Camera Geometry and Calibration

Lu Yang

Email: yanglu@uestc.edu.cn

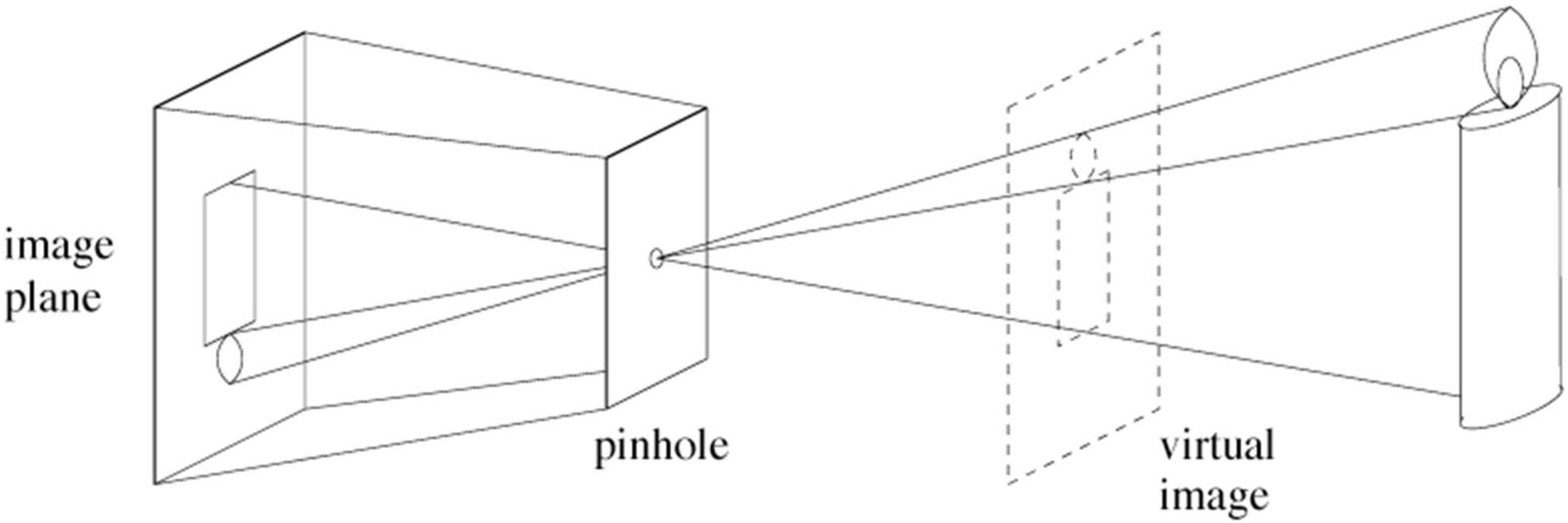
School of Automation Engineering, UESTC

Outline

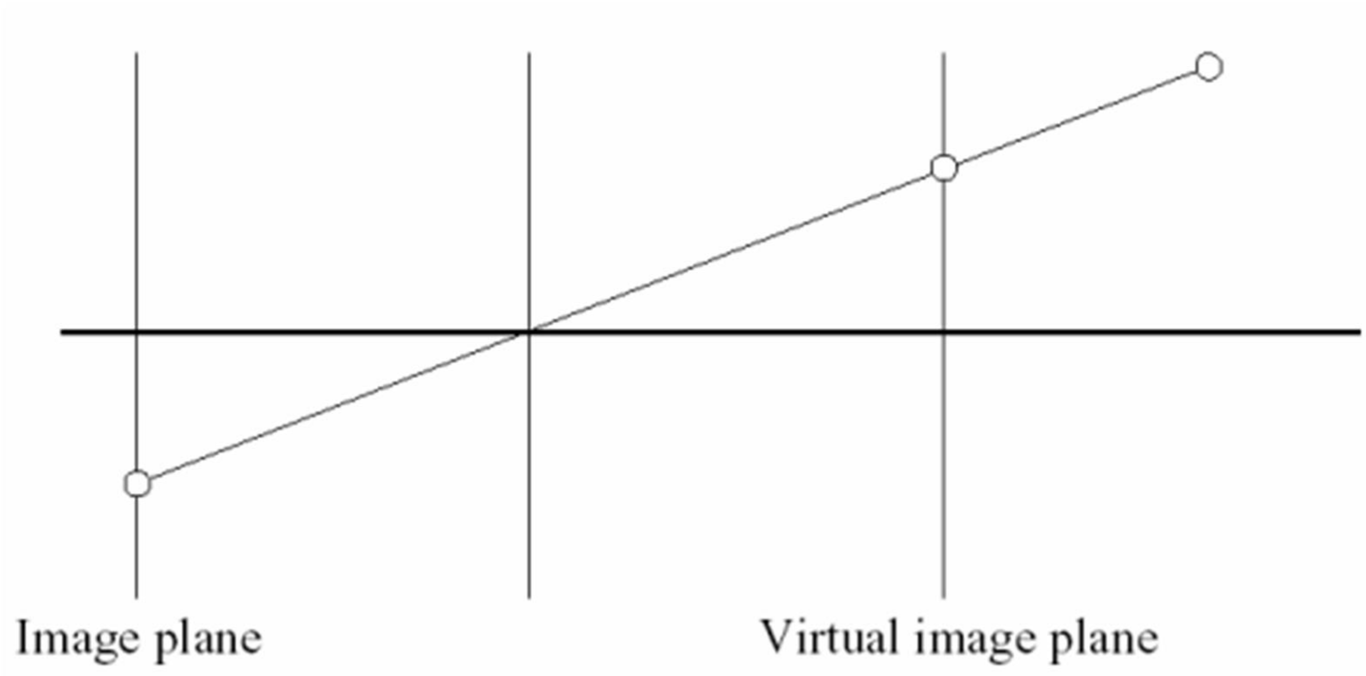
- Ideal pinhole model
- Ideal thin lens model
- General lens model
- Geometric models of cameras

Pinhole Cameras

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



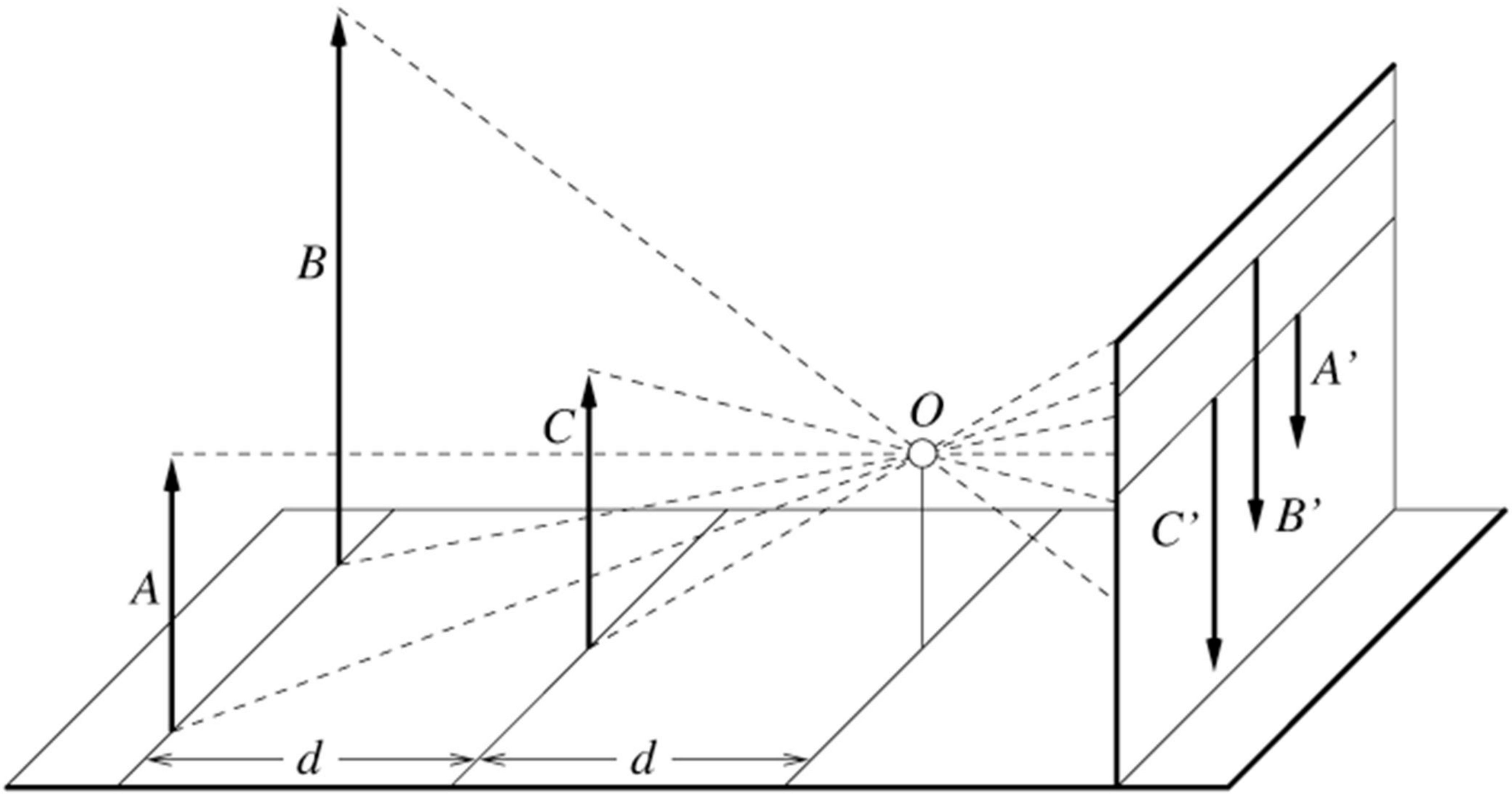
Equivalent Model with Virtual Image Plane



