Why study Computer Vision?

- Images and movies are everywhere
- Fast-growing collection of useful applications
  - building representations of the 3D world from pictures
  - automated surveillance (who's doing what)
  - movie post-processing
  - face finding
- Various deep and attractive scientific mysteries
  - how does object recognition work?
- Greater understanding of human vision

Properties of Vision

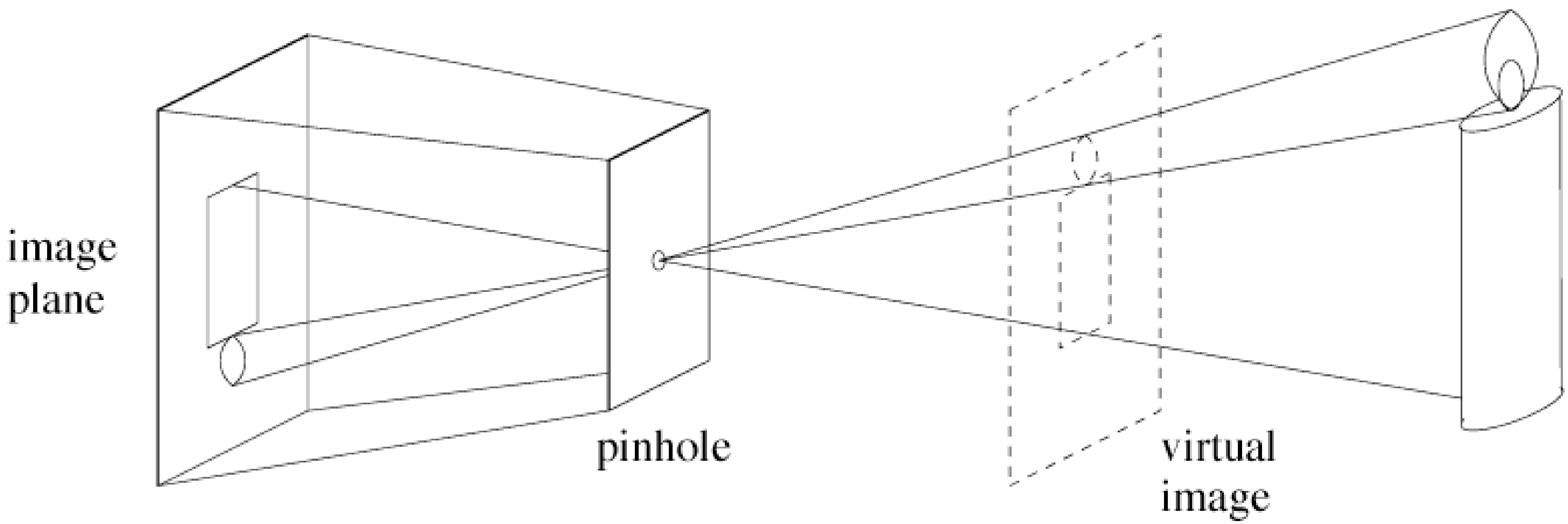
- People draw distinctions between what is seen
  - "Object recognition"
  - This could mean "is this a fish or a bicycle?"
  - It could mean "is this George Washington?"
  - It could mean "is this poisonous or not?"
  - It could mean "is this slippery or not?"
  - It could mean "will this support my weight?"
  - Great mystery
    - How to build programs that can draw useful distinctions based on image properties.

Cameras

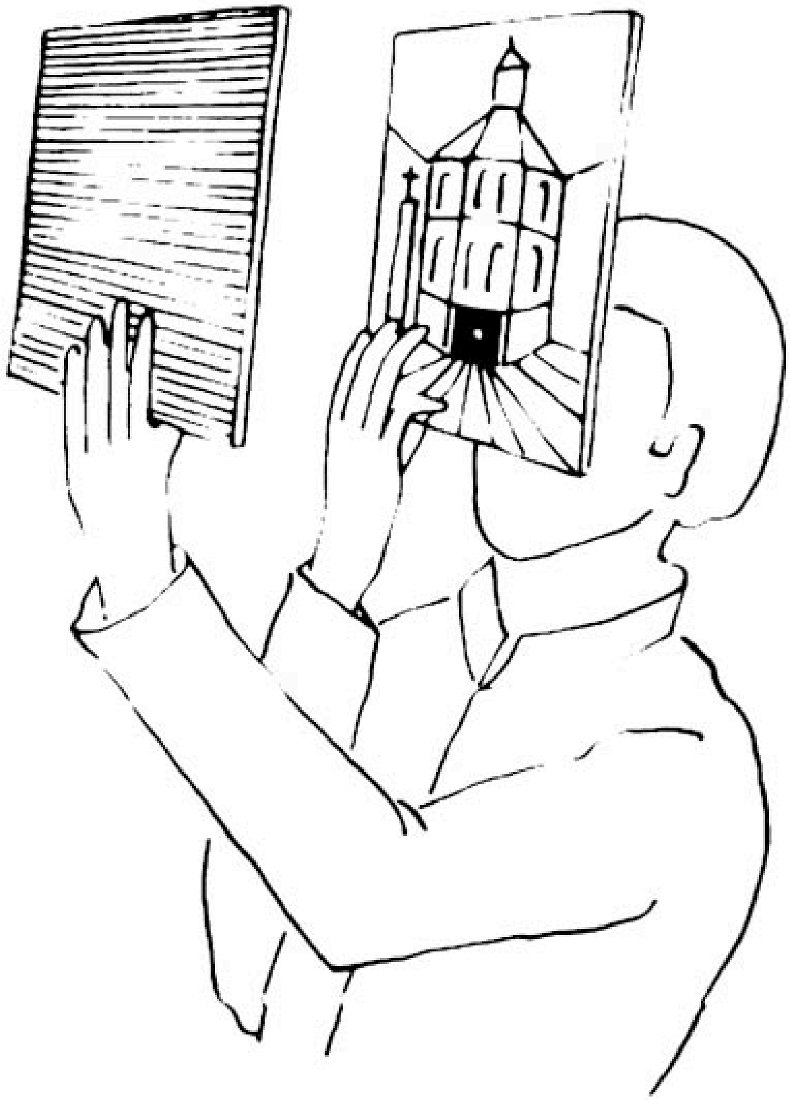
- First photograph due to Niepce
- First on record shown in the book - 1822
- Basic abstraction is the pinhole camera
  - lenses required to ensure image is not too dark
  - various other abstractions can be applied

Pinhole cameras

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



Vasari's Portrait of Filippo Brunelleschi



Edgerton's depiction of the reverse-reflection 'peepshow' of the Baptistry with which Brunelleschi demonstrated perspective in ~1420

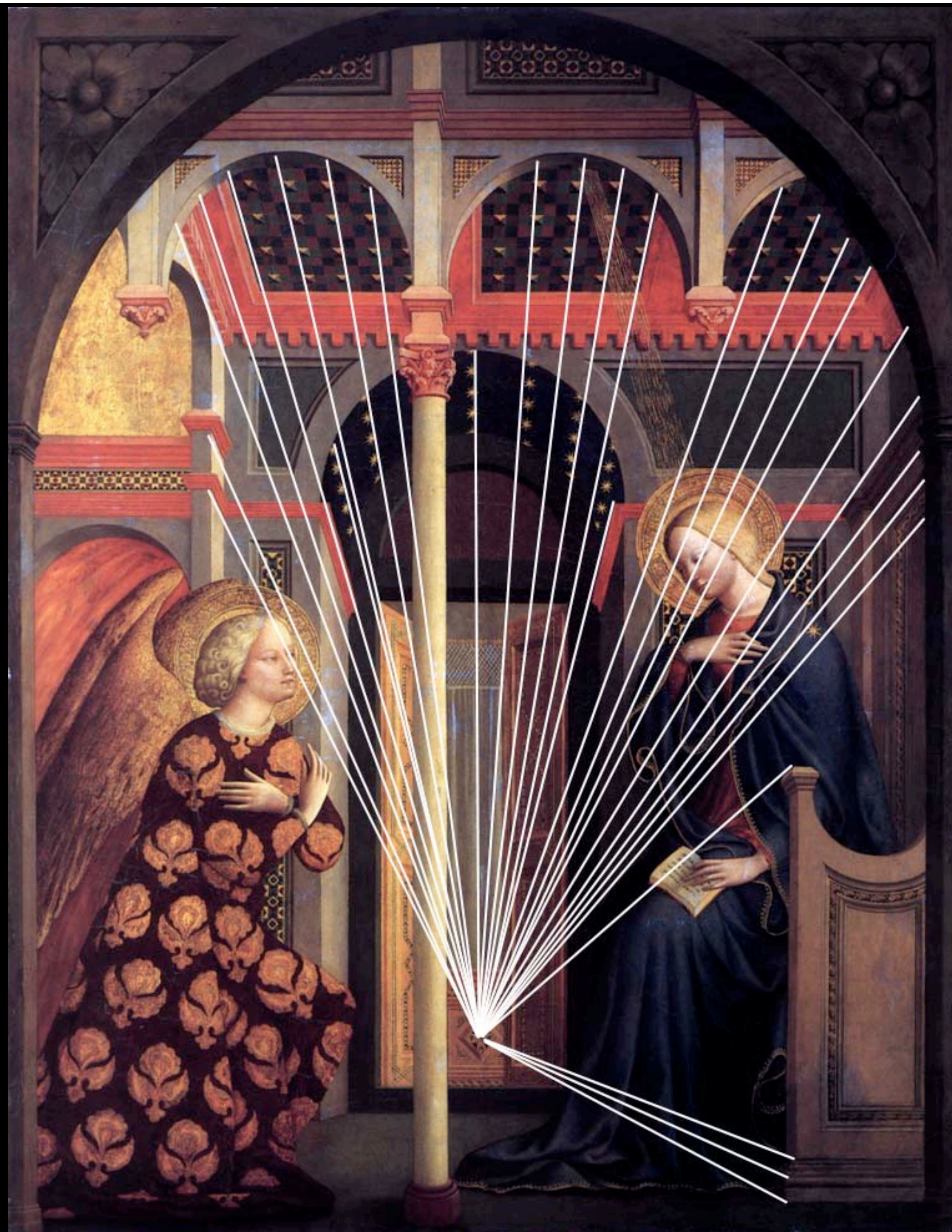
The 'Goldman' Annunciation

by Masolino

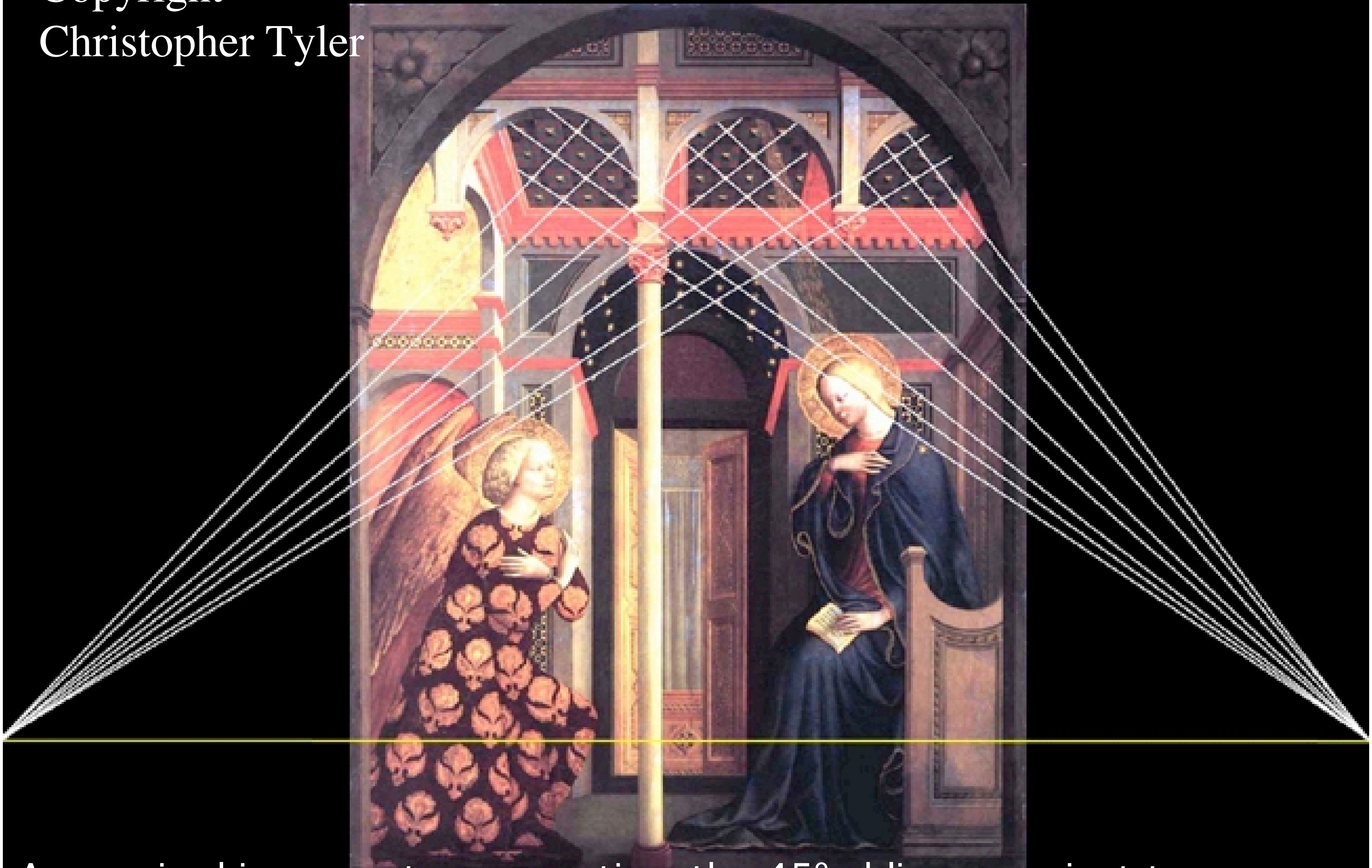
(1424)

This complex architecture is estimated to have predated (and may have inspired) Masaccio's 'Trinity' by four years.

