COLLEGE OF ENGINEERING
ELECTRICAL ENGINEERING & COMPUTER SCIENCE
UNIVERSITY OF MICHIGAN

Geometrical Primitives, Transformations and Image Formation

EECS 598-08 Fall 2014
Foundations of Computer Vision

http://web.eecs.umich.edu/~jjcorso/t/598F14

Instructor: Jason Corso
jjcorso@eecs.umich.edu

Materials on these slides have come from many sources in addition to myself; I am infinitely grateful to these, especially Greg Hager and Silvio Savarese.

Plan

- Geometric Primitives
  - Points, Lines in 2D and 3D
  - Transformations in 2D and 3D
- Basic Image Formation
- Camera Parameters
- Lens Distortion

Geometric Primitives

- 2D points: pixel coordinates

$$\mathbf{x} = \begin{bmatrix} x & y \end{bmatrix}^\mathsf{T} \in \mathbb{R}^2$$

- Using homogeneous coordinates
  - Vectors differing by scale are equivalent.

$$\tilde{\mathbf{x}} = \begin{bmatrix} \tilde{x} & \tilde{y} & \tilde{w} \end{bmatrix}^\mathsf{T} \in \mathbb{P}^2$$

$$\tilde{\mathbf{x}} = \tilde{w} \begin{bmatrix} x & y & 1 \end{bmatrix}^\mathsf{T} = \tilde{w}\bar{\mathbf{x}}$$

$$\bar{\mathbf{x}} = \begin{bmatrix} x & y & 1 \end{bmatrix}$$

augmented vector

2D Projective Space

$$\mathbb{P}^2 = \mathbb{R}^3 - \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^\mathsf{T}$$

  - When the last element $\tilde{w} = 0$, call it an *ideal point.*

Geometric Primitives

- 2D lines with homogeneous coordinates

$$\tilde{\boldsymbol{l}} = \begin{bmatrix} a & b & c \end{bmatrix}^\mathsf{T}$$

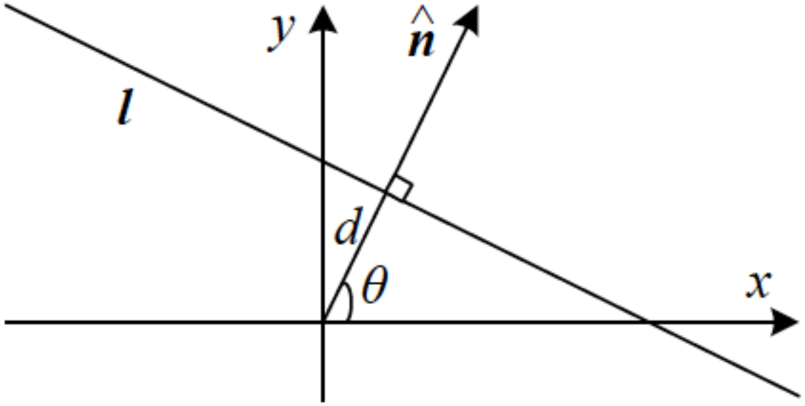
$$\bar{\mathbf{x}}^\mathsf{T} \tilde{\boldsymbol{l}} = ax + by + c = 0$$

- Normalized coordinates

$$\boldsymbol{l} = \begin{bmatrix} \hat{n}_x & \hat{n}_y & d \end{bmatrix}^\mathsf{T} = \begin{bmatrix} \underbrace{\hat{\mathbf{n}}^\mathsf{T}}_{\text{normal vector}} d \end{bmatrix}^\mathsf{T} \quad \text{s.t.} \quad \|\underbrace{\hat{\mathbf{n}}}_{\text{normal vector}}\| = 1$$

- Polar coordinates

$$\begin{aligned} \boldsymbol{l} &= (\theta, d) \\ &= \begin{bmatrix} \cos\theta & \sin\theta & d \end{bmatrix} \end{aligned}$$



The image on this slides is sourced from the Szeliski book.

Geometric Primitives

- Intersection of two lines

$$\tilde{\mathbf{x}} = \tilde{\boldsymbol{l}}_1 \times \tilde{\boldsymbol{l}}_2$$

- Line connecting two points

$$\tilde{\boldsymbol{l}} = \tilde{\mathbf{x}}_1 \times \tilde{\mathbf{x}}_2$$

Geometric Primitives

- 3D points

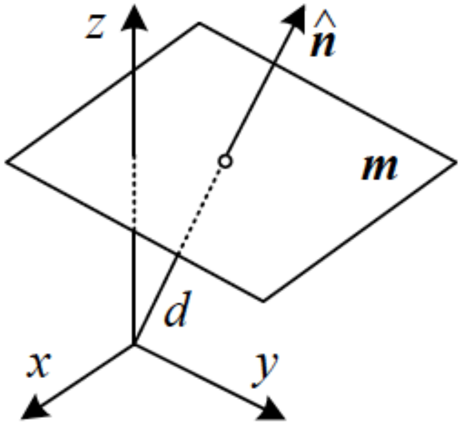
$$\mathbf{X} = \begin{bmatrix} X & Y & Z \end{bmatrix}^\mathsf{T} \in \mathbb{R}^3$$

$$\tilde{\mathbf{X}} = \begin{bmatrix} \tilde{X} & \tilde{Y} & \tilde{Z} & \tilde{W} \end{bmatrix}^\mathsf{T} \in \mathbb{P}^3$$

$$\tilde{\mathbf{X}} = \begin{bmatrix} \tilde{X} & \tilde{Y} & \tilde{Z} & 1 \end{bmatrix}^\mathsf{T} = \tilde{W}\overline{\mathbf{X}}$$

Geometric Primitives

- 3D planes

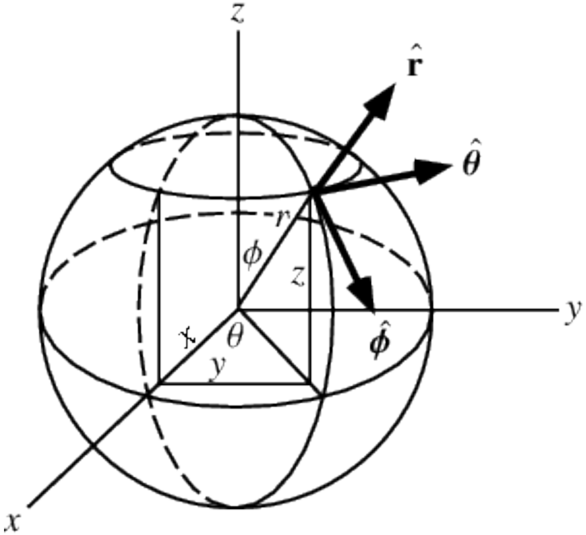


$$\tilde{\mathbf{M}} = \begin{bmatrix} A & B & C & D \end{bmatrix}^{\mathsf{T}}$$

$$\overline{\mathbf{X}}^{\mathsf{T}}\tilde{\mathbf{M}} = AX + BY + CZ + D = 0$$

$$\mathbf{M} = \begin{bmatrix} \hat{N}_X & \hat{N}_Y & \hat{N}_Z & D \end{bmatrix}^{\mathsf{T}} \quad \text{when} \quad \|\hat{\mathbf{N}}\| = 1$$

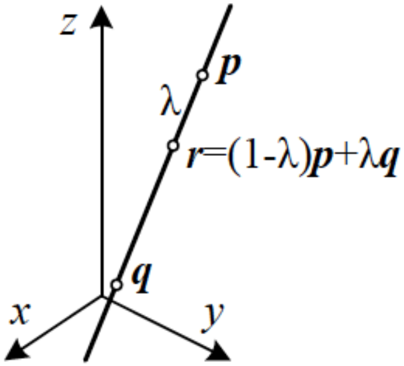
- Spherical coordinates
  - $\hat{\mathbf{N}}$ can be written as a function of two angles (θ, ϕ) .



$$\hat{\mathbf{N}} = \begin{bmatrix} \cos\theta\sin\phi & \sin\theta\cos\phi & \sin\phi \end{bmatrix}^{\mathsf{T}}$$

Images on this slides are sourced from the Szeliski book and Wolfram website.

Geometric Primitives



- 3D lines
  - Consider two points on the line $(\mathbf{P}, \mathbf{Q})$.

$$\boldsymbol{R} = (1-\lambda)\mathbf{P} + \lambda\mathbf{Q}$$

  - For the case of homogeneous coordinates:

$$\tilde{\boldsymbol{R}} = \mu\tilde{\mathbf{P}} + \lambda\tilde{\mathbf{Q}}$$

  - When the second point is at infinity,

$$\tilde{\mathbf{Q}} = \begin{bmatrix} \hat{V}_x & \hat{V}_y & \hat{V}_z & 0 \end{bmatrix}^{\mathsf{T}}$$

$$\boldsymbol{R} = \mathbf{P} + \lambda\tilde{\mathbf{Q}}$$

The image on this slides is sourced from the Szeliski book.

Geometric Transformations

- 2D translation

Identity matrix

$$\mathbf{x}' = \begin{bmatrix} \mathcal{I} & \mathbf{t} \end{bmatrix} \bar{\mathbf{x}}$$

$$\bar{\mathbf{x}}' = \begin{bmatrix} \mathcal{I} & \mathbf{t} \\ \mathbf{0}^{\mathsf{T}} & 1 \end{bmatrix} \bar{\mathbf{x}}$$

- 2D rotation and translation
  - 2D rigid body or Euclidean transformation

$$\mathbf{x}' = \begin{bmatrix} \mathcal{R} & \mathbf{t} \end{bmatrix} \bar{\mathbf{x}}$$

Rotation matrix

$$\mathcal{R} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$$\mathcal{R}\mathcal{R}^{\mathsf{T}} = \mathcal{I}$$

$$|\mathcal{R}| = 1$$

$$\bar{\mathbf{x}}' = \begin{bmatrix} \mathcal{R} & \mathbf{t} \\ \mathbf{0}^{\mathsf{T}} & 1 \end{bmatrix} \bar{\mathbf{x}}$$

Geometric Transformations

- 2D scaled rotation or similarity transform

$$\bar{\mathbf{x}}' = \begin{bmatrix} s\mathcal{R} & \mathbf{t} \\ \mathbf{0}^{\mathsf{T}} & 1 \end{bmatrix} \bar{\mathbf{x}} = \begin{bmatrix} a & -b & t_x \\ b & a & t_y \\ 0 & 0 & 1 \end{bmatrix} \bar{\mathbf{x}}$$

  - Constraint $a^2 + b^2 = 1$ is not enforced.

- 2D affine transformation

$$\bar{\mathbf{x}}' = \mathcal{A}\bar{\mathbf{x}} = \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix} \bar{\mathbf{x}}$$

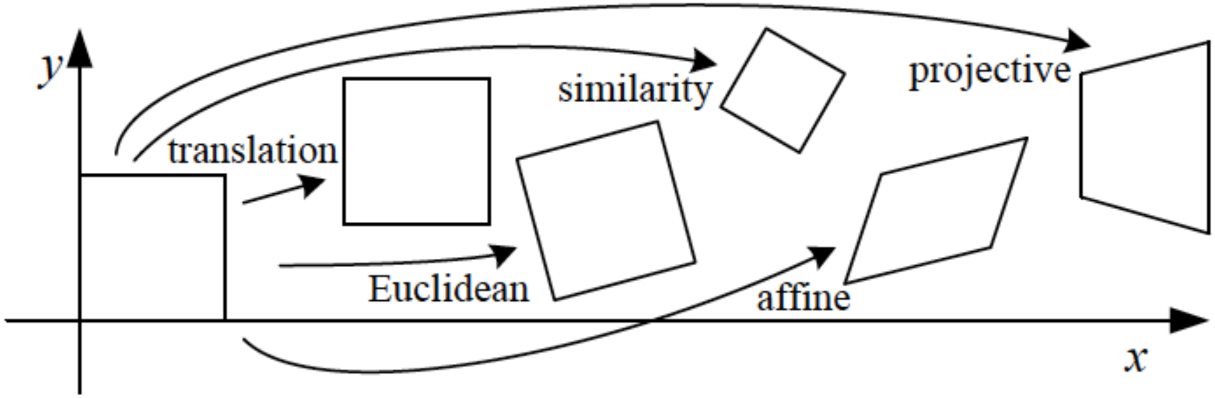
Geometric Transformations

- 2D projective, also called the homography

$$\tilde{\mathbf{x}}' = \tilde{\mathcal{H}}\tilde{\mathbf{x}} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{bmatrix} \tilde{\mathbf{x}}$$

- Projective matrix $\tilde{\mathcal{H}}$ is defined up to scale.
- Inhomogeneous results are computed after homogeneous operation.

