Cameras

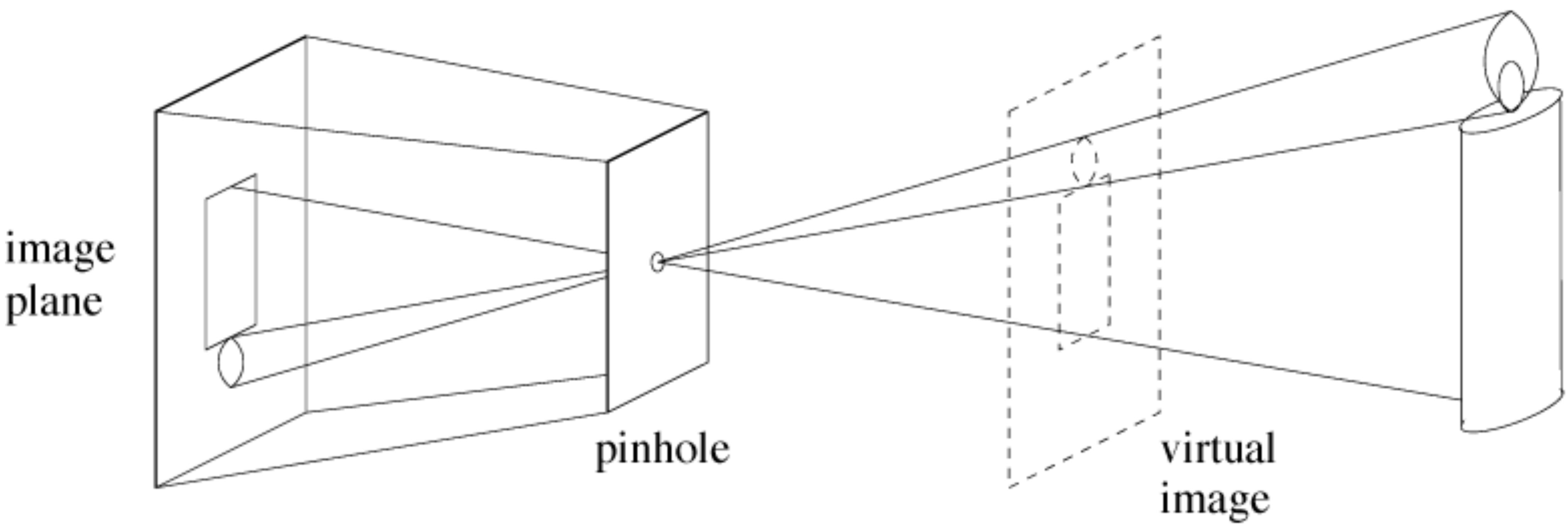
D.A. Forsyth, UIUC

Cameras

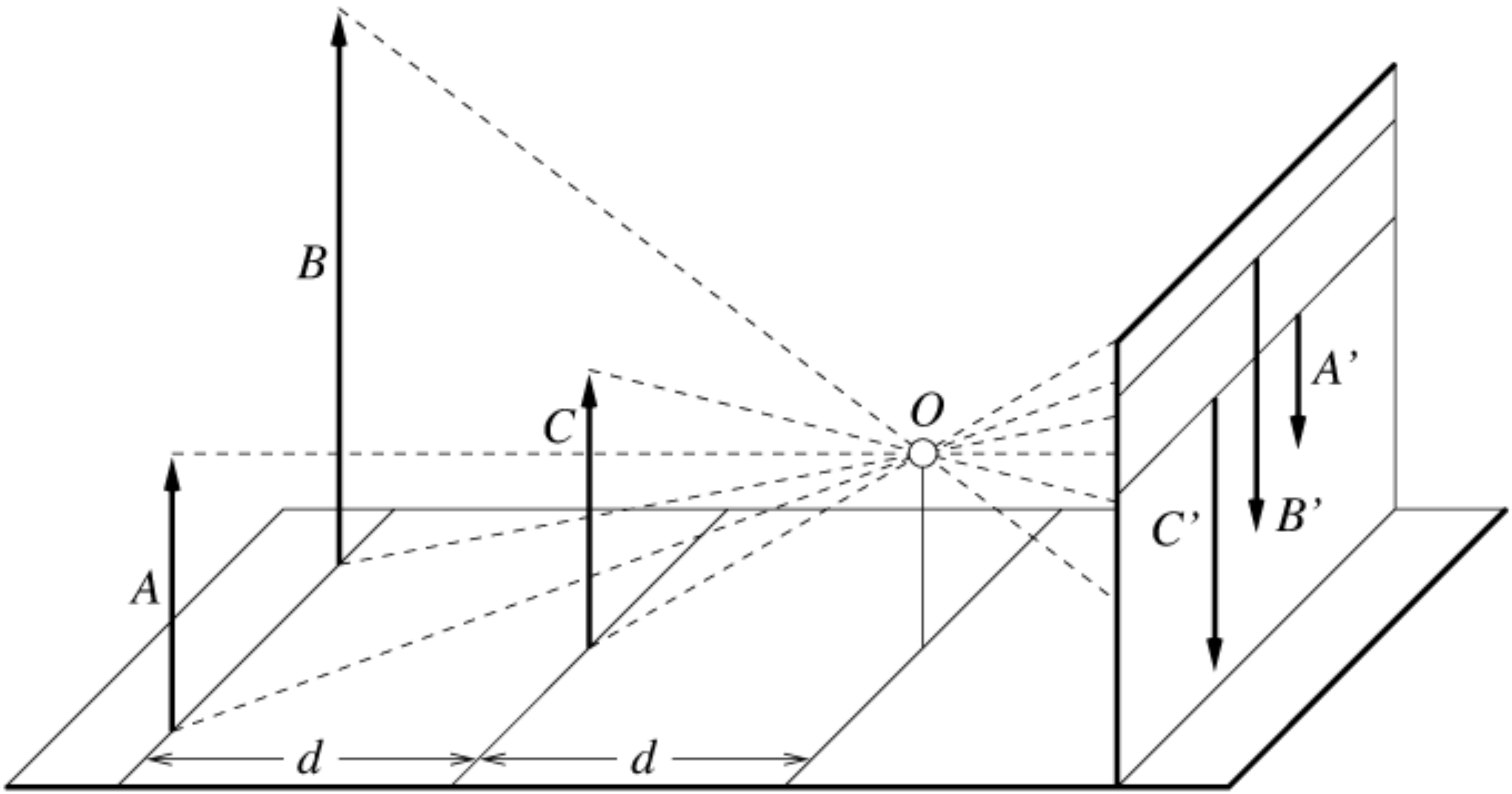
- First photograph due to Niepce
- First on record shown in the book - 1822
- Basic abstraction is the pinhole camera
  - lenses required to ensure image is not too dark
  - various other abstractions can be applied