The accurate perspective extends from the arch in the background to
to the coffers of the portico . . .

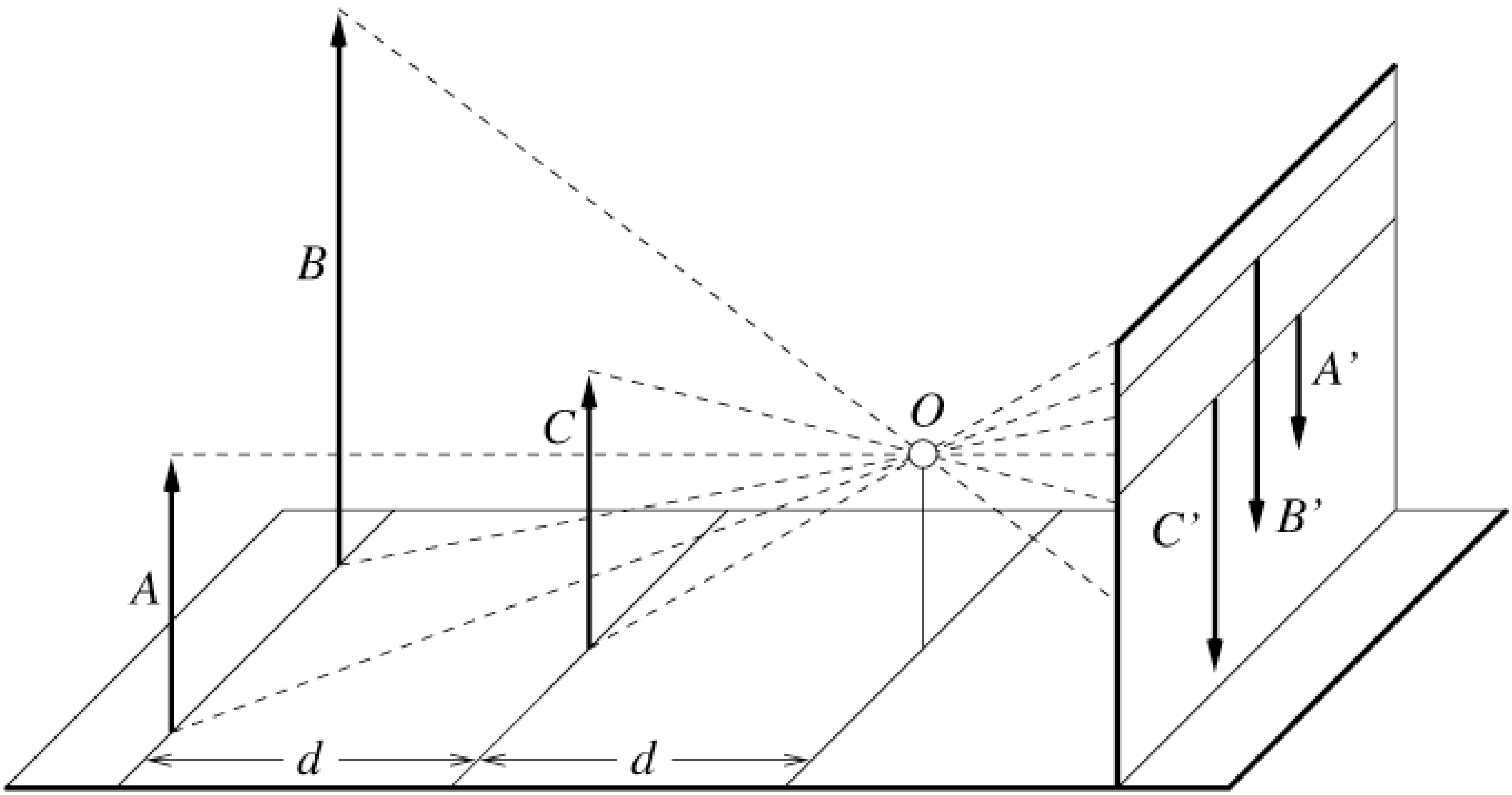


. . . and even details of the Virgin's throne in the right foreground.



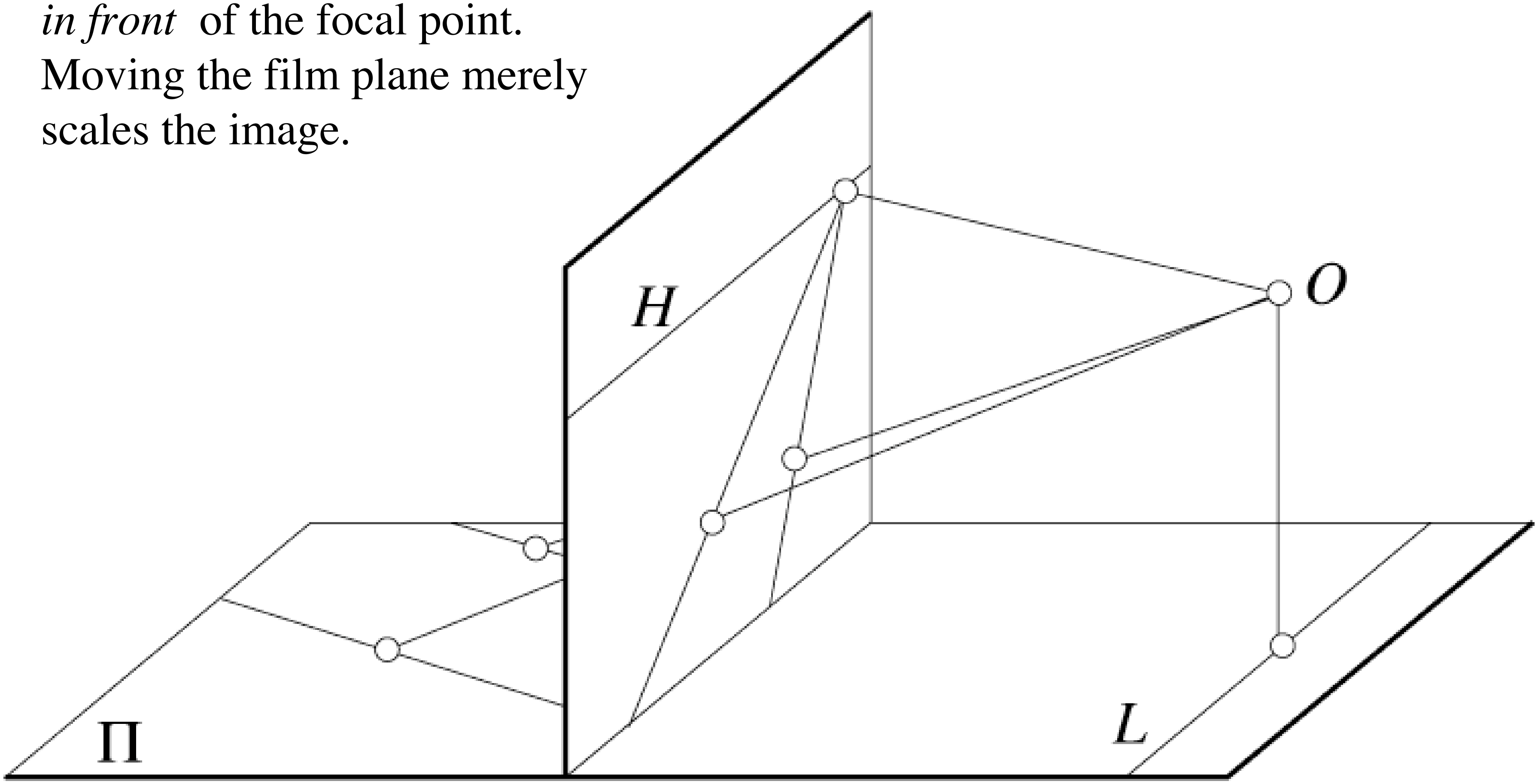
As required in accurate perspective, the 45º obliques project to distance points at the same height as the central vanishing point.

