Computer vision
Cameras

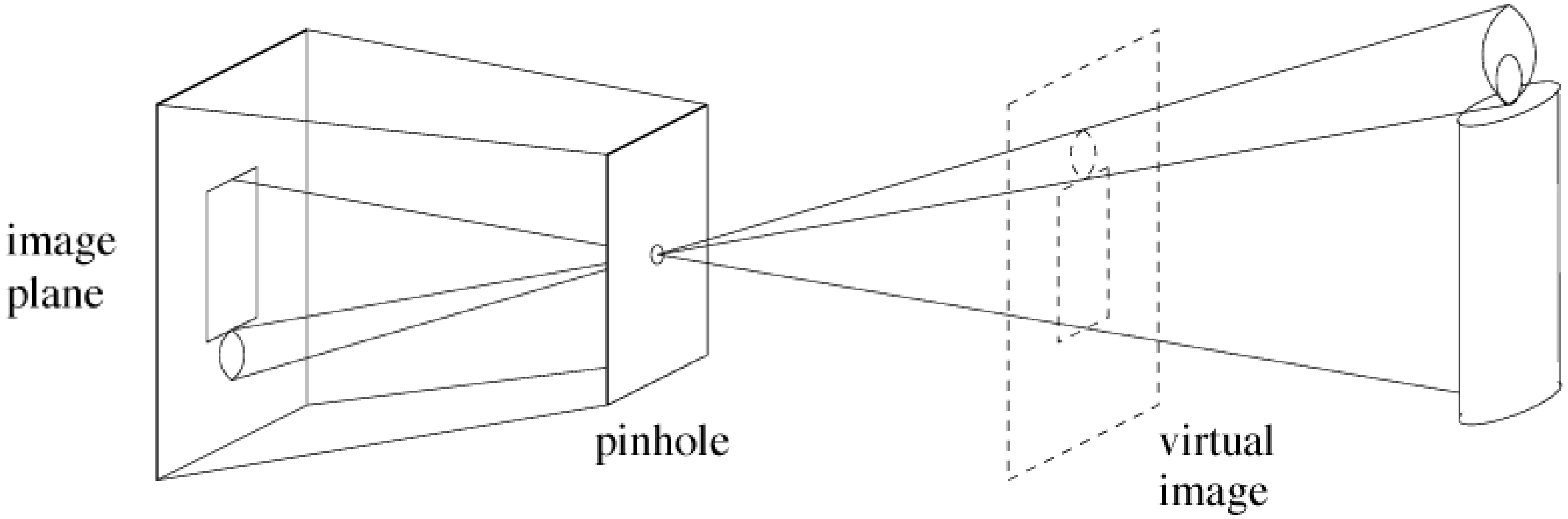
Le Thanh Ha, Ph.D
Assoc. Prof. at University of Engineering and Technology,
Vietnam National University
lthavnu@vnu.edu.vn; lthavnu@gmail.com; 0983 692 592

Cameras

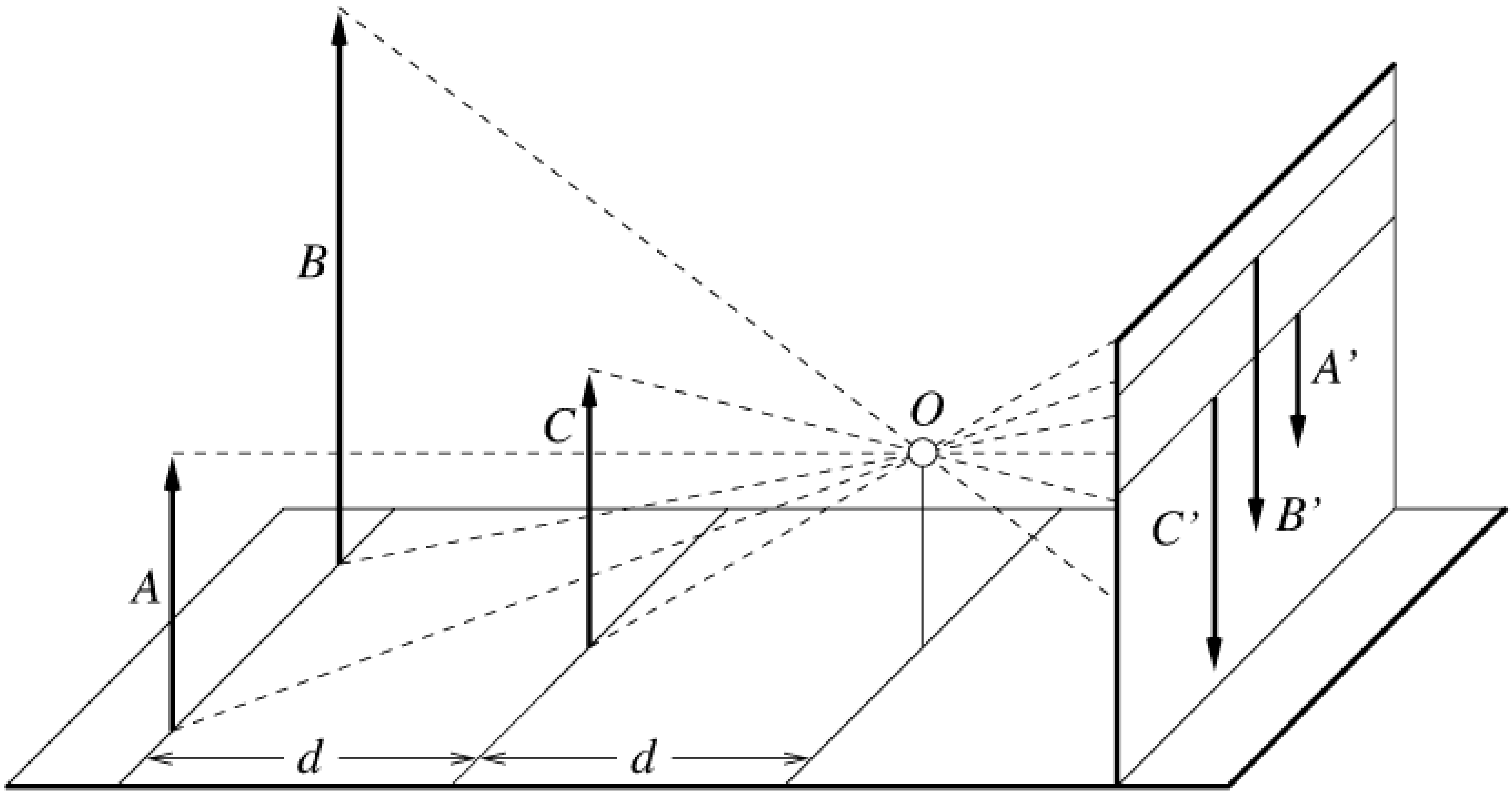
- First photograph due to Niepce
- First on record shown in the book - 1822
- Basic abstraction is the pinhole camera
  - lenses required to ensure image is not too dark
  - various other abstractions can be applied

