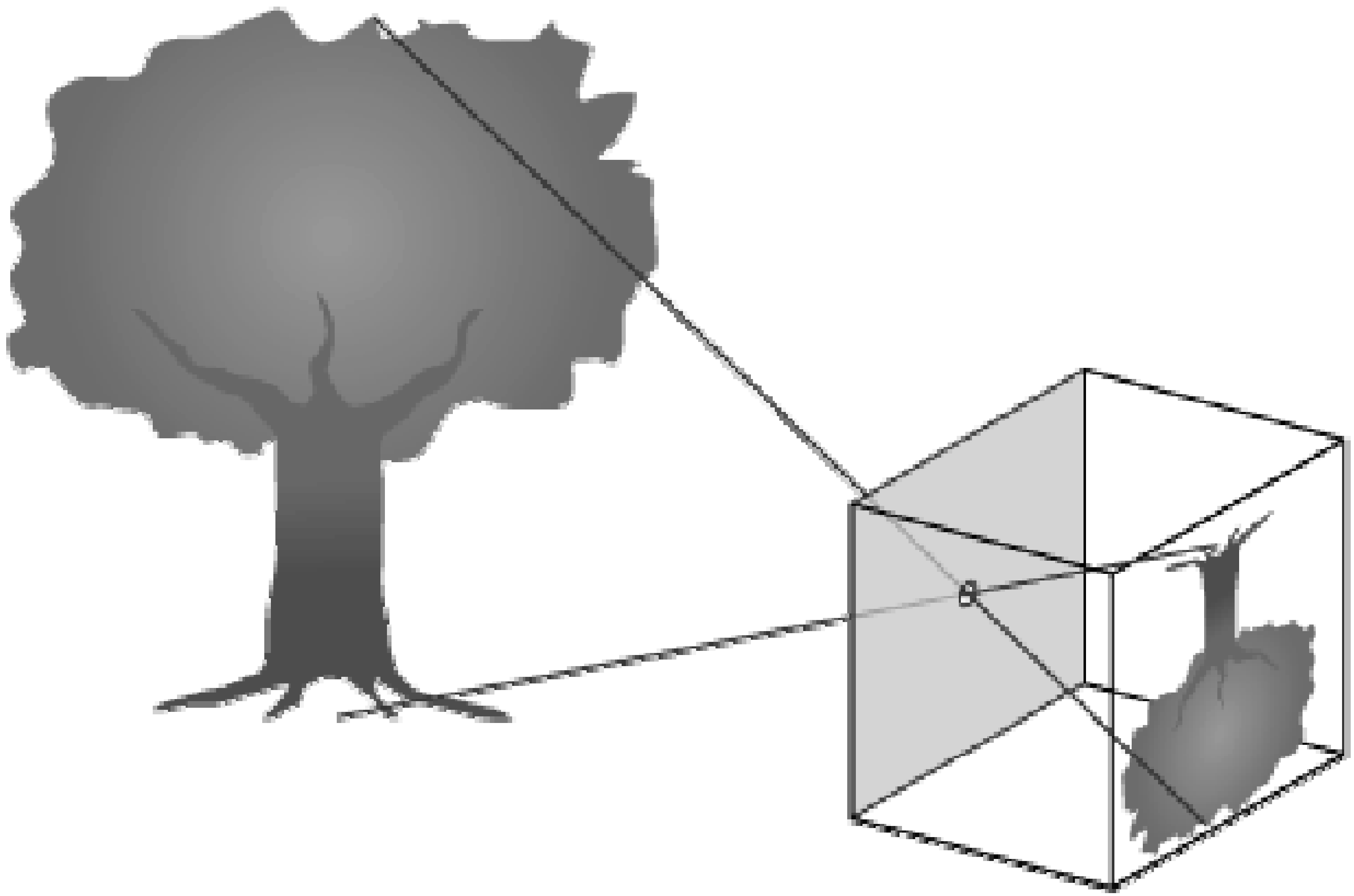
COS429: COMPUTER VISON
CAMERAS AND PROJECTIONS (2 lectures)

- Pinhole cameras
- Analytical Euclidean geometry
- The intrinsic parameters of a camera
- The extrinsic parameters of a camera
- Camera calibration
- Least-squares techniques

Reading: Chapters 1 – 3, Forsyth & Ponce

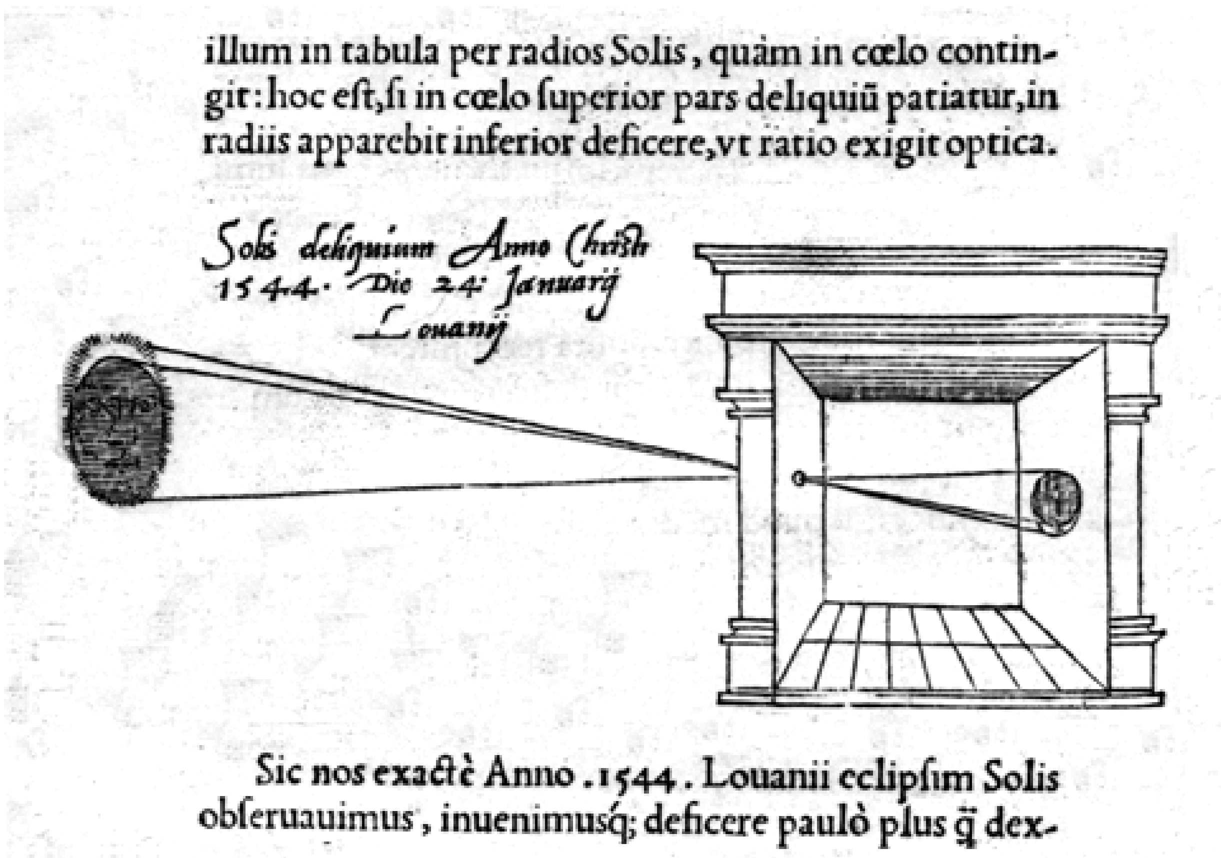
Many of the slides in this lecture are courtesy to Prof. J. Ponce



Some history…

Milestones:

• Leonardo da Vinci (1452-1519): first record of camera obscura


illum in tabula per radios Solis, quàm in cœlo contingit: hoc eſt, ſi in cœlo ſuperior pars deliquiũ patiatur, in radiis apparebit inferior deficere, vt ratio exigit optica.

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

Some history…

Milestones:

• Leonardo da Vinci (1452-1519): first record of camera obscura

• Johann Zahn (1685): first portable camera

Some history…

Milestones:

• Leonardo da Vinci (1452-1519): first record of camera obscura

• Johann Zahn (1685): first portable camera

• Joseph Nicephore Niepce (1822): first photo - birth of photography

Photography (Niepce, “La Table Servie,” 1822)

Some history…

Milestones:

• Leonardo da Vinci (1452-1519): first record of camera obscura
• Johann Zahn (1685): first portable camera
• Joseph Nicephore Niepce (1822): first photo - birth of photography
• Daguerréotypes (1839)
• Photographic Film (Eastman, 1889)
• Cinema (Lumière Brothers, 1895)
• Color Photography (Lumière Brothers, 1908)
• Television (Baird, Farnsworth, Zworykin, 1920s)

Photography (Niepce, “La Table Servie,” 1822)

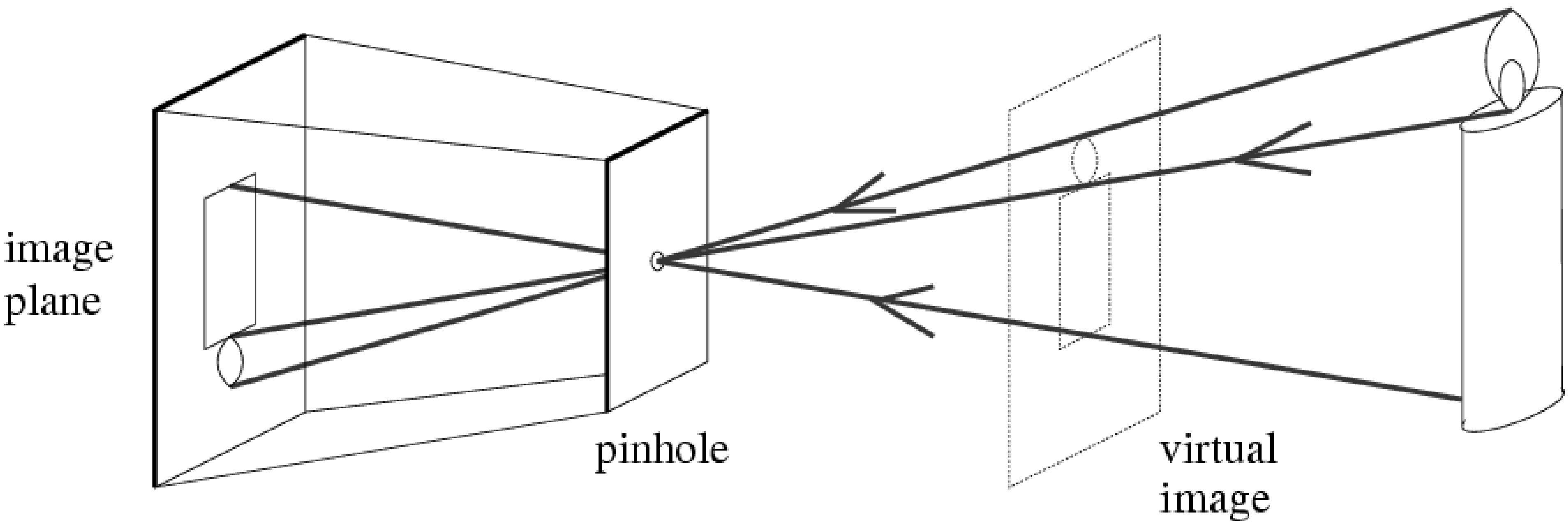
Let’s also not forget…

Motzu
(468-376 BC)

