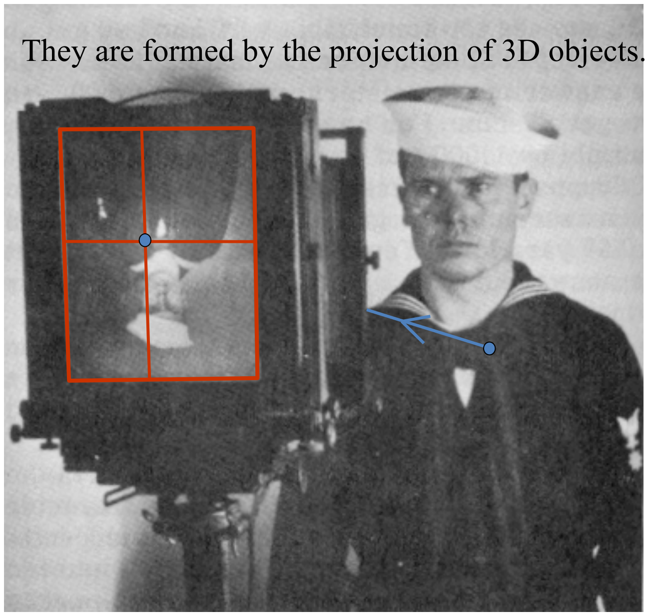
CS 534: Computer Vision
Camera Geometry

Ahmed Elgammal
Dept of Computer Science
Rutgers University

Outlines

- Projective Geometry
- Homogenous coordinates
- Euclidean Geometry
- Rigid Transformations
- Perspective Projection
- Other projection models
  - Weak perspective projection
  - Orthographic projection
- Camera intrinsic and extrinsic parameters



Images are two-dimensional patterns of brightness values.

Images of the 3-D world

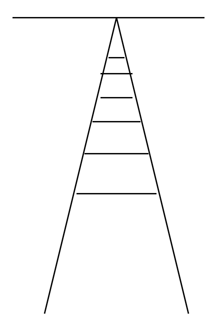
- What is the geometry of the image of a three-dimensional object?
  - Given a point in space, where will we see it in an image?
  - Given a line segment in space, what does its image look like?
  - Why do the images of lines that are parallel in space appear to converge to a single point in an image?
- How can we recover information about the 3-D world from a 2-D image?
  - Given a point in an image, what can we say about the location of the 3-D point in space?
  - Are there advantages to having more than one image in recovering 3-D information?
  - If we know the geometry of a 3-D object, can we locate it in space (say for a robot to pick it up) from a 2-D image?

Projective geometry 101

- Euclidean geometry describes shapes “as they are”
  - properties of objects that are unchanged by rigid motions
    - lengths
    - angles
    - parallelism
- Projective geometry describes objects “as they appear”
  - lengths, angles, parallelism become “distorted” when we look at objects
  - mathematical model for how images of the 3D world are formed

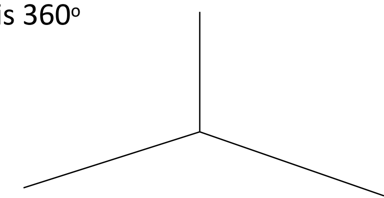
Example 1

- Consider a set of railroad tracks
  - Their actual shape:
    - tracks are parallel
    - ties are perpendicular to the tracks
    - ties are evenly spaced along the tracks
  - Their appearance
    - tracks converge to a point on the horizon
    - tracks don’t meet ties at right angles
    - ties become closer and closer towards the horizon



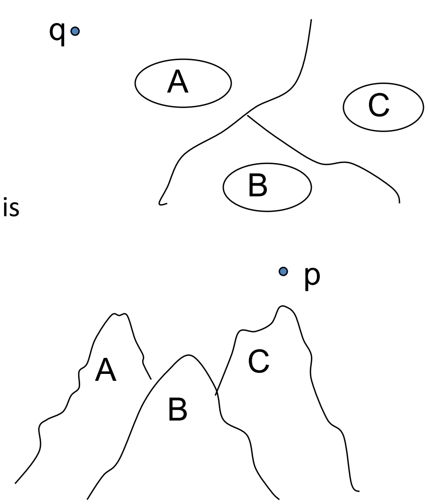
Example 2

- Corner of a room
  - Actual shape
    - three walls meeting at right angles. Total of 270^{o} of angle.
  - Appearance
    - a point on which three lines segments are concurrent. Total angle is 360^{o}



Example 3

- B appears between A and C from point p
- But from point q, A appears between B and C
- Apparent displacement of objects due to change in viewing position is called parallax shift



Projective Geometry

- Non-metrical description
- Less restrictive/ more general than Euclidean
- Describes properties that are invariant under projective transformation
- Invariant properties include
  – Incidence
  – Cross Ratio

	Euclidean	similarity	affine	projective
Transformations				
rotation	X	X	X	X
translation	X	X	X	X
uniform scaling		X	X	X
nonuniform scaling			X	X
shear			X	X
perspective projection				X
composition of projections				X
Invariants				
length	X			
angle	X	X		
ratio of lengths	X	X		
parallelism	X	X	X	
incidence	X	X	X	X
cross ratio	X	X	X	X

Figure 1: The four different geometries, the transformations allowed in each, and the measures that remain invariant under those transformations.

Projective Geometry

- Classical Euclidean geometry: through any point not on a given line, there exists a unique line which is parallel to the given line.
  - For 2,000 years, mathematician tried to "prove" this from Euclid's postulates.
  - In the early 19'th century, geometry was revolutionized when mathematicians asked: What if this were false?
  - That is, what if we assumed that EVERY pair of lines intersected?
  - To do this, we'll have to add points and lines to the standard Euclidean plane.