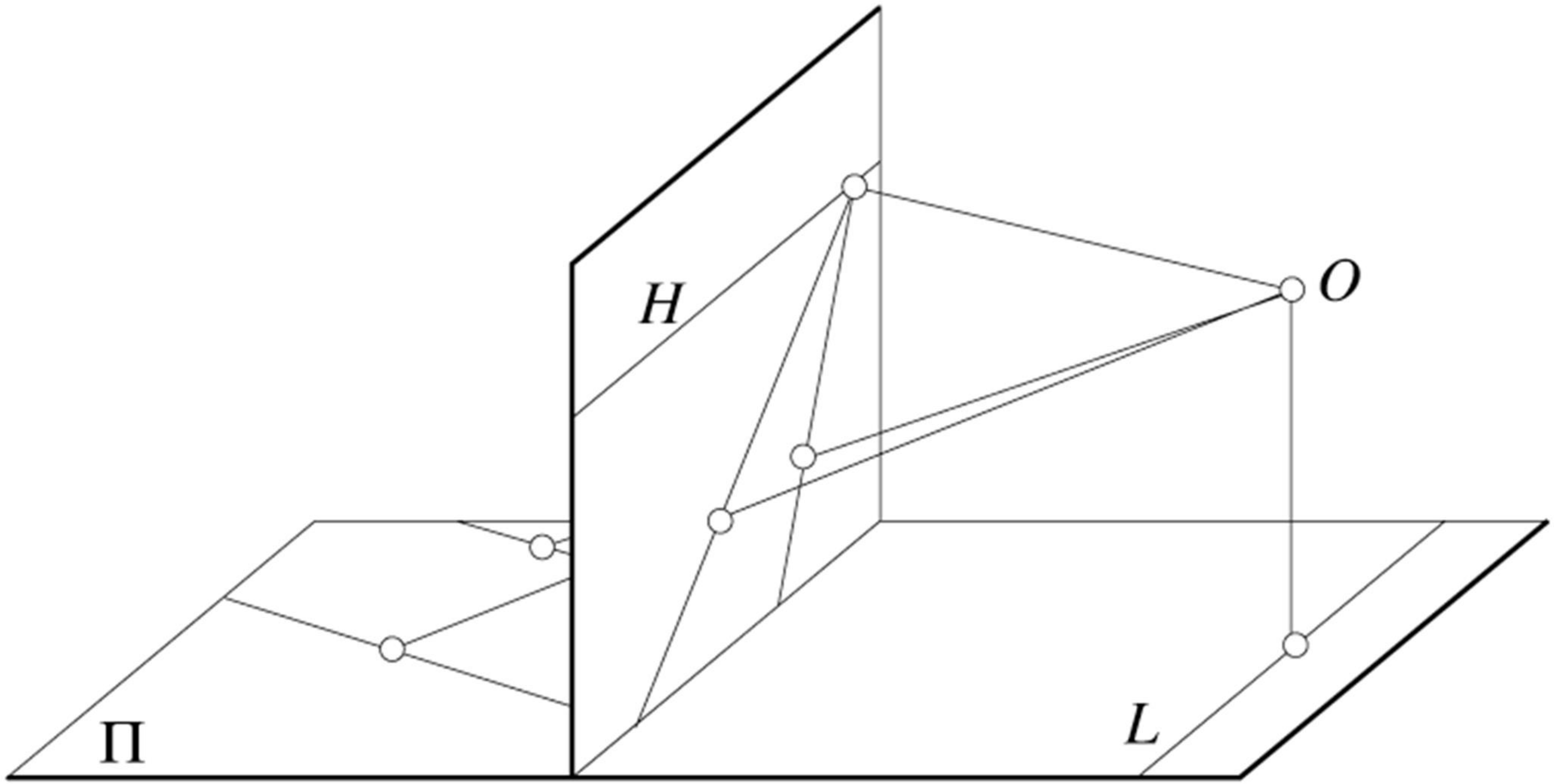
Basic Geometric Properties

- Distant objects are smaller
- Lines project to lines
- The projection of parallel lines meet at a single vanishing point
- Vanishing points of coplanar sets of lines are colinear, form the vanishing line of the plane (horizon)

Distant objects are smaller



Parallel Lines



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.

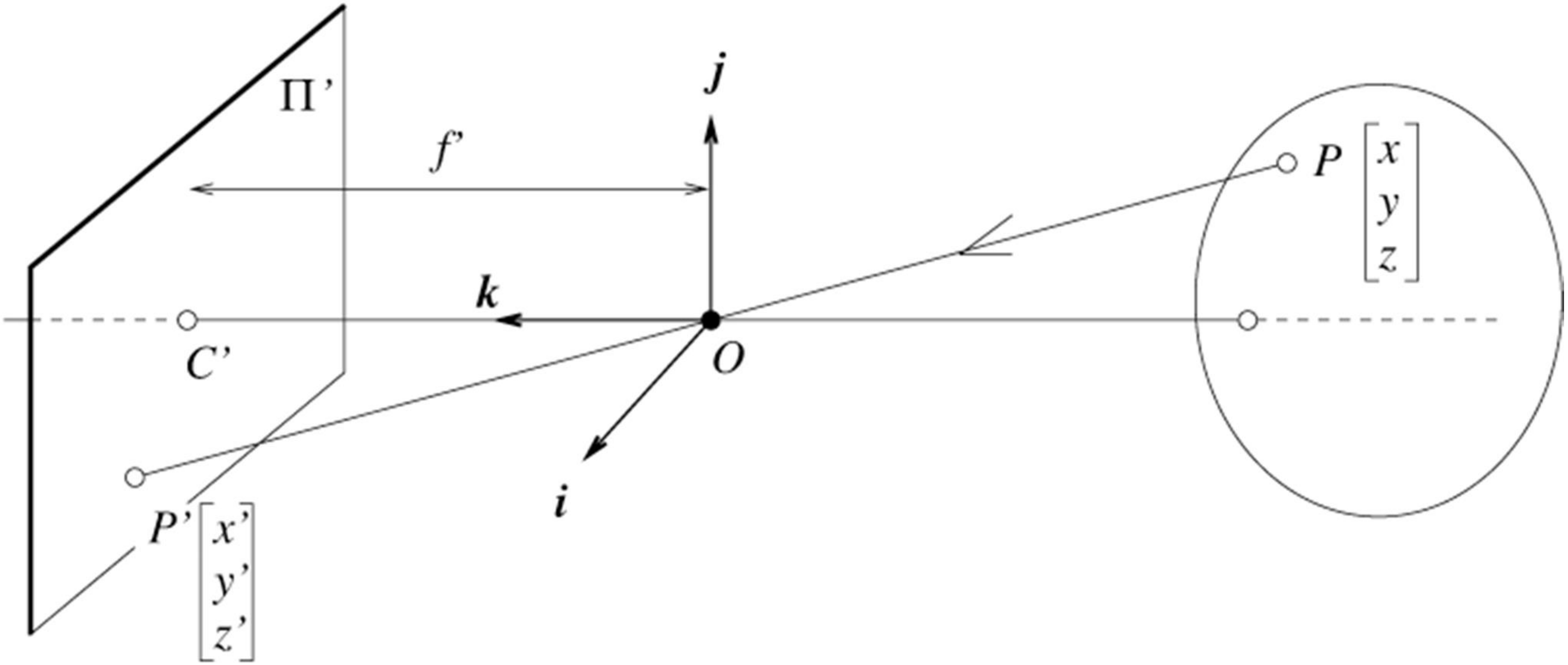
  The line is called the *horizon* for that plane

- Good way to spot faked images

  Scale and perspective don't work

  Vanishing points behave badly

Perspective Projection

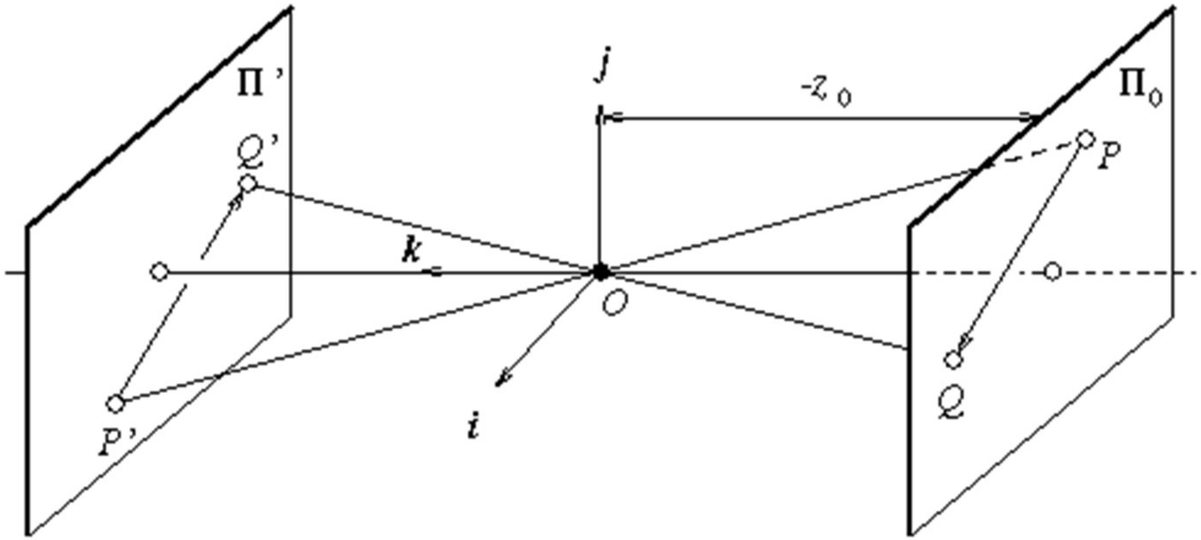


$$\begin{cases} x' = \lambda x \\ y' = \lambda y \\ f' = \lambda z \end{cases} \Leftrightarrow \lambda = \frac{f'}{z}$$

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

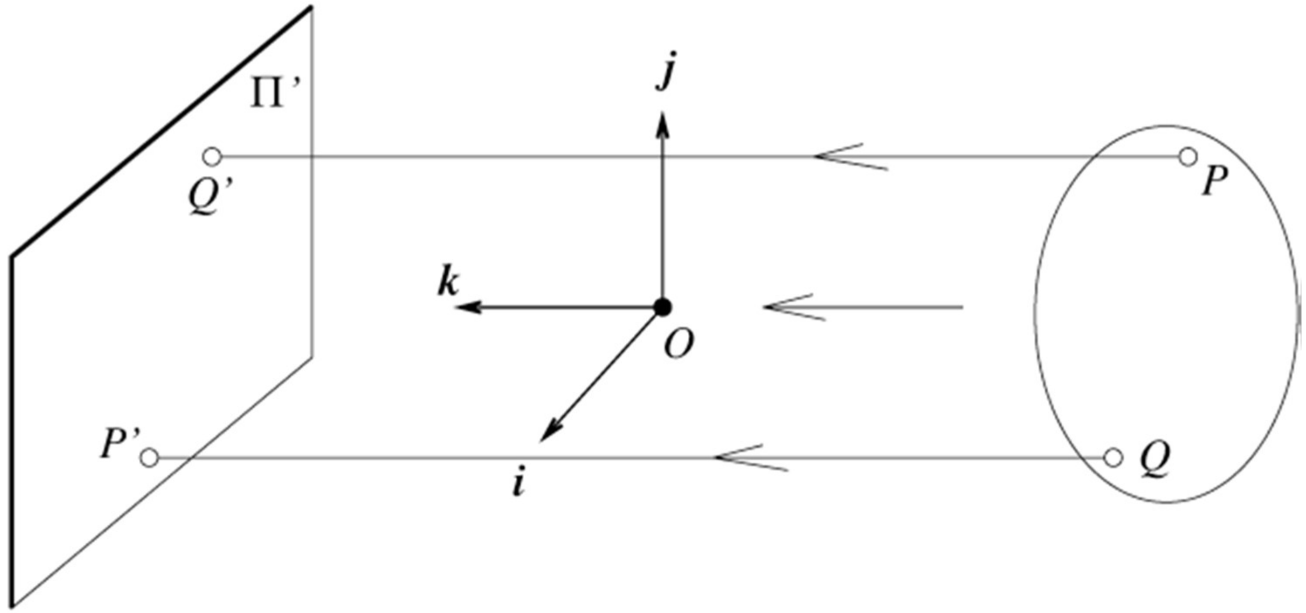
Weak Perspective Project (Affine Projection)