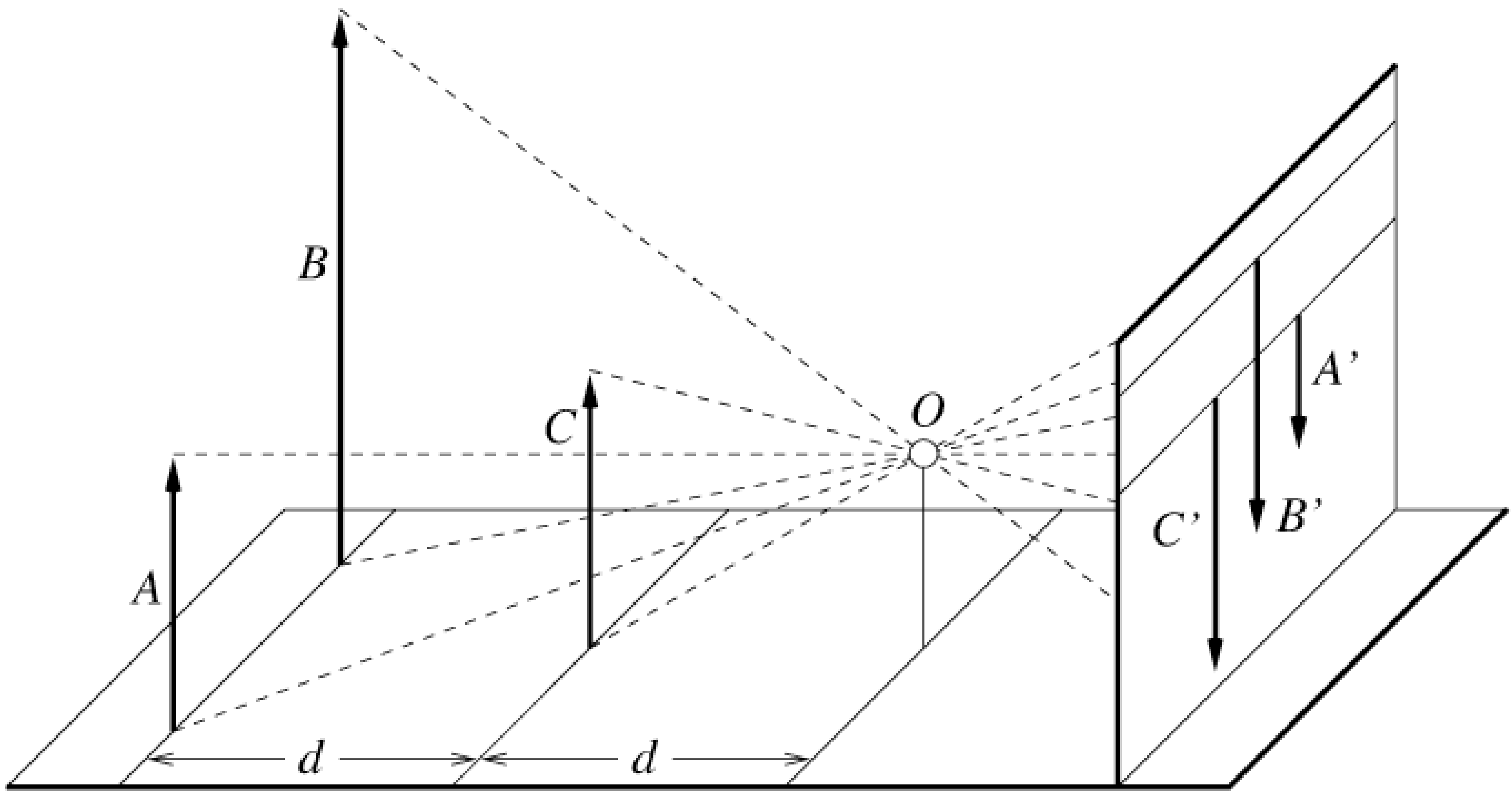
Aristotle
(384-322 BC)

Ibn al-Haitham
(965-1040)

Pinhole perspective projection

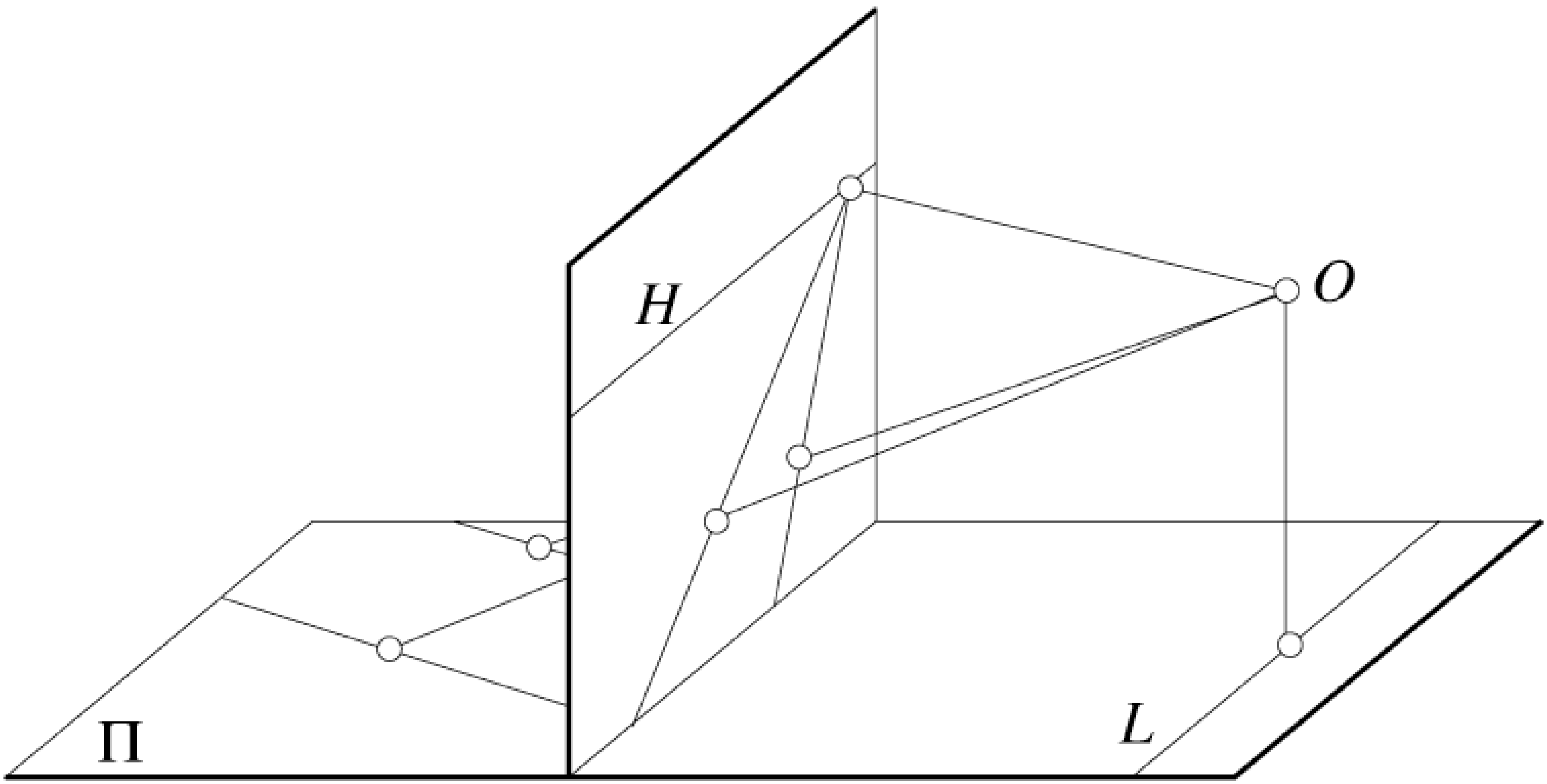


Distant objects are smaller



Parallel lines meet

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



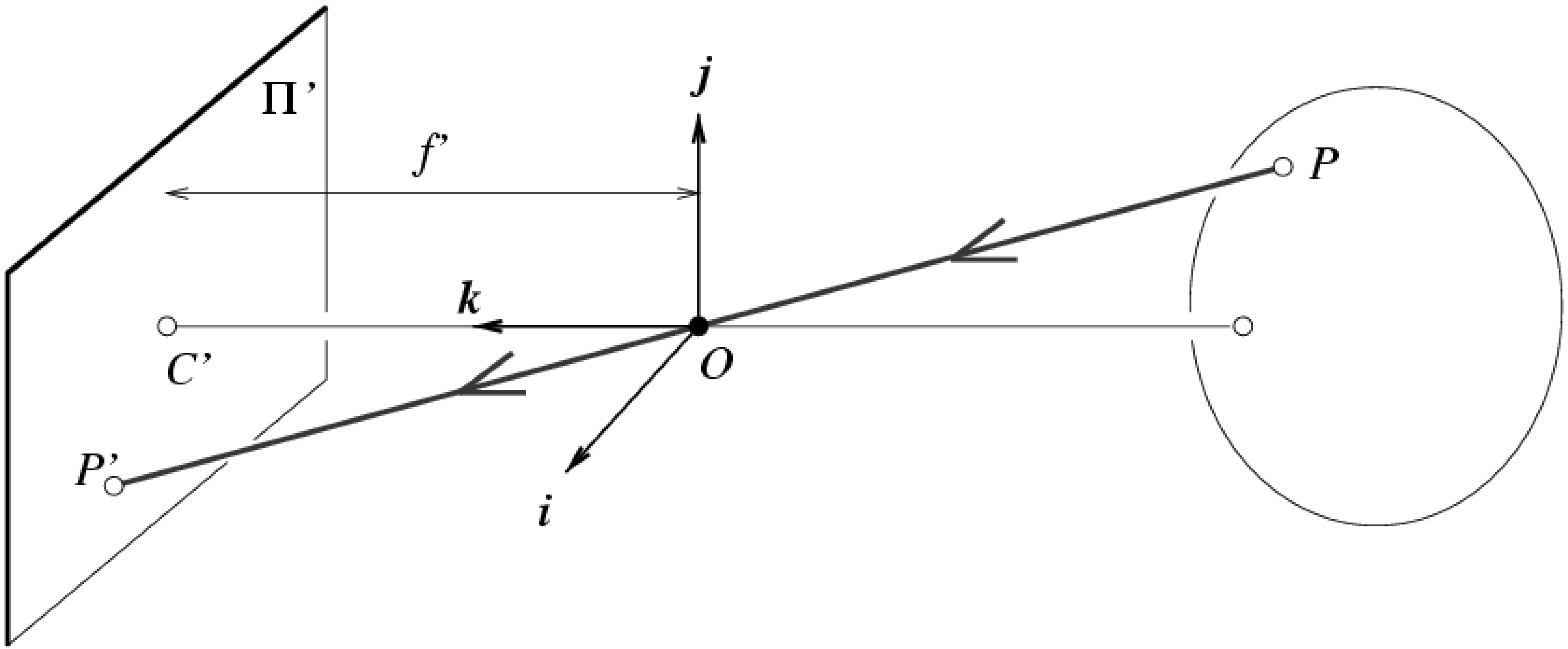
Vanishing points

- Each set of parallel lines 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