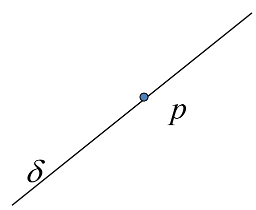
Homogeneous coordinates

- If (x,y) are the rectangular coordinates of a point, P, and if (x1, x2, x3) are any three real numbers such that:
  - x1/x3 = x
  - x2/x3 = y
- then (x1, x2, x3) are a set of homogeneous coordinates for (x,y).
- So, in particular, (x,y,1) are a set of homogeneous coordinates for (x,y)
- Given the homogeneous coordinates, (x1, x2, x3), the rectangular coordinates can be recovered.
- But (x,y) has an infinite number of homogeneous coordinate representations, because if (x1, x2, x3) are homogeneous coordinates of (x,y), then so are (kx1, kx2,k x3) for any k <>0.

Homogenous Coordinates
Lines

- On 2D plane

$$ax + by - d = 0 \Leftrightarrow \delta^T \cdot p = 0$$



$$\delta = \begin{pmatrix} a \\ b \\ -d \end{pmatrix} \qquad p = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

Homogenous Coordinates

- Cross product is the Intersection of two lines

$$x = l \times l'$$

$$proof: l \cdot (l \times l') = 0, l' \cdot (l \times l') = 0$$

- Intersection of parallel lines ?
- Line passing through two points: $p_1 x p_2$
- Three points on the same line: det[p_1 p_2 p_3]=0
- Three line intersecting in a point: det[l_1 l_2 l_3] =0

point	$\mathbf{p} = (X, Y, W)$
incidence	$\mathbf{p}^T\mathbf{u} = 0$
collinearity	$\lvert \mathbf{p}_1 \quad \mathbf{p}_2 \quad \mathbf{p}_3 \rvert = 0$
join of 2 points	$\mathbf{u} = \mathbf{p}_1 \times \mathbf{p}_2$
ideal points	$(X, Y, 0)$

(a)

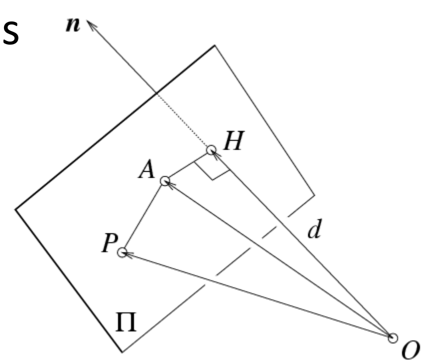
line	$\mathbf{u} = (a, b, c)$
incidence	$\mathbf{p}^T\mathbf{u} = 0$
concurrence	$\lvert \mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3 \rvert = 0$
intersection of 2 lines	$\mathbf{p} = \mathbf{u}_1 \times \mathbf{u}_2$
ideal line	$(0, 0, c)$

(b)

Table from S. Birchfield “An Introduction to Projective Geometry”

Homogenous Coordinates
Planes

$$P = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \qquad n = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$



$$\overrightarrow{AP}.\mathbf{n} = 0 \Leftrightarrow OP \cdot n - OA \cdot n = 0 \Leftrightarrow ax + by + cz - d = 0 \Leftrightarrow \mathbf{\Pi}^{\mathrm{T}}.\mathbf{P} = 0$$

where $\mathbf{\Pi} = \begin{bmatrix} a \\ b \\ c \\ -d \end{bmatrix}$ and $\mathbf{P} = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$ Homogenous coordinates

- Represent both points and lines as 4 numbers
- For 3D points, just add one to obtain its homogenous coordinate.
- Homogenous coordinates are defined up to scale: multiplying by any nonzero scale will not change the equation:

$$(a \quad b \quad c \quad -d)\cdot\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = 0$$

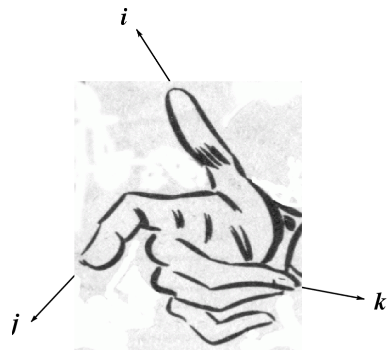
Homogenous Coordinates

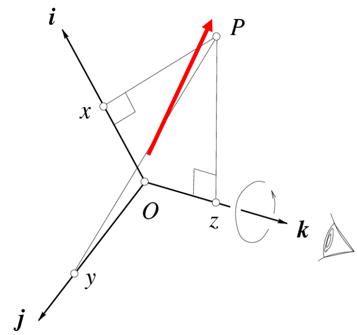
- Spheres

$$x^2 + y^2 + z^2 = R^2$$

$$(x \quad y \quad z \quad 1)\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & R^2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = 0$$

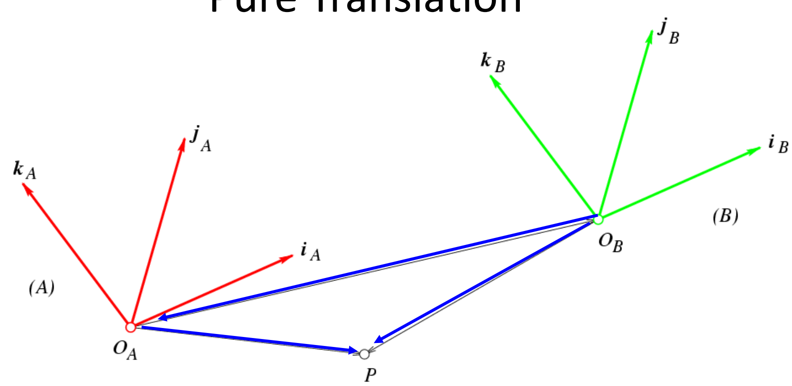
Euclidean Geometry
Coordinate Systems





$$\begin{cases} x = \overrightarrow{OP}.\mathbf{i} \\ y = \overrightarrow{OP}.\mathbf{j} \\ z = \overrightarrow{OP}.\mathbf{k} \end{cases} \Leftrightarrow \quad \overrightarrow{OP} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \quad \Leftrightarrow \quad \mathbf{P} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Coordinate System Change
Pure Translation



