Joint Institute of Engineering

SUN YAT-SEN UNIVERSITY

Carnegie Mellon University

Camera Geometry & Calibration

Forsyth&Ponce: Chap. 1,2,3
Szeliski: Chap. 2.1

Homogeneous Coordinates in 2-D

- Physical point $p = \begin{bmatrix} x \\ y \end{bmatrix}$ represented by three coordinates

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} \qquad \begin{aligned} x &= u/w \\ y &= v/w \end{aligned} \qquad \text{What if } w = 0???$$

- 2 sets of homogeneous coordinate vectors are equivalent if they are proportional to each other:

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} \equiv \begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} \Leftrightarrow \begin{bmatrix} u \\ v \\ w \end{bmatrix} \equiv \lambda \begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} \; \lambda \neq 0$$

Lines in 2-D

- General equation of a line in 2-D:

$$ax + by + c = 0$$

- In homogeneous coordinates:

$$l^{\mathsf{T}} p = l \cdot p = 0 \qquad l = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \qquad p = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Recall : Dot and Cross Products

- What is the formula?
- What is the geometric interpretation?
- See the Math Primer (to be sent out)

Using the Cross-Product

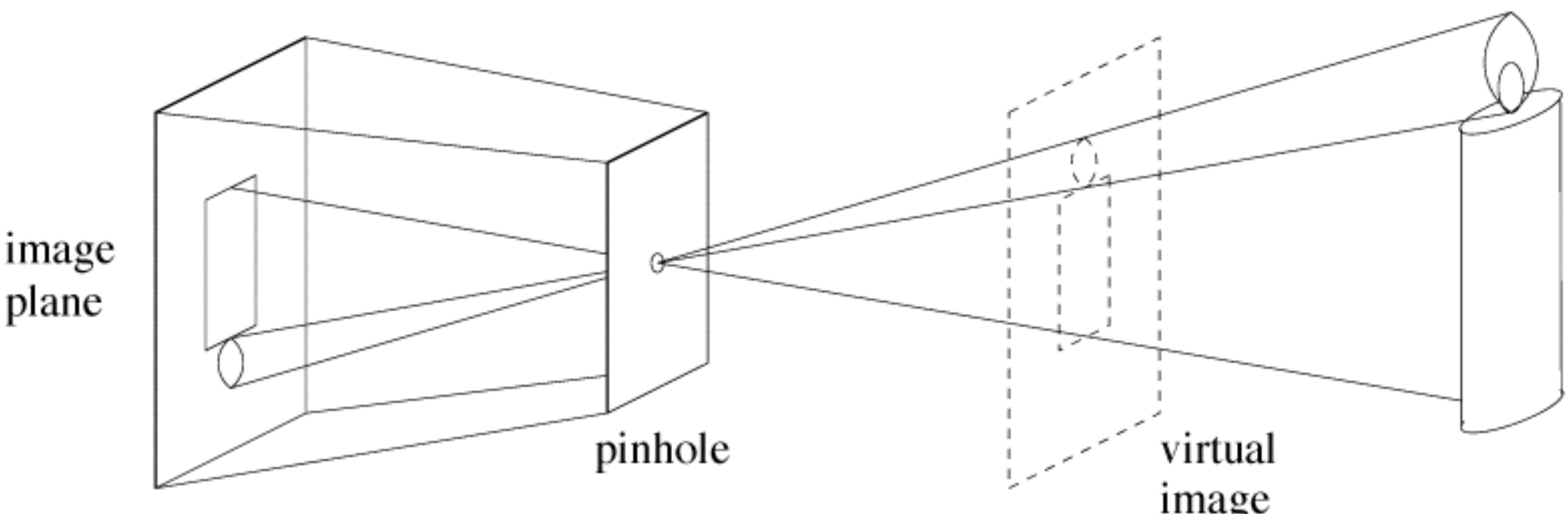
- The cross product of two 3-coordinate vectors p and q is : $p \times q$
- Properties: $p \cdot (p \times q) = q \cdot (p \times q) = p \times p = 0$
- Questions:
  - Write $p \equiv q$ using the cross-product
  - What is the intersection of 2 lines l_1 and l_2 in homogeneous coordinates?
  - What is the line l going through 2 points in homogeneous coordinates p_1 and p_2 ?