Distant objects are smaller



Parallel lines meet

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

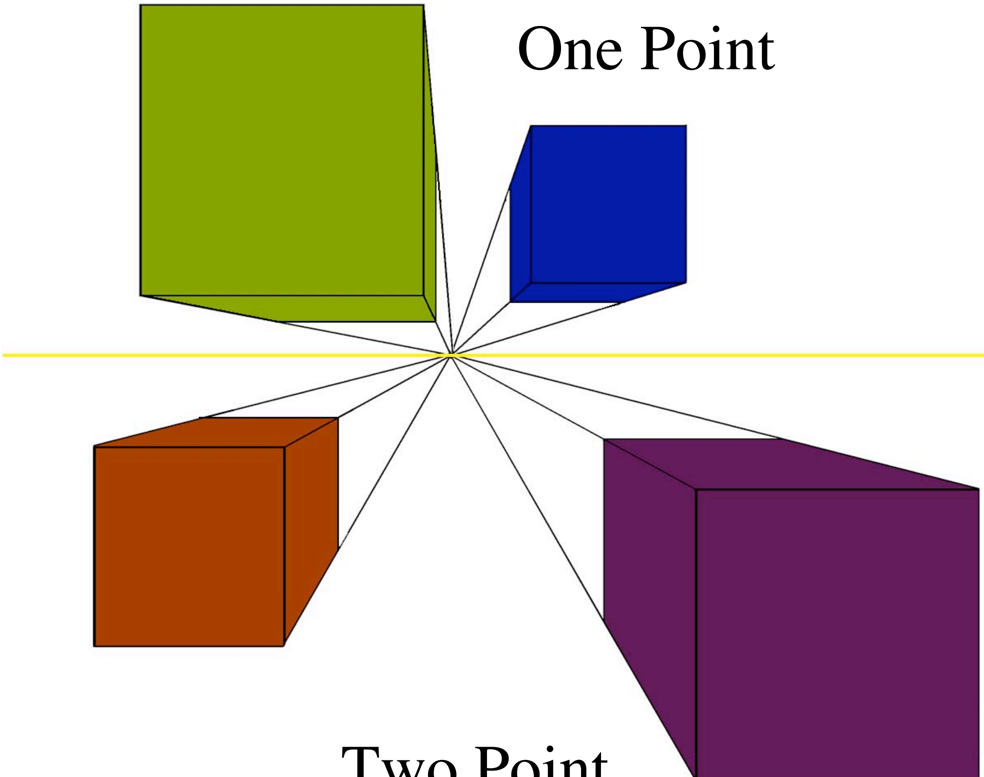


Vanishing points

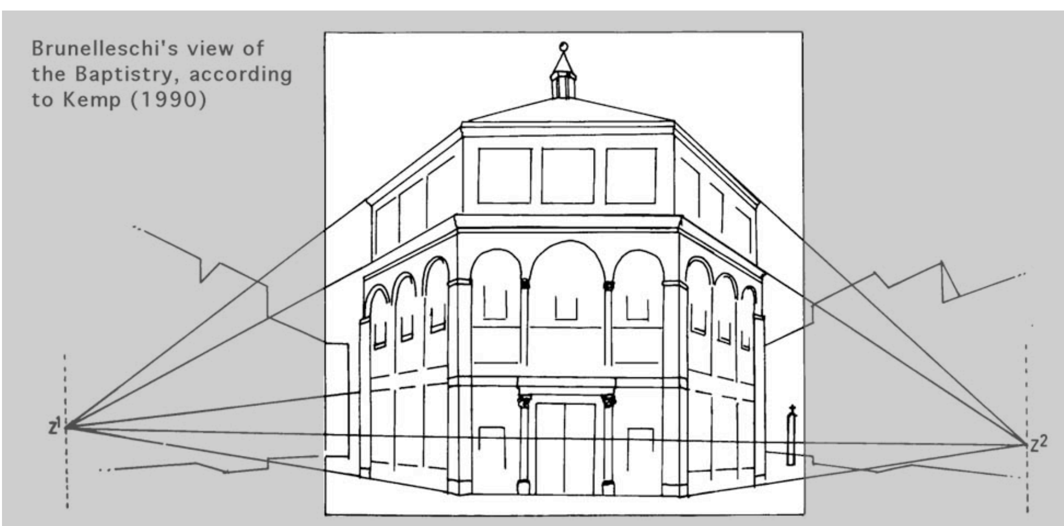
- each set of parallel lines (=direction) meets at a different point
  - The *vanishing point* for this direction
- Sets of parallel lines on the same plane lead to *collinear* vanishing points.
  - The line is called the *horizon* for that plane
- Good ways to spot faked images
  - scale and perspective don't work
  - vanishing points behave badly
  - supermarket tabloids are a great source.

Faked Photos

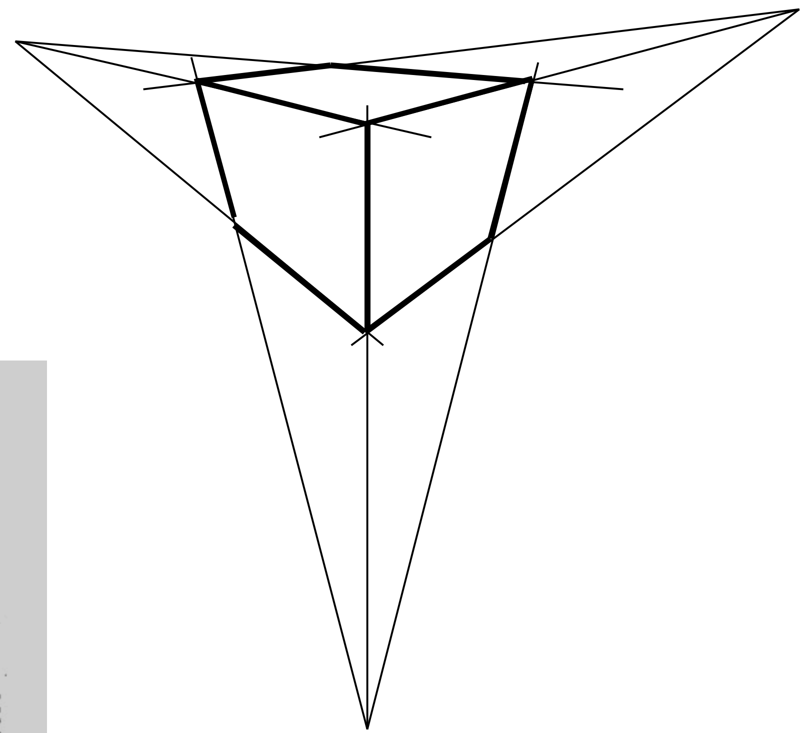
One Point



Two Point

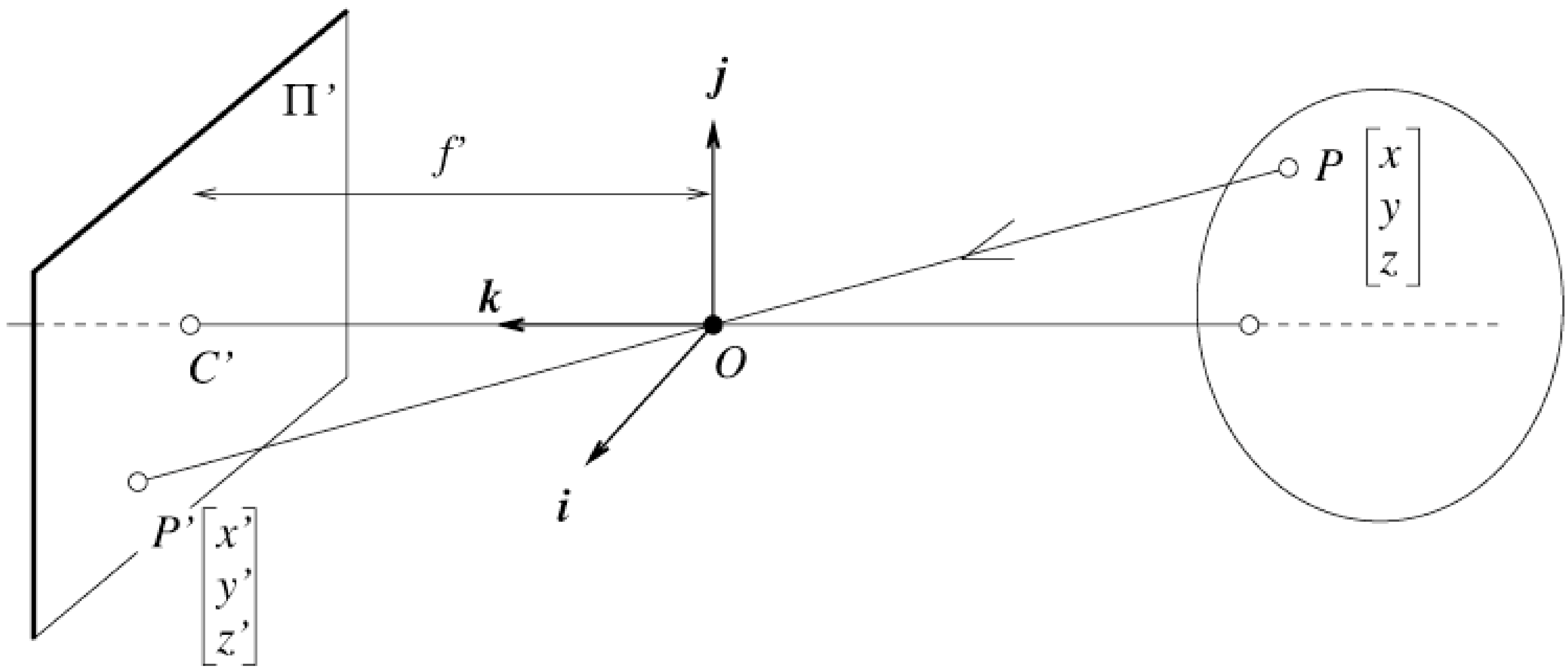


Three Point

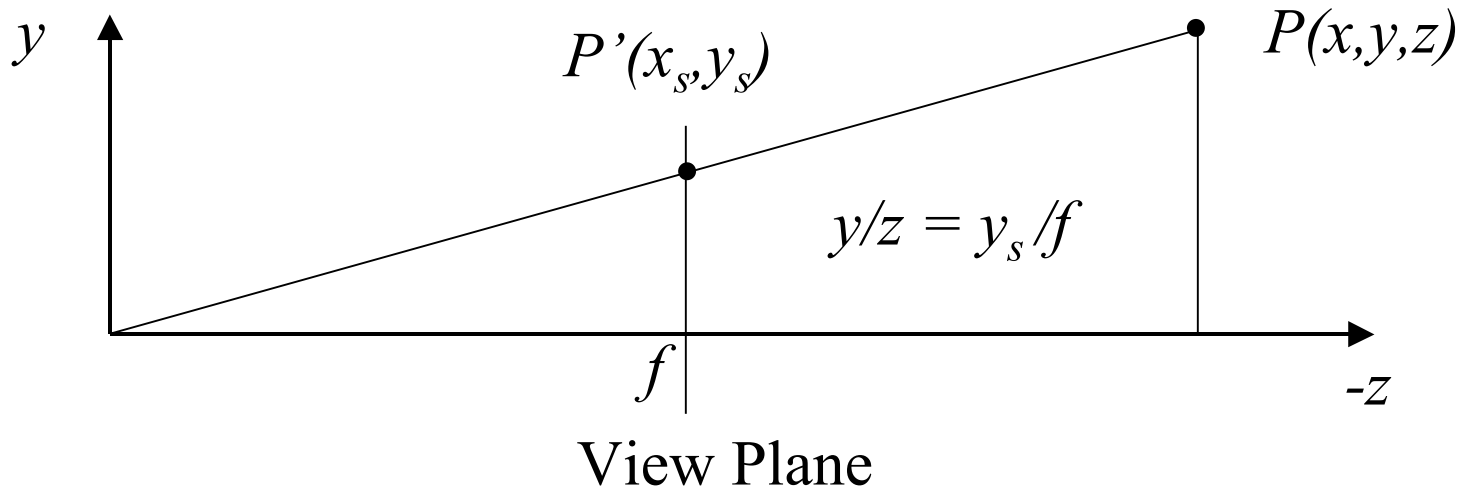


The equation of projection

- If you know *P(x,y,z)* and *f'*, what is *P'?*
  - Where does a point in view space end up on the screen?



The equation of projection



- Cartesian coordinates:
  - We have, by similar triangles, that (x, y, z) -> (f x/z, f y/z, -f)

–Ignore the third coordinate, and get

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

Homogenous coordinates

- Add an extra coordinate and use an equivalence relation
- for 2D
  - equivalence relation k*(X,Y,Z) is the same as (X,Y,Z)
- for 3D
  - equivalence relation k*(X,Y,Z,T) is the same as (X,Y,Z,T)

- Basic notion
  - Possible to represent points "at infinity"
    - Where parallel lines intersect
    - Where parallel planes intersect
  - Possible to write the action of a perspective camera as a matrix

The camera matrix

- Turn previous expression into HC's
  - HC's for 3D point are (X,Y,Z,T)
  - HC's for point in image are (U,V,W)

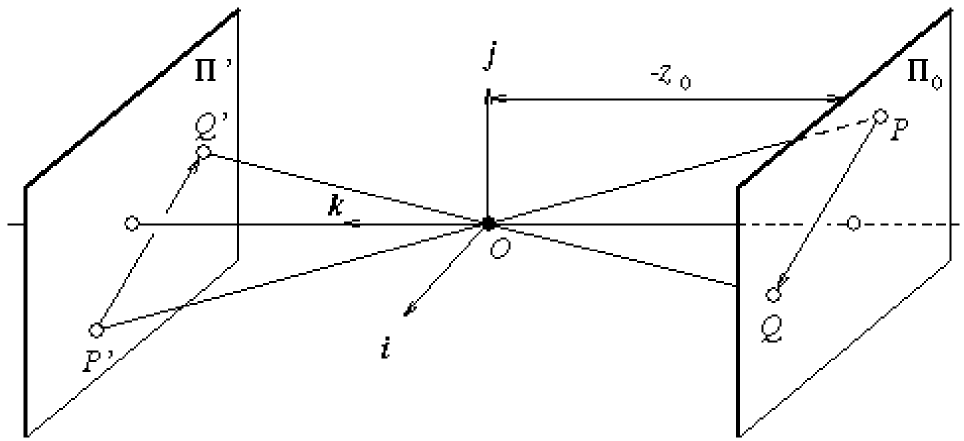
$$\begin{bmatrix} U \\ V \\ W \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ T \end{bmatrix}$$