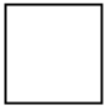
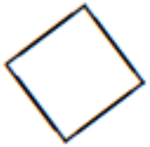
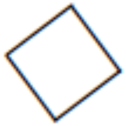
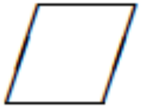
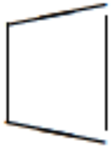
Hierarchy of 2D Planar Transformations



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

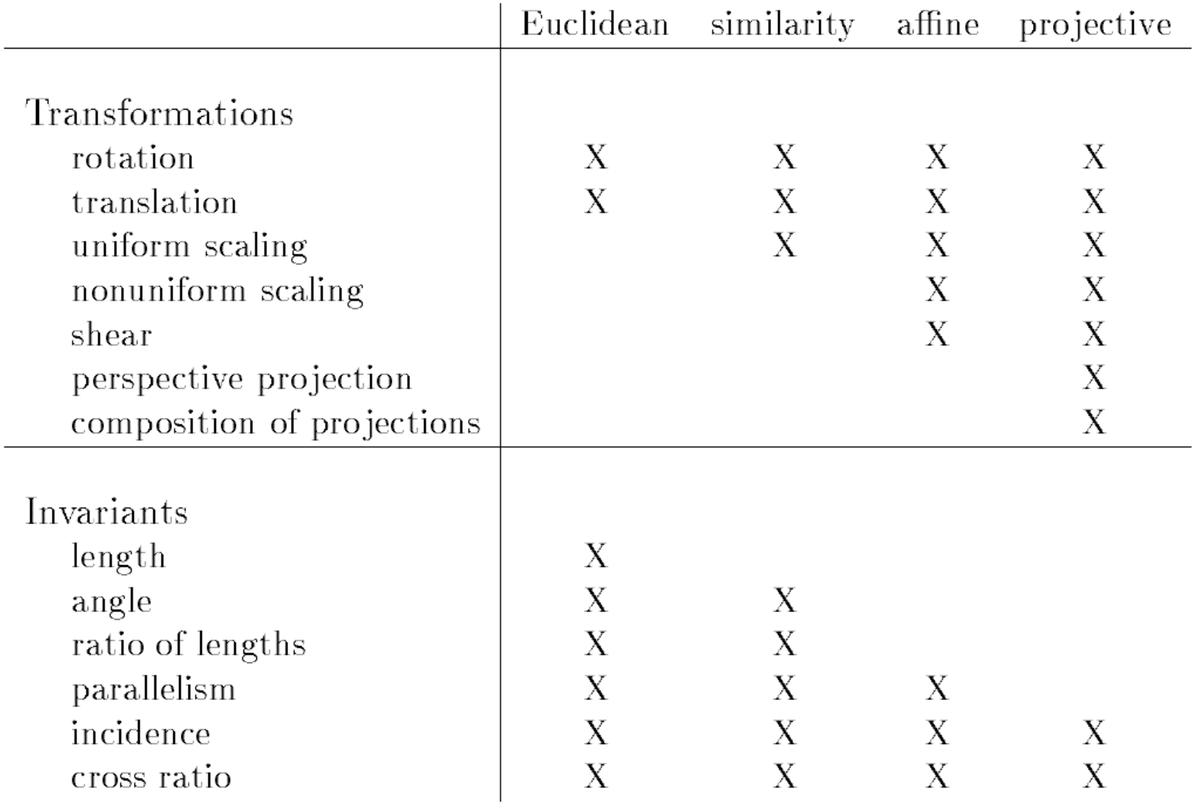
The images on this slides are sourced from the Szeliski book.

Hierarchy of 3D Coordinate Transformations

Transformation	**Matrix**	**# DoF**	**Preserves**	**Icon**
translation	$\left[\begin{array}{c\|c} \boldsymbol{I} & \boldsymbol{t} \end{array} \right]_{3\times 4}$	3	orientation	
rigid (Euclidean)	$\left[\begin{array}{c\|c} \boldsymbol{R} & \boldsymbol{t} \end{array} \right]_{3\times 4}$	6	lengths	
similarity	$\left[\begin{array}{c\|c} s\boldsymbol{R} & \boldsymbol{t} \end{array} \right]_{3\times 4}$	7	angles	
affine	$\left[\begin{array}{c} \boldsymbol{A} \end{array} \right]_{3\times 4}$	12	parallelism	
projective	$\left[\begin{array}{c} \tilde{\boldsymbol{H}} \end{array} \right]_{4\times 4}$	15	straight lines	

Projective Geometry

- These geometry basics are but the surface of an area important to computer vision called **projective geometry**.

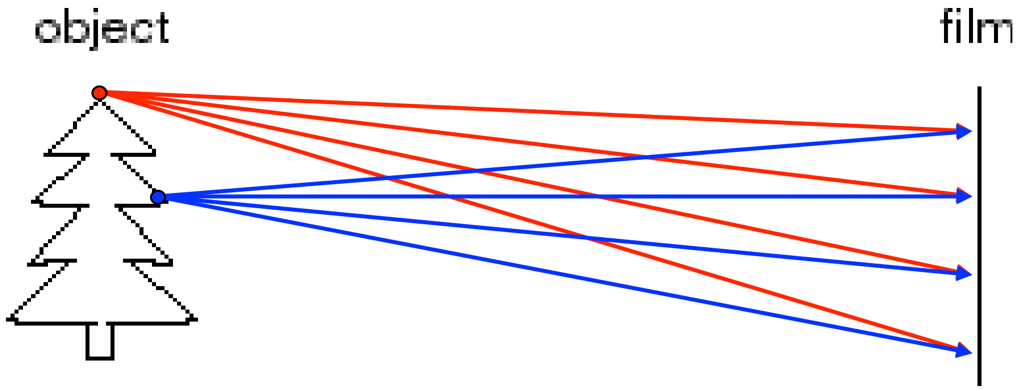


	Euclidean	similarity	affine	projective
Transformations				
rotation	X	X	X	X
translation	X	X	X	X
uniform scaling		X	X	X
nonuniform scaling			X	X
shear			X	X
perspective projection				X
composition of projections				X
Invariants				
length	X			
angle	X	X		
ratio of lengths	X	X		
parallelism	X	X	X	
incidence	X	X	X	X
cross ratio	X	X	X	X

- Further reading: “An Introduction to Projective Geometry” by Stan Birchfield.

Light

- Getting light to the sensor.

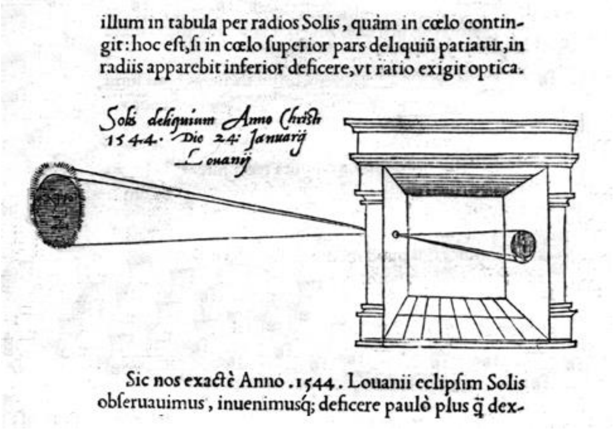


- What does this image look like?

Source: S. Savarese, GD Hager and S Seitz slides.

Light through a pinhole

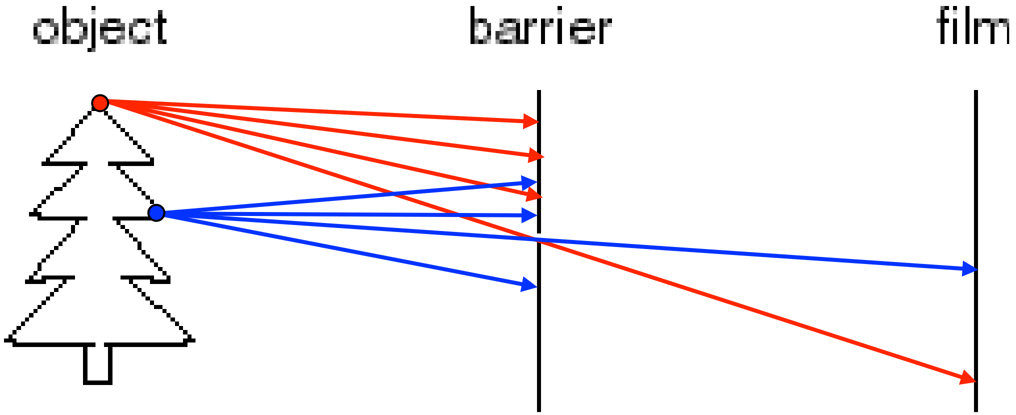
- Place a barrier in front of the film.
- Let a small pinhole of light through.
    - **aperture**



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

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

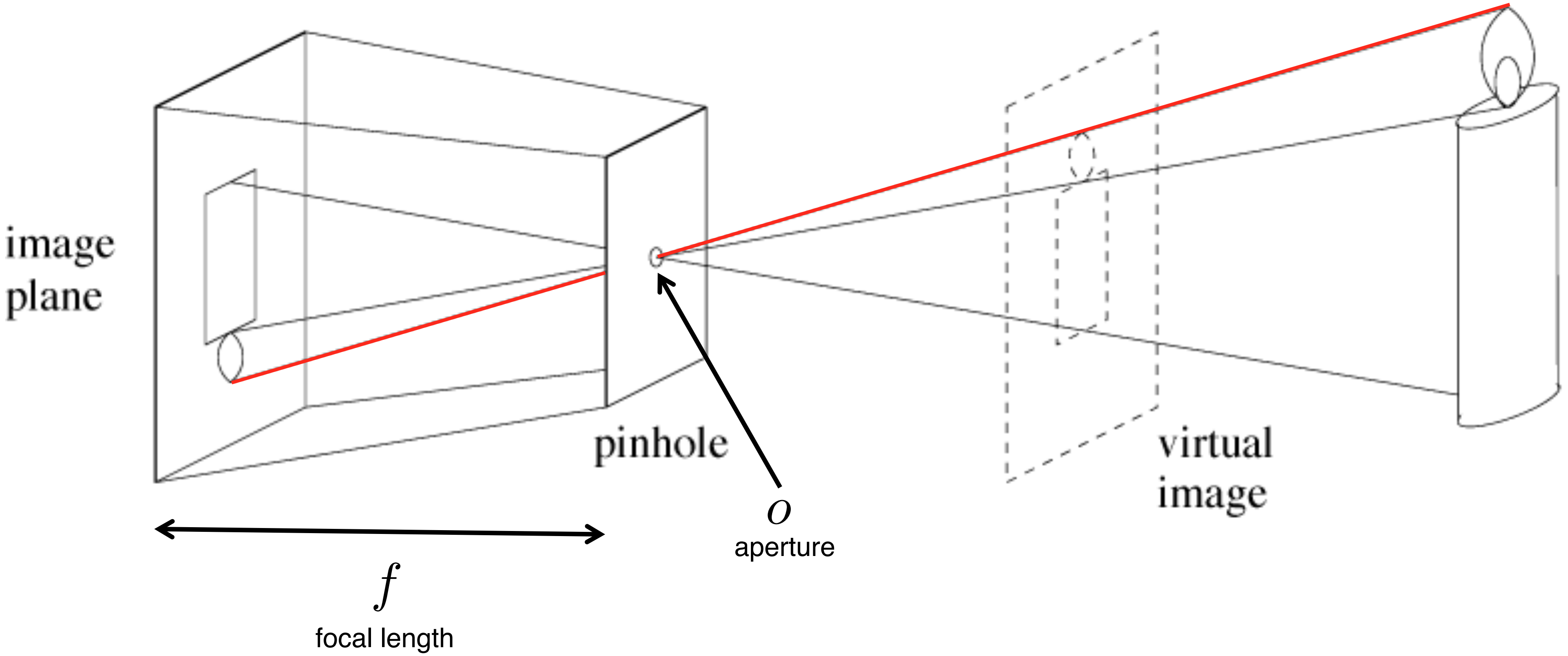
Leonardo da Vinci (1452-1519): Camera Obscura



Source: D. Forsyth, S. Savarese, GD Hager and S Seitz slides.