Pin-hole Model in Homogeneous Coordinates

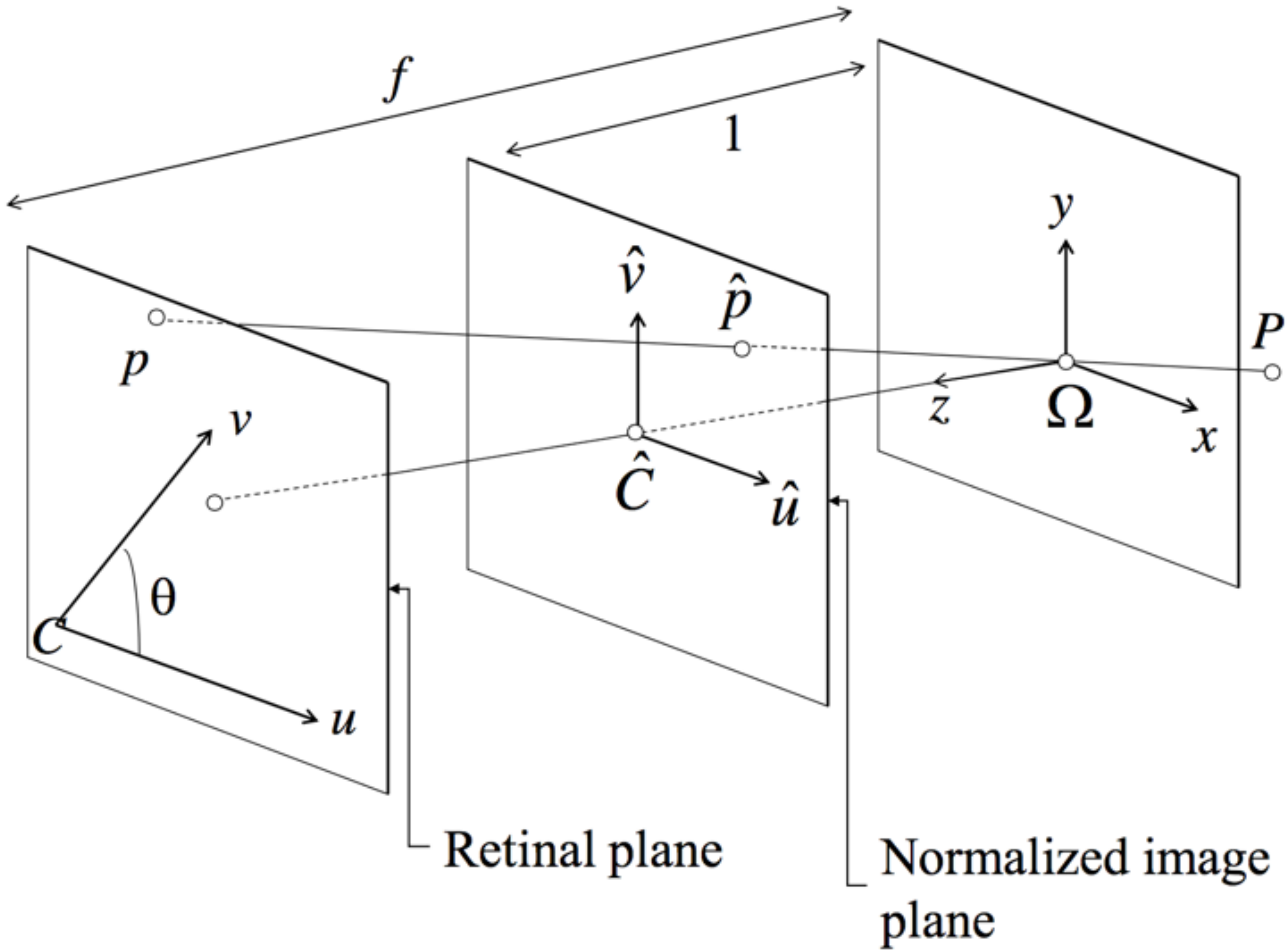
$$x = \frac{fX}{Z}$$

$$y = \frac{fY}{Z}$$

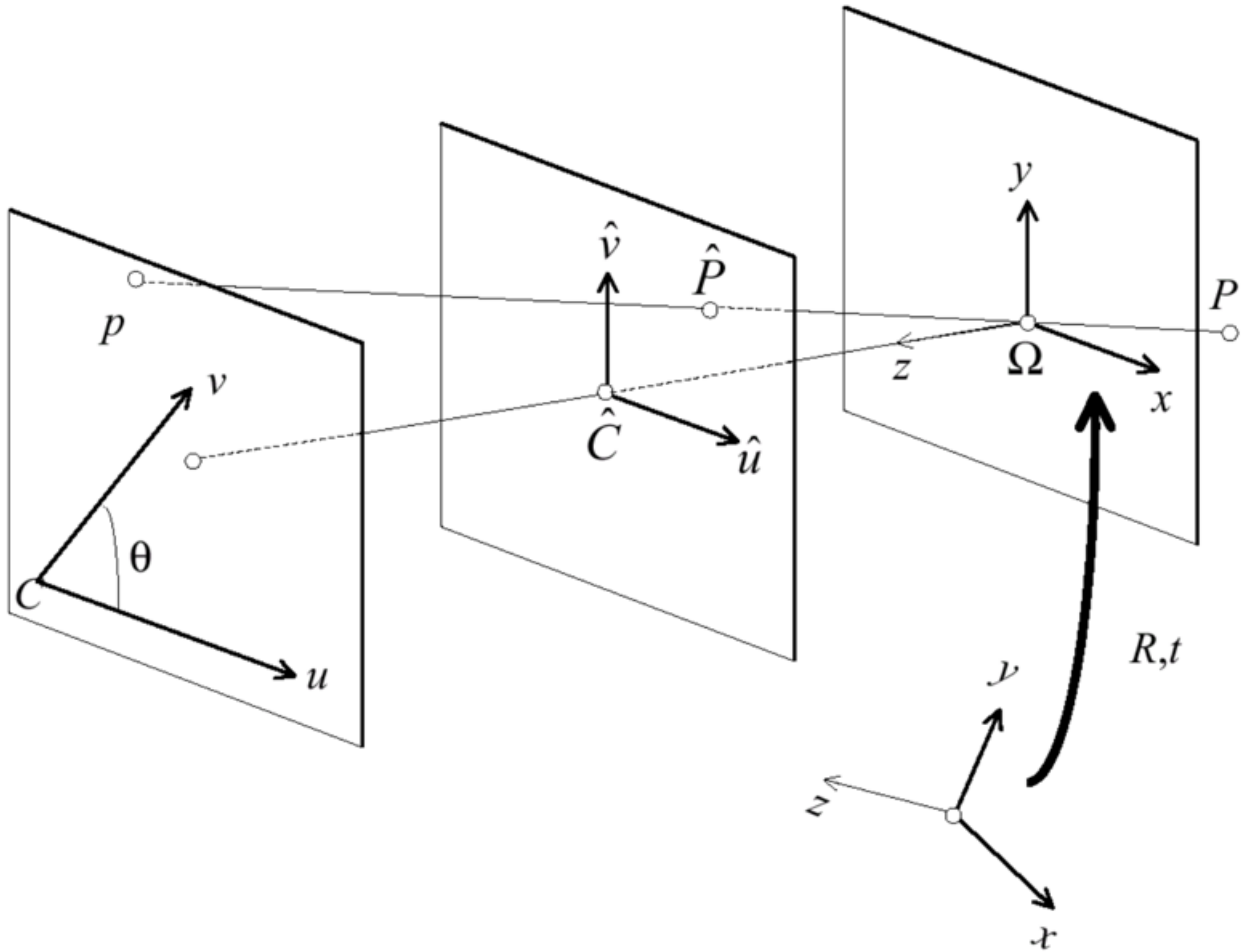
$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$