Relation between HC's and image coordinates

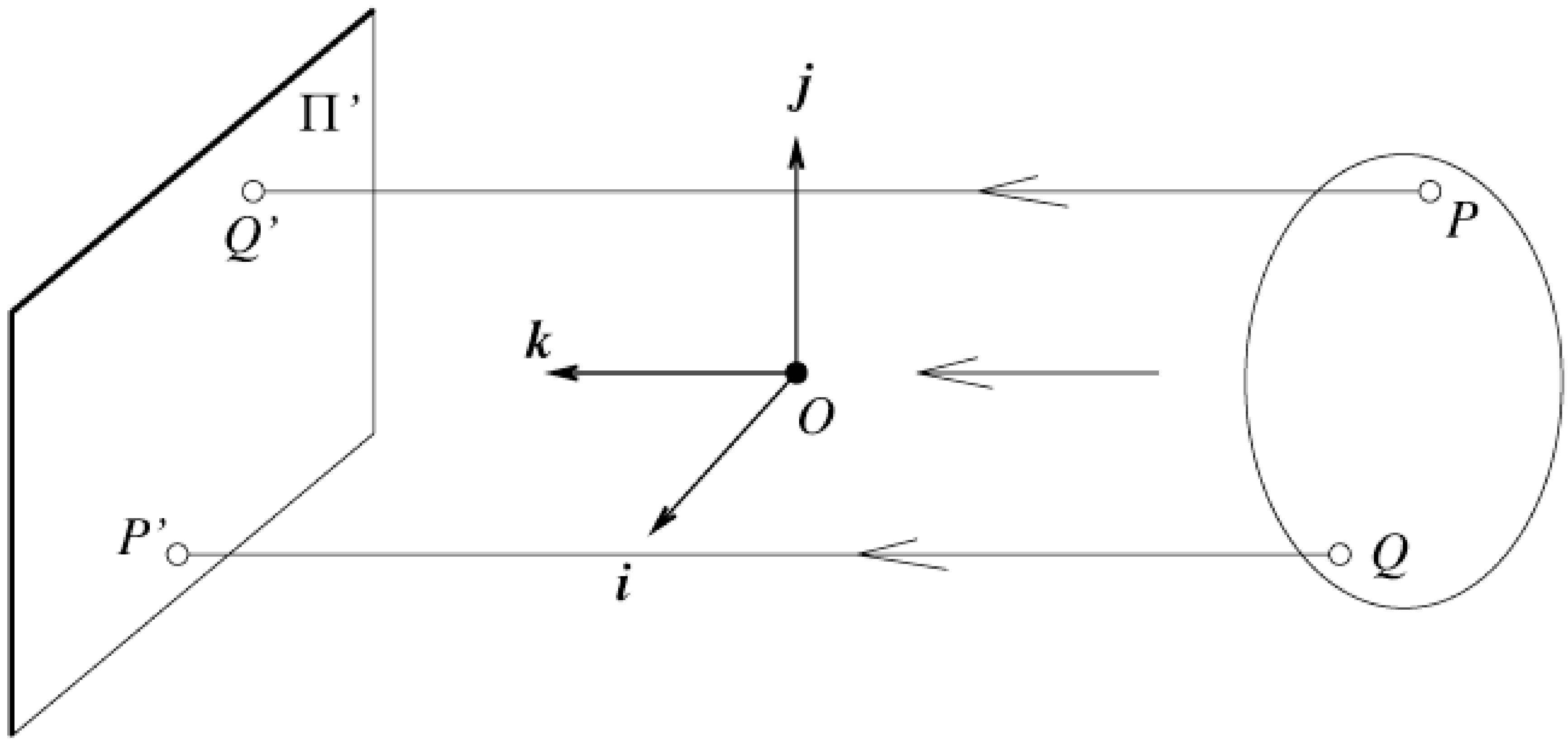
$$x_s = U/W$$
$$y_s = U/W$$

Weak perspective

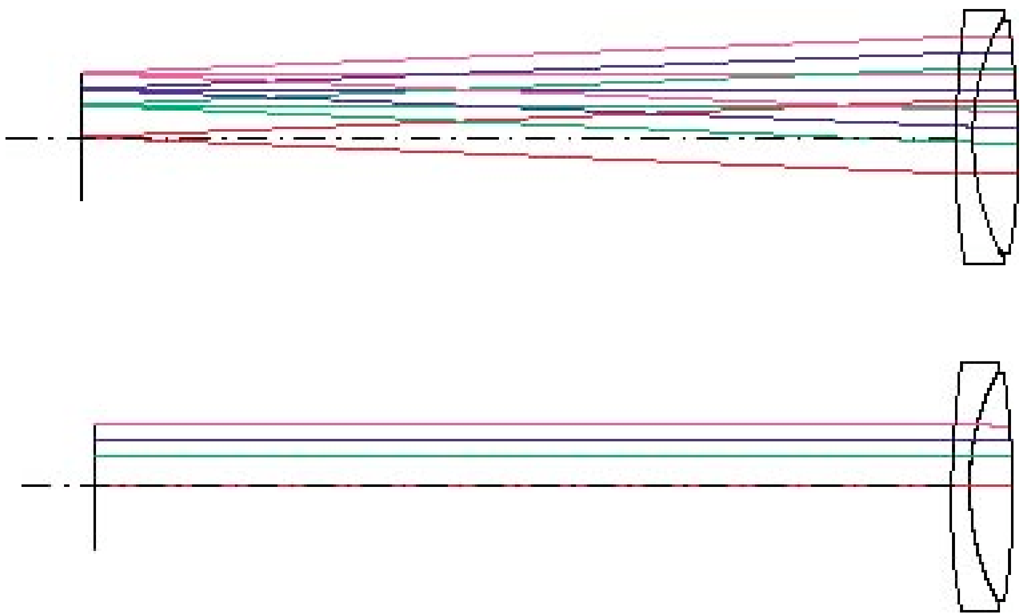
- Issue
  - perspective effects, but not over the scale of individual objects
  - collect points into a group at about the same depth, then divide each point by the depth of its group
  - Adv: easy
  - Disadv: wrong

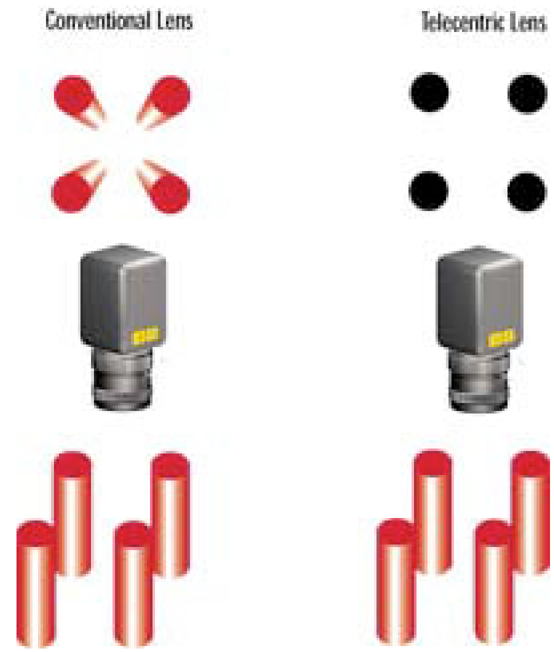


Orthographic projection



Orthographic via Telecentric Lens

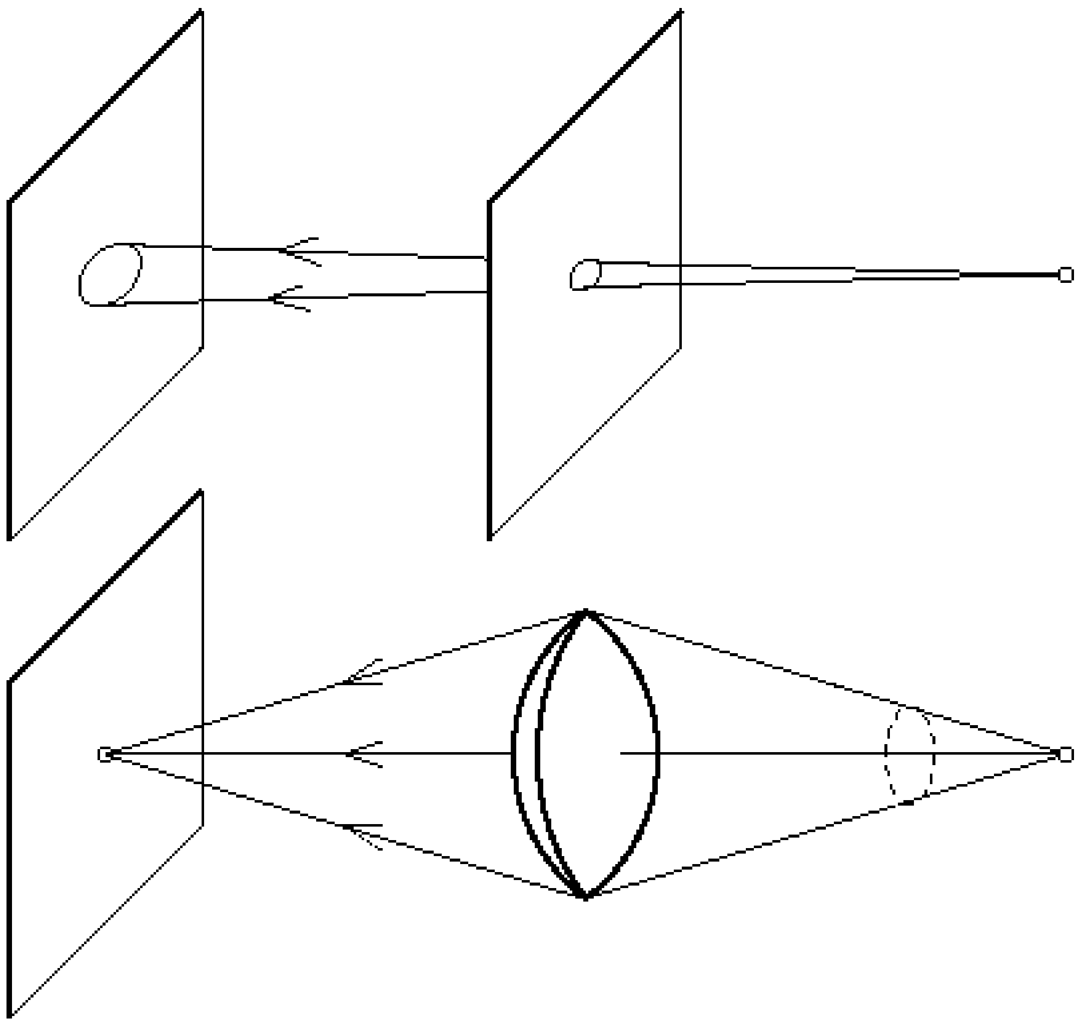




The projection matrix for orthographic projection

$$\begin{bmatrix} U \\ V \\ W \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ T \end{bmatrix}$$