Pinhole cameras

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

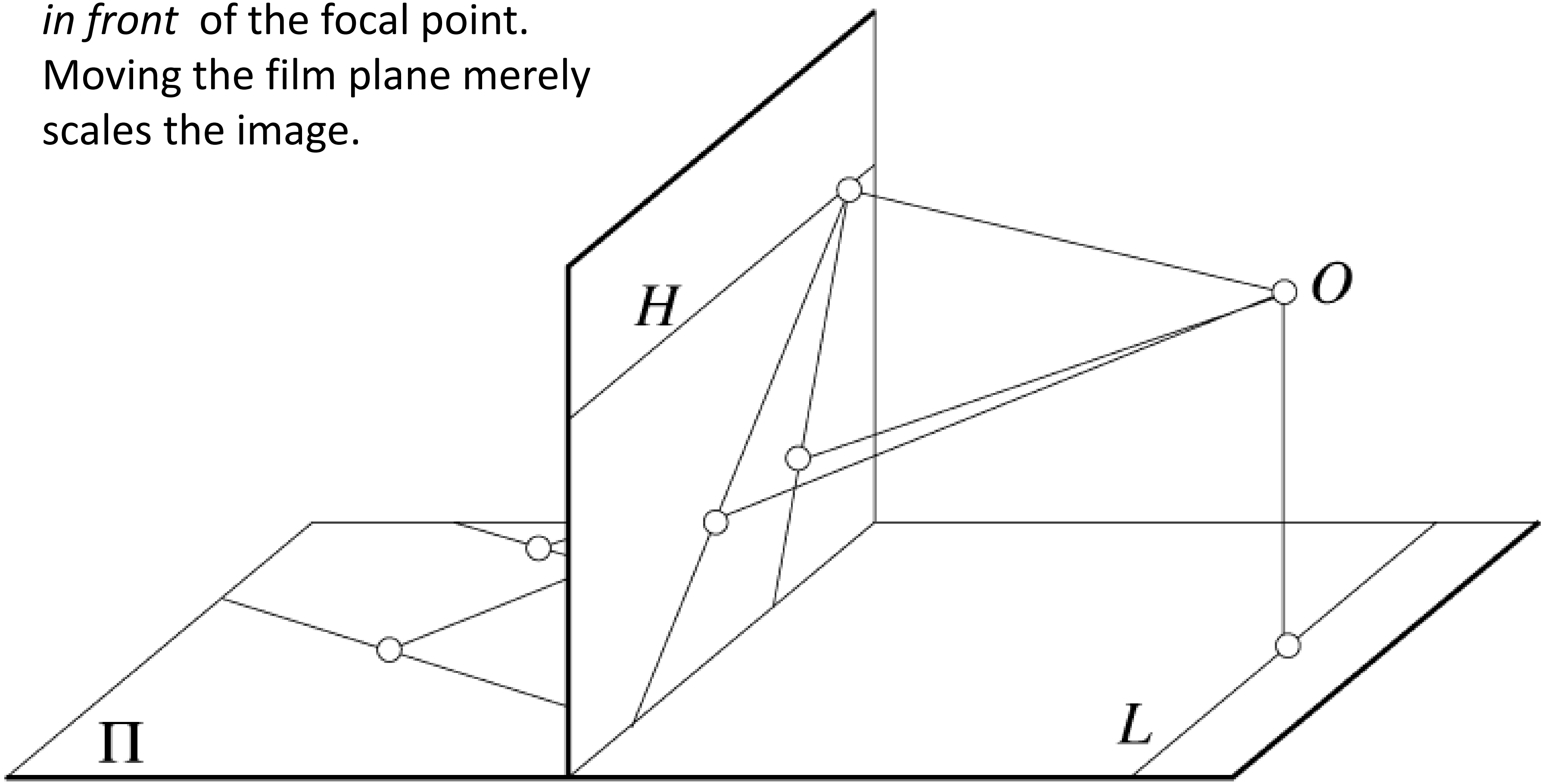


Distant objects are smaller



Parallel lines meet

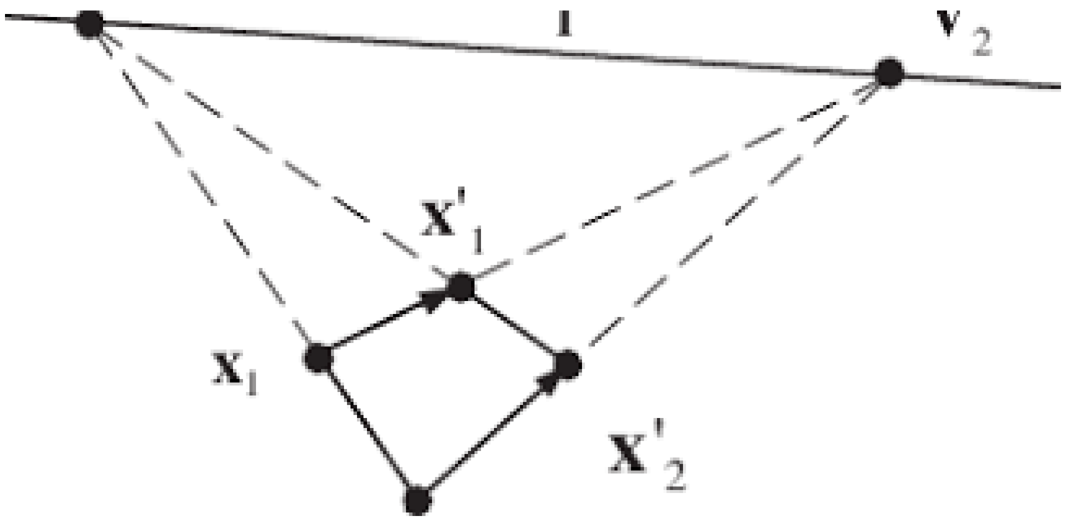
Common to draw film plane
in front of the focal point.
Moving the film plane merely
scales the image.



