Computer Vision

Cameras & Calibration

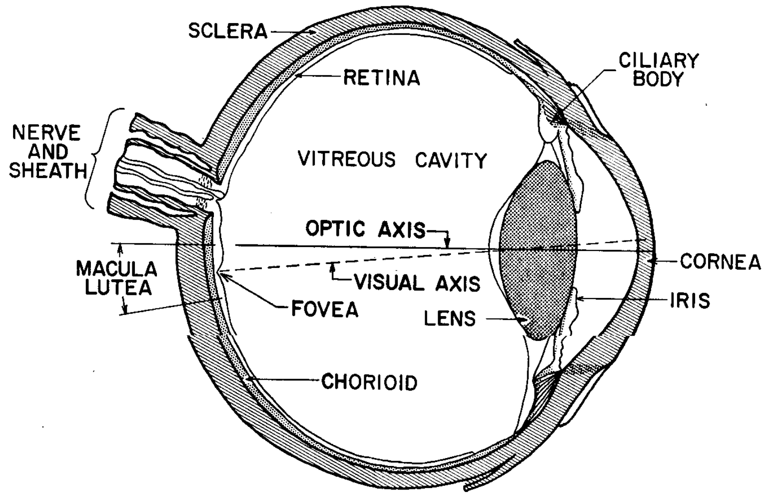
Cameras, lenses, and sensors

Figure 1.16 The first photograph on record, *la table servie*, obtained by Nicéphore Niepce in 1822. *Collection Harlinge–Viollet.*

- Pinhole cameras
- Lenses
- Projection models
- Geometric camera parameters

From <u>Computer Vision</u>, Forsyth and Ponce, Prentice-Hall, 2002.

Reproduced by permission, the American Society of Photogrammetry and Remote Sensing. A.L. Nowicki, “Stereoscopy.” Manual of Photogrammetry, Thompson, Radlinski, and Speert (eds.), third edition, 1966.



Animal eye: a looonnng time ago.

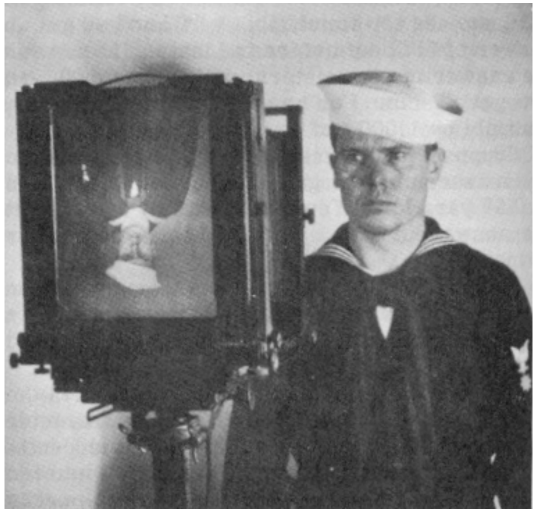
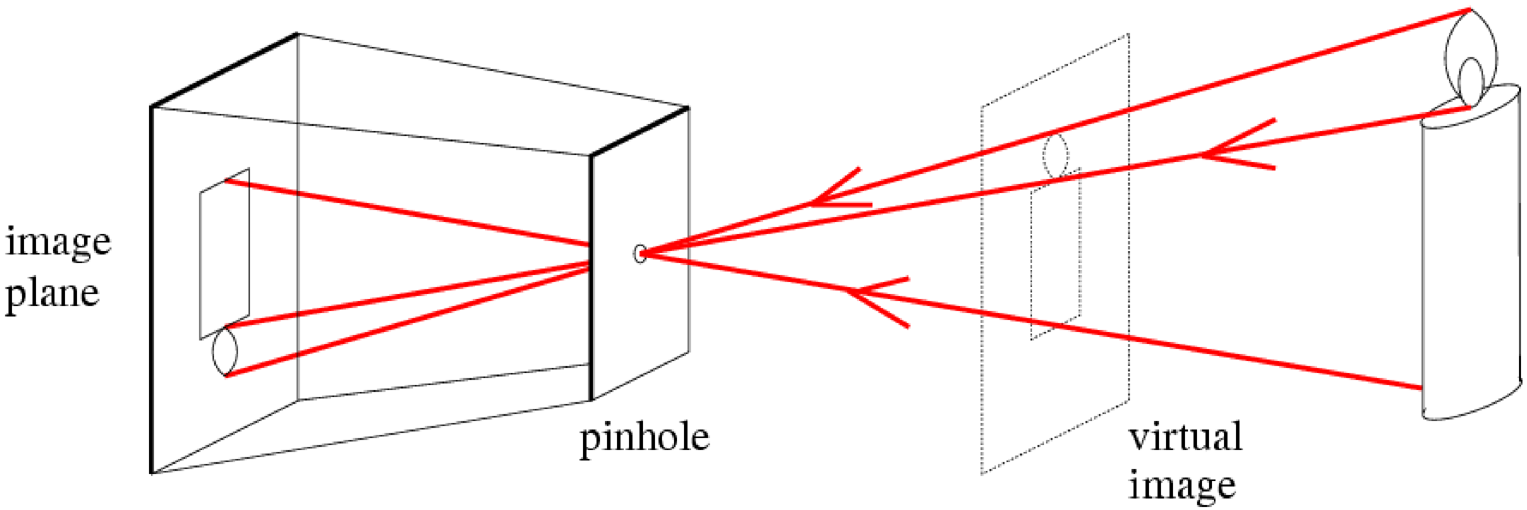


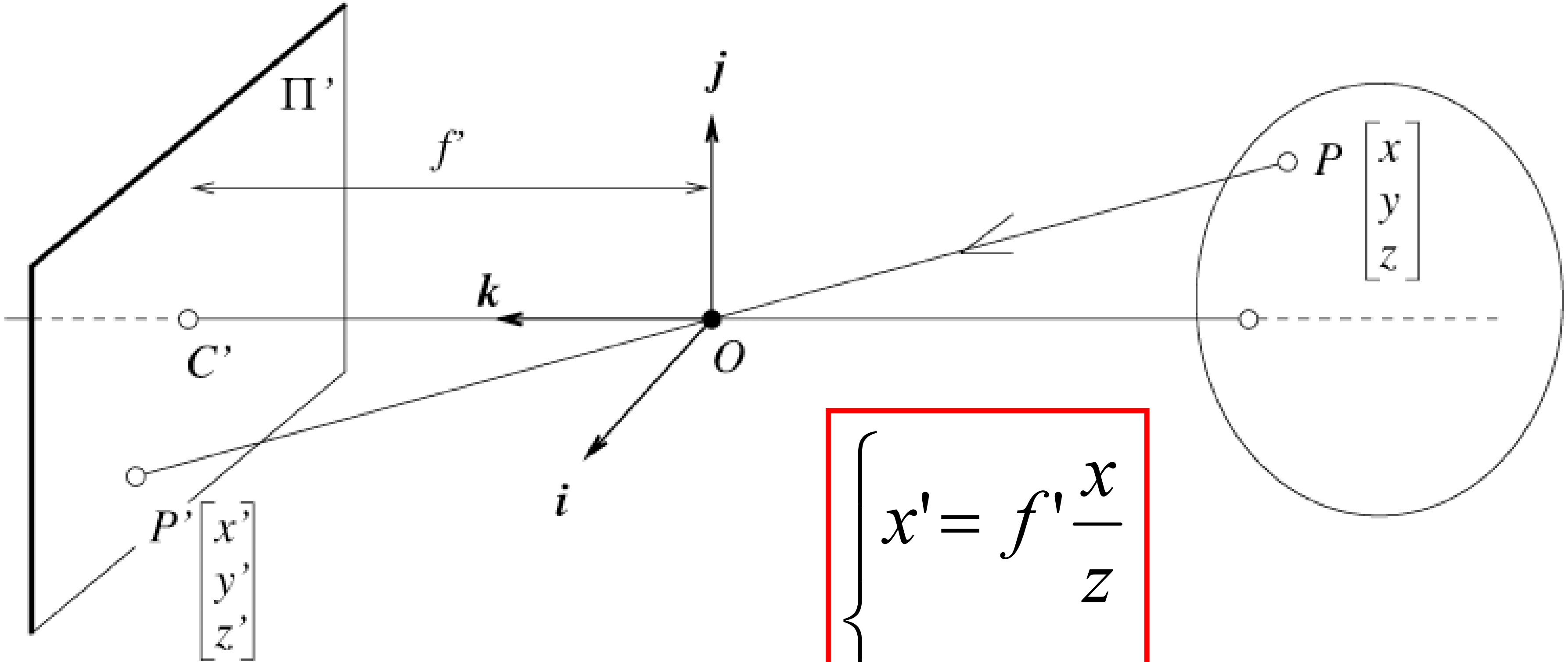
Figure from US Navy Manual of Basic Optics and Optical Instruments, prepared by Bureau of Naval Personnel. Reprinted by Dover Publications, Inc., 1969.

Photographic camera: Niepce, 1816.



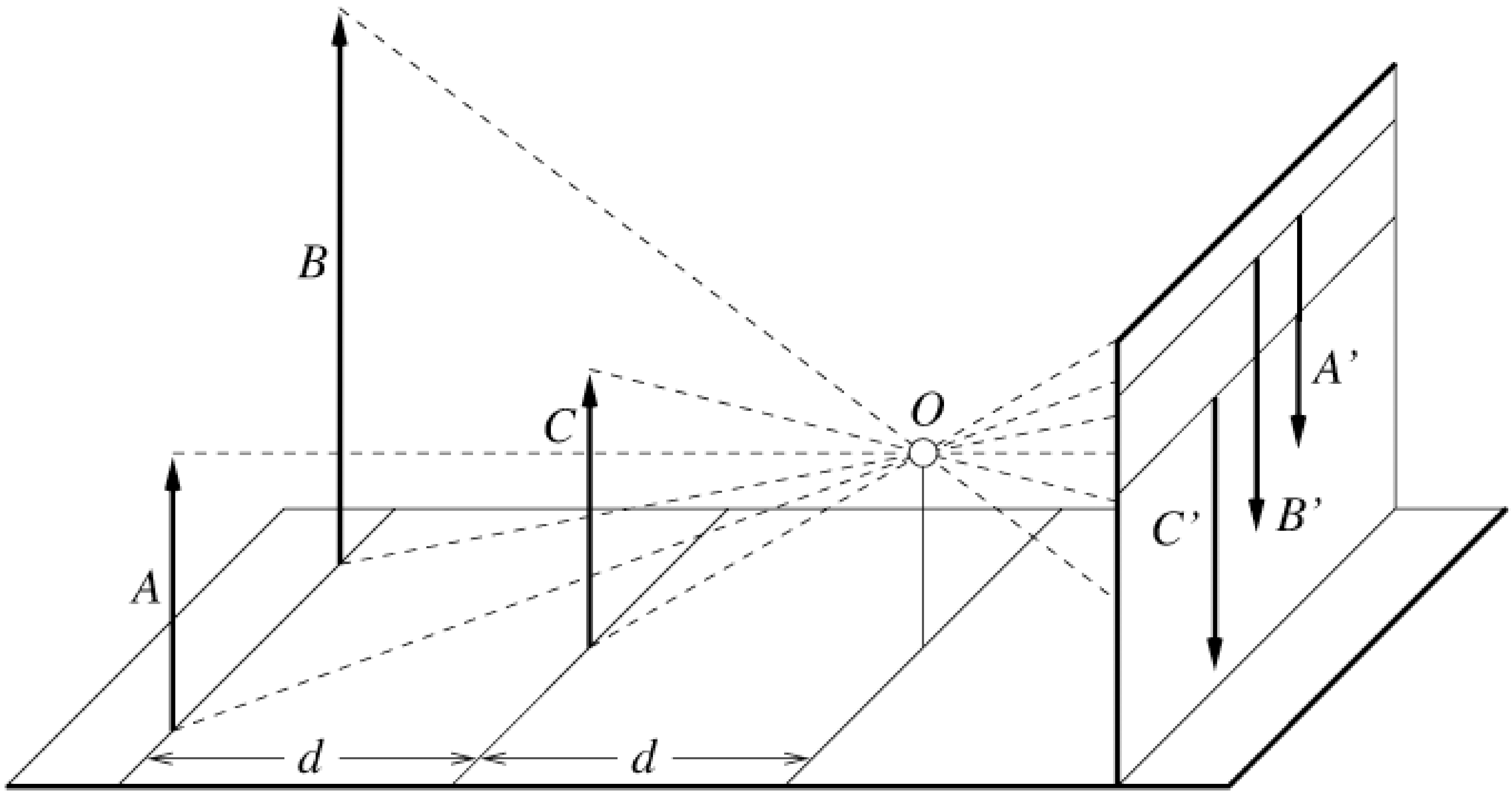
Pinhole perspective projection: Brunelleschi, XVth Century.
Camera obscura: XVIth Century.