- Issue
  - Perspective effects, but not over the scale of individual objects
  - Collect points into a group at about the same depth, then divide each point by the depth of its group
  - If scene points are in a plane, projections are simply magnified by m
  - Justified if scene depth is small relative to average distance from camera



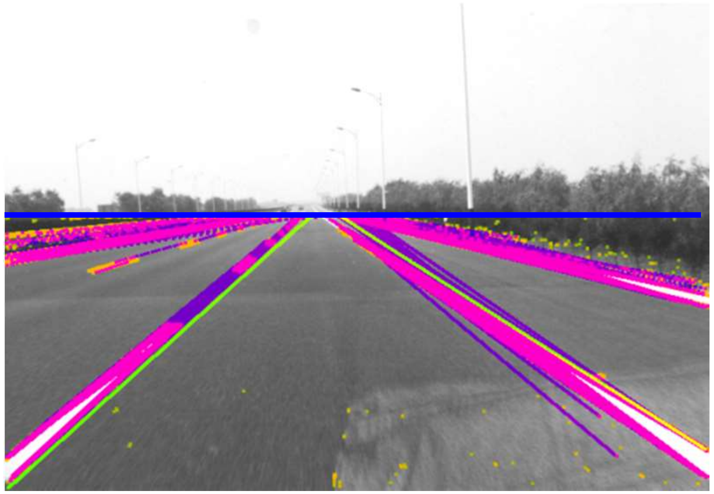
$$\begin{cases} x' \approx -mx \\ y' \approx -my \\ \mathrm{m} = -\dfrac{f'}{z_0} \end{cases}$$

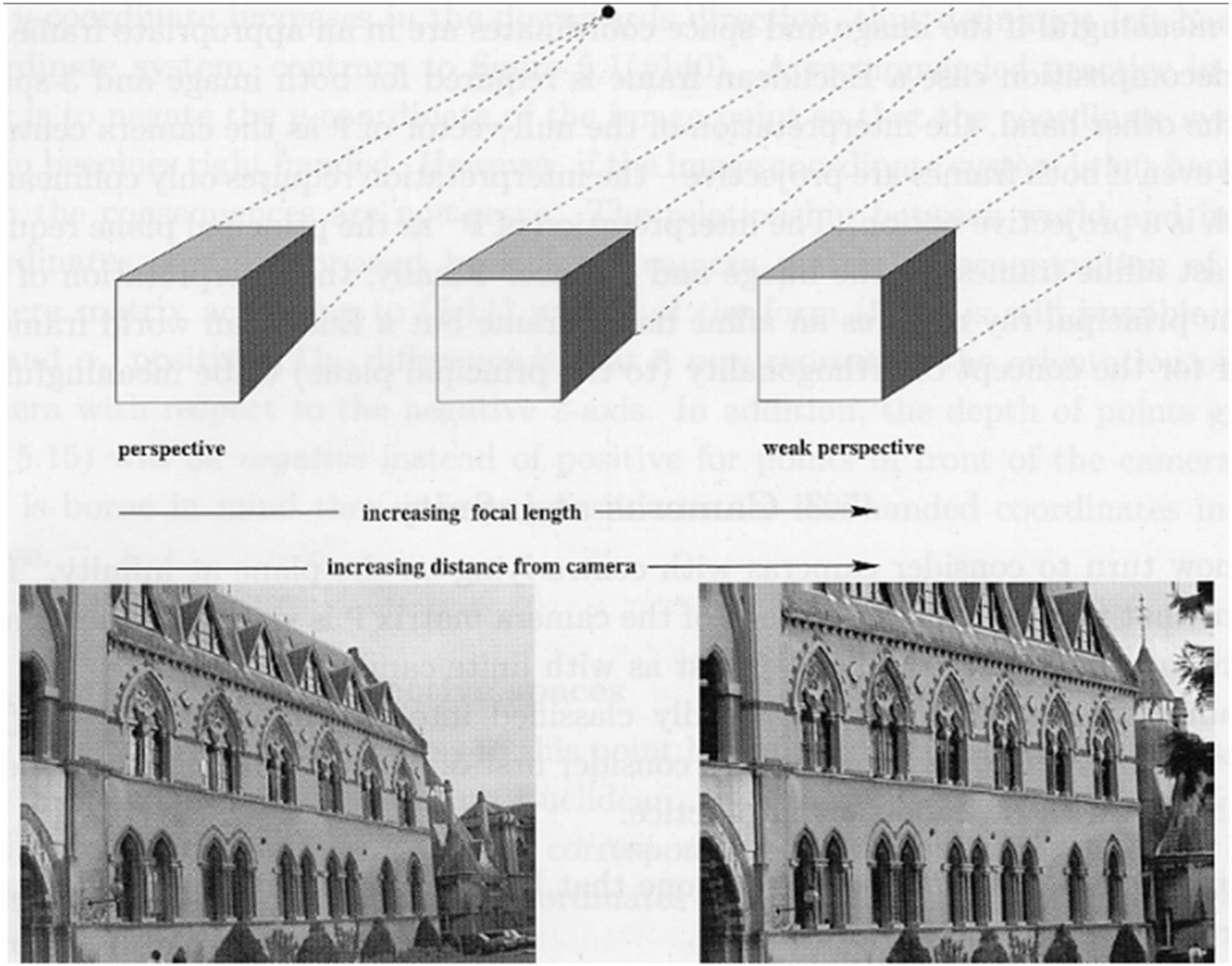
Orthographic projection



Justified if scene depth is small compared to distance from camera and camera remains at approximately constant distance.

Some Images

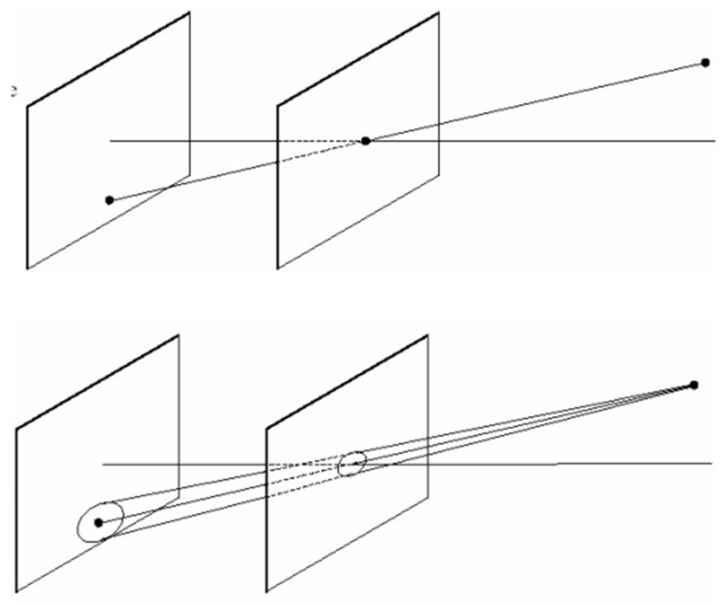


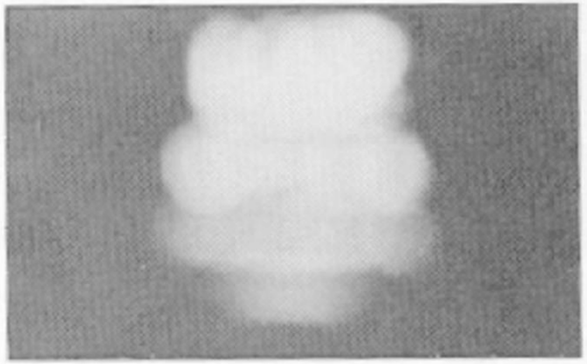


From Zisserman & Hartley

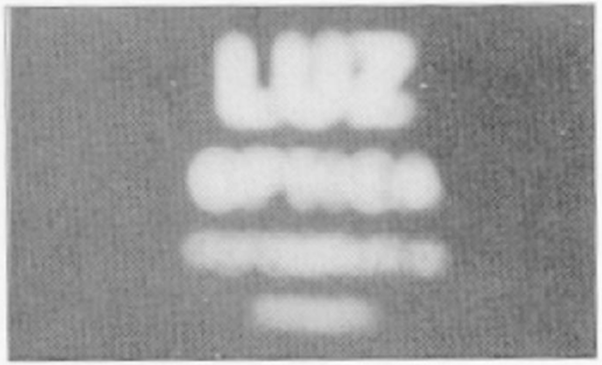
Some Limitations of Pinhole Model

- Ideal pinhole: single scene point generate single image, **diffraction, low light level;**
- Finite-size pinhole: single scene point generate extended image. **Resulting image is blurry**