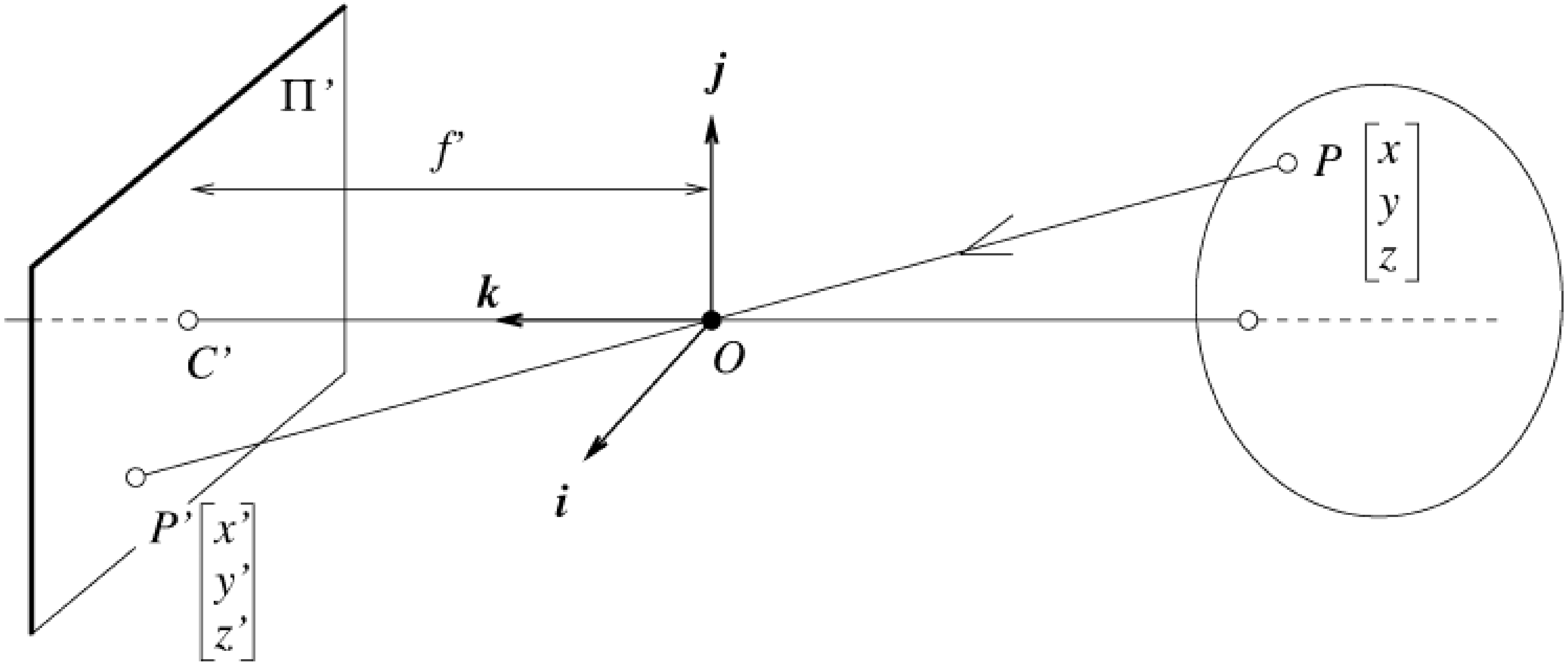
Vanishing points

- 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



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)

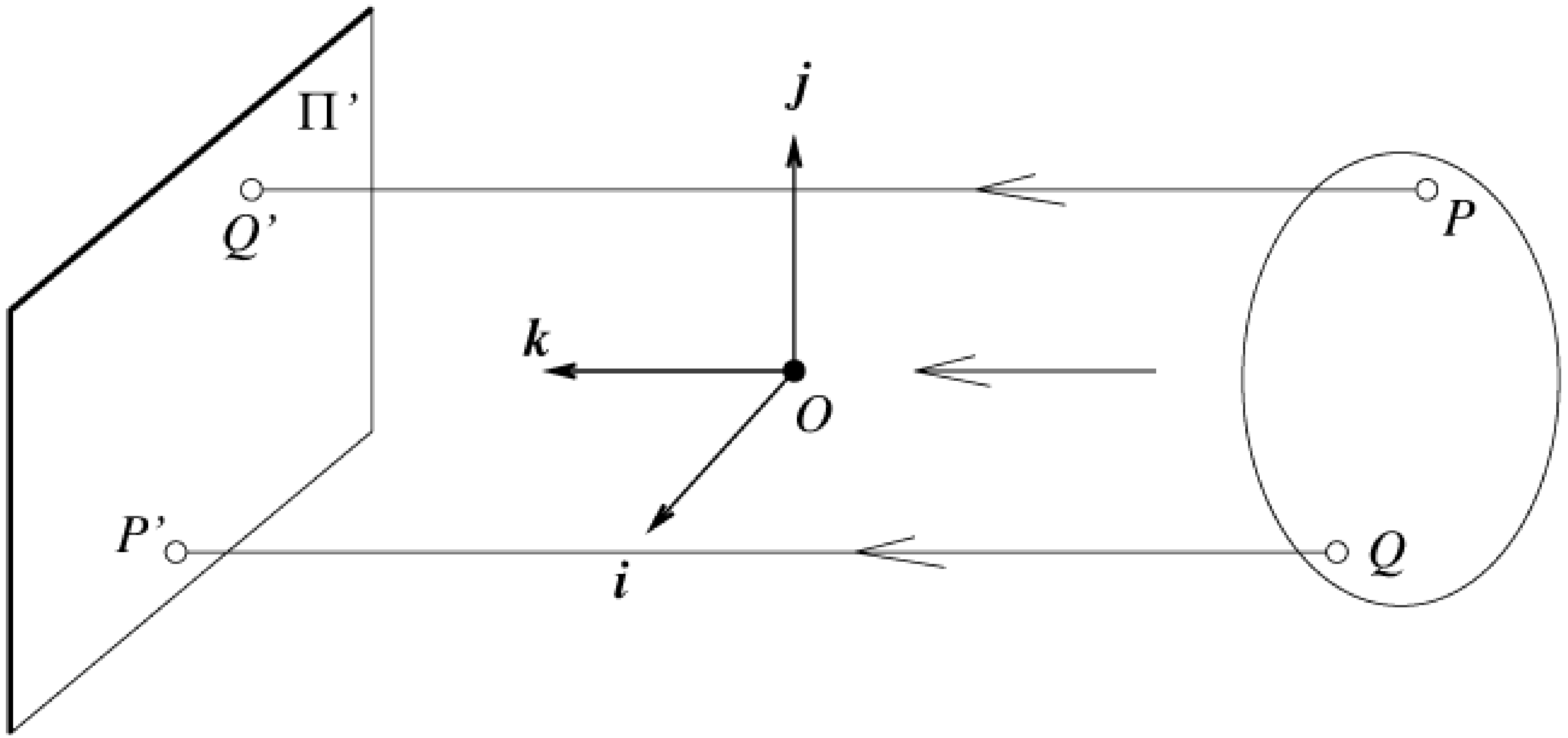
- Basic notion
  - Possible to represent points "at infinity"
    - Where parallel lines intersect
    - Where parallel planes intersect
  - Possible to write the action of a perspective camera as a matrix