Pinhole cameras

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

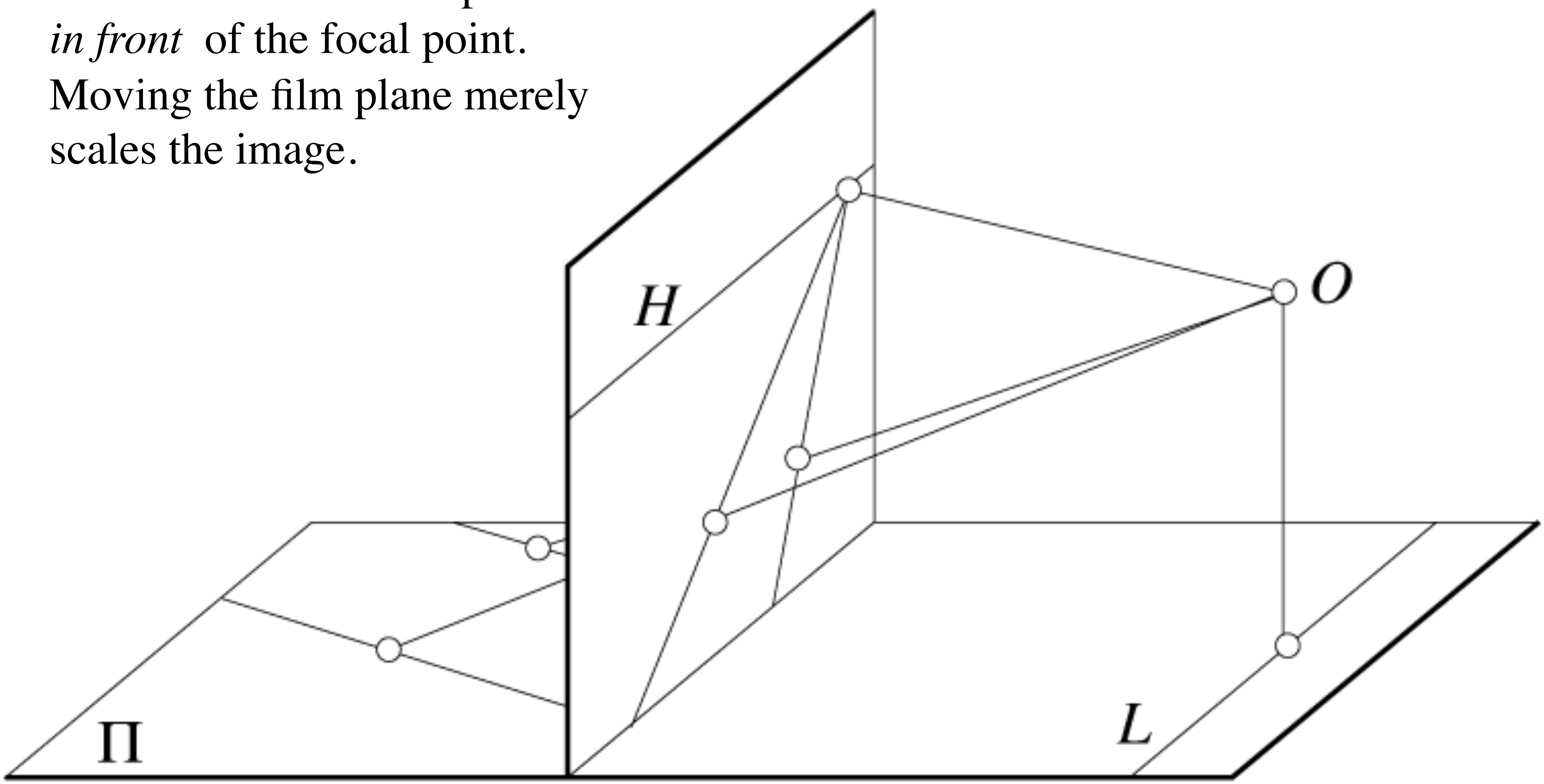


Distant objects are smaller



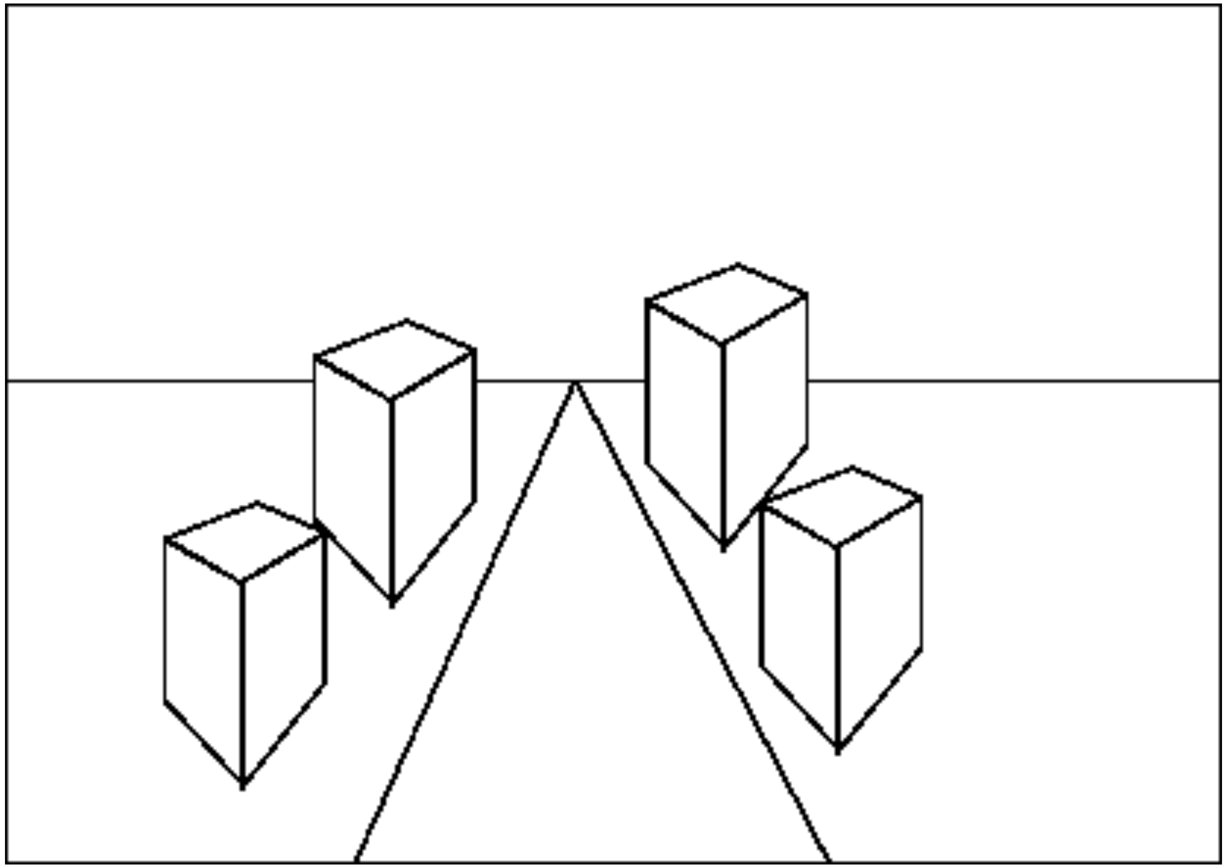
Parallel lines meet

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