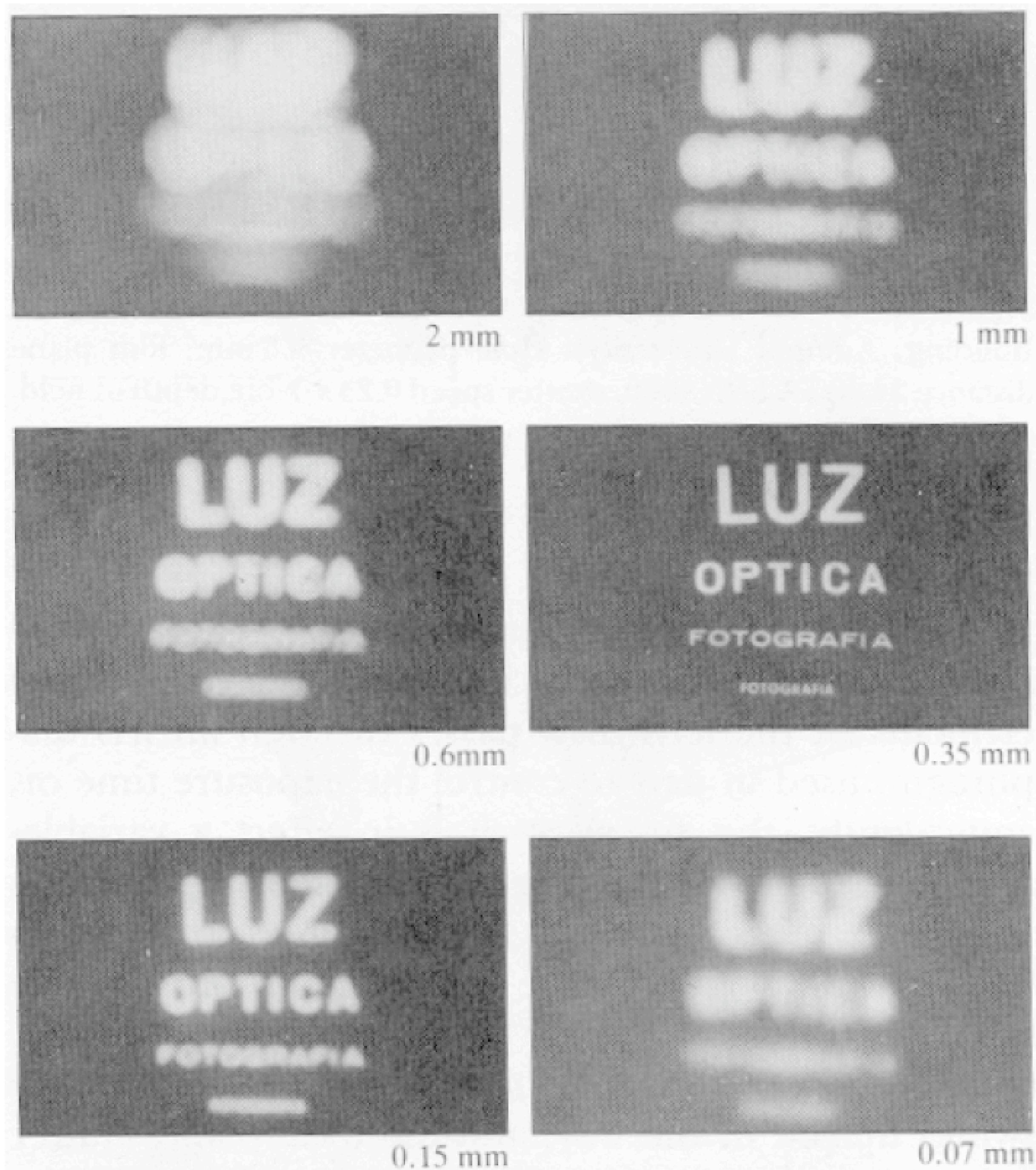
The reason for lenses



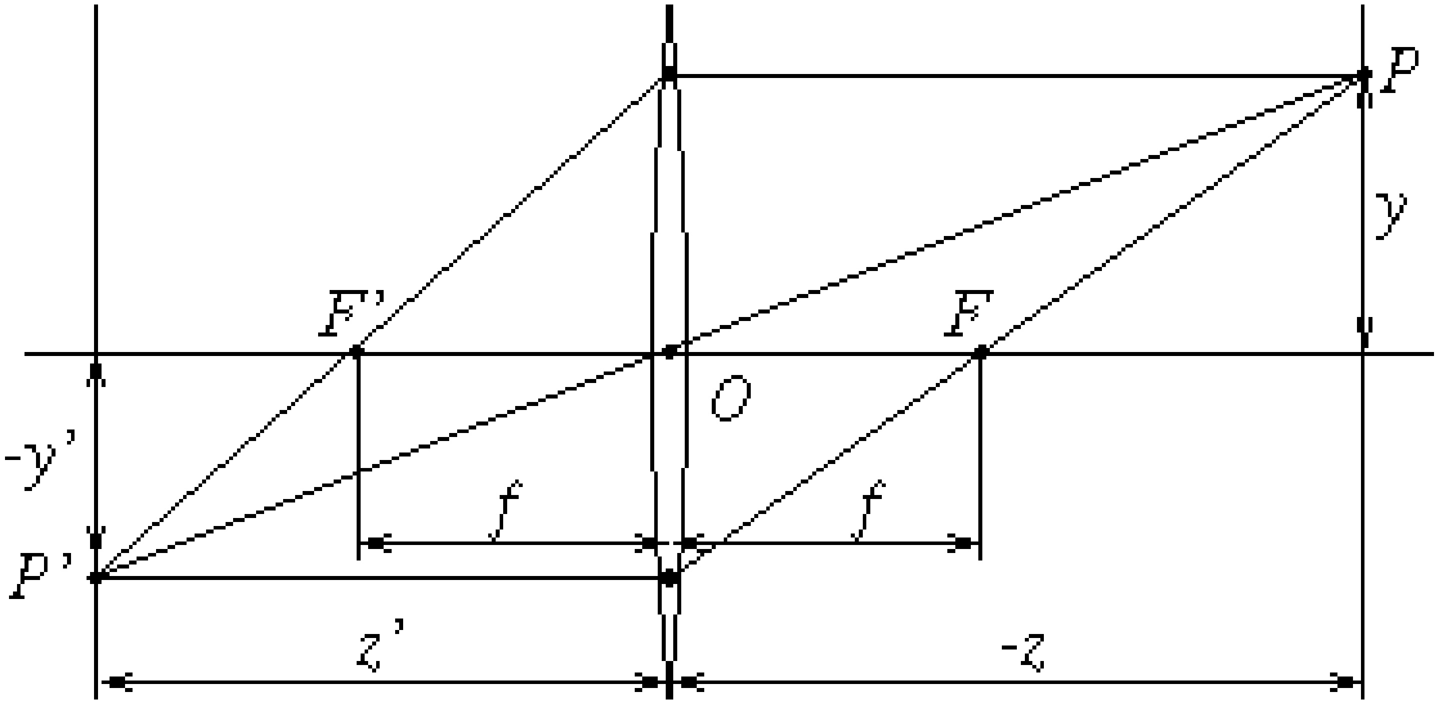
Pinhole too big - many directions are averaged, blurring the image

Pinhole too small- diffraction effects blur the image

Generally, pinhole cameras are *dark*, because a very small set of rays from a particular point hits the screen.



The thin lens



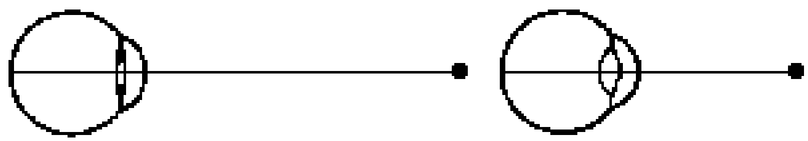
$$\frac{1}{z'} - \frac{1}{z} = \frac{1}{f}$$

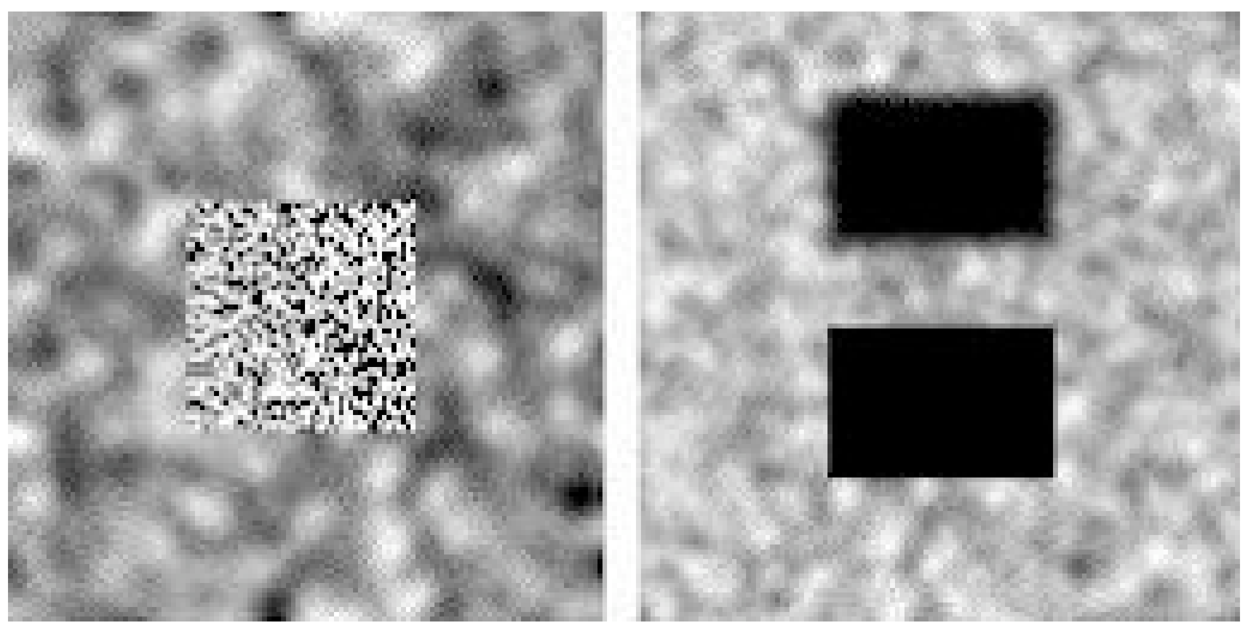
$$f = \frac{R}{2(n-1)}$$

$z' = z\,f/(z+f)$

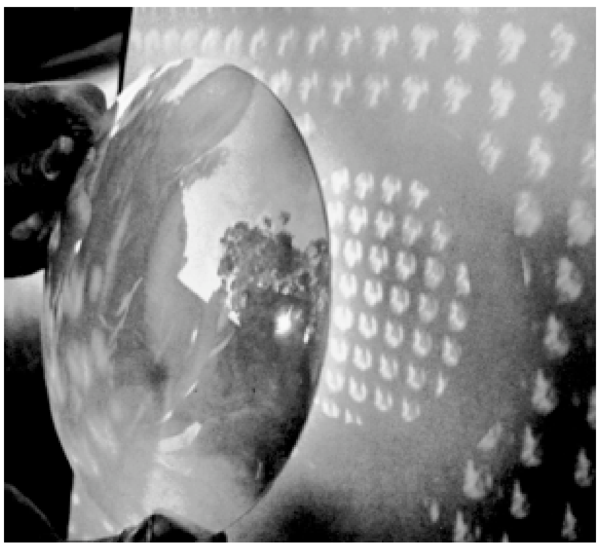
Is limited focus bad?

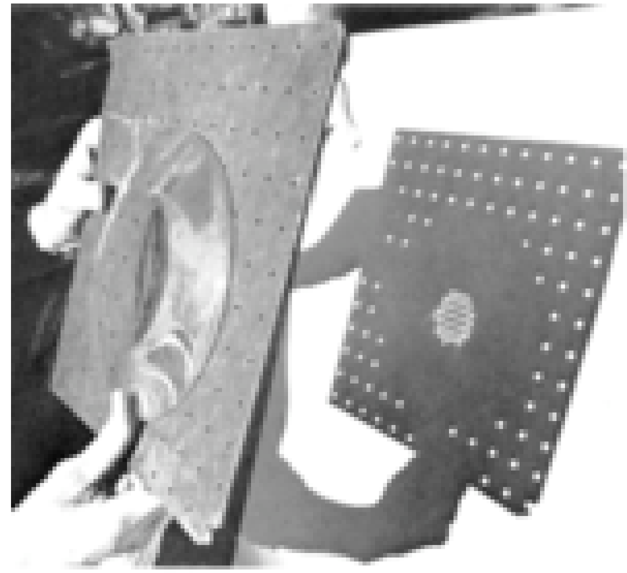
1) Know your current focal length
2) Blur gradient => Depth