2 mm



1 mm

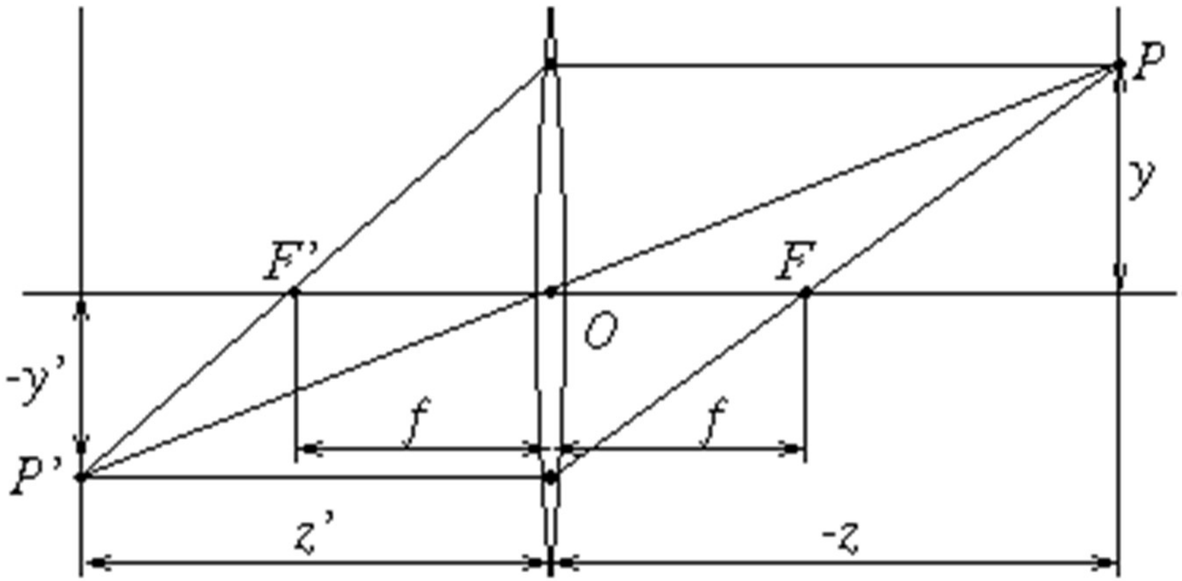
0.6mm

0.35 mm

0.15 mm

0.07 mm

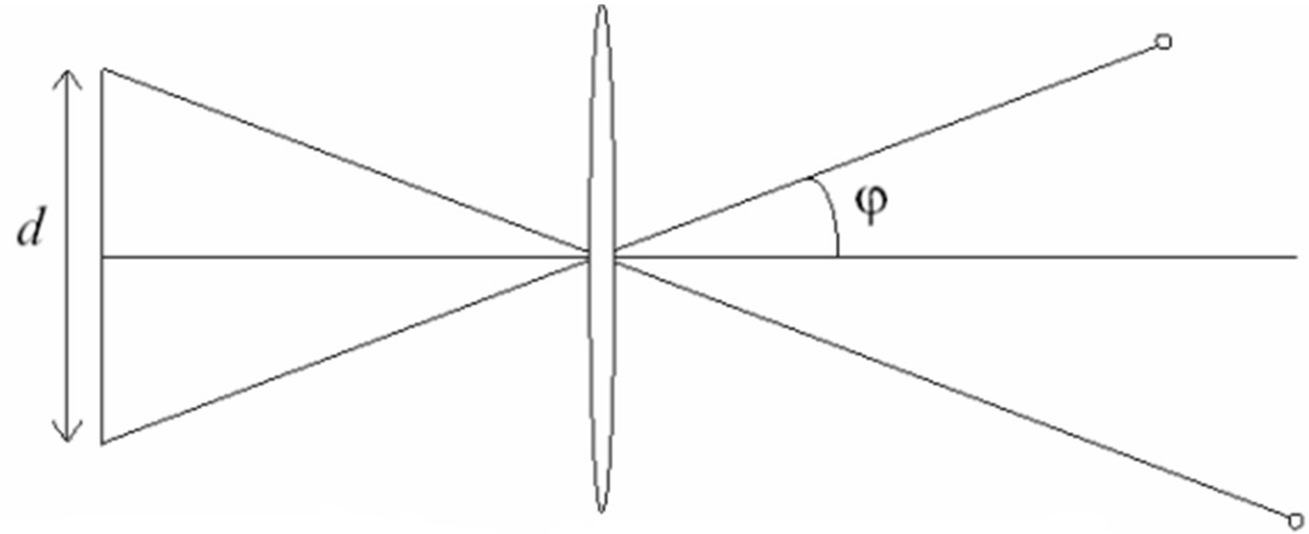
Thin Lens Model



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

- All rays emanating from P converge to a single point P';
- Points at infinity are focused on plane z'=f;
- Idea because: infinite aperture, infinite field of view, infinitely small distance between surfaces.

Finite Field of View

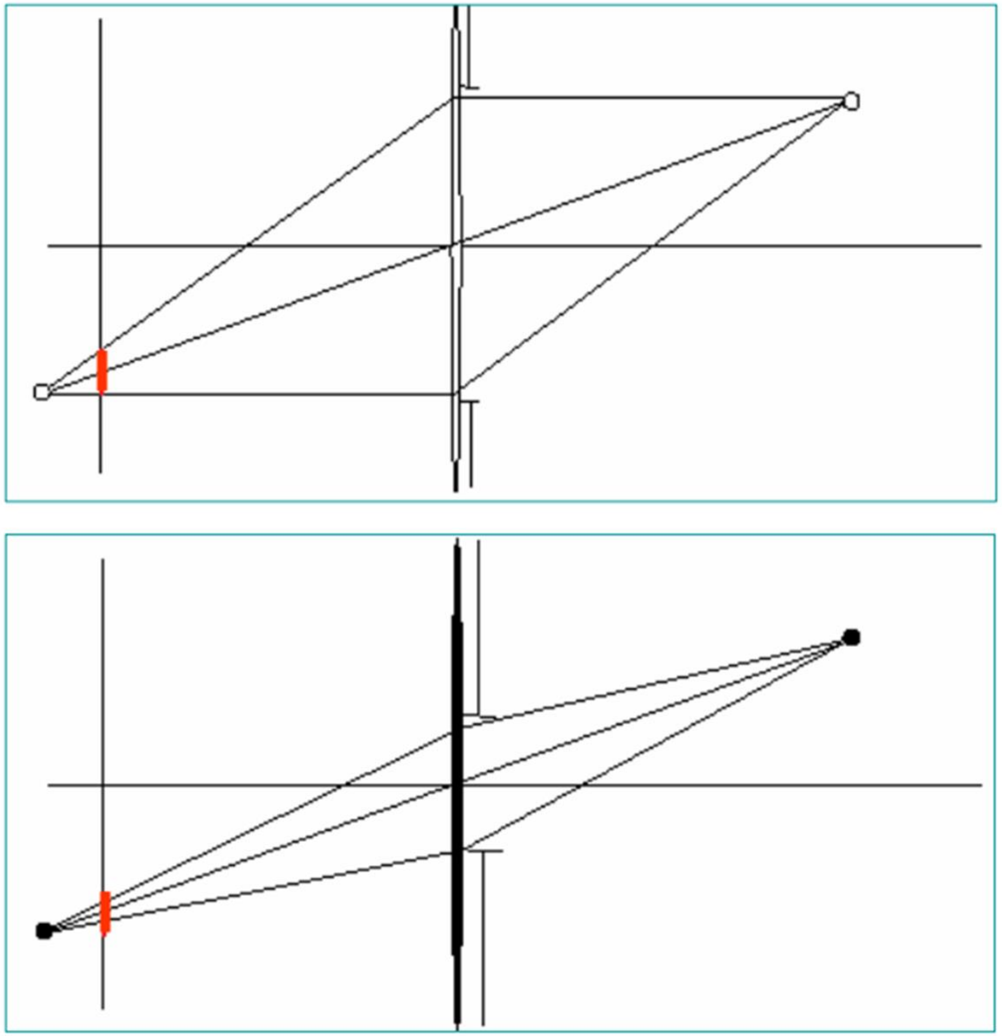


Size of field of view governed by size of the camera retina

$$\varphi = \tan^{-1}\left(\frac{d}{2f}\right)$$

How about human eye?

Finite Aperture/1

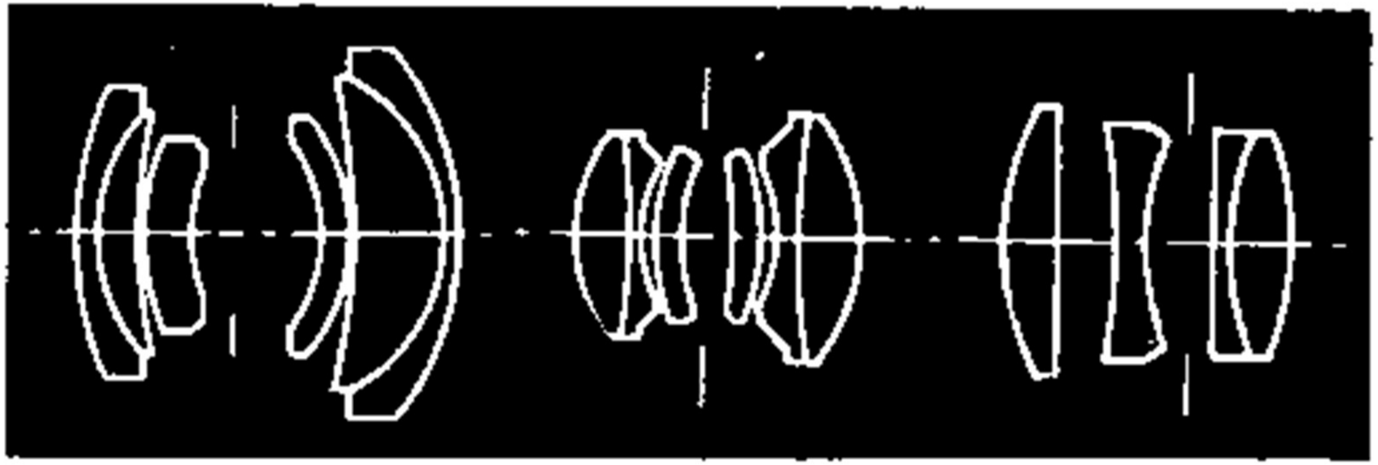


Ideal case: Only points on one plane are in perfect focus;

Finite aperture: points within a region of depth D (depth of field) are in focus.
For a given f, the larger the aperture, the smaller D.

Finite Aperture/2

Real Lenses



- Previous approximation is incorrect

→Aberrations and distortions;

→Blurring and incorrect shape in the image

Snell's Law of Refraction

$$n_1 \sin(\alpha_1) = n_2 \sin(\alpha_2)$$

First-order optics: appropriate for ideal model of thin lens

$$n_1 \alpha_1 \approx n_2 \alpha_2$$

Higher order optics: necessary for real lenses

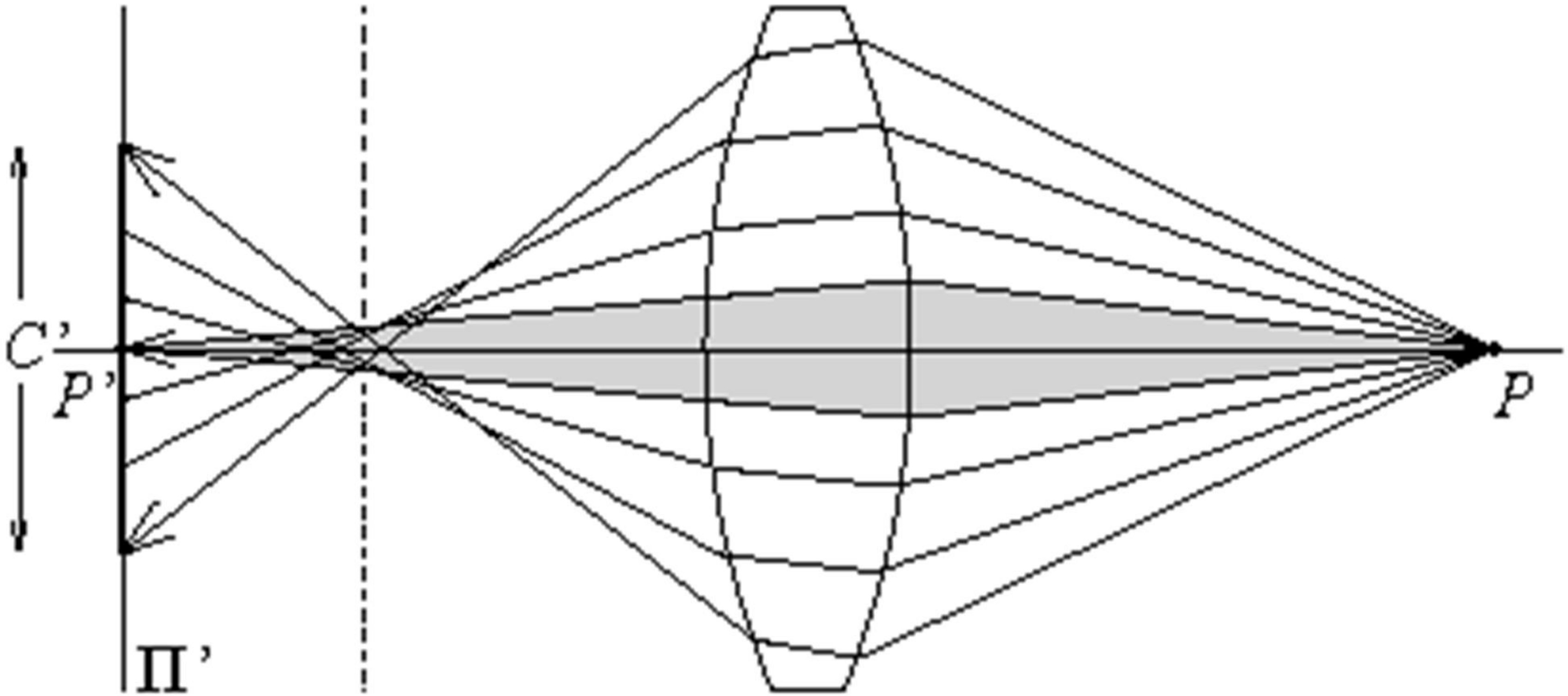
$$\sin \alpha \approx \alpha - \frac{\alpha^3}{6} + ...$$