Normalized Image Projection

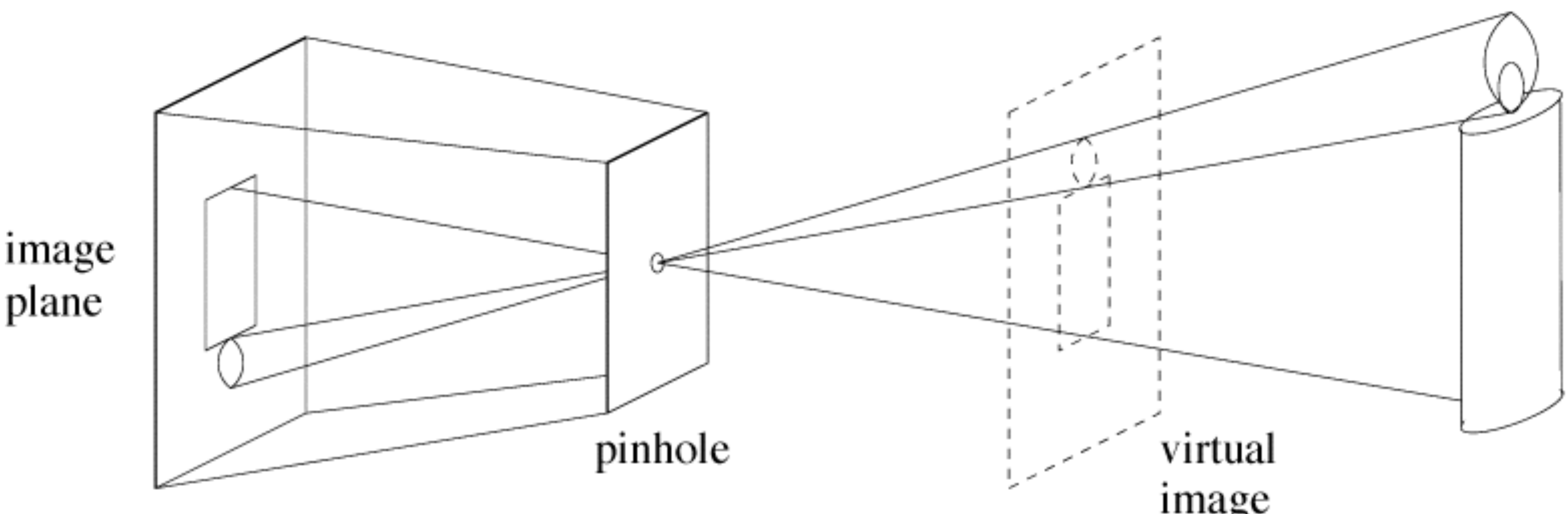


Normalized Image Projection



World Coordinates to Pixel Coordinates

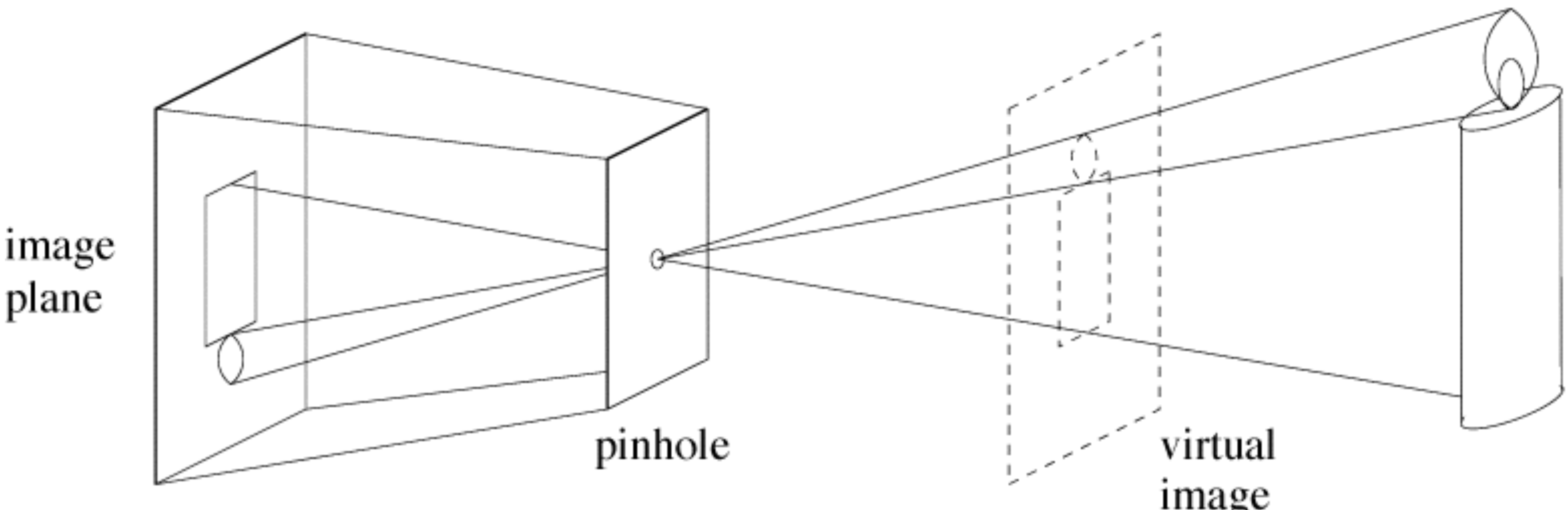
$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$



World Coordinates to Pixel Coordinates

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \alpha & 0 & 0 & 0 \\ 0 & \beta & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

