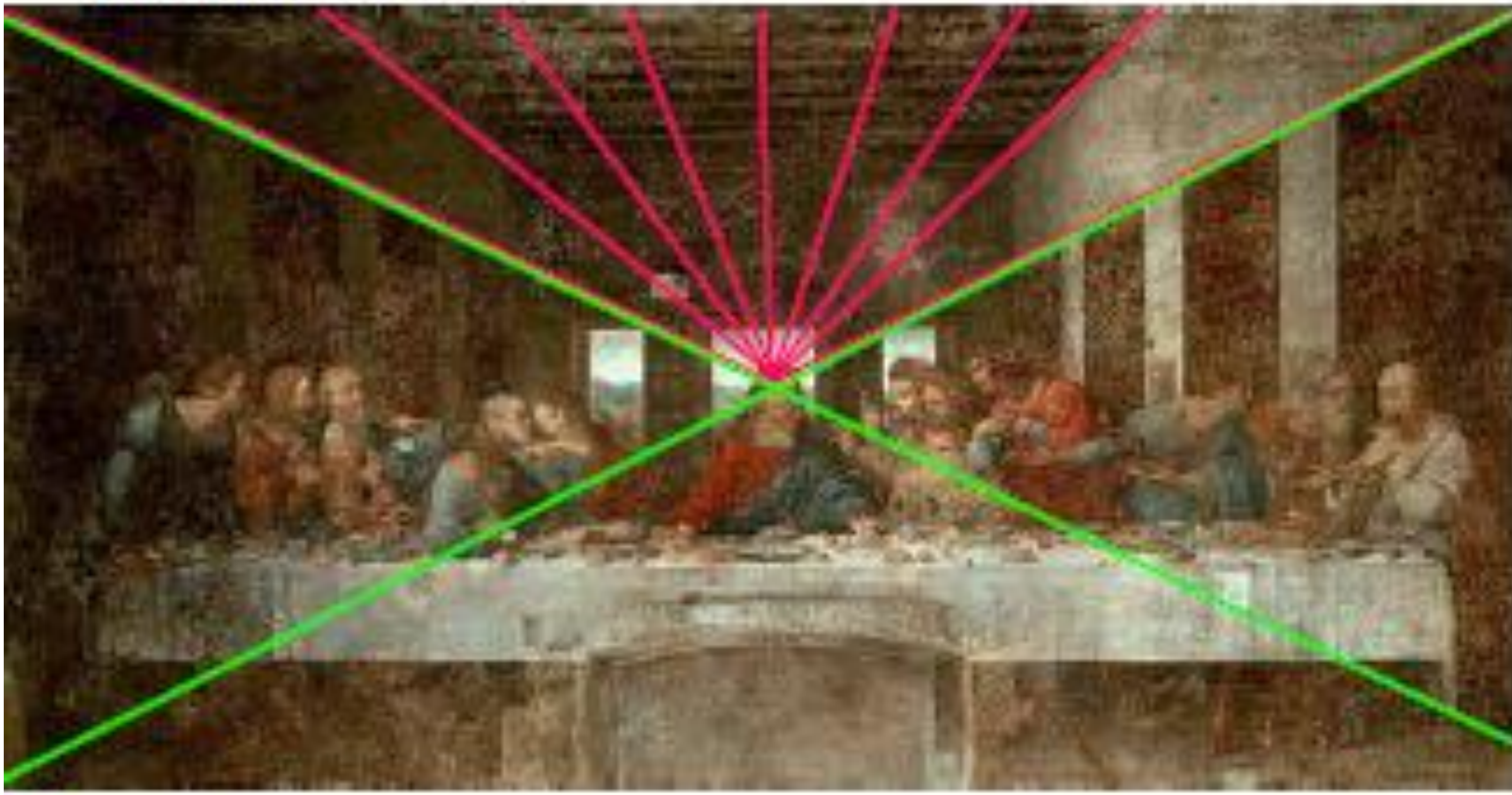
Piero della Francesca, "Flagellation of Christ" (c. 1455)

- Carefully planned
- Strong sense of space
- Low eye level



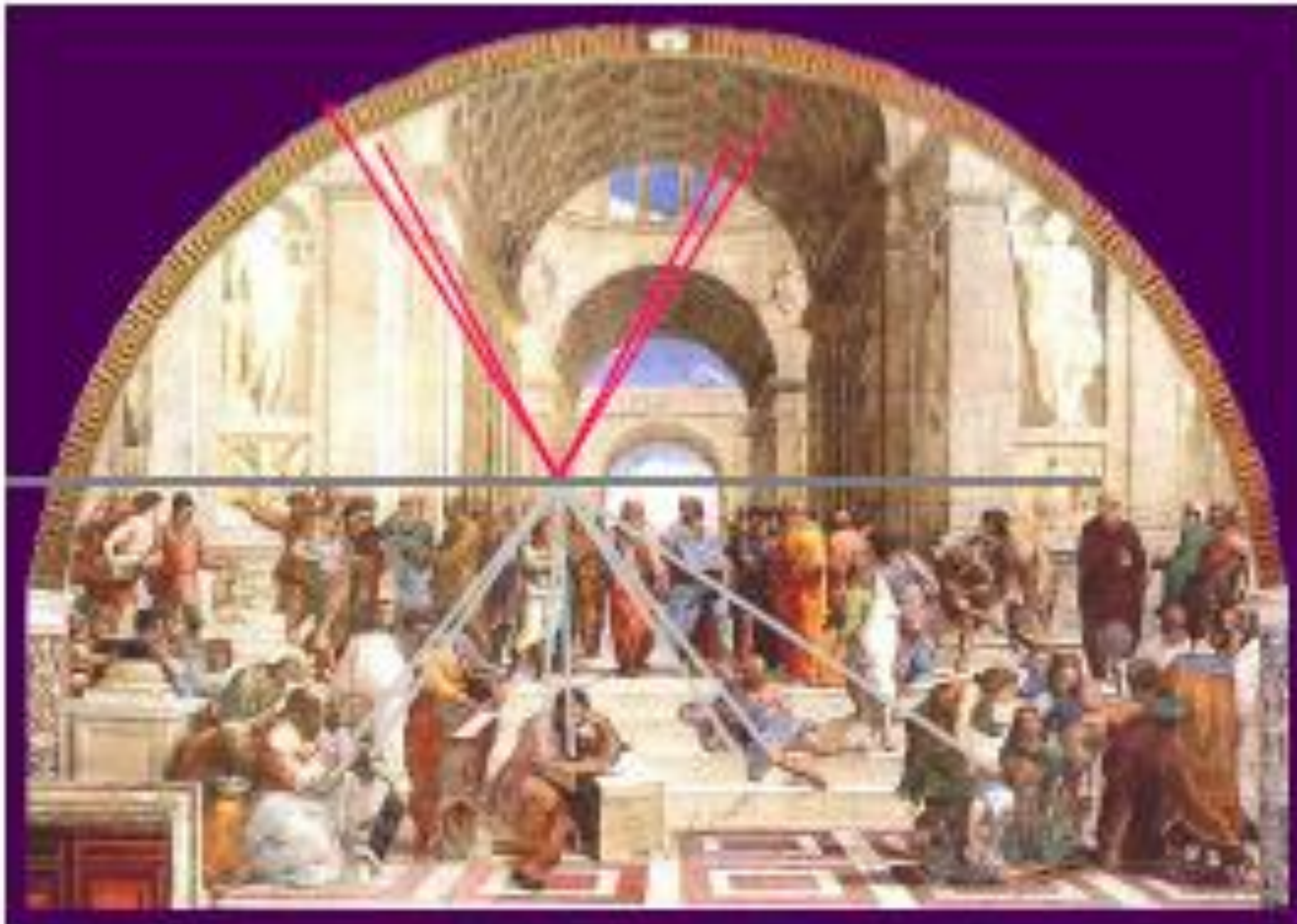
Leonardo da Vinci, "Last Supper" (c. 1497)

- Use of perspective to direct viewer's eye
- Strong perspective lines to corners of image



Raphael, "School of Athens" (1510-11)

Single-point perspective
Central
Strong, coherent space



Painters have used Heuristics to aid in Robust Perception of Perspective

Example: Leonardo's Moderate Distance Rule

To minimize noticeable distortion, use shallow perspective:

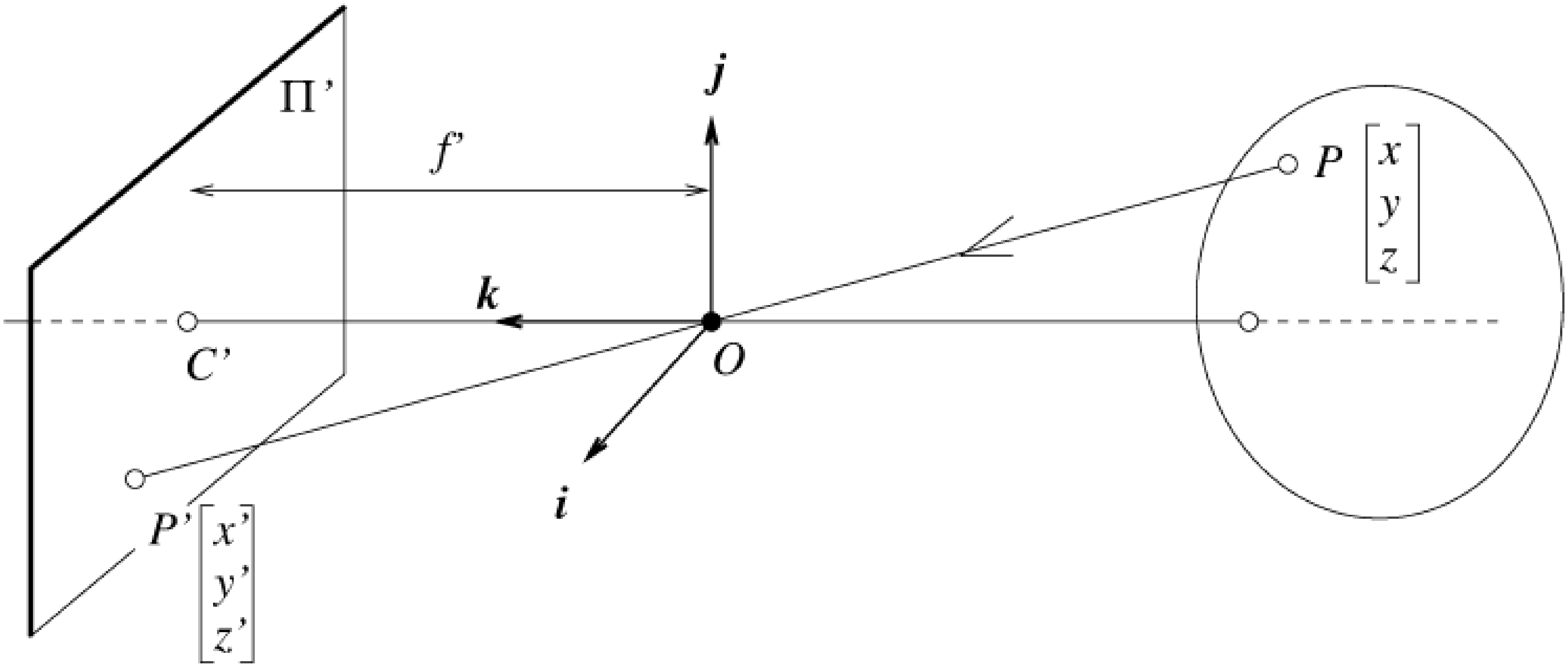
> *"Make your view at least 20 times as far off as the greatest width or height of the objects represented, and this will satisfy any spectator placed anywhere opposite to the picture."*
>
> -- Leonardo

Example: Extreme Viewpoints

Mantegna, *Lamentation over the dead Christ*, 1480

Raphael, *School of Athens, 1511*

The equation of projection



(Forsyth & Ponce)

The equation of projection

- Cartesian coordinates:
  - We have, by similar triangles, that

$$x' = f' \frac{x}{z}$$

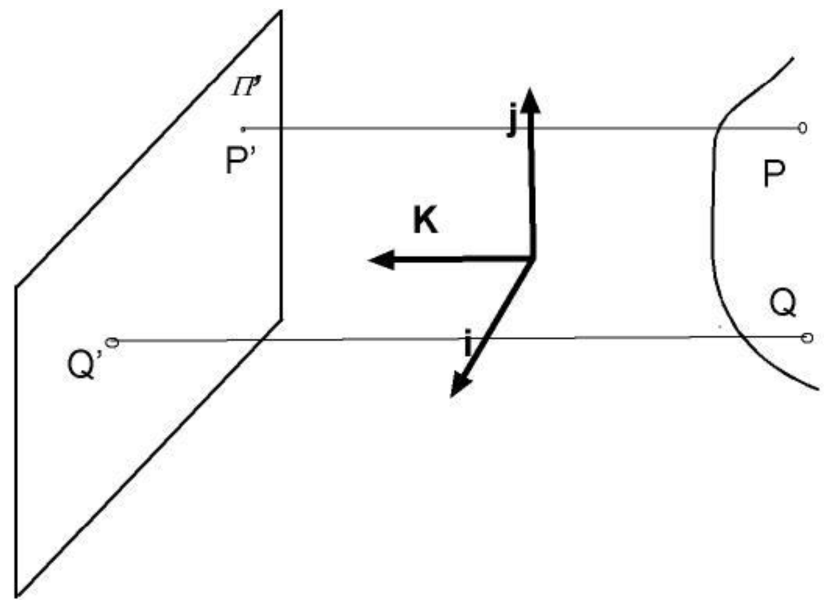
$$y' = f' \frac{y}{z}$$

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

  - Ignore the third coordinate, and get

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

Orthographic projection

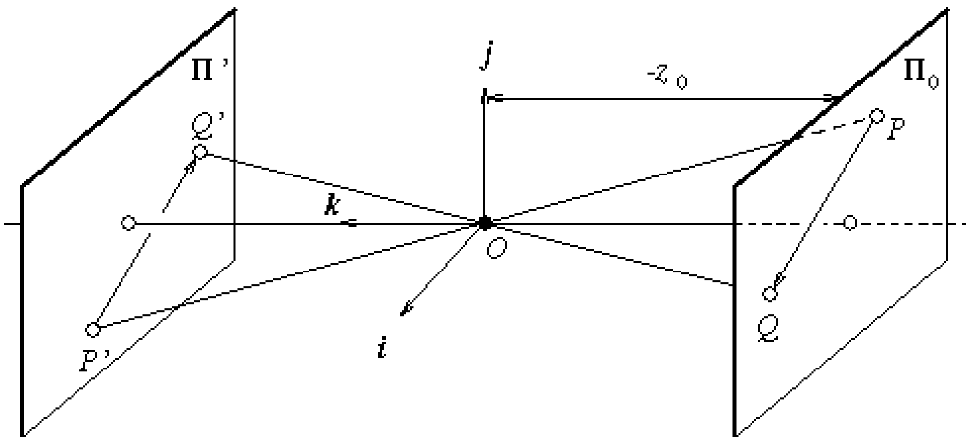


$$x' = x$$

$$y' = y$$

Weak perspective (scaled orthographic projection)

- Issue
  - perspective effects, but not over the scale of individual objects
  - collect points into a group at about the same depth, then divide each point by the depth of its group



(Forsyth & Ponce)