$$O_BP = O_BO_A + O_AP \;, \; {}^BP = {}^AP + {}^BO_A$$

Coordinate System Change
Pure Translation

Notations:
Left Superscript : Coordinate Frame of Reference

Origin of coordinate frame A in coordinate frame B

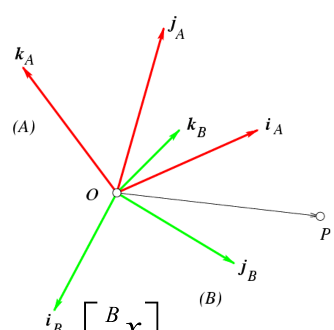
Point P in coordinate Frame A

Point P in coordinate Frame B

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

Coordinate System Change
Pure Rotation

Multiply both sides by $\begin{bmatrix} i_B \\ j_B \\ k_B \end{bmatrix}$

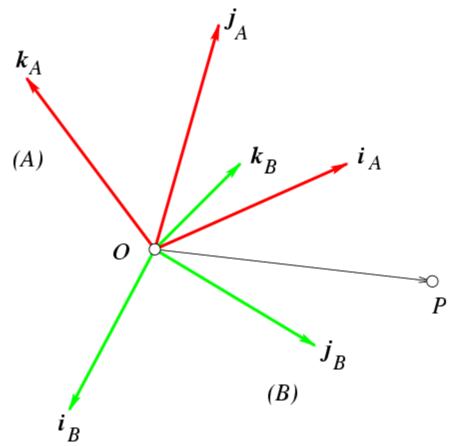


$$\overrightarrow{OP} = \begin{bmatrix} \mathbf{i}_A & \mathbf{j}_A & \mathbf{k}_A \end{bmatrix} \begin{bmatrix} {}^{A}x \\ {}^{A}y \\ {}^{A}z \end{bmatrix} = \begin{bmatrix} \mathbf{i}_B & \mathbf{j}_B & \mathbf{k}_B \end{bmatrix} \begin{bmatrix} {}^{B}x \\ {}^{B}y \\ {}^{B}z \end{bmatrix}$$

$$\Rightarrow \quad {}^{B}P = {}^{B}_{A}R\,{}^{A}P$$

$${}^{B}_{A}R = \begin{bmatrix} \mathbf{i}_A.\mathbf{i}_B & \mathbf{j}_A.\mathbf{i}_B & \mathbf{k}_A.\mathbf{i}_B \\ \mathbf{i}_A.\mathbf{j}_B & \mathbf{j}_A.\mathbf{j}_B & \mathbf{k}_A.\mathbf{j}_B \\ \mathbf{i}_A.\mathbf{k}_B & \mathbf{j}_A.\mathbf{k}_B & \mathbf{k}_A.\mathbf{k}_B \end{bmatrix}$$

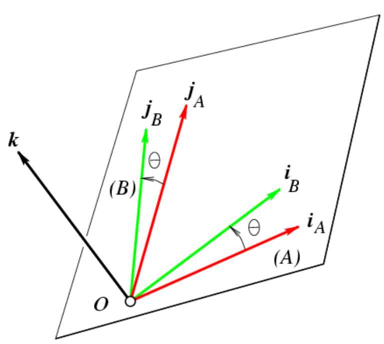
Coordinate Changes: Pure Rotations



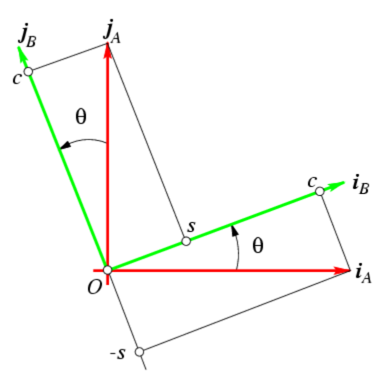
Rotation matrix describing the frame A in coordinate frame B

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

Coordinate Changes: Rotations about the z Axis



$${}^B_A R = \begin{bmatrix} \mathbf{i}_A.\mathbf{i}_B & \mathbf{j}_A.\mathbf{i}_B & \mathbf{k}_A.\mathbf{i}_B \\ \mathbf{i}_A.\mathbf{j}_B & \mathbf{j}_A.\mathbf{j}_B & \mathbf{k}_A.\mathbf{j}_B \\ \mathbf{i}_A.\mathbf{k}_B & \mathbf{j}_A.\mathbf{k}_B & \mathbf{k}_A.\mathbf{k}_B \end{bmatrix}$$



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

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:

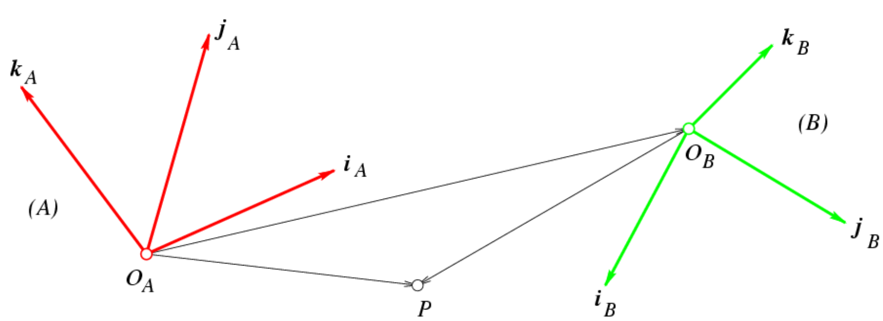
- Its rows (or columns) form a right-handed orthonormal coordinate system.

example

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

- In general, any rotation matrix can be written as the product of three elementary rotations about the i, j, and k vectors

Coordinate Changes: Rigid Transformations



$$ {}^{B}P = {}^{B}_{A}R \; {}^{A}P + {}^{B}O_{A} $$

Block Matrix Multiplication

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

What is AB ?