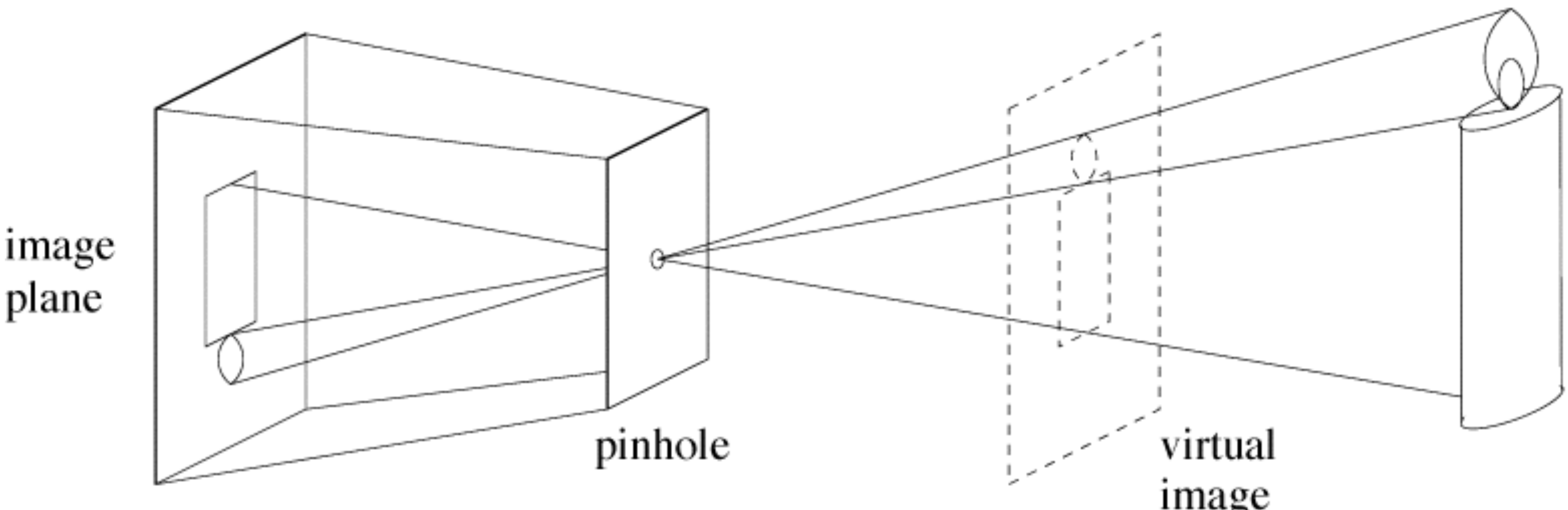
Meters to pixels in x and y



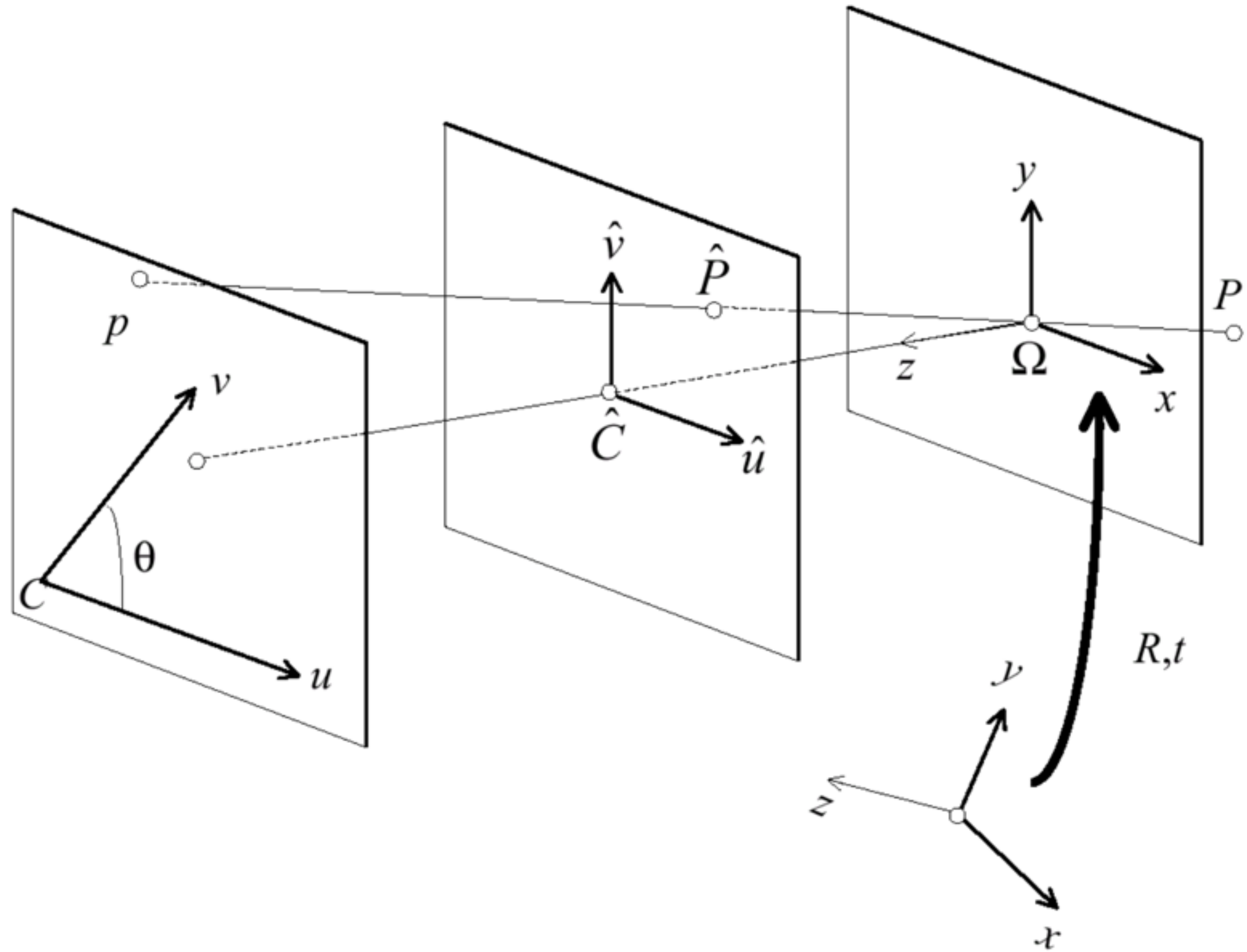
World Coordinates to Pixel Coordinates

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \alpha & 0 & x_0 & 0 \\ 0 & \beta & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

x and y offsets



Details : Skewed Axes, Rotation, Translation



Standard Perspective Camera Model

Scale in x direction between world coordinates and image coordinates

Skew of camera axes. $\theta = 90^{o}$ if the axes are perpendicular

Principal point = Image coordinates of the projection of camera origin on the retina

$$M = \begin{bmatrix} \alpha & -\alpha \cot\theta & u_o \\ 0 & \dfrac{\beta}{\sin\theta} & v_o \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R & t \end{bmatrix}$$

Translation between world coordinate system and camera

Rotation between world coordinate system and camera

Scale in y direction between world coordinates and image coordinates

Basic transformations (reminder) …

Transformation	Vector Coordinates	Homogeneous Coordinates	Degrees of Freedom	Invariants
Translation	$\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{t}$	$\begin{bmatrix} \boldsymbol{I} & \boldsymbol{t} \\ 0\ 0\ 0 & 1 \end{bmatrix}$	3	lengths, angles
Rotation	$\boldsymbol{y} = \boldsymbol{R}\boldsymbol{x}$ $\boldsymbol{R}^T\boldsymbol{R} = \boldsymbol{R}\boldsymbol{R}^T = \boldsymbol{I}$	$\begin{bmatrix} \boldsymbol{R} & \boldsymbol{0} \\ 0\ 0\ 0 & 1 \end{bmatrix}$	3	lengths, angles
Rigid	$\boldsymbol{y} = \boldsymbol{R}\boldsymbol{x} + \boldsymbol{t}$	$\begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \\ 0\ 0\ 0 & 1 \end{bmatrix}$	6	lengths, angles
Affine	$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{t}$	$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{t} \\ 0\ 0\ 0 & 1 \end{bmatrix}$	12	ratios of lengths, parallelism
Projective		4×4 matrix $\boldsymbol{M}$	15	colinearity, incidence

Alternative Notations

By rows:

$$M = \begin{bmatrix} m_1^T \\ m_2^T \\ m_3^T \end{bmatrix}$$

3

4

By components:

Intrinsic parameter matrix

$$M = K\begin{bmatrix} R & t \end{bmatrix}$$

3x3 3x3 3x1

Extrinsic parameter matrix

By blocks:

$$M = \begin{bmatrix} A & b \end{bmatrix}$$

3x3 3x1

$$A = \begin{bmatrix} a_1^T \\ a_2^T \\ a_3^T \end{bmatrix}$$

Q: Is a given 3x4 matrix M the projection matrix of some camera?
A: Yes, if and only if det(A) is not zero

