



Image Formation I

Chapter 1 (Forsyth&Ponce)

Cameras

Guido Gerig

CS 6320 Spring 2015

Acknowledgements:

- Slides used from Prof. Trevor Darrell, (<http://www.eecs.berkeley.edu/~trevor/CS280.html>)
- Some slides modified from Marc Pollefeys, UNC Chapel Hill. Other slides and illustrations from J. Ponce, addendum to course book.



GEOMETRIC CAMERA MODELS

- The Intrinsic Parameters of a Camera
- The Extrinsic Parameters of a Camera
- The General Form of the Perspective Projection Equation
- Line Geometry

Reading: Chapter 1.



Images are two-dimensional patterns of brightness values.

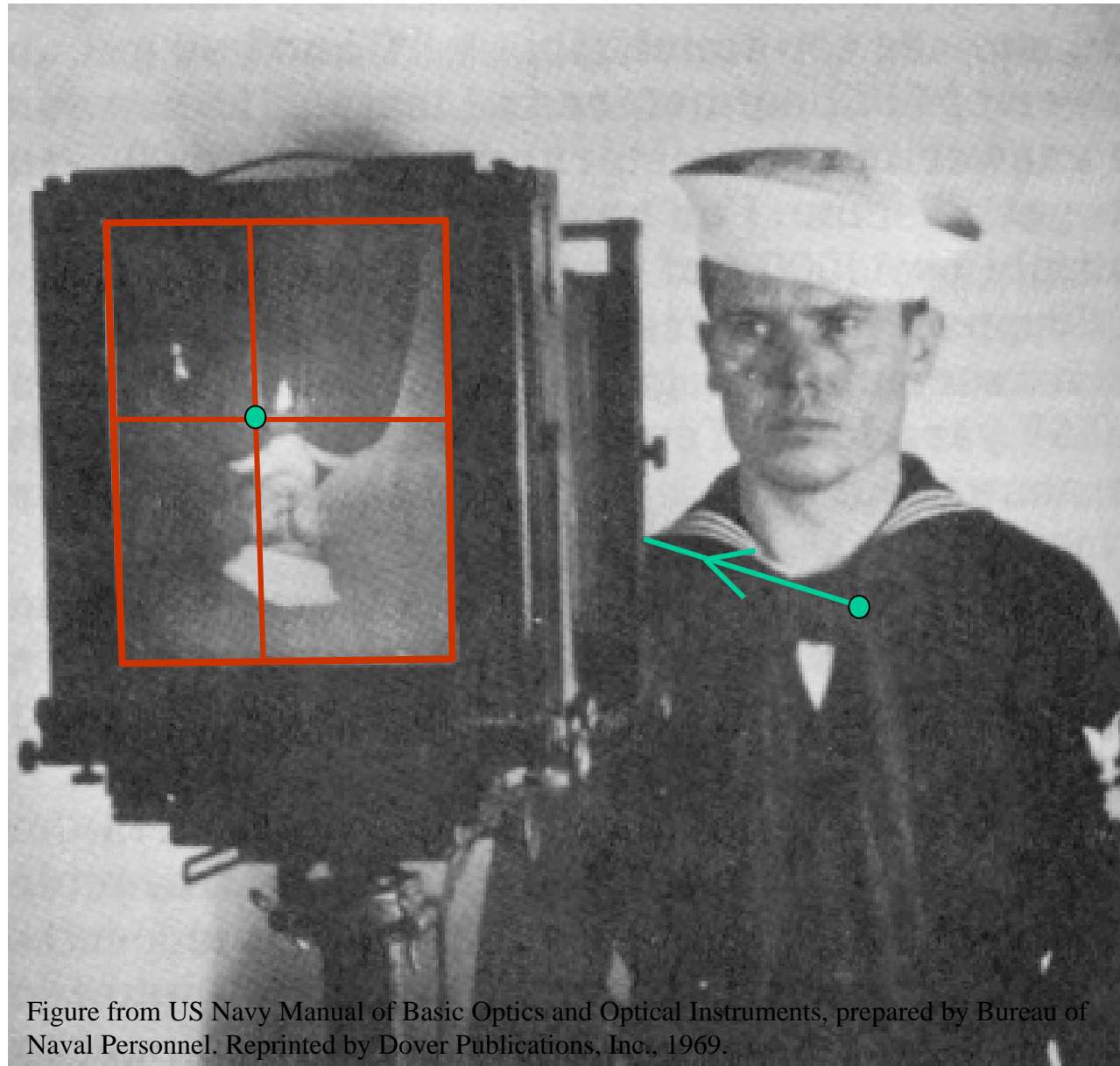
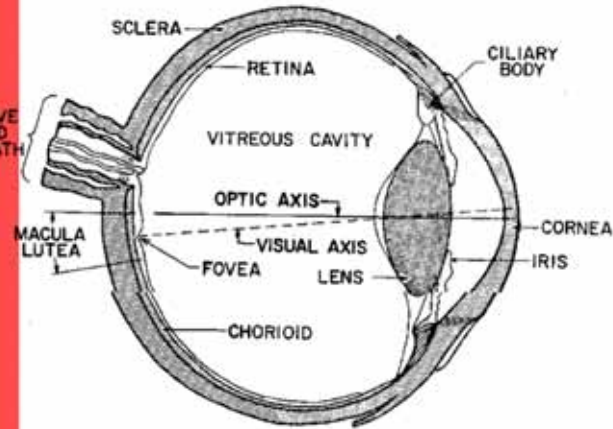
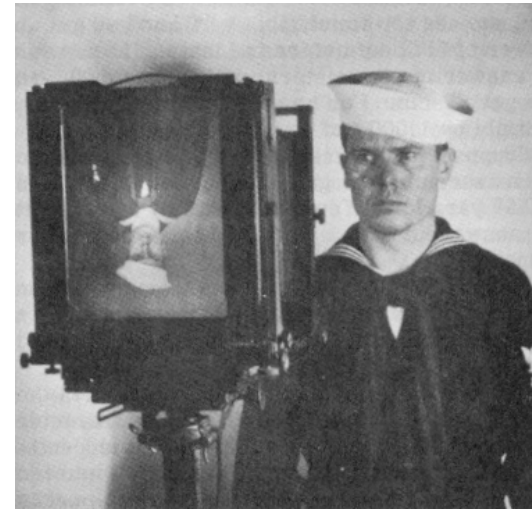


Figure from US Navy Manual of Basic Optics and Optical Instruments, prepared by Bureau of Naval Personnel. Reprinted by Dover Publications, Inc., 1969.

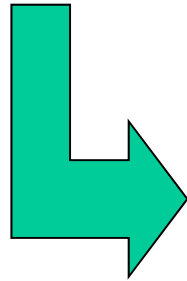
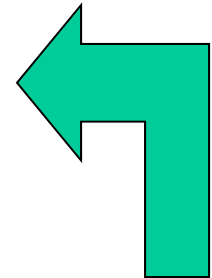
They are formed by the projection of 3D objects.



Animal eye:
a looonng time ago.

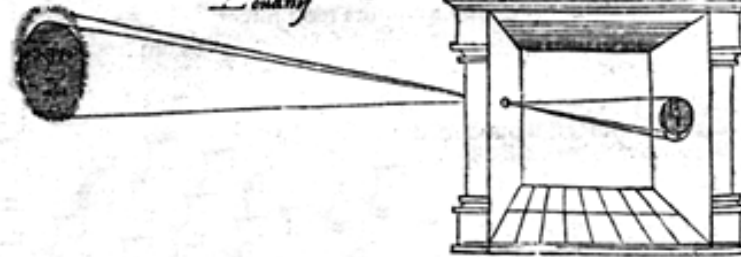


Photographic camera:
Niepce, 1816.



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

*Solis deliquium Anno Christi
1544. Die 24. Januarij
Louanij*



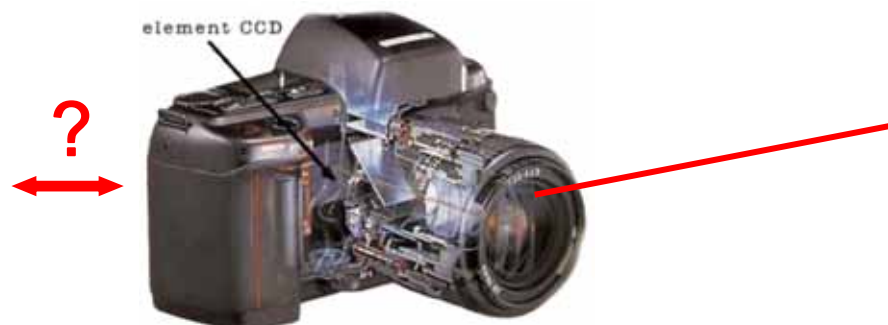
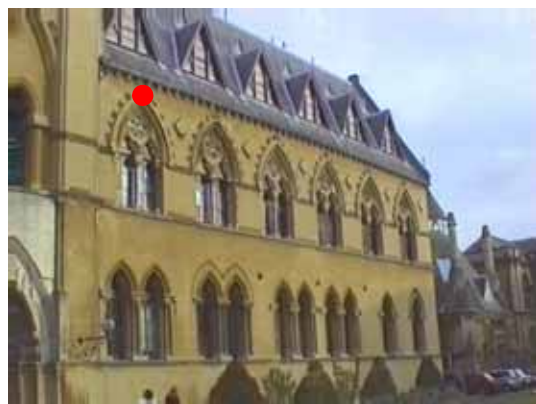
Sic nos exactè Anno .1544. Louanii eclipsim Solis
obseruauimus, inuenimusq; deficere paulò plus q̄ dex-