The Equation of Weak Perspective

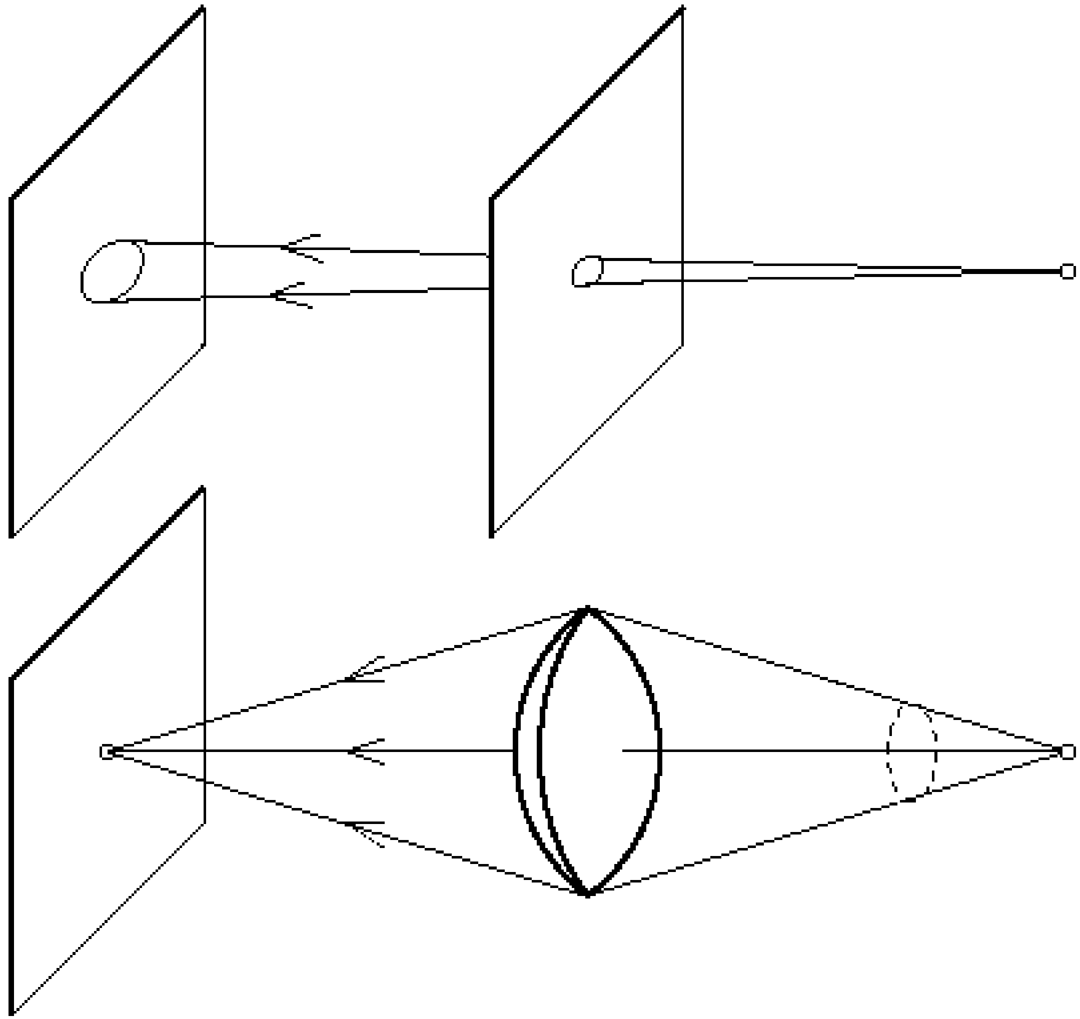
$$(x, y, z) \rightarrow s(x, y)$$

- s is constant for all points.
- Parallel lines no longer converge, they remain parallel.

Pros and Cons of These Models

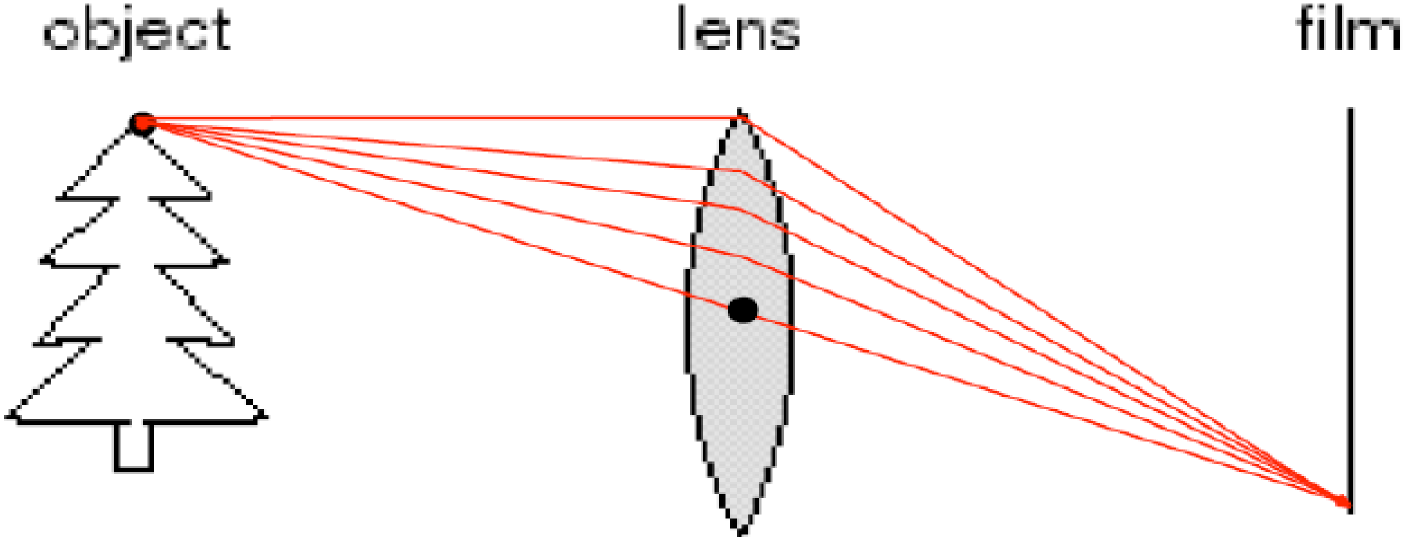
- Weak perspective much simpler math.
  - Accurate when object is small and distant.
  - Most useful for recognition.
- Pinhole perspective much more accurate for scenes.
  - Used in structure from motion.
- When accuracy really matters, must model real cameras.

Cameras with Lenses



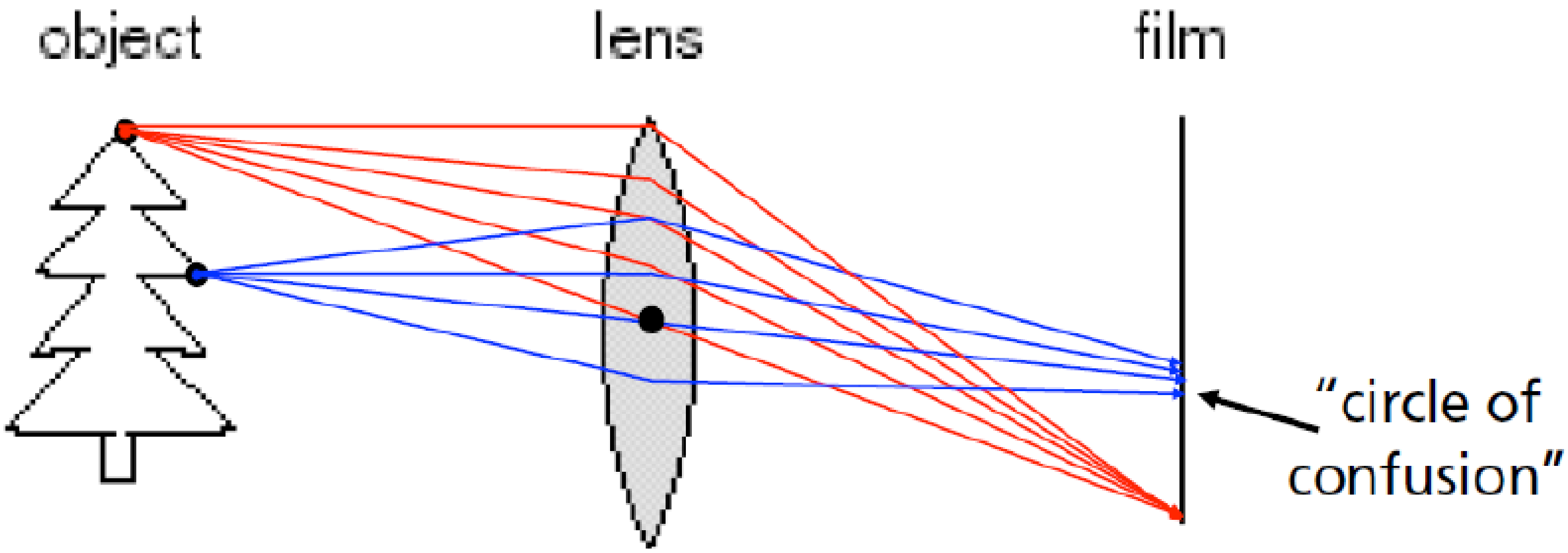
(Forsyth & Ponce)

Adding a lens



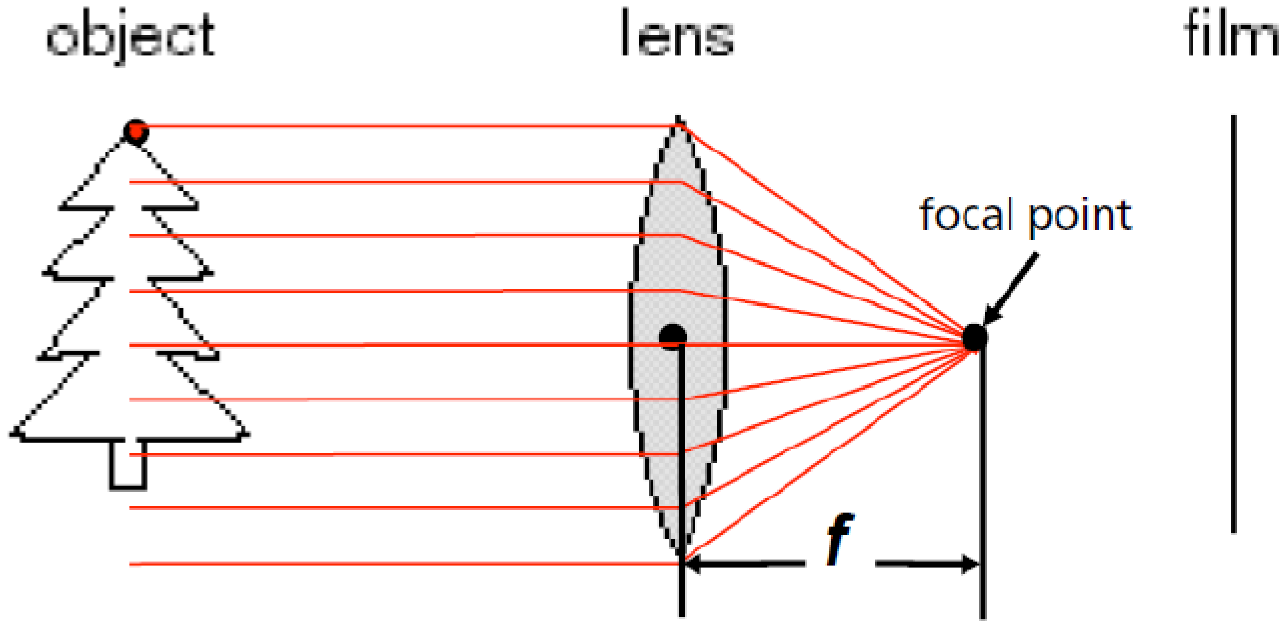
- A lens focuses light onto the film
- Lets enough ligh through
- Rays passing through the center are not deviated

In Focus

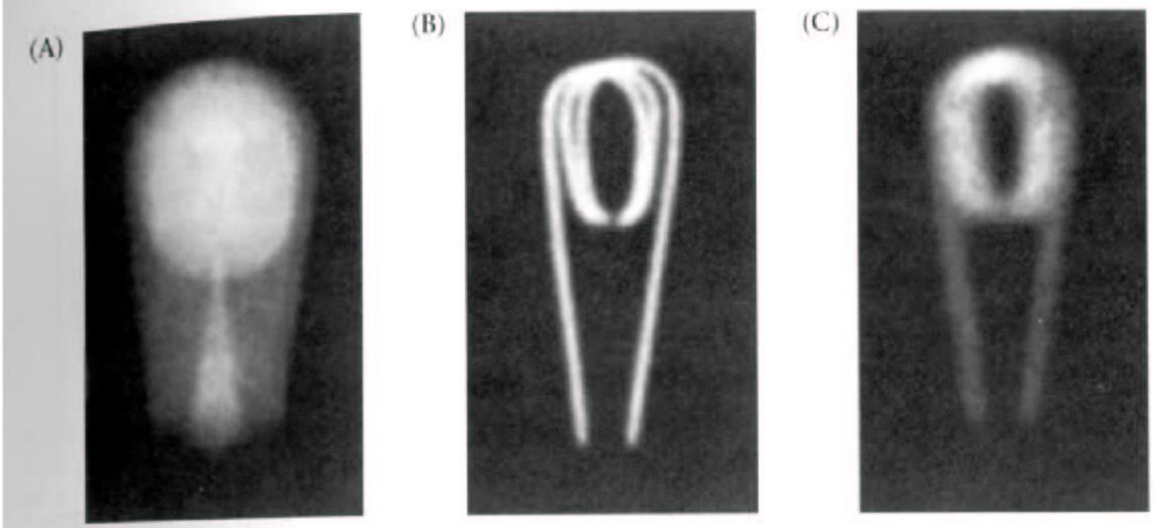


- There is a specific distance at which objects are "in focus"
- other points project to a "circle of confusion" in the image

Focal point



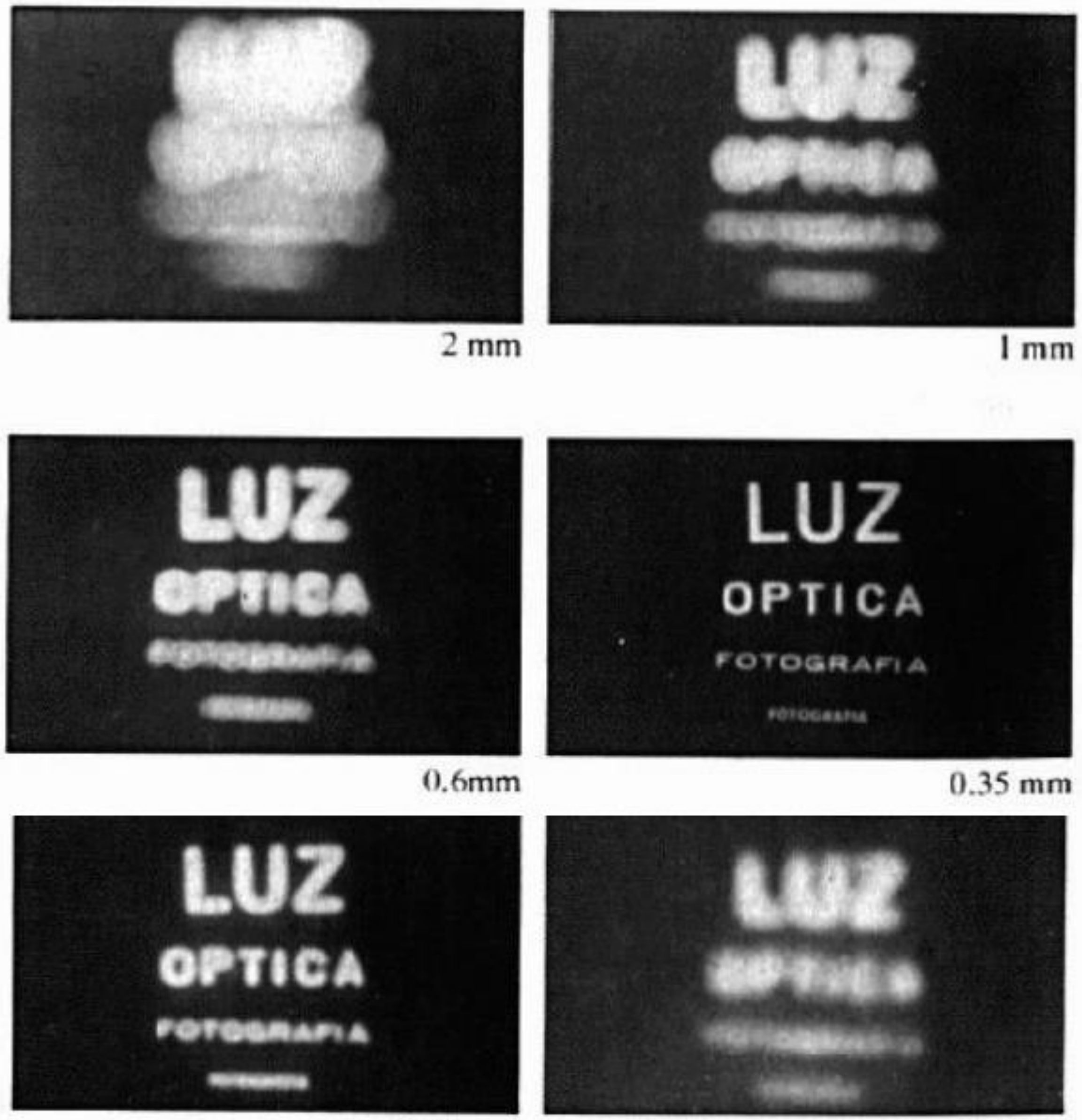
- All parallel rays converge to one point on a plane located at the *focal length f*



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.