The equation of projection



$$\begin{cases} x' = f'\dfrac{x}{z} \\ y' = f'\dfrac{y}{z} \end{cases}$$

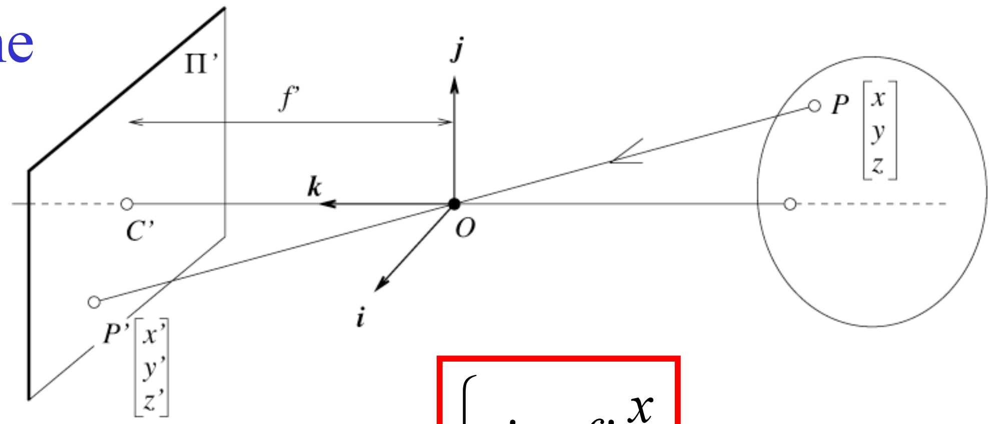
Distant objects are smaller



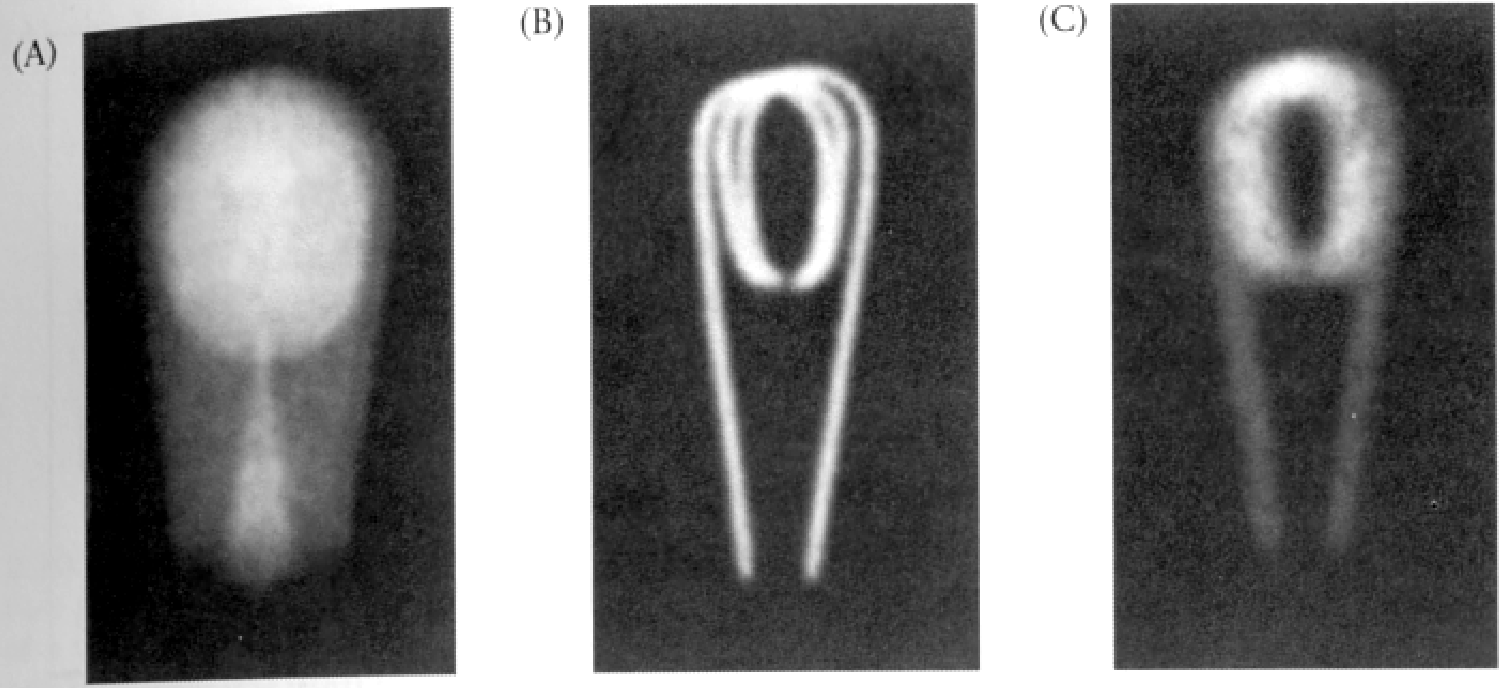
Forsyth&Ponce

Geometric properties of projection

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



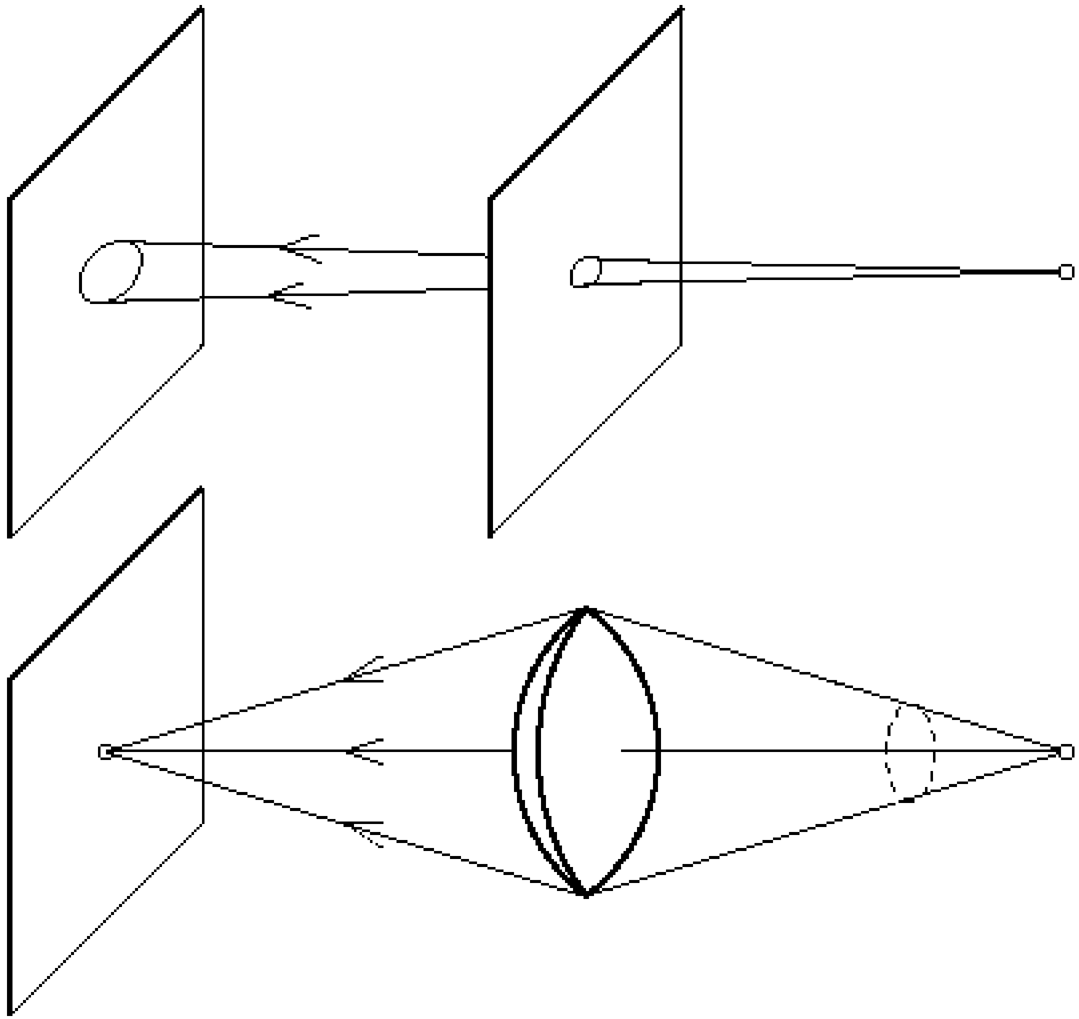
$$\begin{cases} x' = f'\dfrac{x}{z} \\[2ex] y' = f'\dfrac{y}{z} \end{cases}$$



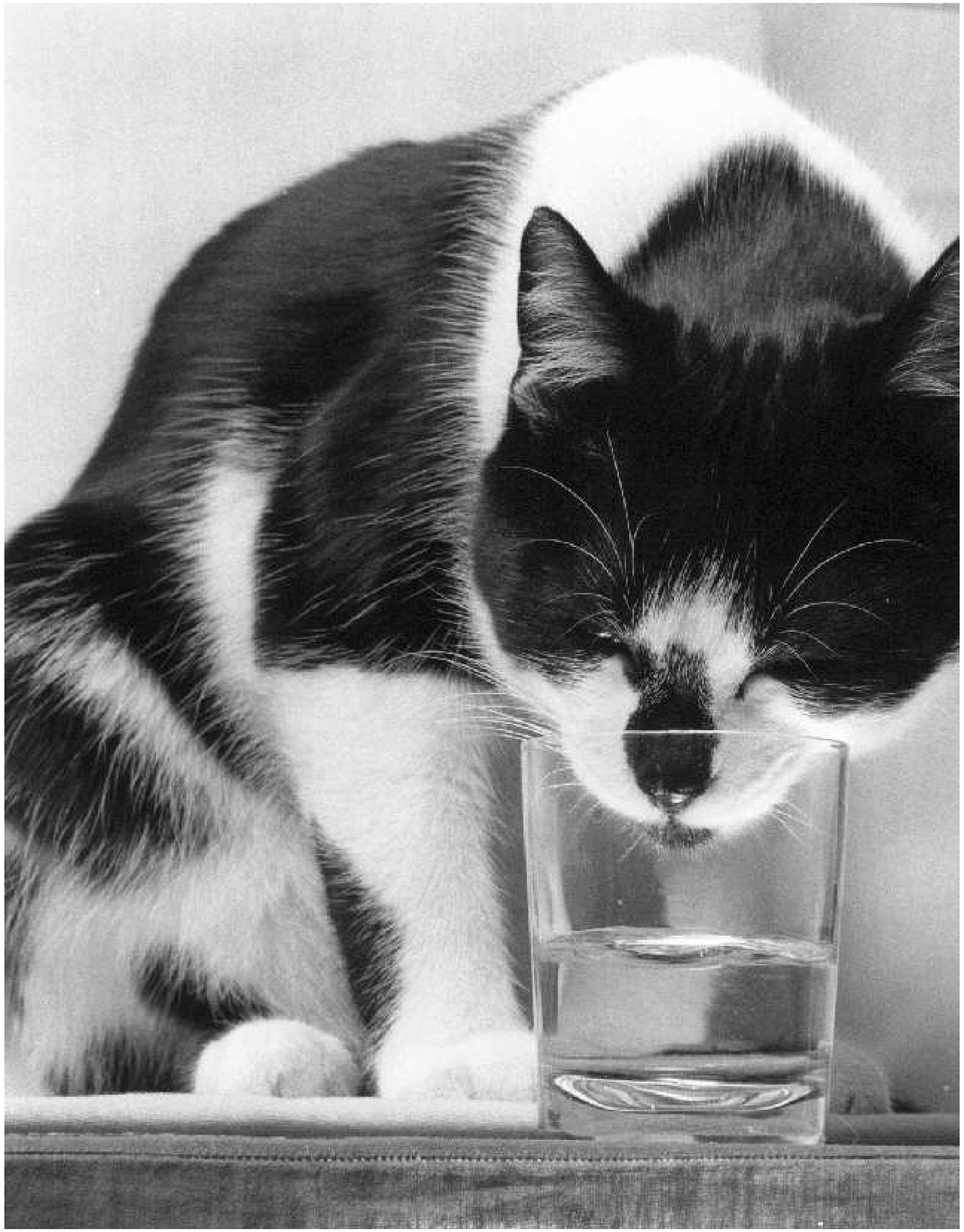
2.18 DIFFRACTION LIMITS THE QUALITY OF PINHOLE OPTICS. These three images of a bulb filament were made using pinholes with decreasing size. (A) When the pinhole is relatively large, the image rays are not properly converged, and the image is blurred. (B) Reducing the size of the pinhole improves the focus. (C) Reducing the size of the pinhole further worsens the focus, due to diffraction. From Ruechardt, 1958.

Wandell, Foundations of Vision, Sinauer, 1995

The reason for lenses

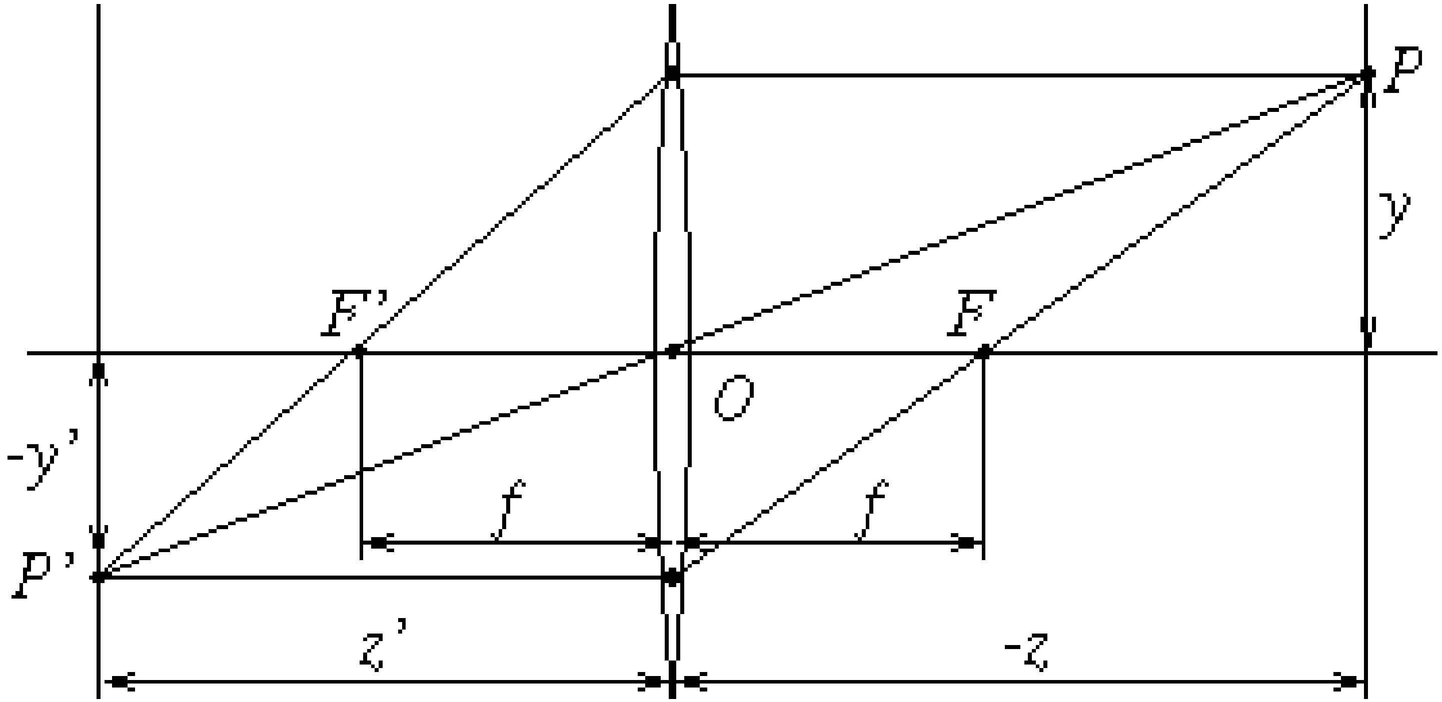


Water glass refraction



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

The thin lens, first order optics



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

All rays through P also pass through P', but only for points at -z: "*depth of field*".