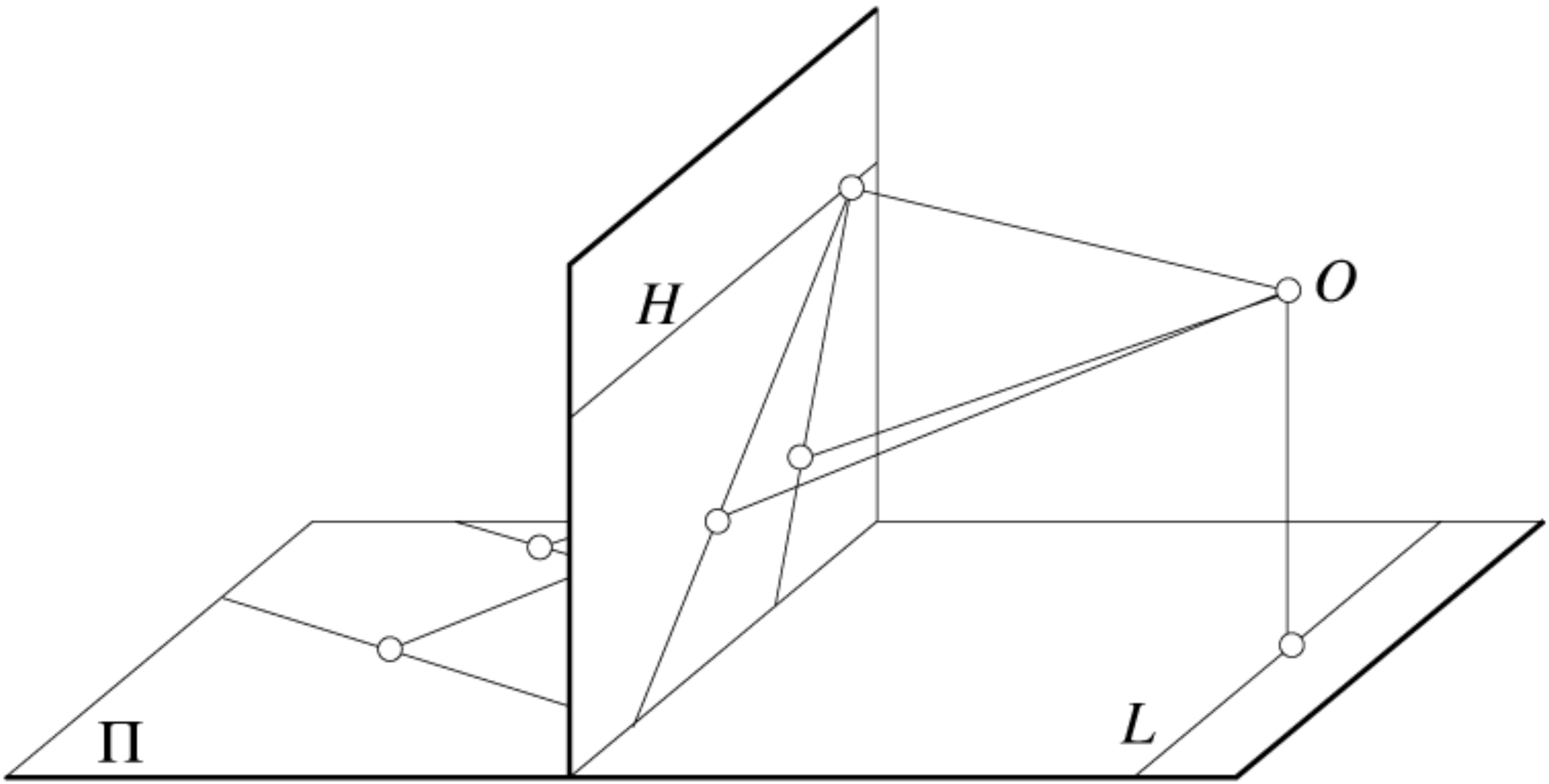


Vanishing points

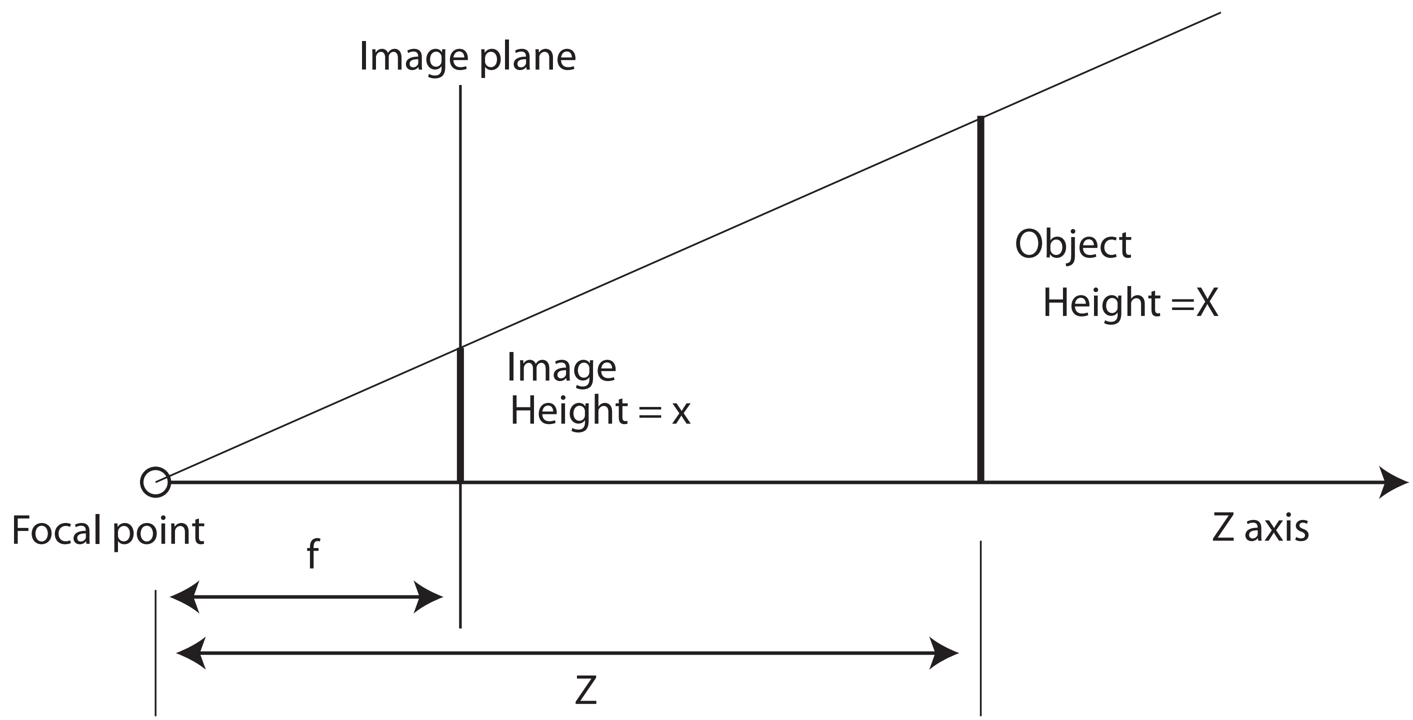
- Each set of parallel lines (=direction) meets at a different point
  - The vanishing point for this direction
- Sets of parallel lines on the same plane lead to collinear vanishing points.
  - The line is called the horizon for that plane
- Good ways to spot faked images
  - scale and perspective don't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



