

CS 534: Computer Vision
Camera Geometry
Camera Calibration

Spring 2004

Ahmed Elgammal
Dept of Computer Science
Rutgers University

CS 534 – Ahmed Elgammal 1

Project ideas

- Evaluation projects : evaluate different approaches on some data sets.
- Research Projects: modify some approach with some new idea.
- Implementation and evaluation Project: implement and evaluate an algorithm.

Examples:

- Shape matching
- Face Detection
- Face Recognition
- Moving Object Detection
- Tracking
- Segmentation

CS 534 – Ahmed Elgammal 2

Outlines

- Perspective Projection
- Other projection models
 - Weak perspective projection
 - Orthographic projection
- Euclidean Geometry
- Homogenous coordinates/ Projective Geometry
- Geometric Transformations
- Camera intrinsic and extrinsic parameters
- Camera Calibration
- Linear least-Squares and Singular Value Decomposition

CS 534 – Ahmed Elgammal 3

They are formed by the projection of 3D objects.

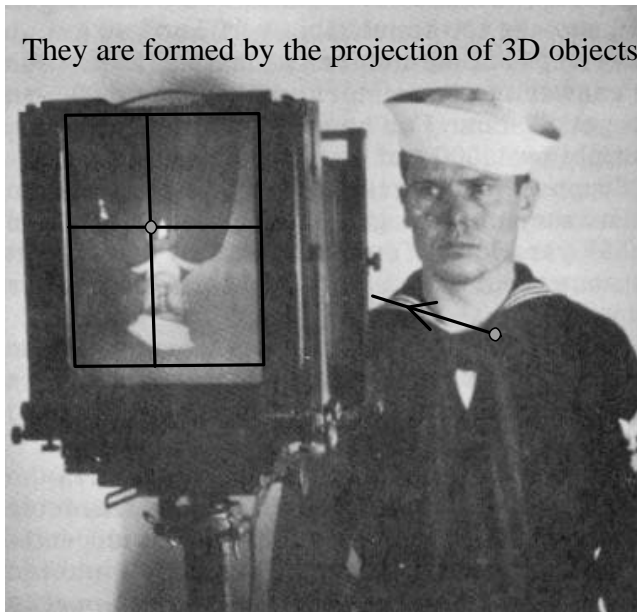


Figure from US Navy Manual of Basic Optics and Optical Instruments, prepared by Bureau of Naval Personnel. Reprinted by Dover Publications, Inc., 1969.

Images are two-dimensional patterns of brightness values.

CS 534 – Ahmed Elgammal 4

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?

CS 534 – Ahmed Elgammal 5

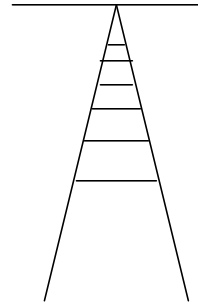
Projective geometry 101

- 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

CS 534 – Ahmed Elgammal 6

Example 1

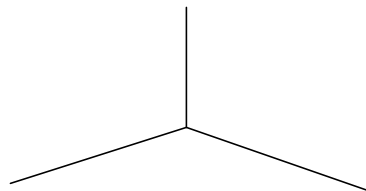
- Consider a set of railroad tracks
 - Their actual shape:
 - tracks are parallel
 - ties are perpendicular to the tracks
 - ties are evenly spaced along the tracks
 - Their appearance
 - tracks converge to a point on the horizon
 - tracks don't meet ties at right angles
 - ties become closer and closer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



CS 534 – Ahmed Elgammal 7

Example 2

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

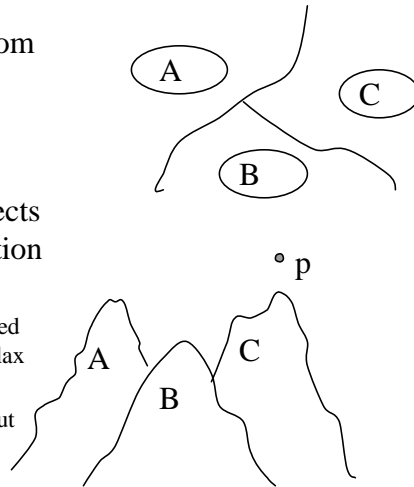


CS 534 – Ahmed Elgammal 8

Example 3

q

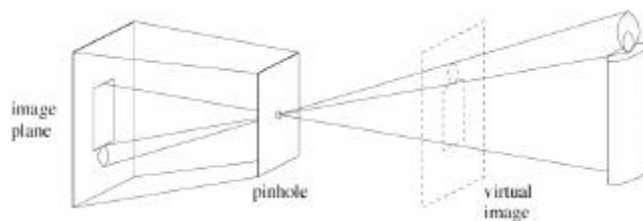
- B appears between A and C from point p
- But from point q, A appears between B and C
- Apparent displacement of objects due to change in viewing position is called **parallax shift**
 - Greeks knew that if the earth revolved around the sun there would be parallax shift of the stars
 - Tycho Brahe looked for this shift, but his instruments were not accurate enough.