Forsyth&Ponce

More accurate models of real lenses

- Finite lens thickness
- Higher order approximation to $\sin(\theta)$
- Chromatic aberration
- Vignetting

Thick lens

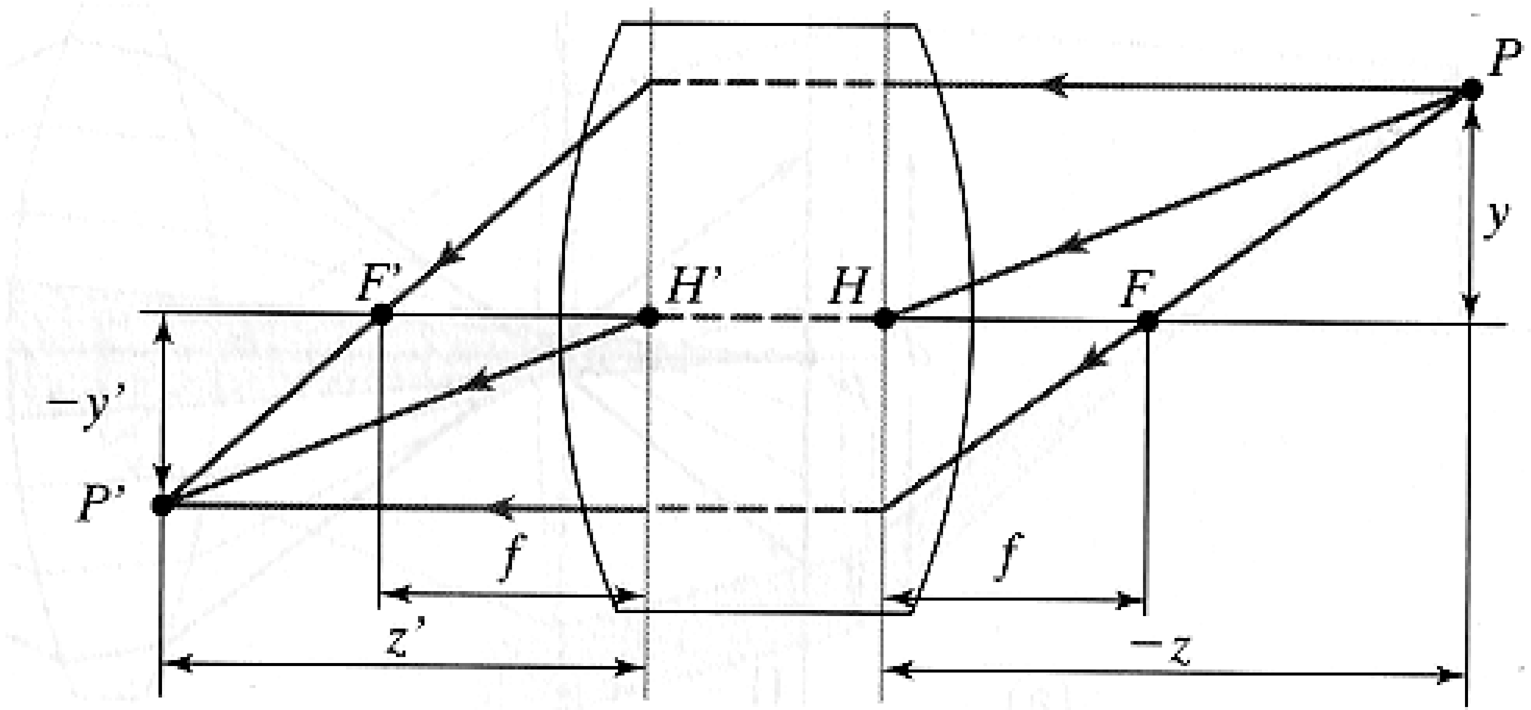
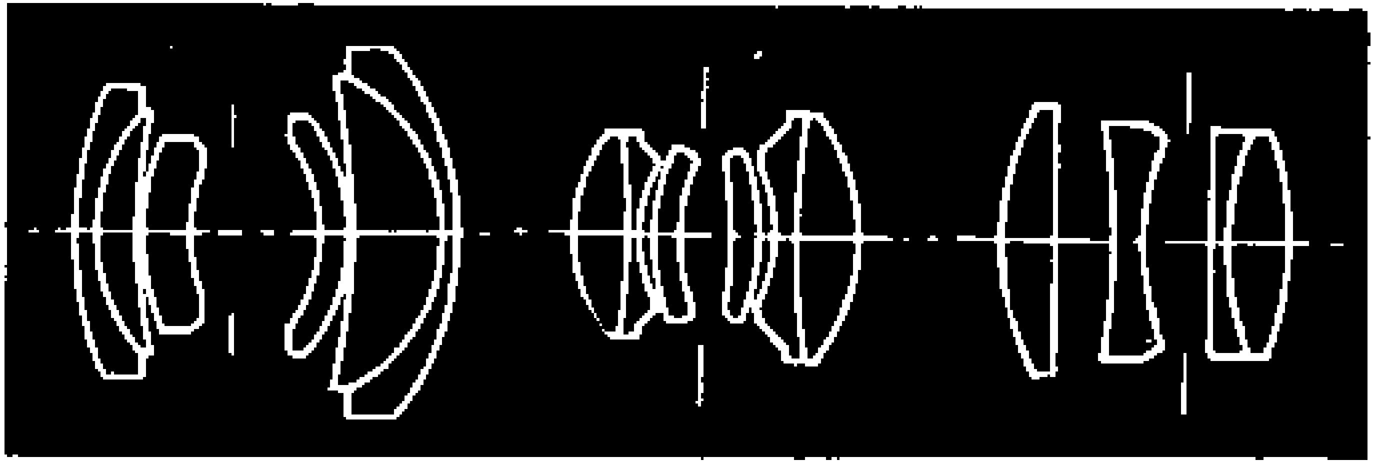


Figure 1.11 A simple thick lens with two spherical surfaces.

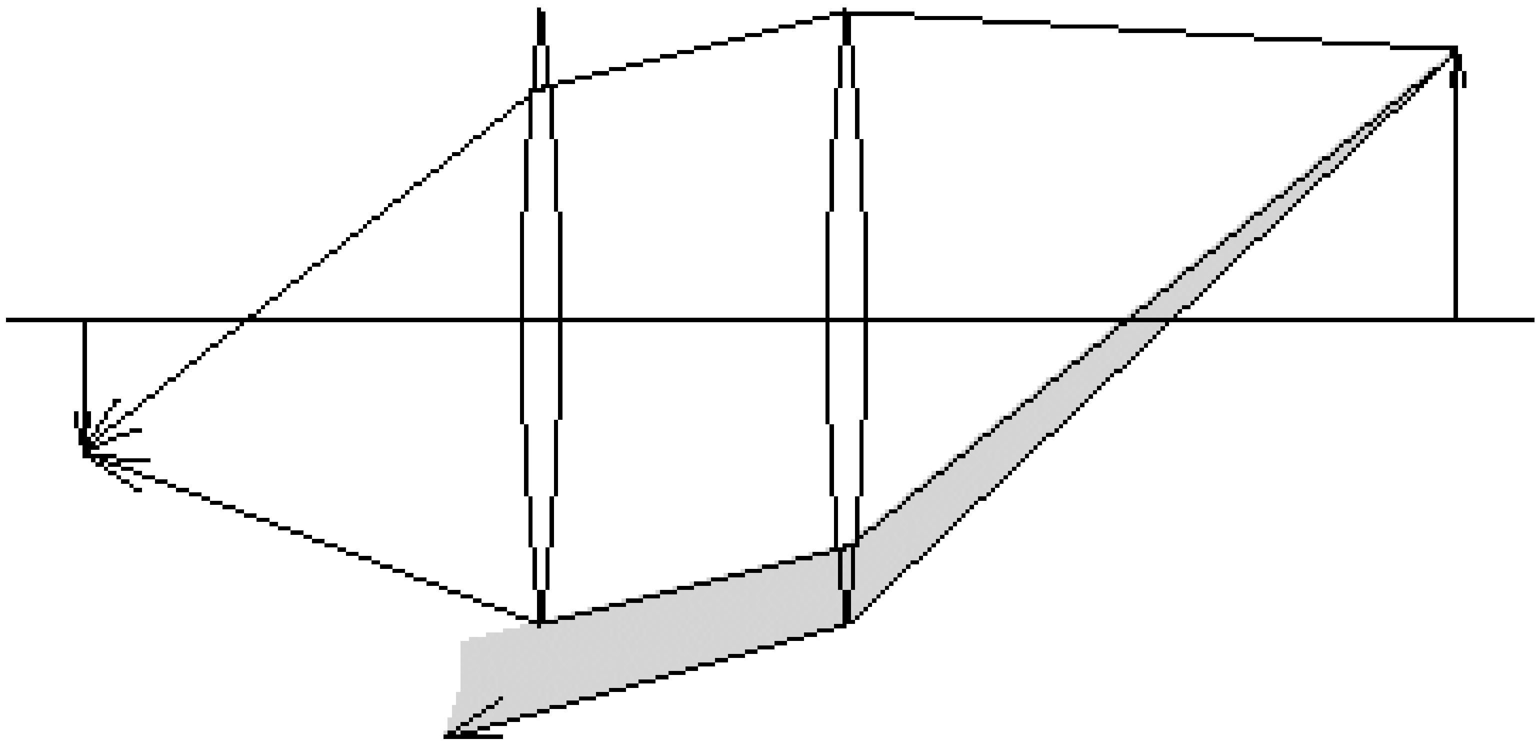
Forsyth&Ponce

Lens systems



Lens systems can be designed to correct for aberrations described by 3rd order optics

Vignetting



Chromatic aberration

(great for prisms, bad for lenses)

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)

Summary so far

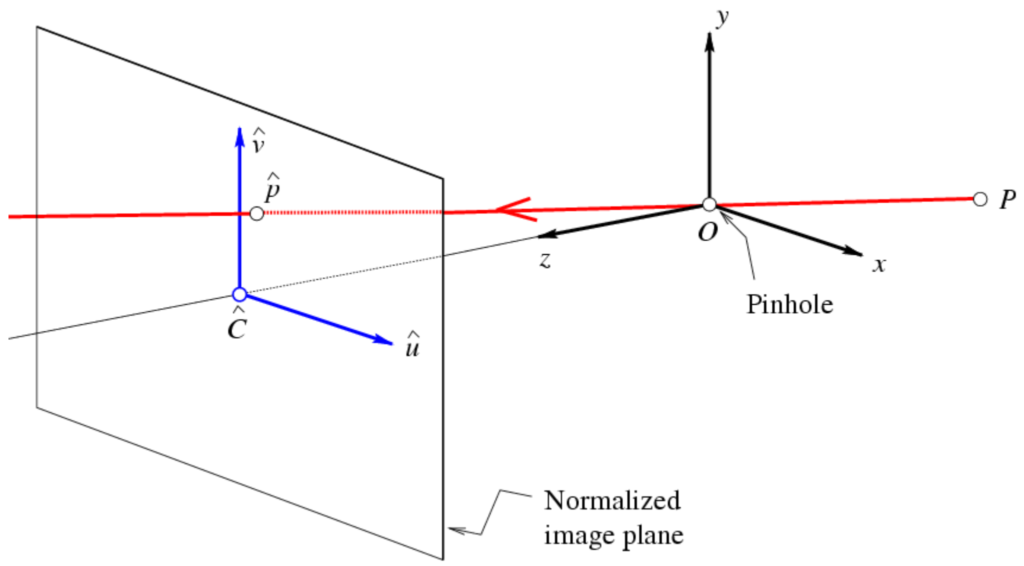
- Want to make images
- Pinhole camera models the geometry of perspective projection
- Lenses make it work in practice
- Models for lenses
  - Thin lens, spherical surfaces, first order optics
  - Thick lens, higher-order optics, vignetting.

Camera calibration

Use the camera to tell you things about the world:

- Relationship between coordinates in the world and coordinates in the image: *geometric camera calibration.*
- (Relationship between intensities in the world and intensities in the image: *photometric camera calibration*, not covered in this course