Properties of Projection

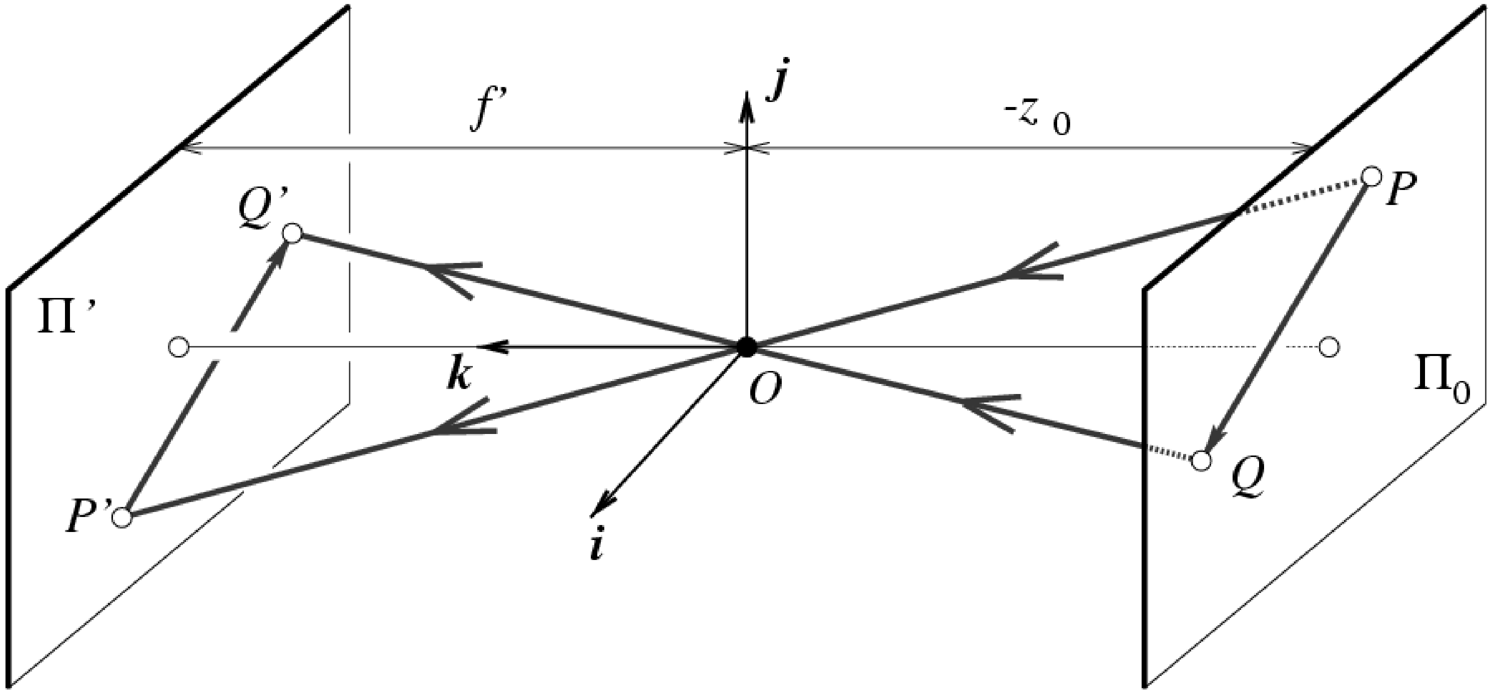
- Points project to points
- Lines project to lines
- Planes project to the whole image or a half image
- Angles are not preserved
- Degenerate cases
  - Line through focal point projects to a point.
  - Plane through focal point projects to line
  - Plane perpendicular to image plane projects to part of the image (with horizon).

Pinhole Perspective Equation



NOTE: *z* is always negative..

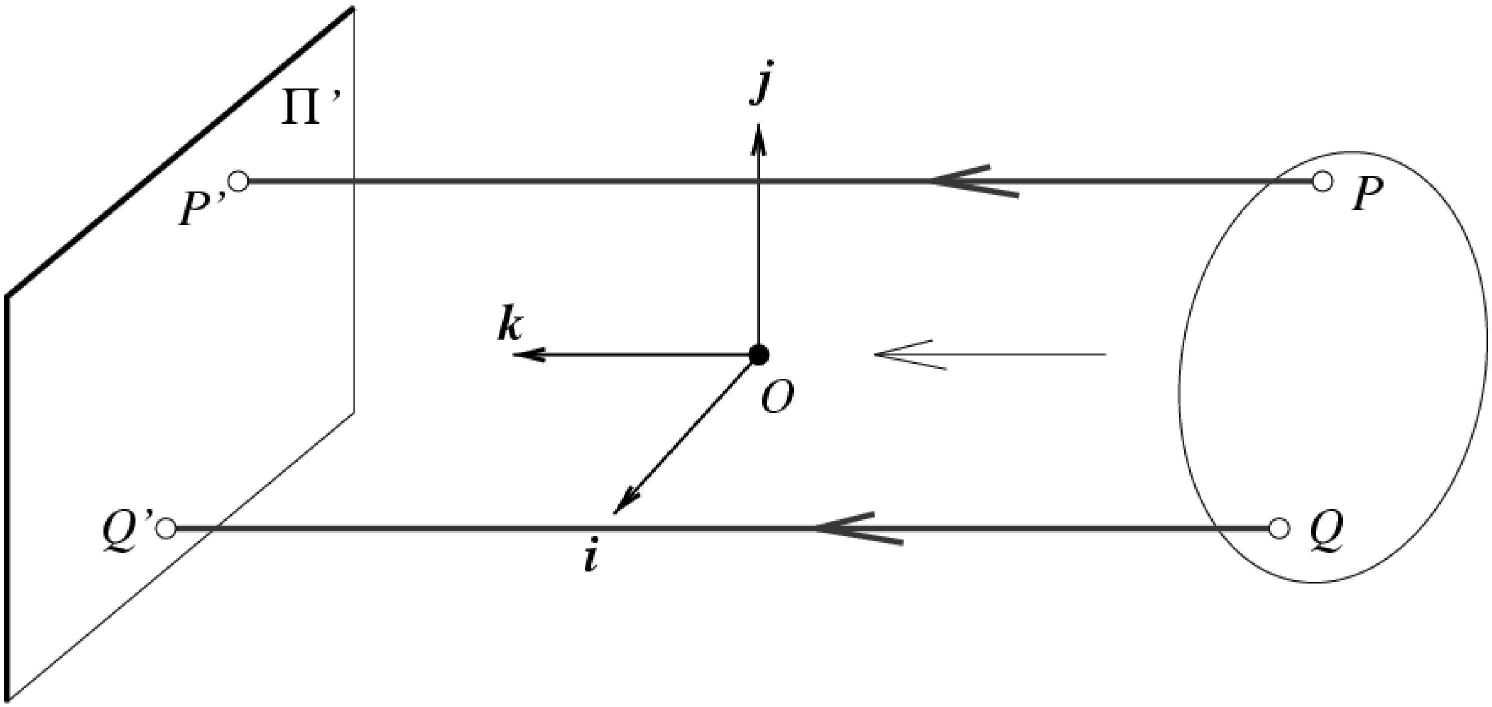
Affine projection models: Weak perspective projection



is the magnification.

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

Affine projection models: Orthographic projection



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

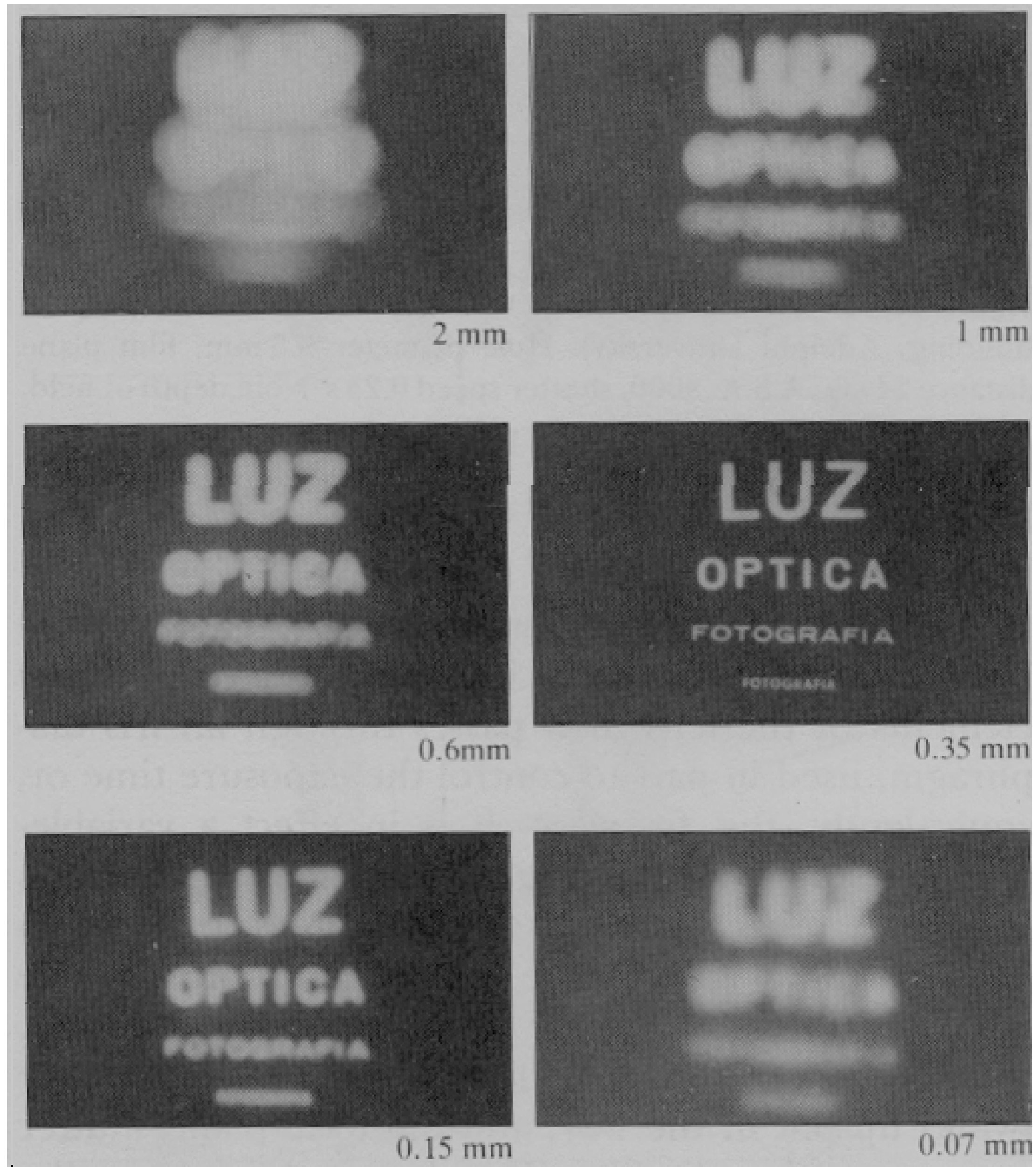
Pros and Cons of These Models

- Weak perspective much simpler math.
  - Accurate when object is small and distant.
  - Most useful for recognition.
- Pinhole perspective much more accurate for scenes.
  - Used in structure from motion.
- When accuracy really matters, must model real cameras.