The camera matrix

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

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

Orthographic projection



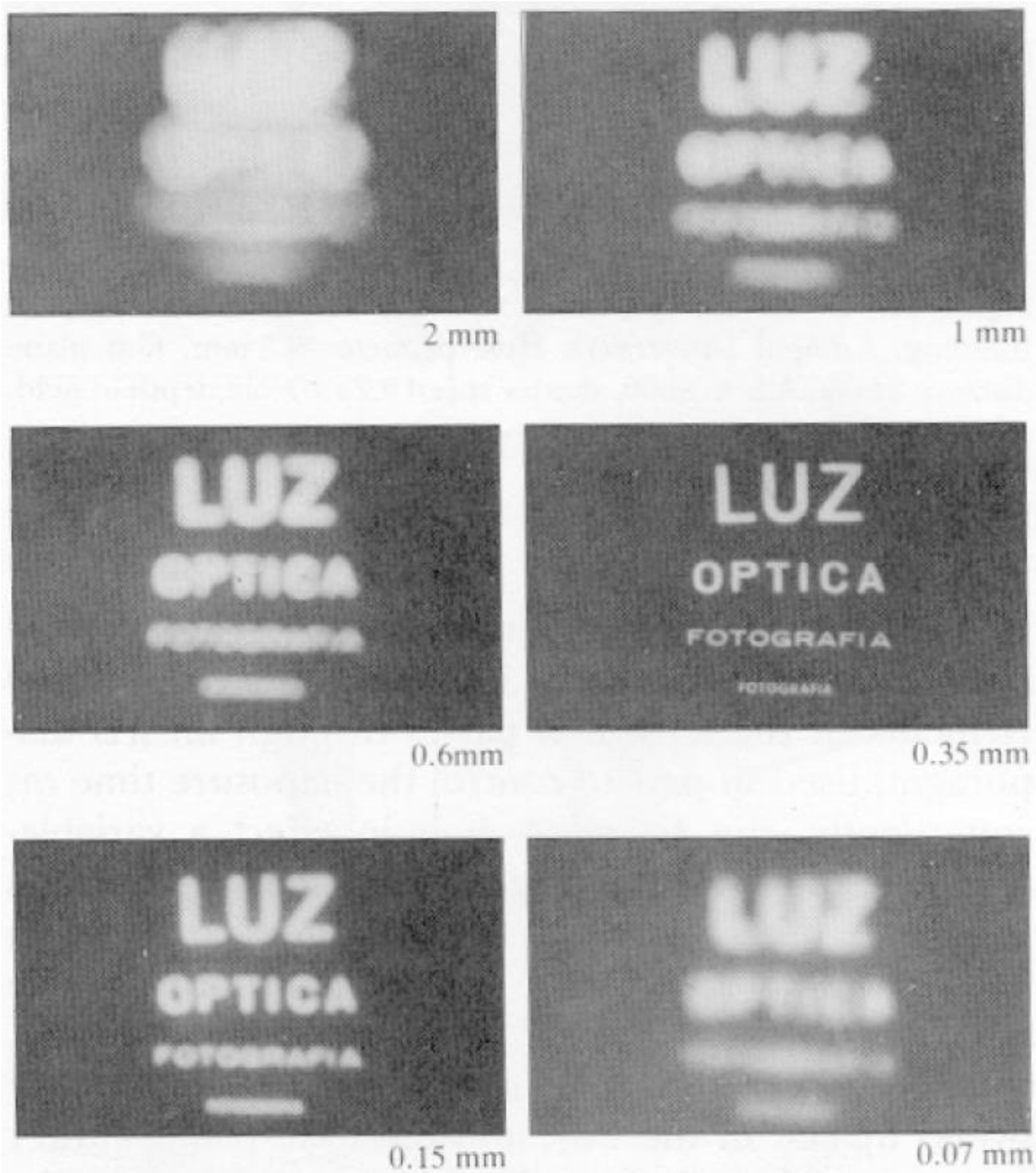
The projection matrix for orthographic projection

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

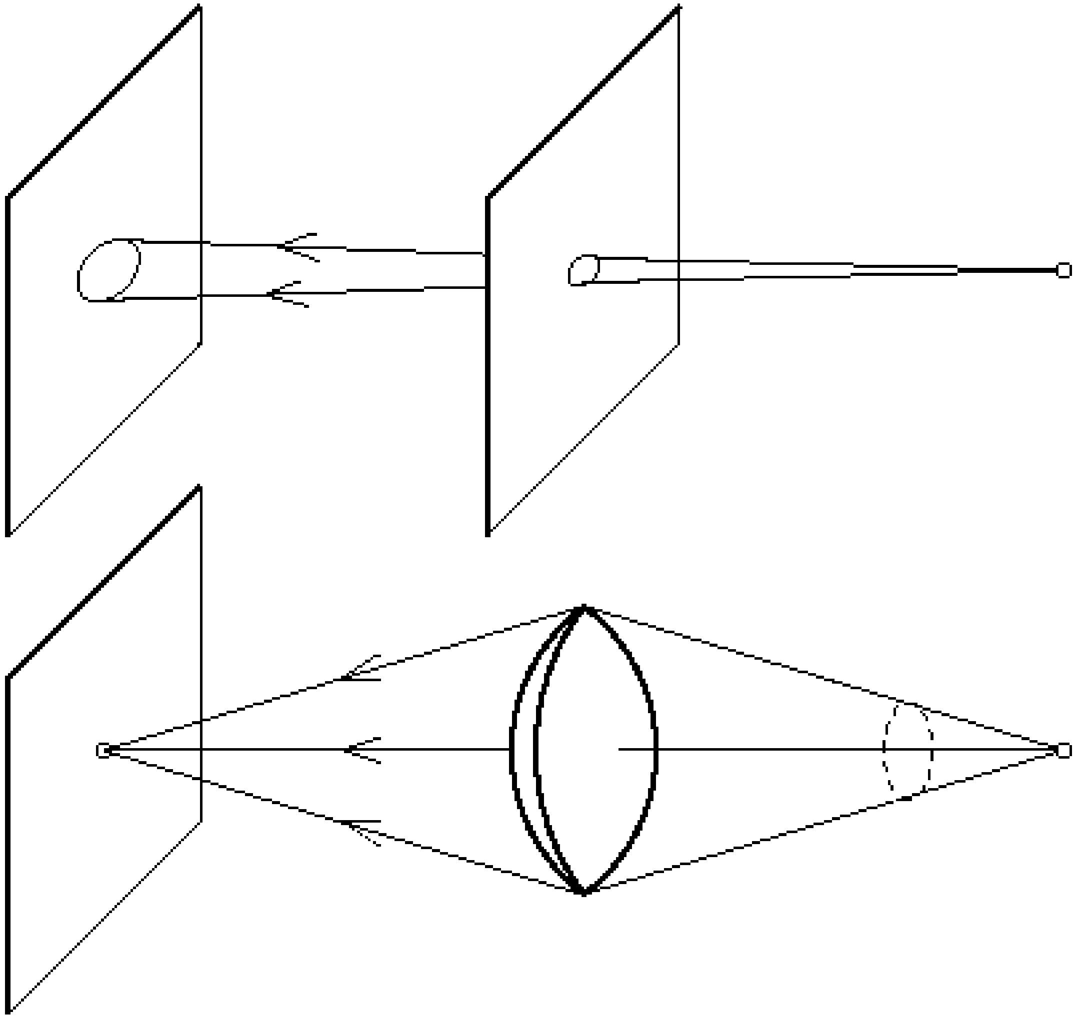
Pinhole too big - many directions are averaged, blurring the image