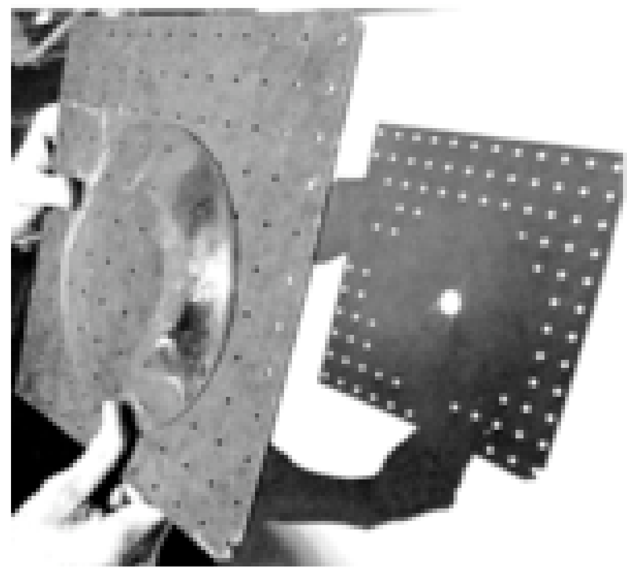




Depth of Field

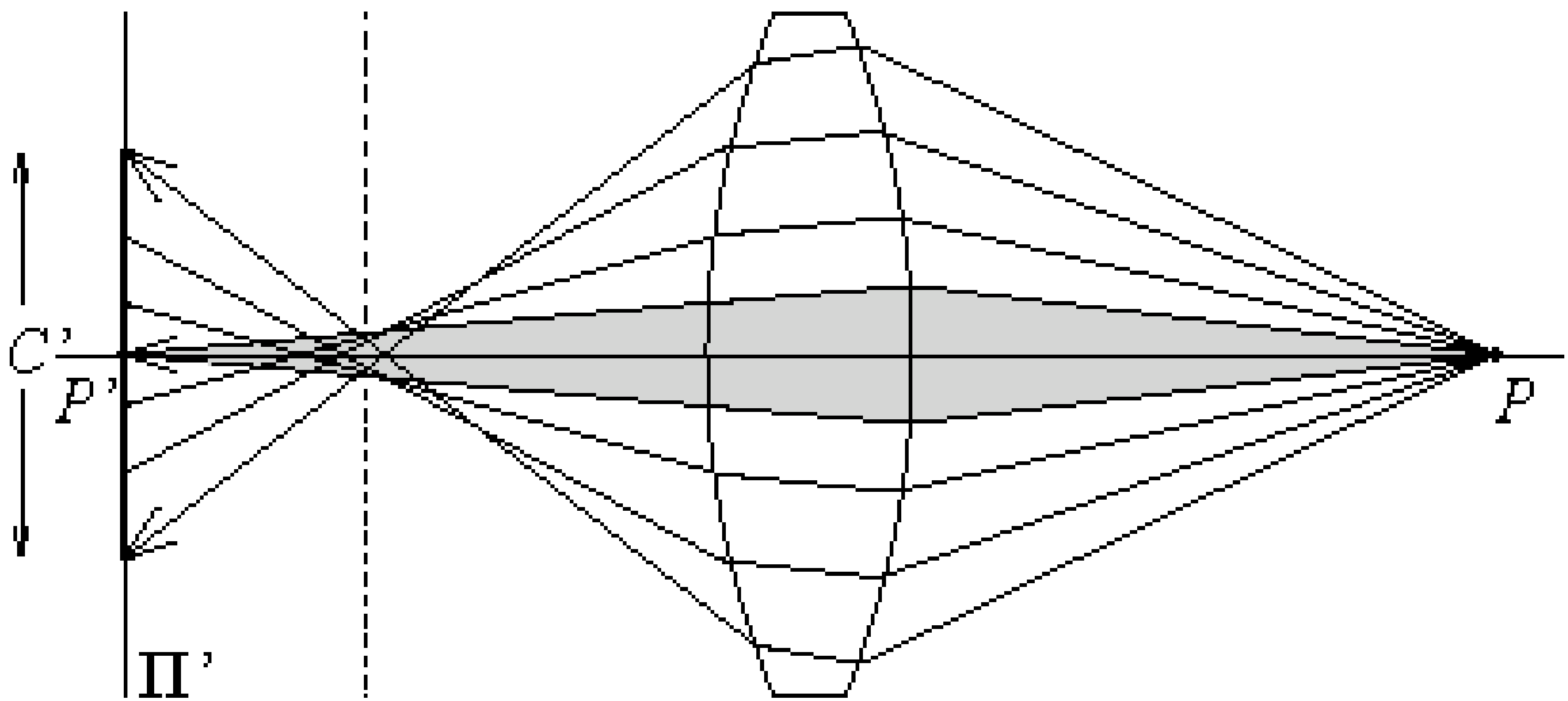




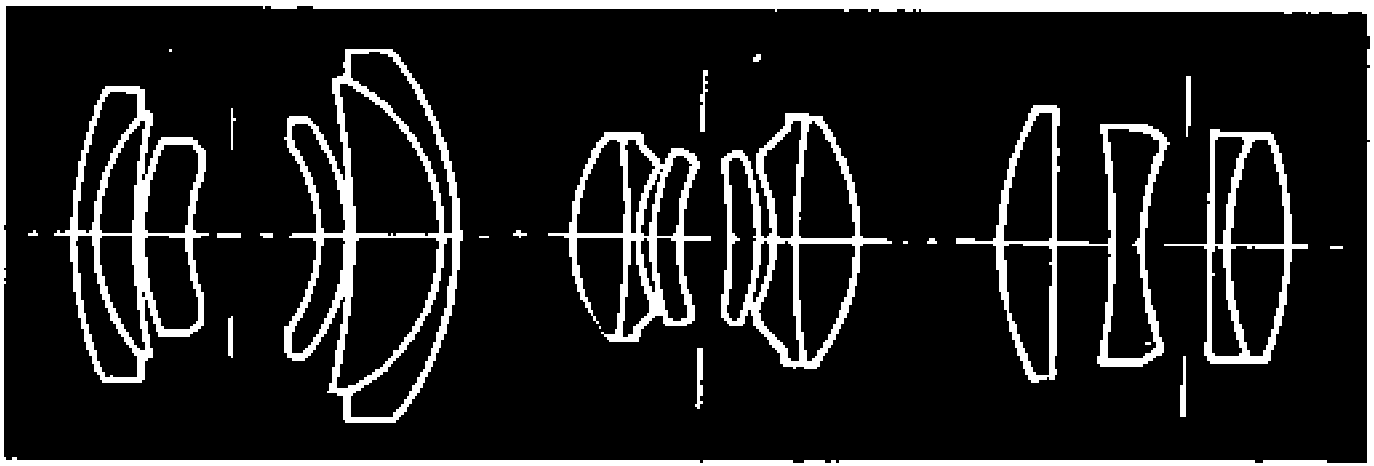


Images from Bob Miller's Light Walk:
http://www.exploratorium.edu/sln/light_walk/index.html