Wandell, Foundations of Vision, Sinauer, 1995

Shrinking the aperture

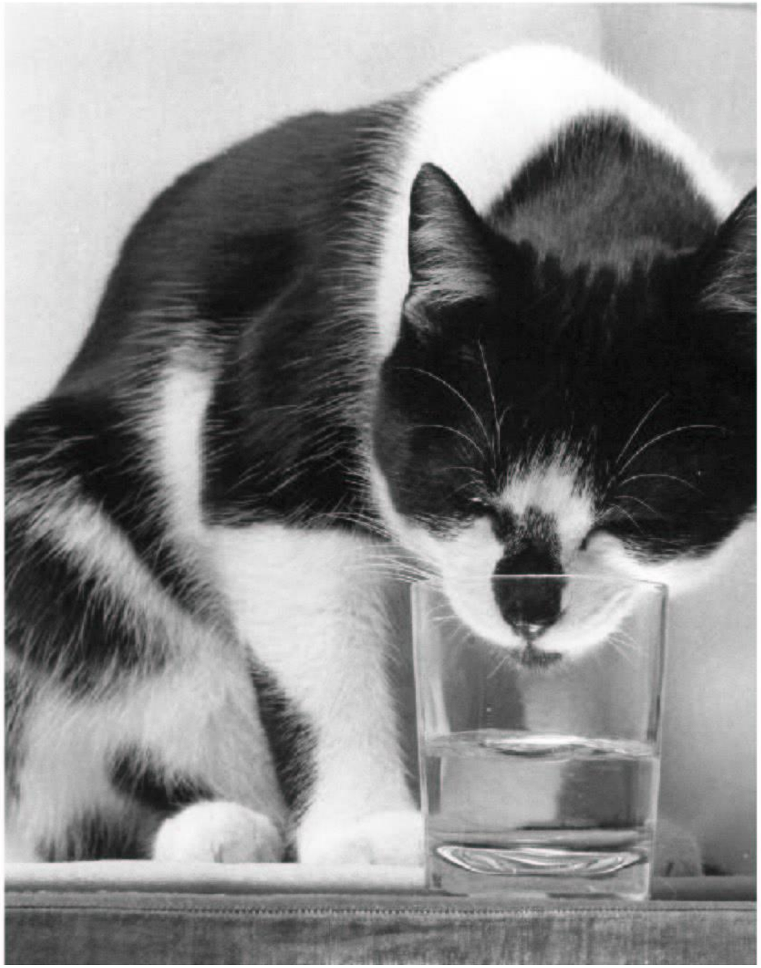


Diffraction effects. Noise due to long exposure

Interaction of light with matter

- Absorption
- Scattering
- Refraction
- Reflection
- Other effects:
  - Diffraction: deviation of straight propagation in the presence of obstacles
  - Fluorescence:absorbtion of light of a given wavelength by a fluorescent molecule causes reemission at another wavelength

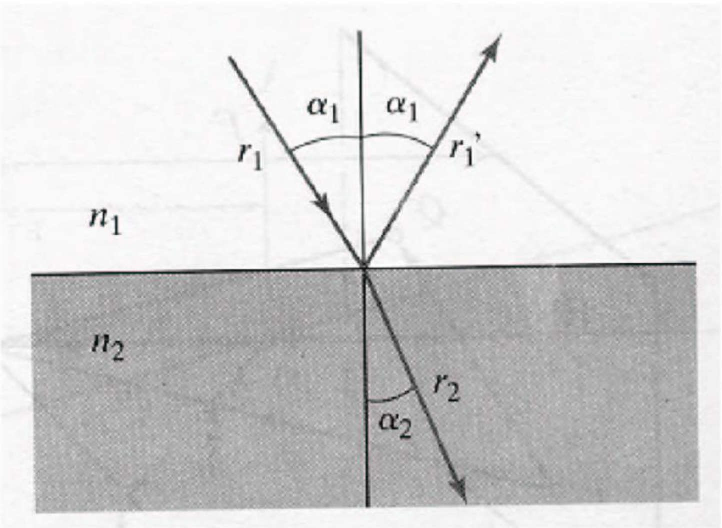
Water glass refraction



http://data.pg2k.hd.org/_exhibits/natural-science/cat-black-and-white-domestic-short-hair-DSH-with-nose-in-glass-of-water-on-bedside-table-tweaked-mono-1-AJHD.jpg

Refraction

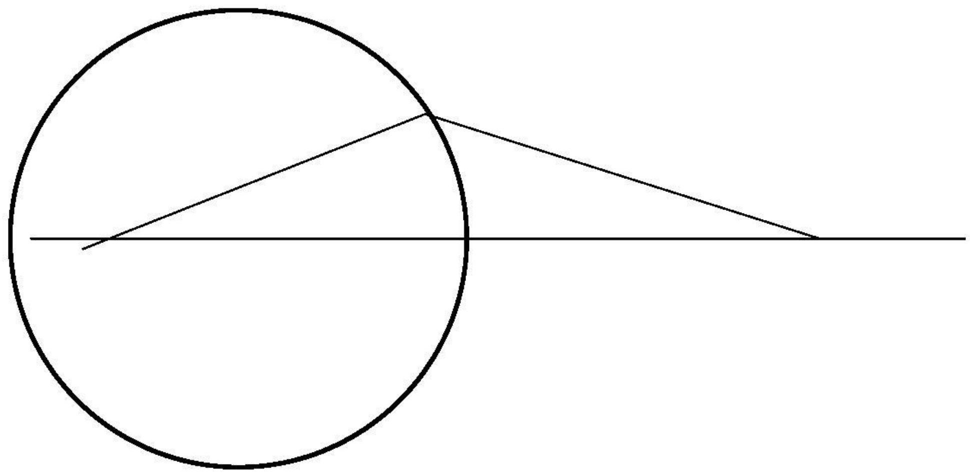
Snell's law



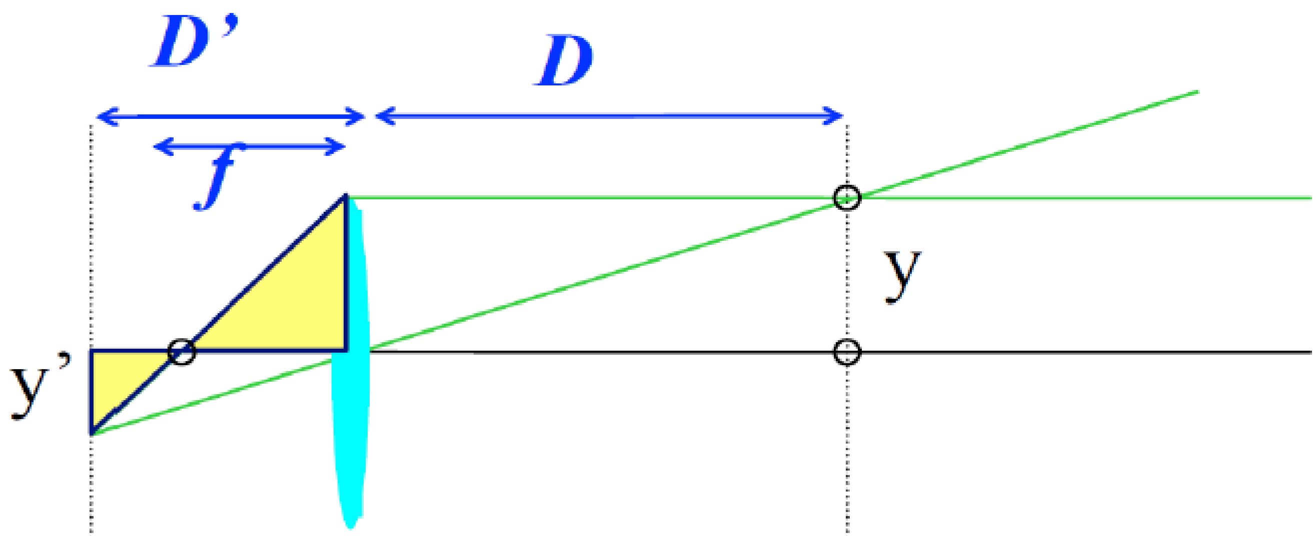
$$n_1 \sin(\alpha_1) = n_2 \sin(\alpha_2)$$

n1, *n2*: indexes of refraction

Spherical lens



Thin lens formula



Green similiar triangles:

$$\frac{y'}{y} = \frac{D'}{D}$$

Yellow similiar triangles:

$$\frac{y'}{y} = \frac{D' - f}{f}$$

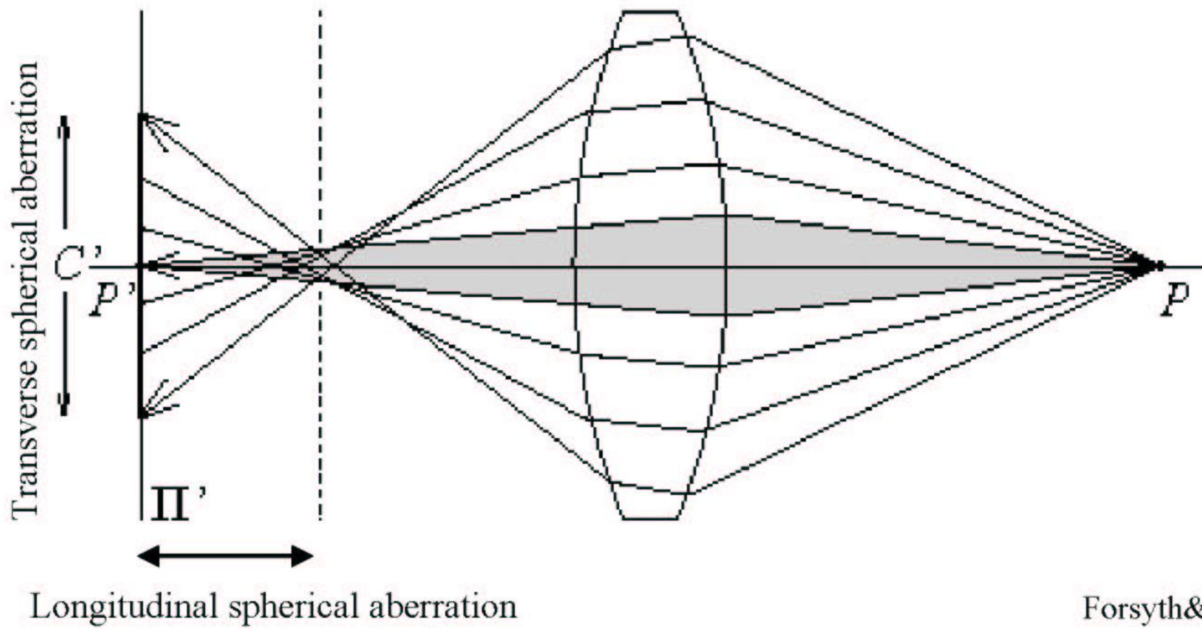
$$\frac{D'}{D} = \frac{D' - f}{f} \Rightarrow \frac{D'}{D} = \frac{D'}{f} - 1 \Rightarrow \boxed{\frac{1}{D} + \frac{1}{D'} = \frac{1}{f}}$$ **Thin lens formula**

Any point satisfying the thin lens equation is in focus.

Assumptions for thin lens equation

- Lens surfaces are spherical
- Incoming light rays make a small angle with the optical axis
- The lens thickness is small compared to the radii of curvature
- The refractive index is the same for the media on both sides of the lens

Spherical aberration (from 3rd order optics



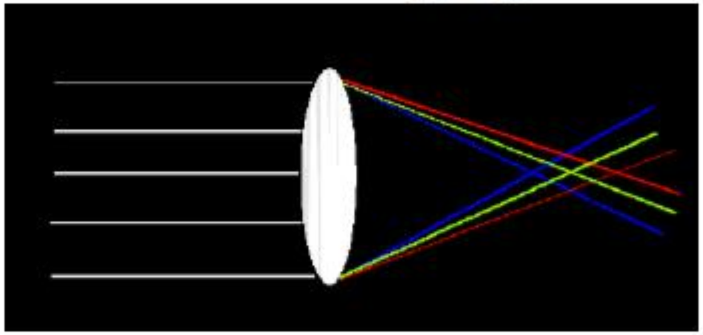
Forsyth&Ponce

Other aberrations

- Astigmatism: unevenness of the cornea
- Distortion : different areas of lens have different focal length
- Coma : point not on optical axis is depicted as asymmetrical comet-shaped blob
- Chromatic aberration

Lens Flaws: Chromatic Aberration

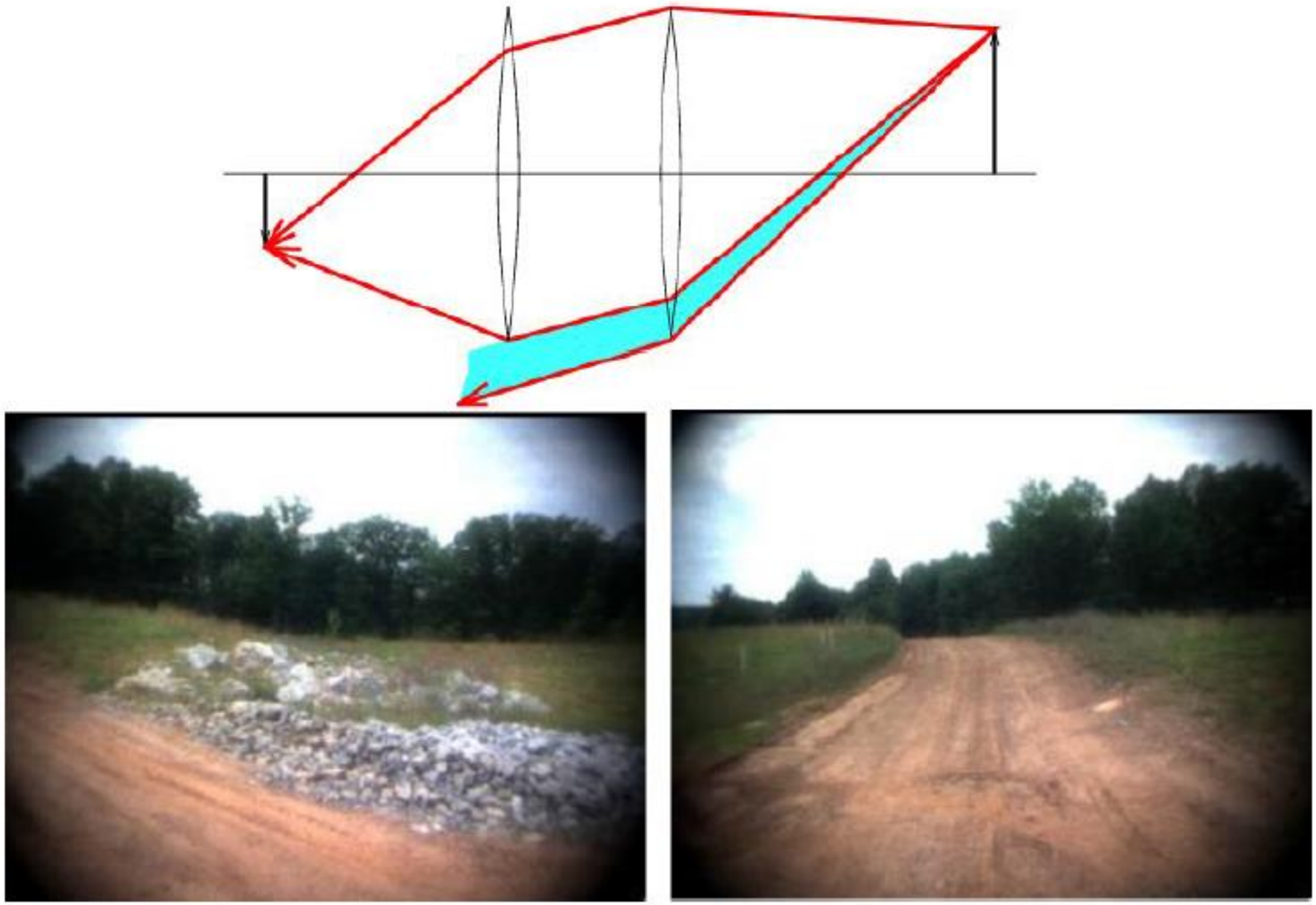
Lens has different refractive indices for different wavelengths: causes color fringing