Pinhole too small- diffraction effects blur the image

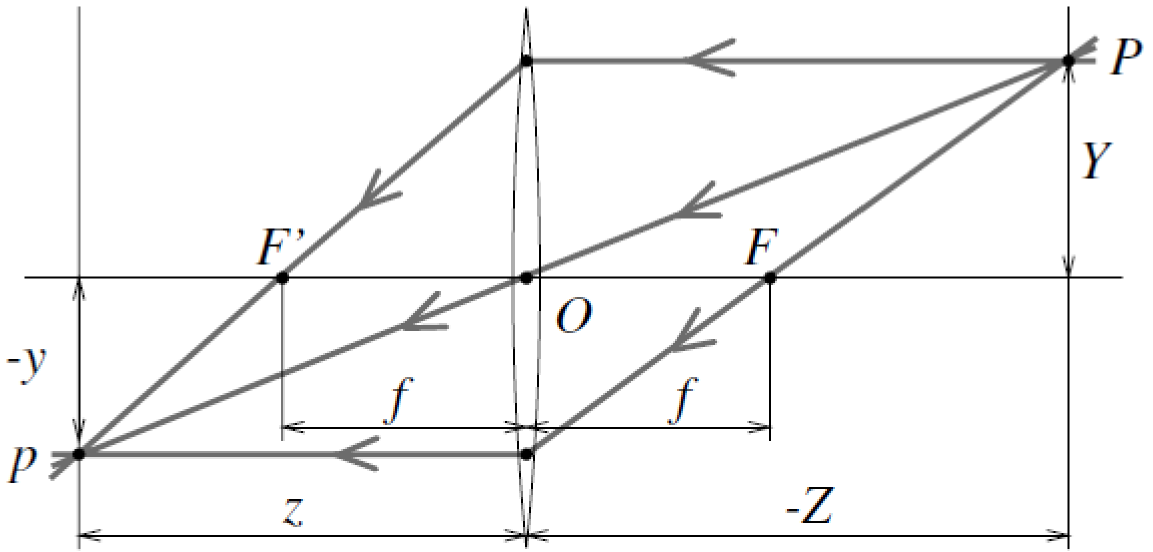
Generally, pinhole cameras are *dark*, because a very small set of rays from a particular point hits the screen.



The reason for lenses



The thin lens

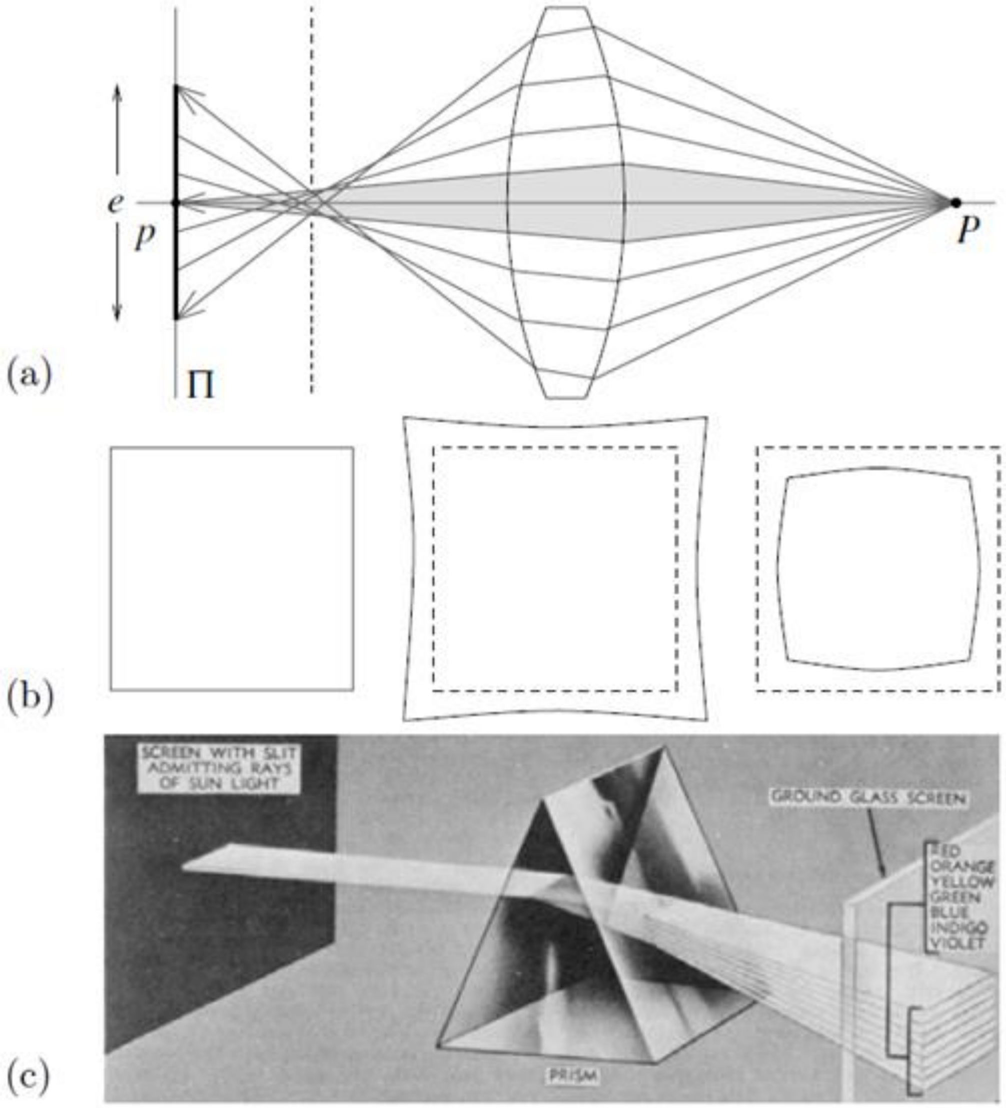


$$\frac{1}{z} - \frac{1}{Z} = \frac{1}{f}$$

$f = \frac{R}{2(n-1)}$ is the *focal length* of the lens

radius R and index of refraction n.

Spherical aberration



Spherical aberration: The gray region is the paraxial zone where the rays issued from P intersect at its paraxial image p. If an image plane π were erected in p, the image of p in that plane would form a circle of confusion of diameter e. The focus plane yielding the circle of least confusion is indicated by a dashed line.

Distortion: From left to right, the nominal image of a fronto-parallel square, pincushion distortion, and barrel distortion.