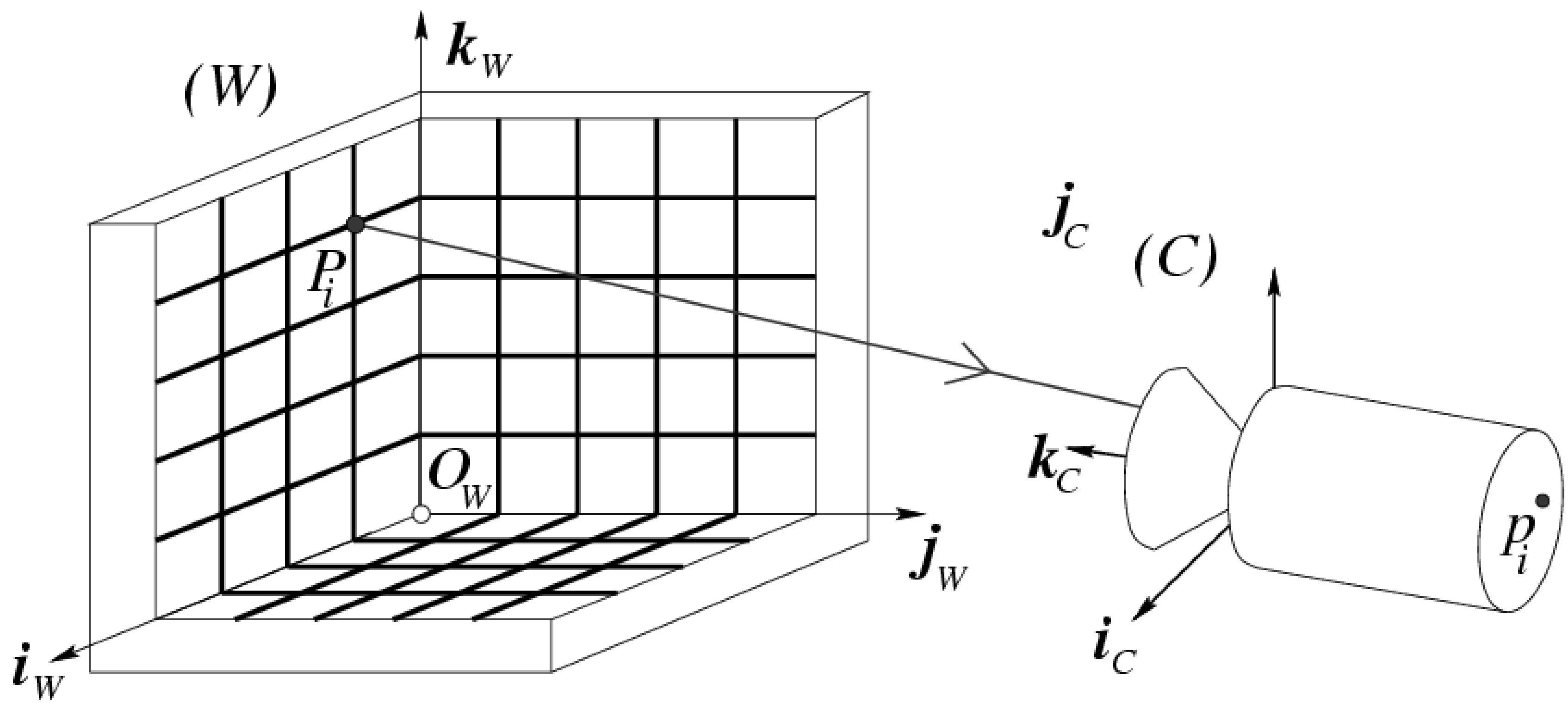
Diffraction effects in pinhole cameras.

Shrinking pinhole size

Use a lens!

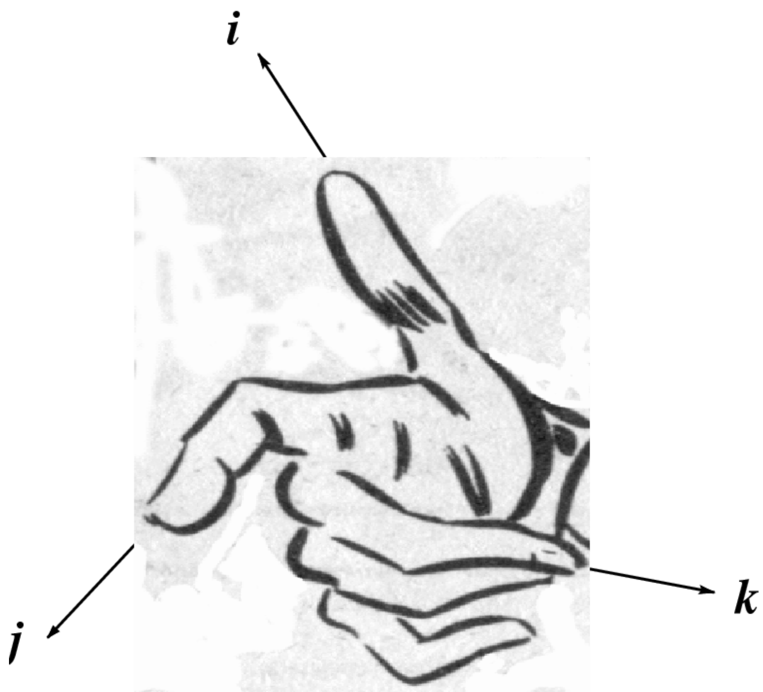


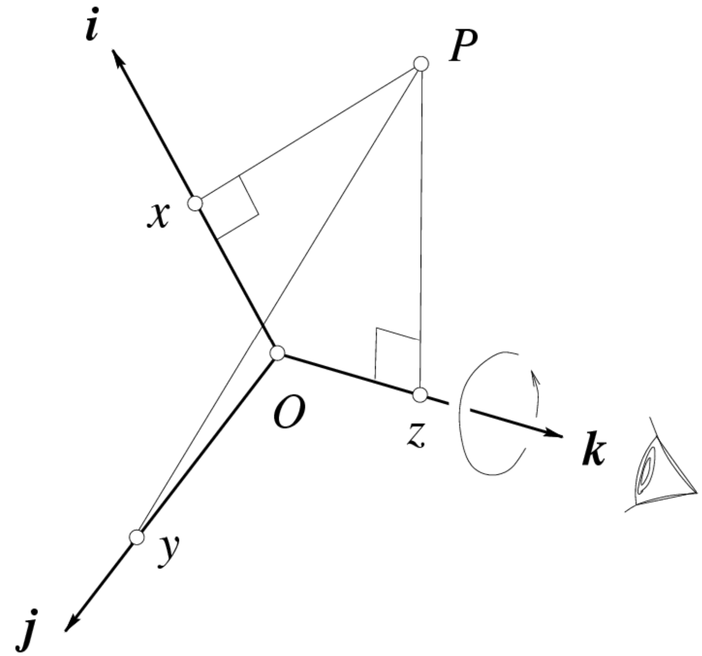
Quantitative Measurements and Calibration