Intrinsic parameters



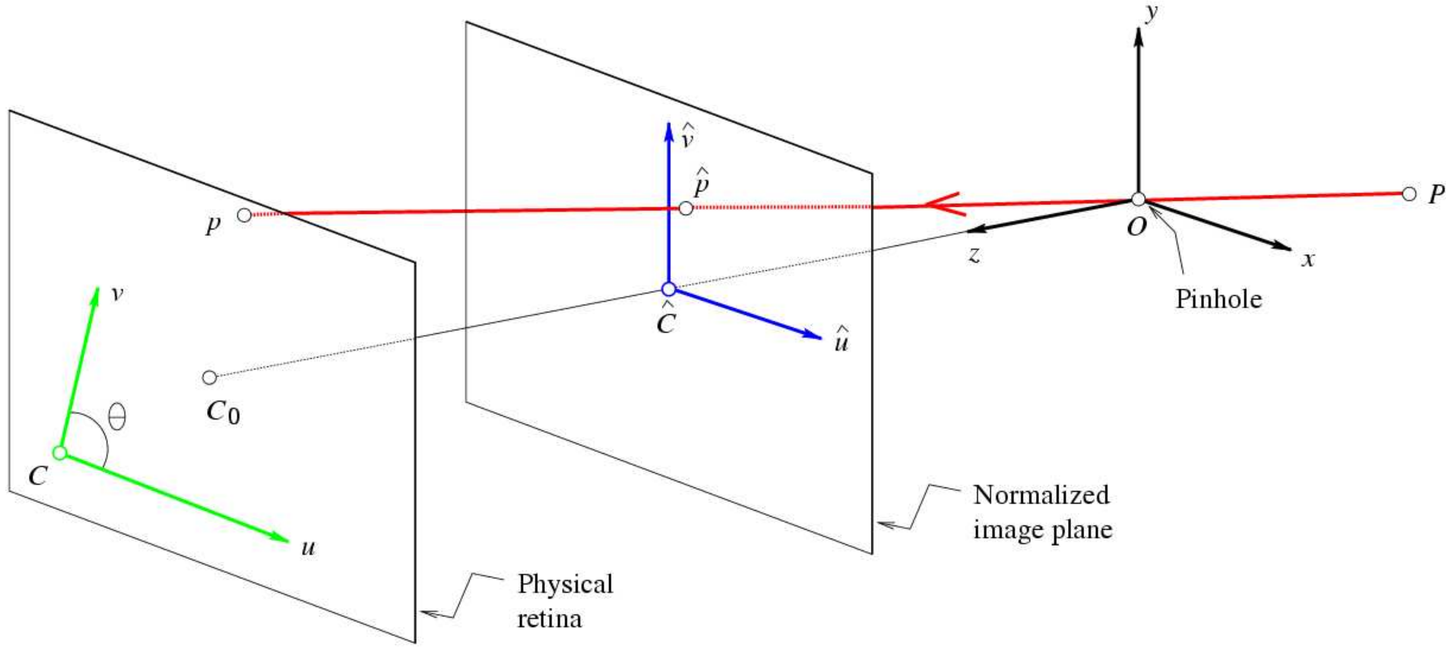
Forsyth&Ponce

Perspective projection

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

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

Intrinsic parameters

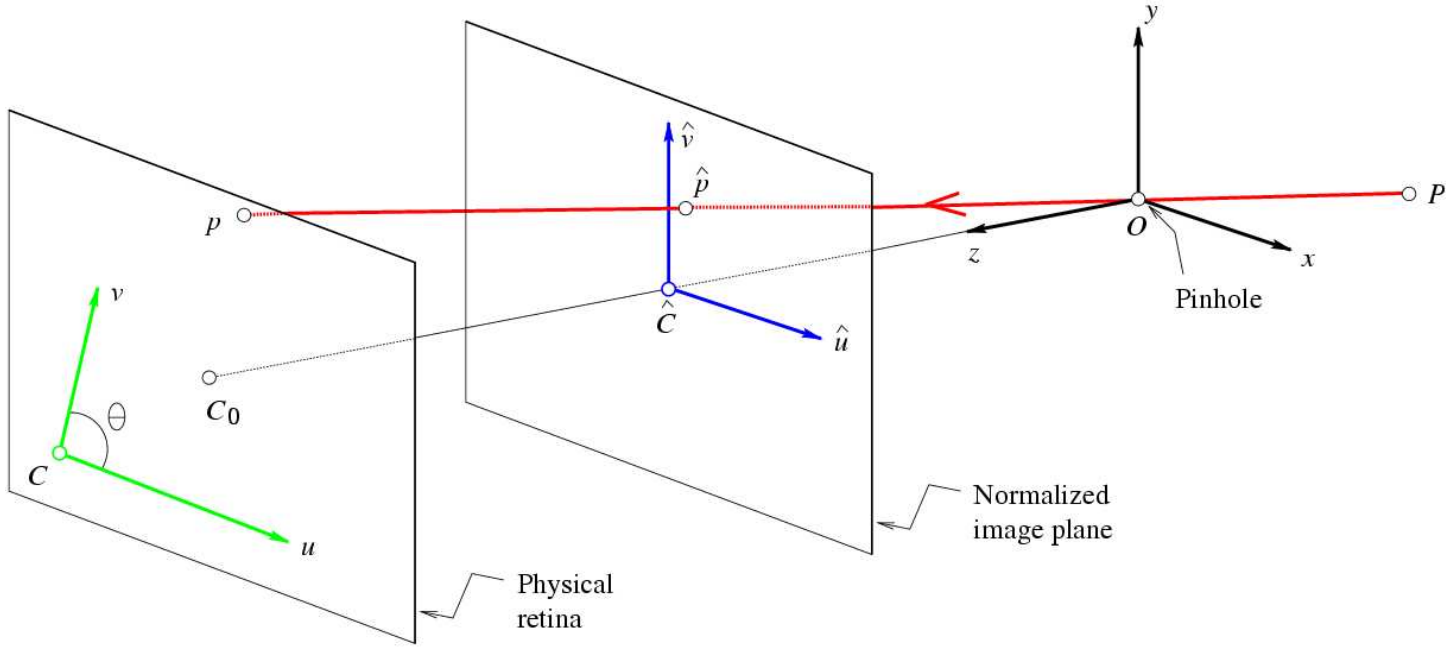


But "pixels" are in some arbitrary spatial units…

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

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

Intrinsic parameters

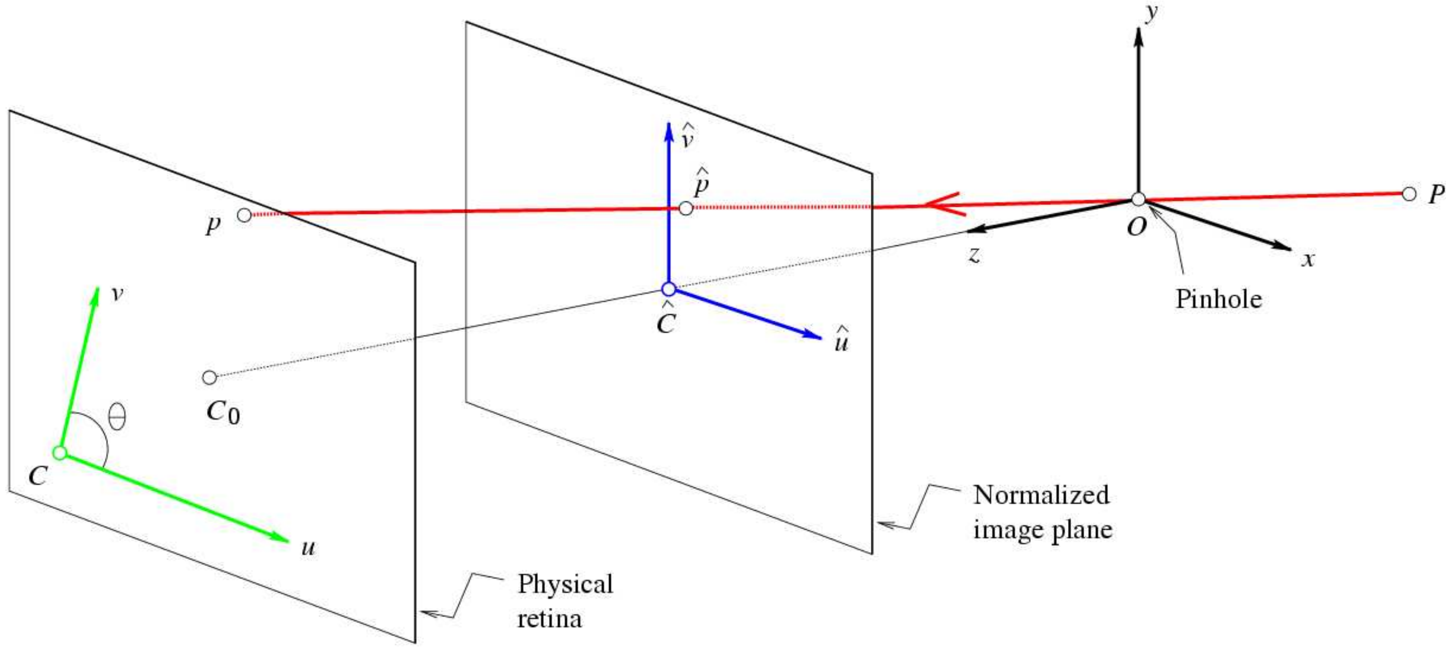


But "pixels" are in some arbitrary spatial units

$$u = \alpha \frac{x}{z}$$

$$v = \alpha \frac{y}{z}$$

Intrinsic parameters

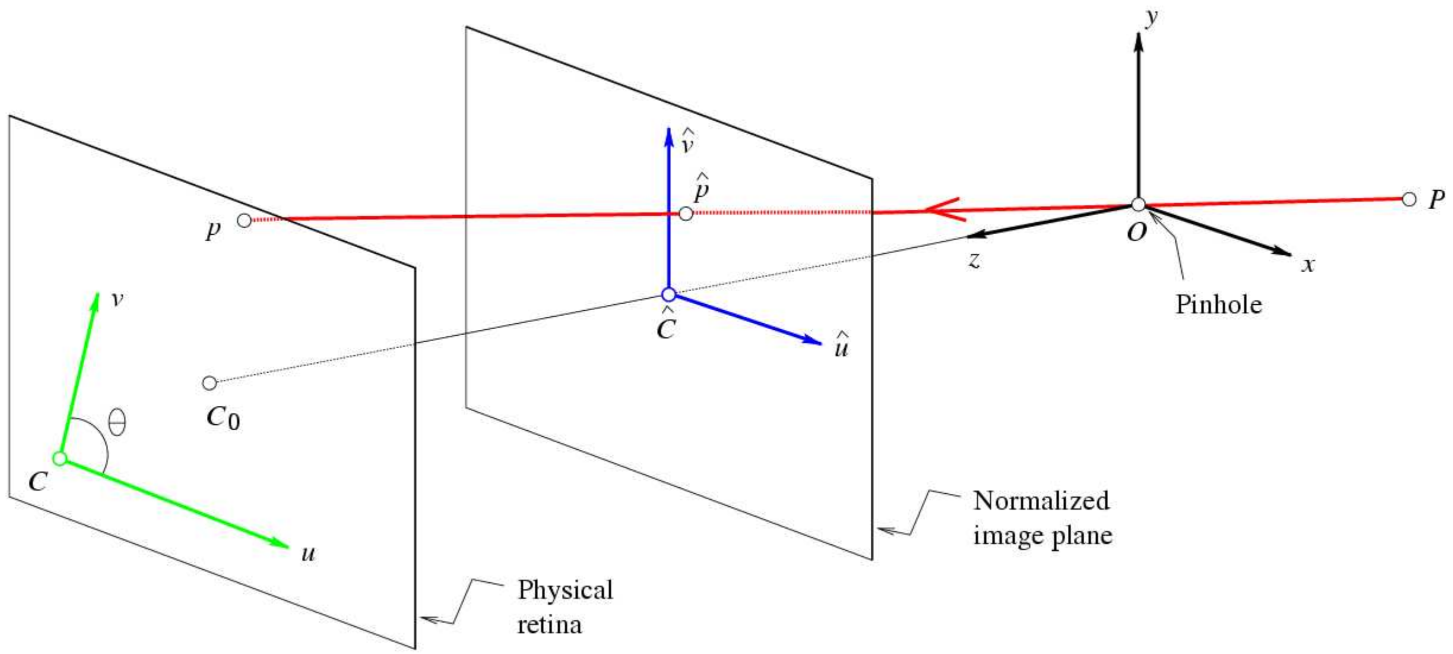


Maybe pixels are not square…

$$u = \alpha \frac{x}{z}$$

$$v = \alpha \frac{y}{z}$$

Intrinsic parameters

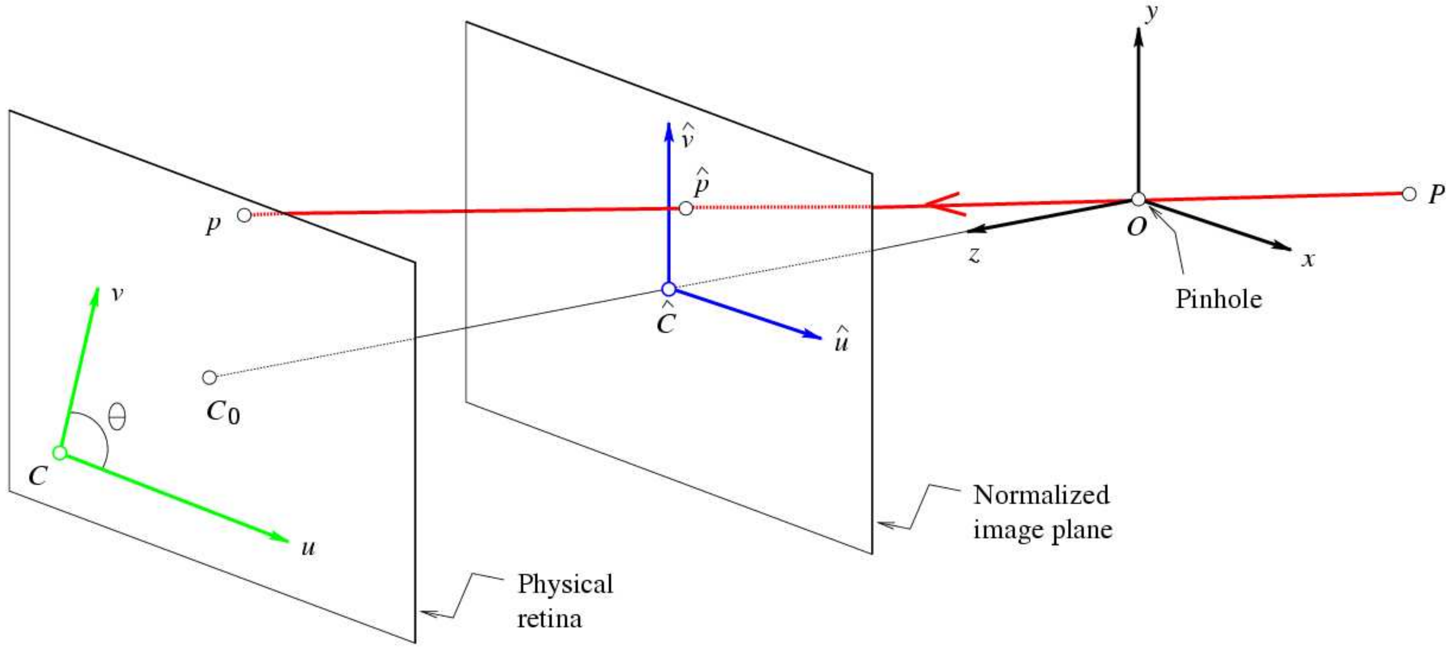


Maybe pixels are not square

$$u = \alpha \frac{x}{z}$$

$$v = \beta \frac{y}{z}$$

Intrinsic parameters

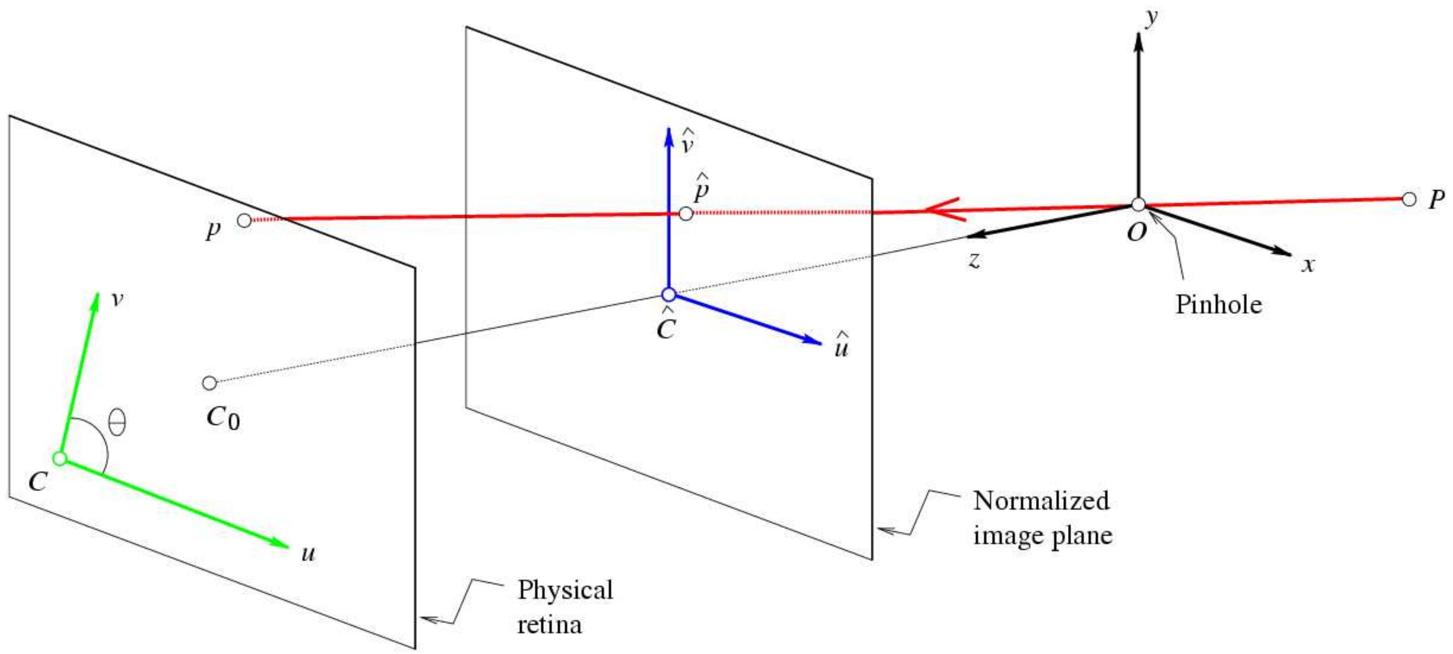


We don't know the origin of our camera pixel coordinates…

$$u = \alpha \frac{x}{z}$$

$$v = \beta \frac{y}{z}$$

Intrinsic parameters

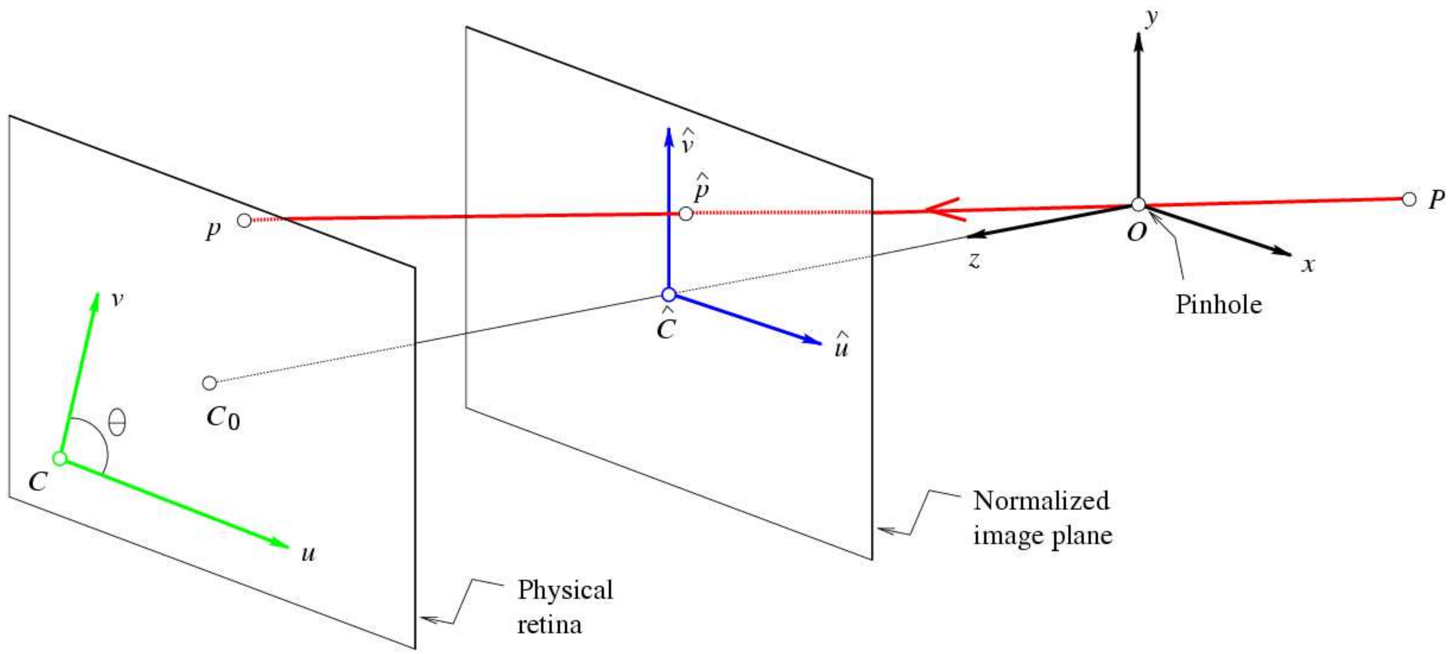


We don’t know the origin of our camera pixel coordinates

$$u = \alpha \frac{x}{z} + u_0$$

$$v = \beta \frac{y}{z} + v_0$$

Intrinsic parameters

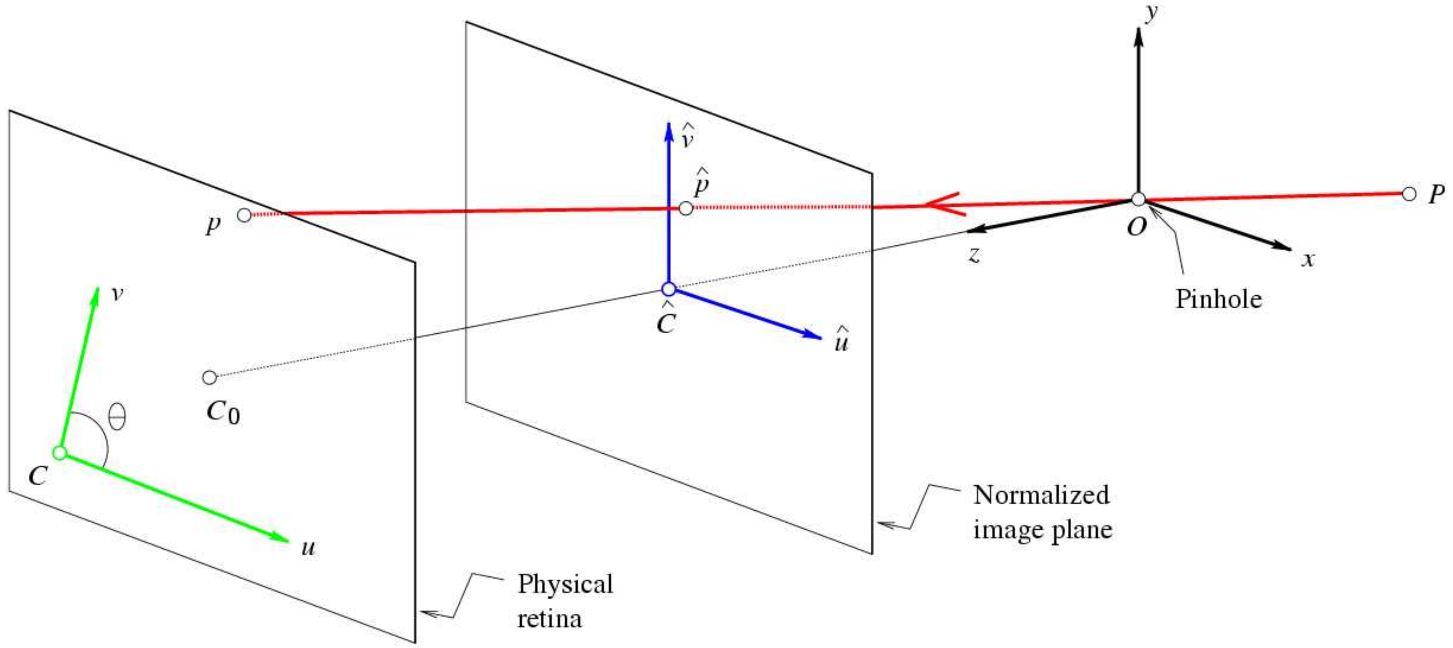


May be skew between camera pixel axes…

$$u = \alpha \frac{x}{z} + u_0$$

$$v = \beta \frac{y}{z} + v_0$$

Intrinsic parameters

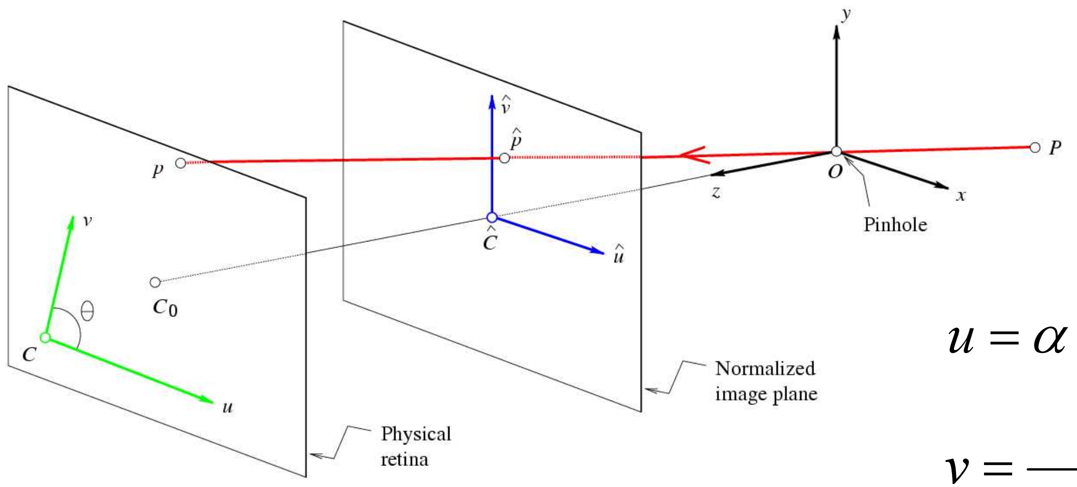


May be skew between camera pixel axes

$$u = \alpha \frac{x}{z} - \alpha \cot(\theta) \frac{y}{z} + u_0$$

$$v = \frac{\beta}{\sin(\theta)} \frac{y}{z} + v_0$$

Intrinsic parameters



$$u = \alpha \frac{x}{z} - \alpha \cot(\theta) \frac{y}{z} + u_0$$

$$v = \frac{\beta}{\sin(\theta)} \frac{y}{z} + v_0$$

Using homogenous coordinates,
we can write this as:

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \frac{1}{z} \begin{pmatrix} \alpha & -\alpha \cot(\theta) & u_0 & 0 \\ 0 & \dfrac{\beta}{\sin(\theta)} & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

or:

$$\boxed{\vec{p} = \frac{1}{z} \begin{pmatrix} K & \vec{0} \end{pmatrix} \vec{P}}$$

Extrinsic parameters: translation and rotation of camera frame

$$^{C}P = {}^{C}_{W}R\ {}^{W}P + {}^{C}O_{W}$$

Non-homogeneous coordinates

$$\begin{pmatrix} C_X \\ C_Y \\ C_Z \\ 1 \end{pmatrix} = \begin{pmatrix} - & - & - & | \\ - & {}^{C}_{W}R & - & {}^{C}O_{W} \\ - & - & - & | \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} W_X \\ W_Y \\ W_Z \\ 1 \end{pmatrix}$$

Homogeneous coordinates

$$\begin{pmatrix} {}^{C}P \\ 1 \end{pmatrix} = \begin{pmatrix} {}^{C}_{W}\mathcal{R} & {}^{C}O_{W} \\ \mathbf{0}^{T} & 1 \end{pmatrix} \begin{pmatrix} {}^{W}P \\ 1 \end{pmatrix}$$

Block matrix form

Combining extrinsic and intrinsic calibration parameters

$$\vec{p} = \frac{1}{z}\begin{pmatrix} K & \vec{0} \end{pmatrix} \vec{P}$$

Intrinsic

$${}^{C}P = {}^{C}_{W}R\ {}^{W}P + {}^{C}O_{W}$$

Extrinsic

$$\vec{p} = \frac{1}{z}\ K\begin{pmatrix} {}^{C}_{W}R & {}^{C}O_{W} \end{pmatrix} \vec{P}$$

$$\vec{p} = \frac{1}{z}\ M\ \vec{P}$$