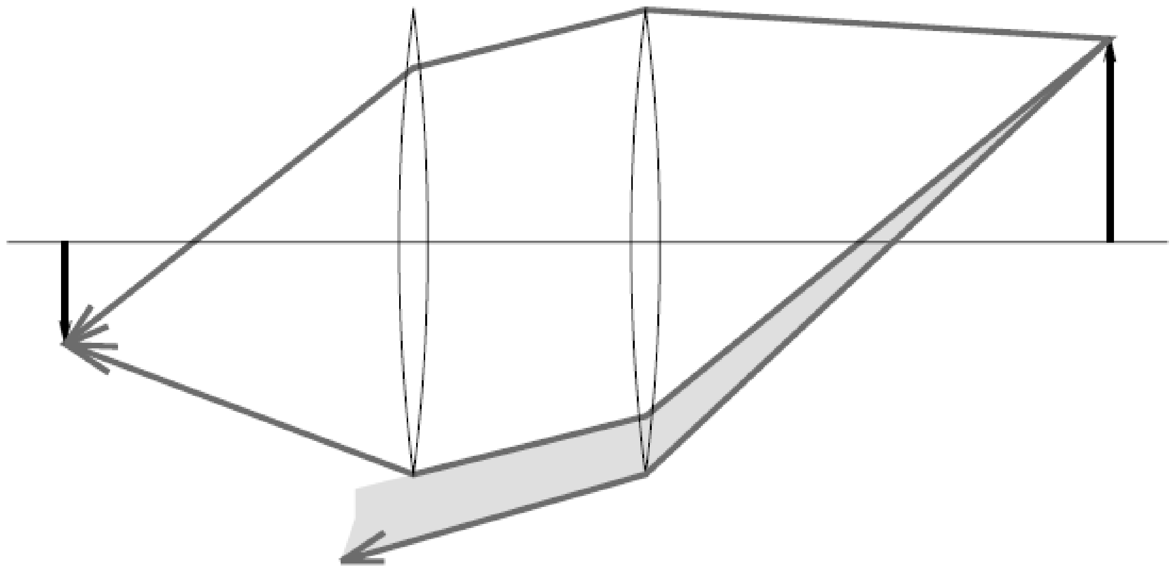
Chromatic aberration: The index of refraction of a transparent medium depends on the wavelength (or color) of the incident light rays. Here, a prism decomposes white light into a palette of colors. Figure from US NAVY MANUAL OF BASIC OPTICS AND OPTICAL INSTRUMENTS, prepared by the Bureau of Naval Perso```nnel, reprinted by Dover Publications, Inc. (1969).

Example of the usual photo of rose without changes and with chromatic aberrations (red-green edging):

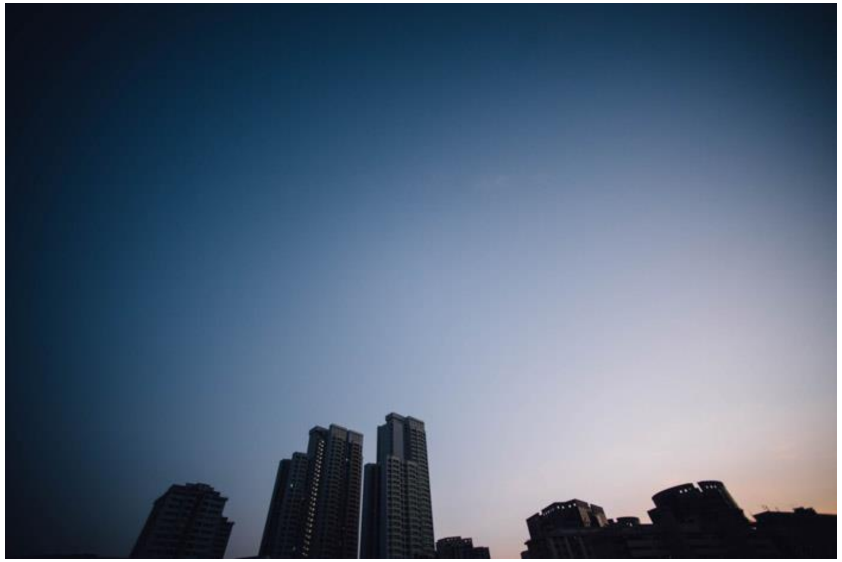
Barrel distortion

Pincushion distortion

Vignetting



Vignetting effect in a two-lens system. The shaded part of the beam never reaches the second lens. Additional apertures and stops in a lens further contribute to vignetting.



Other (possibly annoying) phenomena

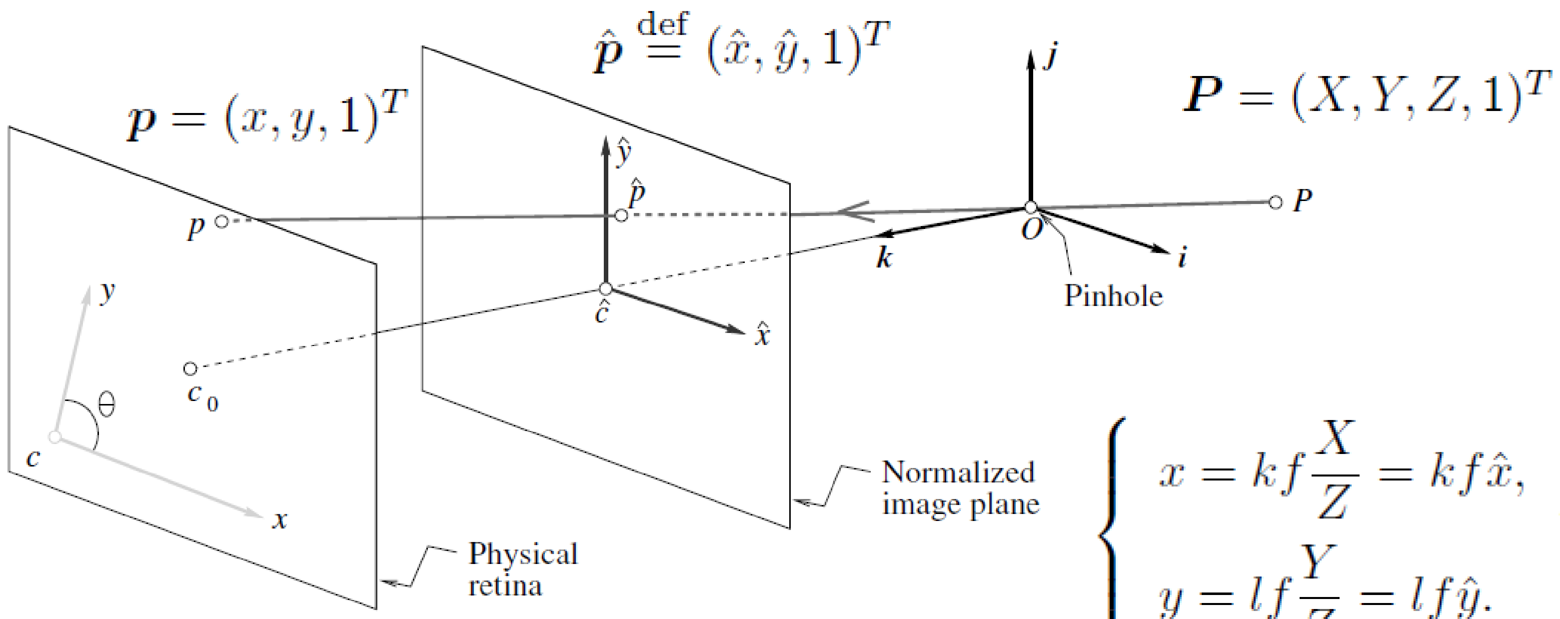
- Chromatic aberration
  - Light at different wavelengths follows different paths; hence, some wavelengths are defocussed
  - Machines: coat the lens
  - Humans: live with it
- Scattering at the lens surface
  - Some light entering the lens system is reflected off each surface it encounters (Fresnel's 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.)