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



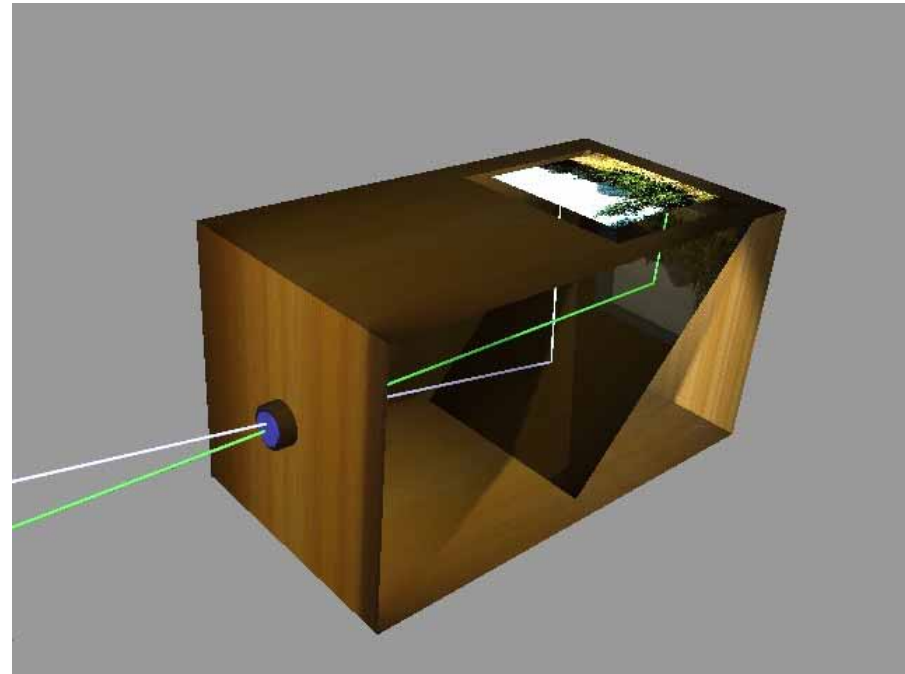
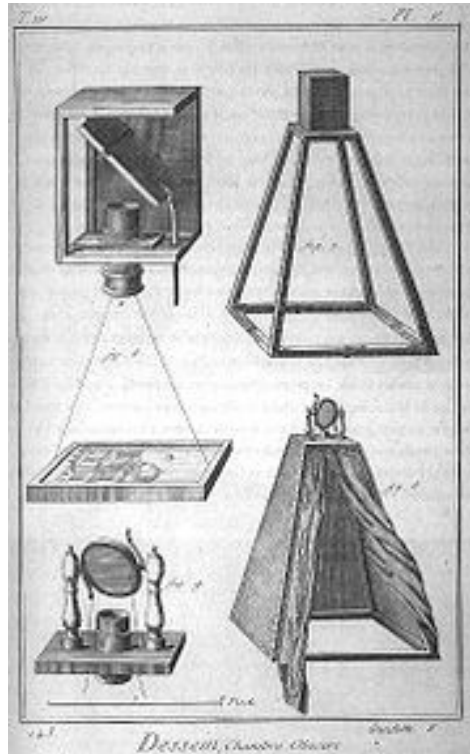
Camera model

Relation between pixels and rays in space





Camera obscura + lens

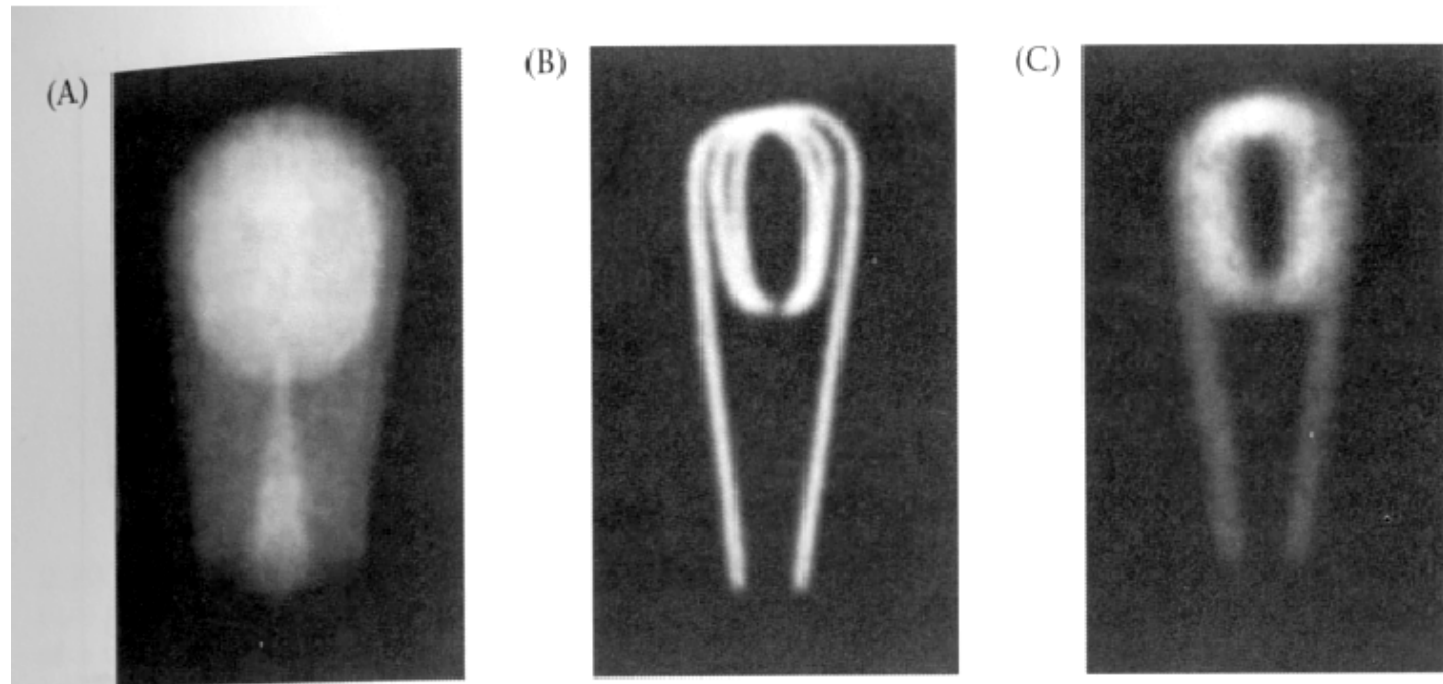


The **camera obscura** (Latin for 'dark room') is an optical device that projects an [image](#) of its surroundings on a screen (source Wikipedia).





Limits for pinhole cameras



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.

Physical parameters of image formation

- Geometric
 - Type of projection
 - Camera pose
- Photometric
 - Type, direction, intensity of light reaching sensor
 - Surfaces' reflectance properties
- Optical
 - Sensor's lens type
 - focal length, field of view, aperture
- Sensor
 - sampling, etc.

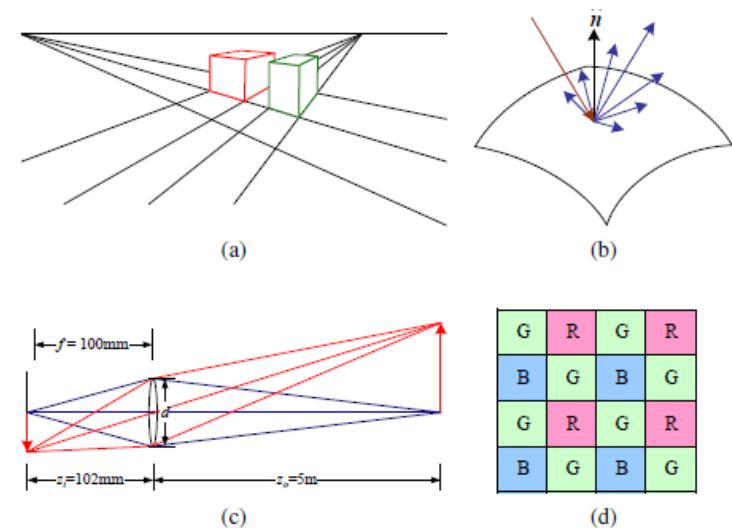


Figure 2.1 A few components of the image formation process: (a) perspective projection; (b) light scattering when hitting a surface; (c) lens optics; (d) Bayer color filter array.

Physical parameters of image formation

- Geometric
 - Type of projection
 - Camera pose
- Optical
 - Sensor's lens type
 - focal length, field of view, aperture
- Photometric
 - Type, direction, intensity of light reaching sensor
 - Surfaces' reflectance properties
- Sensor
 - sampling, etc.

Perspective and art

- Use of correct perspective projection indicated in 1st century B.C. frescoes
- Skill resurfaces in Renaissance: artists develop systematic methods to determine perspective projection (around 1480-1515)