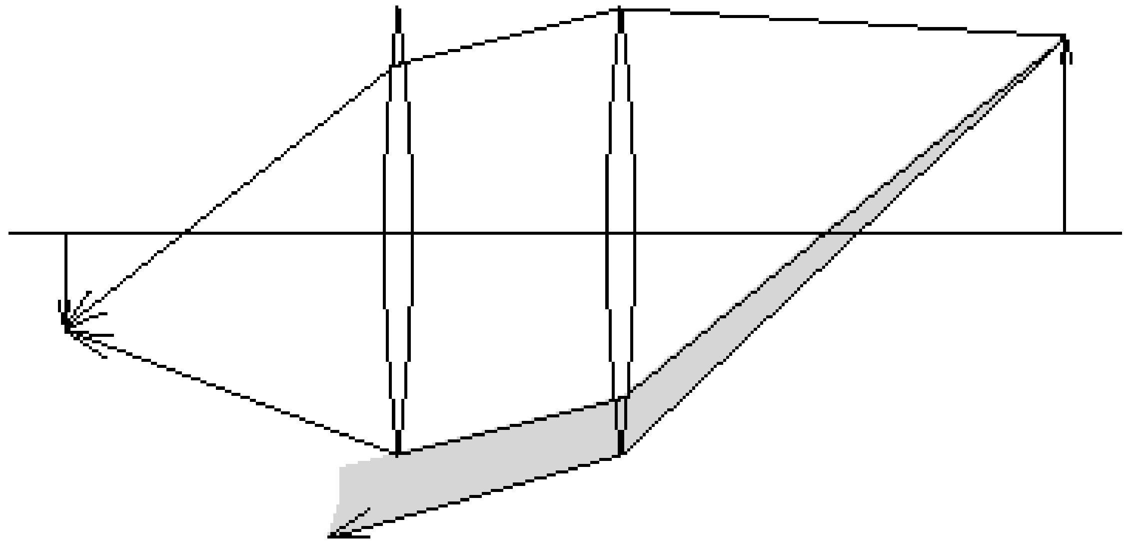
Spherical aberration



Lens systems



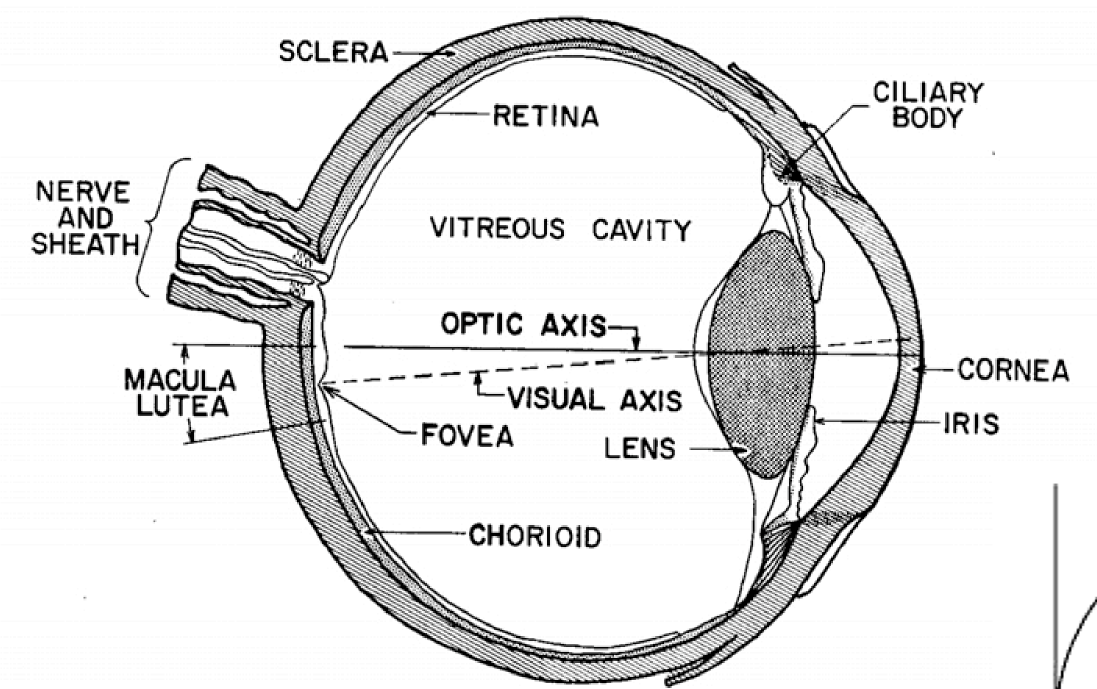
Vignetting

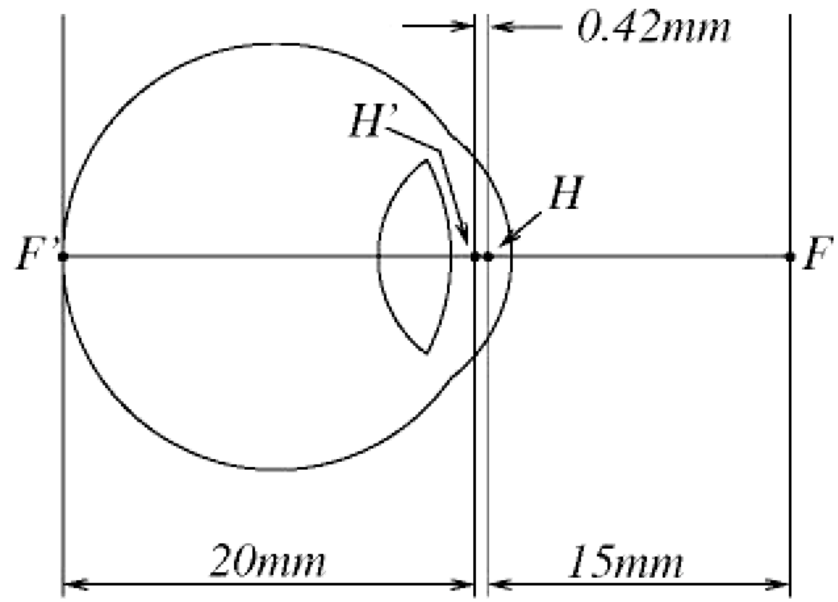


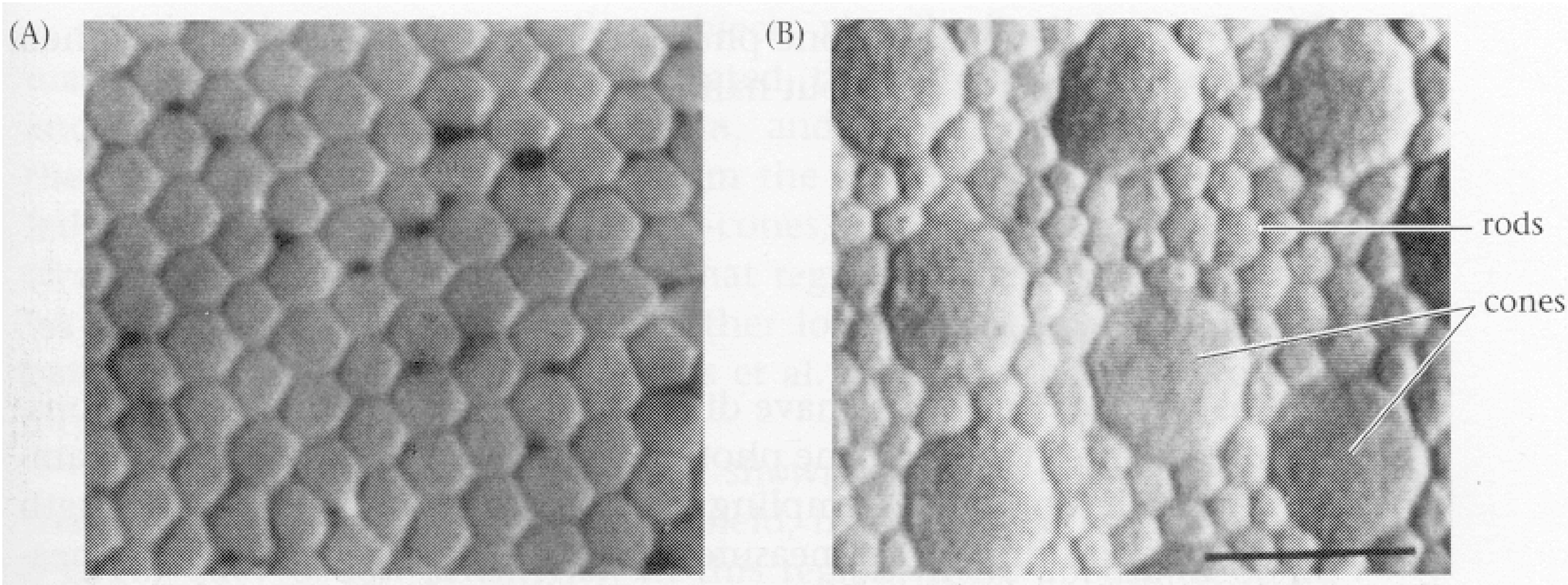
Other (possibly annoying) phenomena

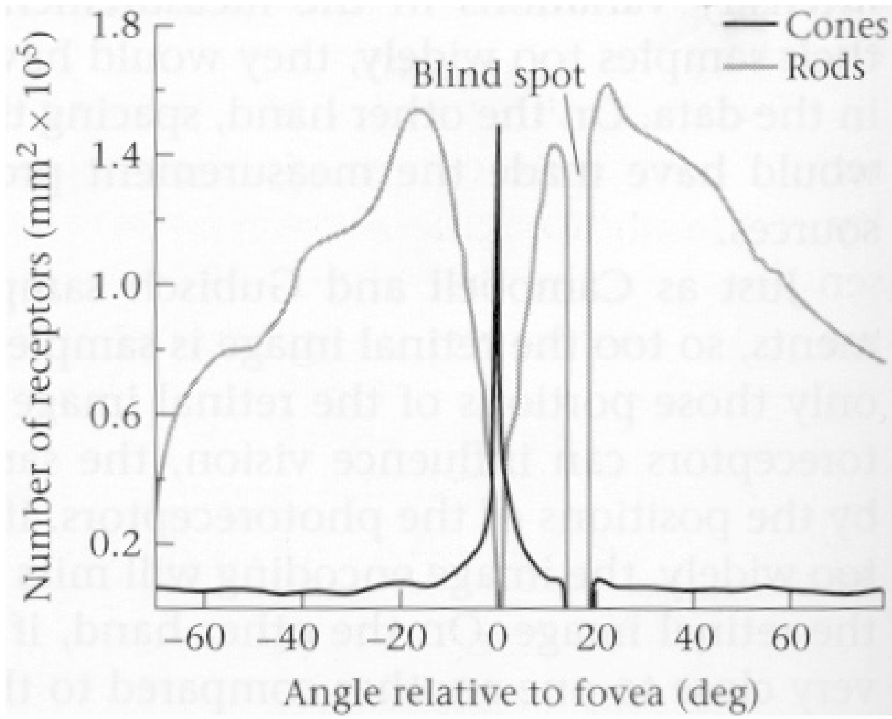
- Chromatic aberration
  - Light at different wavelengths follows different paths; hence, some wavelengths are defocussed
  - Machines: coat the lens
  - Humans: live with it
- Scattering at the lens surface
  - Some light entering the lens system is reflected off each surface it encounters (Fresnel's law gives details)
  - Machines: coat the lens, interior
  - Humans: live with it (various scattering phenomena are visible in the human eye)
- Geometric phenomena (Barrel distortion, etc.)

Human Eye-thick lens

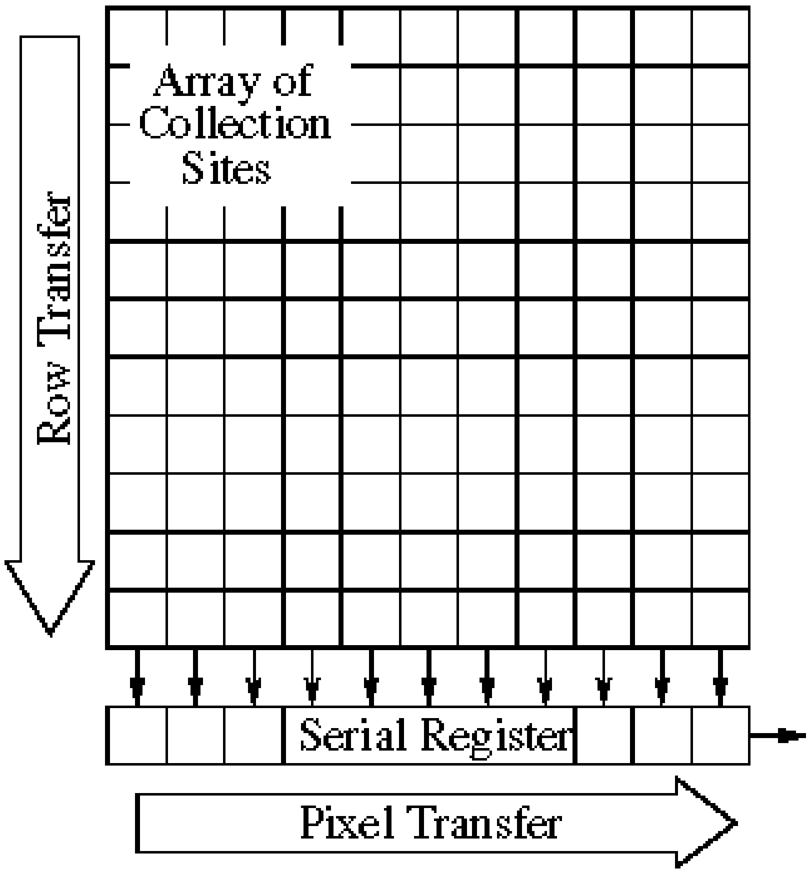





(A)
(B)
rods
cones


Cones
Rods
Blind spot
1.8
1.4
1.0
0.6
0.2
Number of receptors (mm^2 × 10^5)
−60
−40
−20
0
20
40
60
Angle relative to fovea (deg)

CCD Cameras



$$I = \int_t \int_\lambda \int_y \int_x E(x, y, \lambda, t) S(x, y) q(\lambda) dx\, dy\, d\lambda\, dt,$$

Noise sources:

1) Spatial response/Efficiency
2) Thermal fluctuations
3) Shot Noise
4) Read Noise (amplifier)
5) Quantization

Camera parameters

- Issue
  - camera may not be at the origin, looking down the z-axis
    - extrinsic parameters
  - one unit in camera coordinates may not be the same as one unit in world coordinates
    - intrinsic parameters - focal length, principal point, aspect ratio, angle between axes, etc.

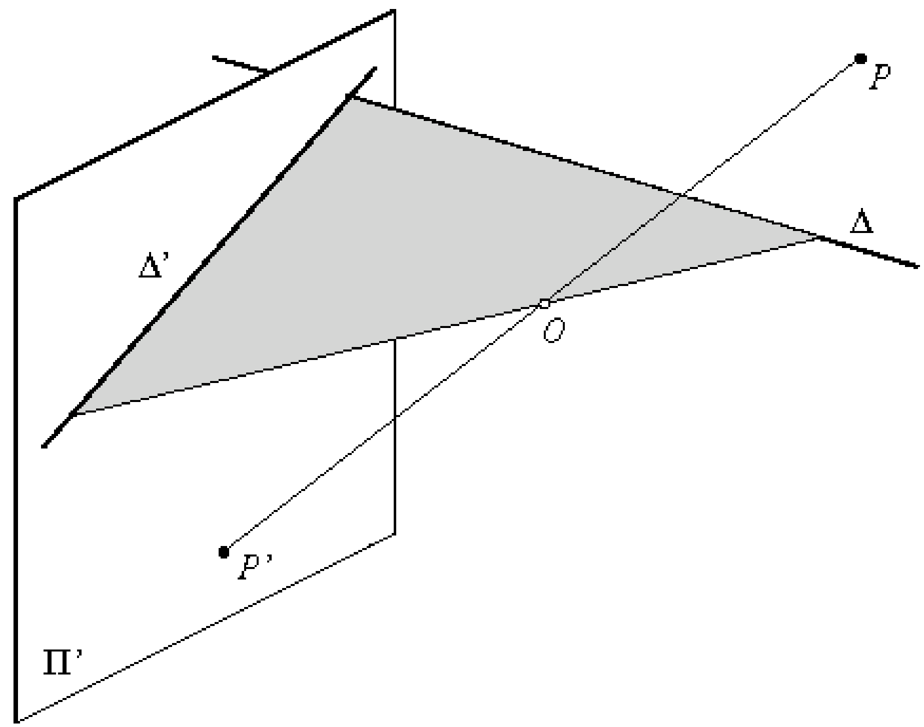
$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} \text{Transformation} \\ \text{representing} \\ \text{intrinsic parameters} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \text{Transformation} \\ \text{representing} \\ \text{extrinsic parameters} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix}$$

Camera calibration

- Issues:
  - what are intrinsic parameters of the camera?
  - what is the camera matrix? (intrinsic+extrinsic)
- General strategy:
  - view calibration object
  - identify image points
  - obtain camera matrix by minimizing error
  - obtain intrinsic parameters from camera matrix
- Error minimization:
  - Linear least squares
    - easy problem numerically
    - solution can be rather bad
  - Minimize image distance
    - more difficult numerical problem
    - solution usually rather good,
    - start with linear least squares
  - Numerical scaling is an issue