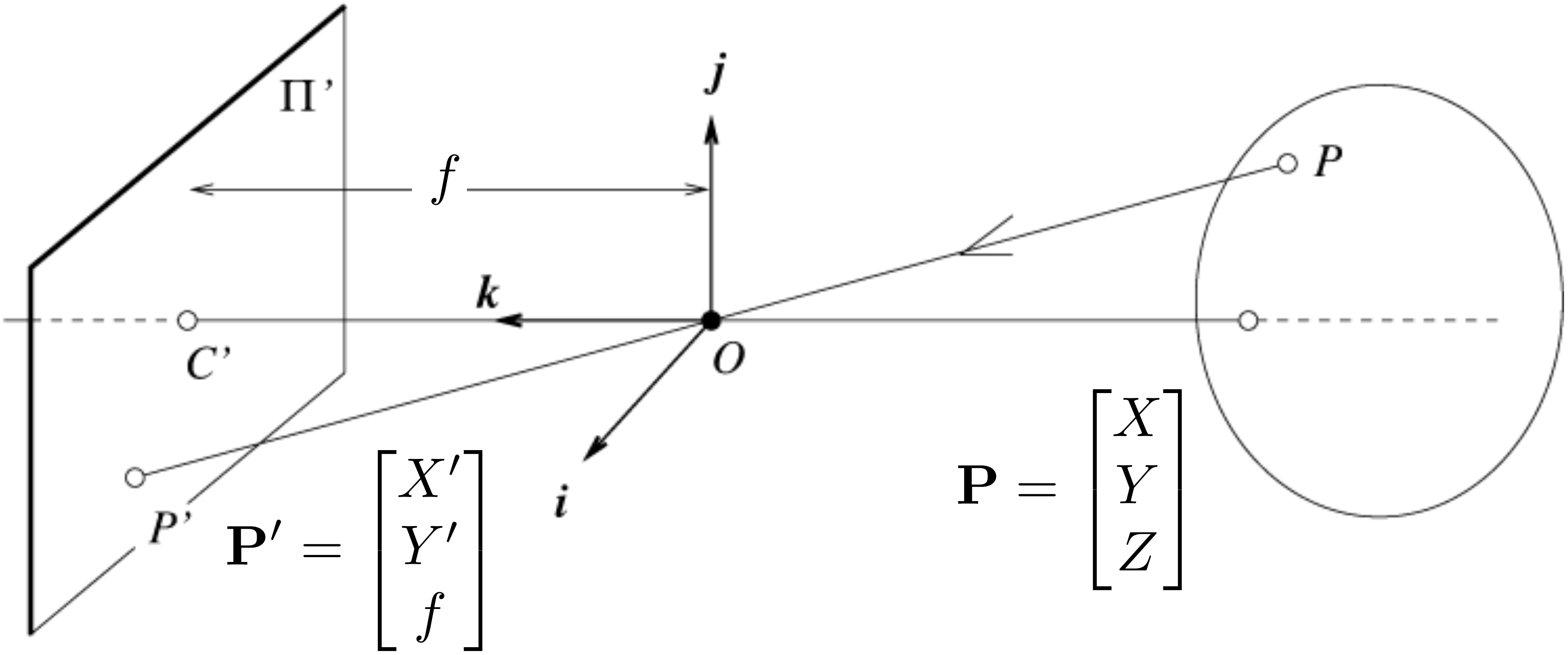
Light through a pinhole

- Pinhole: box with a small hole in it.
  - Abstract model that does indeed work in practice.



Source: D. Forsyth, S. Savarese slides.

Pinhole, or *Central*, Perspective



$$\mathbf{P}' = \begin{bmatrix} X' \\ Y' \\ f \end{bmatrix}$$

$$\mathbf{P} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

- Points P, O, P' are collinear.

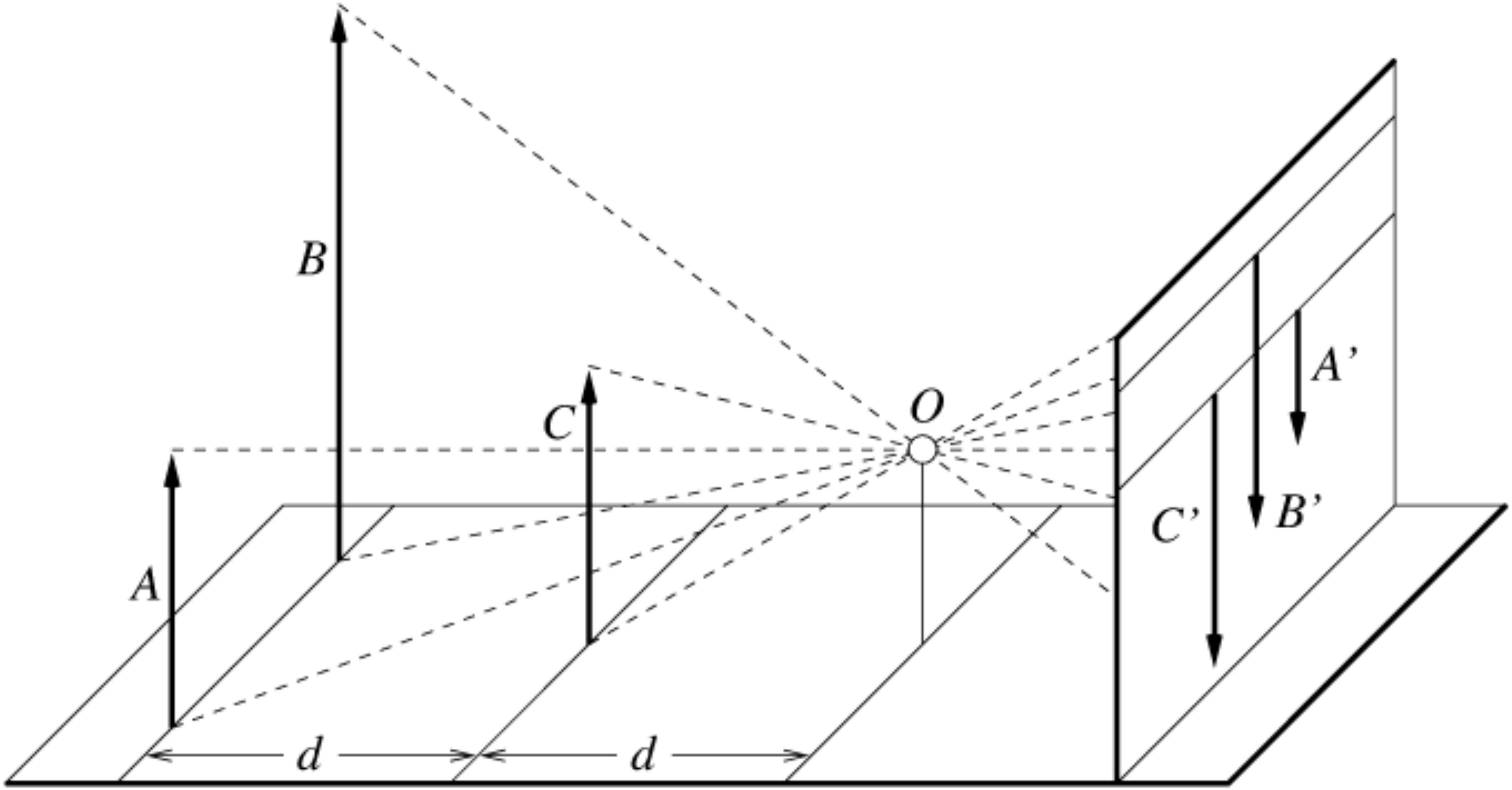
$$\overrightarrow{OP'} = \lambda \overrightarrow{OP} \longrightarrow \lambda = \frac{X'}{X} = \frac{Y'}{Y} = \frac{f}{Z}$$

- Therefore, we have $X' = f\frac{X}{Z}$ and $Y' = f\frac{Y}{Z}$.

Source: D. Forsyth, S. Savarese slides.

Properties of Pinhole Perspective Projection

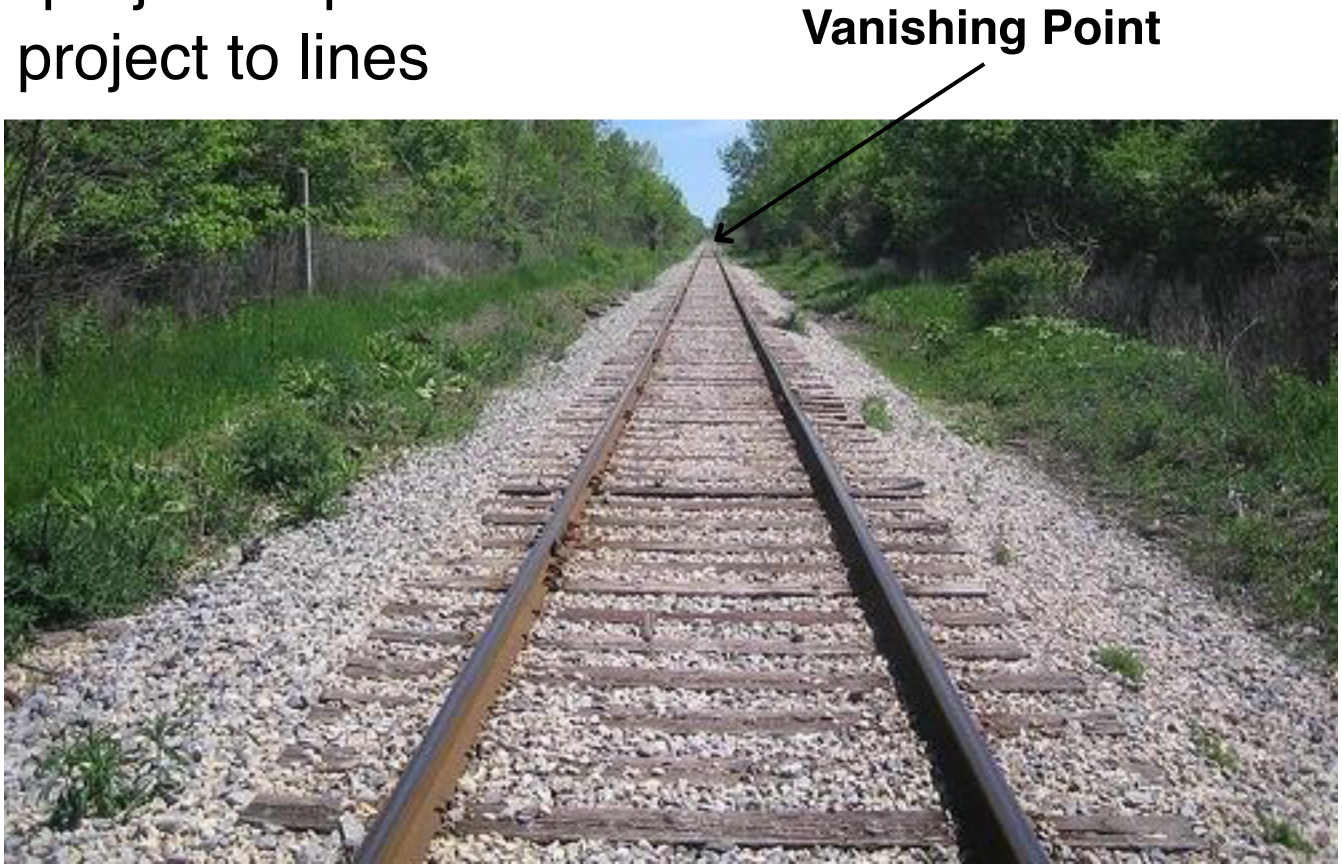
- Distant objects appear smaller



Source: D. Forsyth, S. Savarese slides.