CS 534 – Ahmed Elgammal 9

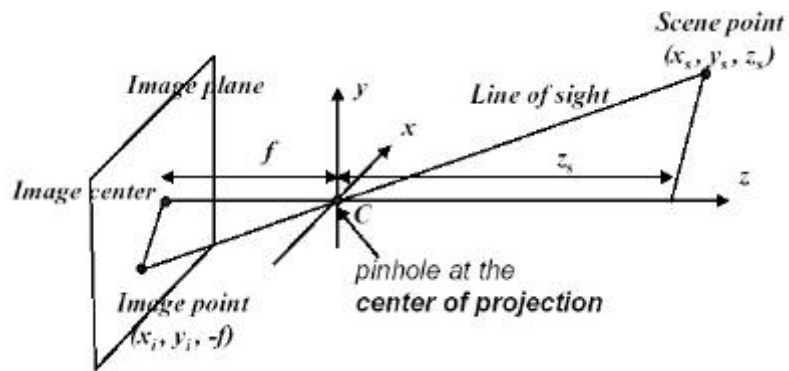
Pinhole Perspective

- Abstract camera model - box with a small hole in it
- Assume a single point pinhole (ideal pinhole):
 - Pinhole (central) perspective projection {Brunelleschi 15th 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.



CS 534 – Ahmed Elgammal 10

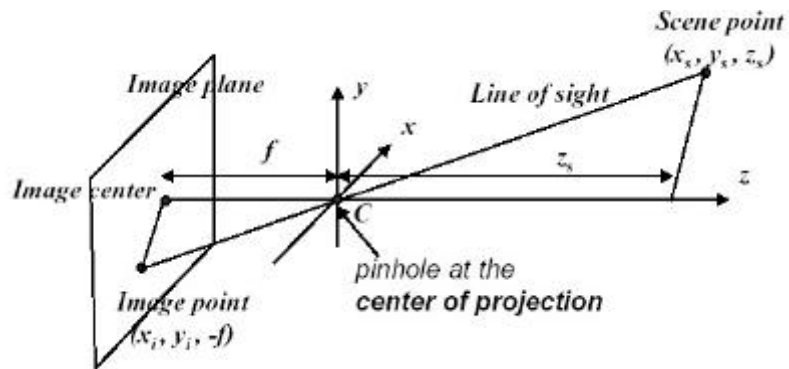
Perspective Projection



- 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

CS 534 – Ahmed Elgammal 11

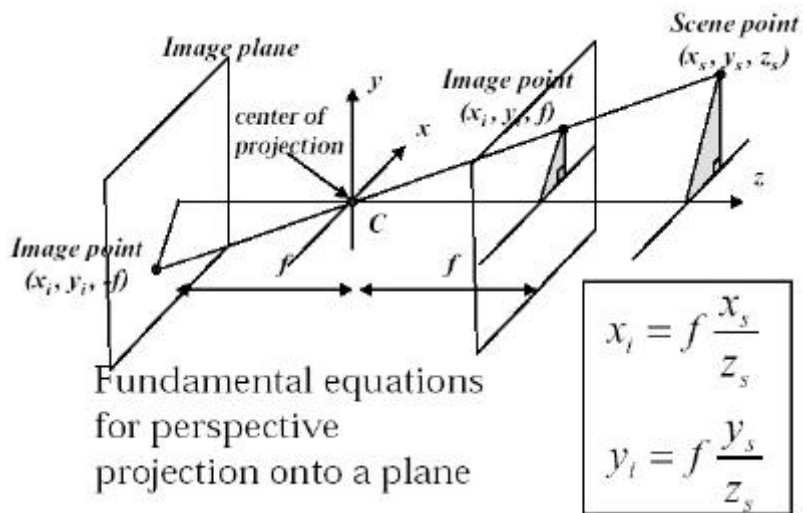
Perspective Projection



- 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

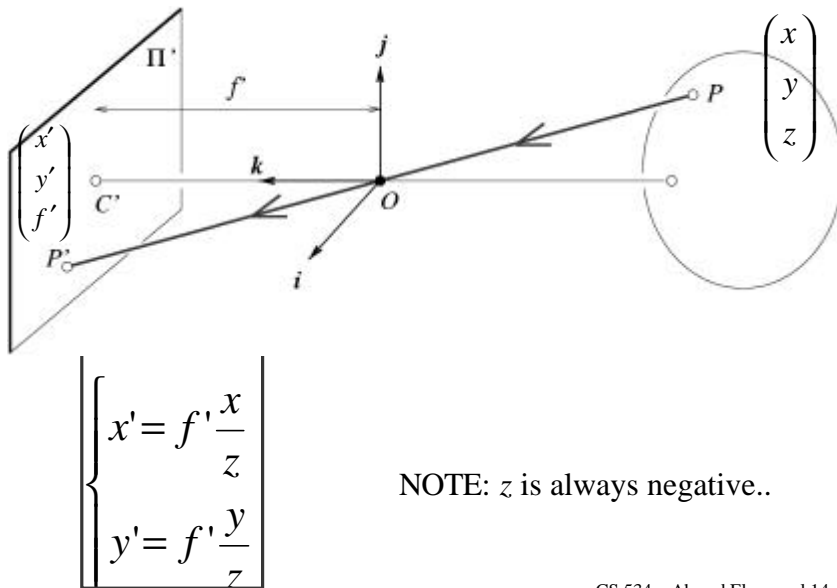
CS 534 – Ahmed Elgammal 12

Perspective Projection



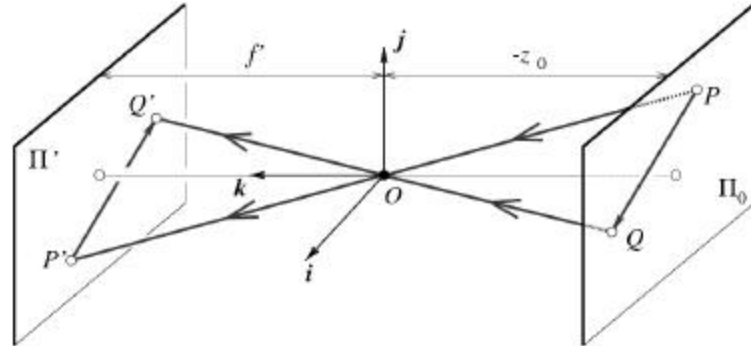