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

$$^{B}P = {}^{B}_{A}R\,{}^{A}P + {}^{B}O_{A}$$

Homogeneous Representation of Rigid Transformations

$$\boxed{\begin{bmatrix} ^{B}P \\ 1 \end{bmatrix} = \begin{bmatrix} {}^{B}_{A}R\,{}^{A}P + {}^{B}O_{A} \\ 1 \end{bmatrix} = \begin{bmatrix} {}^{B}_{A}R & {}^{B}O_{A} \\ \mathbf{0}^{T} & 1 \end{bmatrix} \begin{bmatrix} ^{A}P \\ 1 \end{bmatrix} = {}^{B}_{A}T \begin{bmatrix} ^{A}P \\ 1 \end{bmatrix}}$$

Other transformations

$$T = \begin{pmatrix} A & t \\ 0^{T} & 1 \end{pmatrix}$$

A is 3x3 rotation matrix :
Rigid transformation
Lengths and angels are preserved

A is arbitrary (nonsingular) 3x3 matrix :
Affine transformation
Lengths and angles may not be preserved

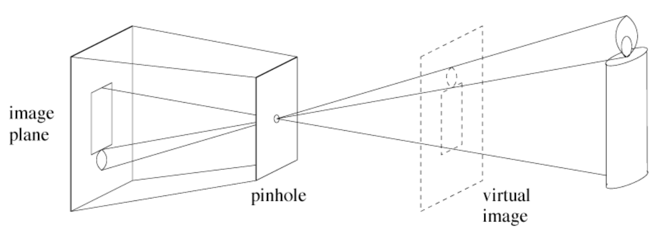
T is arbitrary (nonsingular) 4x4 matrix :
projective transformation
Lengths and angles may not be preserved

	Euclidean	similarity	affine	projective
Transformations				
rotation	X	X	X	X
translation	X	X	X	X
uniform scaling		X	X	X
nonuniform scaling			X	X
shear			X	X
perspective projection				X
composition of projections				X
Invariants				
length	X			
angle	X	X		
ratio of lengths	X	X		
parallelism	X	X	X	
incidence	X	X	X	X
cross ratio	X	X	X	X

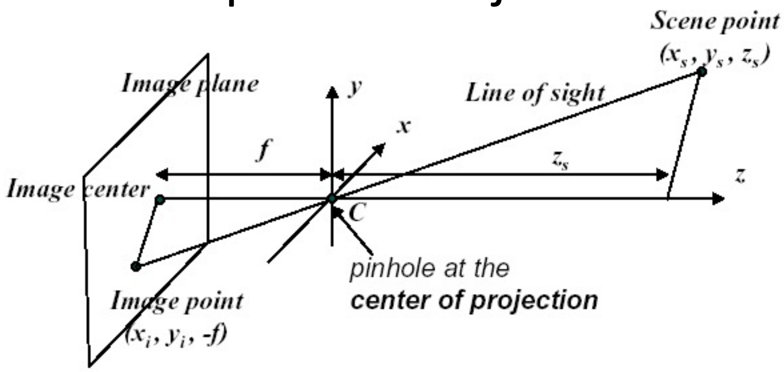
Figure 1: The four different geometries, the transformations allowed in each, and the measures that remain invariant under those transformations.

Pinhole Perspective

- Abstract camera model - box with a small hole in it
- Assume a single point pinhole (ideal pinhole):
  - Pinhole (central) perspective projection {Brunellleschi 15^{th} Century}
  - Extremely simple model for imaging geometry
  - Doesn't strictly apply
  - Mathematically convenient – acceptable approximation.
  - Concepts: image plane, virtual image plane
  - Moving the image plane merely scales the image.

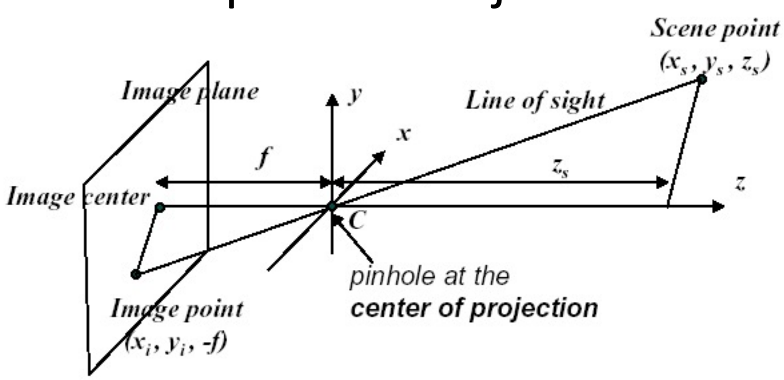


Perspective Projection



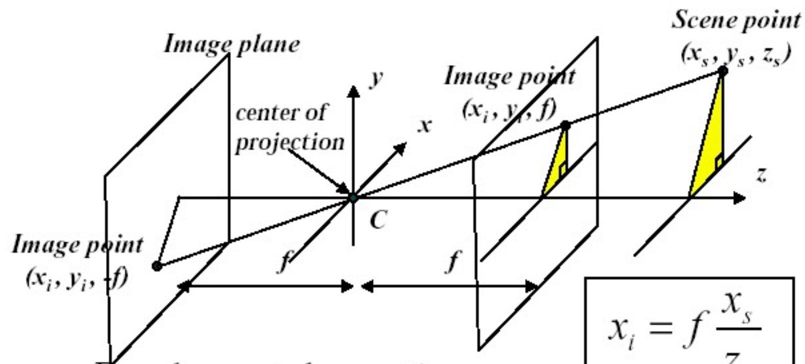
- Coordinate system center at the pinhole (center of projection).
- Image plane parallel to xy plane at distance f (focal length)
- Image center: intersection of z axis with image plane