Properties of Pinhole Perspective Projection

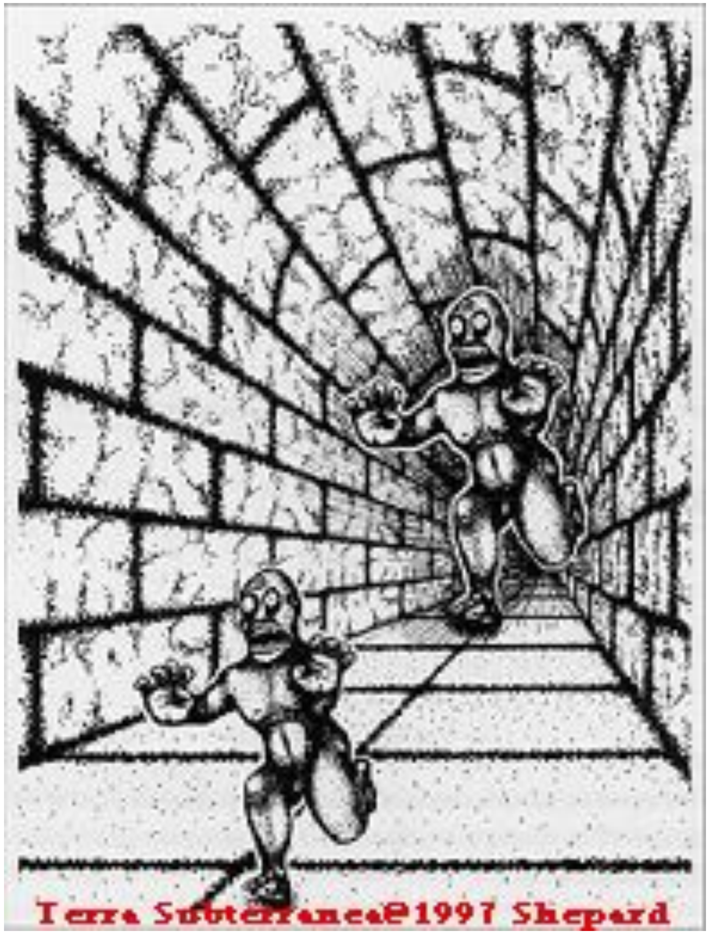
- Points project to points
- Lines project to lines

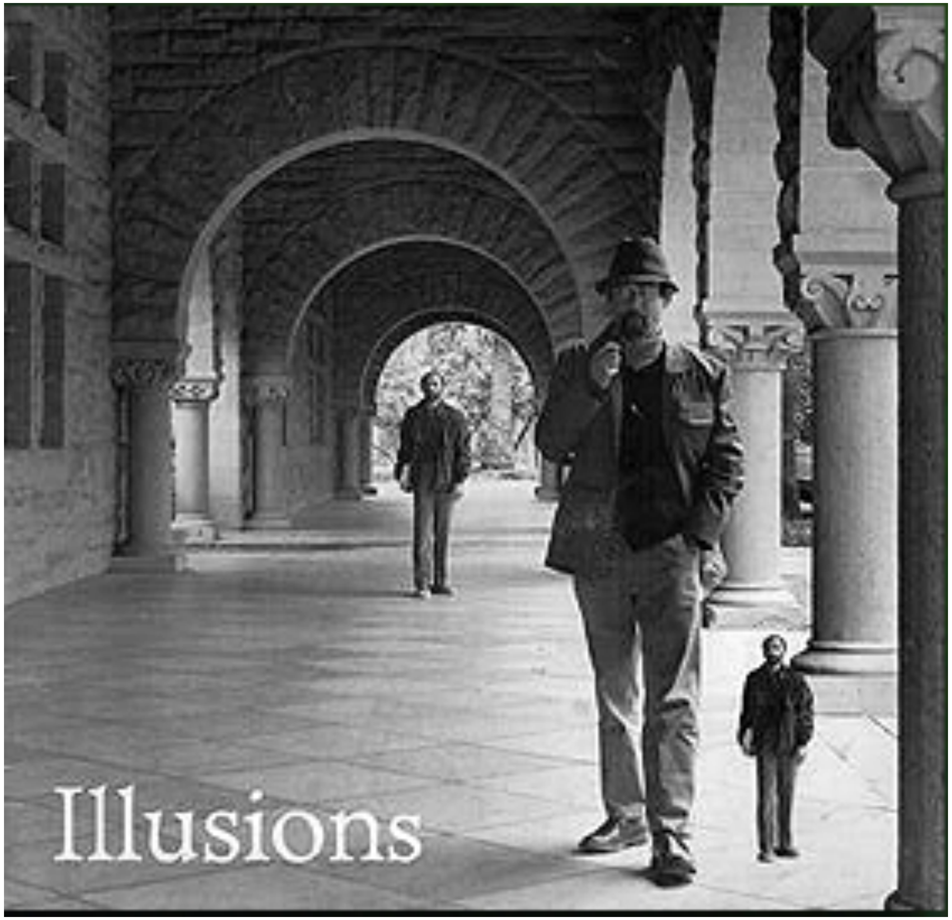


- Angles are not preserved.
- Parallel lines meet!

Source: S. Savarese slides.

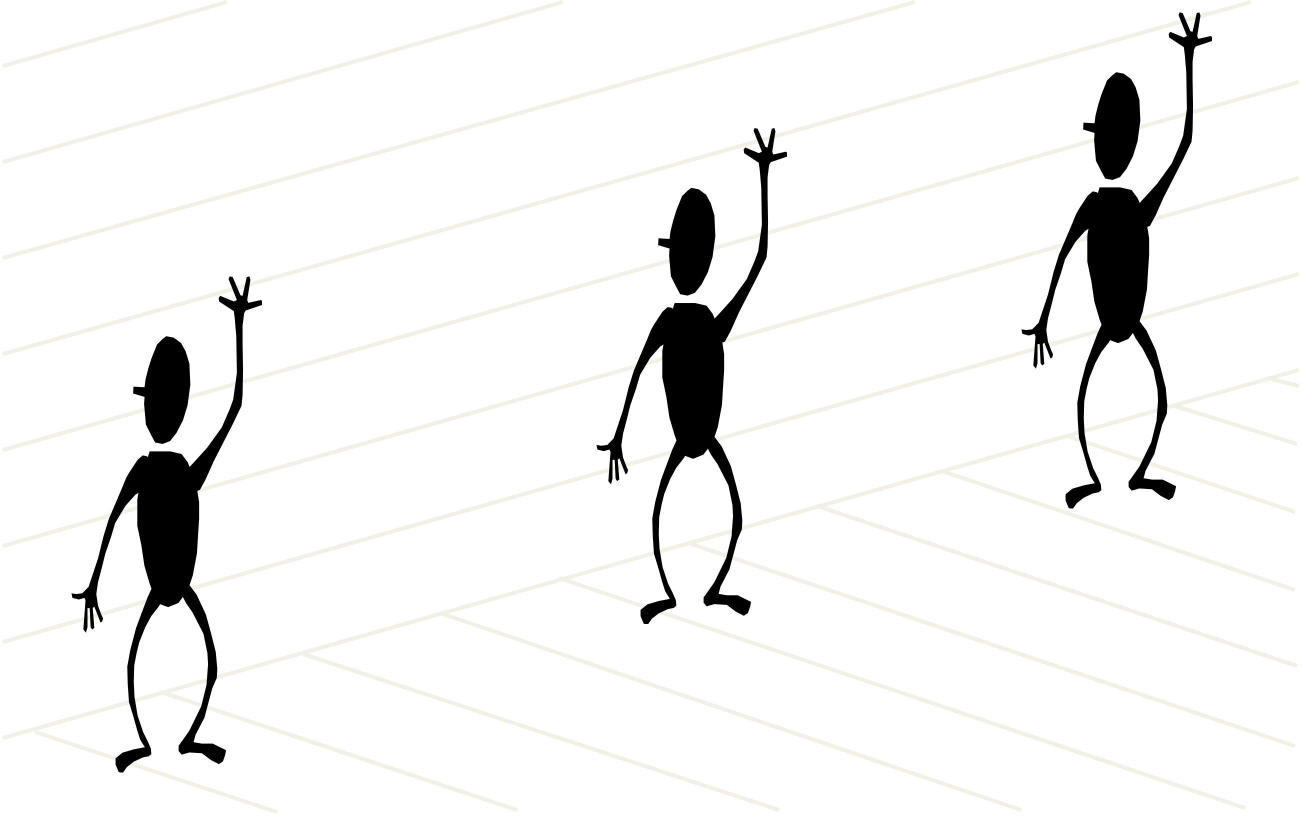
Fun with vanishing points





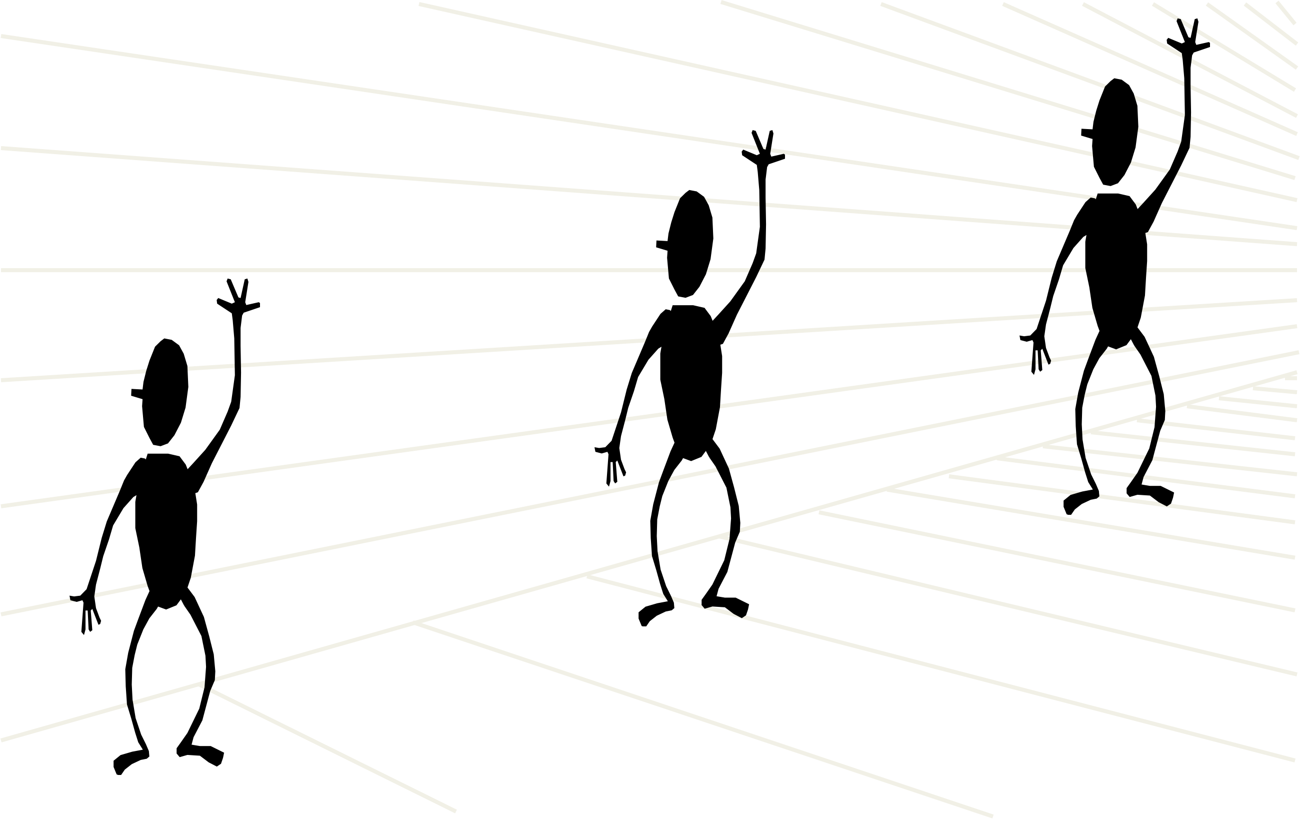
Source: S. Seitz slides.