Q: Is the decomposition unique?
A: There are multiple equivalent solutions

Applying the Projection Matrix

Homogeneous coordinates of point in image

$$p = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$$

Homogeneous coordinates of point in world

$$P = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Homogeneous vector transformation p proportional to MP

$$\boxed{p \equiv MP}$$

Computation of individual coordinates:

$$\boxed{\begin{aligned} u &= \frac{m_1^T P}{m_3^T P} = \frac{m_1 .P}{m_3 .P} \\ v &= \frac{m_2^T P}{m_3^T P} = \frac{m_2 .P}{m_3 .P} \end{aligned}}$$

Geometric Interpretation

Projection equation:

$$u = \frac{m_1^T P}{m_3^T P} = \frac{a_1^T \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + b_1}{a_3^T \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + b_3}$$

Observations:

- $$a_1^T \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + b_1 = 0$$

  is the equation of a plane of normal direction a_1

- From the projection equation, it is also the plane defined by: $u = 0$

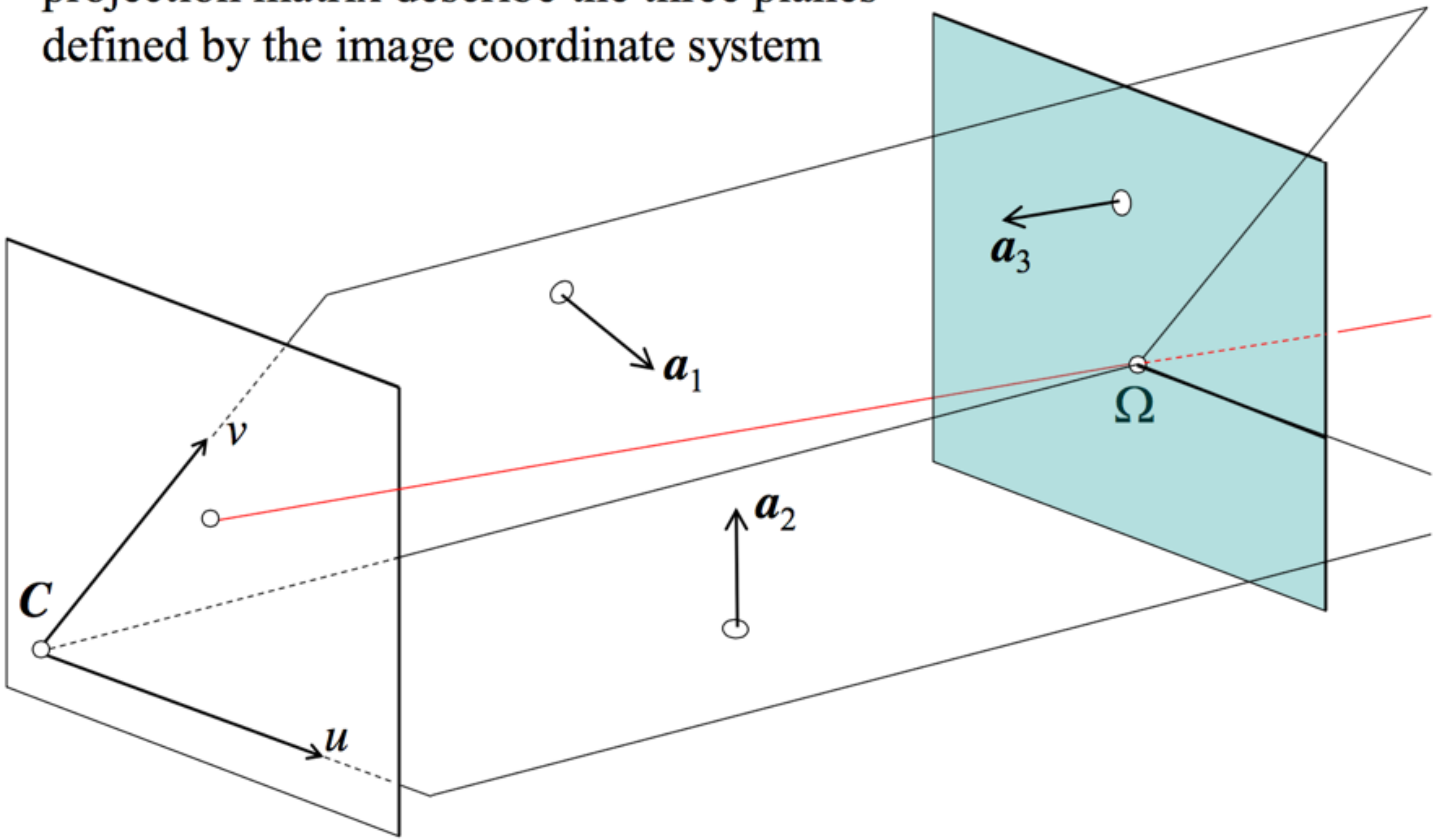
- Similarly:
  - (a_2,b_2) describes the plane defined by: $v = 0$
  - (a_3,b_3) describes the plane defined by:

    $$u = \infty \quad v = \infty$$

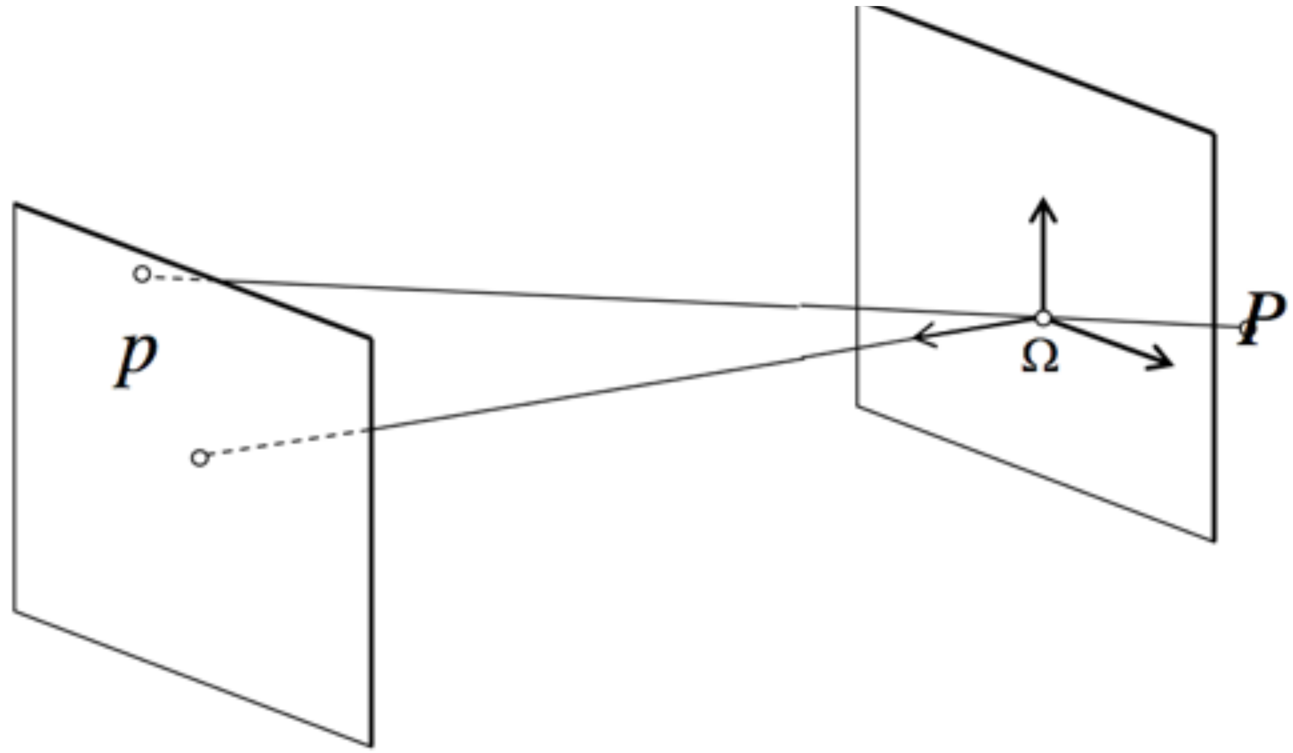
  → That is the plane passing through the pinhole ($z = 0$)

Geometric Interpretation

Geometric Interpretation: The rows of the projection matrix describe the three planes defined by the image coordinate system



Other useful geometric properties….



Q: Given an image point p, what is the direction of the corresponding ray in space?

A: $\boxed{w = A^{-1} p}$

Q: Can we compute the position of the camera center Ω?

A: $\boxed{\Omega = -A^{-1} b}$

Affine Cameras

Note: If the last row is $m_3^T = [0 \quad 0 \quad 0 \quad 1]$
the coordinates equations degenerate to:

$$\boxed{\begin{aligned} u &= m_1^T P = m_1 .P \\ v &= m_2^T P = m_2 .P \end{aligned}}$$

The mapping between world and image coordinates becomes *linear*.
This is an *affine* camera.

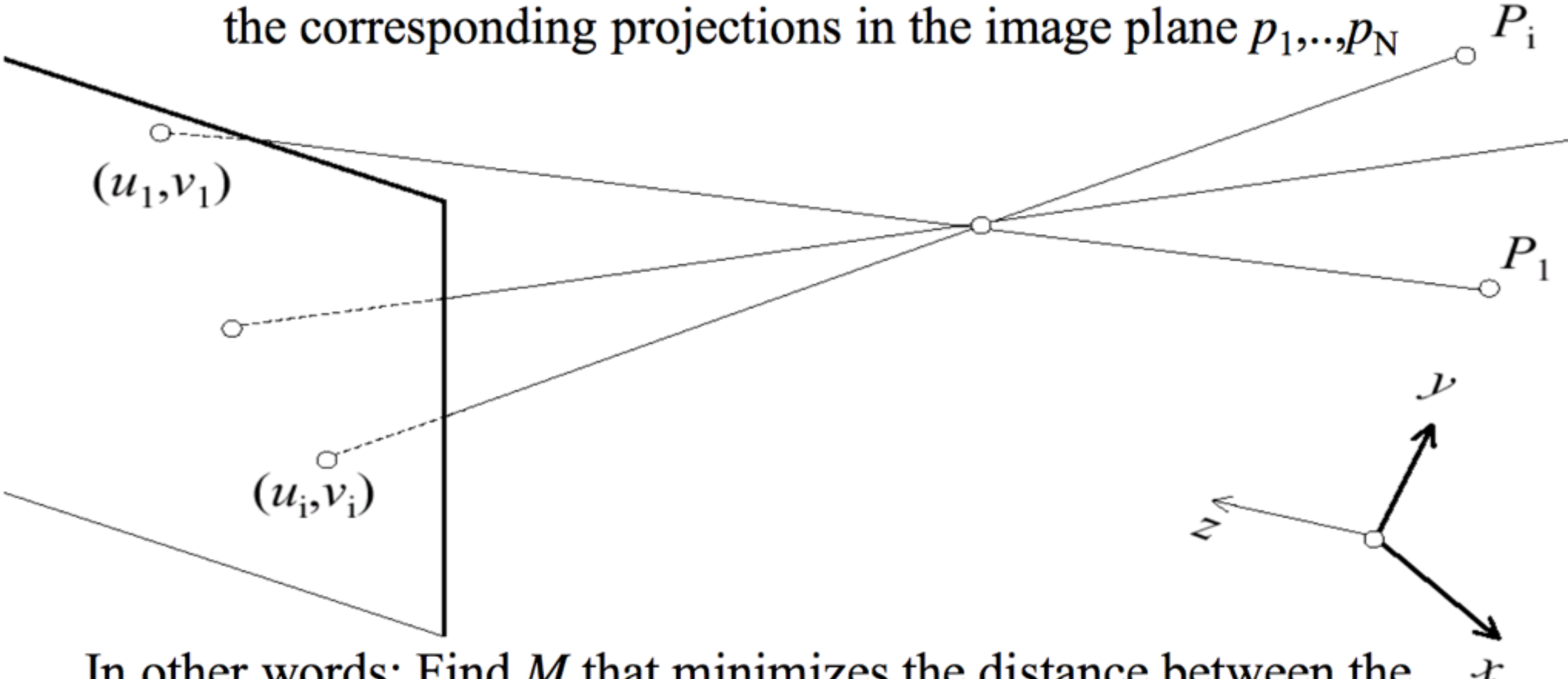
- Example: Weak-perspective projection model
- Projection defined by 8 parameters
- Parallel lines are projected to parallel lines
- The transformation can be written as a direct linear transformation

$$\begin{bmatrix} x \\ y \end{bmatrix} = \underbrace{M_{affine}}_{2x4\ \text{projection matrix}} P = \overbrace{K_2}^{2x2\ \text{intrinsic parameter matrix}} [\underbrace{R_2}_{\substack{2x3\ \text{matrix} = \text{first 2 rows of the rotation matrix} \\ \text{between world and camera frames}}} \quad \overbrace{t_2}^{\substack{\text{First 2 components of the translation} \\ \text{between world and camera frames}}}] P$$

Calibration Problem

- How to obtain M from a set of Points in Space and their corresponding pixel coordinates in an image.