Forsyth&Ponce

Other ways to write the same equation

pixel coordinates

world coordinates

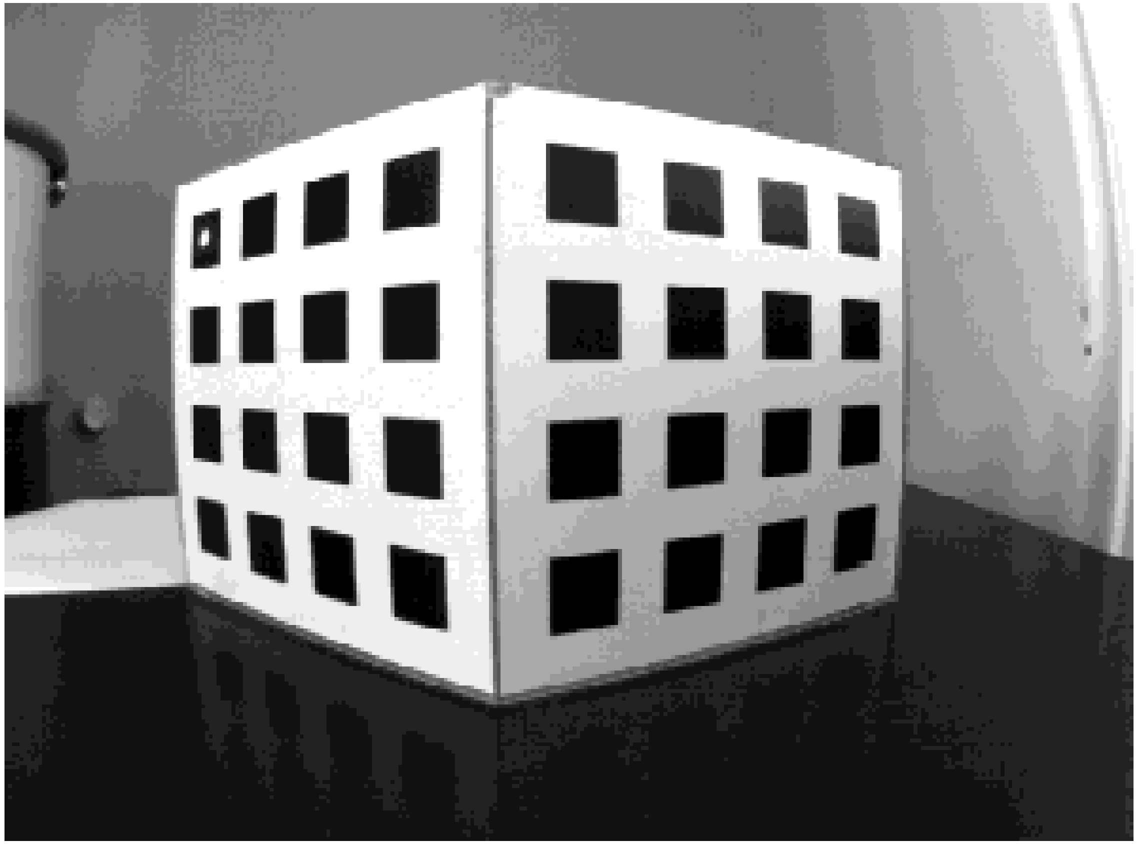
$$\vec{p} = \frac{1}{z} M\vec{P}$$

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \frac{1}{z} \begin{pmatrix} . & m_1^T & . & . \\ . & m_2^T & . & . \\ . & m_3^T & . & . \end{pmatrix} \begin{pmatrix} W_x \\ W_y \\ W_z \\ 1 \end{pmatrix}$$

$$\begin{cases} u = \dfrac{m_1 \cdot \vec{P}}{m_3 \cdot \vec{P}} \\ v = \dfrac{m_2 \cdot \vec{P}}{m_3 \cdot \vec{P}} \end{cases}$$

z is in the *camera* coordinate system, but we can solve for that, since $1 = \frac{m_3 \cdot \vec{P}}{z}$, leading to:

Calibration target



The Opti-CAL Calibration Target Image

http://www.kinetic.bc.ca/CompVision/opti-CAL.html

Camera calibration

From before, we had these equations relating image positions, u,v, to points at 3-d positions P (in homogeneous coordinates):

$$u = \frac{m_1 \cdot \vec{P}}{m_3 \cdot \vec{P}}$$

$$v = \frac{m_2 \cdot \vec{P}}{m_3 \cdot \vec{P}}$$

So for each feature point, i, we have:

$$(m_1 - u_i m_3) \cdot \vec{P}_i = 0$$

$$(m_2 - v_i m_3) \cdot \vec{P}_i = 0$$

Camera calibration

Stack all these measurements of i=1…n points

$$(m_1 - u_i m_3) \cdot \vec{P}_i = 0$$

$$(m_2 - v_i m_3) \cdot \vec{P}_i = 0$$

into a big matrix:

$$\begin{pmatrix} P_1^T & 0^T & -u_1 P_1^T \\ 0^T & P_1^T & -v_1 P_1^T \\ \cdots & \cdots & \cdots \\ P_n^T & 0^T & -u_n P_n^T \\ 0^T & P_n^T & -v_n P_n^T \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$$

Camera calibration

In vector form:

$$\begin{pmatrix} P_1^T & 0^T & -u_1 P_1^T \\ 0^T & P_1^T & -v_1 P_1^T \\ \cdots & \cdots & \cdots \\ P_n^T & 0^T & -u_n P_n^T \\ 0^T & P_n^T & -v_n P_n^T \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$$

Showing all the elements:

$$\begin{pmatrix} P_{1x} & P_{1y} & P_{1z} & 1 & 0 & 0 & 0 & 0 & -u_1 P_{1x} & -u_1 P_{1y} & -u_1 P_{1z} & -u_1 \\ 0 & 0 & 0 & 0 & P_{1x} & P_{1y} & P_{1z} & 1 & -v_1 P_{1x} & -v_1 P_{1y} & -v_1 P_{1z} & -v_1 \\ & & & & & & \cdots & \cdots & \cdots & & & \\ P_{nx} & P_{ny} & P_{nz} & 1 & 0 & 0 & 0 & 0 & -u_n P_{nx} & -u_n P_{ny} & -u_n P_{nz} & -u_n \\ 0 & 0 & 0 & 0 & P_{nx} & P_{ny} & P_{nz} & 1 & -v_n P_{nx} & -v_n P_{ny} & -v_n P_{nz} & -v_n \end{pmatrix} \begin{pmatrix} m_{11} \\ m_{12} \\ m_{13} \\ m_{14} \\ m_{21} \\ m_{22} \\ m_{23} \\ m_{24} \\ m_{31} \\ m_{32} \\ m_{33} \\ m_{34} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$$

Camera calibration

$$\begin{pmatrix} P_{1x} & P_{1y} & P_{1z} & 1 & 0 & 0 & 0 & 0 & -u_1P_{1x} & -u_1P_{1y} & -u_1P_{1z} & -u_1 \\ 0 & 0 & 0 & 0 & P_{1x} & P_{1y} & P_{1z} & 1 & -v_1P_{1x} & -v_1P_{1y} & -v_1P_{1z} & -v_1 \\ & & & & & & \cdots & \cdots & \cdots & & & \\ P_{nx} & P_{ny} & P_{nz} & 1 & 0 & 0 & 0 & 0 & -u_nP_{nx} & -u_nP_{ny} & -u_nP_{nz} & -u_n \\ 0 & 0 & 0 & 0 & P_{nx} & P_{ny} & P_{nz} & 1 & -v_nP_{nx} & -v_nP_{ny} & -v_nP_{nz} & -v_n \end{pmatrix} \begin{pmatrix} m_{11} \\ m_{12} \\ m_{13} \\ m_{14} \\ m_{21} \\ m_{22} \\ m_{23} \\ m_{24} \\ m_{31} \\ m_{32} \\ m_{33} \\ m_{34} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$$