Perspective cues



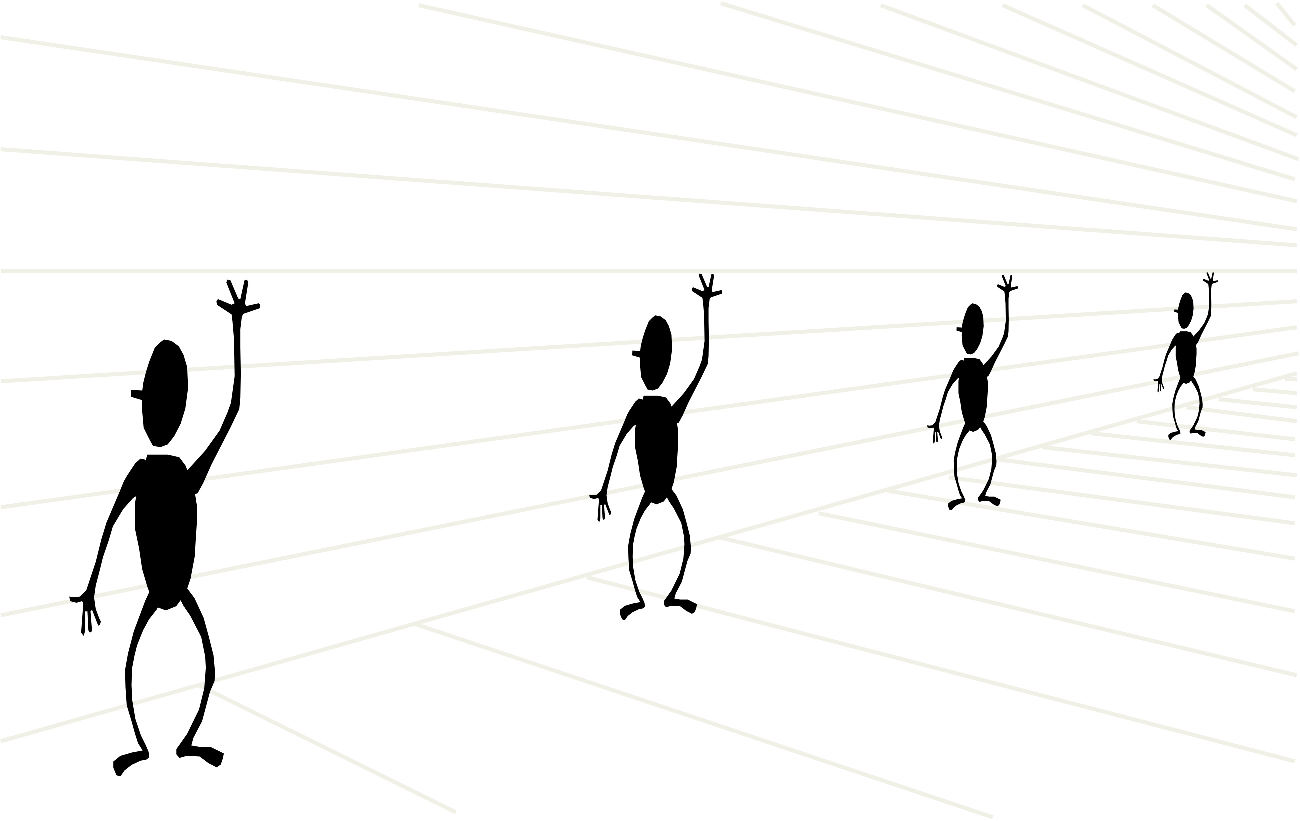
Source: S. Seitz slides.

Perspective cues



Source: S. Seitz slides.

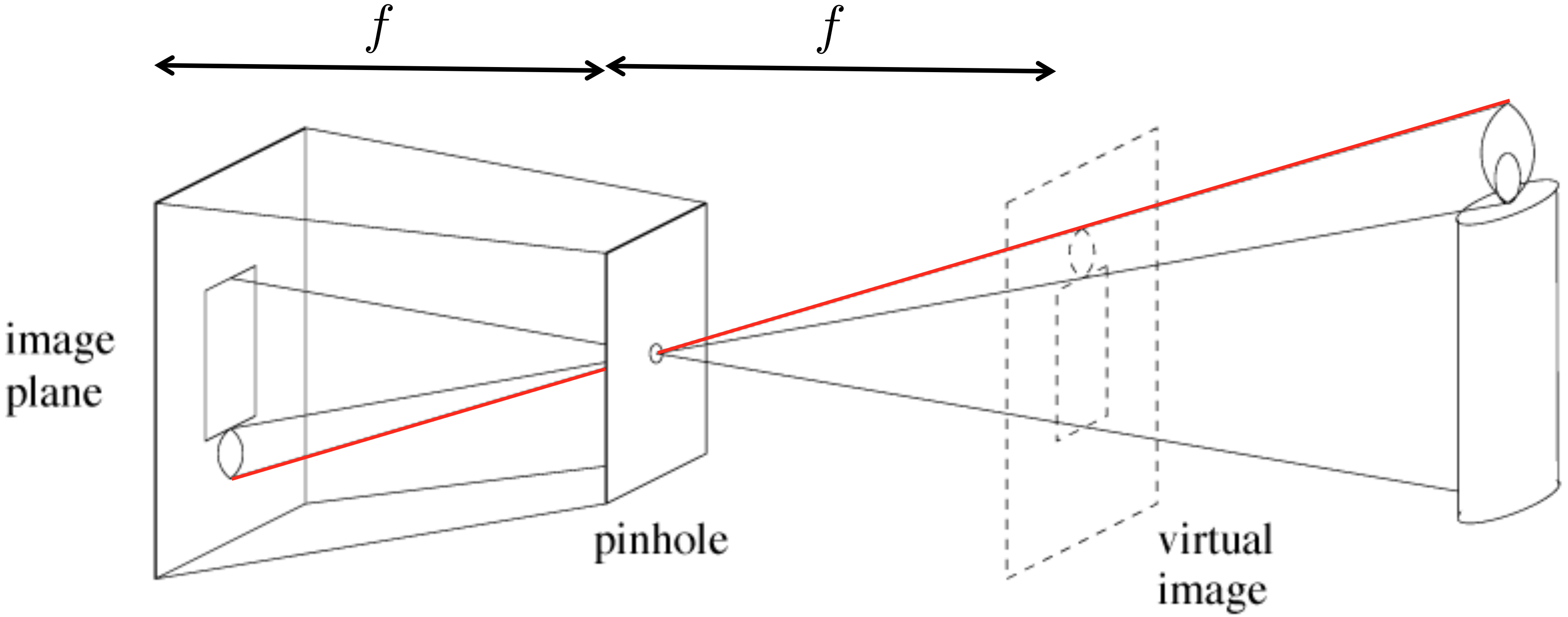
Perspective cues



Source: S. Seitz slides.

Pinhole, or *Central*, Perspective

- It is common to draw the image plane in front of the focal point.



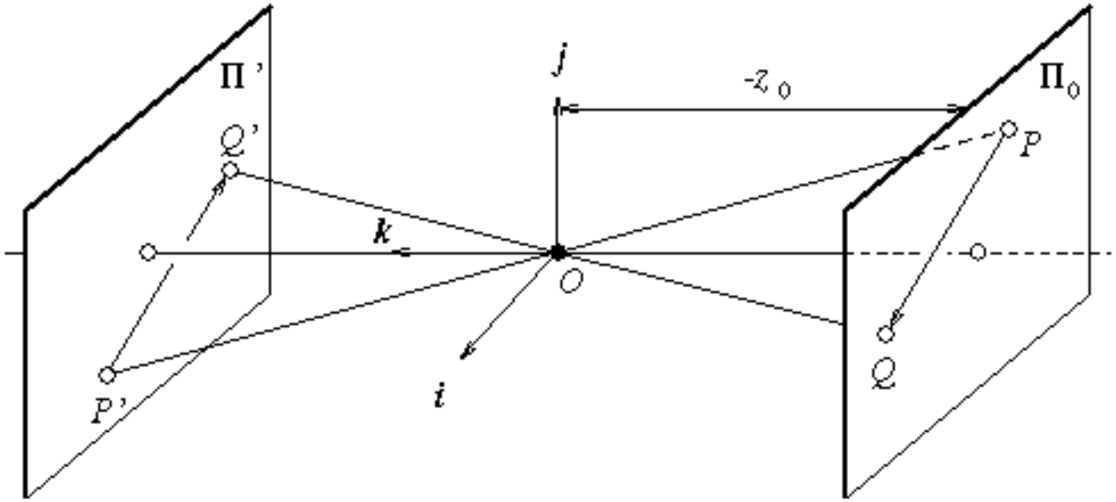
Source: D. Forsyth, S. Savarese slides.

Weak Perspective

- A coarser approximation to image formation is called **weak perspective**, or scaled orthography.
- Consider a fronto-parallel plane Π_0 defined by $Z = Z_0$.
- Rewrite projection equations for any point in Π_0

$$X' = -mX \qquad Y' = -mY$$

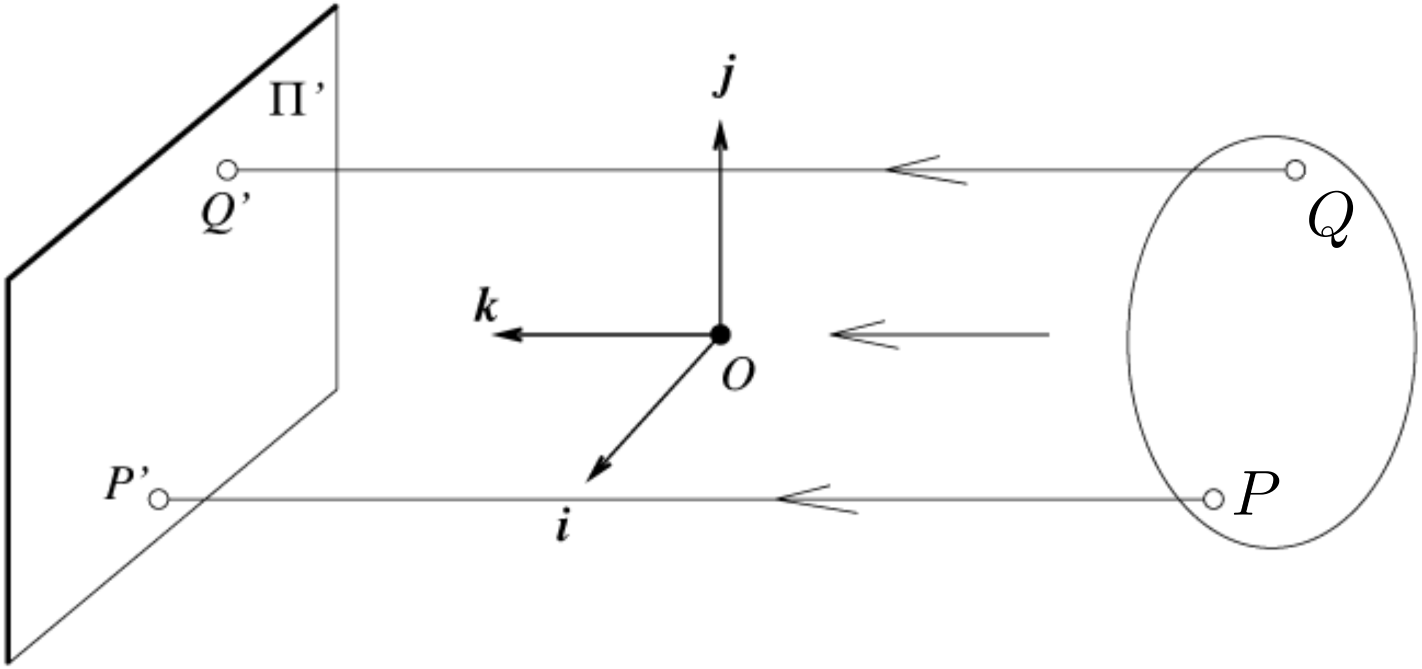
$$X' = -\frac{f}{Z_0}X \qquad Y' = -\frac{f}{Z_0}Y$$



Source: D. Forsyth, S. Savarese slides.

Orthographic Projections

- Further, when the camera will be at a fixed distance from the scene, we can further normalize the coordinates.
  - Make $m = -1$
  - Then $X' = X$ and $Y' = Y$



- Almost never an acceptable model for image formation.

Source: D. Forsyth, S. Savarese slides.

Projection Matrices

- Can formulate the perspective projections as matrix operations with homogeneous coordinates.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{1}{d} & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ -\frac{z}{f} \end{bmatrix} \Longrightarrow \begin{bmatrix} -f\frac{X}{Z} & -f\frac{Y}{Z} \end{bmatrix}^{\mathsf{T}}$$

- Why are homogeneous coordinates necessary here?
- Can also formulate as a 4x4 projection.