Aberrations:

Blurring: e.g., spherical aberrations

Geometric distortion

Spherical aberration

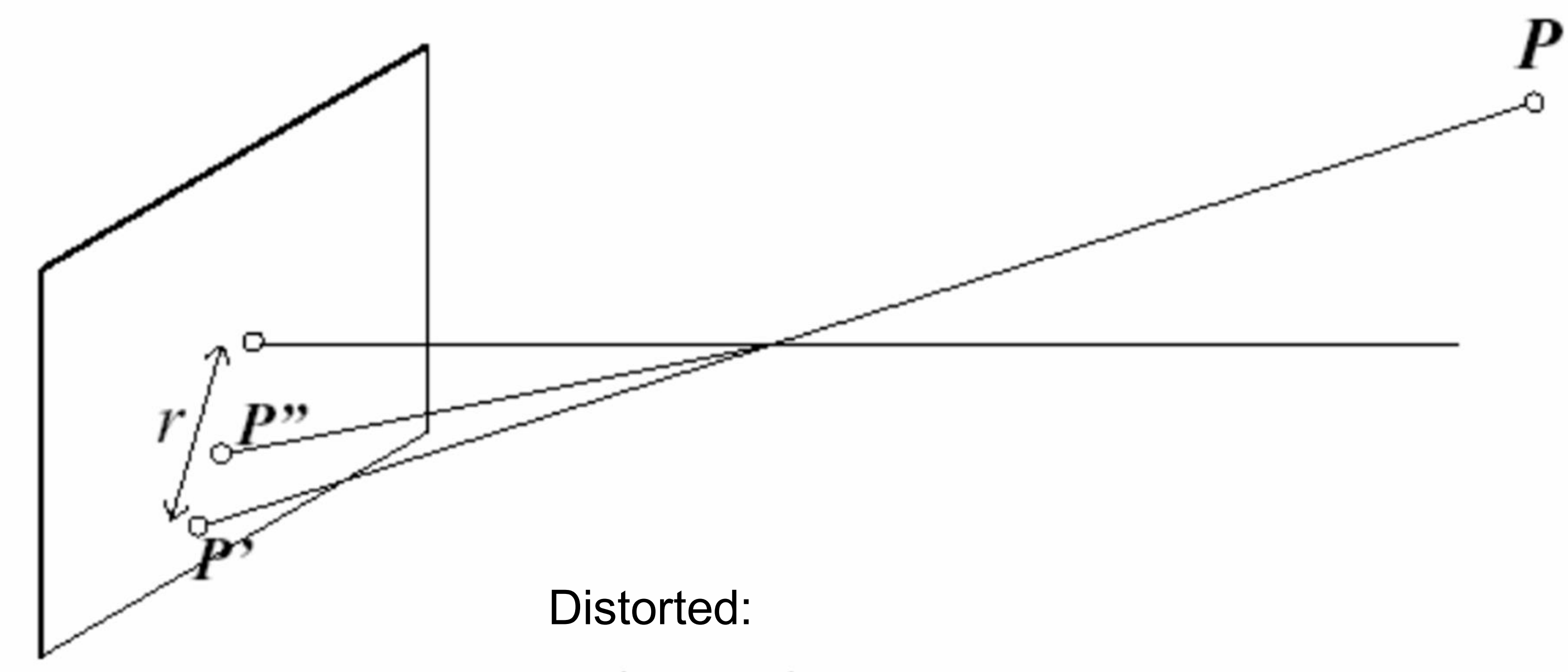


Geometric Distortion

(a) (b) (c)

Examples of radial lens distortion: (a) barrel; (b) pincushion, and (c) fisheye.

Radial Distortion Model



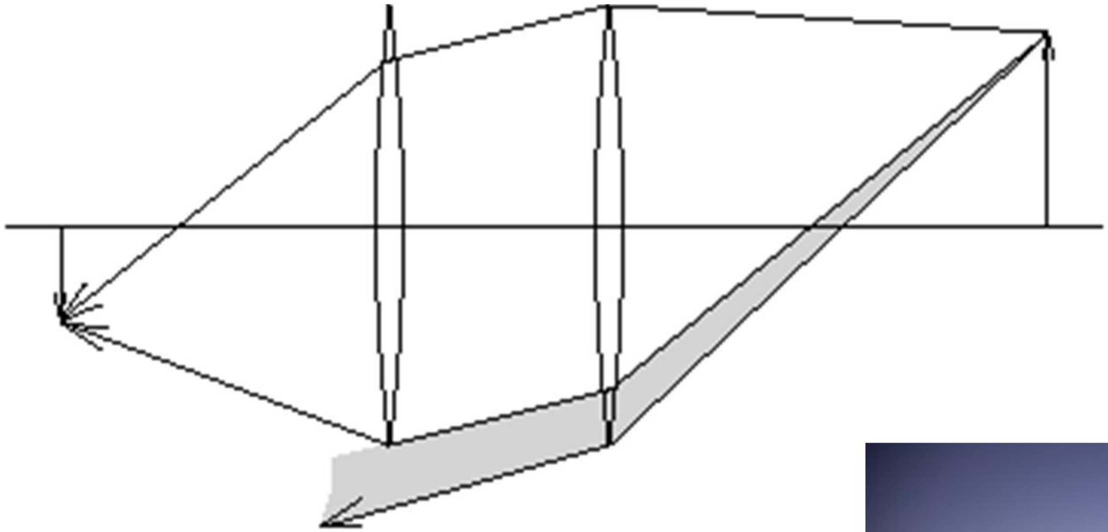
Distorted:

$$\begin{cases} x'' = \dfrac{1}{\lambda} x' \\ y'' = \dfrac{1}{\lambda} y' \end{cases} \qquad \lambda = 1 + k_1 r^2 + k_2 r^4 + \cdots$$

Ideal:

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

Vignetting



Perspective Projection	$x' = f\dfrac{x}{z}$ $y' = f\dfrac{y}{z}$	x,y: World coordinates x',y': Image coordinates f: pinhole-to-retina distance
Weak-Perspective Projection (Affine)	$x' \approx -mx$ $y' \approx -my$ $\quad m = -\dfrac{f}{\bar{z}}$	x,y: World coordinates x',y': Image coordinates m: magnification
Orthographic Projection (Affine)	$x' \approx x$ $y' \approx y$	x,y: World coordinates x',y': Image coordinates
Common distortion model	$x'' = \dfrac{1}{\lambda}x'$ $y'' = \dfrac{1}{\lambda}y'$ $\lambda = 1 + k_1 r^2 + k_2 r^4 + \cdots$	x',y': Ideal image coordinates x'',y'': Actual image coordinates

Camera Geometry and Calibration

Homogeneous coordinates in 2D

- Physical point $\mathrm{p} = \begin{bmatrix} x \\ y \end{bmatrix}$ represented by three coordinates $\begin{bmatrix} \mathrm{u} \\ \mathrm{v} \\ \mathrm{w} \end{bmatrix}$ $\begin{matrix} x = u / w \\ y = v / w \end{matrix}$ what if $w = 0$?

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

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

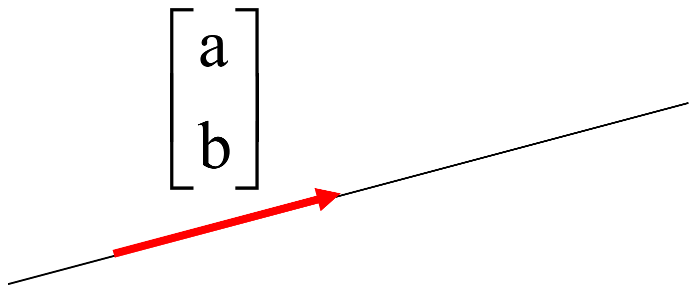
Lines in 2D

- General equation of a line in 2D

$$ax + by + c = 0$$

In homogeneous coordinates

$$l^{\mathrm{T}} p = l \bullet p = 0,\ l = \begin{bmatrix} a \\ b \\ c \end{bmatrix},\ p = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$



Homogeneous coordinates in 3D

- Physical point $\mathrm{p} = \begin{bmatrix} \mathrm{x} \\ \mathrm{y} \\ \mathrm{z} \end{bmatrix}$ represented by four coordinates $\begin{bmatrix} \mathrm{u} \\ \mathrm{v} \\ \mathrm{w} \\ \mathrm{t} \end{bmatrix}$

$$x = u / t$$
$$y = v / t \text{ what if } w = 0\,?$$
$$z = w / t$$

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

$$\begin{bmatrix} \mathrm{u} \\ \mathrm{v} \\ \mathrm{w} \\ \mathrm{t} \end{bmatrix} \equiv \begin{bmatrix} u' \\ v' \\ w' \\ t' \end{bmatrix} \Leftrightarrow \begin{bmatrix} \mathrm{u} \\ \mathrm{v} \\ \mathrm{w} \\ \mathrm{t} \end{bmatrix} \equiv \lambda \begin{bmatrix} u' \\ v' \\ w' \\ t' \end{bmatrix} \quad \lambda \neq 0$$