Perspective Projection



- The point on the image plane that corresponds to a particular point in the scene is found by following the line that passes through the scene point and the center of projection.
- *Line of sight* to a point in the scene is the line through the center of projection to that point

Perspective Projection

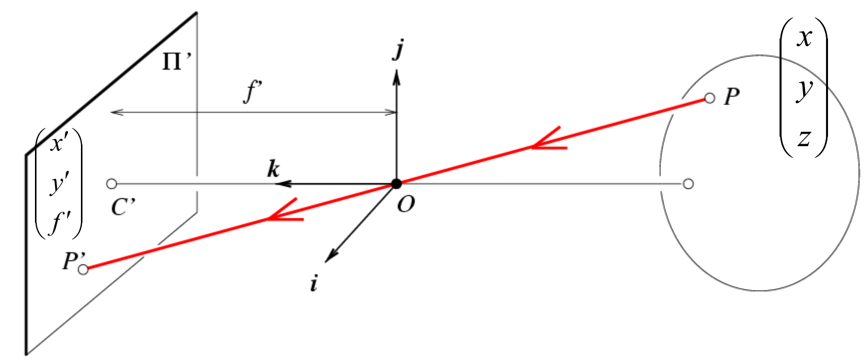


Fundamental equations for perspective projection onto a plane

$$x_i = f\,\frac{x_s}{z_s}$$

$$y_i = f\,\frac{y_s}{z_s}$$

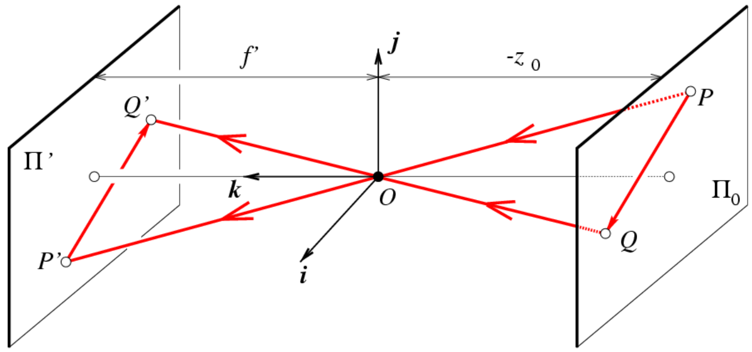
Pinhole Perspective Equation



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

NOTE: z is always negative..

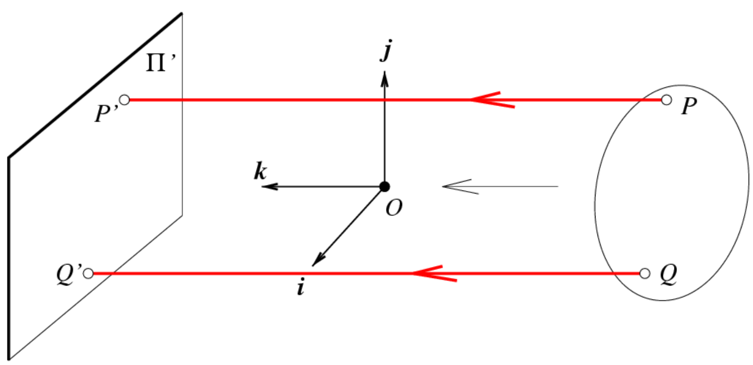
Affine projection models: Weak perspective projection



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

- When the scene depth is small compared its distance from the Camera, we can assume everything is on one plane,
- m can be taken constant: weak perspective projection
- also called scaled orthography (everything is a scaled version of the scene)

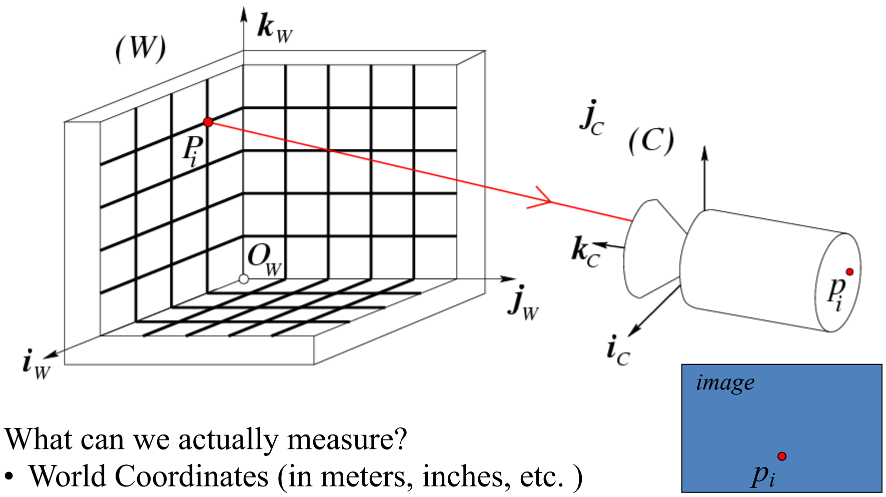
Affine projection models
Orthographic projection



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

When the camera is at a (roughly constant) distance from the scene, take m=1.
All rays are parallel to k axis.