The equation of projection - I

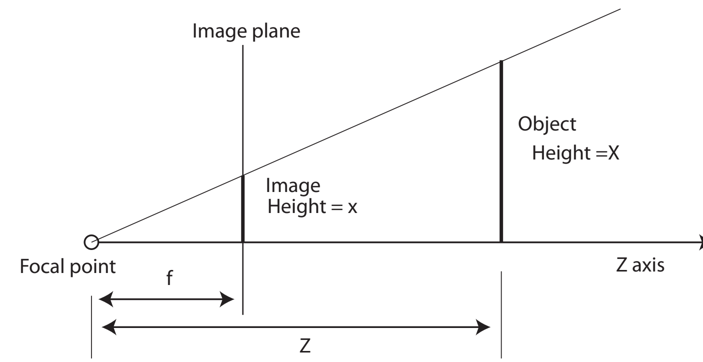


The equation of projection - II



The equation of projection - III

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



  - Ignore the third coordinate, and get

$$(X, Y, Z) \to (fX/Z, fY/Z)$$

  - notice we could have sign changes, etc. depending on
    - whether there is a right handed/left handed coordinate system
    - whether image plane is in front of/behind focal point

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
- Can 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)

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

The camera matrix - II

- 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)

Transforms points from object coordinates into world coordinates most likely a rotation and translation

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

Transforms camera coordinates (f is hidden in here)

Usual forms

Pixel aspect ratio

Camera intrinsic parameters

$$\mathcal{C} = \begin{pmatrix} fa & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{pmatrix}$$

Location of the camera center (where the z-axis pierces the image plane)

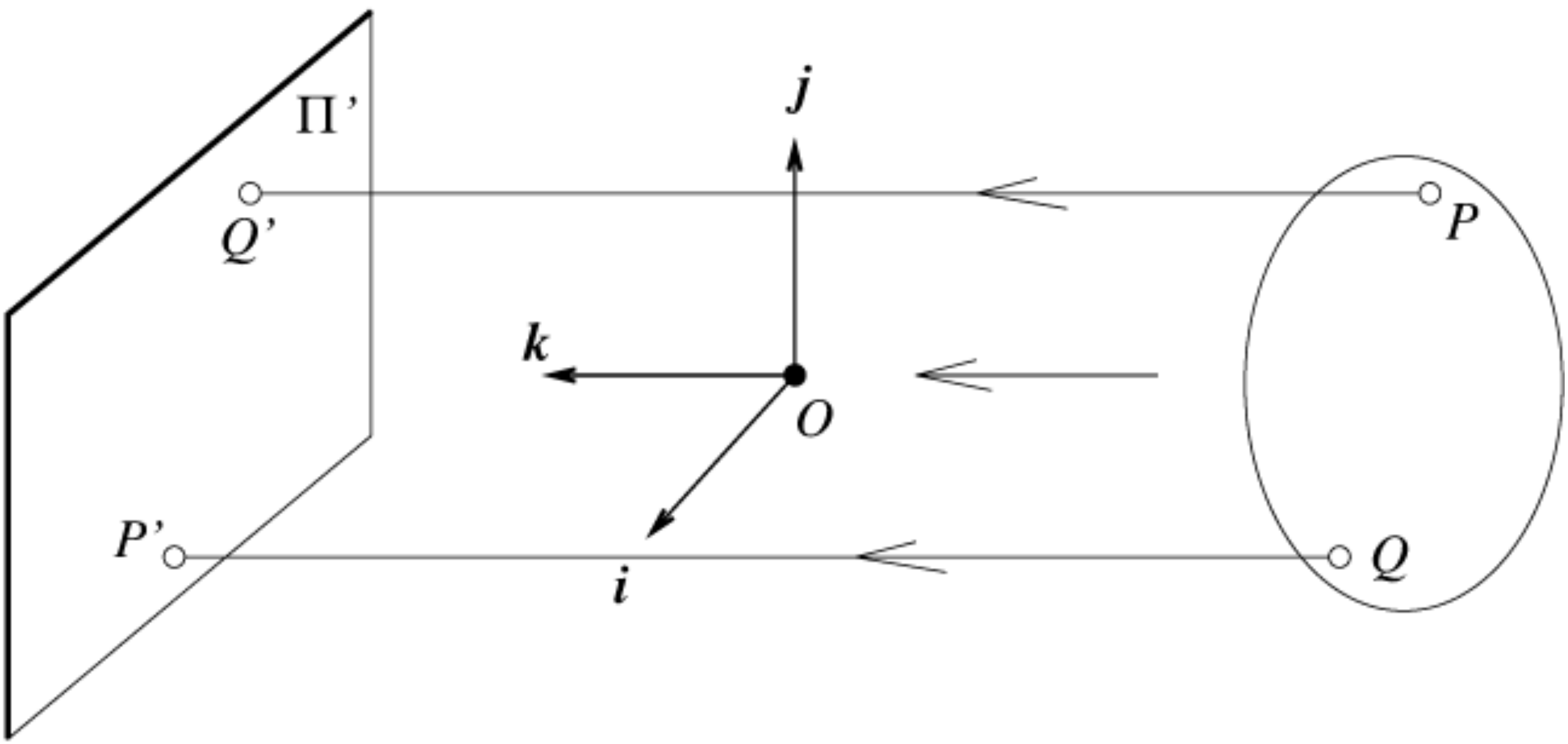
A scale, incorporating focal length and pixel size