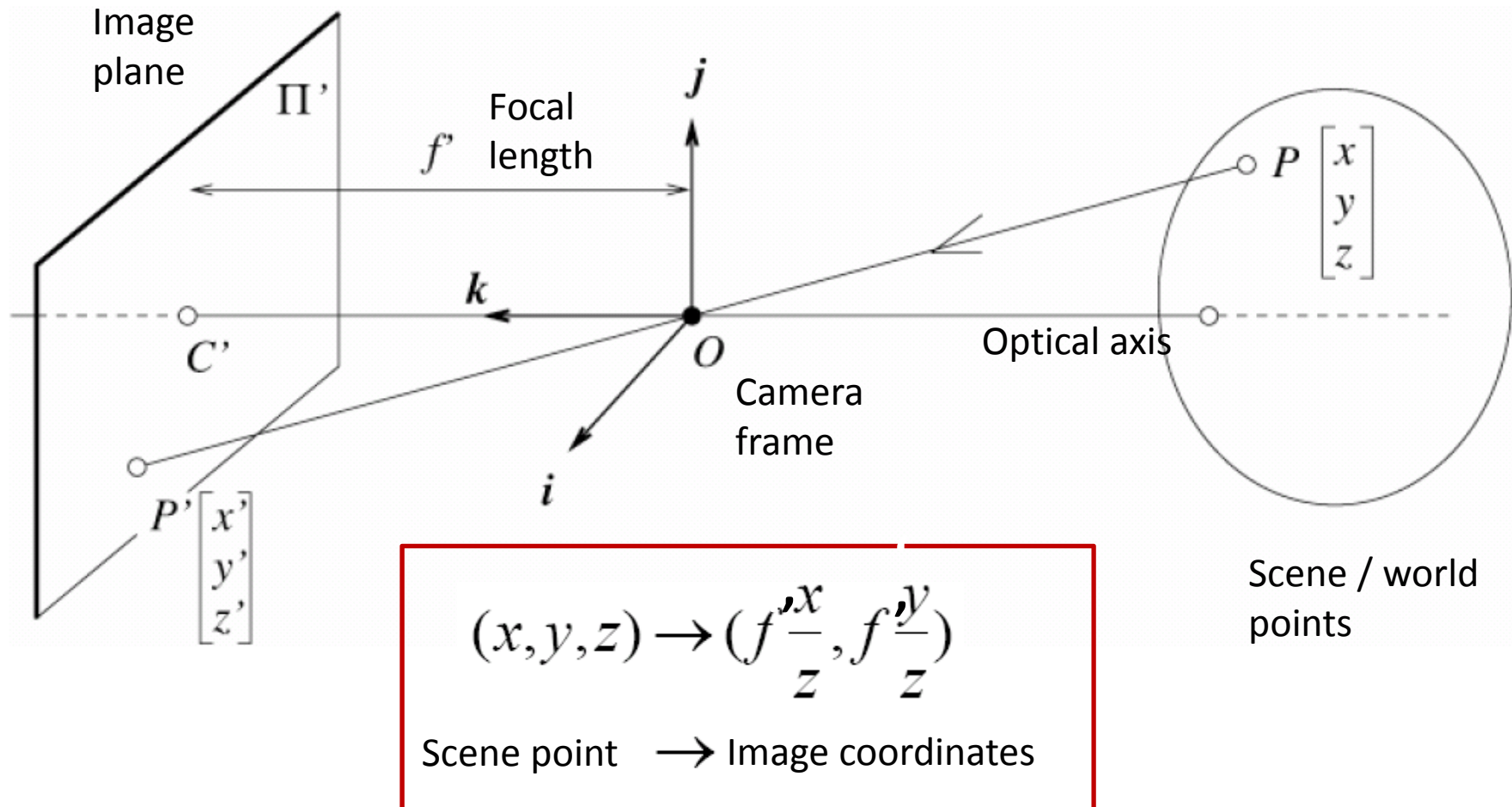
Raphael



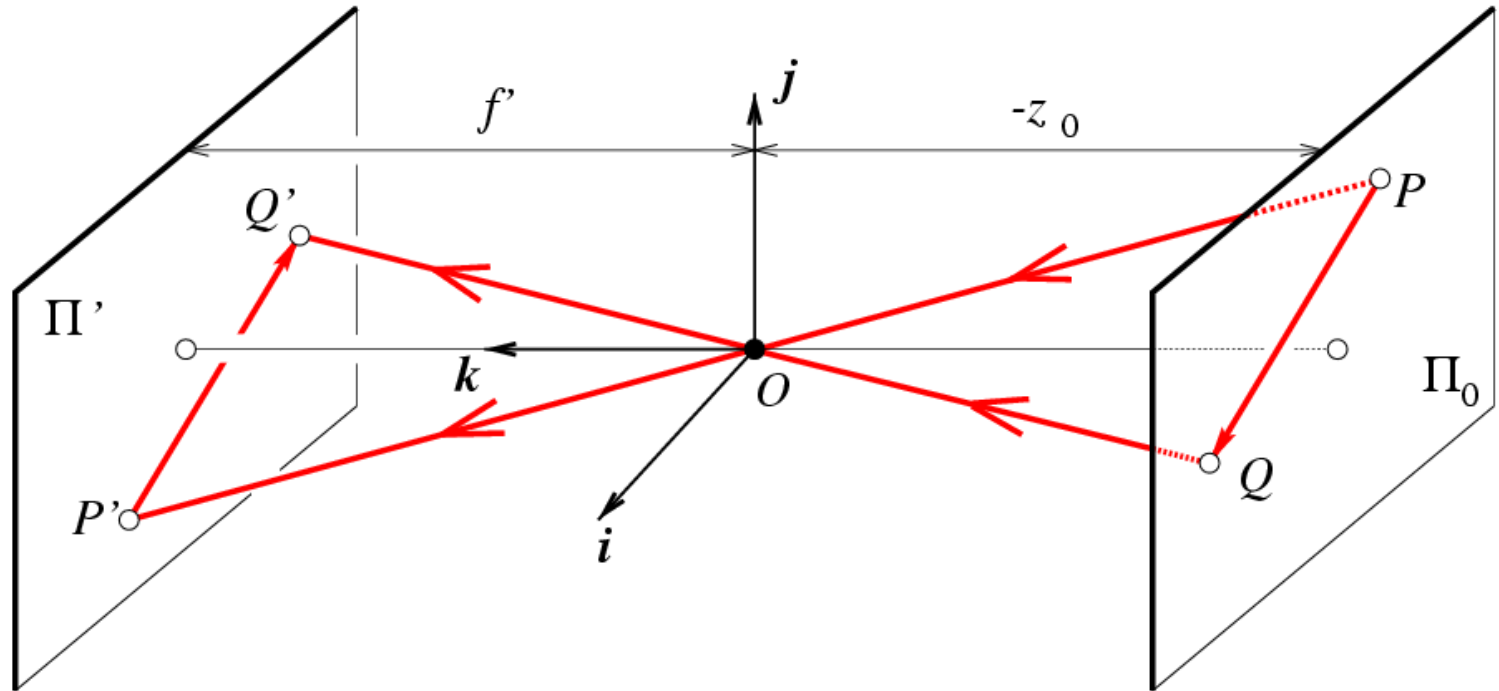
Durer, 1525

Perspective projection equations

- 3d world mapped to 2d projection in image plane



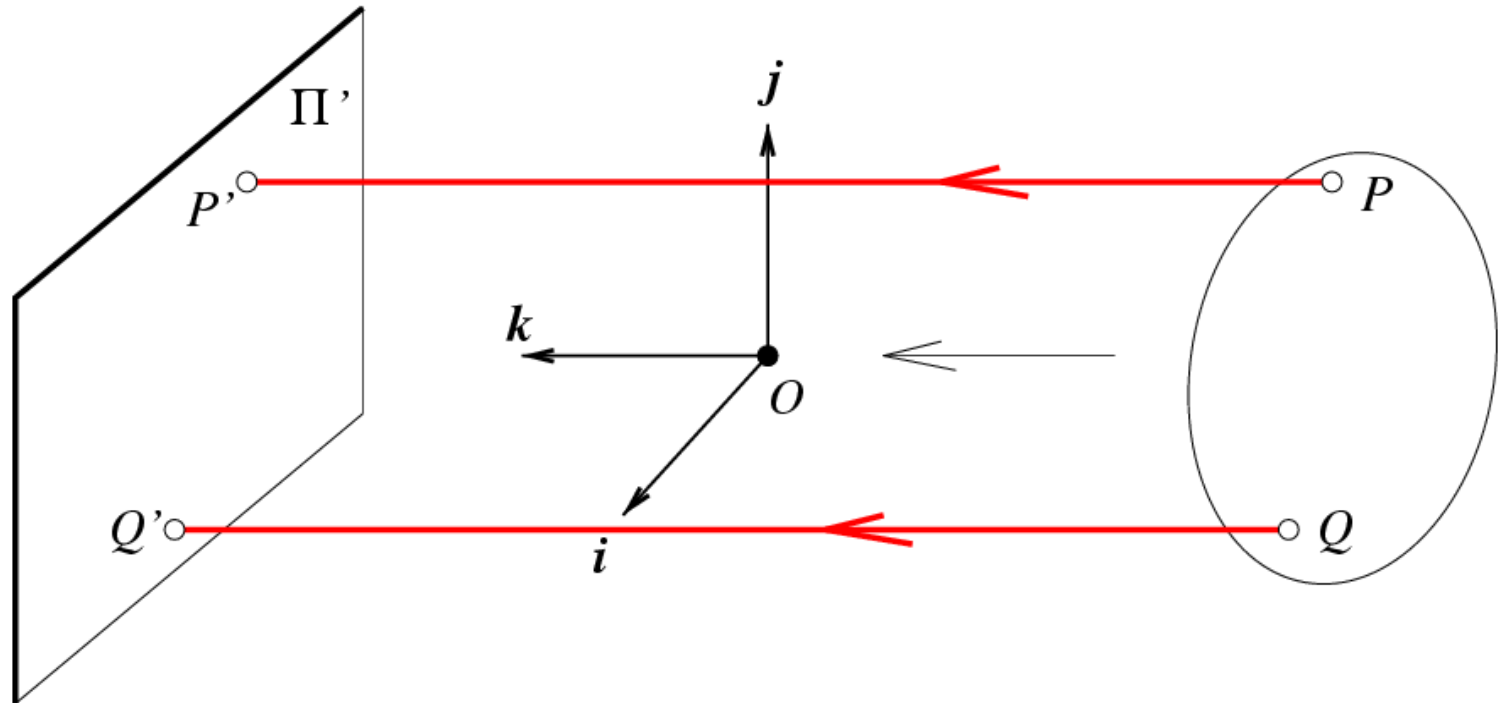
Affine projection models: Weak perspective projection



$$\begin{cases} x' = -mx \\ y' = -my \end{cases} \quad \text{where} \quad m = \frac{f'}{z_0} \quad \text{is the magnification.}$$

When the scene distance variation is small compared to its distance from the Camera, m can be taken constant: weak perspective projection.

Affine projection models: Orthographic projection



$$\begin{cases} x' = x \\ y' = y \end{cases}$$

When the camera is at a (roughly constant) distance from the scene, take $m=1$.

Homogeneous coordinates

Is this a linear transformation?