High quality lens (top)
low quality lens (bottom)
blur + green edges

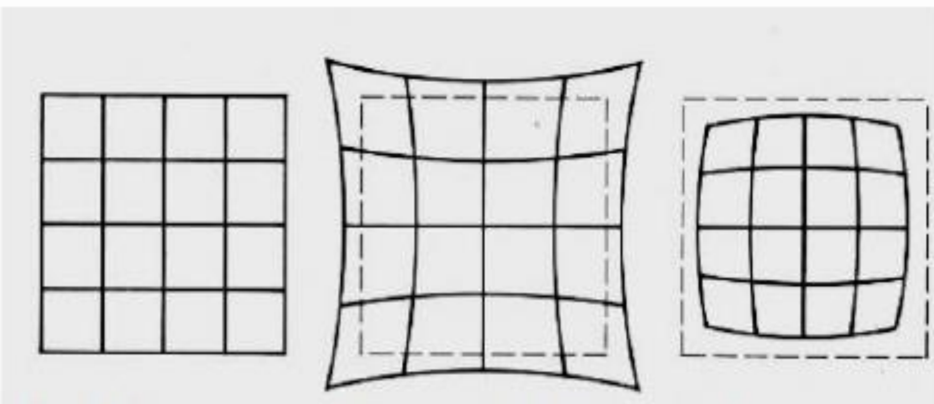
Purple fringing

Lens flaws: Vignetting



Lens distortion

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

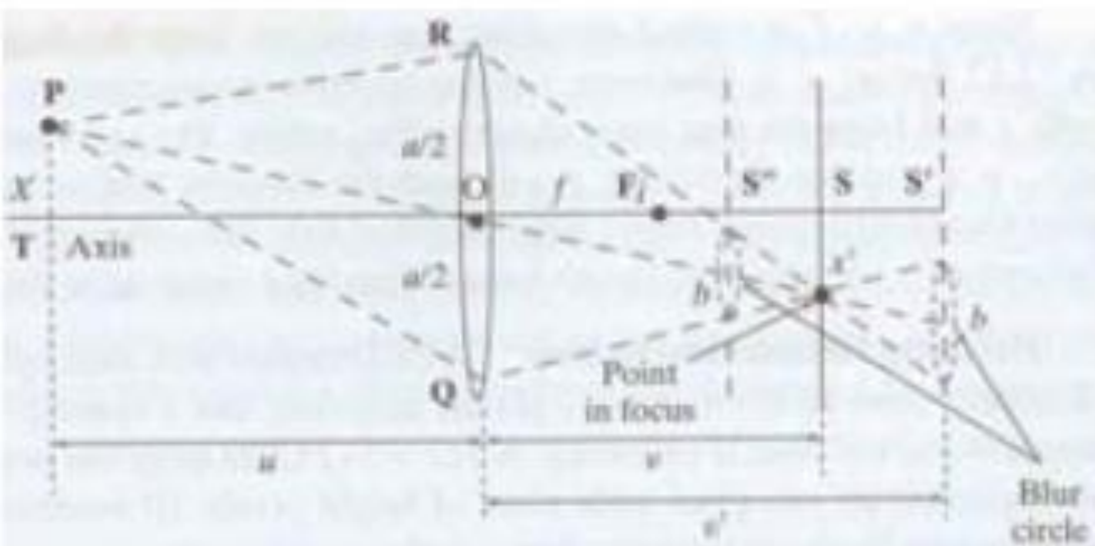


No distortion Pin cushion Barrel

Focus and depth of field

Focus and depth of field

- Depth of field: distance between image planes where blur is tolerable



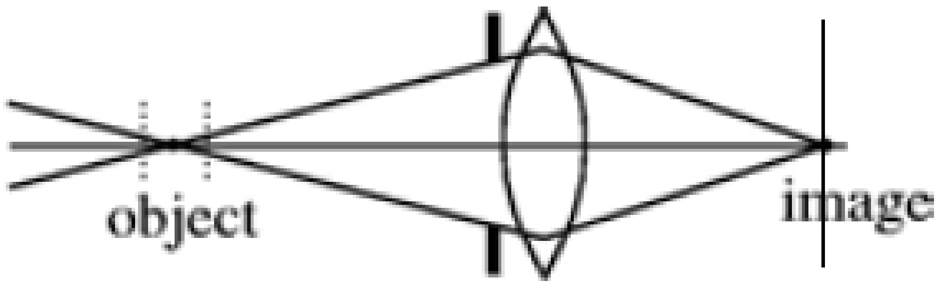
"circles of confusion"

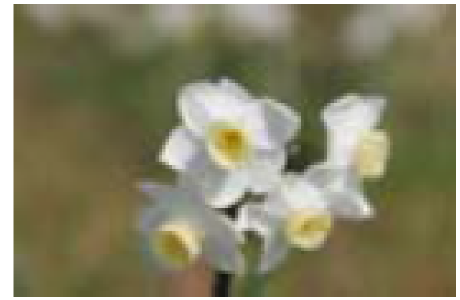
Thin lens: scene points at distinct depths come in focus at different image planes.

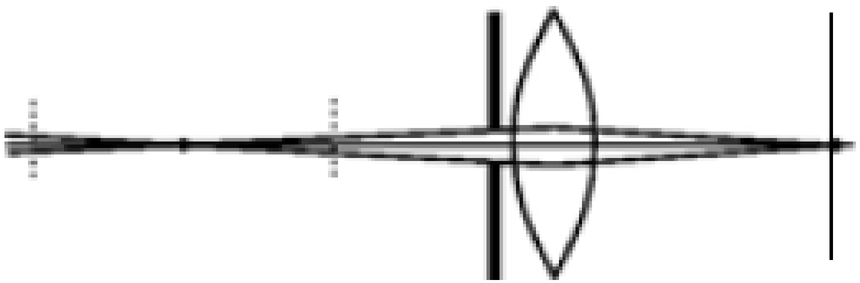
(Real camera lens systems have greater depth of field.)

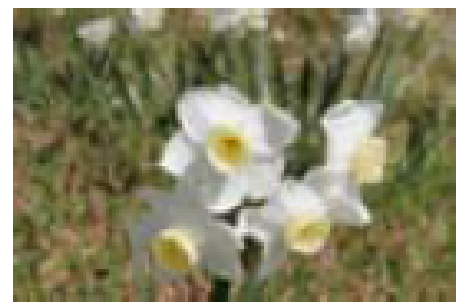
Focus and depth of field

- How does the aperture affect the depth of field?



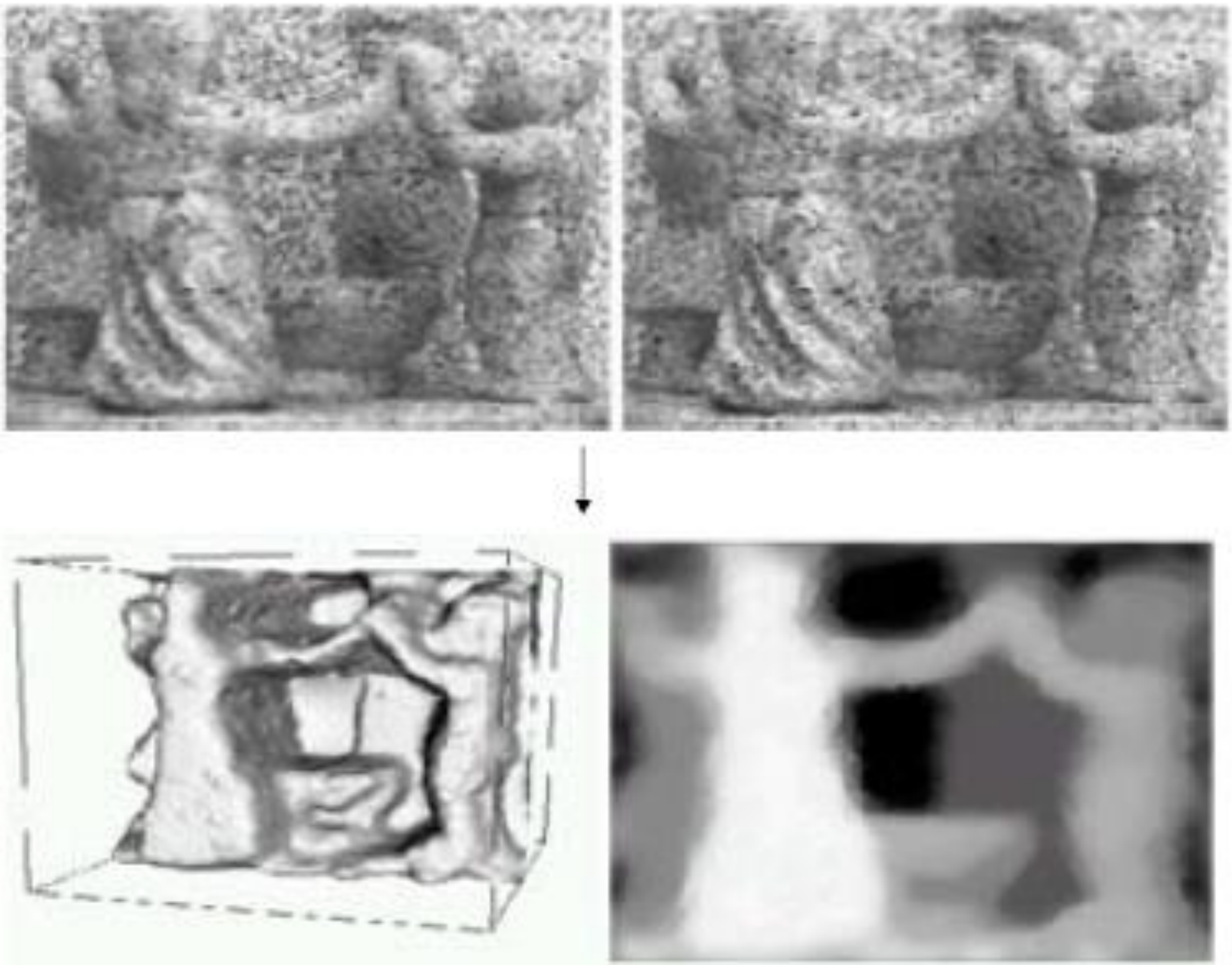






- A smaller aperture increases the range in which the object is approximately in focus

Depth from focus



Images from same point of view, different camera parameters

3d shape / depth estimates

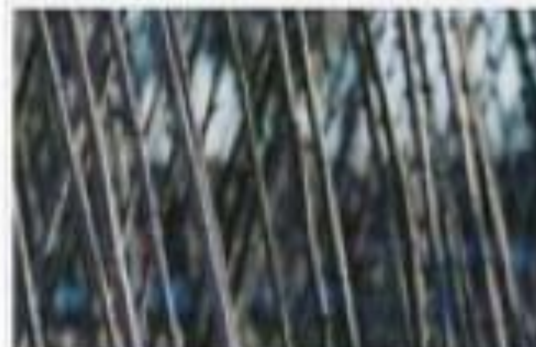
Field of view

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

28 mm lens, 65.5° × 46.4°

50 mm lens, 39.6° × 27.0°

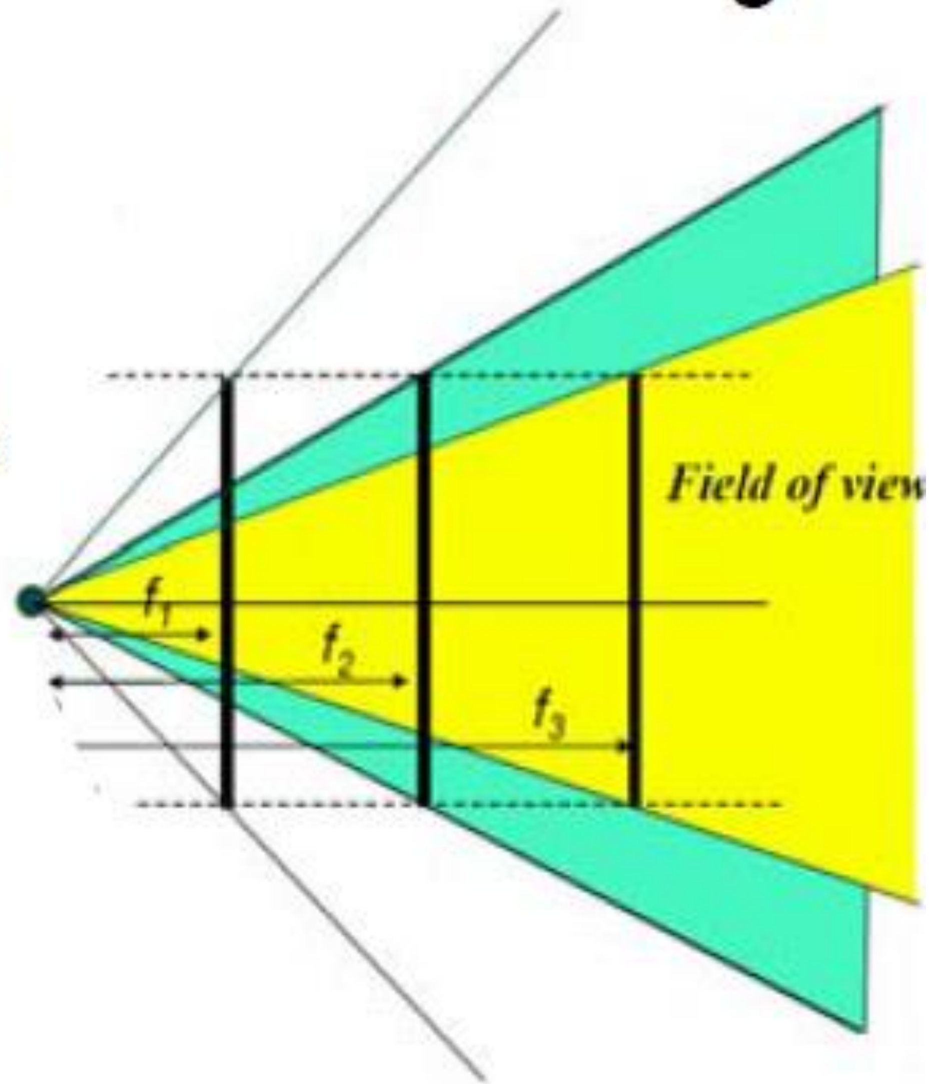
70 mm lens, 28.9° × 19.5°



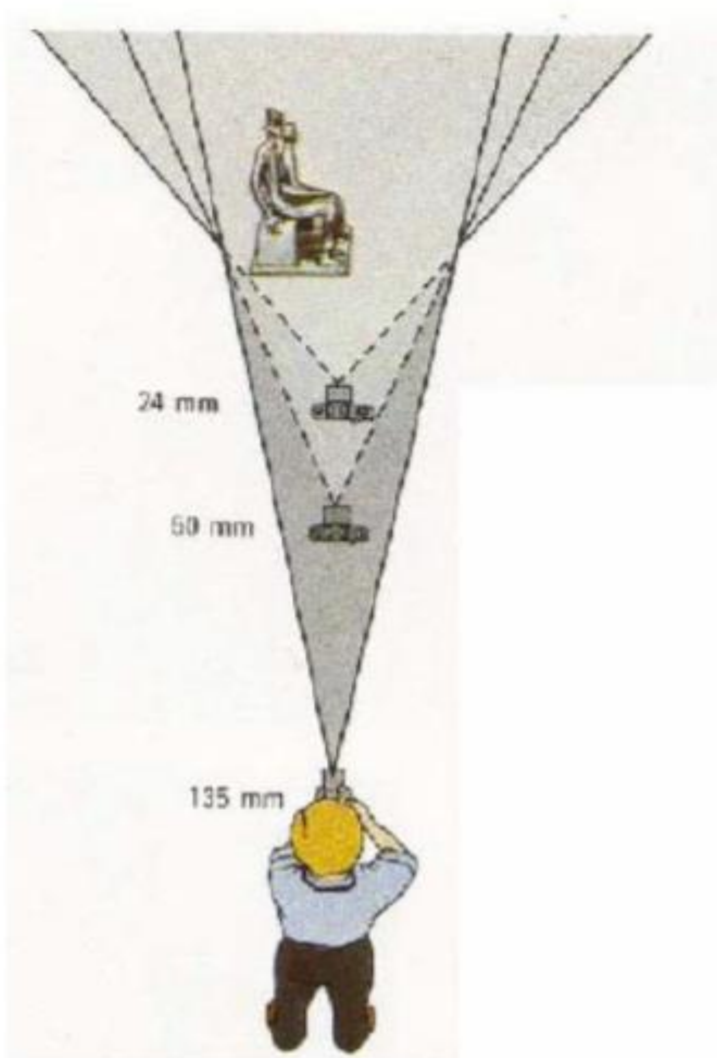
210 mm lens, 9.8° × 6.5°

Field of view depends on focal length

- As ***f*** gets smaller, image becomes more *wide angle*
  - more world points project onto the finite image plane
- As ***f*** gets larger, image becomes more *telescopic*
  - smaller part of the world projects onto the finite image plane