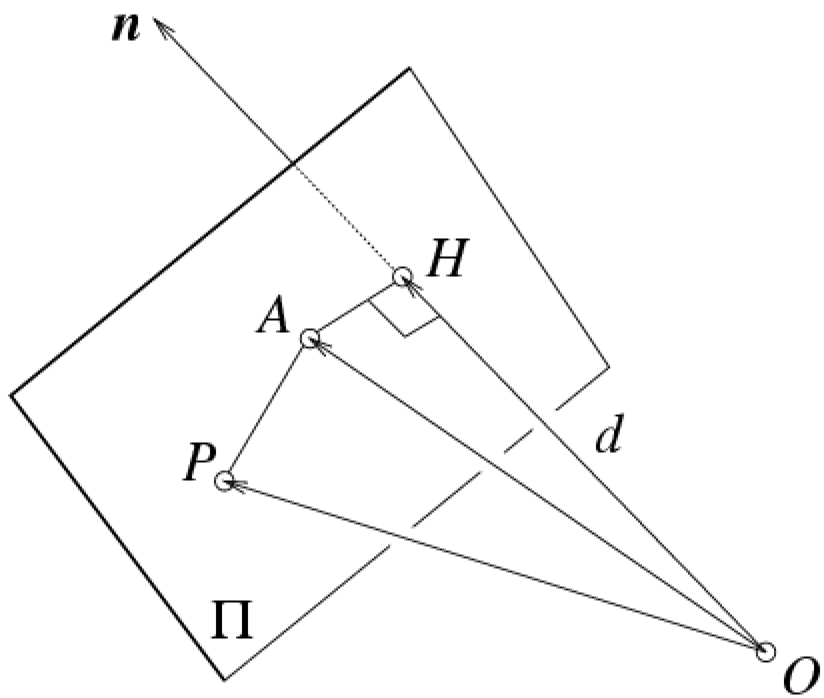
Euclidean Geometry

Euclidean Coordinate Systems

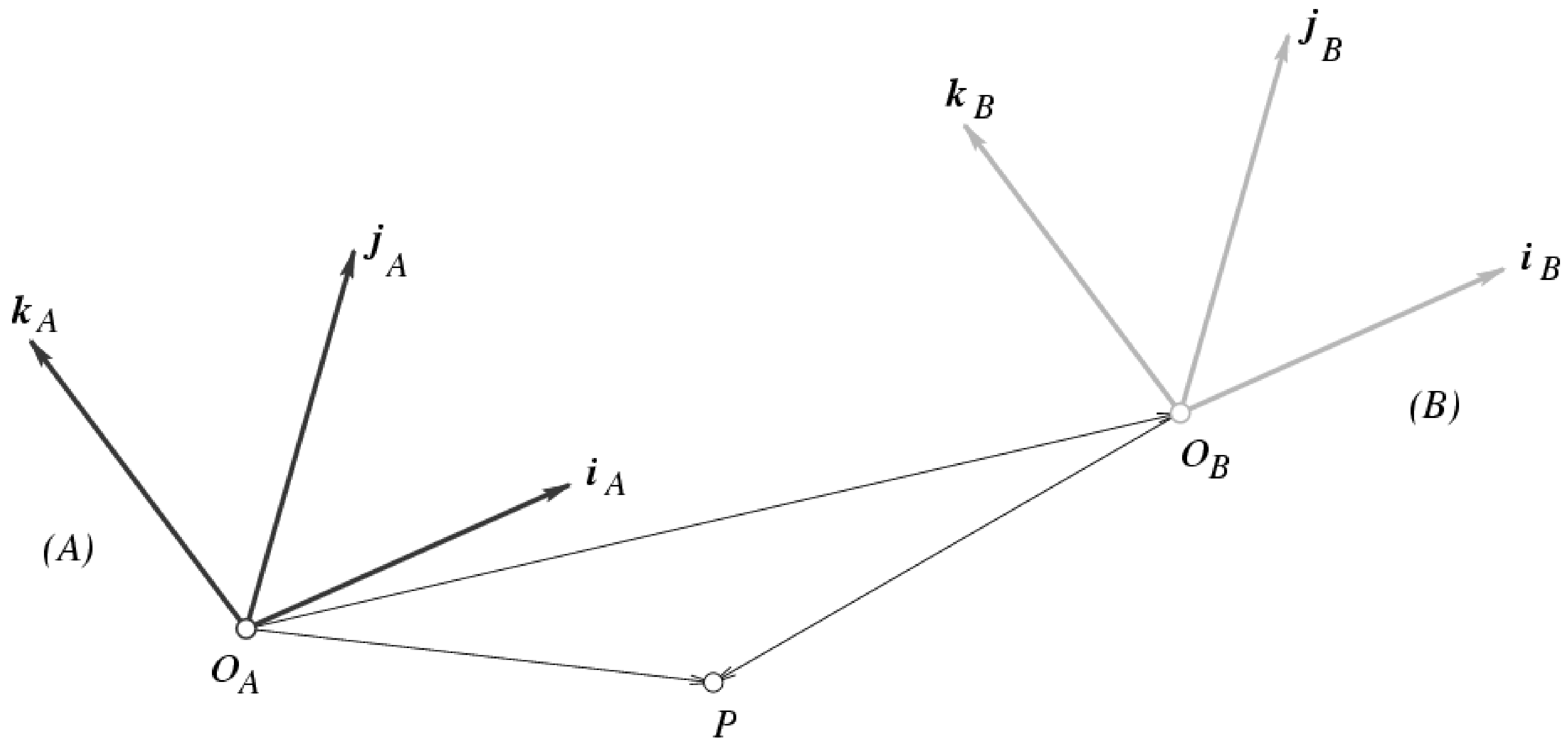




Planes

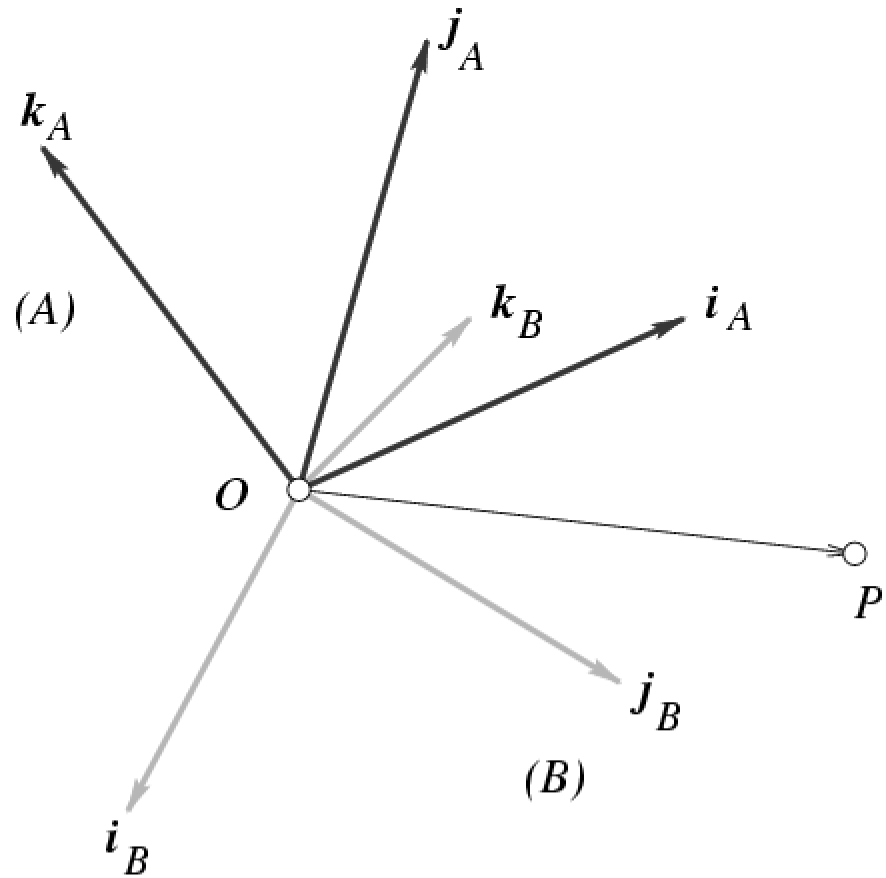


Coordinate Changes: Pure Translations

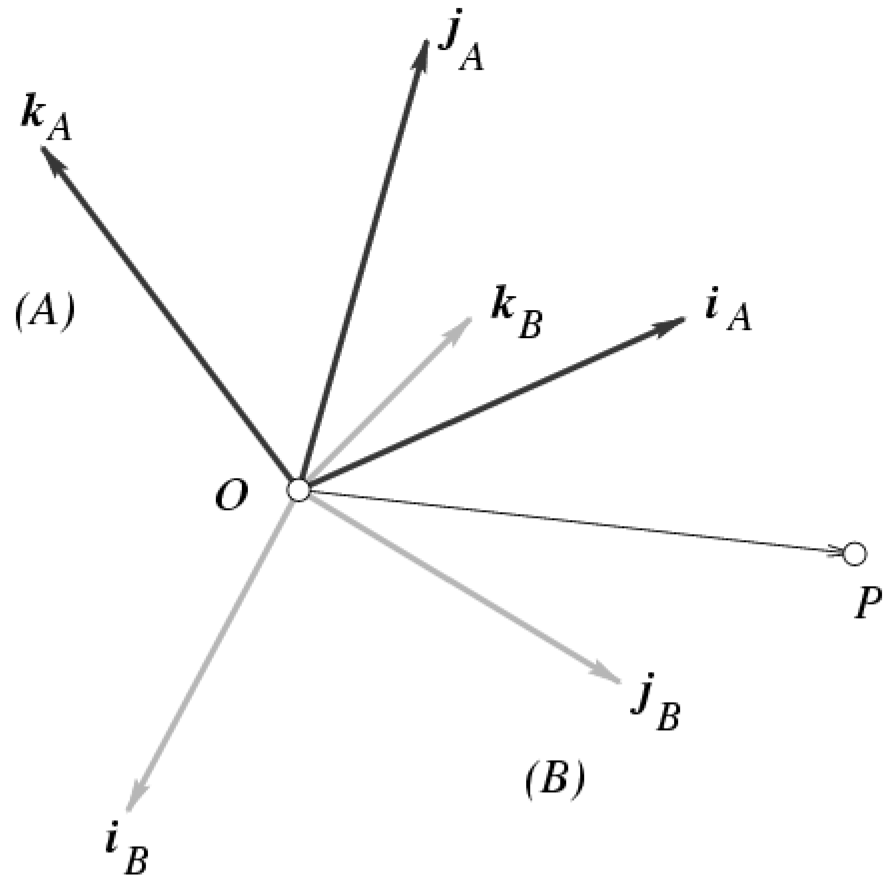


$$\overrightarrow{O_B P} = \overrightarrow{O_B O_A} + \overrightarrow{O_A P} \ , \quad {}^B P = {}^A P + {}^B O_A$$