- no—division by z is nonlinear

Trick: add one more coordinate:

$$(x, y) \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

homogeneous image
coordinates

$$(x, y, z) \Rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

homogeneous scene
coordinates

Converting *from* homogeneous coordinates

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x/w, y/w)$$

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \Rightarrow (x/w, y/w, z/w)$$

Perspective Projection Matrix

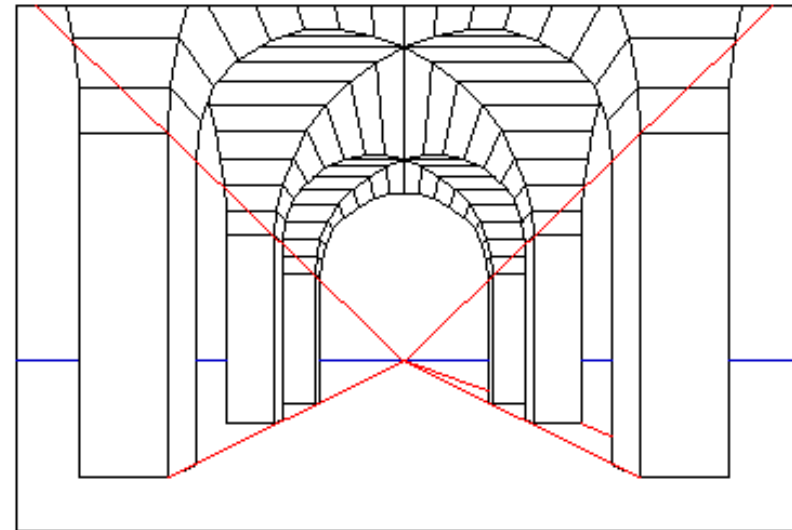
- Projection is a matrix multiplication using homogeneous coordinates:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f' & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z/f' \end{bmatrix} \Rightarrow \left(f' \frac{x}{z}, f' \frac{y}{z} \right)$$

divide by the third coordinate
to convert back to non-
homogeneous coordinates

Complete mapping from world points to image pixel
positions?

Points at infinity, vanishing points



Points from infinity represent rays into camera which are close to the optical axis.

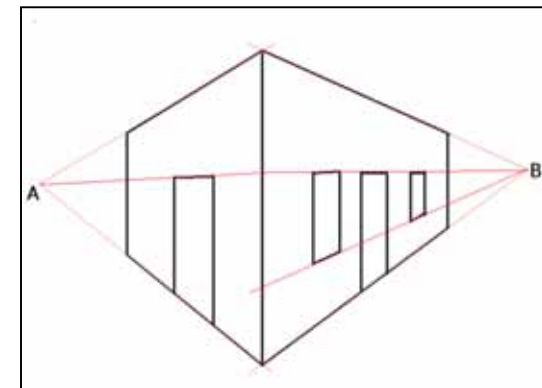
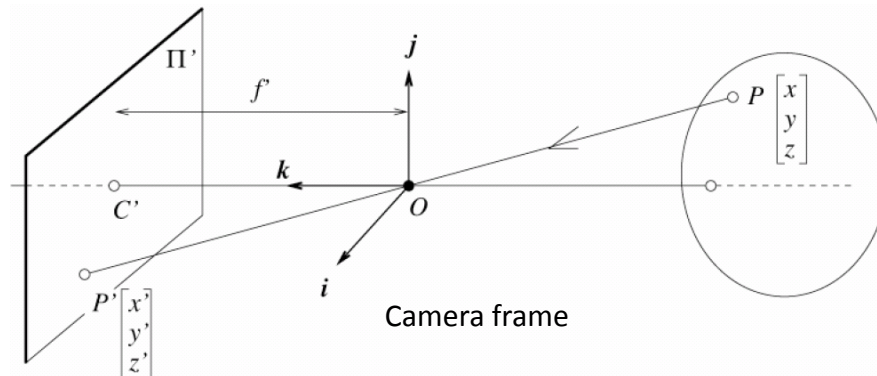


Image source: [wikipedia](#)

Perspective projection & calibration

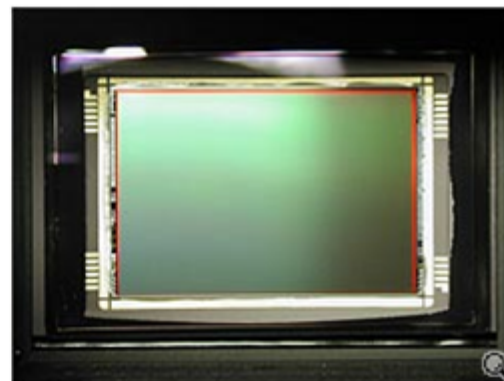
- Perspective equations so far in terms of *camera's* reference frame....
- Camera's *intrinsic* and *extrinsic* parameters needed to calibrate geometry.



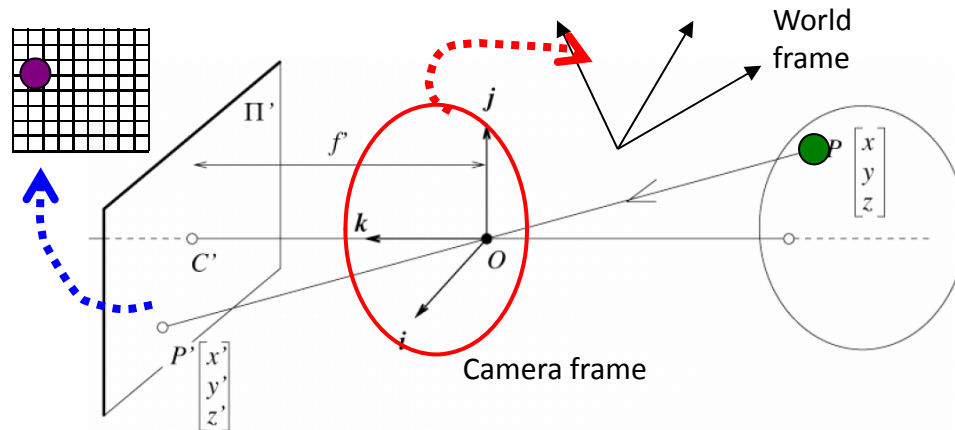


The CCD camera

CCD camera



Perspective projection & calibration



Extrinsic:

Camera frame \leftrightarrow World frame

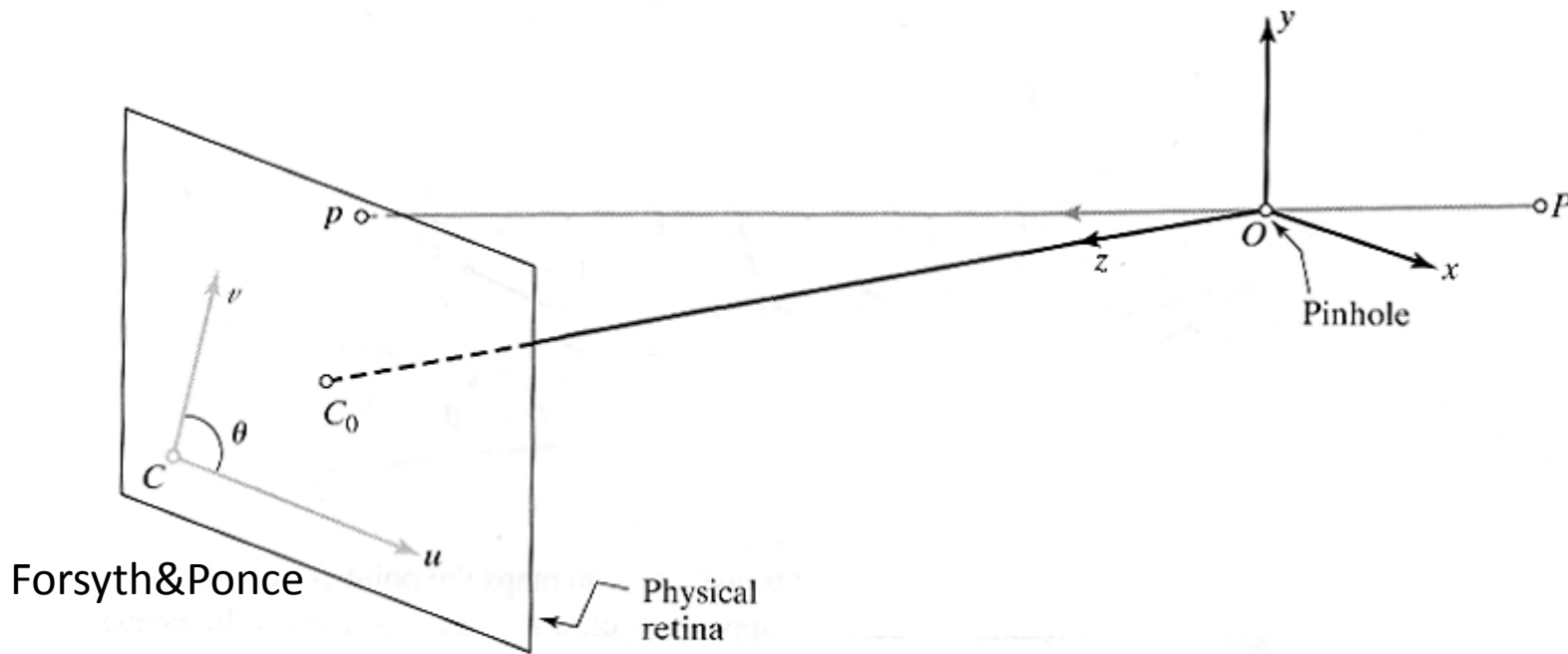
Intrinsic:

Image coordinates relative to camera

\leftrightarrow Pixel coordinates

3D
point
(4x1)