Field of View / Focal Length




Large FOV, small f
Camera close to car


Small FOV, large f
Camera far from the car

Same effect for faces

wide-angle

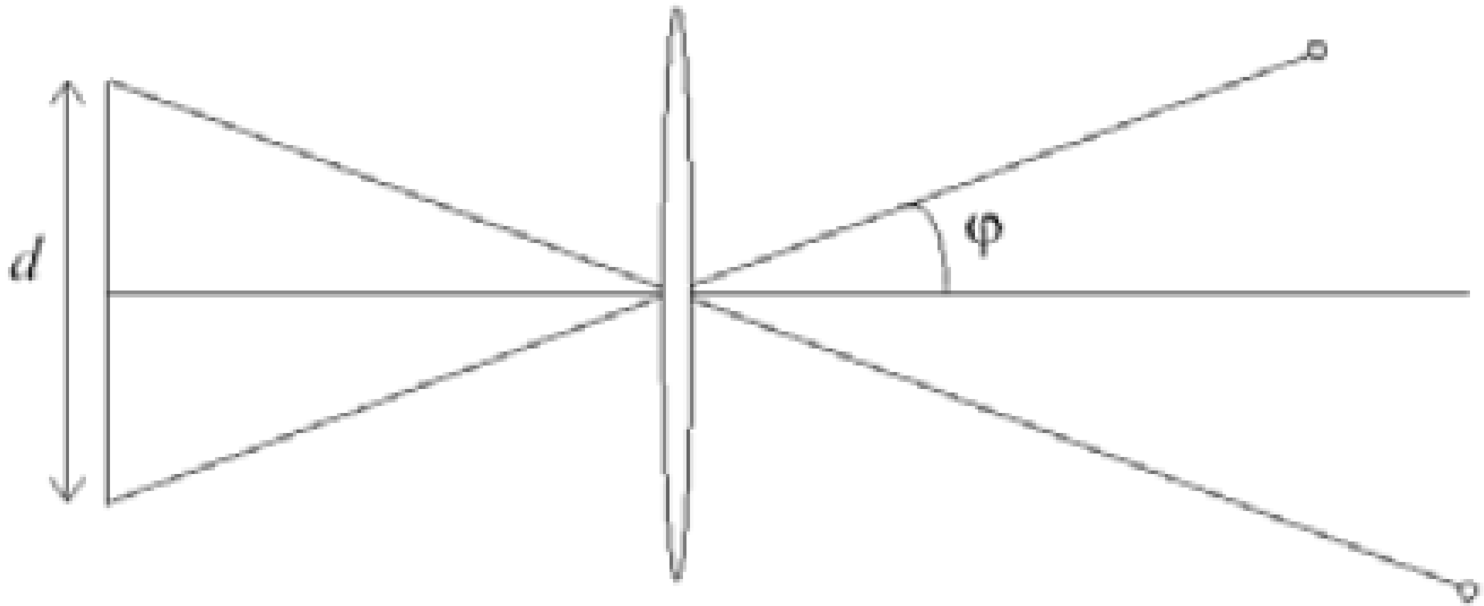
Small f

standard

telephoto

Large f

Field of view depends on focal length

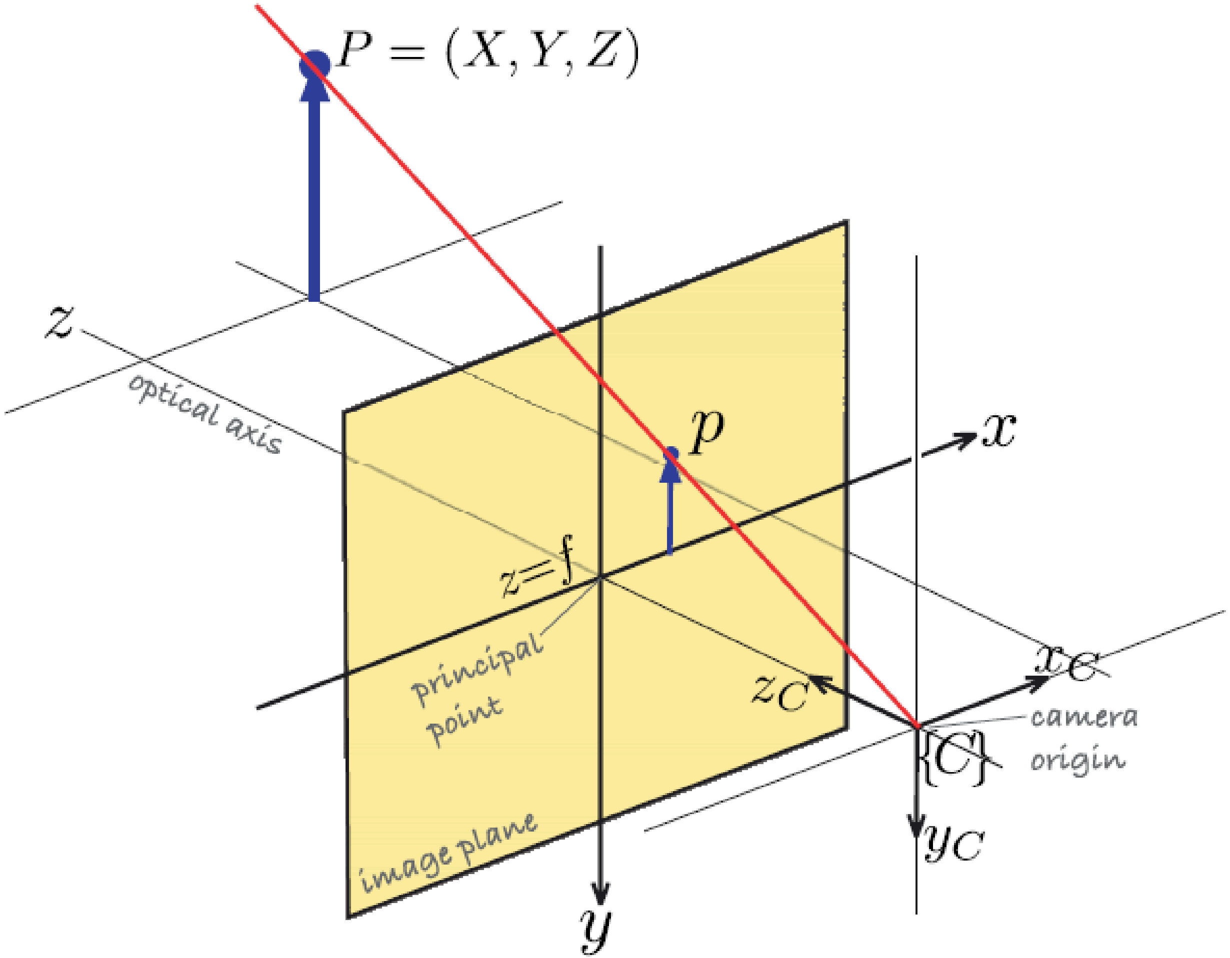


Size of field of view govemed by size of the camera retina:

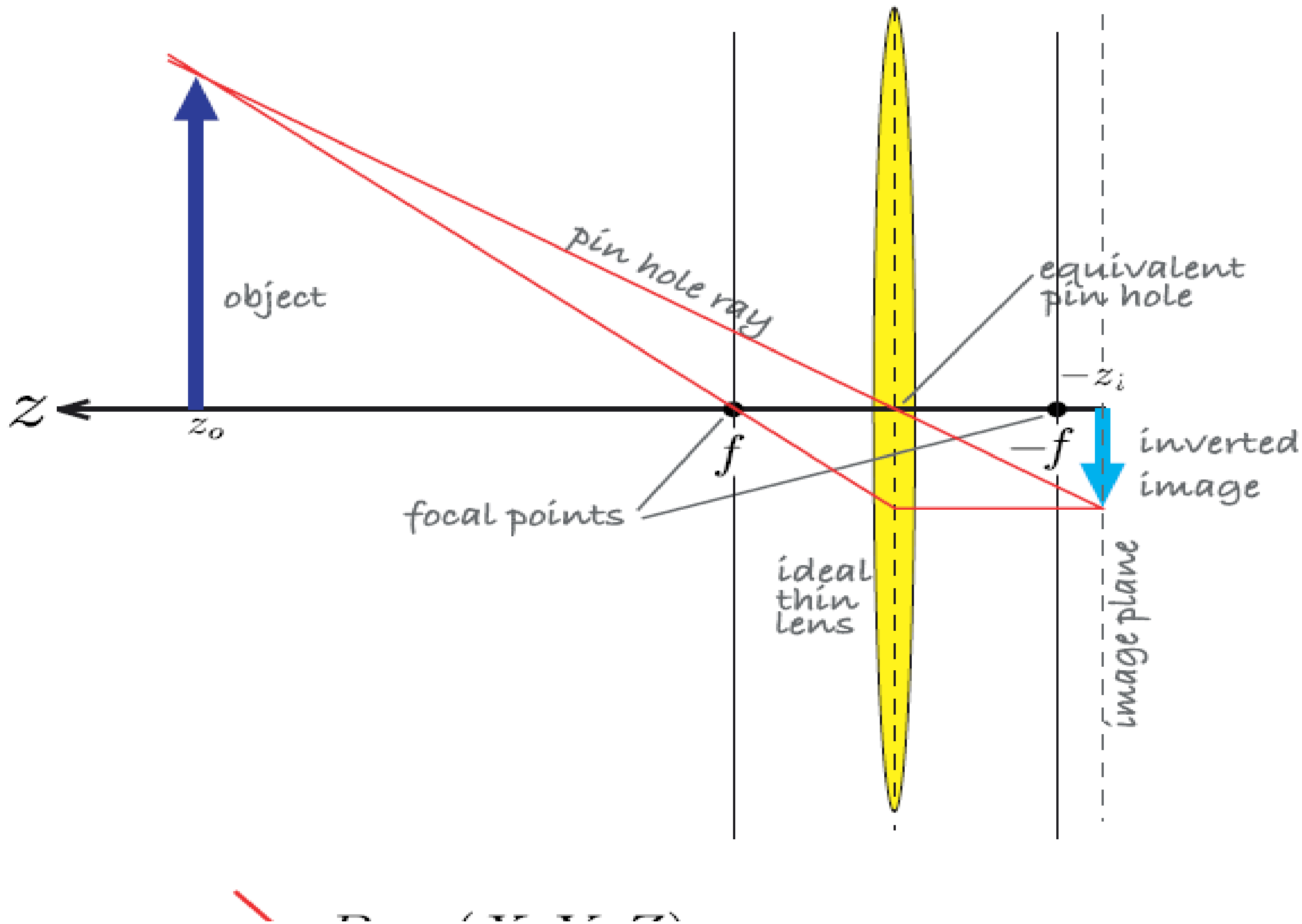
$$\varphi = \tan^{-1}\left(\frac{d}{2f}\right)$$

Smaller FOV = larger Focal Length

Summary



$P = (X, Y, Z)$
z
optical axis
p
x
$z=f$
principal point
x_C
z_C
camera origin
$\{C\}$
y_C
image plane
y

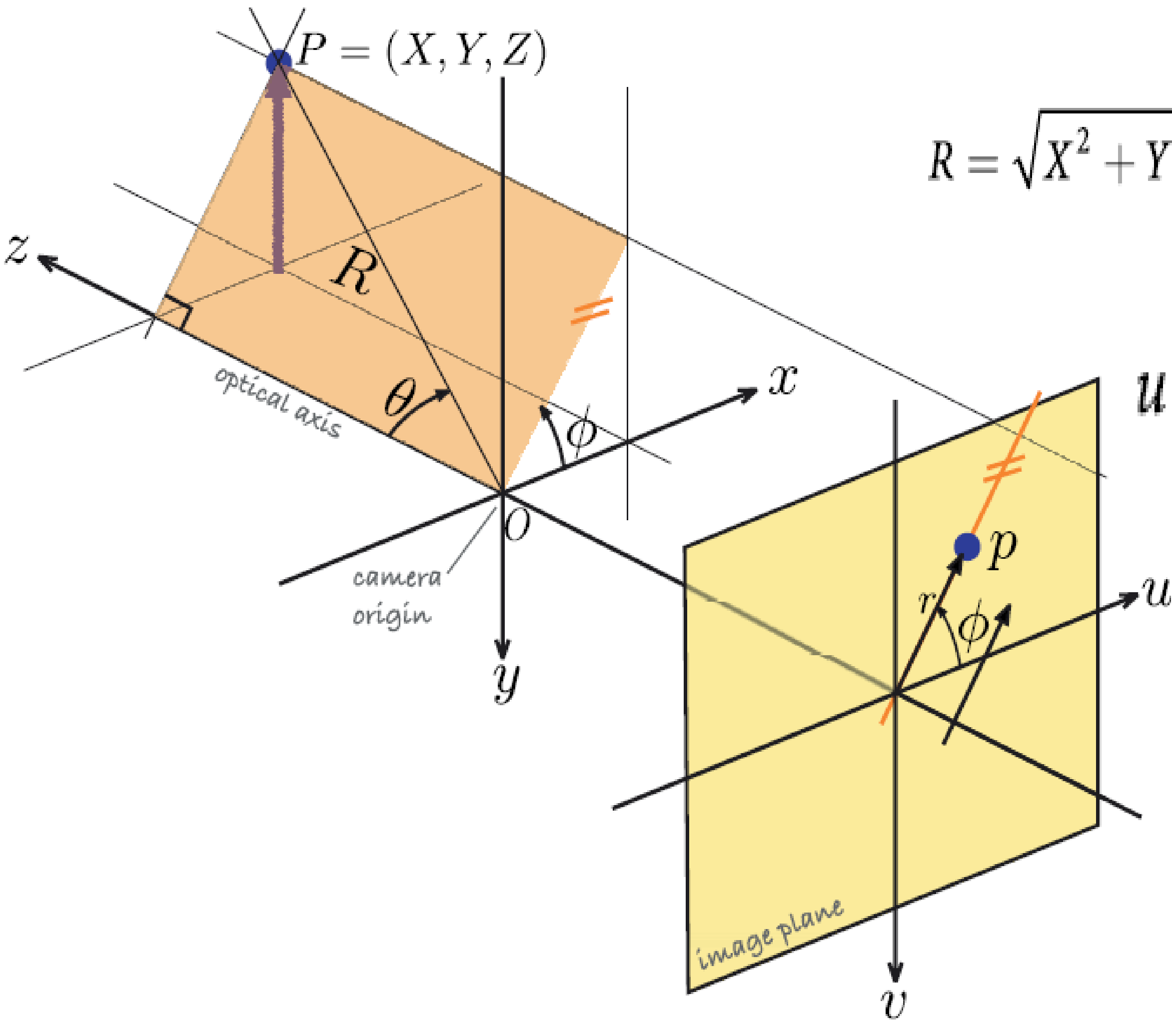

pin hole ray
equivalent pin hole
object
Z
z_o
f
$-z_i$
$-f$
inverted image
focal points
ideal thin lens
image plane