3D rotation matrix

3D translation vector

$$\mathcal{W} = \begin{pmatrix} \mathcal{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{pmatrix}$$

Orthographic projection



The projection matrix for orthographic projection

- Almost never encounter orthographic projection

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

Weak perspective or Affine Cameras

- Assume
  - The range of depths over points is small compared to distance to points

$$Z_i = Z_0 + \delta Z_i$$

$$\frac{1}{Z_i} = \frac{1}{Z_0}\left(\frac{1}{1+\frac{\delta Z_i}{Z_0}}\right) \approx \frac{1}{Z_0}(1 - \frac{\delta Z_i}{Z_0}) \approx \frac{1}{Z_0}$$

- So you could scale all points with one scale
  - Scaled orthography

Affine Cameras - I

- And this becomes (for the relevant group of points)

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

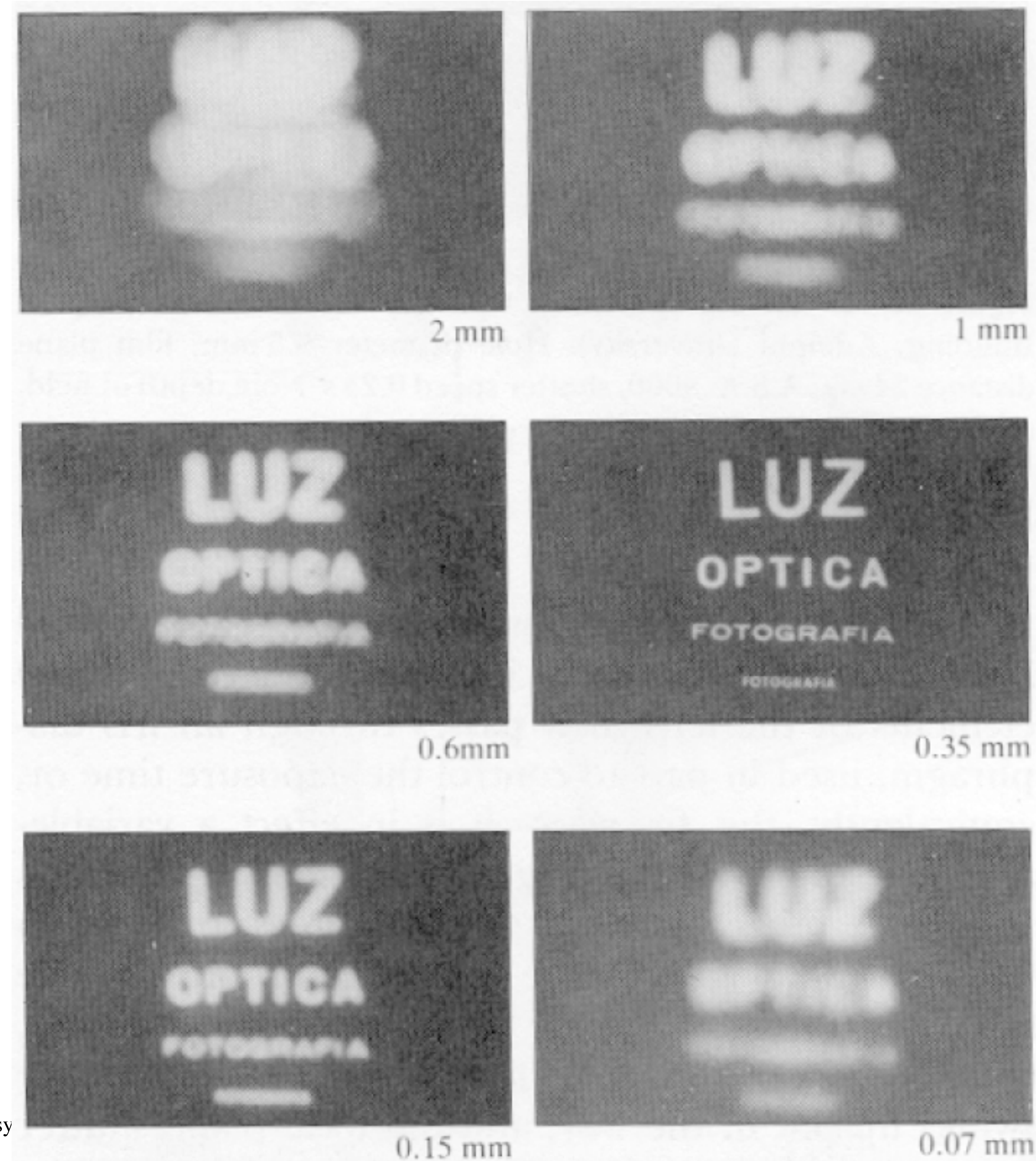
Hide the scale in here

- We will see further simplifications soon

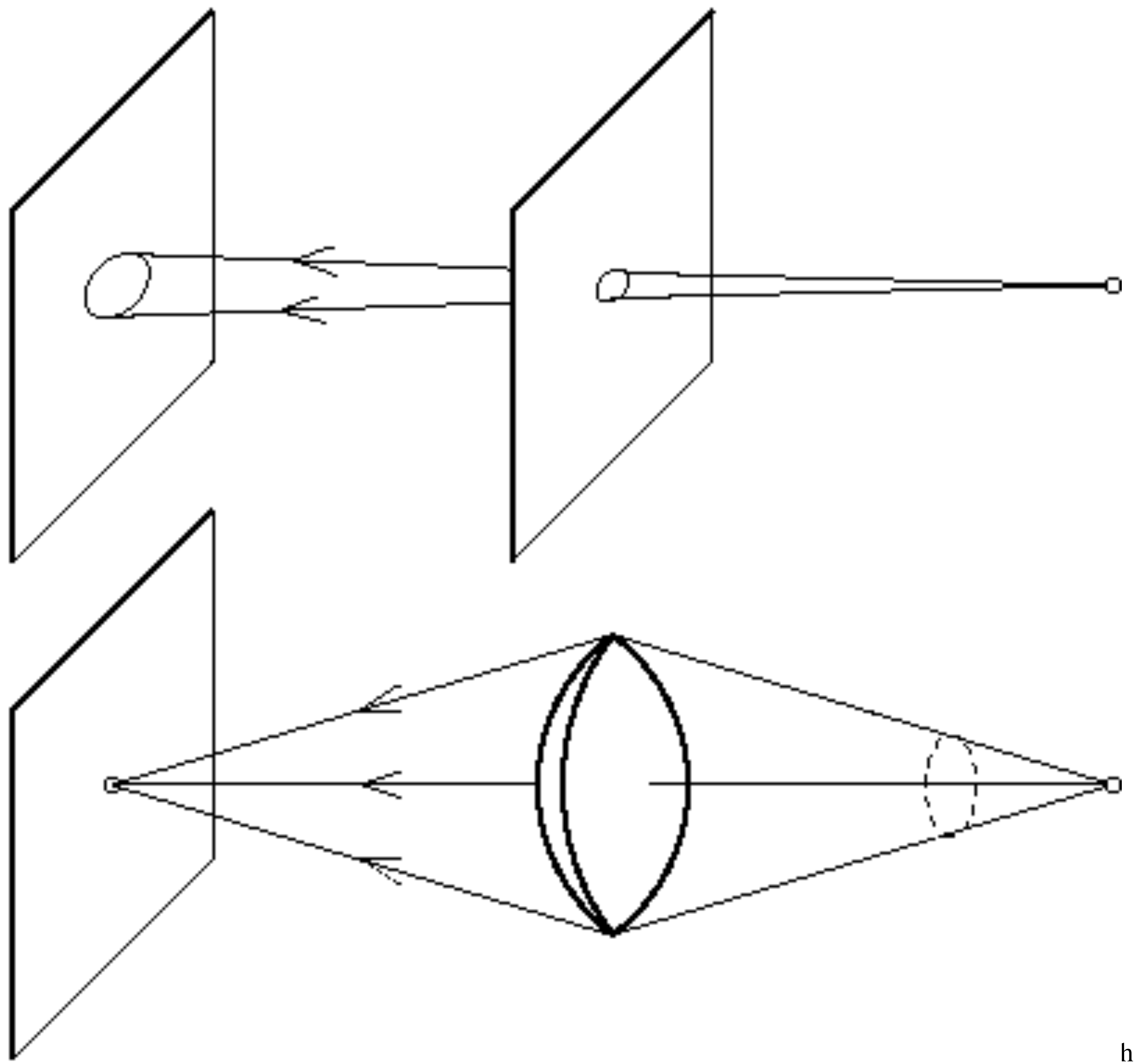
Pinhole too big - many directions are averaged, blurring the image

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.