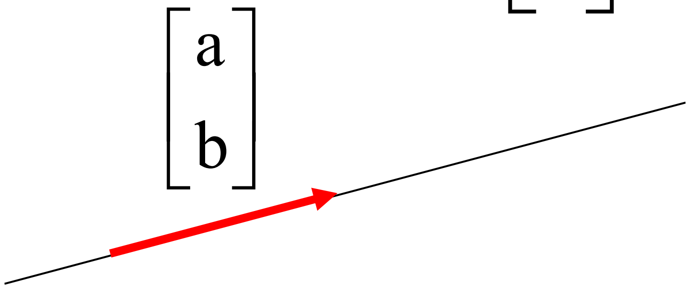
Lines in 3D

- General equation of a line in 3D

$$ax + by + cz + \mathrm{d} = 0$$

In homogeneous coordinates

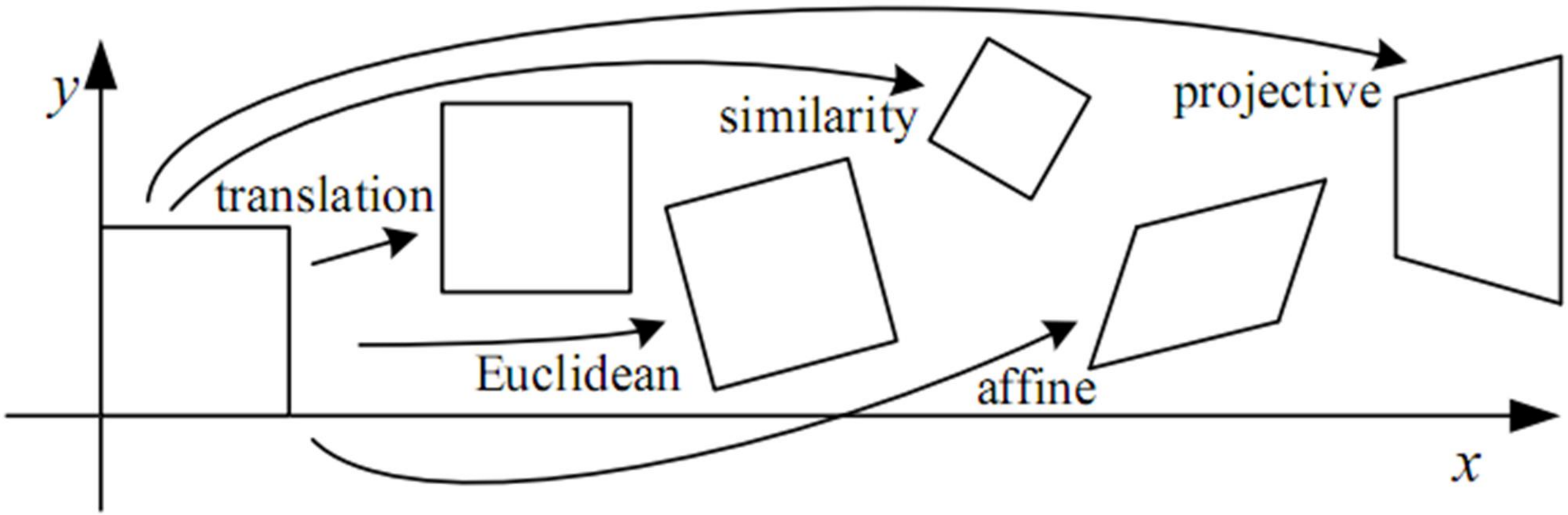
$$\Pi^{\mathrm{T}} p = \Pi \bullet p = 0, \Pi = \begin{bmatrix} \mathrm{a} \\ \mathrm{b} \\ \mathrm{c} \\ \mathrm{d} \end{bmatrix}, p = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

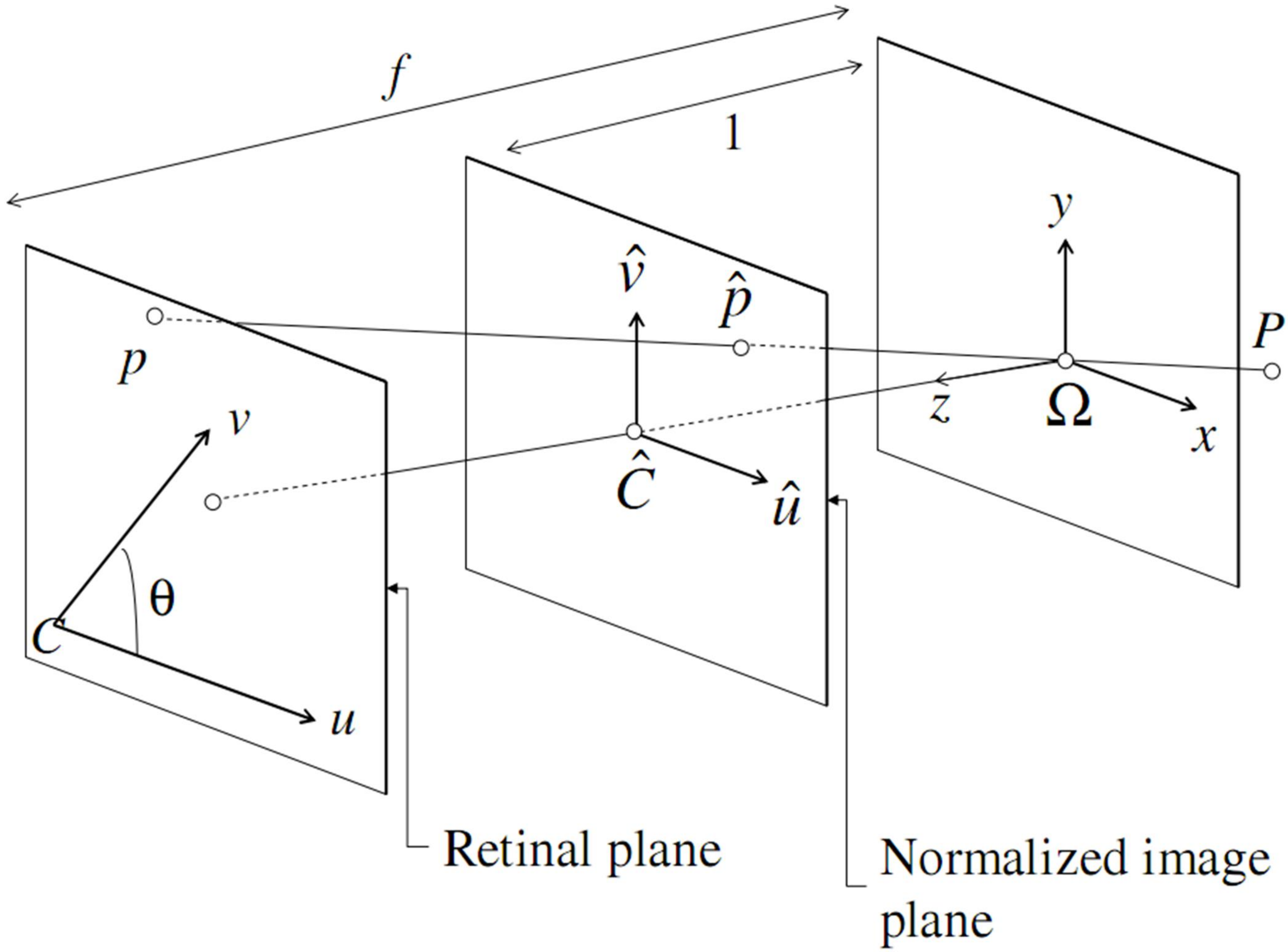


Basic Transformations

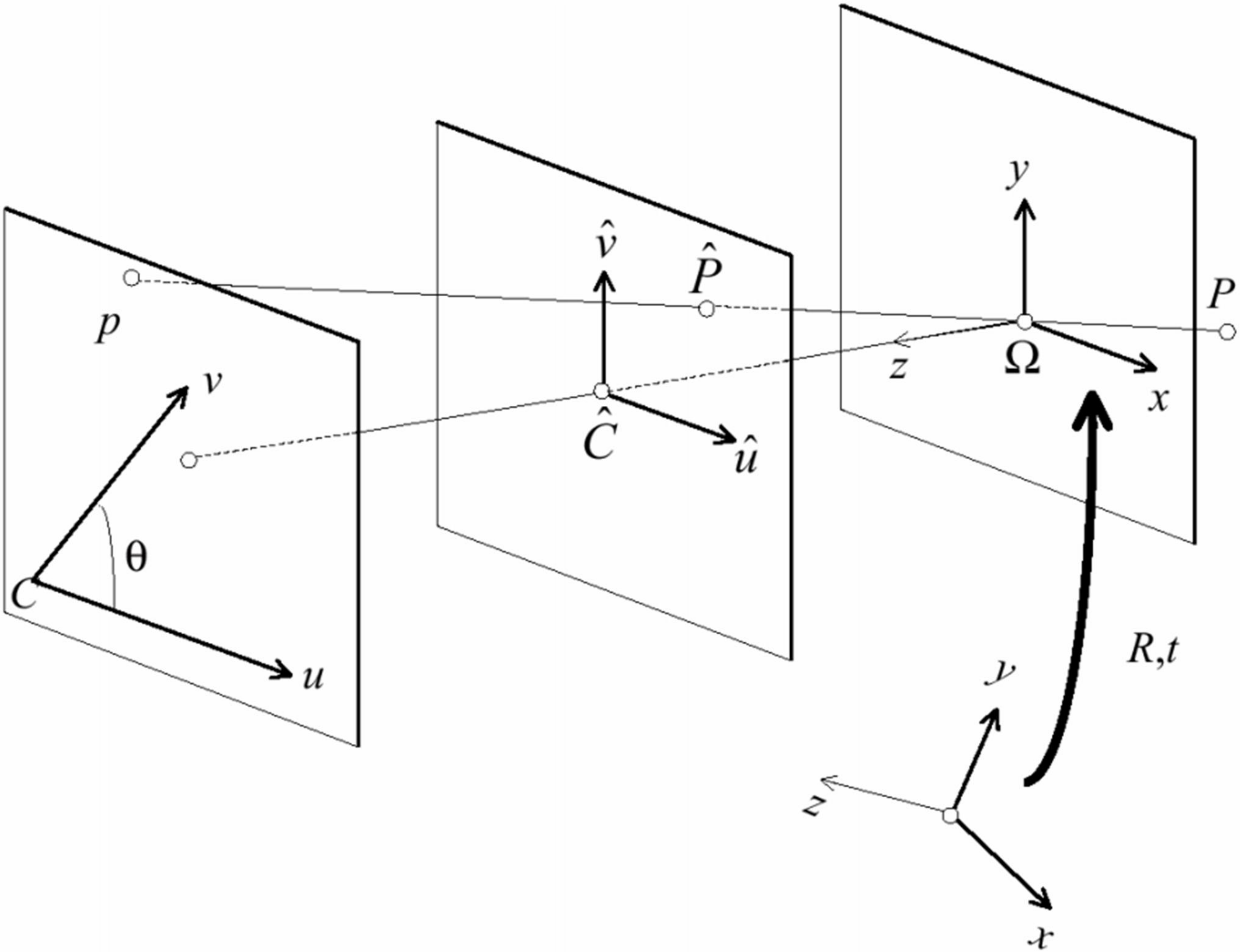
Transformation	Vector Coordinates	Homogeneous Coordinates	Degrees of Freedom	Invariants
Translation	$\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{t}$	$\begin{bmatrix} \boldsymbol{I} & \boldsymbol{t} \\ 0\ 0\ 0 & 1 \end{bmatrix}$	3	lengths, angles
Rotation	$\boldsymbol{y} = \boldsymbol{R}\boldsymbol{x}$ $\boldsymbol{R}^T\boldsymbol{R} = \boldsymbol{R}\boldsymbol{R}^T = \boldsymbol{I}$ $\mathrm{Det}(\boldsymbol{R}) = +1$	$\begin{bmatrix} \boldsymbol{R} & \boldsymbol{0} \\ 0\ 0\ 0 & 1 \end{bmatrix}$	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

Basic Set of 2D Planar Transformations





f
1
y
v̂
p̂
P
p
v
z
Ω
x
Ĉ
û
θ
C
u
Retinal plane
Normalized image plane



$\hat{v}$
$\hat{P}$
y
P
p
v
z
Ω
x
$\hat{C}$
$\hat{u}$
θ
C
u
R,t
y
z
x

Standard Perspective Projection Model

Scale in x direction between world coordinates and image coordinates

Skew of camera axes. $\theta = 90^{\circ}$ 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