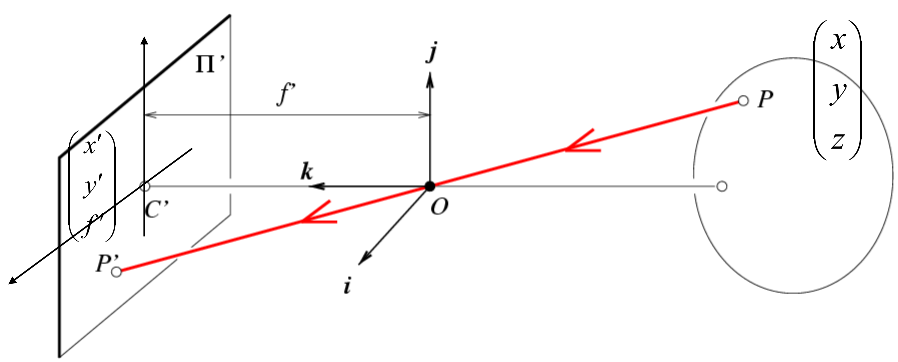
Quantitative Measurements and Calibration



What can we actually measure?

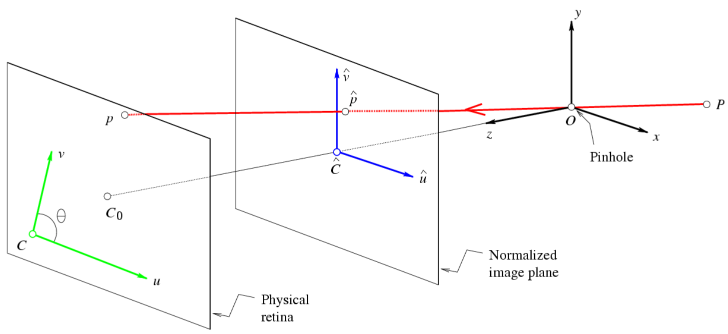
- World Coordinates (in meters, inches, etc.)
- Image Coordinates (in pixels)
- How to relate these measurements ?

Pinhole Perspective Equation



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

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \frac{1}{z}\begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

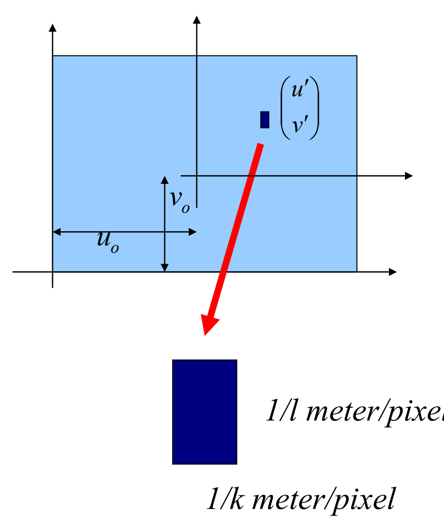


- We don't know where is the image center
- Pixels are rectangular
- Image axes are not necessary perpendicular (skew)

- Pixels are rectangular with scale parameters *k,l*

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

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



- Pixels are rectangular with scale parameters *k,l*

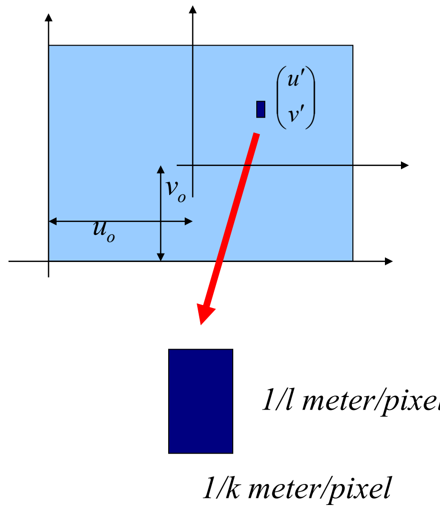
$$u' = kf\frac{x}{z} = \alpha\frac{x}{z}$$

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

- Move coordinate system to the corner

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

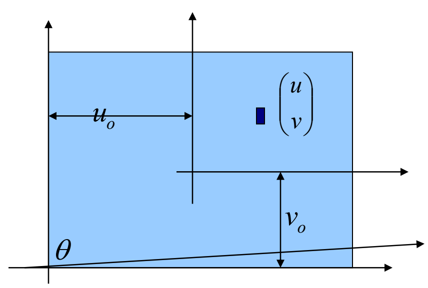
$$v = \beta\frac{y}{z} + v_o$$



- Pixel grid may not be exactly orthogonal
- $\theta \simeq 90$ but not exactly

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

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



$$u = \alpha\frac{x}{z} - \alpha\cot\theta\frac{y}{z} + u_o$$

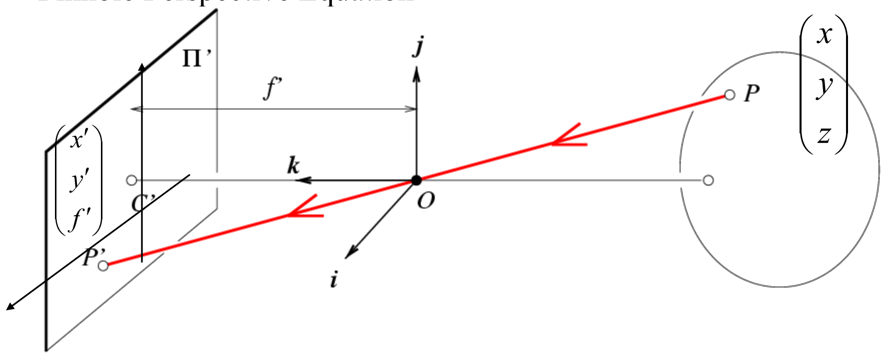
$$v = \frac{\beta}{\sin\theta}\frac{y}{z} + v_o$$