Projection Matrices

- How does scaling affect the projection?

$$s\begin{bmatrix}X & Y & Z & 1\end{bmatrix}^{\mathsf{T}}$$

$$\begin{bmatrix} s & 0 & 0 & 0 \\ 0 & s & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} sX \\ sY \\ -z \end{bmatrix} \Longrightarrow \begin{bmatrix} -s\frac{X}{Z} & -s\frac{Y}{Z} \end{bmatrix}^{\mathsf{T}}$$

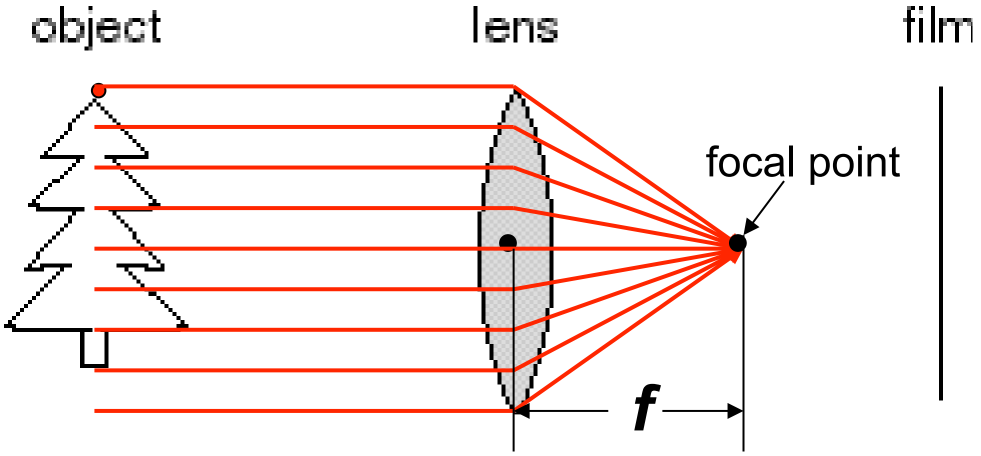
Role of aperture size

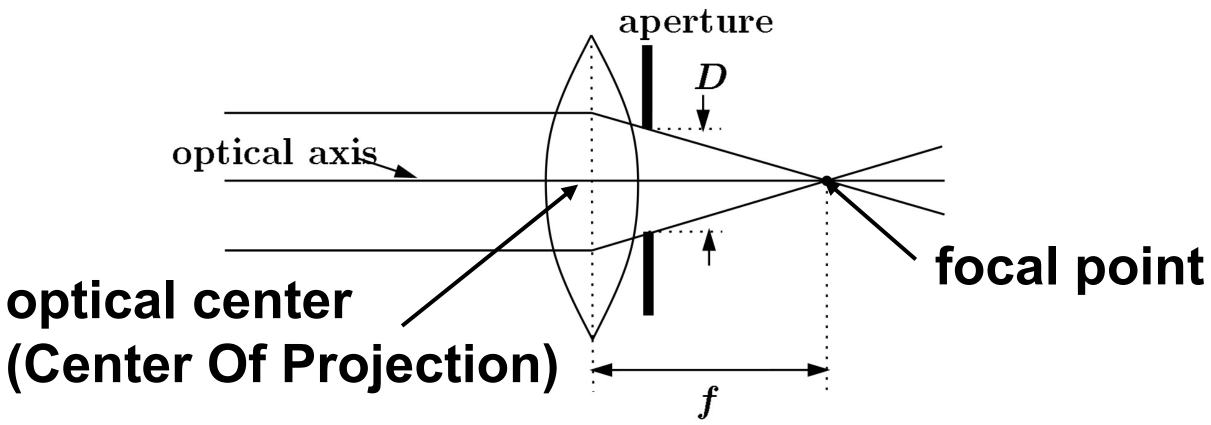
- When aperture is big, what happens?
- Why not make the aperture as small as possible?
  - Not enough light gets through.
  - Diffraction.

Source: S. Seitz, S. Savarese slides.

Adding a lens

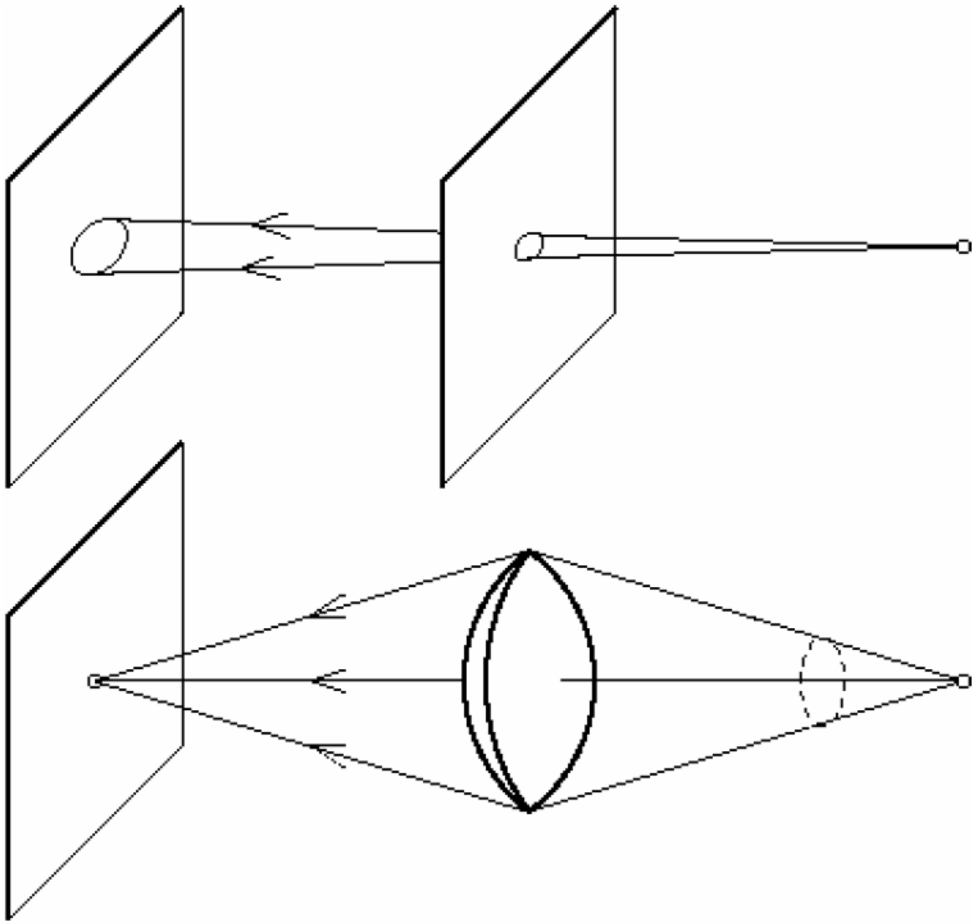
- A lens focuses the light onto the film/CCD.
- Rays passing through the center are not deviated.
- All parallel rays converge to one point on a plane located at the **focal length** *f.*





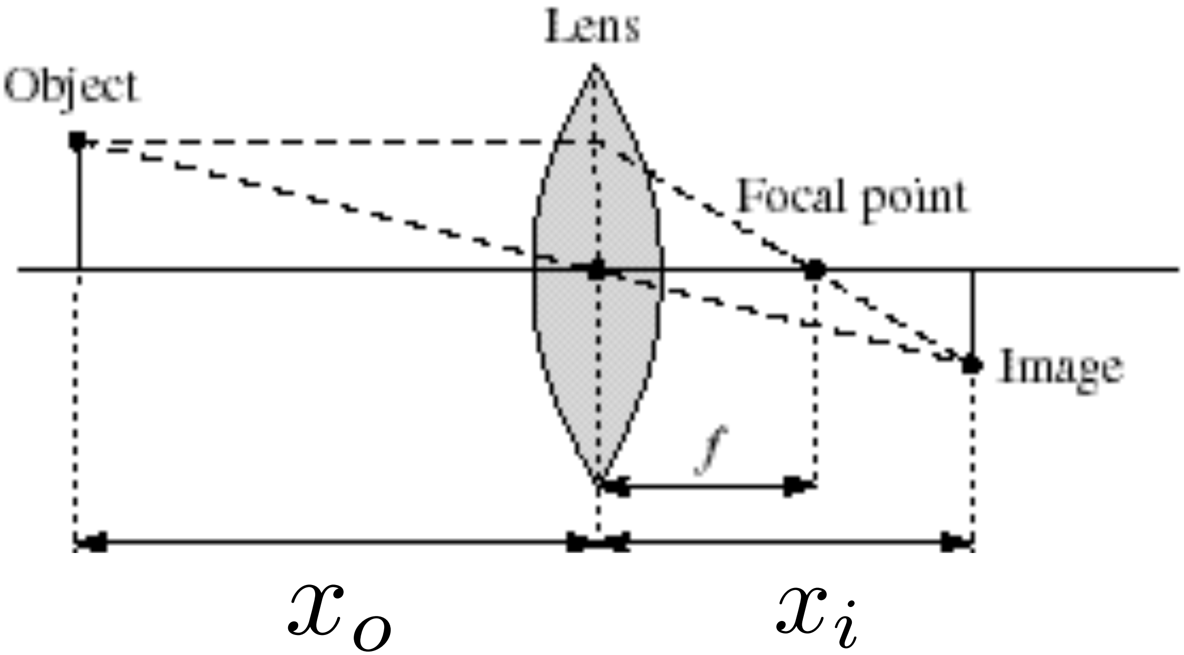
Source: GD Hager and S Seitz slides.

Pinhole vs. lens



Source: GD Hager and S Seitz slides.

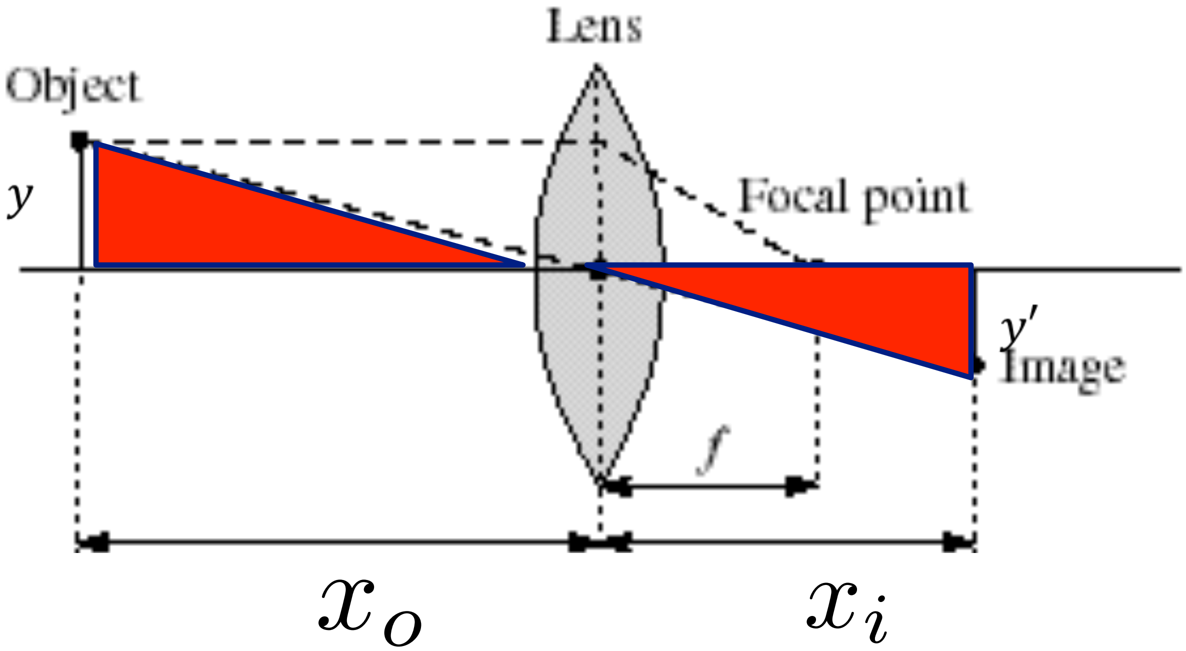
Thin lens equation



- How to relate distance of object from optical center (x_o) to the distance at which it will be in focus (x_i), given focal length f ?

Source: GD Hager and S Seitz slides.