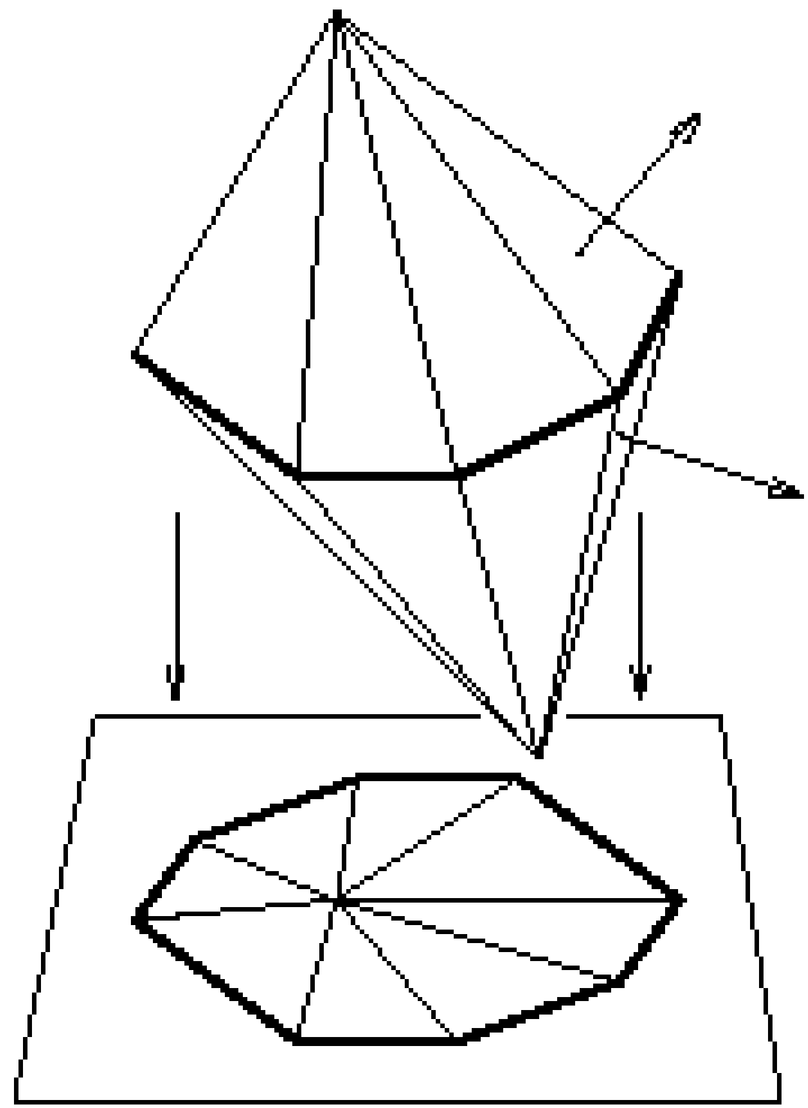
Geometric properties of projection

- Points go to points
- Lines go to lines
- Planes go to whole image
- Polygons go to polygons
- Degenerate cases
  - line through focal point to point
  - plane through focal point to line



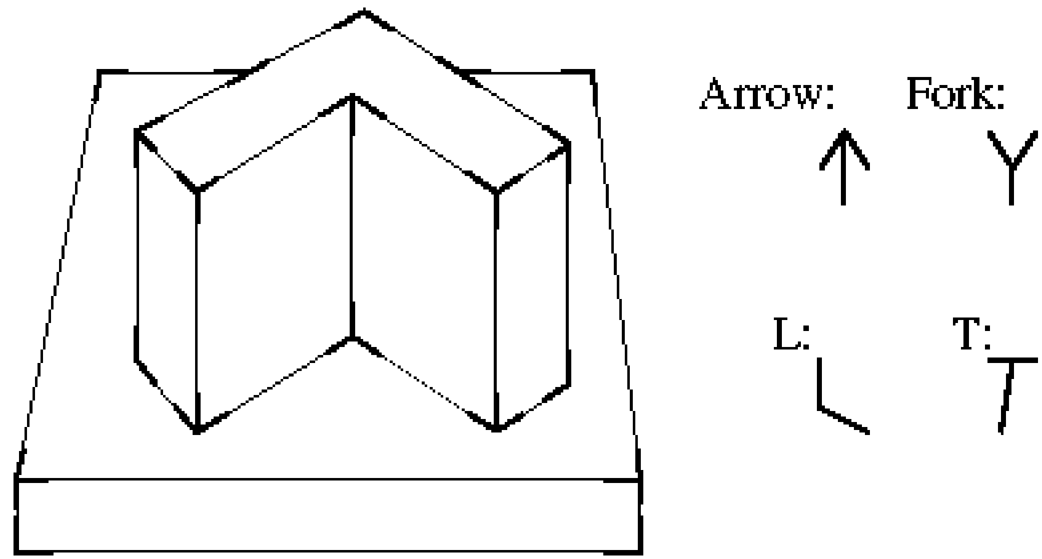
Polyhedra project to polygons

- (because lines project to lines)



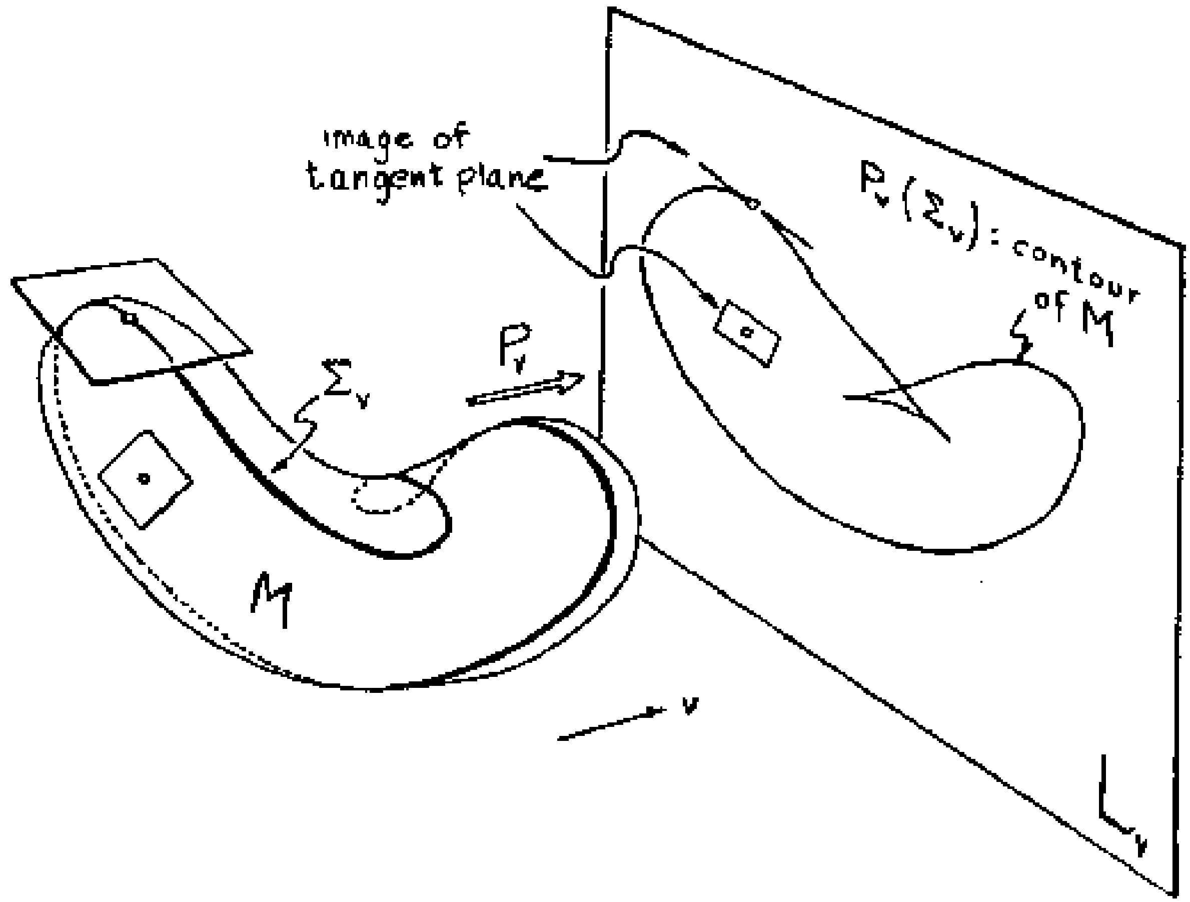
Junctions are constrained

- This leads to a process called “line labelling”
  - one looks for consistent sets of labels, bounding polyhedra
  - disadv - can’t get the lines and junctions to label from real images



Curved surfaces are much more interesting

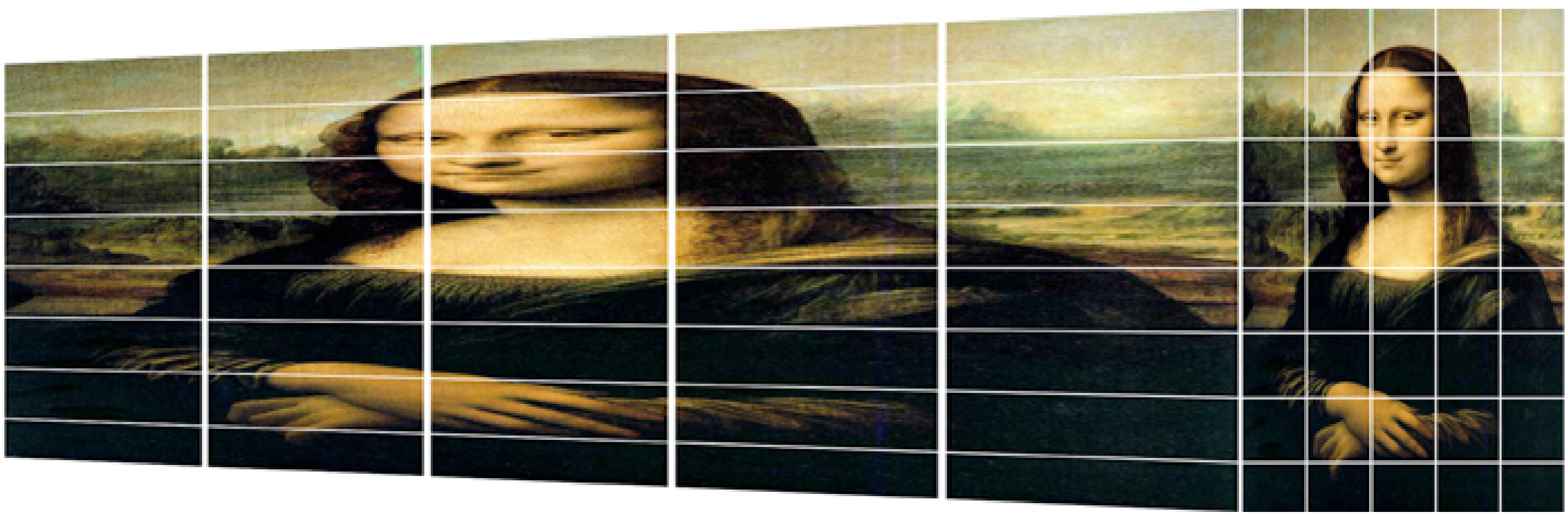
- Crucial issue:
  outline is the set of points where the viewing direction is tangent to the surface
- This is a projection of a space curve, which varies from view to view of the surface



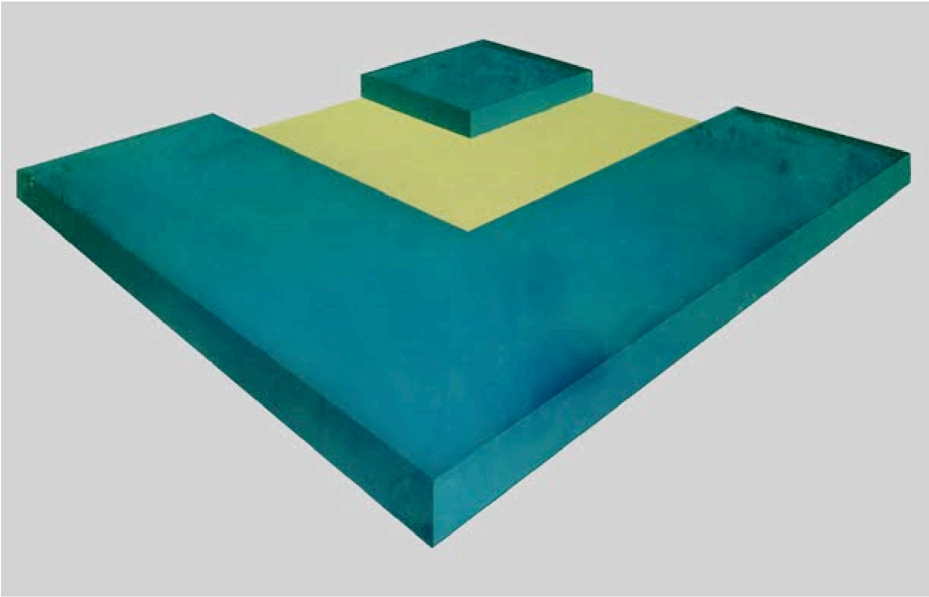
Viewpoint dependent Projections

Samuel von Hoogstraten





What's wrong?



Ron Davis
Six-Ninths Blue, 1966

Ames Room