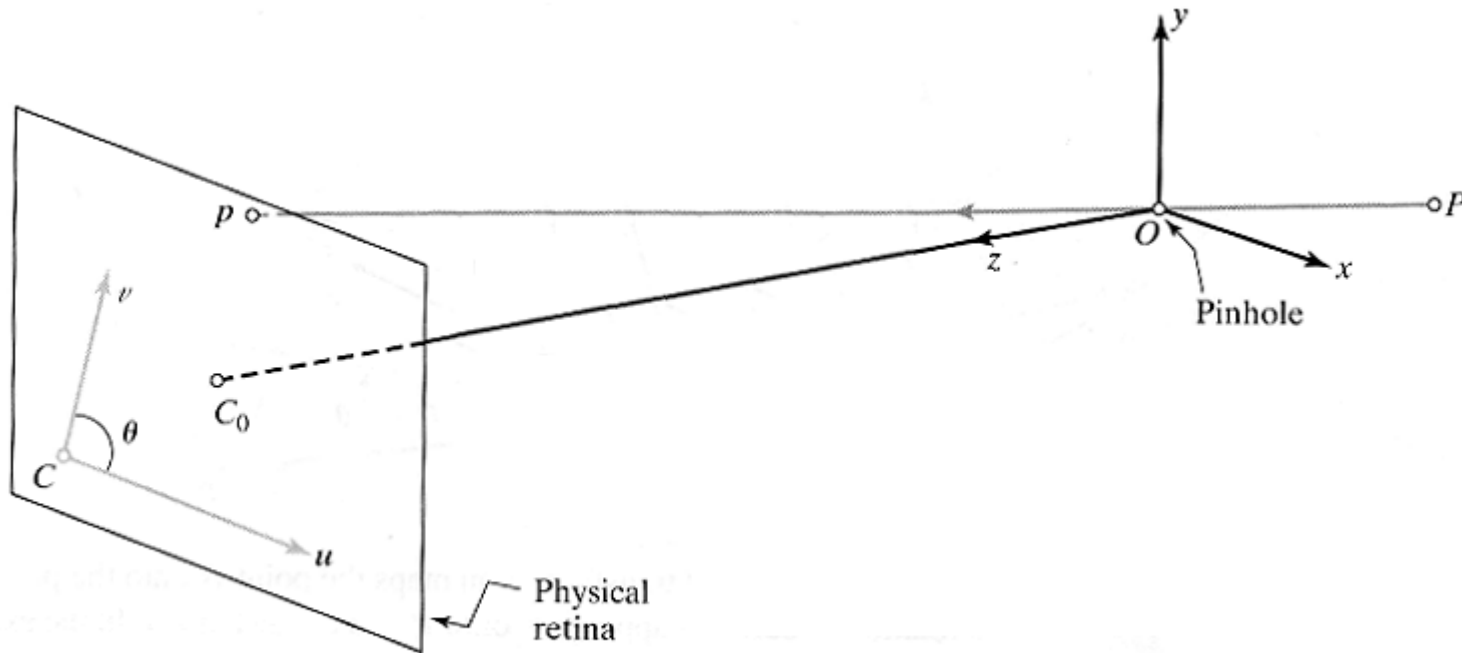
Intrinsic parameters: from idealized world coordinates to pixel values