Non perspective models

Fish-eye lens camera

$$R = \sqrt{X^2 + Y^2 + Z^2},\ \theta = \cos^{-1}\frac{R}{Z},\ \phi = \tan^{-1}\frac{Y}{X}$$

$$u = r\cos\phi,\ v = r\sin\phi$$

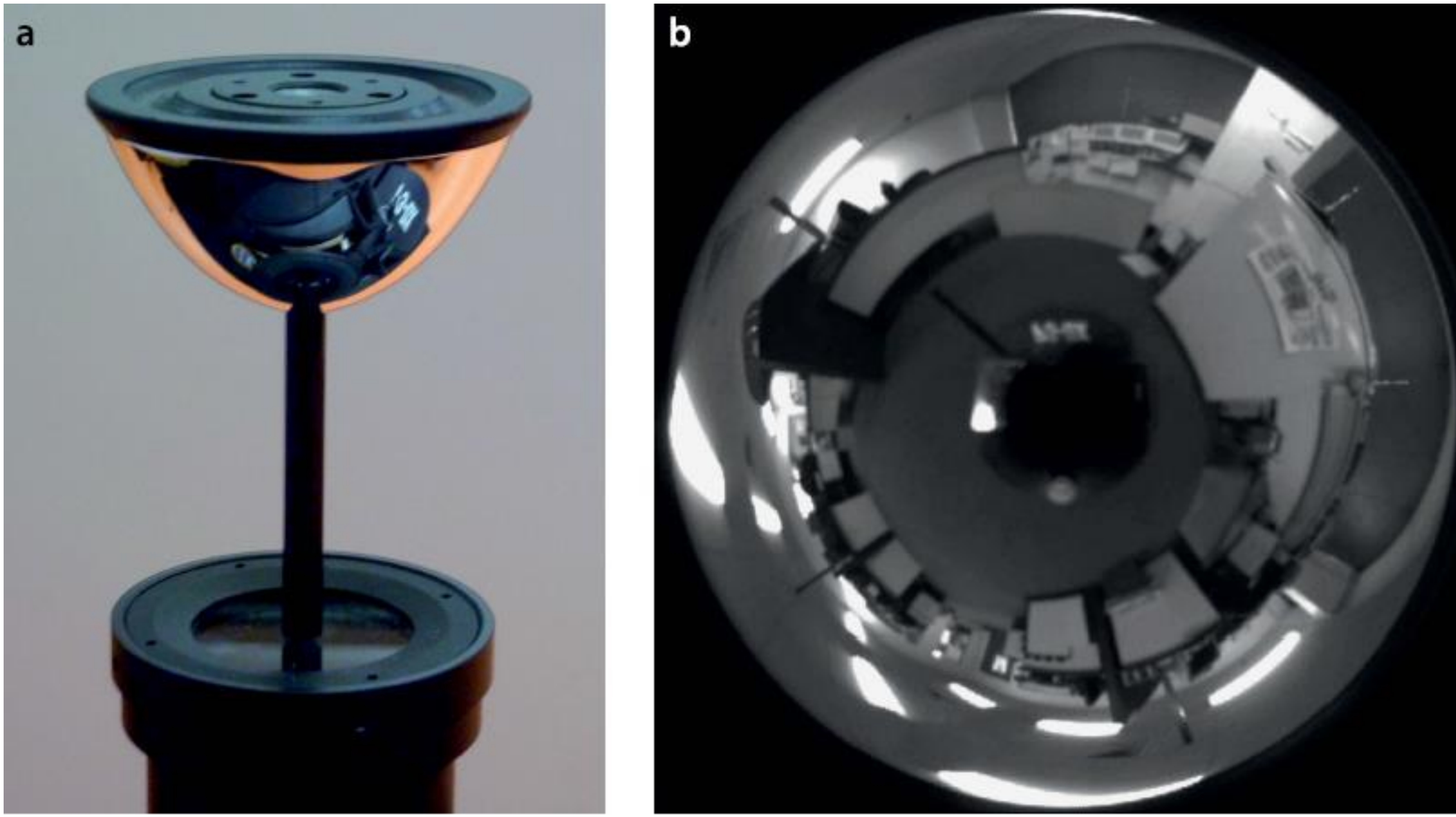


Different models

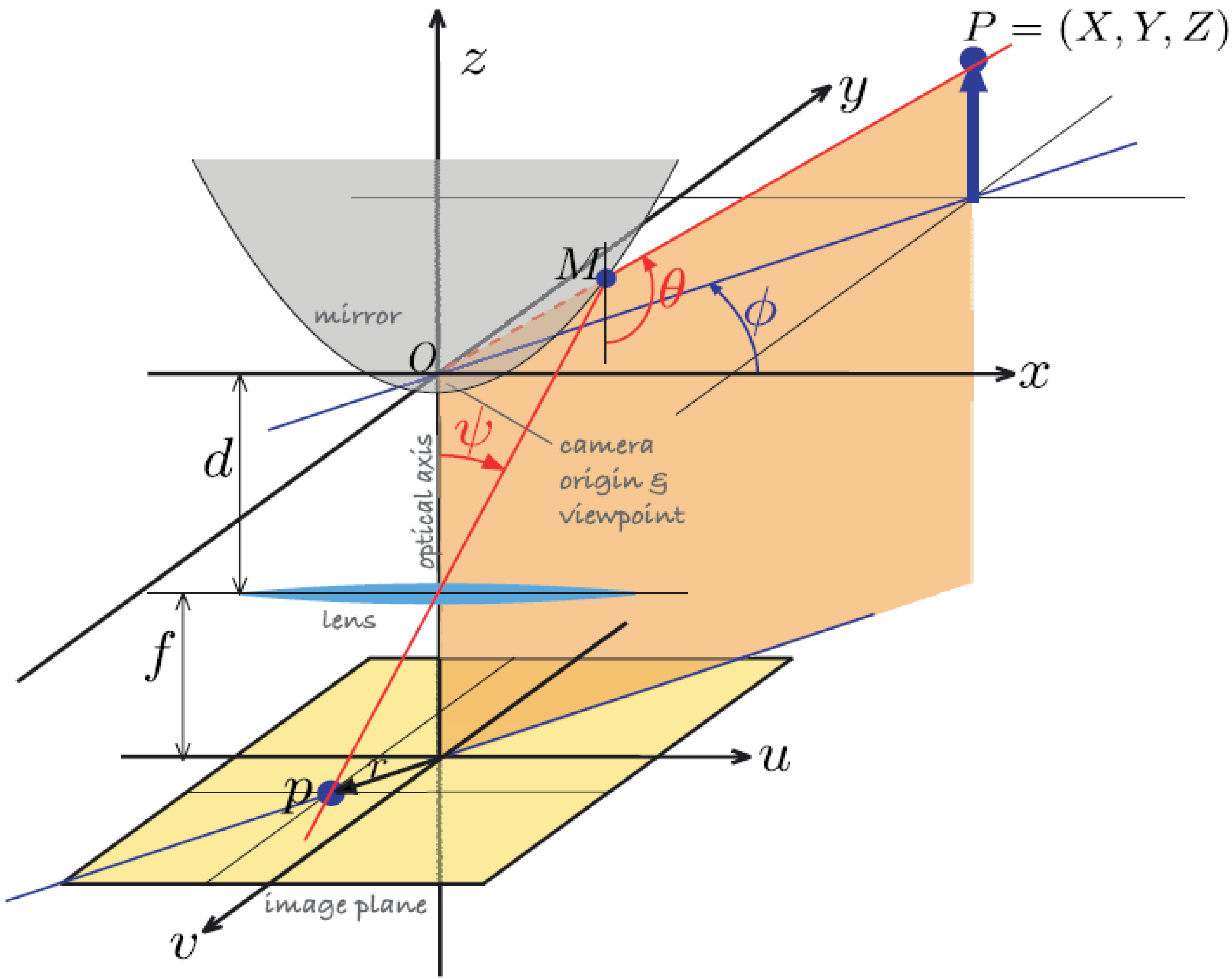
Mapping	Equation
Equiangular	$r = k\theta$
Stereographic	$r = k\tan(\theta/2)$
Equisolid	$r = k\sin(\theta/2)$
Polynomial	$r = k_1\theta + k_2\theta^2 + \cdots$

Catadioptric sensors

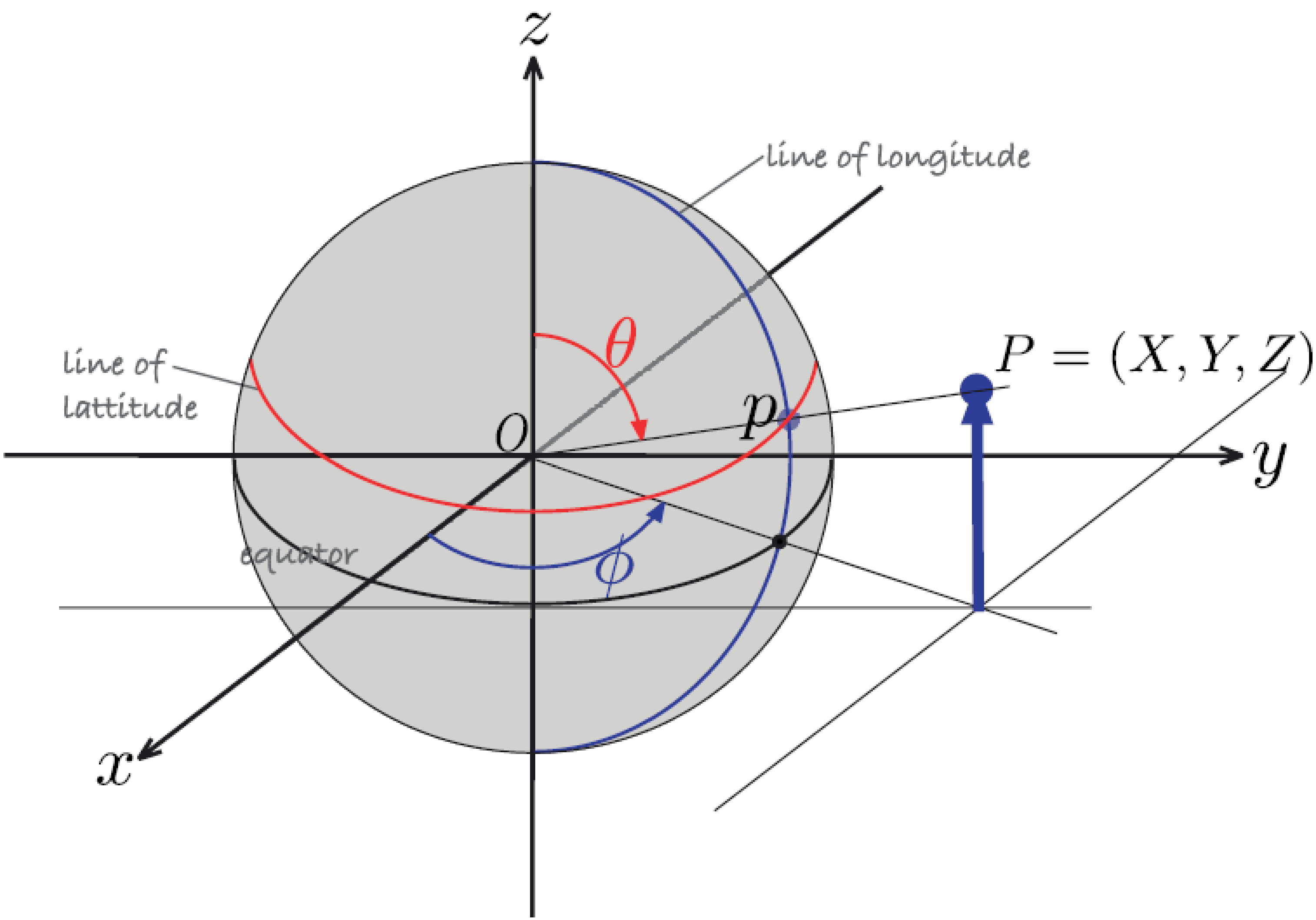
Catadioptric sensor and its image



Catadioptric geometry



Spherical Eye



Spherical Eye equations

$$p = (x, y, z)$$

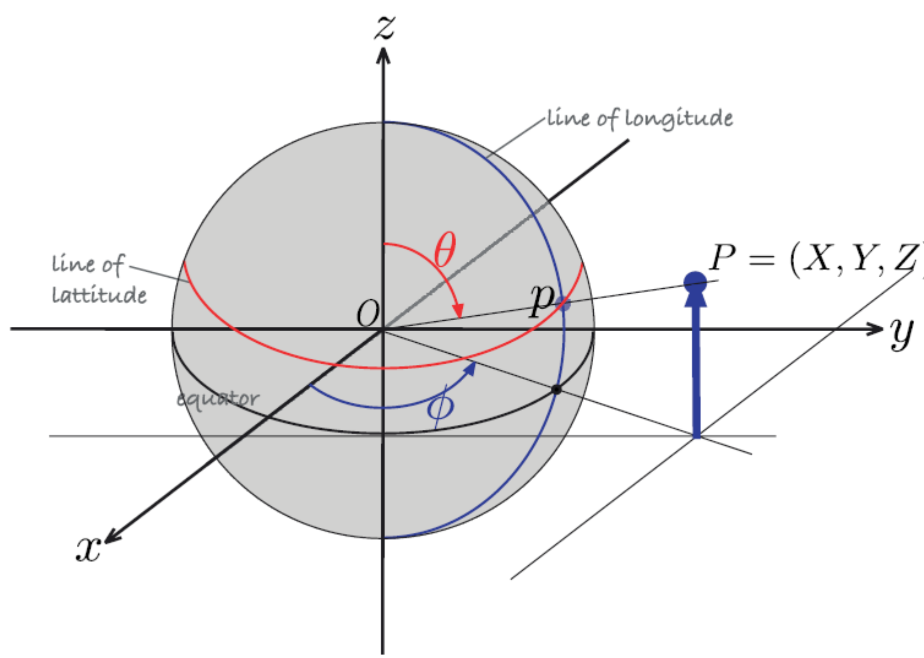
$$x = \frac{X}{R},\ y = \frac{Y}{R},\ \text{and } z = \frac{Z}{R}$$