Thin lens equation

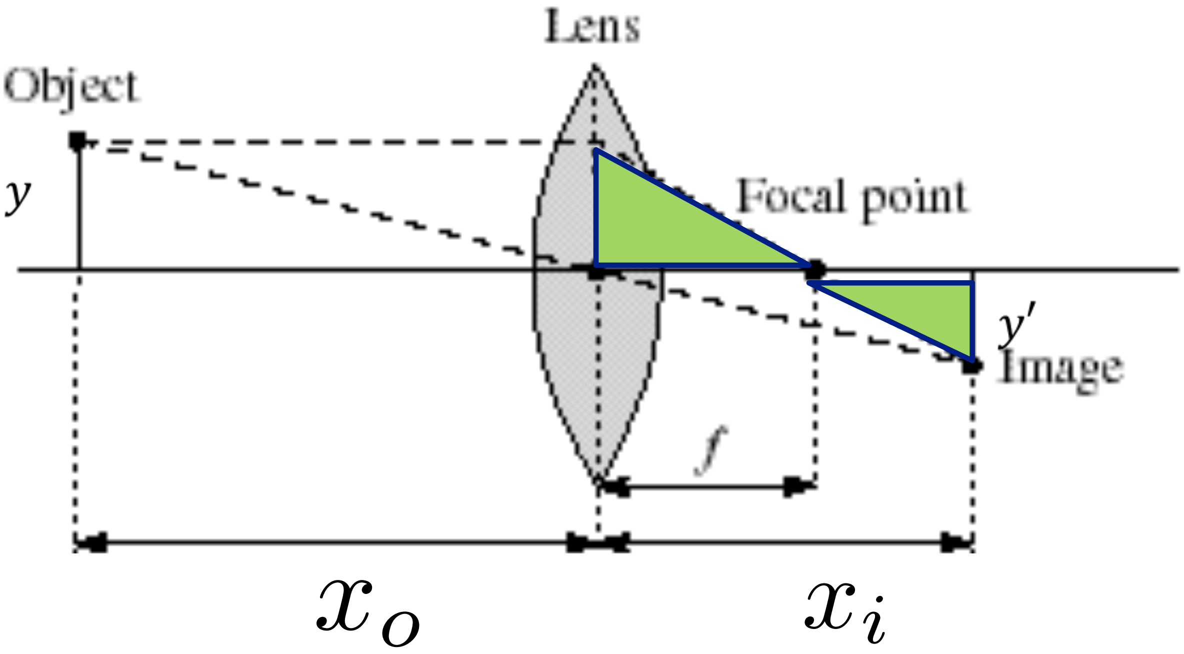


$$\frac{y'}{y} = \frac{x_i}{x_o}$$

- How to relate distance of object from optical center (x_o) to the distance at which it will be in focus (x_i), given focal length f ?

Source: GD Hager and S Seitz slides.

Thin lens equation



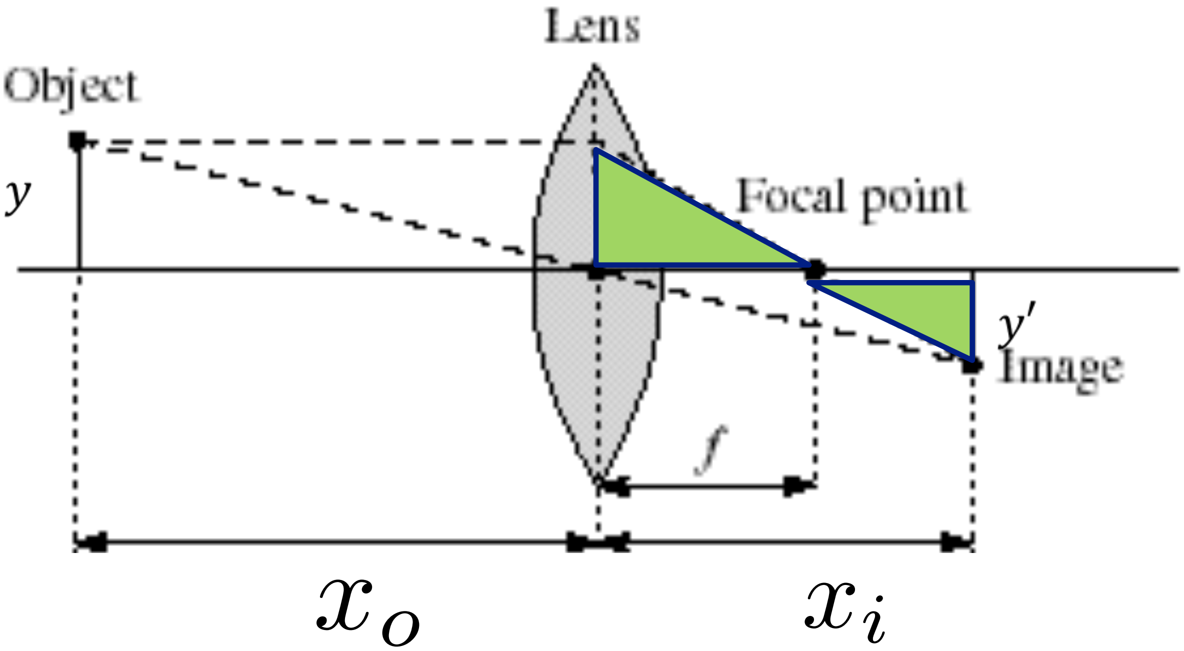
$$\frac{y'}{y} = \frac{x_i}{x_o}$$

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

- How to relate distance of object from optical center (x_o) to the distance at which it will be in focus (x_i), given focal length f ?

Source: GD Hager and S Seitz slides.

Thin lens equation



$$\frac{y'}{y} = \frac{x_i}{x_o}$$

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

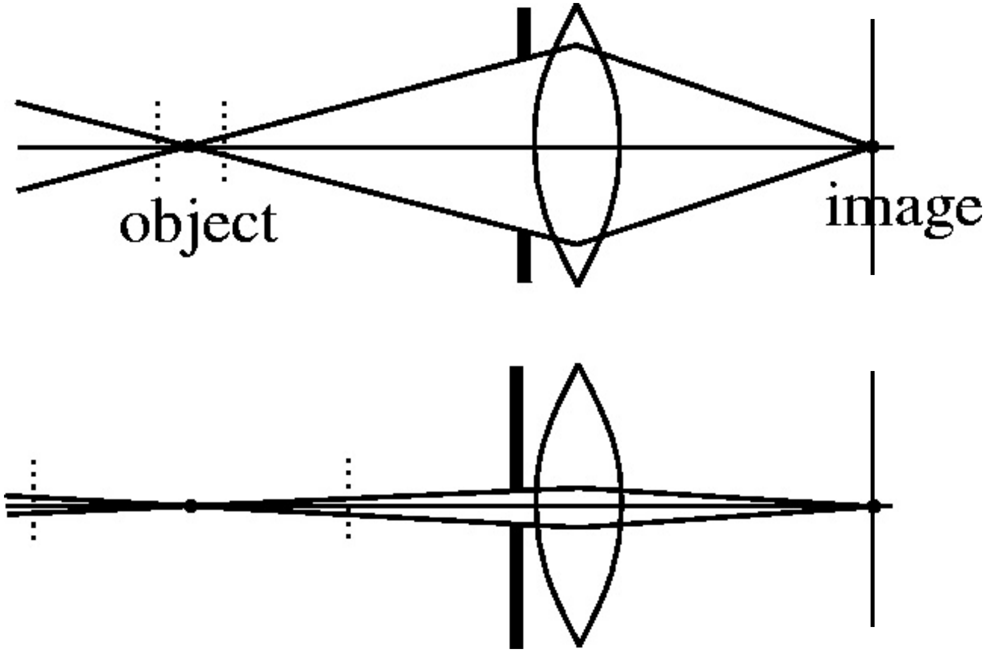
$$\downarrow$$

- Any object point satisfying this equation is in focus

$$\frac{1}{f} = \frac{1}{x_o} + \frac{1}{x_i}$$

Source: GD Hager and S Seitz slides.

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

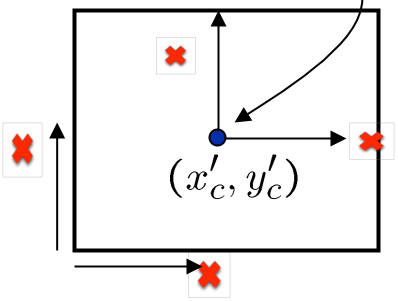
Flower images from Wikipedia http://en.wikipedia.org/wiki/Depth_of_field

Source: S Seitz slides.

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

$$\mathbf{x} = \begin{bmatrix} sx \\ sy \\ s \end{bmatrix} = \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \mathbf{\Pi X}$$

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

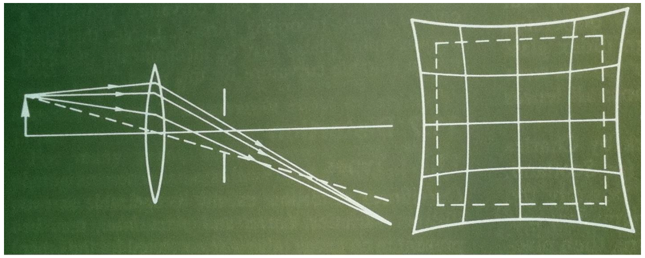
$$\mathbf{\Pi} = \underbrace{\begin{bmatrix} -fs_x & 0 & x'_c \\ 0 & -fs_y & y'_c \\ 0 & 0 & 1 \end{bmatrix}}_{\text{intrinsics}} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\text{projection}} \underbrace{\begin{bmatrix} \mathbf{R}_{3x3} & \mathbf{0}_{3x1} \\ \mathbf{0}_{1x3} & 1 \end{bmatrix}}_{\text{rotation}} \underbrace{\begin{bmatrix} \overset{\text{identity matrix}}{\mathbf{I}}_{3x3} & \mathbf{T}_{3x1} \\ \mathbf{0}_{1x3} & 1 \end{bmatrix}}_{\text{translation}}$$

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

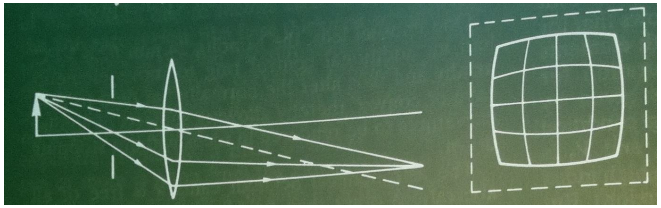
Source: S Seitz slides.

Radial Distortion

- Pin Cushion

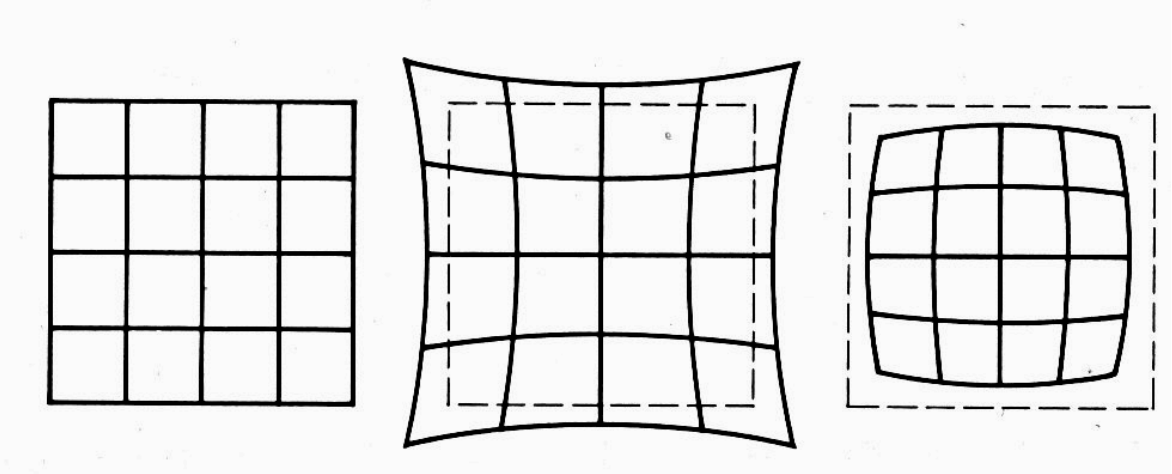


- Barrel / Fisheye



Source: S Savarese slides.