Approximation
s is skew parameter

$$u = \alpha\frac{x}{z} + s\frac{y}{z} + u_o$$

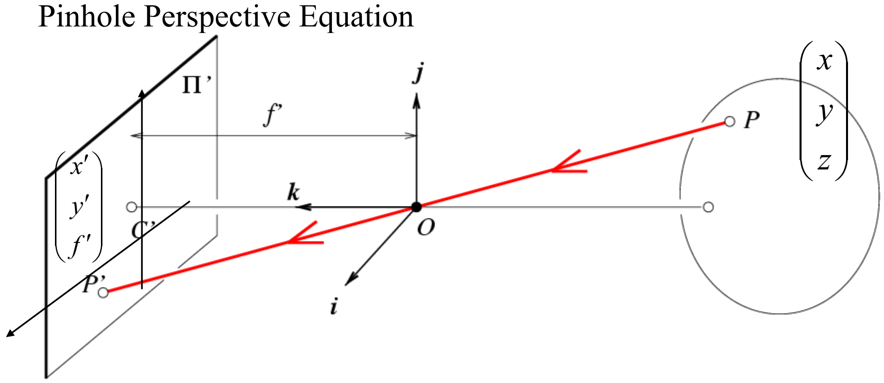
$$v = \beta\frac{y}{z} + v_o$$

Pinhole Perspective Equation



$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \frac{1}{z}\begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \qquad \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z}\begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Pinhole Perspective Equation



$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z}\begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \qquad \begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z}\begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z}\begin{bmatrix} kf & 0 & u_0 & 0 \\ 0 & lf & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z}\begin{bmatrix} \alpha & 0 & u_o & 0 \\ 0 & \beta & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \Rightarrow$$

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z}\begin{bmatrix} \alpha & -\alpha\cot\theta & u_o & 0 \\ 0 & \frac{\beta}{\sin\theta} & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad \text{OR} \quad \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z}\begin{bmatrix} \alpha & s & u_o & 0 \\ 0 & \beta & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z}\begin{bmatrix} \alpha & -\alpha\cot\theta & u_o & 0 \\ 0 & \frac{\beta}{\sin\theta} & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad \text{OR} \quad \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z}\begin{bmatrix} \alpha & s & u_o & 0 \\ 0 & \beta & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

- Five intrinsic camera parameters:
  - Magnification α, β *(in pixels)*
  - Image center location u_o, v_o *(in pixels)*
  - Skew measured as θ or s

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z} \begin{bmatrix} \alpha & -\alpha\cot\theta & u_o & 0 \\ 0 & \frac{\beta}{\sin\theta} & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad \text{OR} \quad \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z} \begin{bmatrix} \alpha & s & u_o & 0 \\ 0 & \beta & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

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

- Five intrinsic camera parameters:
  - Magnification α, β *(in pixels)*
  - Image center location u_o, v_o *(in pixels)*
  - Skew measured as θ or *s*

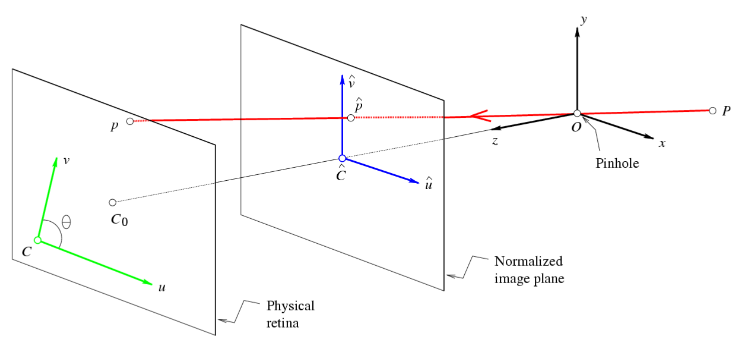
The Intrinsic Parameters of a Camera

Units:

k,l : pixel/m

f : m

α,β : pixel



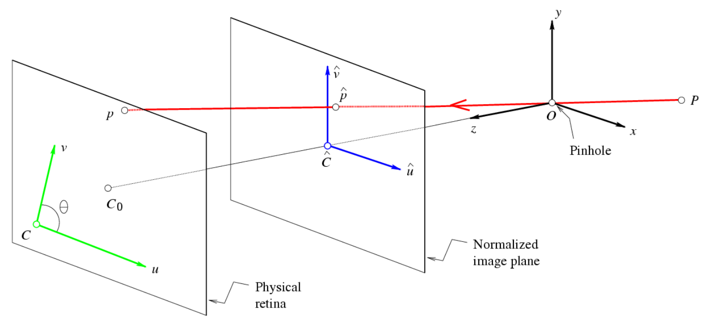
$$\begin{cases} \hat{u} = \dfrac{x}{z} \\ \hat{v} = \dfrac{y}{z} \end{cases} \iff \hat{\boldsymbol{p}} = \frac{1}{z}(\text{Id} \quad \mathbf{0})\begin{pmatrix} \boldsymbol{P} \\ 1 \end{pmatrix}$$