Calibration: Recover M from scene points $P_1,..,P_N$ and the corresponding projections in the image plane $p_1,..,p_N$

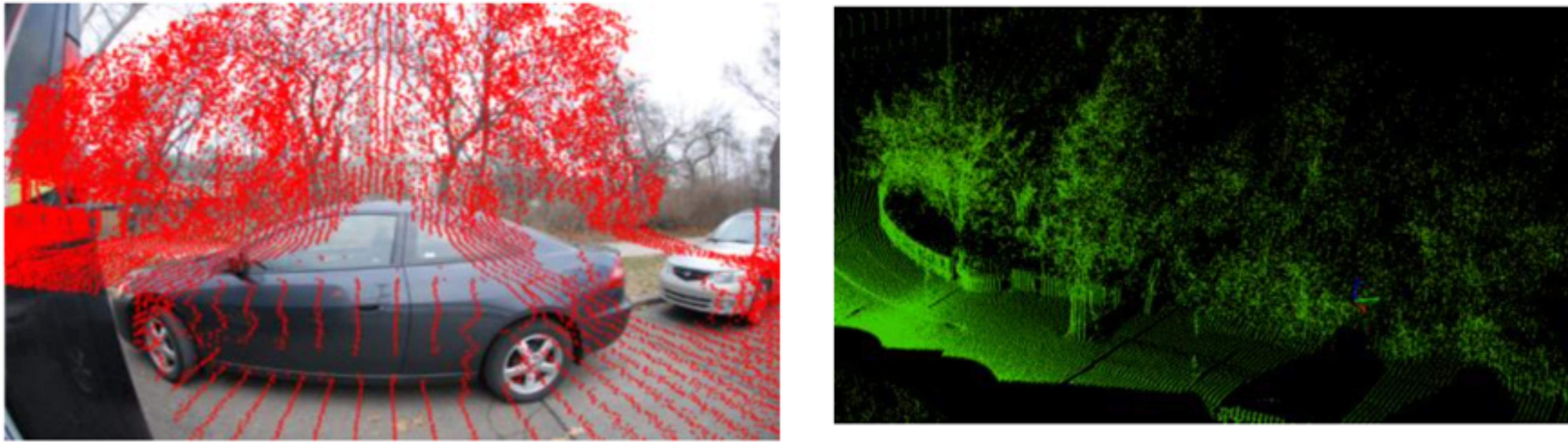


In other words: Find M that minimizes the distance between the actual points in the image, p_i, and their predicted projections MP_i

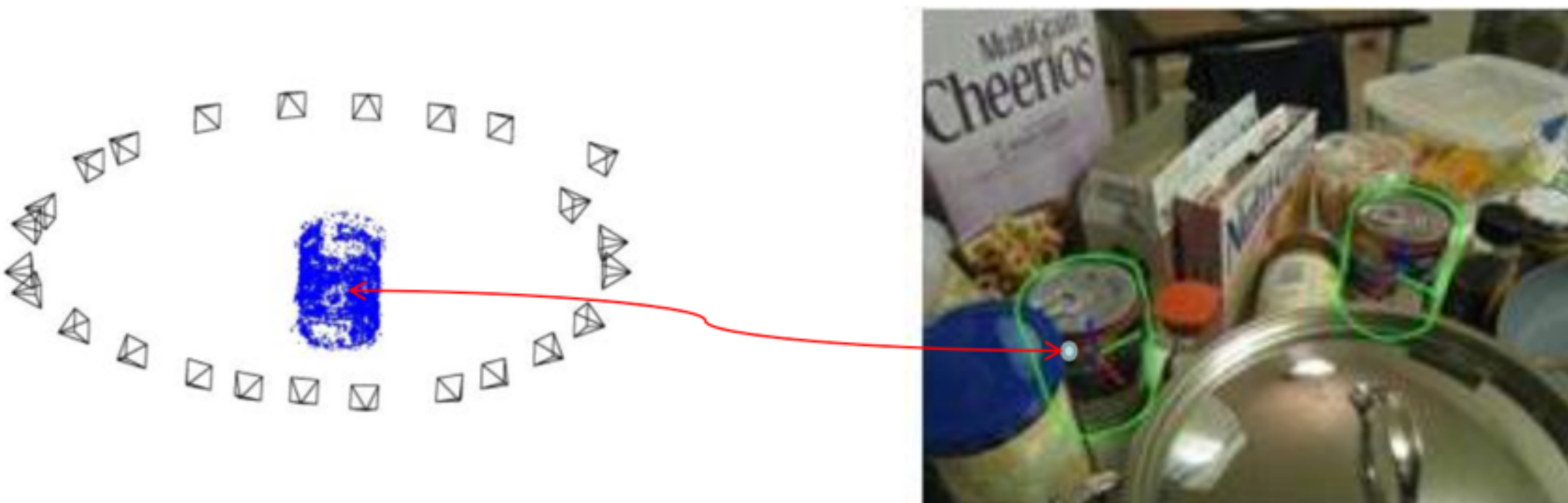
Problems:

- The projection is (in general) non-linear
- M is defined up to an arbitrary scale factor

Examples



3D + image data



Object recognition and localization

Camera Calibration Recipe

- Follows a template used in many (most) problems involving camera geometry in CV (so pay attention).

1. Write relation between image point, projection matrix, and point in space

2. Write non-linear relations between coordinates

3. Make them linear

4. Write in matrix form

5. Place into a single matrix

6. Minimize

Write relation between image point, projection matrix, and point in space:

$$\boxed{p_i \equiv M P_i}$$

Write non-linear relations between coordinates:

$$\boxed{u_i = \frac{m_1{}^T P_i}{m_3{}^T P_i} \qquad v_i = \frac{m_2{}^T P_i}{m_3{}^T P_i}}$$

Make them linear:

$$\boxed{\begin{aligned} m_1{}^T P_i - (m_3{}^T P_i)\, u_i &= 0 \\ m_2{}^T P_i - (m_3{}^T P_i)\, v_i &= 0 \end{aligned}}$$

Write them in matrix form:

$$\begin{bmatrix} P_i^T & 0 & -u_i P_i^T \\ 0 & P_i^T & -v_i P_i^T \end{bmatrix} m = 0 \qquad m = \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix}$$