Radial 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

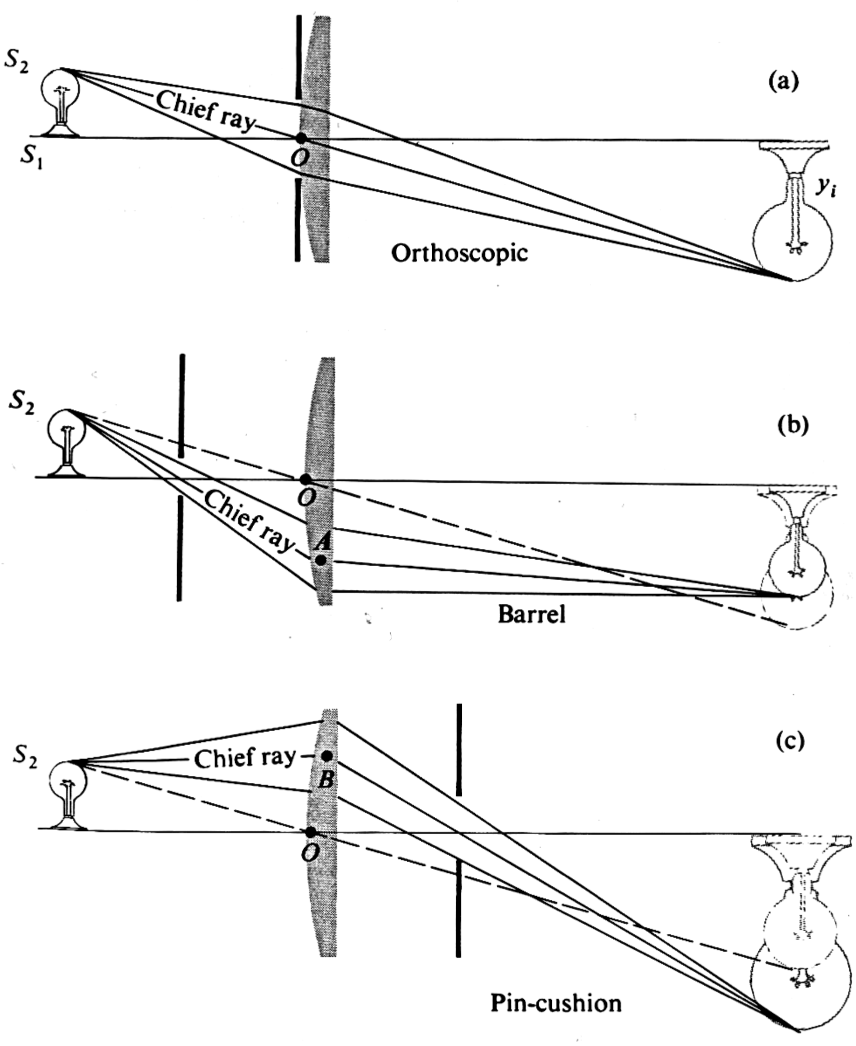
Source: S Seitz slides.

Correcting radial distortion

from Helmut Dersch

Source: S Seitz slides.

Distortion



Source: S Seitz slides.

Modeling distortion

Pro$(\widehat{x}, \widehat{y}, \widehat{z})$
to "normalized"
image coordinates

$$x'_n = \widehat{x}/\widehat{z}$$
$$y'_n = \widehat{y}/\widehat{z}$$

Apply radial distortion

$$r^2 = {x'_n}^2 + {y'_n}^2$$
$$x'_d = x'_n(1 + \kappa_1 r^2 + \kappa_2 r^4)$$
$$y'_d = y'_n(1 + \kappa_1 r^2 + \kappa_2 r^4)$$

Apply focal length
translate image center

$$x' = f x'_d + x_c$$
$$y' = f y'_d + y_c$$

- To model lens distortion
  - – Use above projection operation instead of standard projection matrix multiplication

Source: S Seitz slides.

Other types of projection

- Lots of intriguing variants…
- (I’ll just mention a few fun ones)

Source: S Seitz slides.

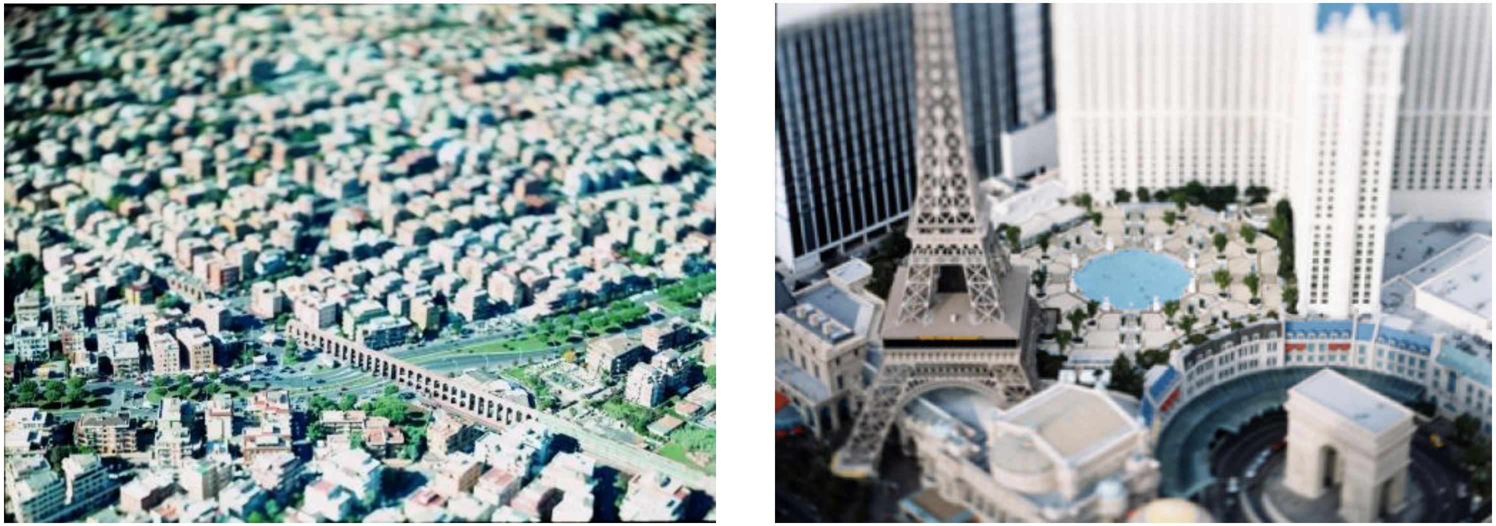
360 degree field of view…

- Basic approach
  - Take a photo of a parabolic mirror with an orthographic lens (Nayar)
  - Or buy one a lens from a variety of omnicam manufacturers…
    - See http://www.cis.upenn.edu/~kostas/omni.html

Source: S Seitz slides.

Tilt-shift

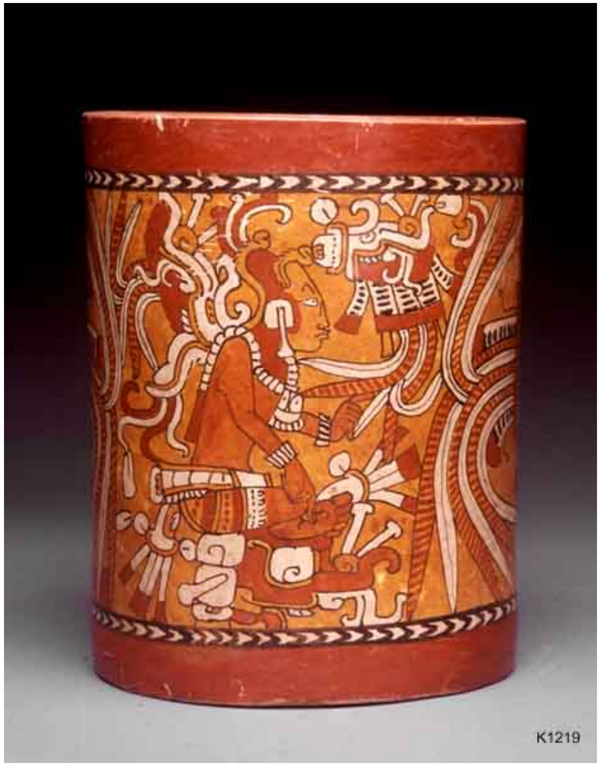
http://www.northlight-images.co.uk/article_pages/tilt_and_shift_ts-e.html



Titlt-shift images from Olivo Barbieri
and Photoshop imitations

Source: S Seitz slides.

Rotating sensor (or object)

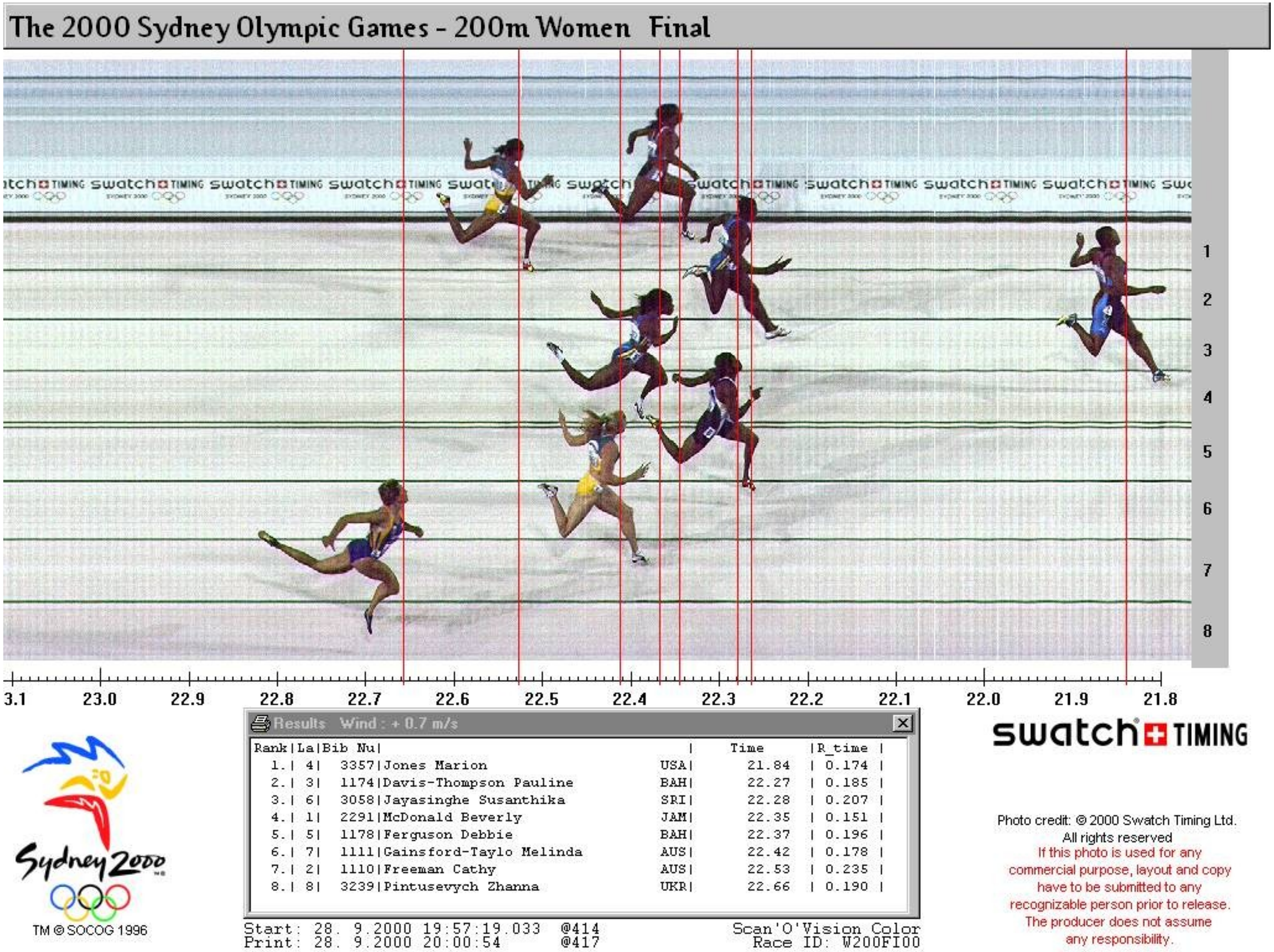


Rollout Photographs © Justin Kerr
http://research.famsi.org/kerrmaya.html

Also known as “cyclographs”, “peripheral images”

Source: S Seitz slides.

Photofinish



The 2000 Sydney Olympic Games - 200m Women Final

```
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                  Scan'O'Vision Color
Print:  28. 9.2000 20:00:54       @417                    Race ID: W200FI00
```

Photo credit: © 2000 Swatch Timing Ltd.

Source: S Seitz slides.