INTRINSIC AND EXTRINSIC PARAMETERS

- The world and camera coordinate systems are related by a set of physical parameters: focal lens, the size of the pixels, the position of the image center, and the position and orientation of the camera.
- **Intrinsic parameters**, which relate the camera’s coordinate system to the idealized coordinate system.
- **Extrinsic parameters**, which relate the camera’s coordinate system to a fixed world coordinate system and specify its position and orientation in space.

Intrinsic parameters

$$p = \frac{1}{Z}\mathcal{M}P.$$



$$\begin{cases} \hat{x} = \dfrac{X}{Z} \\ \hat{y} = \dfrac{Y}{Z} \end{cases} \iff \hat{p} = \frac{1}{Z}(\text{Id} \quad 0)P,$$

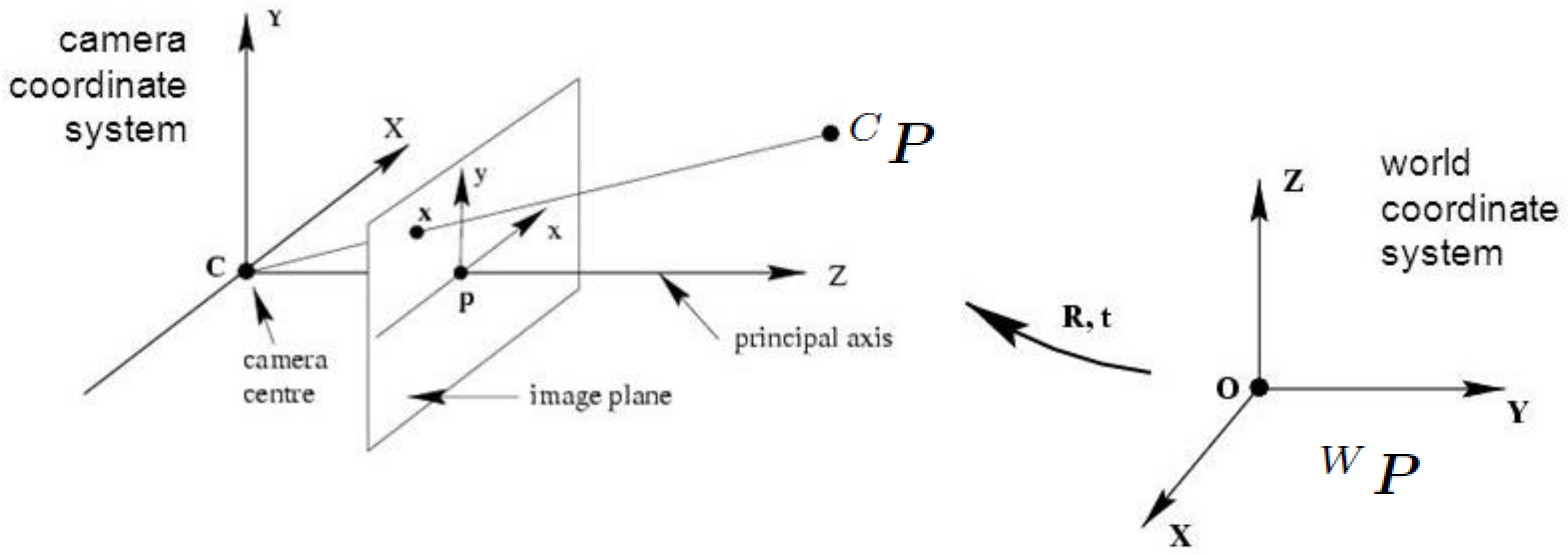
$$\begin{cases} x = kf\dfrac{X}{Z} = kf\hat{x}, \\ y = lf\dfrac{Y}{Z} = lf\hat{y}. \end{cases} \quad \begin{cases} x = \alpha\hat{x} + x_0, \\ y = \beta\hat{y} + y_0. \end{cases} \quad \begin{cases} x = \alpha\hat{x} - \alpha\cot\theta\hat{y} + x_0, \\ y = \dfrac{\beta}{\sin\theta}\hat{y} + y_0. \end{cases}$$

$$p = \frac{1}{Z}\mathcal{K}(\text{Id} \quad 0)P = \frac{1}{Z}\mathcal{M}P, \quad \text{where} \quad \mathcal{M} \stackrel{\text{def}}{=} (\mathcal{K} \quad 0)$$

$$\boxed{\mathcal{K} \stackrel{\text{def}}{=} \begin{pmatrix} \alpha & -\alpha\cot\theta & x_0 \\ 0 & \dfrac{\beta}{\sin\theta} & y_0 \\ 0 & 0 & 1 \end{pmatrix}}$$

Intrinsic matrix

Extrinsic parameters



Transformation from world coor. to image plane coor.

Extrinsic matric

$$p = \frac{1}{Z}\mathcal{M}\,{}^{C}P \qquad {}^{C}P = \begin{pmatrix} \mathcal{R} & t \\ 0^T & 1 \end{pmatrix} {}^{W}P$$

$$p = \boxed{\frac{1}{Z}\mathcal{M}}P, \quad \text{where} \quad \mathcal{M} = \mathcal{K}\boxed{\begin{pmatrix} \mathcal{R} & t \end{pmatrix}}$$

Camera parameters

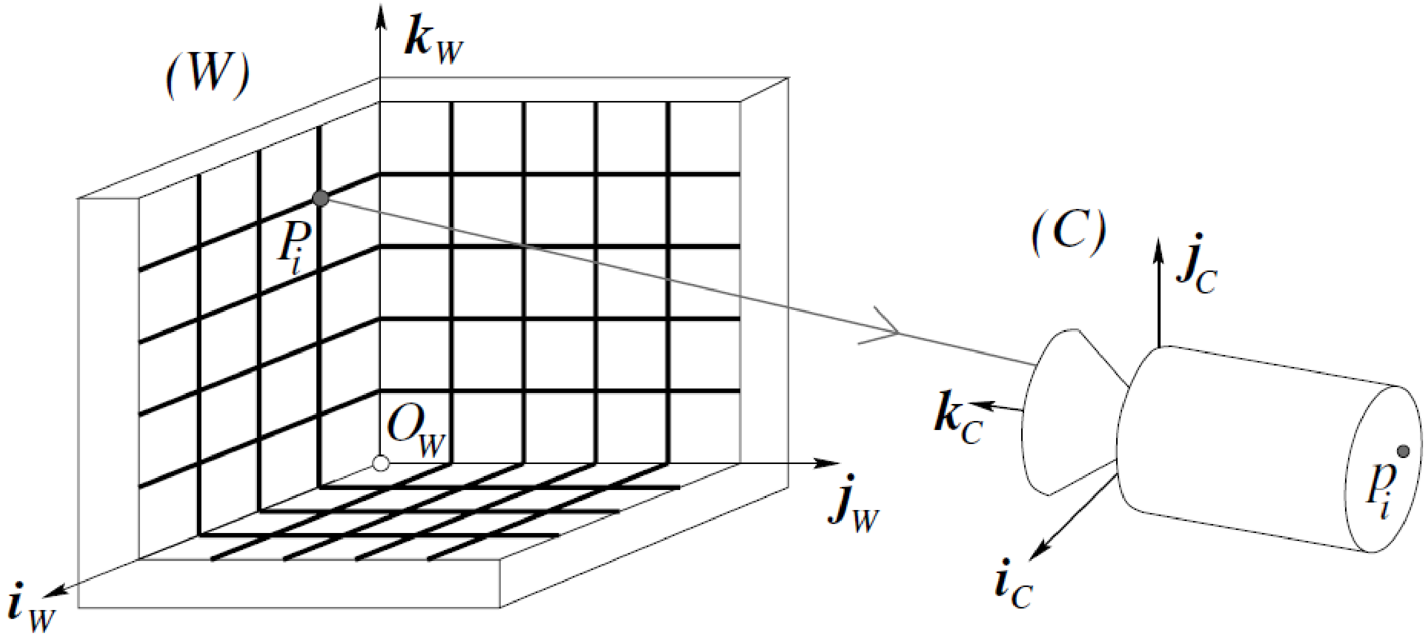
- There are 16 unknowns in intrinsic and extrinsic matrices
- Non-linear
- Estimating these parameters from experiments called camera calibration