P m = 0

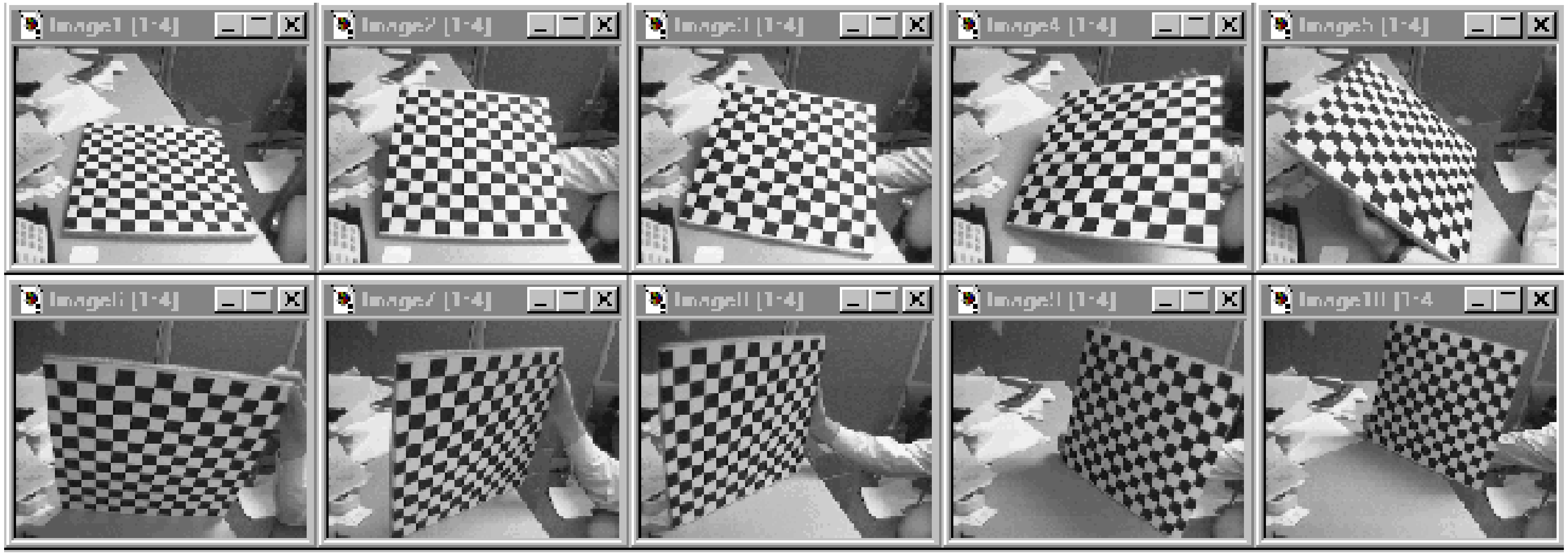
We want to solve for the unit vector m (the stacked one) that minimizes $\left|Pm\right|^2$

The minimum eigenvector of the matrix P^TP gives us that (see Forsyth&Ponce, 3.1)

Camera calibration

Once you have the M matrix, can recover the intrinsic and extrinsic parameters as in Forsyth&Ponce, sect. 3.2.2.

Multi-plane calibration

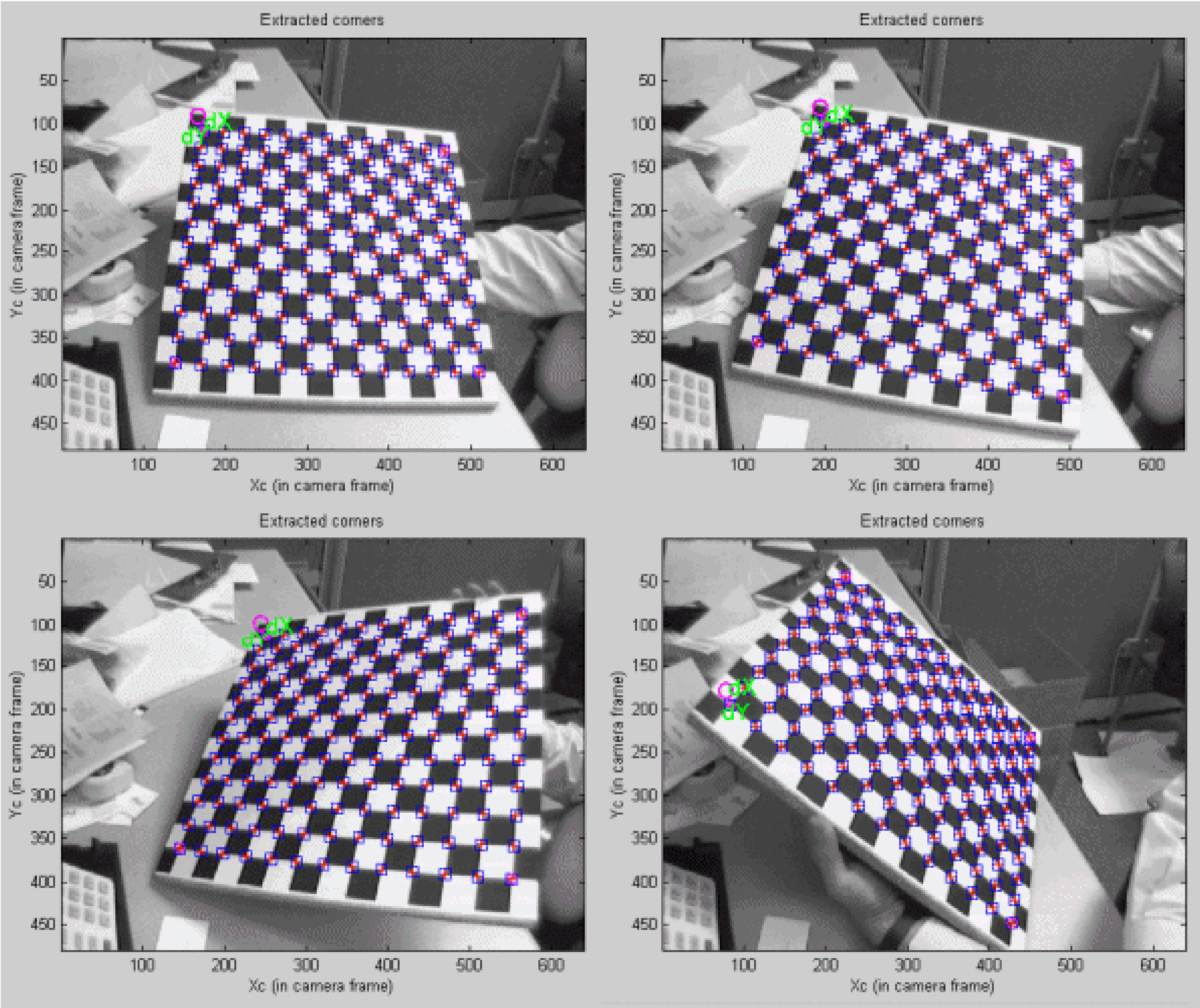


Images courtesy Jean-Yves Bouguet, Intel Corp.

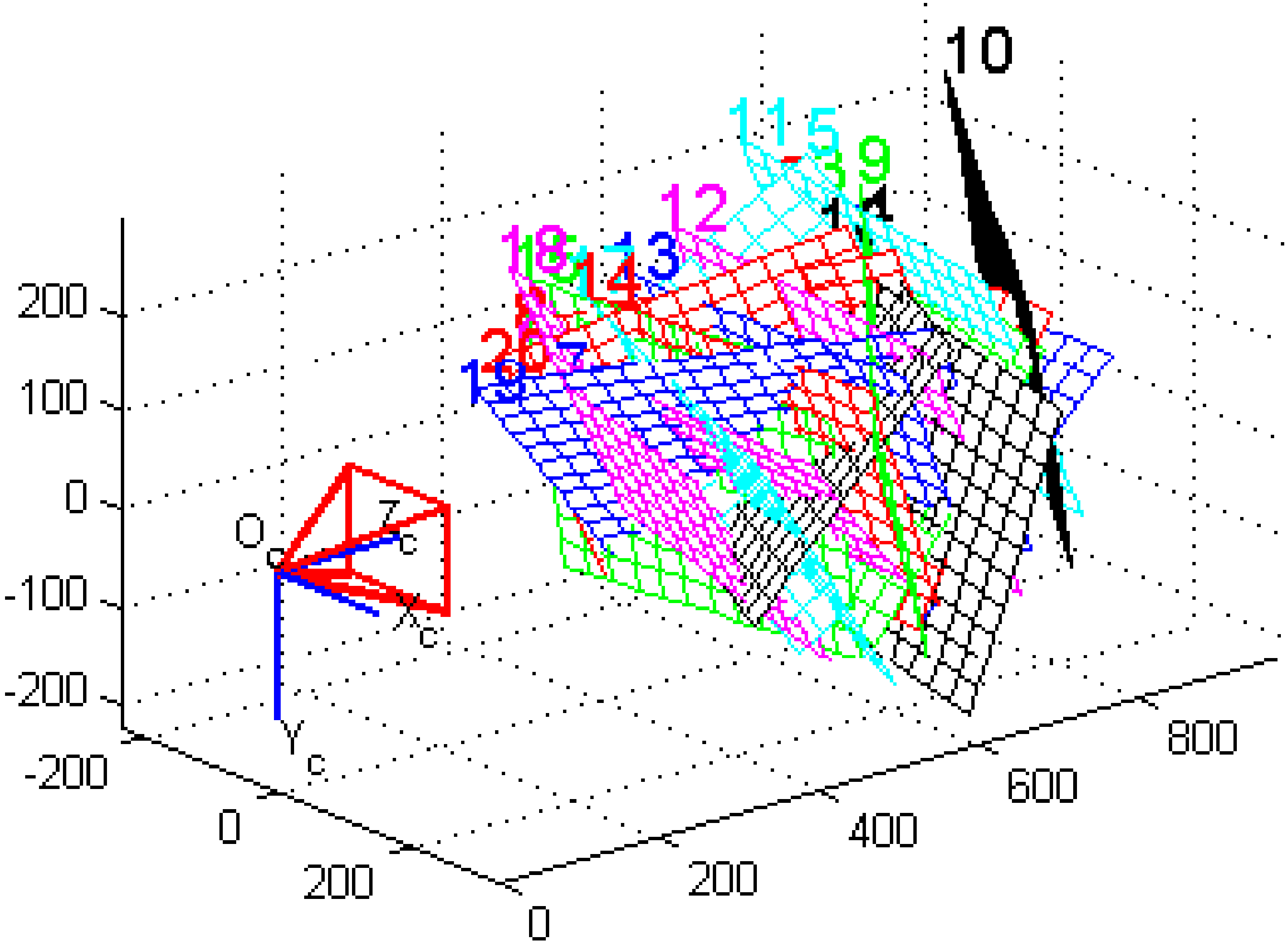
Advantage

- Only requires a plane
- Don’t have to know positions/orientations
- Good code available online!
  - Intel’s OpenCV library: http://www.intel.com/research/mrl/research/opencv/
  - Matlab version by Jean-Yves Bouget: http://www.vision.caltech.edu/bouguetj/calib_doc/index.html
  - Zhengyou Zhang’s web site: http://research.microsoft.com/~zhang/Calib/