Normalized Image Coordinates

Physical Image Coordinates

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

The Intrinsic Parameters of a Camera



Calibration Matrix

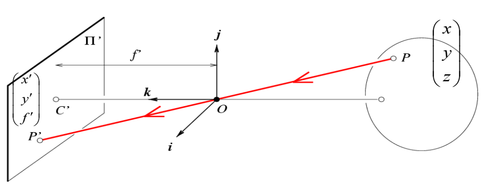
$$\boldsymbol{p} = \mathcal{K}\hat{\boldsymbol{p}}, \quad \text{where} \quad \boldsymbol{p} = \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \quad \text{and} \quad \mathcal{K} \stackrel{\text{def}}{=} \begin{pmatrix} \alpha & -\alpha\cot\theta & u_0 \\ 0 & \dfrac{\beta}{\sin\theta} & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

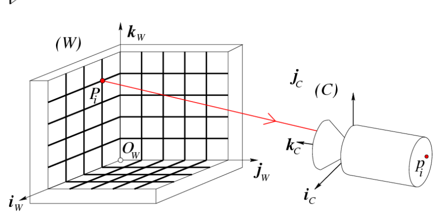
The Perspective Projection Equation

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

Extrinsic Parameters:

- Everything in the world so far is measured as if the pinhole is the coordinate center.
- Let's move to a real world coordinate system
- Where is the pinhole in the world coordinate system? [translation – 3 parameters]
- What is the orientation of the camera ? [rotation – 3 parameters]





$$\begin{pmatrix} {}^{C}P \\ 1 \end{pmatrix} = \begin{pmatrix} {}^{C}_{W}R_{3\times3} & {}^{C}O_{W\,3\times1} \\ 0^{T}{}_{1\times3} & 1 \end{pmatrix}_{4\times4} \begin{pmatrix} {}^{W}P \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} {}^{C}P \\ 1 \end{pmatrix} = \begin{pmatrix} {}^{C}_{W}R_{3\times3} & {}^{C}O_{W\,3\times1} \\ 0^{T}{}_{1\times3} & 1 \end{pmatrix}_{4\times4} \begin{pmatrix} {}^{W}P \\ 1 \end{pmatrix}$$

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} \alpha & -\alpha\cot\theta & u_o & 0 \\ 0 & \frac{\beta}{\sin\theta} & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} K_{3\times3} & 0_{3\times1} \end{bmatrix} \begin{pmatrix} {}^{C}_{W}R_{3\times3} & {}^{C}O_{W\,3\times1} \\ 0^{T}{}_{1\times3} & 1 \end{pmatrix}_{4\times4} \begin{pmatrix} {}^{W}P \\ 1 \end{pmatrix}$$

$$\begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} = \begin{bmatrix} K_{3\times3} & 0_{3\times1} \end{bmatrix} \begin{pmatrix} {}^{C}_{W}R_{3\times3} & {}^{C}O_{W\,3\times1} \\ 0^{T}{}_{1\times3} & 1 \end{pmatrix}_{4\times4} \begin{pmatrix} {}^{W}P \\ 1 \end{pmatrix}$$

$$\begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} = K_{3\times3} \begin{pmatrix} {}^{C}_{W}R_{3\times3} & {}^{C}O_{W\,3\times1} \end{pmatrix}_{3\times4} \begin{pmatrix} {}^{W}P \\ 1 \end{pmatrix} \Rightarrow p = MP$$

Rotation *Translation*

Intrinsic parameters

Extrinsic parameters

Projection matrix

$$\begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} = K_{3\times3}\begin{pmatrix} {}^C_W R_{3\times3} & {}^C O_{W\,3\times1} \end{pmatrix}_{3\times4} \begin{pmatrix} {}^W P \\ 1 \end{pmatrix} \Rightarrow p = MP$$

Rotation *Translation*

Intrinsic parameters

Extrinsic parameters

Projection matrix 3x4

$$M = K_{3\times3}\begin{pmatrix} {}^C_W R_{3\times3} & {}^C O_{W\,3\times1} \end{pmatrix}_{3\times4}$$

Only 11 free parameters (not 12):
5 intrinsic, 3 for rotation, 3 for translation
Perspective Projection Matrix

$$\begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} = K_{3\times3}\begin{pmatrix} {}^C_W R_{3\times3} & {}^C O_{W\,3\times1} \end{pmatrix}_{3\times4} \begin{pmatrix} {}^W P \\ 1 \end{pmatrix} \Rightarrow p = MP$$

$$M = K_{3\times3}\begin{pmatrix} {}^C_W R_{3\times3} & {}^C O_{W\,3\times1} \end{pmatrix}_{3\times4}$$

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.

$\boxed{M \text{ is only defined up to scale in this setting!!}}$

Explicit Form of the Projection Matrix

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

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

$$p = MP = \begin{pmatrix} m_1^T \\ m_2^T \\ m_3^T \end{pmatrix} P \quad \Rightarrow \quad u = \frac{\boldsymbol{m}_1^T \cdot \boldsymbol{P}}{\boldsymbol{m}_3^T \cdot \boldsymbol{P}} \qquad v = \frac{\boldsymbol{m}_2^T \cdot \boldsymbol{P}}{\boldsymbol{m}_3^T \cdot \boldsymbol{P}}$$

Replacing M by λM doesn't change u or v

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

Camera Calibration

- Find the intrinsic and extrinsic parameters of a camera
- VERY large literature on the subject
- Work of Roger Tsai influential
- Basic idea: Given a set of world points P_i and their image coordinates (u_i, v_i) find the projection matrix and then find intrinsic and extrinsic parameters.

Sources

- Forsyth and Ponce, Computer Vision a Modern approach: 1.1, 2.1,2.2
- Fougeras, Three-dimensional Computer Vision
- Slides by J. Ponce UIUC
- Slides by L. S. Davis UMD