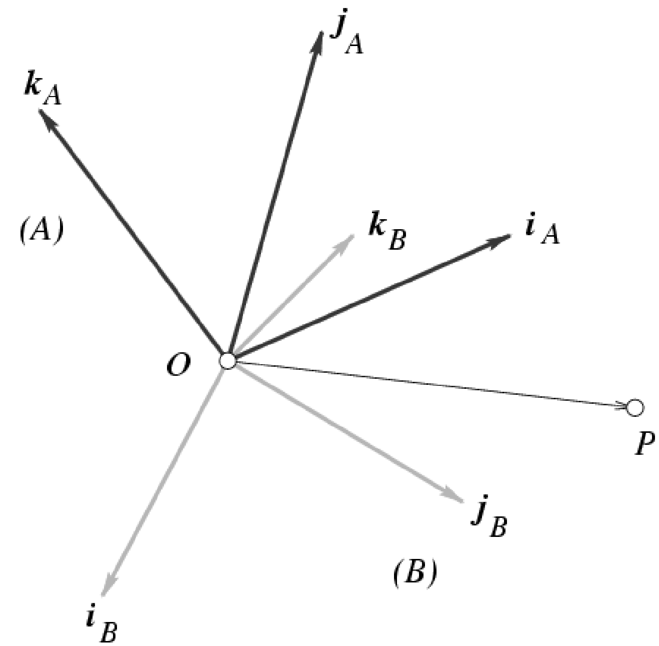
Coordinate Changes: Pure Rotations



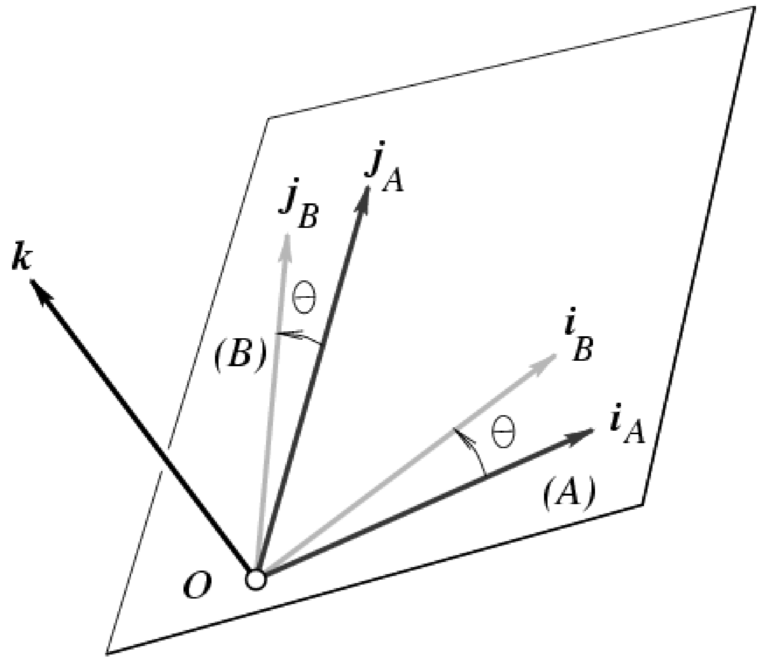
Coordinate Changes: Pure Rotations

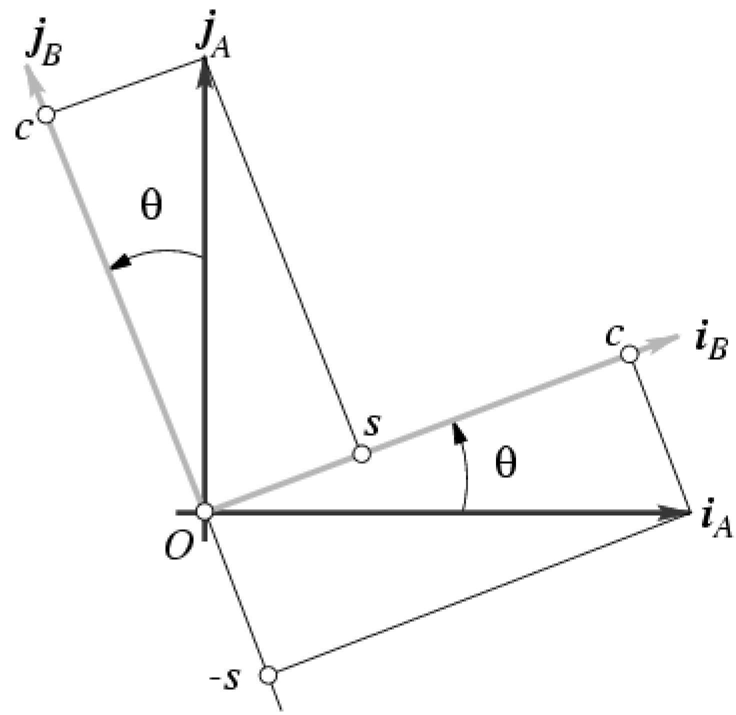


Coordinate Changes: Pure Rotations



Coordinate Changes: Rotations about the z Axis





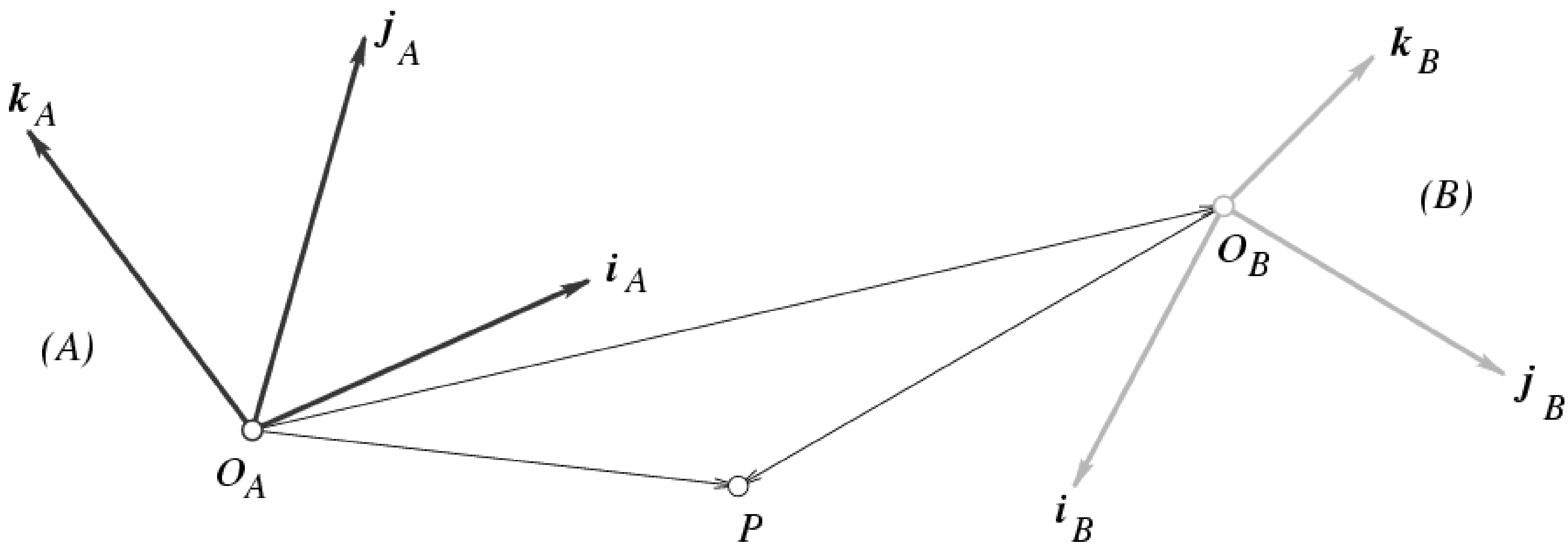
A rotation matrix is characterized by the following properties:

- 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: Rigid Transformations

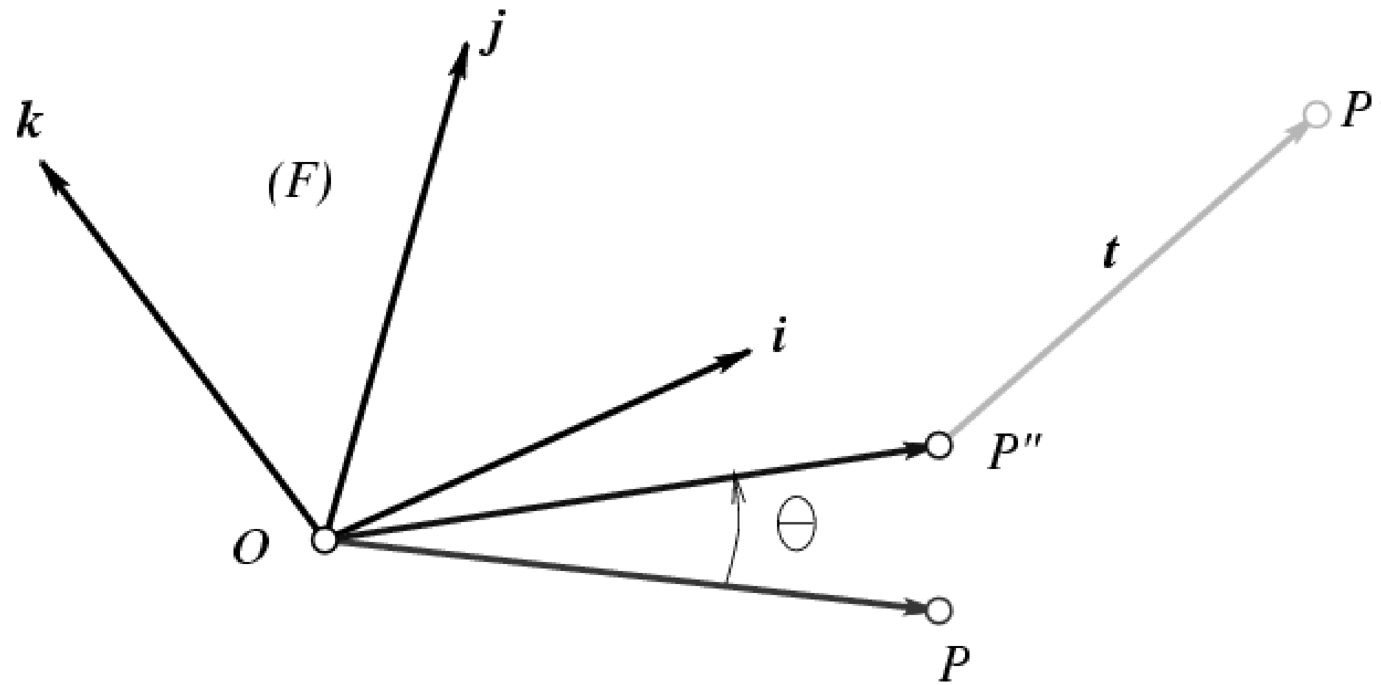


Block Matrix Multiplication

What is AB ?

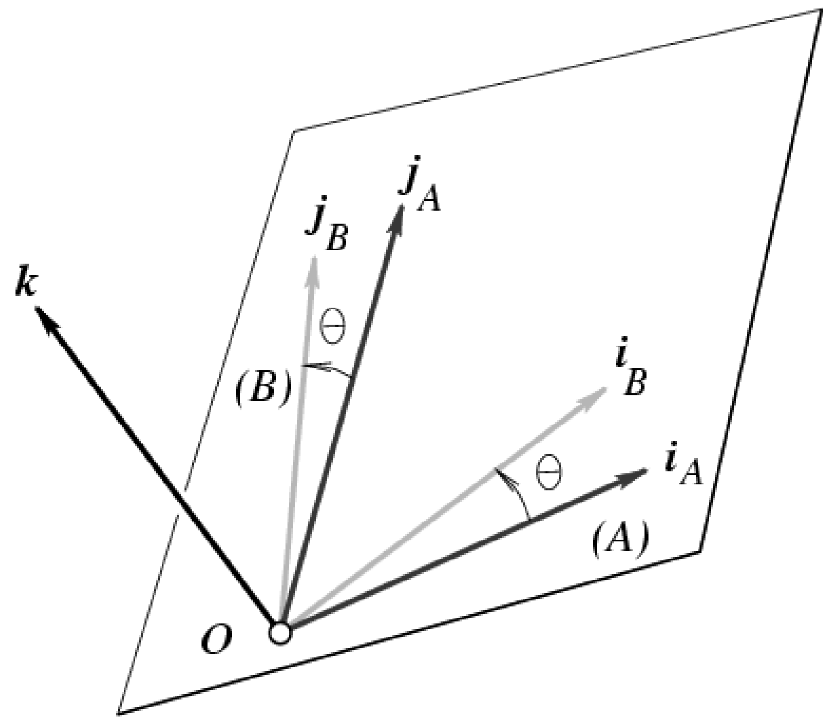
Homogeneous Representation of Rigid Transformations

Rigid Transformations as Mappings



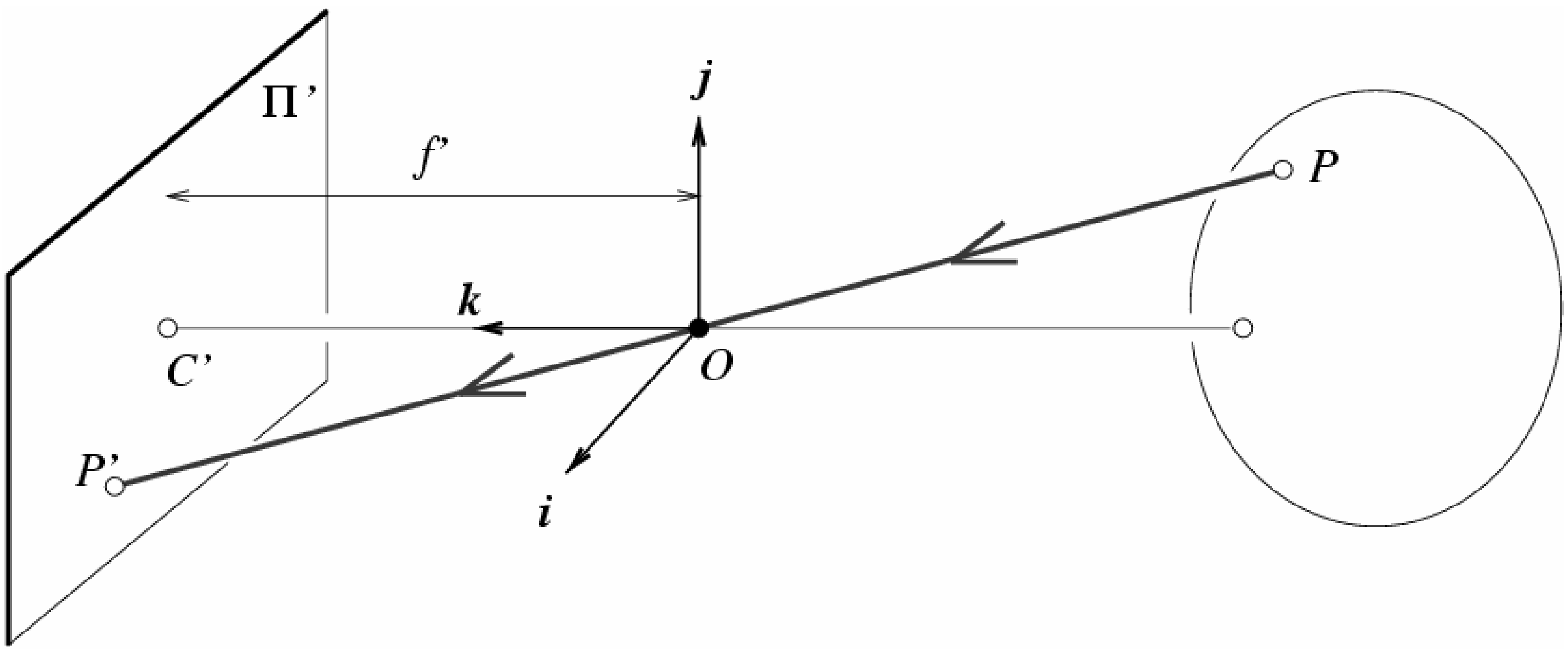
$$^{F}P' = \mathcal{R}\,{}^{F}P + \boldsymbol{t} \iff \begin{pmatrix} {}^{F}P' \\ 1 \end{pmatrix} = \begin{pmatrix} \mathcal{R} & \boldsymbol{t} \\ \mathbf{0}^T & 1 \end{pmatrix} \begin{pmatrix} {}^{F}P \\ 1 \end{pmatrix}$$

Rigid Transformations as Mappings: Rotation about the **k** Axis



$$
{}^F P' = \mathcal{R}\,{}^F P, \quad \text{where} \quad \mathcal{R} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}
$$