Put all the relations for all the points into a single matrix:

$$\begin{bmatrix} P_1^T & 0 & -u_1 P_1^T \\ 0 & P_1^T & -v_1 P_1^T \\ \vdots & \vdots & \vdots \\ P_N{}^T & 0 & -u_N P_N{}^T \\ 0 & P_N{}^T & -v_N P_N{}^T \end{bmatrix} m = 0$$

Solve by minimizing:

$$|Lm|^2 = m^T L^T L m \qquad L = \begin{bmatrix} P_1^T & 0 & -u_1 P_1^T \\ 0 & P_1^T & -v_1 P_1^T \\ \vdots & \vdots & \vdots \\ P_N{}^T & 0 & -u_N P_N{}^T \\ 0 & P_N{}^T & -v_N P_N{}^T \end{bmatrix}$$

Subject to:

$$|m| = 1$$

Slight detour : Homogeneous Least Squares

Suppose that we want to estimate the best vector of parameters X from matrices V_i computed from input data such that $VX_i = 0$

We can do this by minimizing:

$$\sum_i |V_i X|^2 = \sum_i X^T V_i^T V_i \ X = X^T V X$$

Since V=0 is a trivial solution, we need to constrain the magnitude of V to be non-zero, for example: $|V| = 1$. The problem becomes:

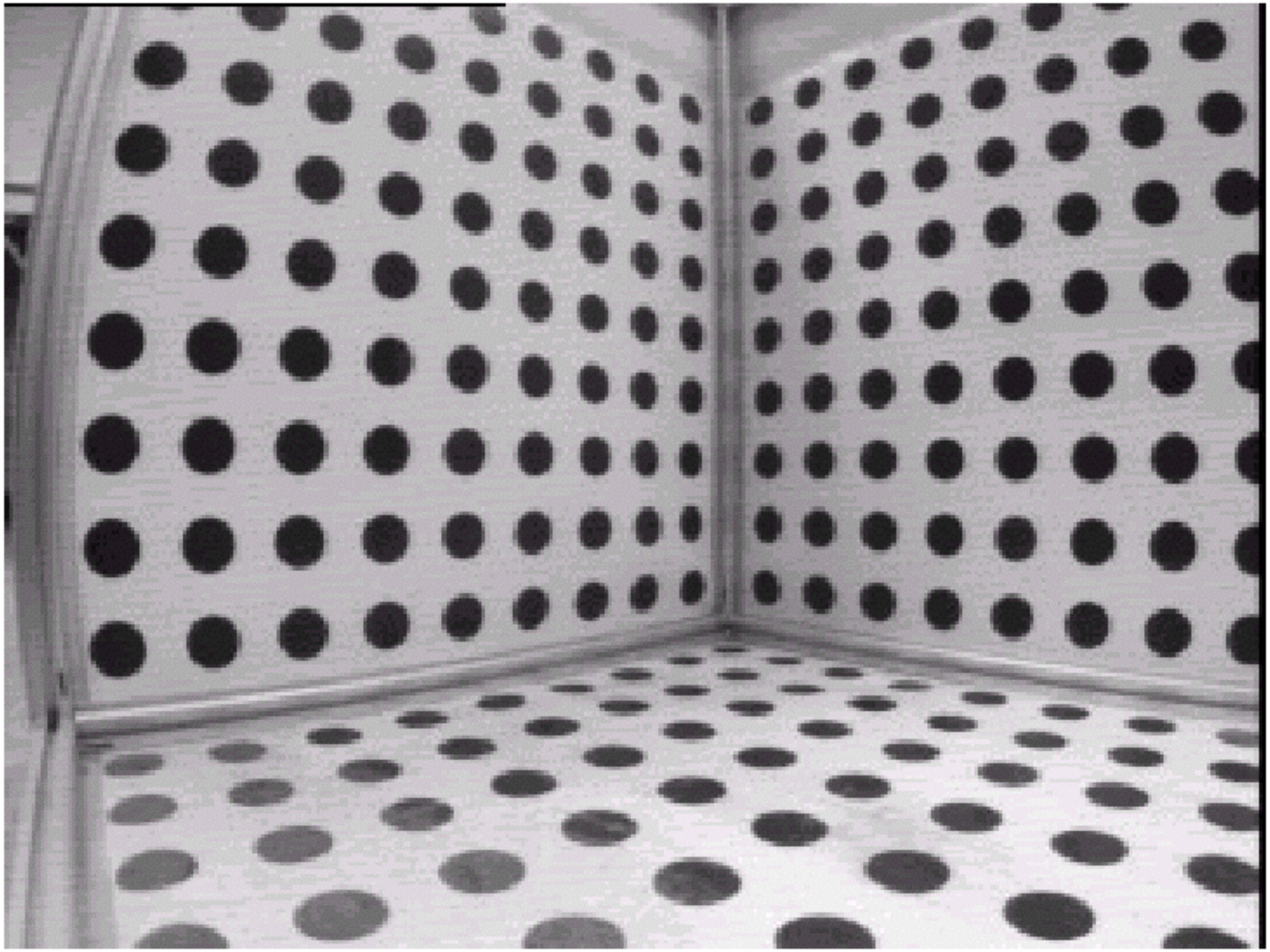
$$Min \ X^T V X$$

$$|X| = 1$$

The key result (which we will use 50 times in this class) is:

For any symmetric matrix V, the minimum of $X^T VX$ is reached at X = eigenvector of V corresponding to the smallest eigenvalue.

Calibration



Key Results

Projection matrix: $p = MP$

$$M = \begin{bmatrix} \alpha & -\alpha\cot\theta & u_o \\ 0 & \dfrac{\beta}{\sin\theta} & v_o \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R & t \end{bmatrix} = \begin{bmatrix} A & b \end{bmatrix}$$

5 intrinsic parameters

6 extrinsic parameters

Existence/unicity: $\det(A) \neq 0$

Zero skew: $(a_1 \times a_3).(a_2 \times a_3) = 0$

Aspect ratio 1: $|a_1 \times a_3| = |a_2 \times a_3|$

Key Results

Planes: $m_i.P = 0$

Optical center: $\Omega = -A^{-1}b$

Viewing ray: $w = A^{-1}b$

Calibration: minimum 6 non-coplanar points

Linear: $\underset{|m|=1}{\text{Min}}\ m^T L^T L m =$ Eigenvector of smallest eigenvalue of $L^T L$

Non-linear: $$\underset{m}{\text{Min}} \sum_i \left(u_i - \frac{m_1.P_i}{m_3.P_i} \right)^2 + \left(v_i - \frac{m_2.P_i}{m_3.P_i} \right)^2$$

More Calibration

- More general calibration approaches (see “Geometric Based Methods in Computer Vision”)
- MATLAB calibration toolbox (Bouguet)
- OpenCV calibration package

Examples