Perspective projection

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

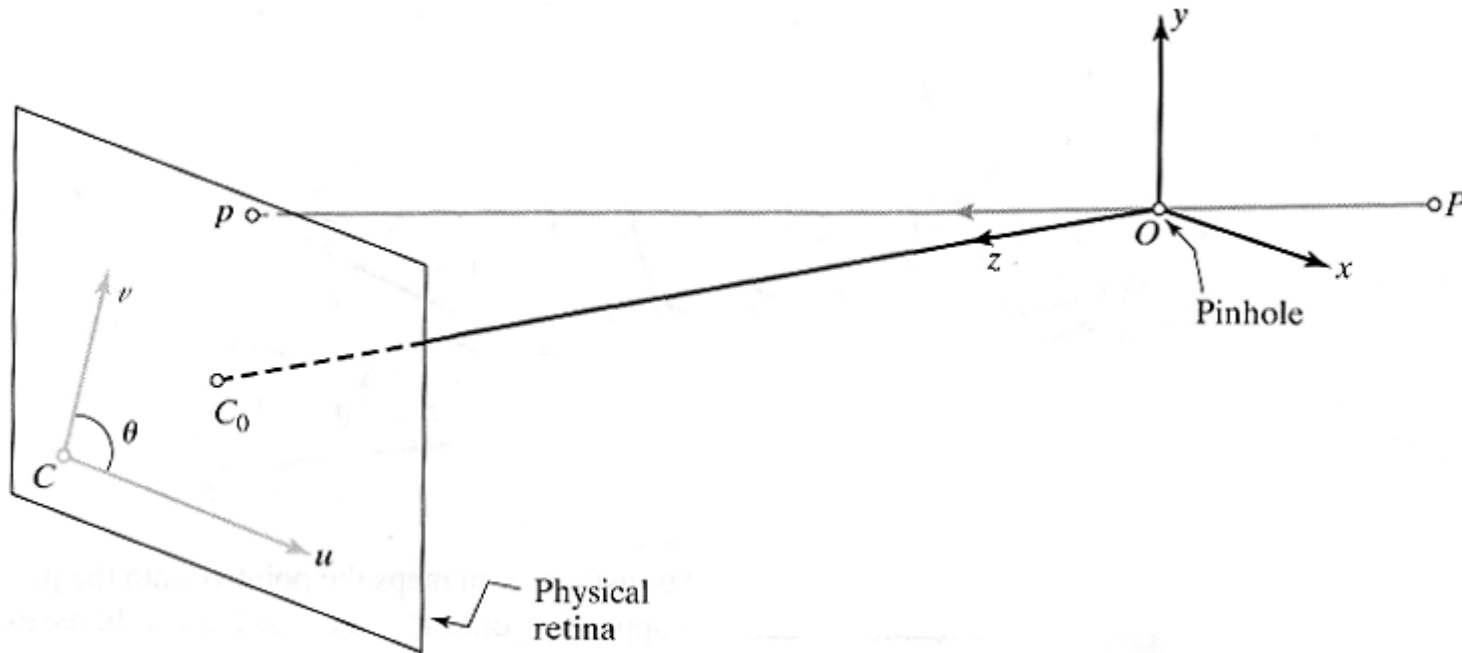
Intrinsic parameters



But “pixels” are in some arbitrary spatial units

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

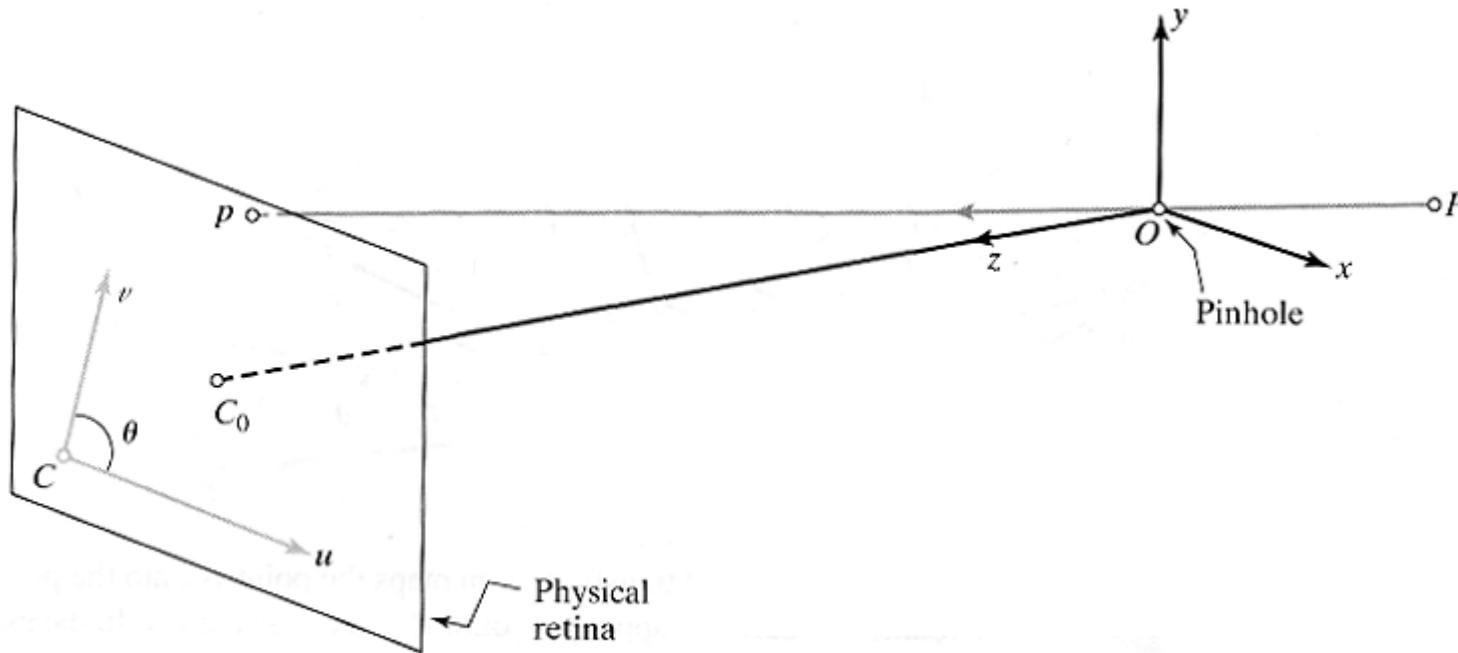
Intrinsic parameters



Maybe pixels are not square

$$u = \alpha \frac{x}{z}$$
$$v = \beta \frac{y}{z}$$

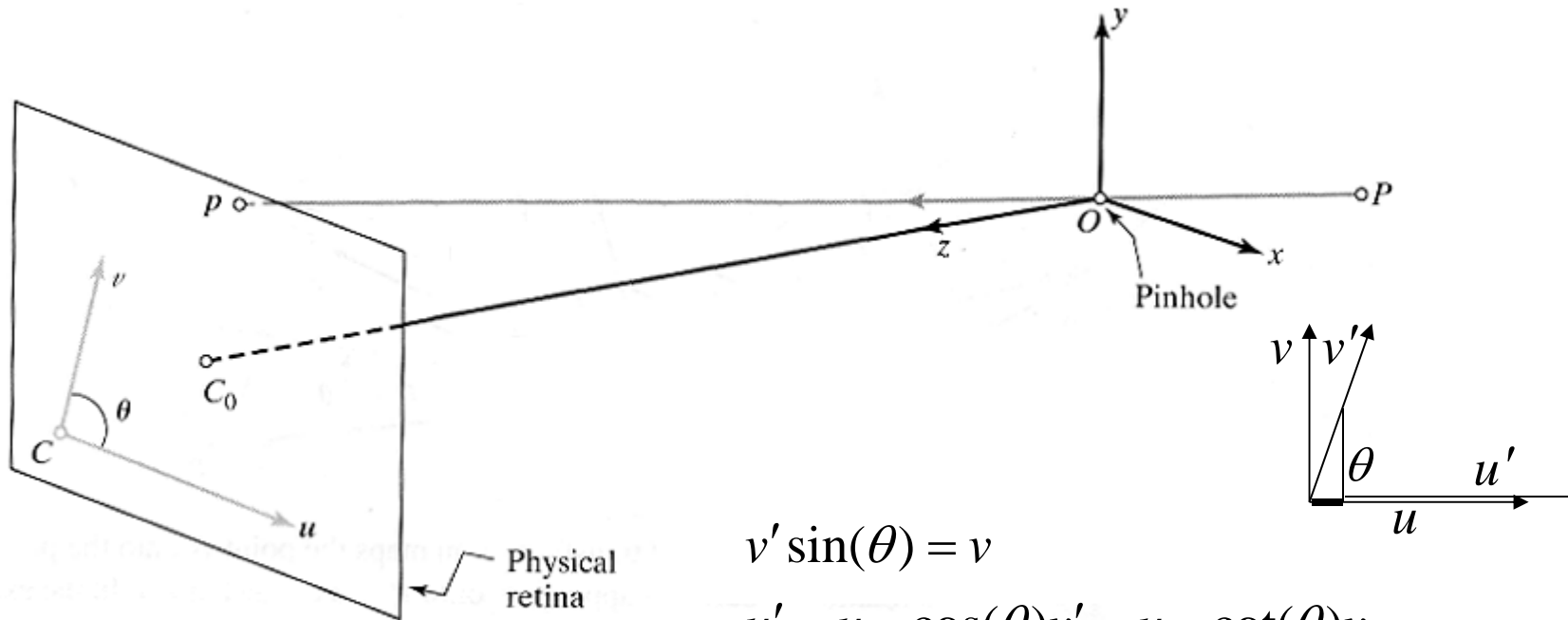
Intrinsic parameters



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



$$v' \sin(\theta) = v$$

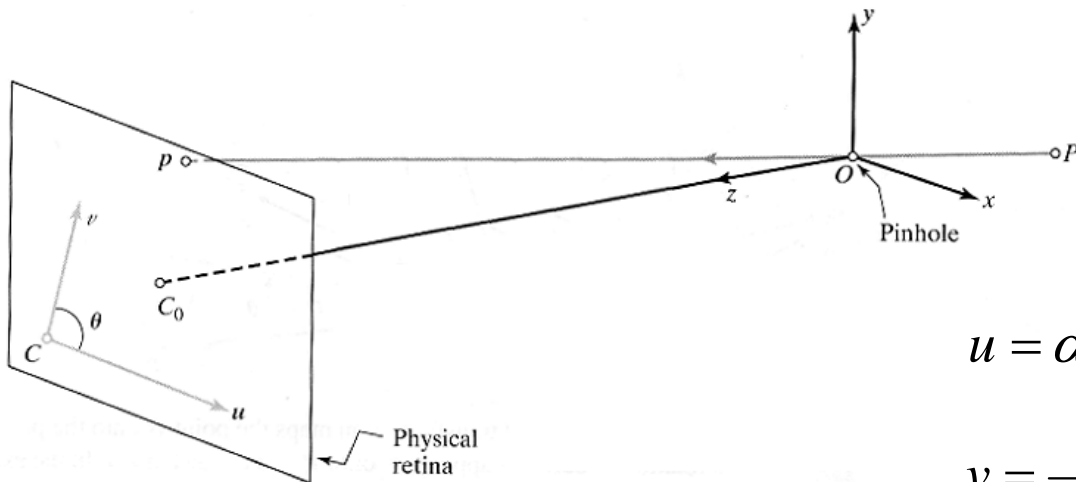
$$u' = u - \cos(\theta)v' = u - \cot(\theta)v$$

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, homogeneous coordinates



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

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

Using homogenous coordinates,
we can write this as:

or:

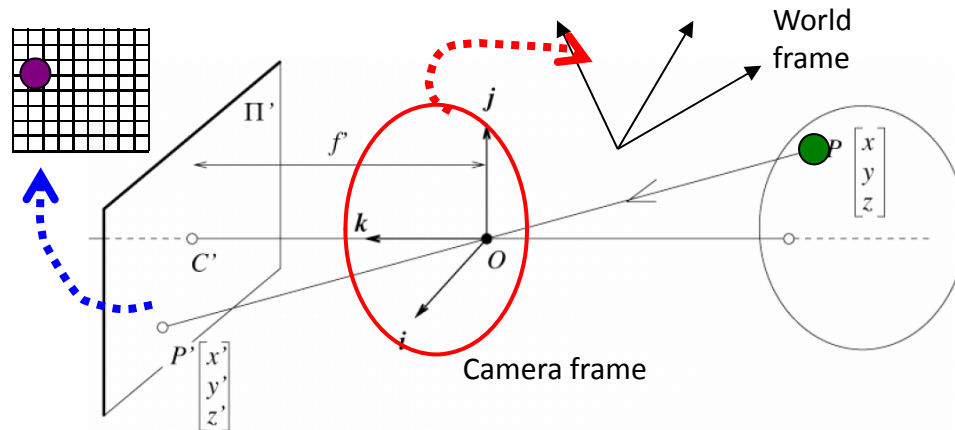
$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \frac{1}{z} \begin{pmatrix} \alpha & -\alpha \cot(\theta) & u_0 & 0 \\ 0 & \frac{\beta}{\sin(\theta)} & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

In pixels

$$\vec{p} = \frac{1}{z} (\mathbf{K}) {}^c \vec{p}$$

In camera-based coords

Perspective projection & calibration



Extrinsic:

Camera frame \leftrightarrow World frame

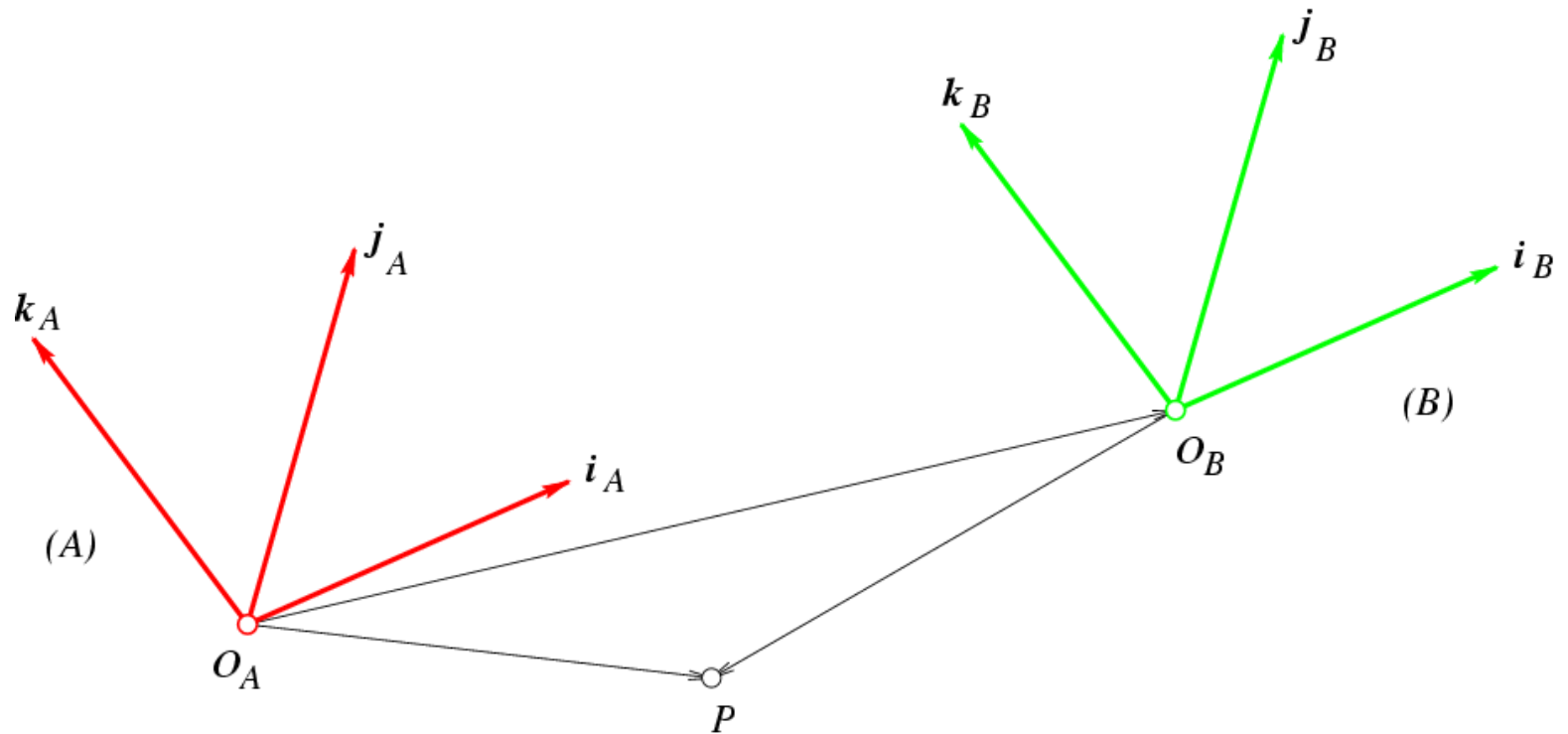
Intrinsic:

Image coordinates relative to camera

\leftrightarrow Pixel coordinates

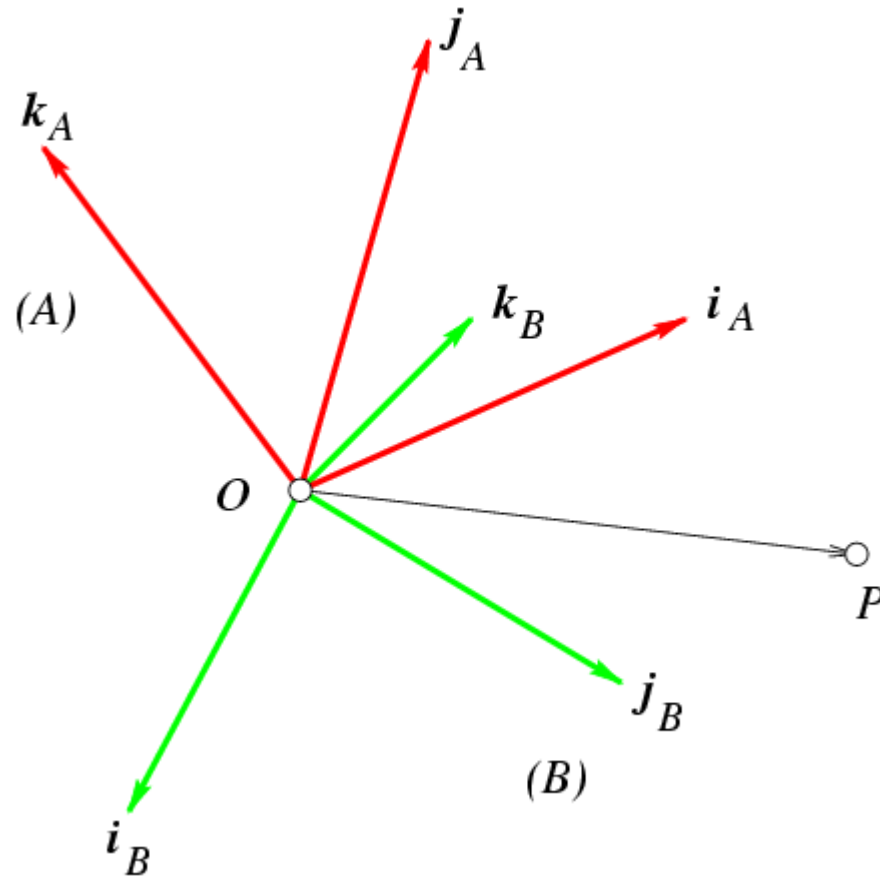
3D
point
(4x1)

Coordinate Changes: Pure Translations



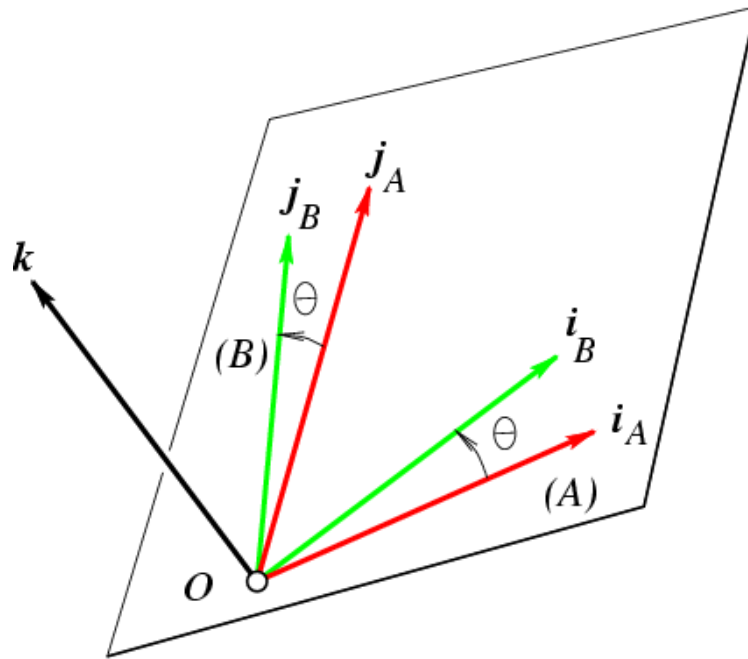
$$\vec{O_B P} = \vec{O_B O_A} + \vec{O_A P} , \quad {}^B P = {}^A P + {}^B O_A$$

Coordinate Changes: Pure Rotations

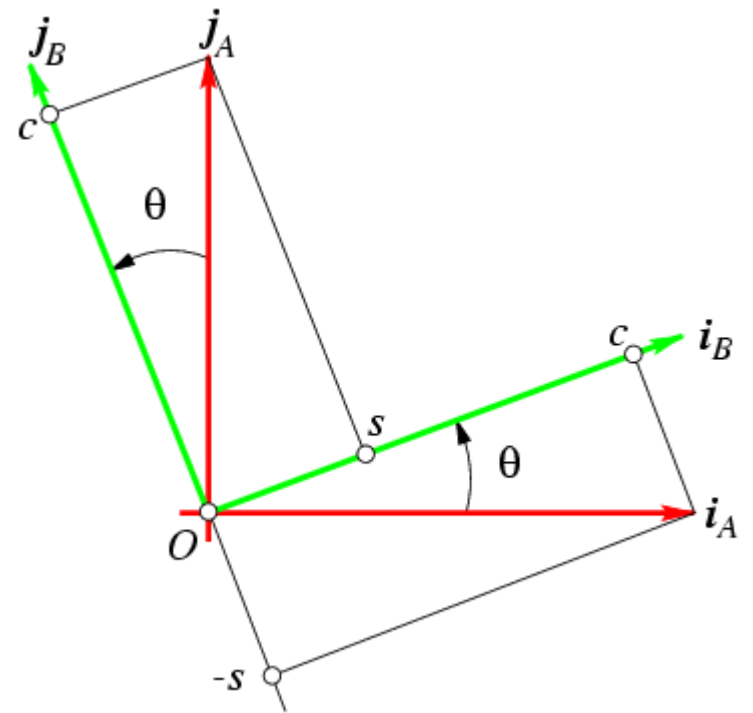


$${}^B_A R = \begin{bmatrix} \mathbf{i}_A \cdot \mathbf{i}_B & \mathbf{j}_A \cdot \mathbf{i}_B & \mathbf{k}_A \cdot \mathbf{i}_B \\ \mathbf{i}_A \cdot \mathbf{j}_B & \mathbf{j}_A \cdot \mathbf{j}_B & \mathbf{k}_A \cdot \mathbf{j}_B \\ \mathbf{i}_A \cdot \mathbf{k}_B & \mathbf{j}_A \cdot \mathbf{k}_B & \mathbf{k}_A \cdot \mathbf{k}_B \end{bmatrix} = ({}^B \mathbf{i}_A, {}^B \mathbf{j}_A, {}^B \mathbf{k}_A) = \begin{bmatrix} {}^A \mathbf{i}_B^T \\ {}^A \mathbf{j}_B^T \\ {}^A \mathbf{k}_B^T \end{bmatrix}$$

Coordinate Changes: Rotations about the k Axis



$${}^B_A R = \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



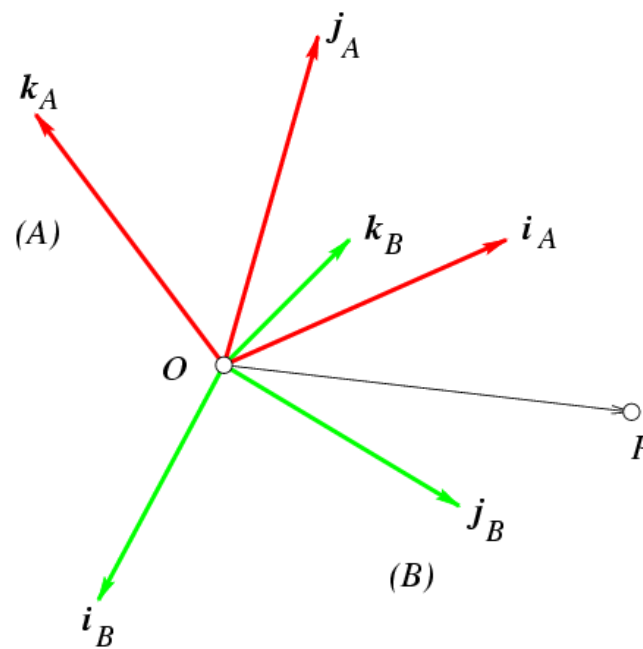
A rotation matrix is characterized by the following properties:

- Its inverse is equal to its transpose, and
- its determinant is equal to 1.

Or equivalently:

- Its rows (or columns) form a right-handed orthonormal coordinate system.

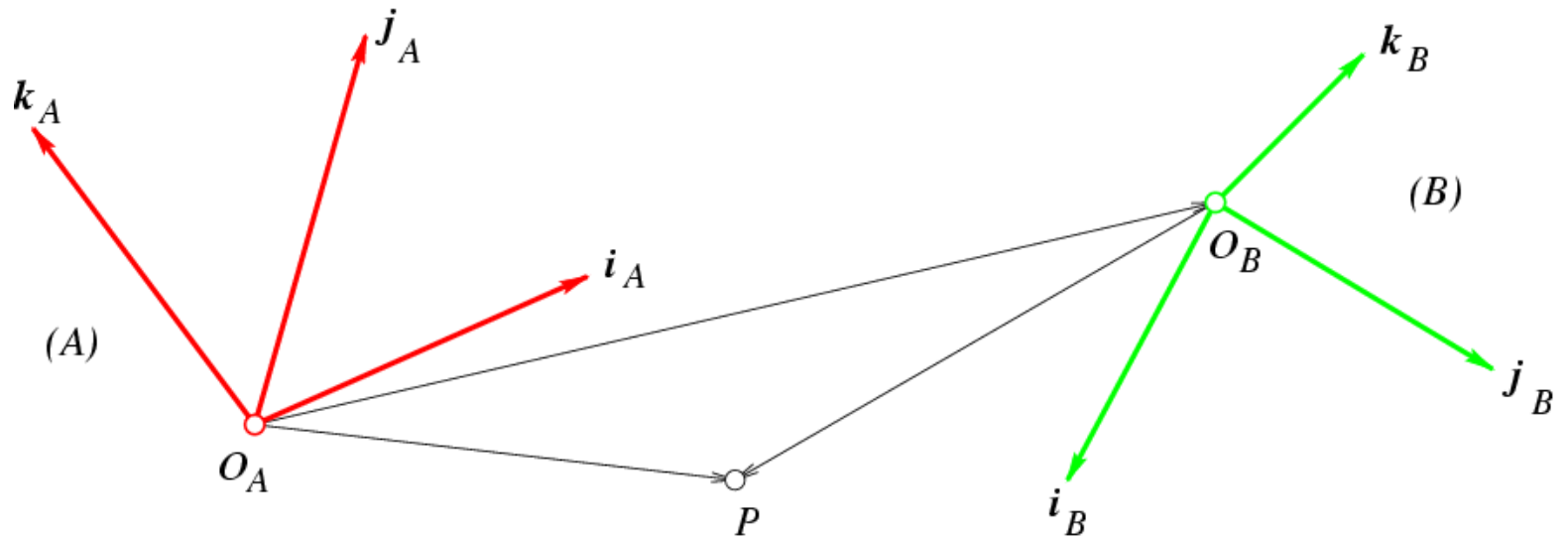
Coordinate Changes: Pure Rotations



$$\overrightarrow{OP} = [\mathbf{i}_A \quad \mathbf{j}_A \quad \mathbf{k}_A] \begin{bmatrix} {}^A x \\ {}^A y \\ {}^A z \end{bmatrix} = [\mathbf{i}_B \quad \mathbf{j}_B \quad \mathbf{k}_B] \begin{bmatrix} {}^B x \\ {}^B y \\ {}^B z \end{bmatrix}$$

$$\Rightarrow {}^B P = {}^B R^A P$$

Coordinate Changes: Rigid Transformations



$${}^B P = {}^B R {}^A P + {}^B O_A$$

Block Matrix Multiplication

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

What is AB ?