Alternate Notations

By rows:

$$M = \begin{array}{cc} \begin{bmatrix} m_1^T \\ m_2^T \\ m_3^T \end{bmatrix} & 3 \\ 4 & \end{array}$$

By components:

Intrinsic parameter matrix

$$M = \underset{3x3}{K}\begin{bmatrix} \underset{3x3}{R} & \underset{3x1}{t} \end{bmatrix}$$

Extrinsic parameter matrix

By blocks:

$$M = \begin{bmatrix} \underset{3x3}{A} & \underset{3x1}{b} \end{bmatrix}$$

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

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

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

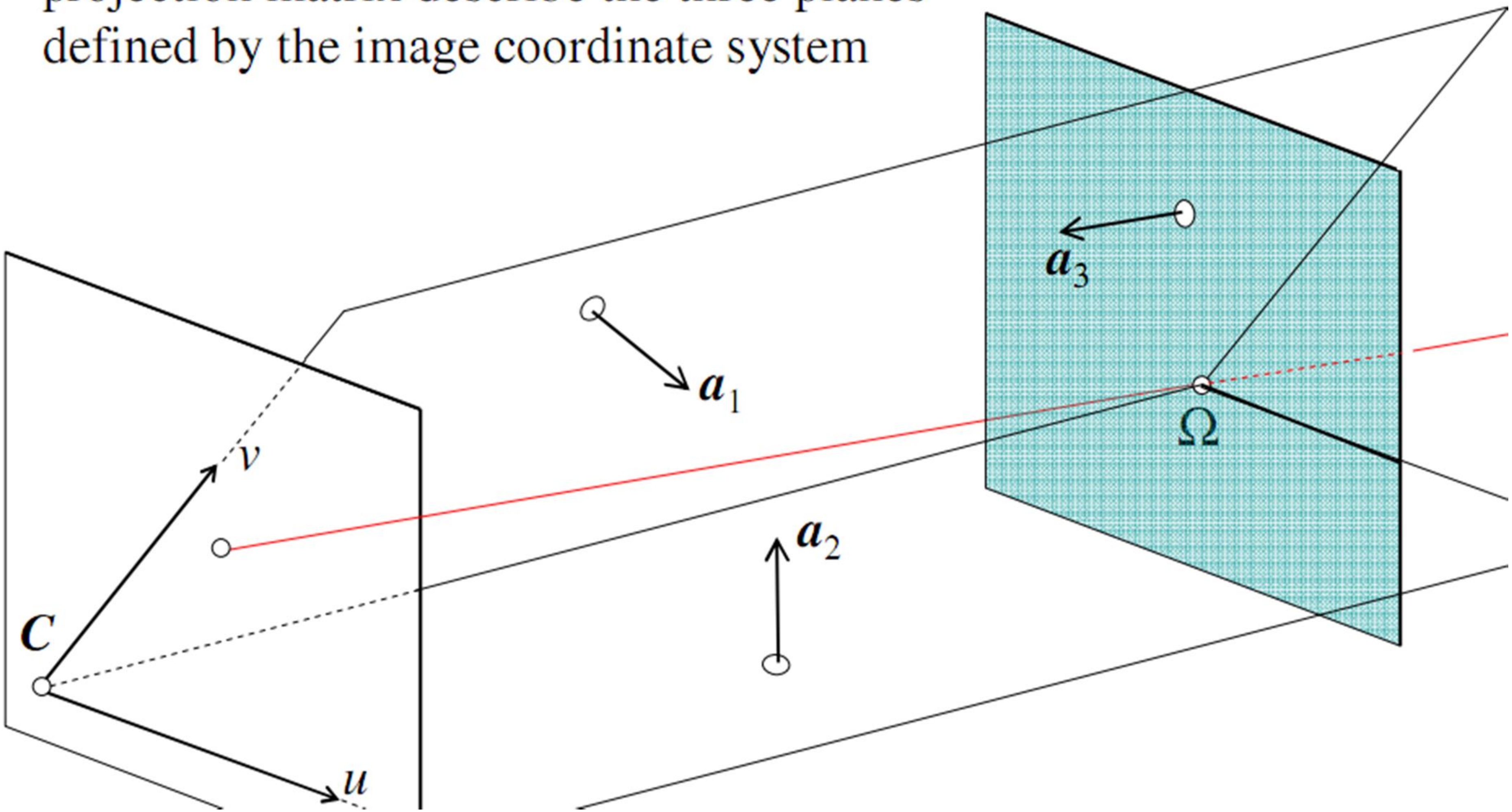
Homogeneous vector transformatio
p proportional to MP

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

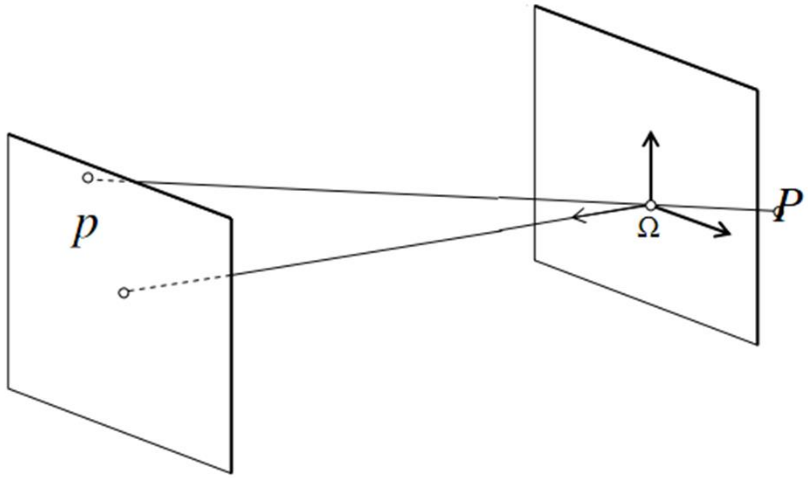
Computation of individual coordinates:

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

Geometric Interpretation: 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

2x2 intrinsic parameter matrix

First 2 components of the translation between world and camera frames

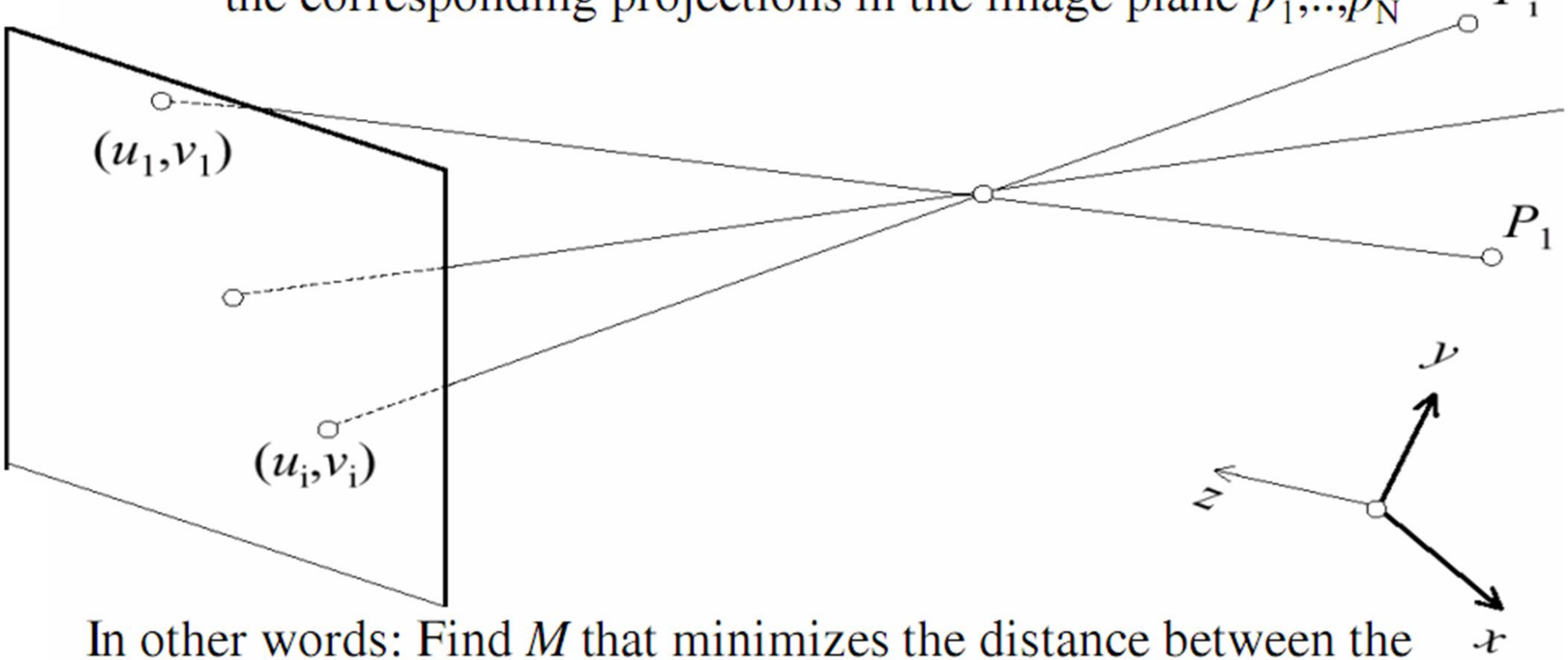
$$\begin{bmatrix} x \\ y \end{bmatrix} = M_{affine} P = K_2 [R_2 \quad t_2] P$$

2x4 projection matrix

2x3 matrix = first 2 rows of the rotation matrix between world and camera frames

Camera Calibration

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

The math for the calibration procedure follows a recipe that is used in many (most?) problems involving camera geometry, so it's worth remembering:

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 digression: 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:

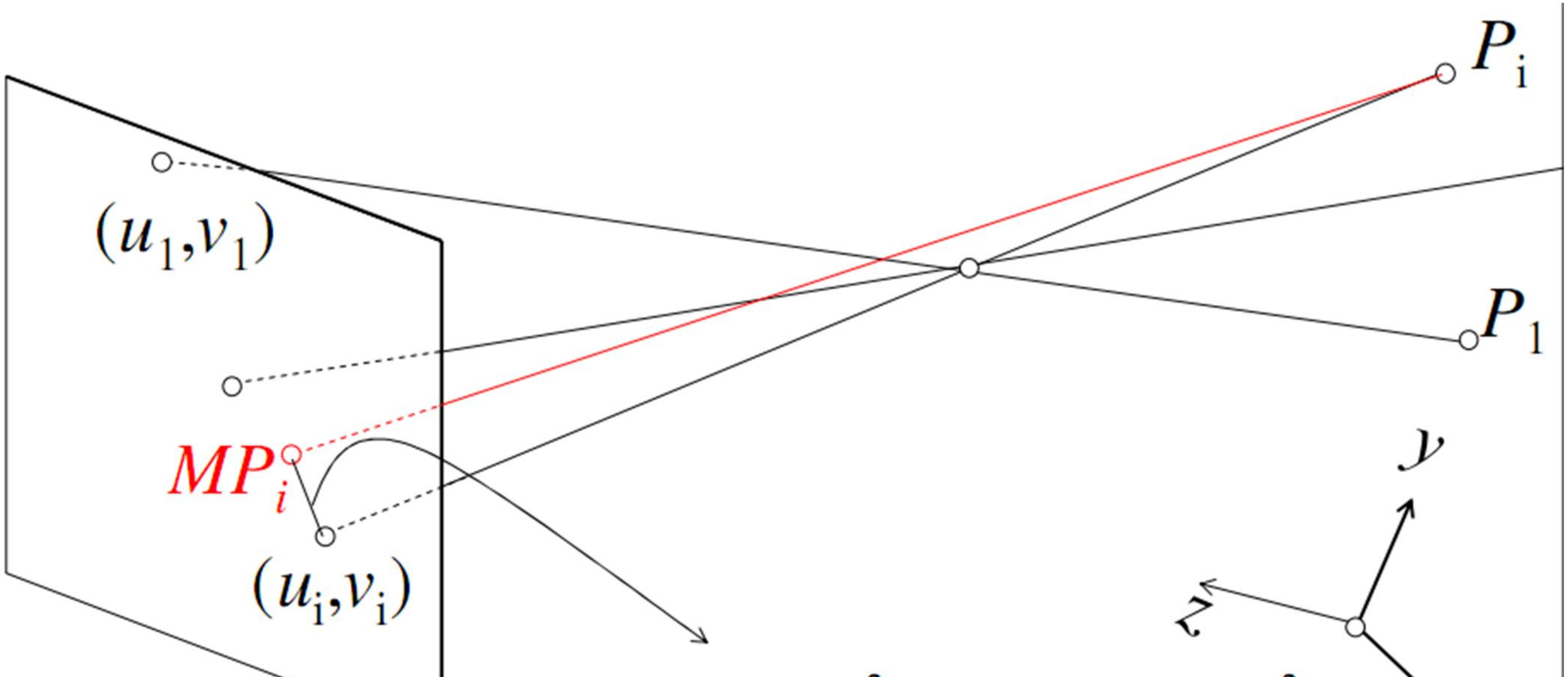
$$Min \ X^T V X$$

$$|X| = 1$$

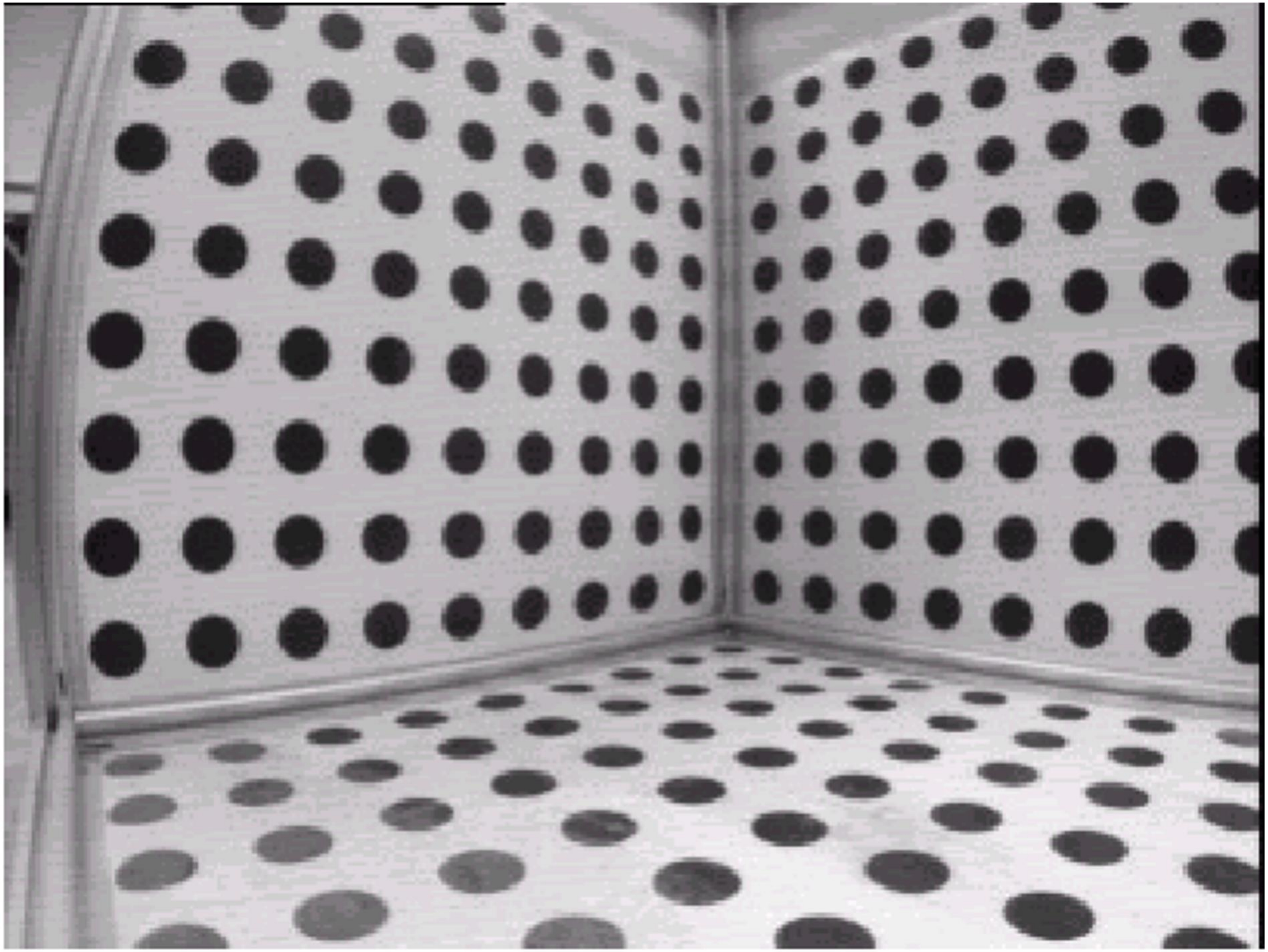
The key result (which we will use 50 times in this class) is:

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

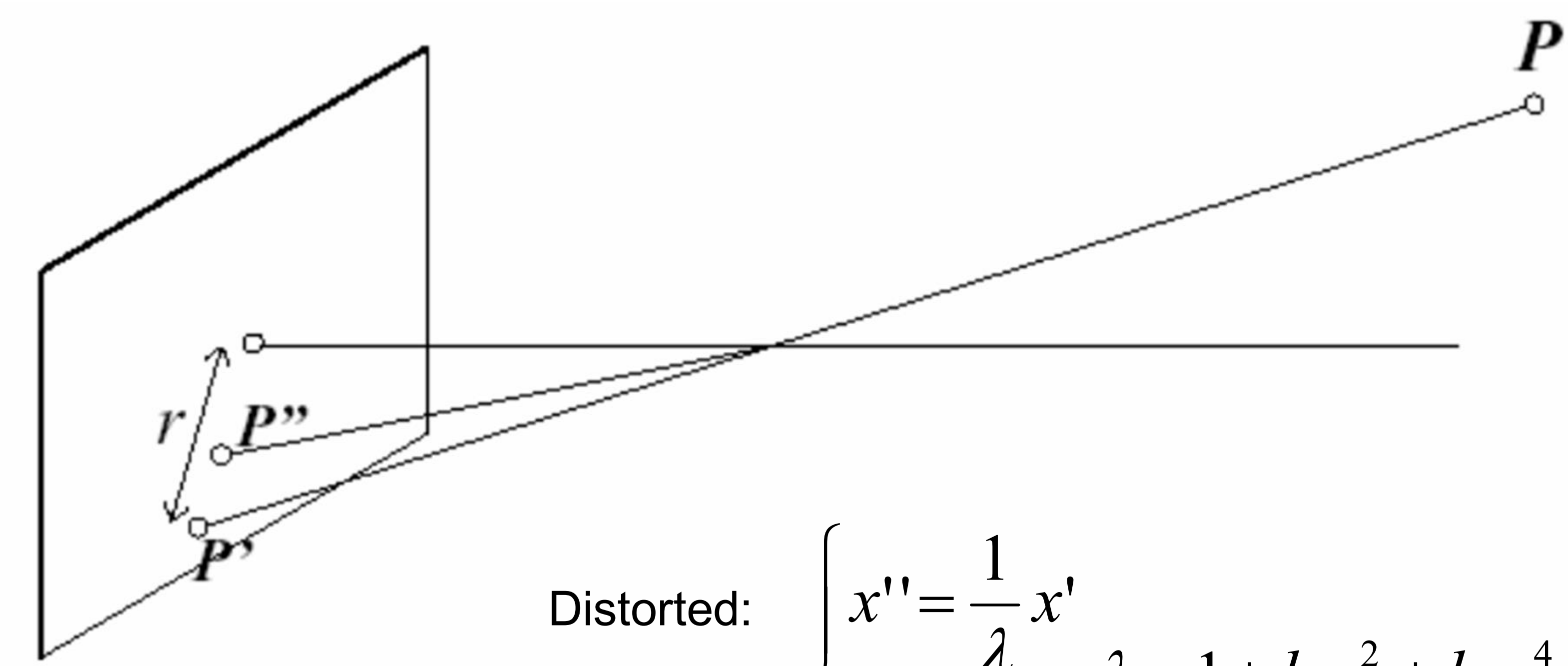
Calibration



$$\left(u_i - \frac{m_1 \cdot P_i}{m_3 \cdot P_i}\right)^2 + \left(v_i - \frac{m_2 \cdot P_i}{m_3 \cdot P_i}\right)^2$$



Radial Distortion Model



Distorted: $$\begin{cases} x'' = \dfrac{1}{\lambda} x' \\ y'' = \dfrac{1}{\lambda} y' \end{cases} \quad \lambda = 1 + k_1 r^2 + k_2 r^4 + \cdots$$

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

We can follow exactly the same recipe with non-linear distortion:

Write non-linear relations between coordinates:

$$\boxed{u_i = \frac{1}{\lambda}\frac{m_1{}^T P_i}{m_3{}^T P_i} \qquad v_i = \frac{1}{\lambda}\frac{m_2{}^T P_i}{m_3{}^T P_i}}$$

Make them linear:

$$\boxed{v_i\left(m_1{}^T P_i\right) - u_i\left(m_2{}^T P_i\right) = 0}$$

Write them in matrix form:

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

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

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

Solve by minimizing: $|Lm|^2 = m^T L^T L m \quad L = \begin{bmatrix} v_1 P_1^T & -u_1 P_1^T \\ \vdots & \vdots \\ \vdots & \vdots \\ \vdots & \vdots \\ v_N P_N{}^T & -u_N P_N{}^T \end{bmatrix}$

Subject to: $|m| = 1$

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: $|a_1 \times a_3| = |a_2 \times a_3|$

Planes: $m_i.P = 0$

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

Viewing ray: $w = A^{-1}\, p$

Calibration: minimum 6 non-coplanar points

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

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

Website for Camera Calibration

- http://www.vision.caltech.edu/bouguetj/calib_doc/
- http://www.ai.sri.com/~konolige/svs/svs.htm
- MATLAB calibration toolbox
- Intel OpenCV calibration package
- Omni-Camera calibration
  http://asl.epfl.ch/~scaramuz/research/Davide_Scaramuzza_files/Research/OcamCalib_Tutorial.htm