$$R = \sqrt{(X^2 + Y^2 + Z^2)}$$

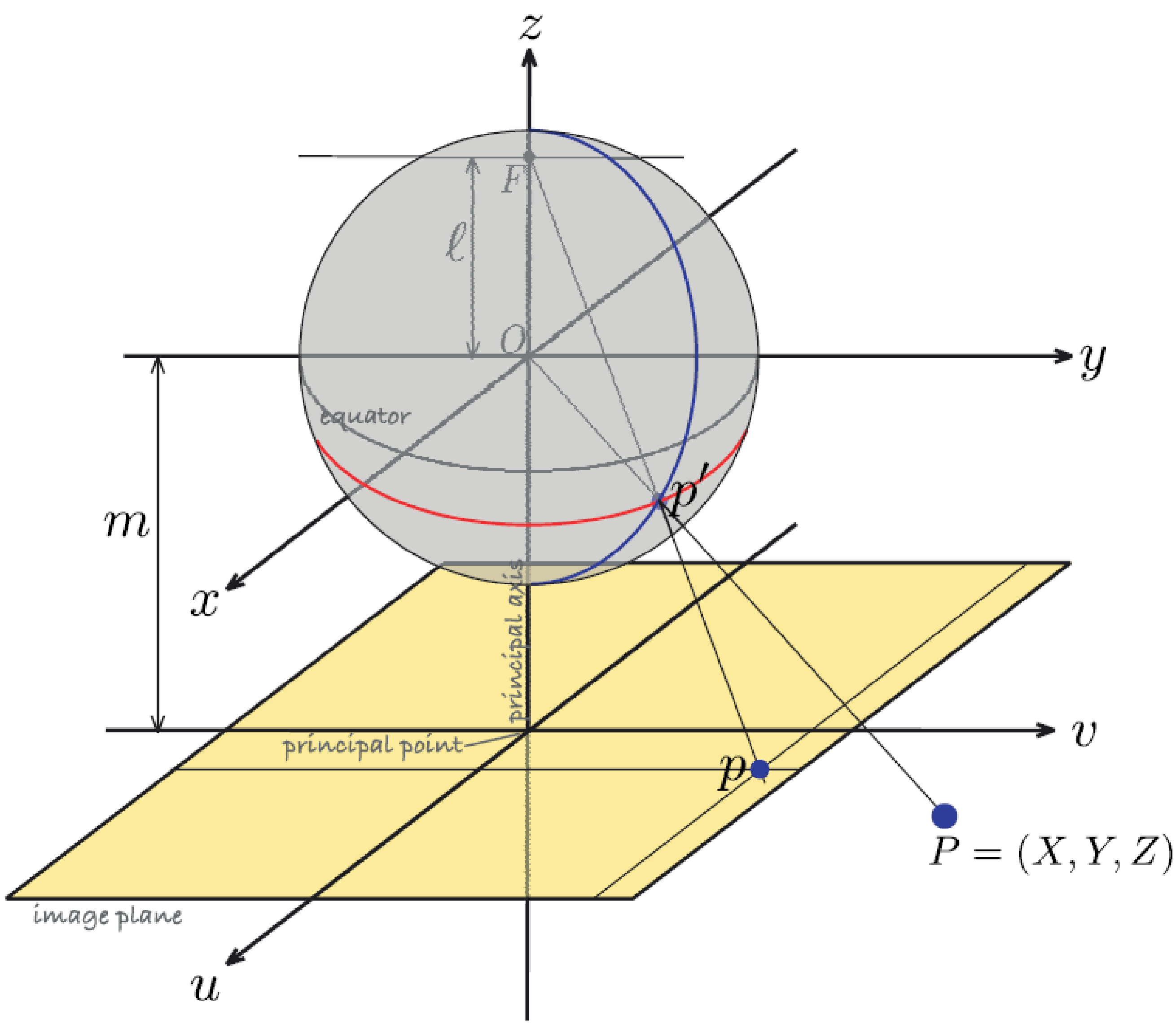
$$\theta = \sin^{-1} r,\ \ \theta \in [0, \pi]$$

where $r = \sqrt{x^2 + y^2}$, and the azimuth angle (or longitude)

$$\phi = \tan^{-1}\frac{y}{x},\ \ \phi \in [-\pi, \pi)$$



Unified imaging

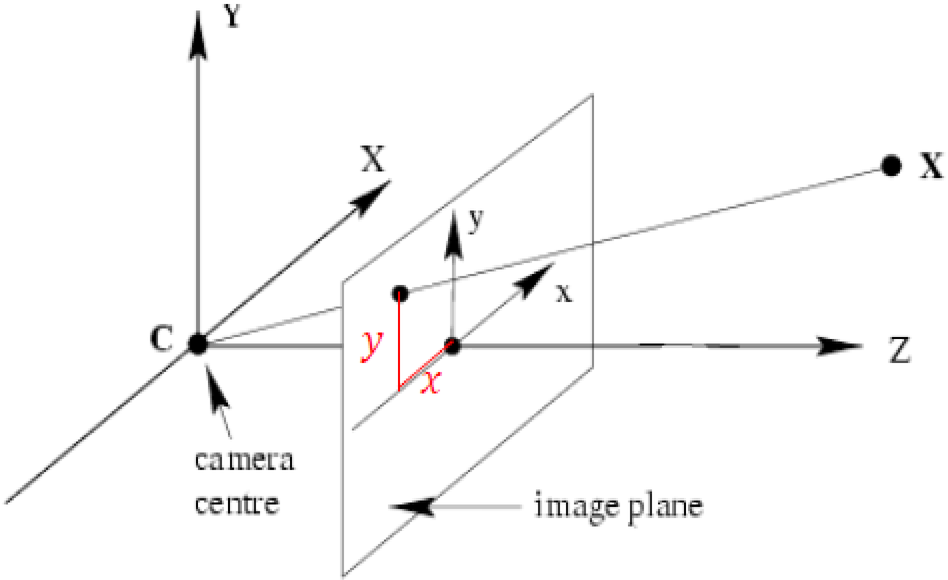


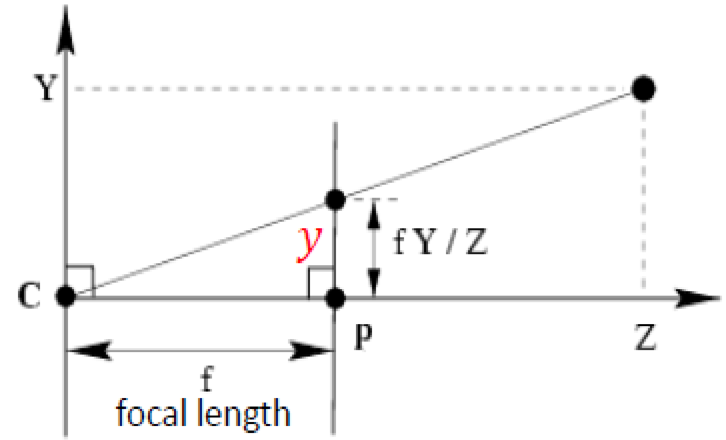
Unified imaging

Imaging	ℓ	m
Perspective	0	f
Stereographic	1	f
Fisheye	>1	f
Catadioptric (elliptical, $0 < \varepsilon < 1$)	$\frac{2\varepsilon}{1+\varepsilon^2}$	$\frac{2\varepsilon(2p-1)}{1+\varepsilon^2}$
Catadioptric (parabolic, $\varepsilon = 1$)	1	$2p-1$
Catadioptric (hyperbolic, $\varepsilon > 1$)	$\frac{2\varepsilon}{1+\varepsilon^2}$	$\frac{2\varepsilon(2p-1)}{1+\varepsilon^2}$

Back to equations

Pinhole camera model – in maths

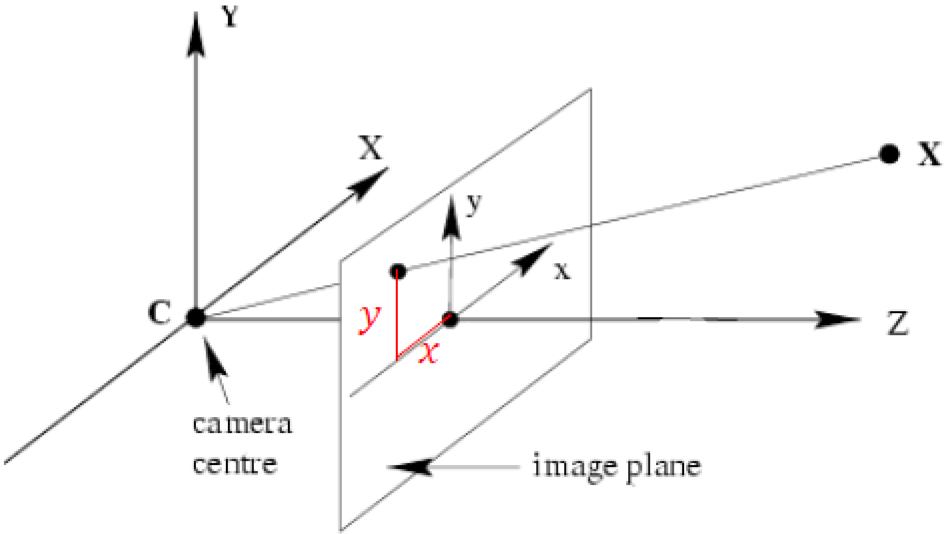


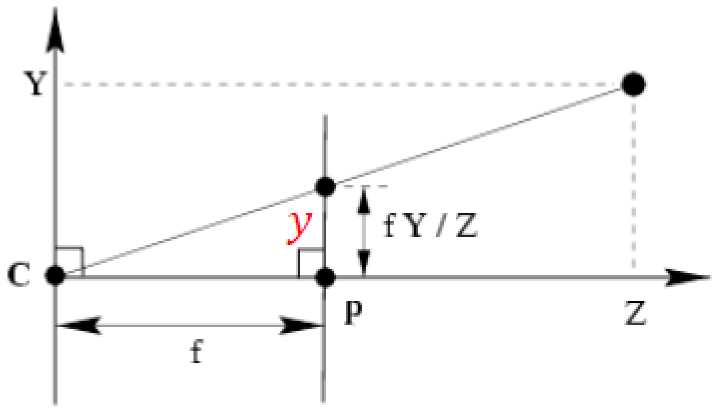


- Similar trinagles: $\dfrac{y}{f} = \dfrac{Y}{Z}$
- That gives: $y = f\frac{Y}{Z}$ and $x = f\frac{X}{Z}$

That gives: $\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$

Pinhole camera model – in maths





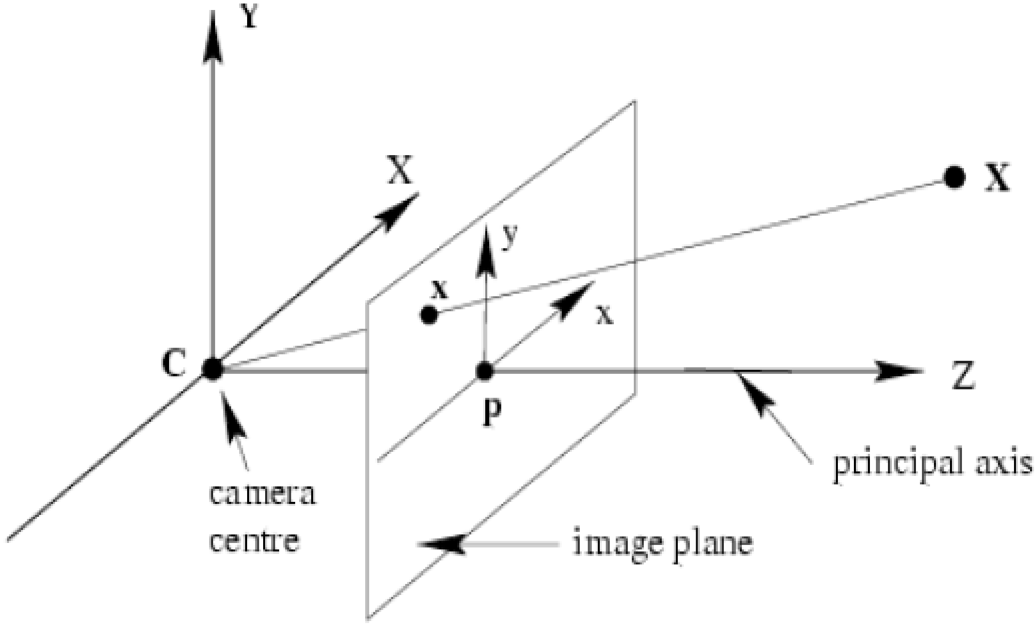
That gives: $\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Calibration matrix } \boldsymbol{K}} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$

In short $\boldsymbol{x} = \boldsymbol{K}\,\tilde{\boldsymbol{X}}$ (here $\tilde{\boldsymbol{X}}$ means inhomogeneous coordinates)

Intrinsic Camera Calibration means we know $\boldsymbol{K}$ (we do that later)

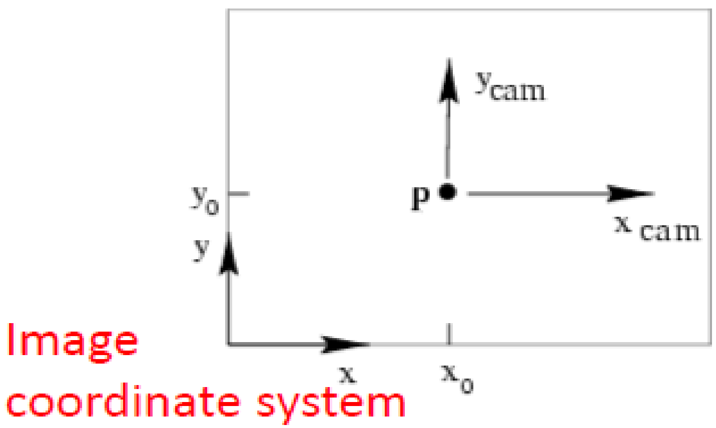
We can go from image points into the 3D world: $\tilde{\boldsymbol{X}} = \boldsymbol{K}^{-1}\,\boldsymbol{x}$

Pinhole camera - definitions



- Principal axis: line from the camera center perpendicular to the image plane
- Normalized (camera) coordinate system: camera center is at the origin and the principal axis is the z-axis
- Principal point (p): point where principal axis intersects the image plane (origin of normalized coordinate system)

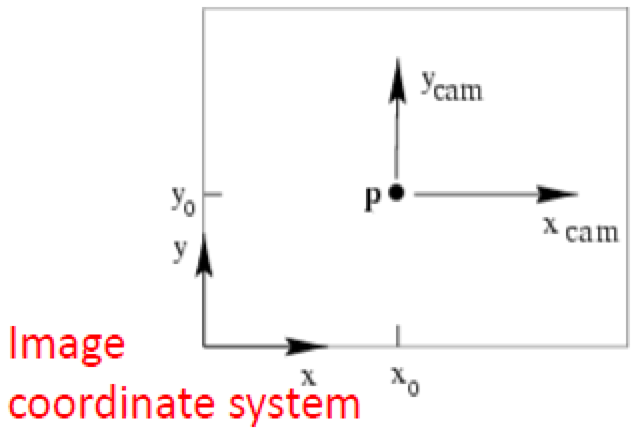
Principal Point



Principal point (p_x, p_y)

- Camera coordinate system: origin is at the principal point
- Image coordinate system: origin is in the corner
  In practice: principal point in center of the image

Adding principal point into $\boldsymbol{K}$

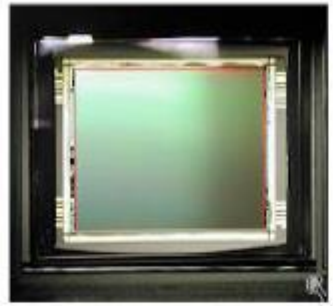


Principal point (p_x, p_y)

Projection with principal point : $y = f\frac{Y}{Z} + p_y = \frac{fY+Zp_y}{Z}$ and $x = f\frac{X}{Z} + p_x = \frac{fX+Zp_x}{Z}$

That gives: $\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$

Pixel Size and Shape



- m_x pixels per unit (m,mm,inch,...) in horizontal direction
- m_y pixels per unit (m,mm,inch,...) in vertical direction
- s' skew of a pixel
- *In practice (close to): m=1 s = 0*

That gives: $$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} m_x & s' & 0 \\ 0 & m_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

Simplified to: $$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} f & s & p_x \\ 0 & mf & p_y \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Final calibration matrix } \boldsymbol{K}} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