$$AB = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}$$

Homogeneous Representation of Rigid Transformations

$$\begin{bmatrix} {}^B P \\ 1 \end{bmatrix} = \begin{bmatrix} {}^B_A R & {}^B O_A \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} {}^A P \\ 1 \end{bmatrix} = \begin{bmatrix} {}^B_A R {}^A P + {}^B O_A \\ 1 \end{bmatrix} = {}^B_A T \begin{bmatrix} {}^A P \\ 1 \end{bmatrix}$$

Extrinsic parameters: translation and rotation of camera frame

$${}^C \vec{p} = {}^C R_W {}^W \vec{p} + {}^C \vec{t}$$

Non-homogeneous
coordinates

$$\begin{pmatrix} {}^C \vec{p} \end{pmatrix} = \begin{pmatrix} - & - & - \\ - & {}^C R_W & - \\ - & - & - \\ \hline 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} | \\ {}^C \vec{t} \\ | \\ \hline 1 \end{pmatrix} \begin{pmatrix} {}^W \vec{p} \end{pmatrix}$$

Homogeneous
coordinates

Combining extrinsic and intrinsic calibration parameters, in homogeneous coordinates

pixels \rightarrow

$$\vec{p} = \frac{1}{z} \mathbf{K} \ ^c \vec{p}$$

Intrinsic

World coordinates \rightarrow

Camera coordinates \rightarrow

$$\begin{pmatrix} \ ^c \vec{p} \\ 1 \end{pmatrix} = \begin{pmatrix} \boxed{\begin{matrix} - & - & - \\ - & \ ^c R & - \\ - & - & - \end{matrix}} & \begin{pmatrix} \ ^c \vec{t} \\ 1 \end{pmatrix} \\ \hline 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \ ^w \vec{p} \\ 1 \end{pmatrix}$$

Extrinsic

$$\vec{p} = \frac{1}{z} \mathbf{K} \begin{pmatrix} \ ^c R & \ ^c \vec{t} \\ 0,0,0 & 1 \end{pmatrix} \ ^w \vec{p}$$

$$\vec{p} = \frac{1}{z} \mathbf{M} \ ^w \vec{p}$$

Other ways to write the same equation

pixel coordinates

world coordinates

$$\vec{p} = \frac{1}{z} M \quad {}^w \vec{p}$$

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \frac{1}{z} \begin{pmatrix} \cdot & m_1^T & \cdot & \cdot \\ \cdot & m_2^T & \cdot & \cdot \\ \cdot & m_3^T & \cdot & \cdot \end{pmatrix} \begin{pmatrix} {}^w p_x \\ {}^w p_y \\ {}^w p_z \\ 1 \end{pmatrix}$$

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

Conversion back from homogeneous coordinates
leads to (note that $z = m_3^T \cdot P$):

Extrinsic Parameters

- When the camera frame (C) is different from the world frame (W),

$$\begin{pmatrix} {}^C P \\ 1 \end{pmatrix} = \begin{pmatrix} {}^C_W \mathcal{R} & {}^C O_W \\ \mathbf{0}^T & 1 \end{pmatrix} \begin{pmatrix} {}^W P \\ 1 \end{pmatrix}.$$

- Thus,

$$\boxed{\mathbf{p} = \frac{1}{z} \mathcal{M} \mathbf{P},}$$
 where

$$\begin{cases} \mathcal{M} = \mathcal{K}(\mathcal{R} \quad \mathbf{t}), \\ \mathcal{R} = {}^C_W \mathcal{R}, \\ \mathbf{t} = {}^C O_W, \\ \mathbf{P} = \begin{pmatrix} {}^W P \\ 1 \end{pmatrix}. \end{cases}$$

- Note: z is *not* independent of \mathcal{M} and \mathbf{P} :

$$\mathcal{M} = \begin{pmatrix} \mathbf{m}_1^T \\ \mathbf{m}_2^T \\ \mathbf{m}_3^T \end{pmatrix} \implies z = \mathbf{m}_3 \cdot \mathbf{P}, \quad \text{or} \quad \begin{cases} u = \frac{\mathbf{m}_1 \cdot \mathbf{P}}{\mathbf{m}_3 \cdot \mathbf{P}}, \\ v = \frac{\mathbf{m}_2 \cdot \mathbf{P}}{\mathbf{m}_3 \cdot \mathbf{P}}. \end{cases}$$

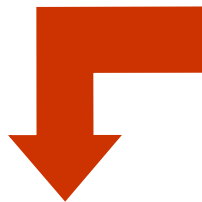
Explicit Form of the Projection Matrix

$$\mathcal{M} = \begin{pmatrix} \alpha \mathbf{r}_1^T - \alpha \cot \theta \mathbf{r}_2^T + u_0 \mathbf{r}_3^T & \alpha t_x - \alpha \cot \theta t_y + u_0 t_z \\ \frac{\beta}{\sin \theta} \mathbf{r}_2^T + v_0 \mathbf{r}_3^T & \frac{\beta}{\sin \theta} t_y + v_0 t_z \\ \mathbf{r}_3^T & t_z \end{pmatrix}$$

Note: If $\mathcal{M} = (\mathcal{A} \quad \mathbf{b})$ then $|\mathbf{a}_3| = 1$.

Replacing \mathcal{M} by $\lambda \mathcal{M}$ in

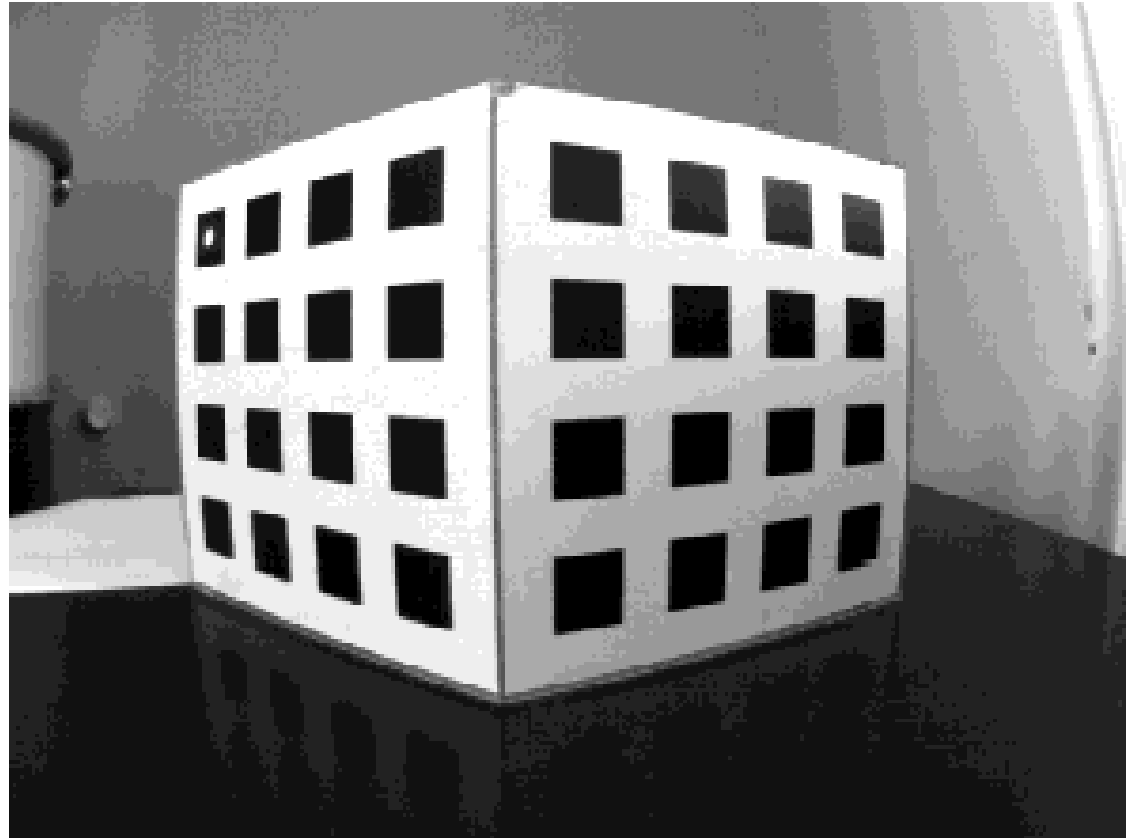
$$\begin{cases} u = \frac{\mathbf{m}_1 \cdot \mathbf{P}}{\mathbf{m}_3 \cdot \mathbf{P}} \\ v = \frac{\mathbf{m}_2 \cdot \mathbf{P}}{\mathbf{m}_3 \cdot \mathbf{P}} \end{cases}$$



does not change u and v .

M is only defined up to scale in this setting!!

Calibration target



The Opti-CAL Calibration Target Image

Find the position, u_i and v_i , in pixels, of each calibration object feature point.

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