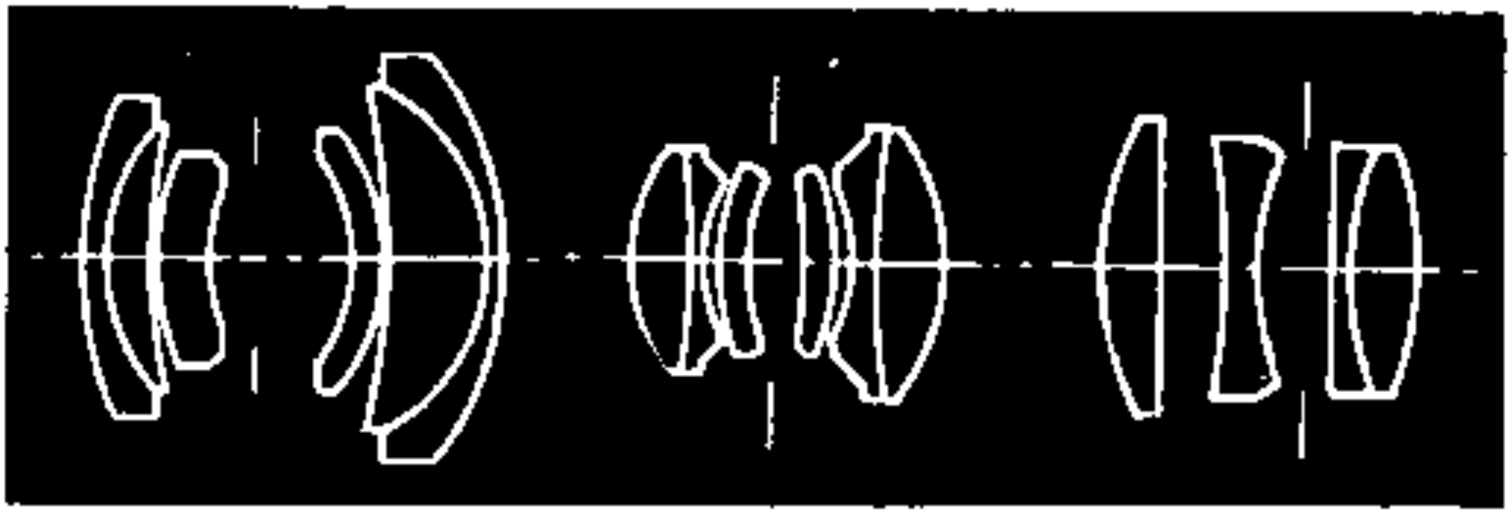
The reason for lenses



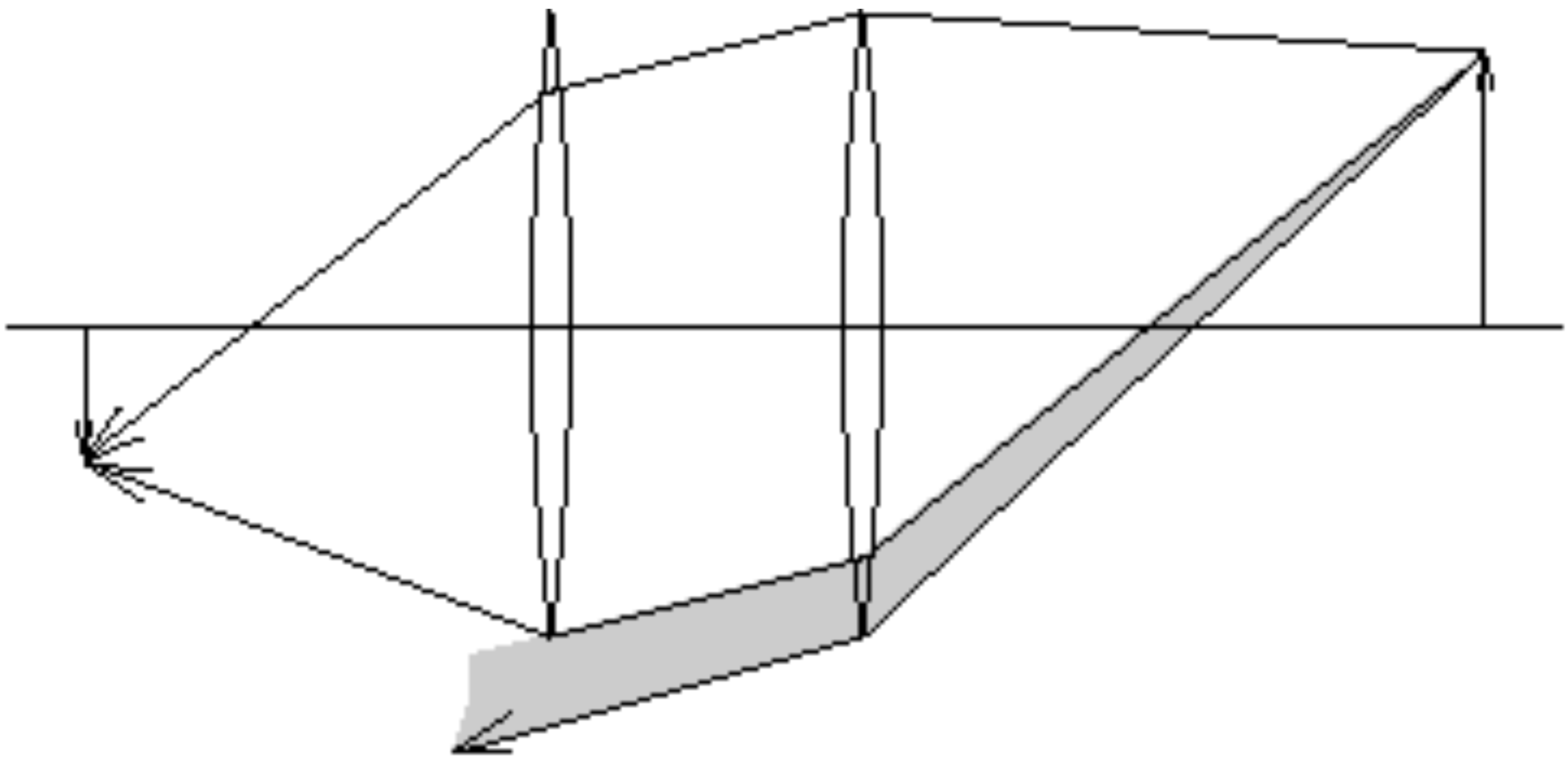
Lenses come with problems

- Spherical aberration
  - Lens is not a perfect thin lens, and point is defocused

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.)

Camera calibration - I

- Issues:
  - what are intrinsic parameters of the camera?
  - what is the camera matrix? (intrinsic+extrinsic)
- General strategy:
  - view calibration object; identify image points in image
  - obtain camera matrix by minimizing error
  - obtain intrinsic parameters from camera matrix

Camera calibration - II

- Error minimization:
  - Linear least squares
    - easy problem numerically
    - solution can be rather bad
  - Minimize image distance
    - more difficult numerical problem
    - solution usually rather good
    - Numerical scaling is an issue
  - Strategy:
    - start with linear least squares, then minimize image distance

Cameras - crucial points

- Pinhole camera is a simple, effective model
  - With important effects
    - distant objects are smaller
    - parallel lines meet in image
- Alternative model: orthographic projection
  - distant objects are not smaller
  - parallel lines do not meet
- Each has straightforward mathematical form
- Most cameras have lenses
  - otherwise they'd be too dark