Pinhole Perspective Equation



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

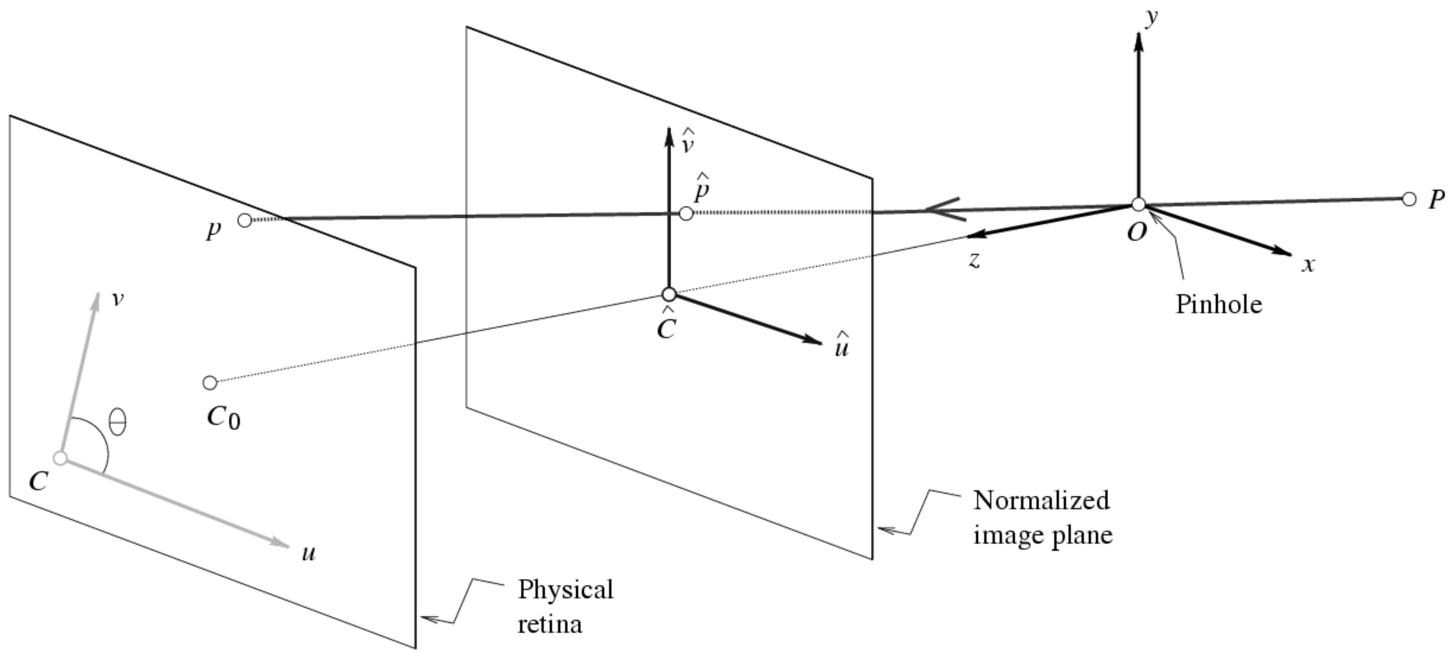
The Intrinsic Parameters of a Camera

Units:

k,l : pixel/m

f : m

α,β : pixel



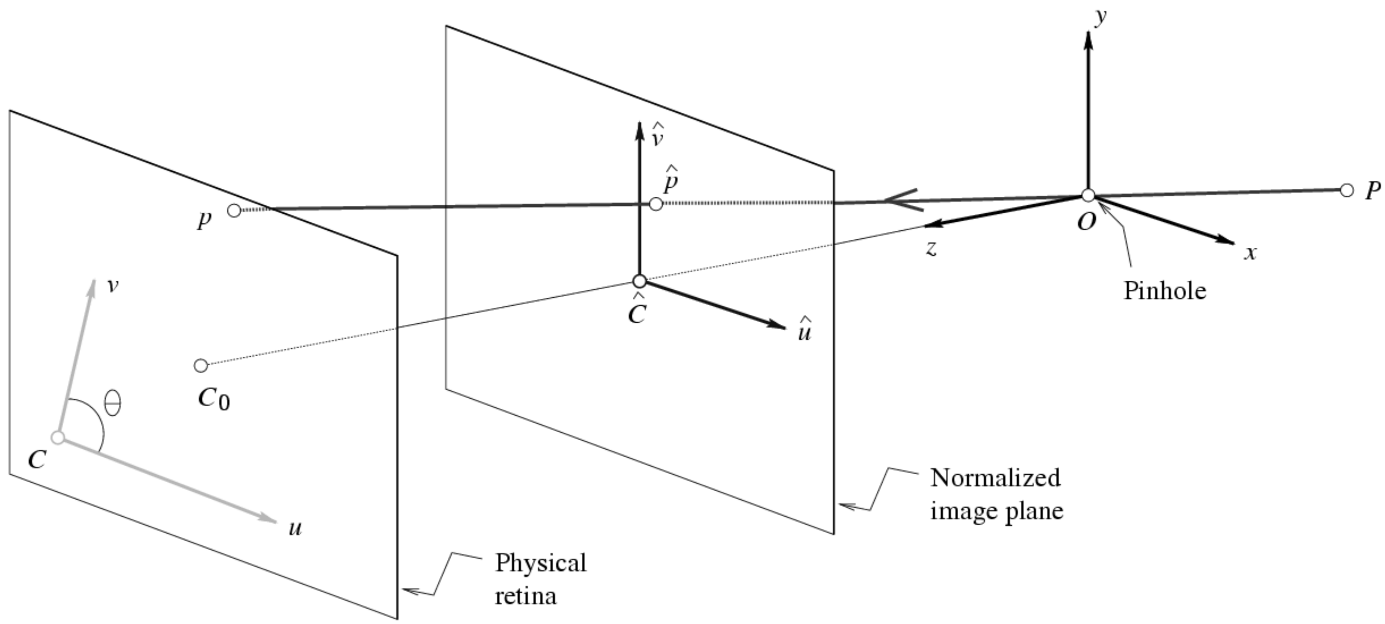
$$\begin{cases} \hat{u} = \dfrac{x}{z} \\ \hat{v} = \dfrac{y}{z} \end{cases} \Longleftrightarrow \hat{\boldsymbol{p}} = \frac{1}{z} \begin{pmatrix} \mathrm{Id} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{P} \\ 1 \end{pmatrix}$$

Normalized Image Coordinates

Physical Image Coordinates

$$\begin{cases} u = kf\dfrac{x}{z} \\ v = lf\dfrac{y}{z} \end{cases} \to \begin{cases} u = \alpha\dfrac{x}{z} + u_0 \\ v = \beta\dfrac{y}{z} + v_0 \end{cases} \to \begin{cases} u = \alpha\dfrac{x}{z} - \alpha\cot\theta\dfrac{y}{z} + u_0 \\ v = \dfrac{\beta}{\sin\theta}\dfrac{y}{z} + v_0 \end{cases}$$

The Intrinsic Parameters of a Camera



Calibration Matrix

$$\boldsymbol{p} = \mathcal{K}\hat{\boldsymbol{p}}, \quad \text{where} \quad \boldsymbol{p} = \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \quad \text{and} \quad \mathcal{K} \stackrel{\text{def}}{=} \begin{pmatrix} \alpha & -\alpha\cot\theta & u_0 \\ 0 & \dfrac{\beta}{\sin\theta} & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

The Perspective Projection Equation

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

The Extrinsic Parameters of a Camera

- 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,

$$\boxed{\boldsymbol{p} = \frac{1}{z}\mathcal{M}\boldsymbol{P},} \quad \text{where} \quad \begin{cases} \mathcal{M} = \mathcal{K}\,(\,\mathcal{R} \quad \boldsymbol{t}\,)\,, \\ \mathcal{R} = {}^C_W\mathcal{R}, \\ \boldsymbol{t} = {}^C O_W, \\ \boldsymbol{P} = \begin{pmatrix} {}^W P \\ 1 \end{pmatrix}. \end{cases}$$

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

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

Explicit Form of the Projection Matrix

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

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

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

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

does not change u and v.

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

Theorem (Faugeras, 1993)

Let $\mathcal{M} = (\mathcal{A} \quad \boldsymbol{b})$ be a 3×4 matrix and let $\boldsymbol{a}_i^T$ $(i = 1, 2, 3)$ denote the rows of the matrix $\mathcal{A}$ formed by the three leftmost columns of $\mathcal{M}$.

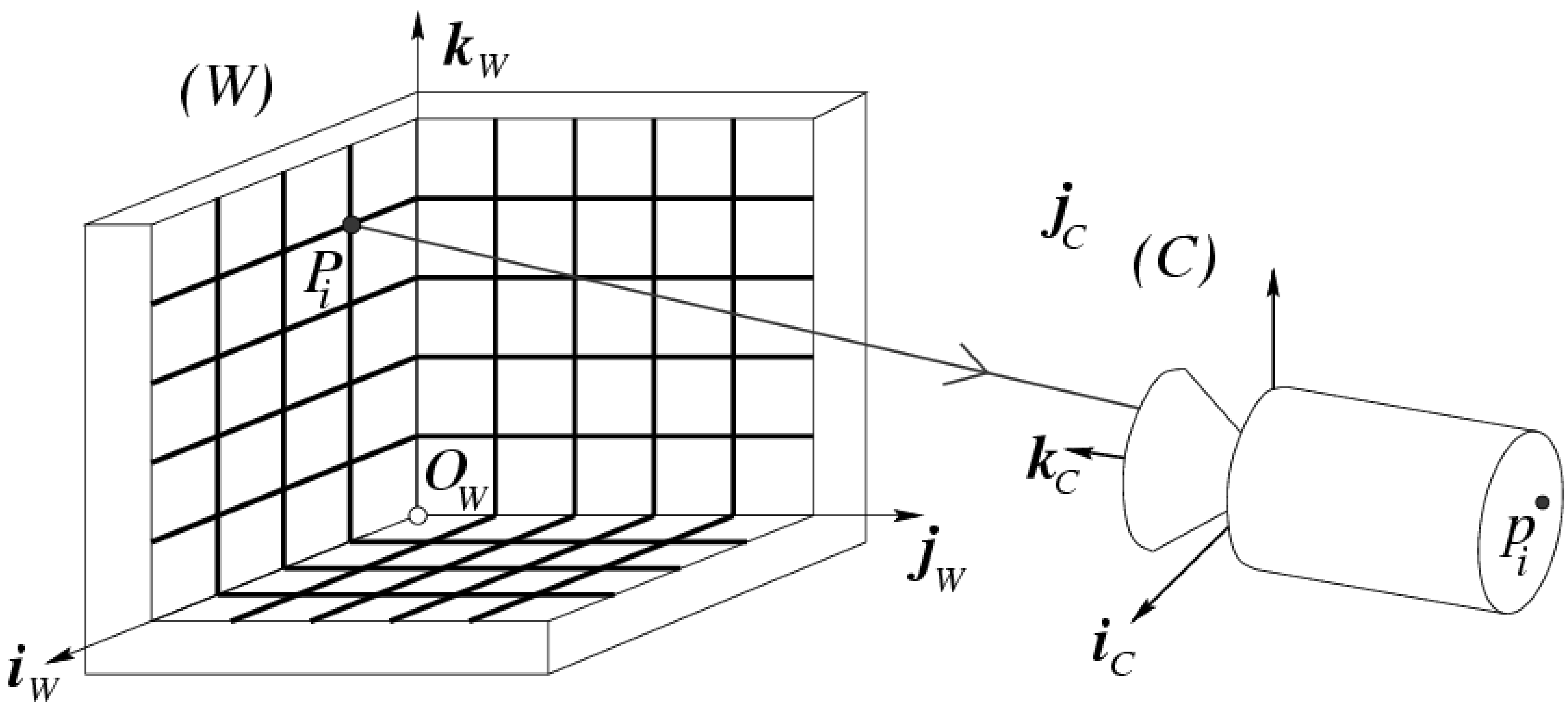
- A necessary and sufficient condition for $\mathcal{M}$ to be a perspective projection matrix is that $\mathrm{Det}(\mathcal{A}) \neq 0$.
- A necessary and sufficient condition for $\mathcal{M}$ to be a zero-skew perspective projection matrix is that $\mathrm{Det}(\mathcal{A}) \neq 0$ and

$$(\boldsymbol{a}_1 \times \boldsymbol{a}_3) \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3) = 0.$$

- A necessary and sufficient condition for $\mathcal{M}$ to be a perspective projection matrix with zero skew and unit aspect-ratio is that $\mathrm{Det}(\mathcal{A}) \neq 0$ and

$$\begin{cases} (\boldsymbol{a}_1 \times \boldsymbol{a}_3) \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3) = 0, \\ (\boldsymbol{a}_1 \times \boldsymbol{a}_3) \cdot (\boldsymbol{a}_1 \times \boldsymbol{a}_3) = (\boldsymbol{a}_2 \times \boldsymbol{a}_3) \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3). \end{cases}$$

Quantitative Measurements and the Calibration Problem



Calibration Procedure

- Calibration target : 2 planes at right angle with checkerboard (Tsai grid)
- We know positions of corners of grid with respect to a coordinate system of the target
- Obtain from images the corners
- Using the equations (relating pixel coordinates to world coordinates) we obtain the camera parameters (the internal parameters and the external (pose) as a side effect)

Estimation procedure

- First estimate M from corresponding image points and scene points (solving homogeneous equation)
- Second decompose M into internal and external parameters
- Use estimated parameters as starting point to solve calibration parameters non-linearly.