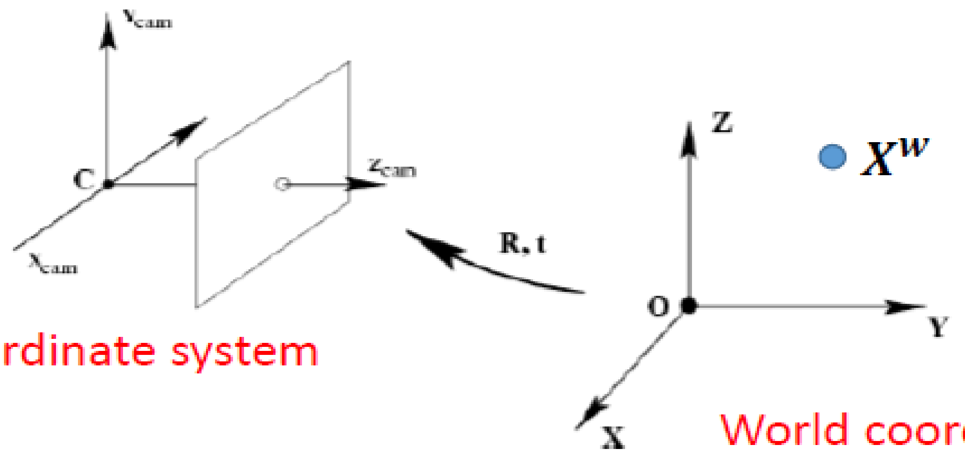
$\boxed{f \text{ now in units of pixels}}$

Camera intrinsic parameters - Summary

- Intrinsic parameters
  - Principal point coordinates (p_x, p_y)
  - Focal length f
  - Pixel magnification factors m
  - Skew (non-rectangular pixels) s

$$\boldsymbol{K} = \begin{bmatrix} f & s & p_x \\ 0 & mf & p_y \\ 0 & 0 & 1 \end{bmatrix}$$

Putting the camera into the world



camera coordinate system

World coordinate system

Given a 3D homogenous point X^w in world coordinate system

1) Translate from world to camera coordinate system:

$$\tilde{X}^{c'} = \tilde{X}^w - \tilde{C}$$

$$\tilde{X}^{c'} = \underbrace{(I_{3\times3} \mid -\tilde{C})}_{3\times4 \text{ matrix}} X^w$$ where $I_{3\times3}$ is 3x3 identity matrix

2) Rotate world coordinate system into camera coordinate system

$$\tilde{X}^c = R\,(I_{3\times3} \mid -\tilde{C})\,X^w$$

3) Apply camera matrix

$$x = K\,R\,(I_{3\times3} \mid -\tilde{C})\,X$$

Camera matrix

- Camera matrix $\boldsymbol{P}$ is defined as:

$$\boldsymbol{x} = \underbrace{\boldsymbol{K}\,\boldsymbol{R}\,(\boldsymbol{I}_{3\times 3} \mid -\tilde{\boldsymbol{C}})}_{\boldsymbol{P}\,(3\times 4)\text{ camera matrix has 11 DoF}}\,\boldsymbol{X}$$

- In short we write: $\boldsymbol{x} = \boldsymbol{P}\,\boldsymbol{X}$

- The camera center is the (right) nullspace of P

$\boldsymbol{P}\,\boldsymbol{C} = \boldsymbol{K}\,\boldsymbol{R}\,(\tilde{\boldsymbol{C}} - \tilde{\boldsymbol{C}}) = 0$

Camera parameters - Summary

- Camera matrix P has 11 DoF

$$\boldsymbol{x} = \boldsymbol{P}\,\boldsymbol{X}$$

$$\boldsymbol{x} = \boldsymbol{K}\,\boldsymbol{R}\,(\boldsymbol{I}_{3\times3}\,|\,-\tilde{\boldsymbol{C}})\,\boldsymbol{X}$$

$$\boldsymbol{K} = \begin{bmatrix} f & s & p_x \\ 0 & mf & p_y \\ 0 & 0 & 1 \end{bmatrix}$$

- Intrinsic parameters
  - Principal point coordinates (p_x, p_y)
  - Focal length f
  - Pixel magnification factors m
  - Skew (non-rectangular pixels) s

- Extrinsic parameters
  - Rotation $\boldsymbol{R}$ (3DoF) and translation $\tilde{\mathbf{C}}$ (3DoF) relative to world coordinate system

Properties of P: the projective camera

- **P** is a general 3X4 matrix of rank 3. If it has rank less than 3 then the projection is line or point, not the whole plane.

- **C** is the camera center if and only if **PC=0.** Indeed, if **PC=0,** consider the line containing **C** and any point **A** in 3-space. Points on this line are: **X**(k)=k**A**+(1-k)**C.** Mapping any such point with P we get: **PX**=k**PA.** All points on the line map to **PA**.

The rows and columns of P

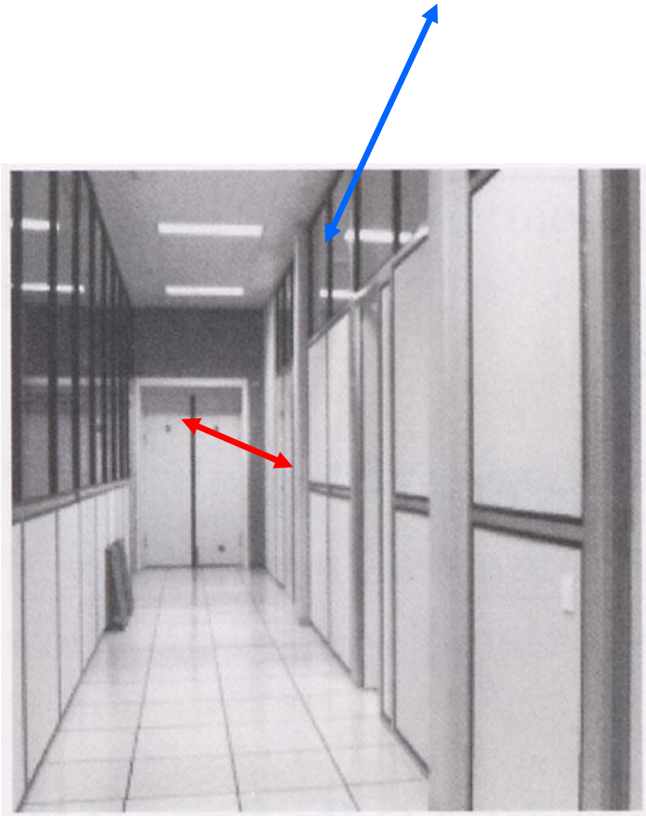
$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix}$$

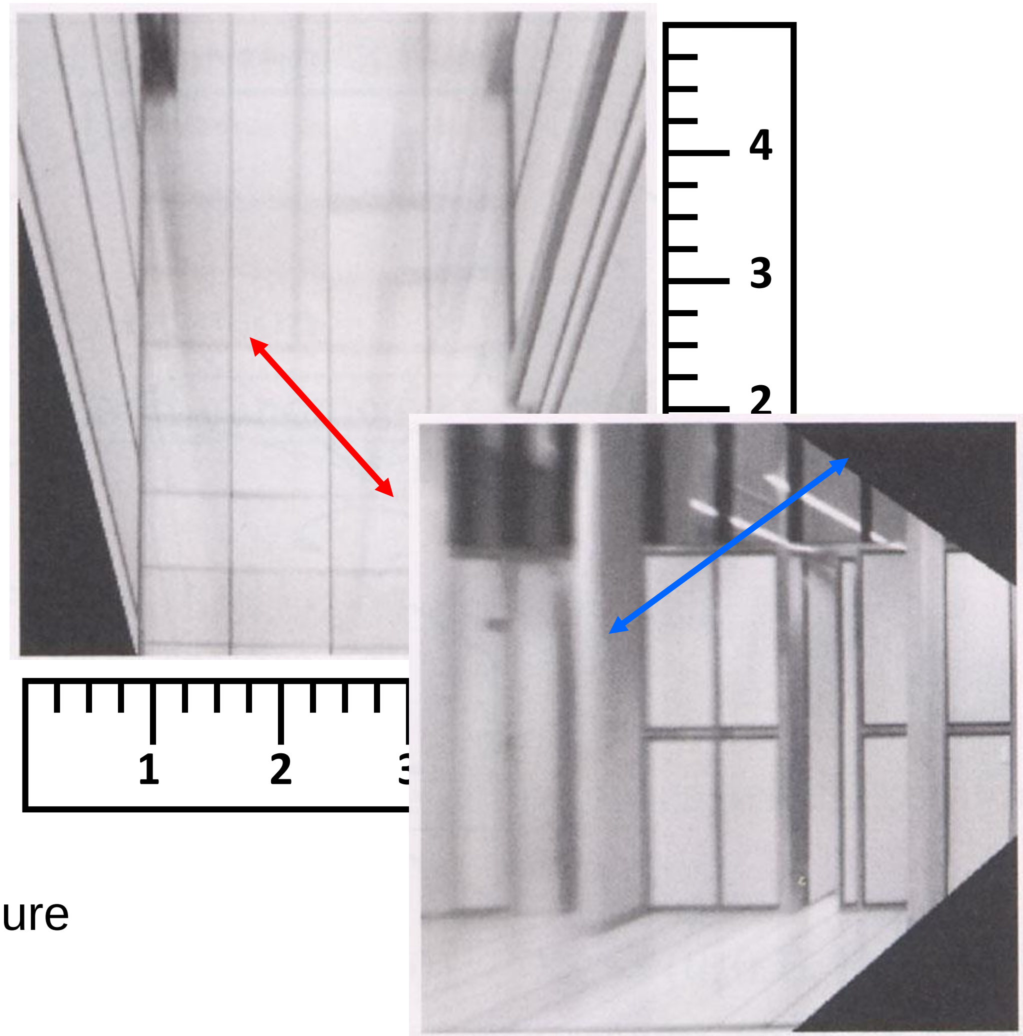
$$P = \begin{bmatrix} p^{1T} \\ p^{2T} \\ p^{3T} \end{bmatrix} \qquad P = \begin{bmatrix} p_1 \; p_2 \; p_3 \; p_4 \end{bmatrix}$$

Columns=image points, rows=world planes

- The first three columns are the vanishing points of the world axes X,Y,Z; X axis has direction D=(1,0,0,0) and **PD=p1**, etc. The last column is the image of the world origin.
- The rows represent planes:
- **The principal plane** (through camera center parallel to the image): If **X** lies on this plane then **PX=(x,y,0), or $\boldsymbol{p^{3T}}$X=0.**
- Consider points **X** on plane p^{1T} . Then p^{1T}**X =0, so PX=(0,y,w),** the points on the image x-axis. Since also p^{1T}**C=0,** plane p^{1T} passes from the camera center and the y image axis. Similarly for the second plane.

Measurements on planes

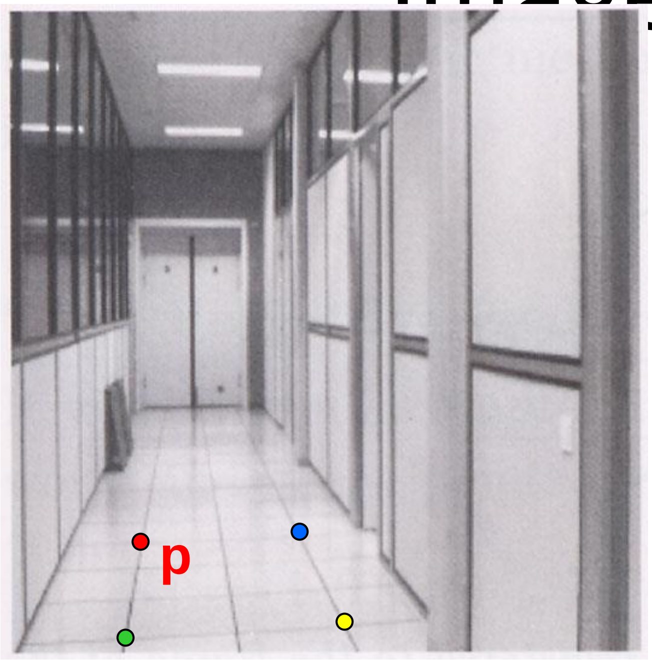


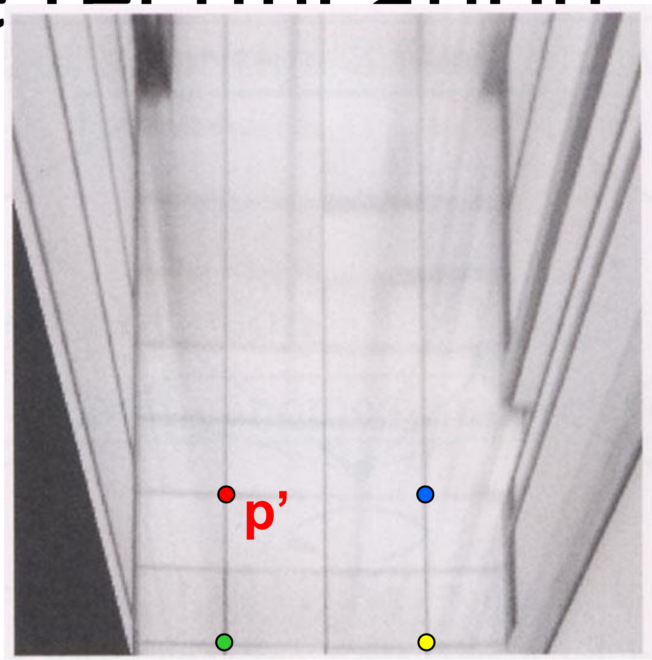


Approach: unwarp then measure

What kind of warp is this?

Image rectification





To unwarp (rectify) an image

- solve for homography **H** given **p** and **p’**
- solve equations of the form: w**p’** = **Hp**
  - linear in unknowns: w and coefficients of **H**
  - H is defined up to an arbitrary scale factor
  - how many points are necessary to solve for **H**?

Solving for homographies

$$\begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix} \cong \begin{bmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}$$

$$\begin{aligned} x'_i &= \frac{h_{00}x_i + h_{01}y_i + h_{02}}{h_{20}x_i + h_{21}y_i + h_{22}} \\ y'_i &= \frac{h_{10}x_i + h_{11}y_i + h_{12}}{h_{20}x_i + h_{21}y_i + h_{22}} \end{aligned}$$

$$\begin{aligned} x'_i(h_{20}x_i + h_{21}y_i + h_{22}) &= h_{00}x_i + h_{01}y_i + h_{02} \\ y'_i(h_{20}x_i + h_{21}y_i + h_{22}) &= h_{10}x_i + h_{11}y_i + h_{12} \end{aligned}$$

$$\begin{bmatrix} x_i & y_i & 1 & 0 & 0 & 0 & -x'_i x_i & -x'_i y_i & -x'_i \\ 0 & 0 & 0 & x_i & y_i & 1 & -y'_i x_i & -y'_i y_i & -y'_i \end{bmatrix} \begin{bmatrix} h_{00} \\ h_{01} \\ h_{02} \\ h_{10} \\ h_{11} \\ h_{12} \\ h_{20} \\ h_{21} \\ h_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Solving for homographies

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x'_1 x_1 & -x'_1 y_1 & -x'_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -y'_1 x_1 & -y'_1 y_1 & -y'_1 \\ & & & & \vdots & & & & \\ x_n & y_n & 1 & 0 & 0 & 0 & -x'_n x_n & -x'_n y_n & -x'_n \\ 0 & 0 & 0 & x_n & y_n & 1 & -y'_n x_n & -y'_n y_n & -y'_n \end{bmatrix} \begin{bmatrix} h_{00} \\ h_{01} \\ h_{02} \\ h_{10} \\ h_{11} \\ h_{12} \\ h_{20} \\ h_{21} \\ h_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

A **h** **0**

2n × 9 **9** **2n**

Linear least squares

- Since **h** is only defined up to scale, solve for unit vector **ĥ**
- Minimize $\|\mathbf{A}\hat{\mathbf{h}}\|^2$

$$\|\mathbf{A}\hat{\mathbf{h}}\|^2 = (\mathbf{A}\hat{\mathbf{h}})^T\mathbf{A}\hat{\mathbf{h}} = \hat{\mathbf{h}}^T\mathbf{A}^T\mathbf{A}\hat{\mathbf{h}}$$

- Solution: **ĥ** = eigenvector of $\mathbf{A}^T\mathbf{A}$ with smallest eigenvalue
- Works with 4 or more points