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r1 & r2 & r3 & t1 \\ r4 & r5 & r6 & t2 \\ r7 & r8 & r9 & t3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Camera calibration

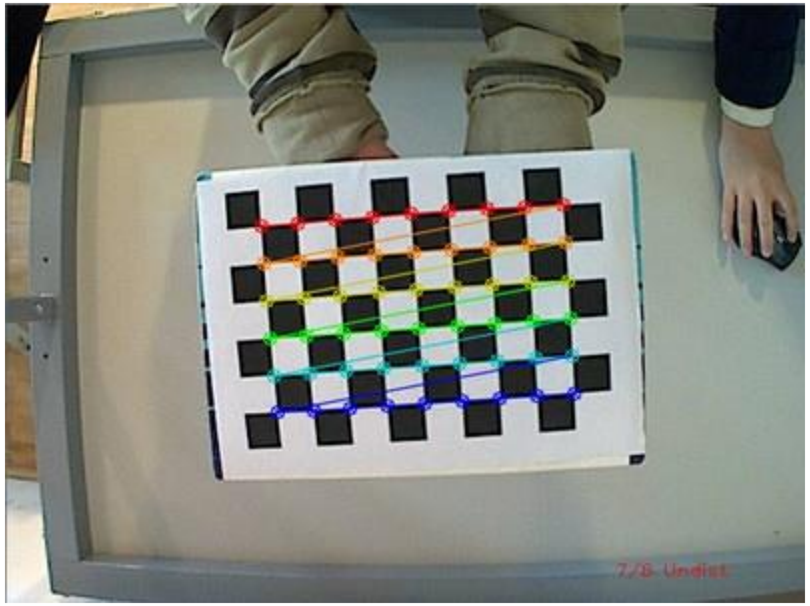
- Issues:
  - what are intrinsic parameters of the camera?
  - what is the camera matrix? (intrinsic+extrinsic)
- General strategy:
  - view calibration object
  - identify image points
  - obtain camera matrix by minimizing error
  - obtain intrinsic parameters from camera matrix
- Error minimization: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

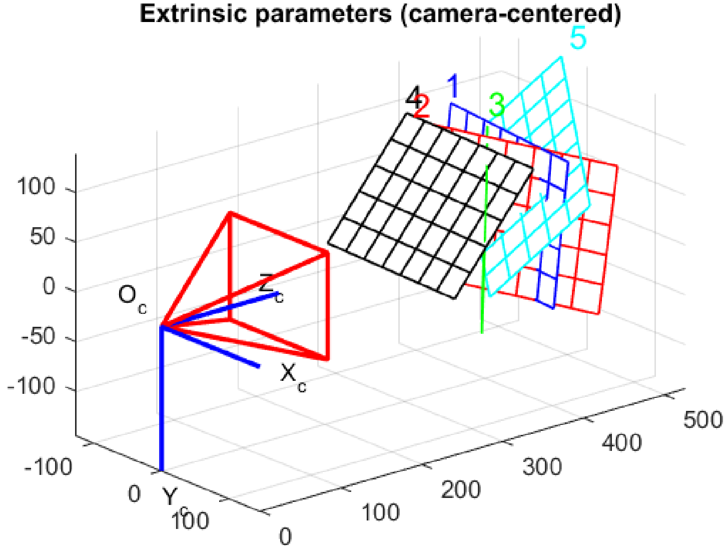
Camera calibration

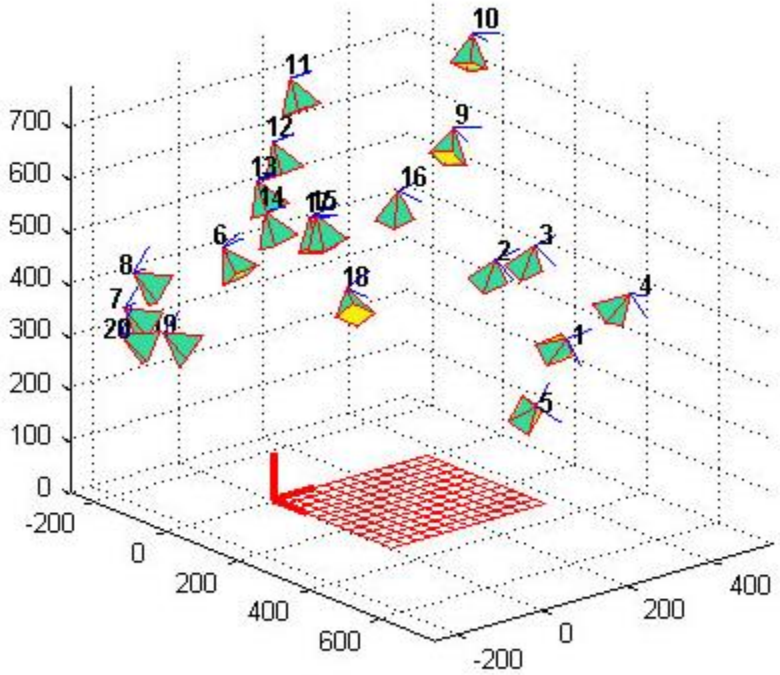


Camera calibration setup: In this example, the calibration rig is formed by three grids drawn in orthogonal planes. Other patterns could be used as well, and they may involve lines or other geometric figures.

Camera calibration







The chessboard is at the origin

Camera calibration

- Do exercise camera calibration with OpenCV
- https://docs.opencv.org/2.4/doc/tutorials/calib3d/camera_calibration/camera_calibration.html