Homogeneous Linear Systems

$$A \boldsymbol{x} = \boldsymbol{0}$$

Square system:

- unique solution: 0
- unless Det(A)=0

$$A \boldsymbol{x} = \boldsymbol{0}$$

Rectangular system ??

- 0 is always a solution

Minimize $|Ax|^2$
under the constraint $|x|^2=1$

How do you solve overconstrained homogeneous linear equations ??

$$E = |\mathcal{U}\boldsymbol{x}|^2 = \boldsymbol{x}^T(\mathcal{U}^T\mathcal{U})\boldsymbol{x}$$

- Orthonormal basis of eigenvectors: $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_q$.
- Associated eigenvalues: $0 \leq \lambda_1 \leq \ldots \leq \lambda_q$.
- Any vector can be written as

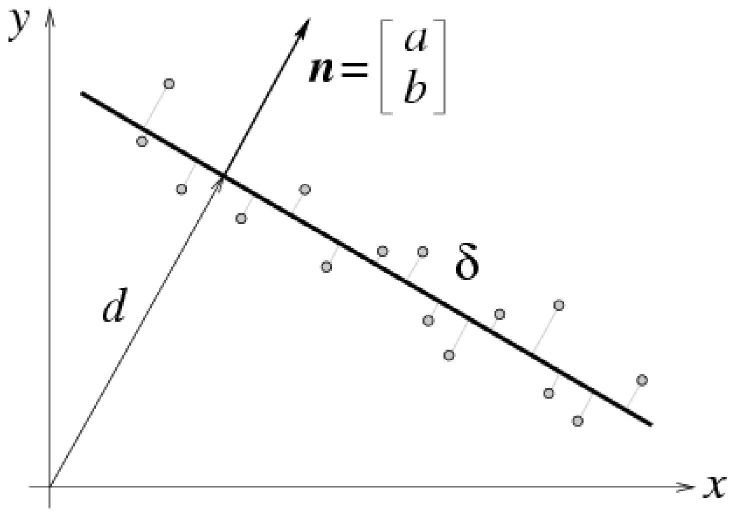
$$\boldsymbol{x} = \mu_1\boldsymbol{e}_1 + \ldots + \mu_q\boldsymbol{e}_q$$

for some μ_i $(i = 1, \ldots, q)$ such that $\mu_1^2 + \ldots + \mu_q^2 = 1$.

$$\begin{aligned} E(\boldsymbol{x})-E(\boldsymbol{e}_1) &= \boldsymbol{x}^T(U^TU)\boldsymbol{x}-\boldsymbol{e}_1^T(U^TU)\boldsymbol{e}_1 \\ &= \lambda_1\mu_1^2+ \ldots +\lambda_q\mu_q^2-\lambda_1 \\ &\geq \lambda_1(\mu_1^2+ \ldots +\mu_q^2-1)=0 \end{aligned}$$

$\boxed{\text{The solution is } \boldsymbol{e}_1 \text{.}}$

Example: Line Fitting



Problem: minimize

$$E(a,b,d) = \sum_{i=1}^{n} (ax_i + by_i - d)^2$$

with respect to (a,b,d).

- Minimize E with respect to d:

$$\frac{\partial E}{\partial d} = 0 \Longrightarrow d = \sum_{i=1}^{n} \frac{ax_i + by_i}{n} = a\bar{x} + b\bar{y}$$

- Minimize E with respect to a,b:

$$E = \sum_{i=1}^{n} [a(x_i - \bar{x}) + b(y_i - \bar{y})]^2 = |\mathcal{U}\boldsymbol{n}|^2 \quad \text{where} \quad \mathcal{U} = \begin{pmatrix} x_1 - \bar{x} & y_1 - \bar{y} \\ \dots & \dots \\ x_n - \bar{x} & y_n - \bar{y} \end{pmatrix}$$

- Done !!

Note:

$$\mathcal{U}^T\mathcal{U} = \begin{pmatrix} \sum_{i=1}^{n} x_i^2 - n\bar{x}^2 & \sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y} \\ \sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y} & \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 \end{pmatrix}$$

- Matrix of second moments of inertia

- Axis of least inertia

Linear Camera Calibration

Given n points $P_1, \ldots, P_n$ with *known* positions and their images $p_1, \ldots, p_n$

$$\begin{pmatrix} u_i \\ v_i \end{pmatrix} = \begin{pmatrix} \dfrac{\boldsymbol{m}_1 \cdot \boldsymbol{P}_i}{\boldsymbol{m}_3 \cdot \boldsymbol{P}_i} \\ \dfrac{\boldsymbol{m}_2 \cdot \boldsymbol{P}_i}{\boldsymbol{m}_3 \cdot \boldsymbol{P}_i} \end{pmatrix} \Longleftrightarrow \begin{pmatrix} \boldsymbol{m}_1 - u_i \boldsymbol{m}_3 \\ \boldsymbol{m}_2 - v_i \boldsymbol{m}_3 \end{pmatrix} \boldsymbol{P}_i = 0$$

$$\mathcal{P}\boldsymbol{m} = 0 \text{ with } \mathcal{P} \stackrel{\text{def}}{=} \begin{pmatrix} \boldsymbol{P}_1^T & \boldsymbol{0}^T & -u_1 \boldsymbol{P}_1^T \\ \boldsymbol{0}^T & \boldsymbol{P}_1^T & -v_1 \boldsymbol{P}_1^T \\ \ldots & \ldots & \ldots \\ \boldsymbol{P}_n^T & \boldsymbol{0}^T & -u_n \boldsymbol{P}_n^T \\ \boldsymbol{0}^T & \boldsymbol{P}_n^T & -v_n \boldsymbol{P}_n^T \end{pmatrix} \text{ and } \boldsymbol{m} \stackrel{\text{def}}{=} \begin{pmatrix} \boldsymbol{m}_1 \\ \boldsymbol{m}_2 \\ \boldsymbol{m}_3 \end{pmatrix} = 0$$

Once *M* is known, you still got to recover the intrinsic and extrinsic parameters !!!

This is a decomposition problem, not an estimation problem.

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

- Intrinsic parameters
- Extrinsic parameters

Degenerate Point Configurations

Are there other solutions besides *M* ??

$$\mathbf{0} = \mathcal{P}\boldsymbol{l} = \begin{pmatrix} \boldsymbol{P}_1^T & \mathbf{0}^T & -u_1\boldsymbol{P}_1^T \\ \mathbf{0}^T & \boldsymbol{P}_1^T & -v_1\boldsymbol{P}_1^T \\ \dots & \dots & \dots \\ \boldsymbol{P}_n^T & \mathbf{0}^T & -u_n\boldsymbol{P}_n^T \\ \mathbf{0}^T & \boldsymbol{P}_n^T & -v_n\boldsymbol{P}_n^T \end{pmatrix} \begin{pmatrix} \boldsymbol{\lambda} \\ \boldsymbol{\mu} \\ \boldsymbol{\nu} \end{pmatrix} = \begin{pmatrix} \boldsymbol{P}_1^T\boldsymbol{\lambda} - u_1\boldsymbol{P}_1^T\boldsymbol{\nu} \\ \boldsymbol{P}_1^T\boldsymbol{\mu} - v_1\boldsymbol{P}_1^T\boldsymbol{\nu} \\ \dots \\ \boldsymbol{P}_n^T\boldsymbol{\lambda} - u_n\boldsymbol{P}_n^T\boldsymbol{\nu} \\ \boldsymbol{P}_n^T\boldsymbol{\mu} - v_n\boldsymbol{P}_n^T\boldsymbol{\nu} \end{pmatrix}$$

$$\Downarrow$$

$$\begin{cases} \boldsymbol{P}_i^T\boldsymbol{\lambda} - \dfrac{\boldsymbol{m}_1^T\boldsymbol{P}_i}{\boldsymbol{m}_3^T\boldsymbol{P}_i}\boldsymbol{P}_i^T\boldsymbol{\nu} = 0 \\ \boldsymbol{P}_i^T\boldsymbol{\mu} - \dfrac{\boldsymbol{m}_2^T\boldsymbol{P}_i}{\boldsymbol{m}_3^T\boldsymbol{P}_i}\boldsymbol{P}_i^T\boldsymbol{\nu} = 0 \end{cases} \Rightarrow \begin{cases} \boldsymbol{P}_i^T(\boldsymbol{\lambda}\boldsymbol{m}_3^T - \boldsymbol{m}_1\boldsymbol{\nu}^T)\boldsymbol{P}_i = 0 \\ \boldsymbol{P}_i^T(\boldsymbol{\mu}\boldsymbol{m}_3^T - \boldsymbol{m}_2\boldsymbol{\nu}^T)\boldsymbol{P}_i = 0 \end{cases}$$

- Coplanar points: $(\lambda,\mu,\nu) = (\Pi,0,0)$ or $(0,\Pi,0)$ or $(0,0,\Pi)$
- Points lying on the intersection curve of two quadric surfaces = straight line + twisted cubic

Does not happen for 6 or more random points!

Analytical Photogrammetry

Given n points $P_1, \ldots, P_n$ with *known* positions and their images $p_1, \ldots, p_n$

Find $\boldsymbol{i}$ and $\boldsymbol{e}$ such that

$$\sum_{i=1}^{n} \left[\left(u_i - \frac{\boldsymbol{m}_1(\boldsymbol{i}, \boldsymbol{e}) \cdot \boldsymbol{P}_i}{\boldsymbol{m}_3(\boldsymbol{i}, \boldsymbol{e}) \cdot \boldsymbol{P}_i} \right)^2 + \left(v_i - \frac{\boldsymbol{m}_2(\boldsymbol{i}, \boldsymbol{e}) \cdot \boldsymbol{P}_i}{\boldsymbol{m}_3(\boldsymbol{i}, \boldsymbol{e}) \cdot \boldsymbol{P}_i} \right)^2 \right] \quad \text{is minimized}$$

Non-Linear Least-Squares Methods

- Newton
- Gauss-Newton
- Levenberg-Marquardt

Iterative, quadratically convergent in favorable situations

Mobile Robot Localization (Devy *et al.*, 1997)