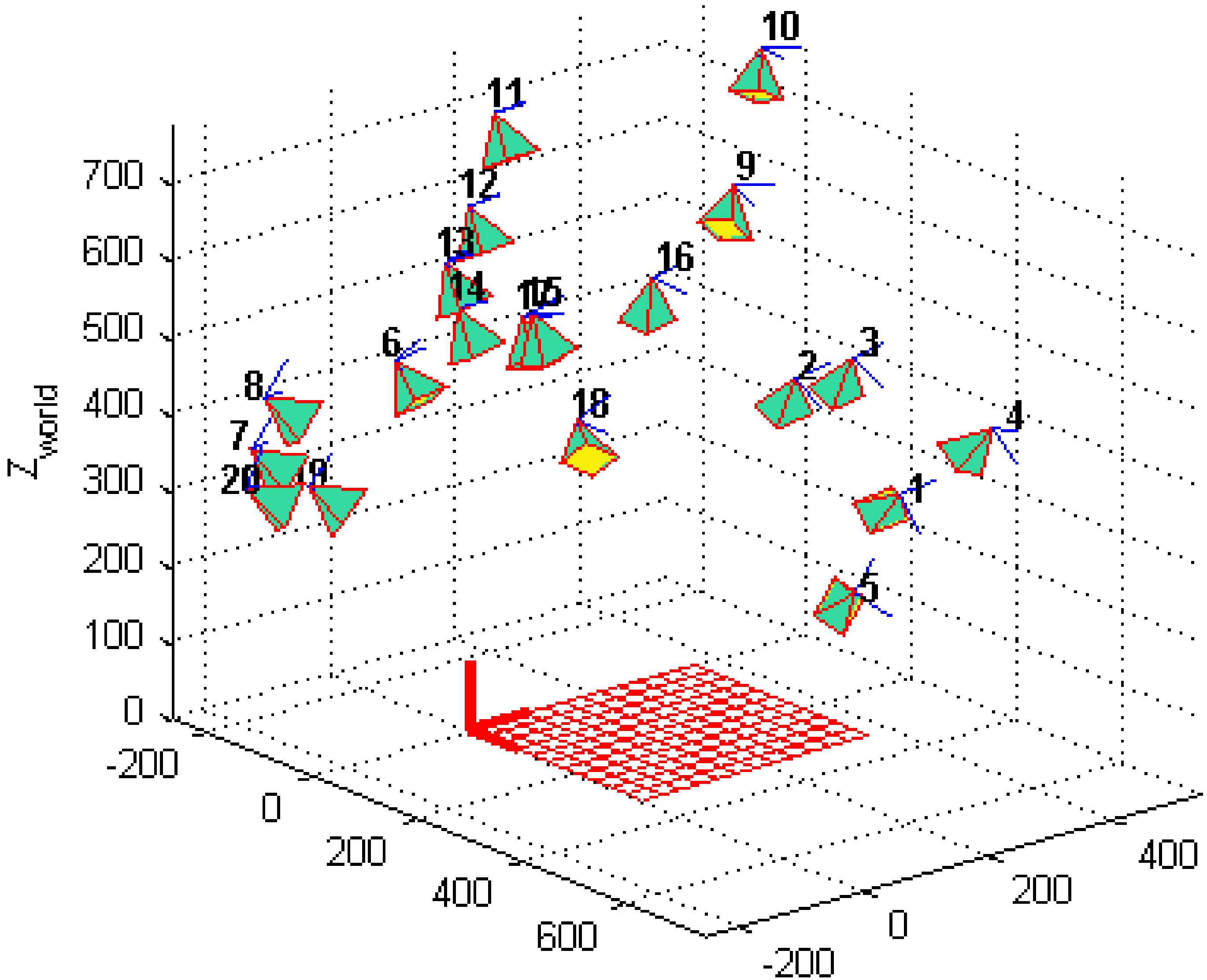
corner extraction



camera calibration



camera calibration



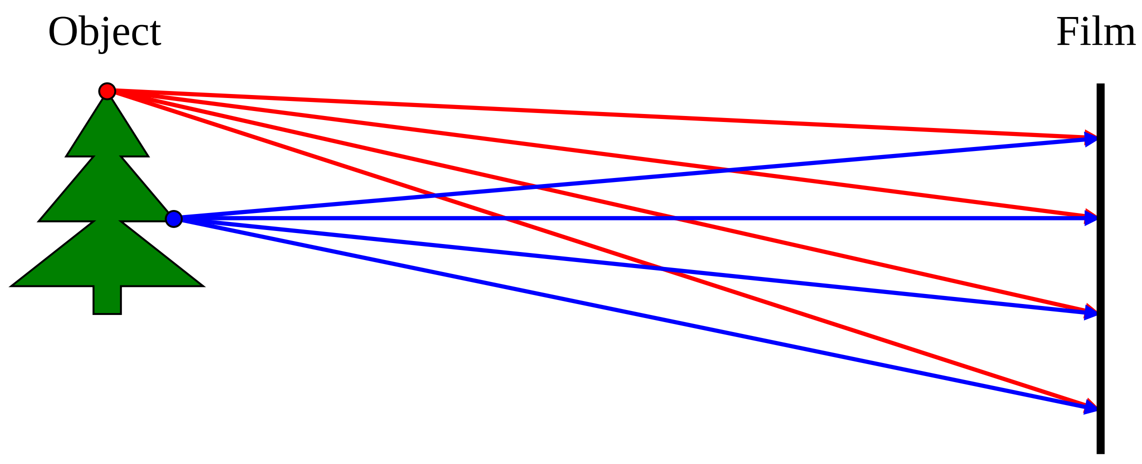
Computer Vision

Cameras & Calibration

summary

Photography

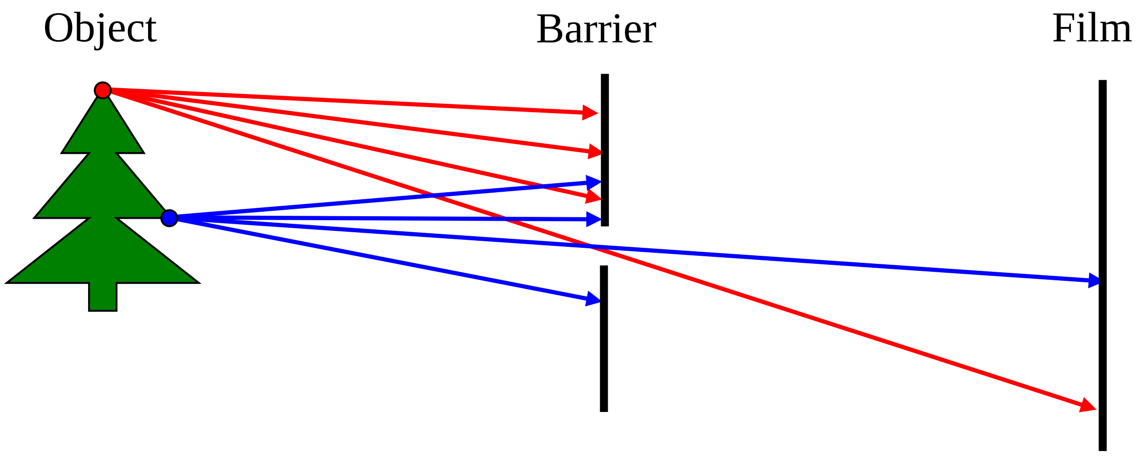
Image formation



Let's design a camera

- Idea 1: put a piece of film in front of an object
- Do we get a reasonable image?

Pinhole camera

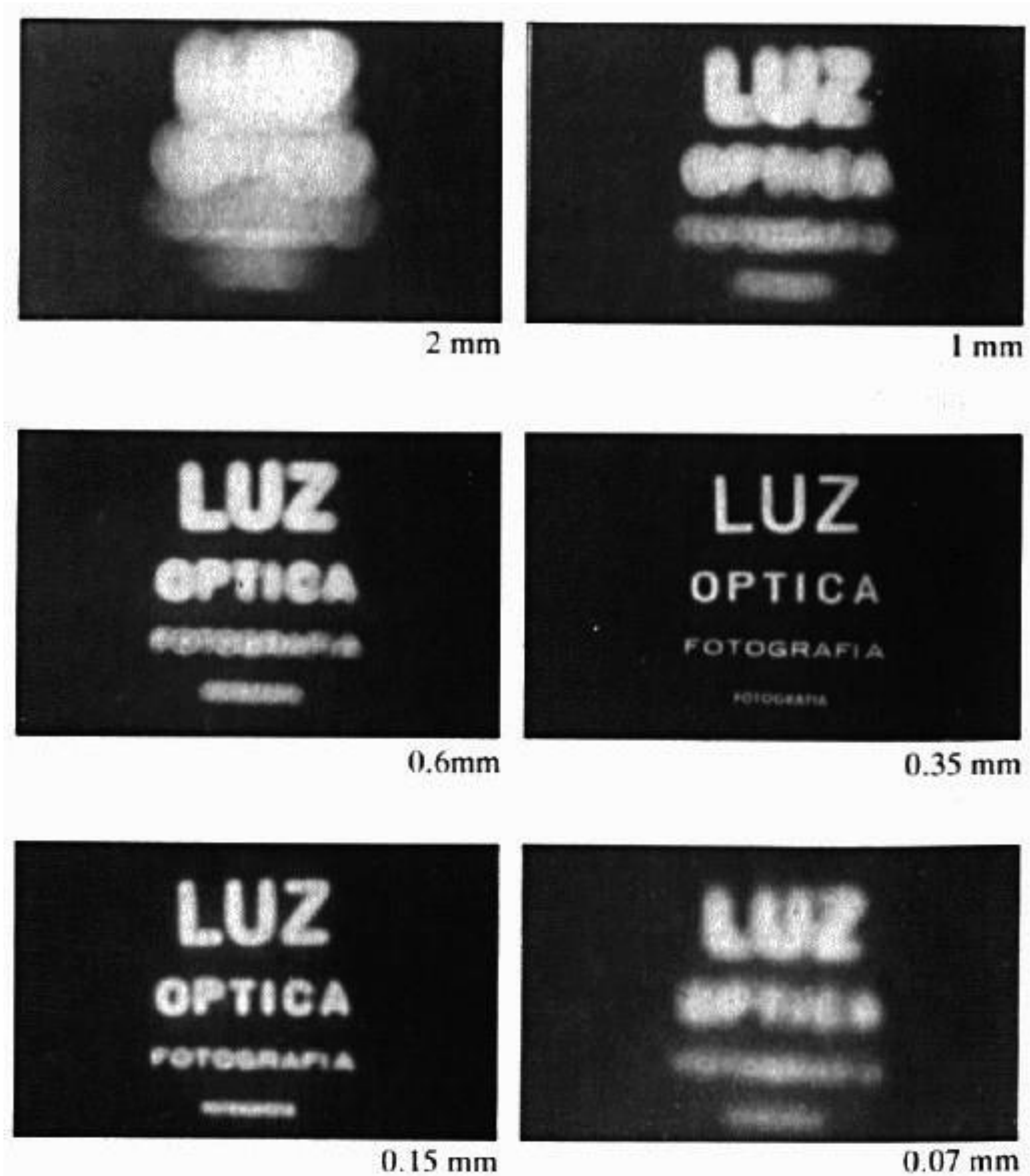


Add a barrier to block off most of the rays

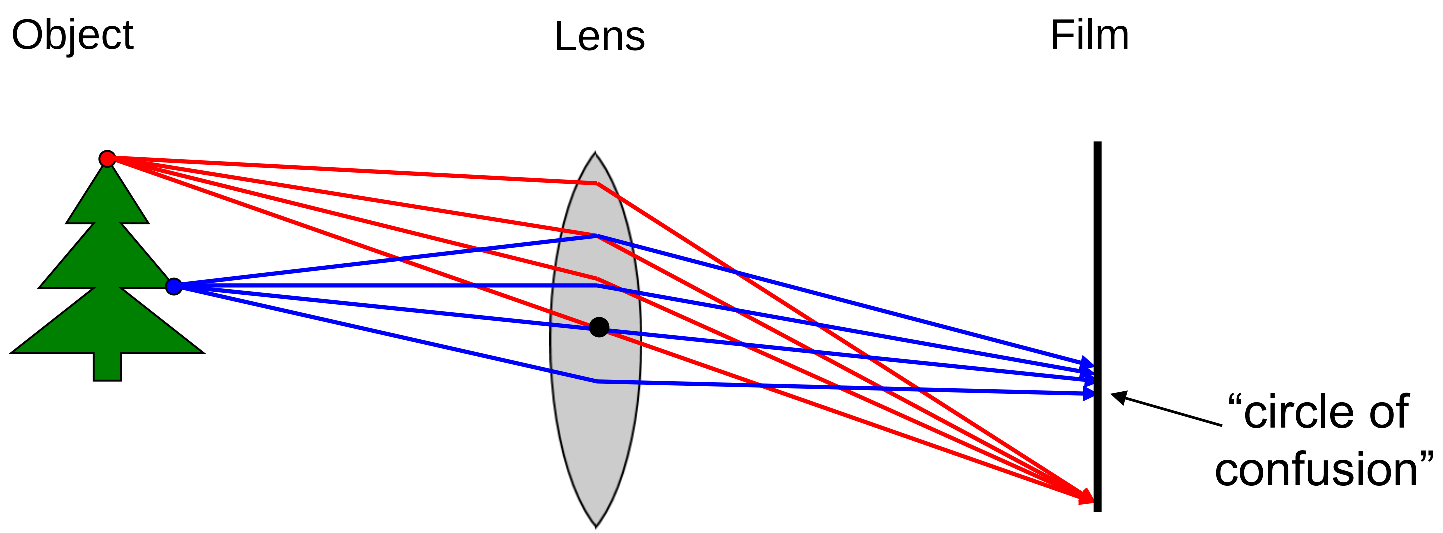
- This reduces blurring
- The opening known as the **aperture**
- How does this transform the image?

Shrinking the aperture of a pinhole camera

(This first camera was known to Aristotle as camera obscura)



Adding a lens

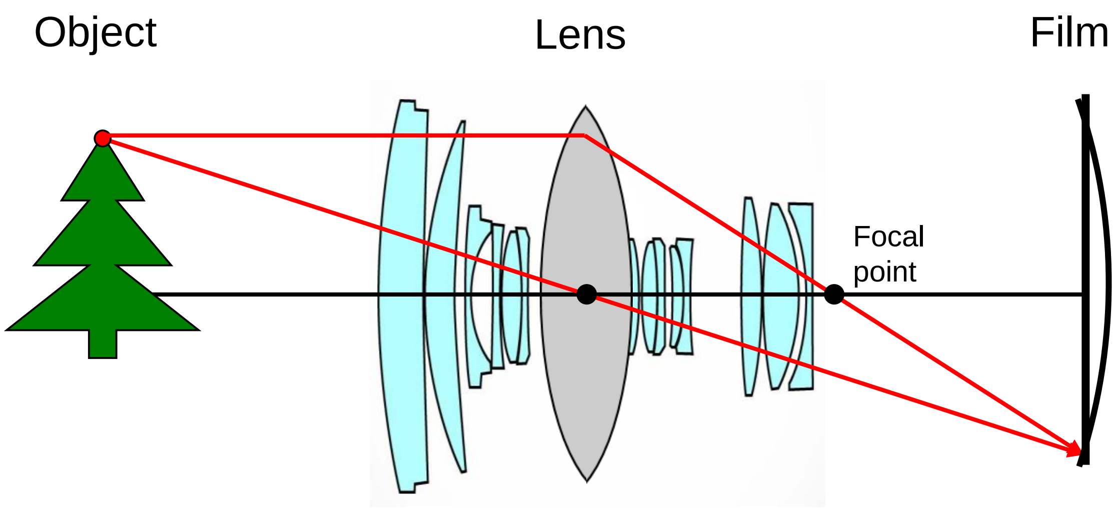


A lens focuses light onto the film

- There is a specific distance at which objects are “in focus”
  - other points project to a “circle of confusion” in the image
- Changing the shape of the lens changes this distance

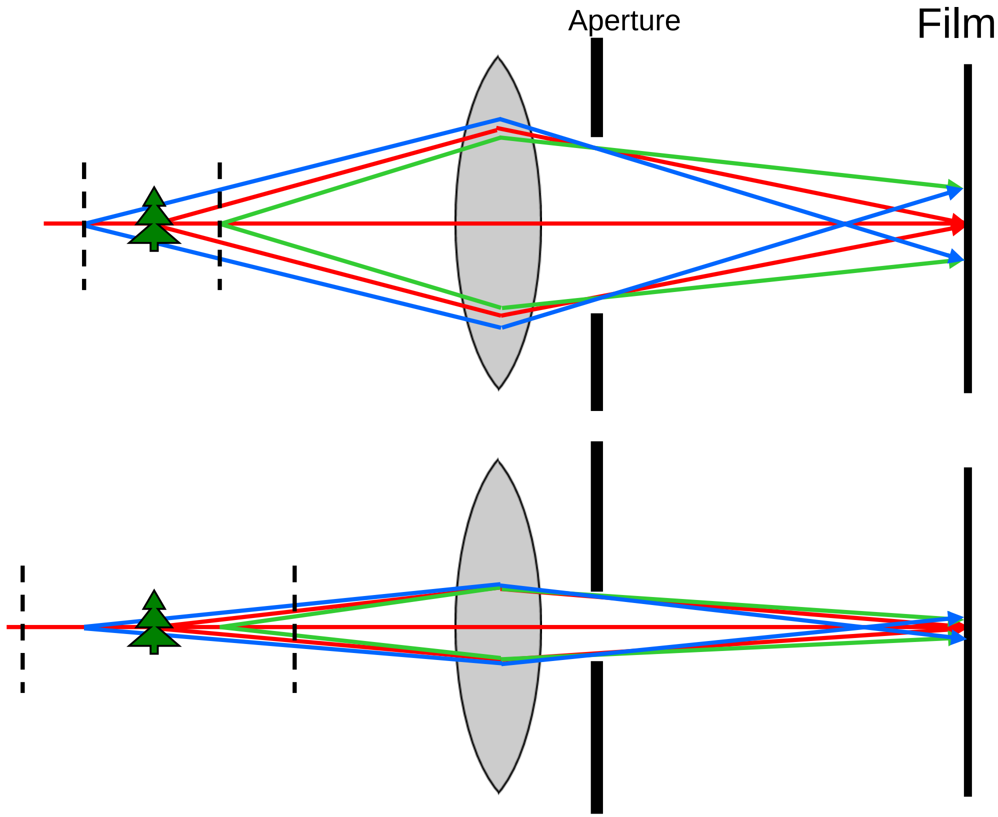
Thin lens assumption

The thin lens assumption assumes the lens has no thickness, but this isn't true…



By adding more elements to the lens, the distance at which a scene is in focus can be made roughly planar.

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

Camera parameters

Focus – Shifts the depth that is in focus.

Focal length – Adjusts the zoom, i.e., wide angle or telephoto lens.

Aperture – Adjusts the depth of field **and amount of light** let into the sensor.

Exposure time – How long an image is exposed. The longer an image is exposed the more light, but could result in motion blur.

ISO – Adjusts the sensitivity of the “film”. Basically a gain function for digital cameras. Increasing ISO also increases noise.

Sport photography

Why do they have such big lenses?

[Dirkus Maximus]

Digital Cameras

Digital camera

CCD vs. CMOS

CCD

- Low-noise images
- Consume more power
- More and higher quality pixels

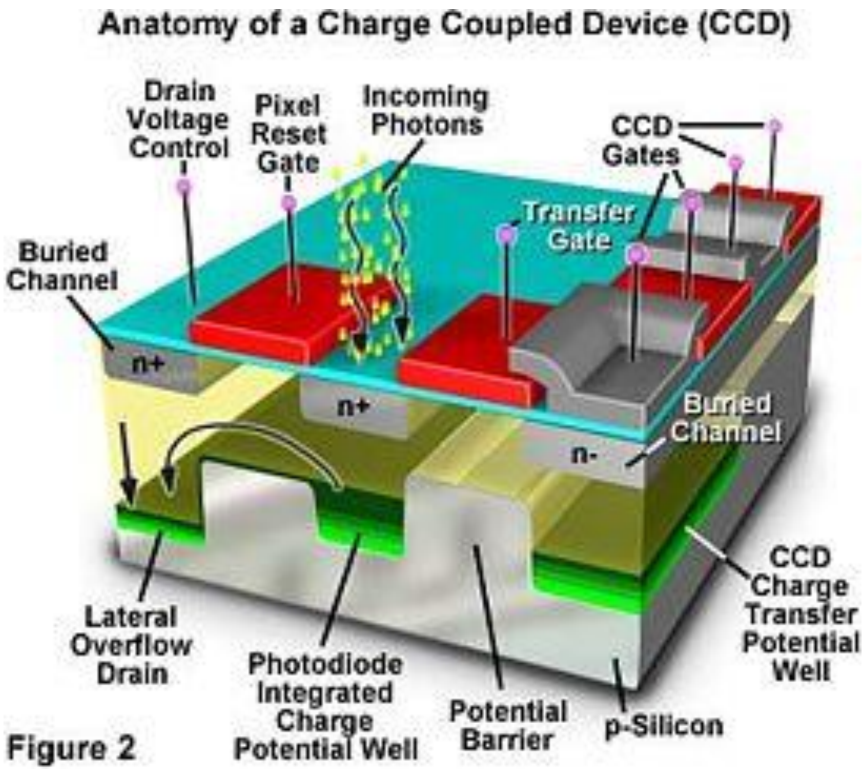
CMOS

- More noise (sensor area is smaller)
- Consume much less power
- Popular in camera phones
- Getting better all the time

http://electronics.howstuffworks.com/digital-camera.htm

Mega-pixels

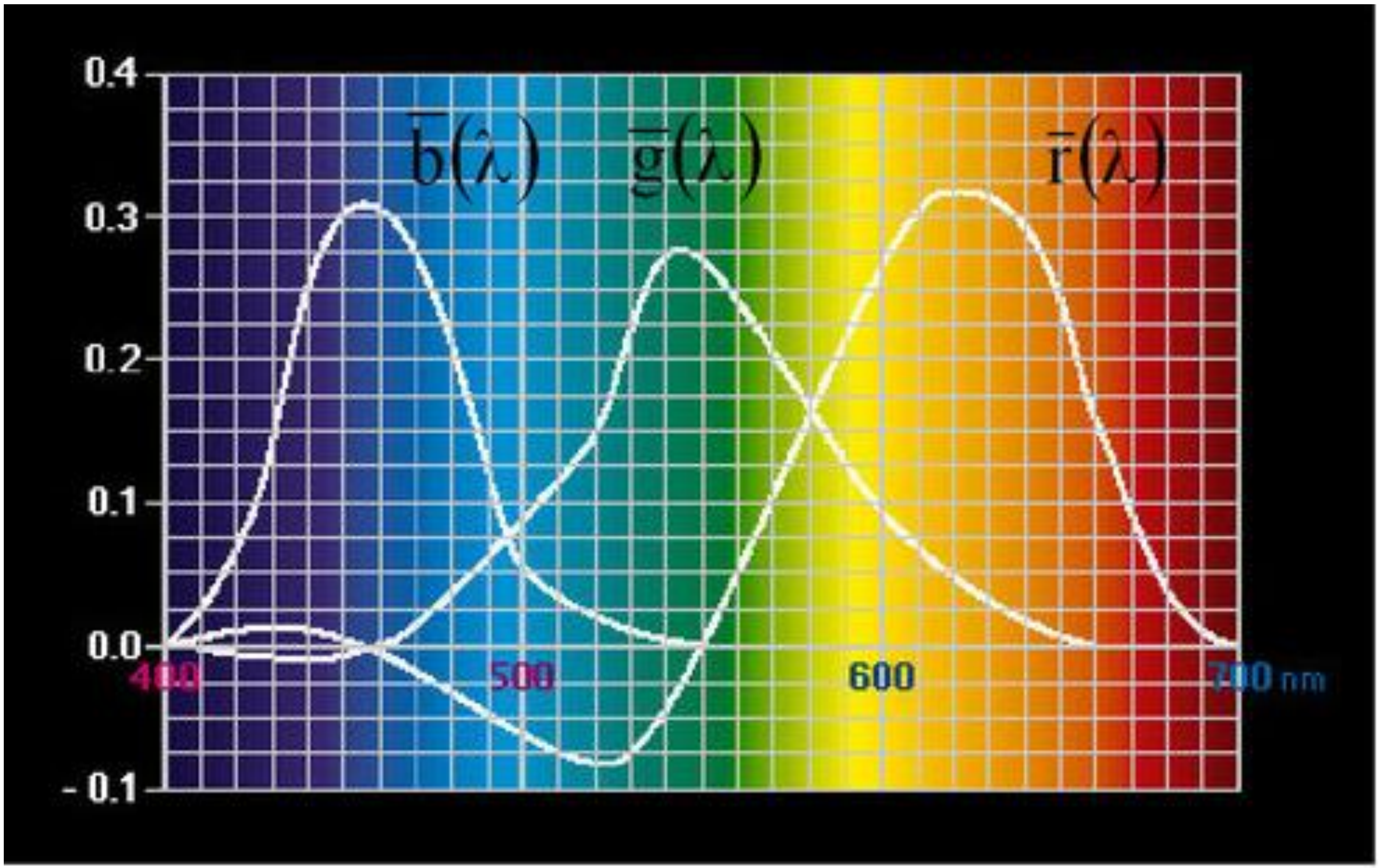
Are more mega-pixels better?



More mega-pixels require higher quality lens.

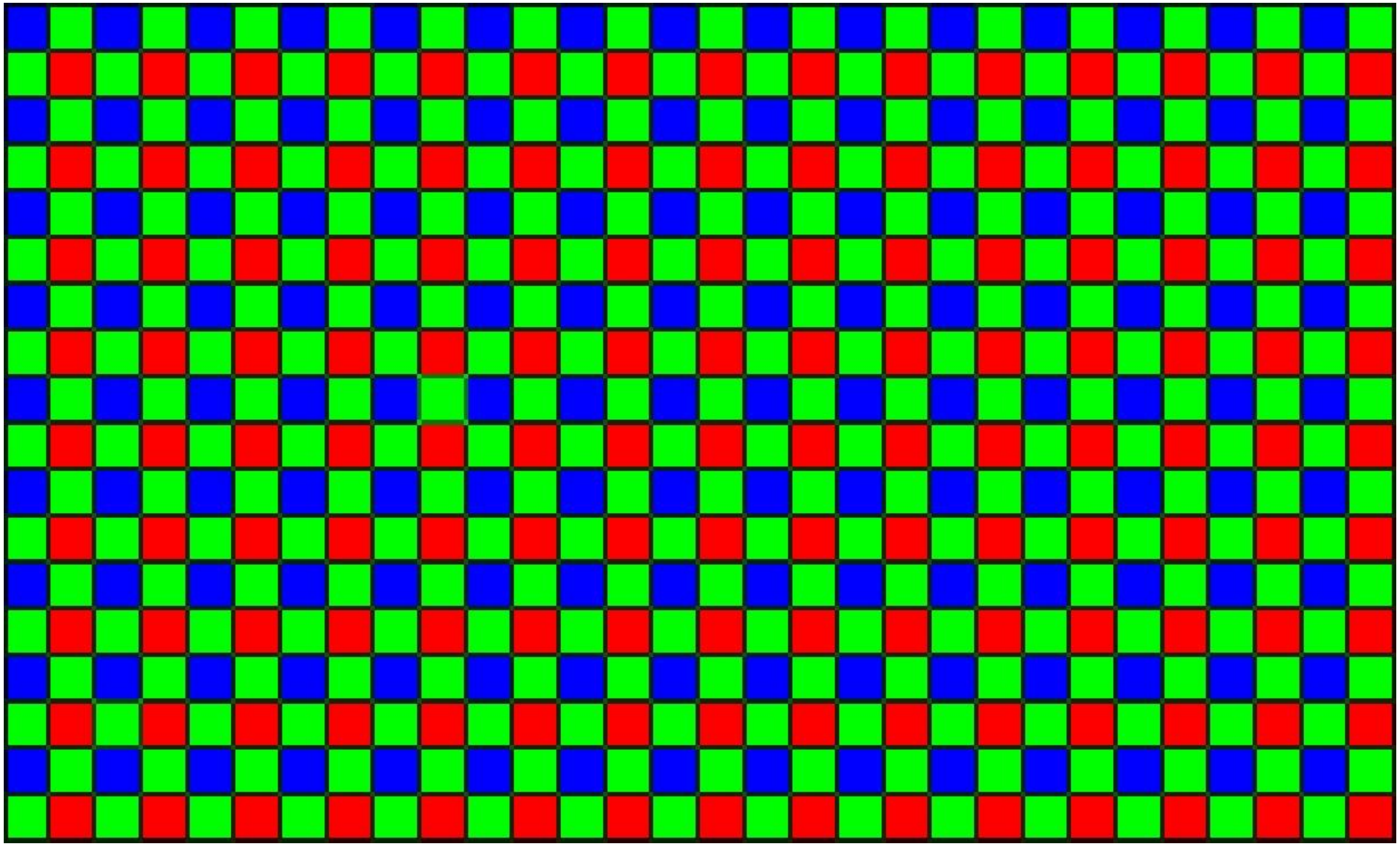
Colors

What colors do humans see?



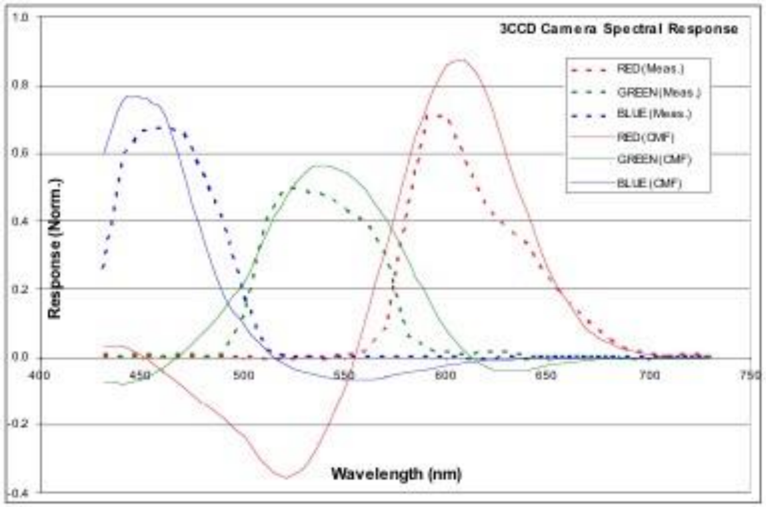
RGB tristimulus values, 1931 RGB CIE

Bayer pattern

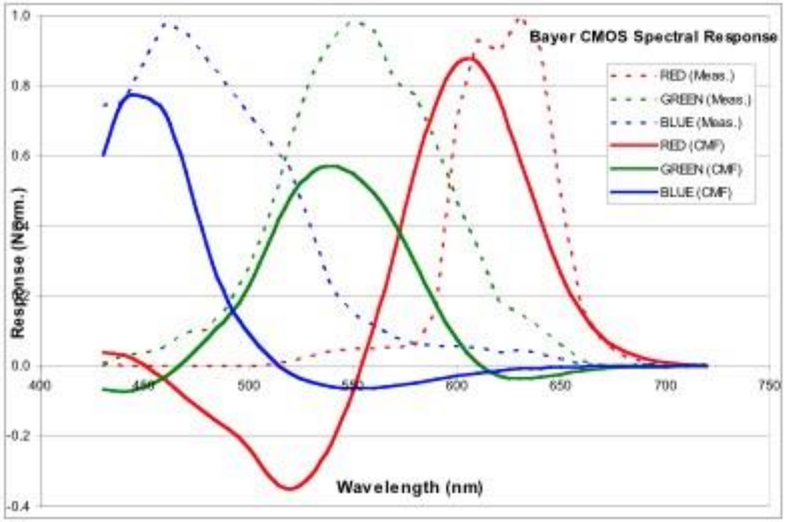


Some high end video cameras have 3 CCD chips.

Spectral response



3 chip CCD



Bayer CMOS

http://www.definitionmagazine.com/journal/2010/5/7/capturing-colour.html

Blooming

The buckets overflow…

Chromatic aberration

Different wavelengths have different refractive indices…

Interlacing

Some video cameras read even lines then odd…

Rolling shutter

Some cameras read out one line at a time:

Vignetting

The corners of images are darker than the middle:

Projection

Projection

Readings

- Szeliski 2.1

Projection

Readings

- Szeliski 2.1

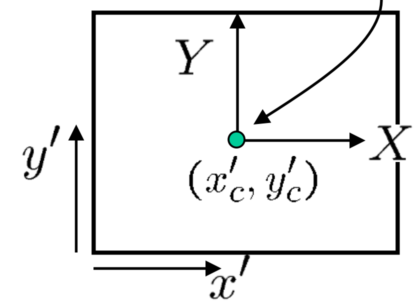
Camera parameters

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

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



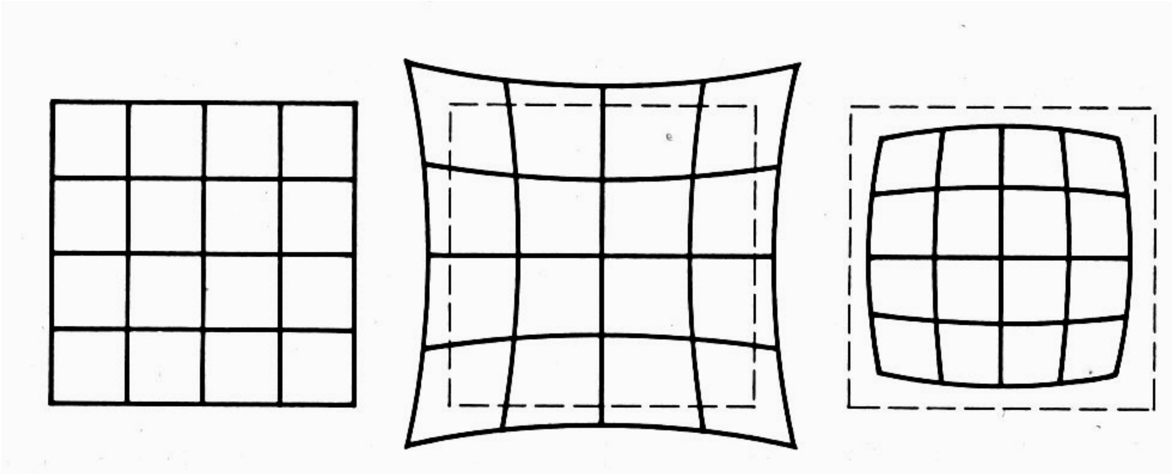
- The projection matrix models the cumulative effect of all parameters
- Useful to decompose into a series of operations

identity matrix

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

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

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

Correcting radial distortion

from [Helmut Dersch]