CS 534 – Ahmed Elgammal 13

Pinhole Perspective Equation



CS 534 – Ahmed Elgammal 14

Affine projection models: Weak perspective projection

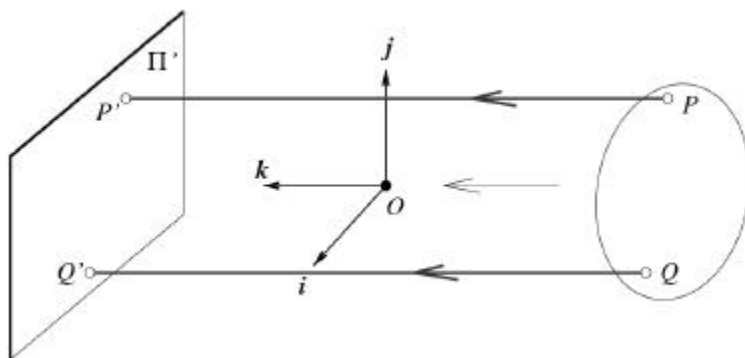


$$\begin{cases} x' = f' \frac{x}{z} \\ y' = f' \frac{y}{z} \end{cases} \Rightarrow \begin{cases} x' = -mx \\ y' = -my \end{cases} \text{ where } m = -\frac{f'}{z_0} \text{ is the magnification.}$$

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

CS 534 – Ahmed Elgammal 15

Affine projection models: Orthographic projection

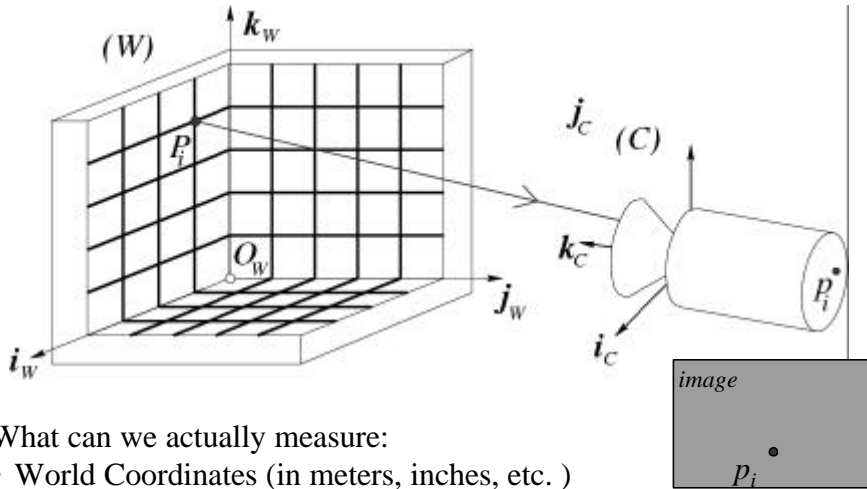


$$\begin{cases} x' = x \\ y' = y \end{cases}$$

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

CS 534 – Ahmed Elgammal 16

Quantitative Measurements and Calibration

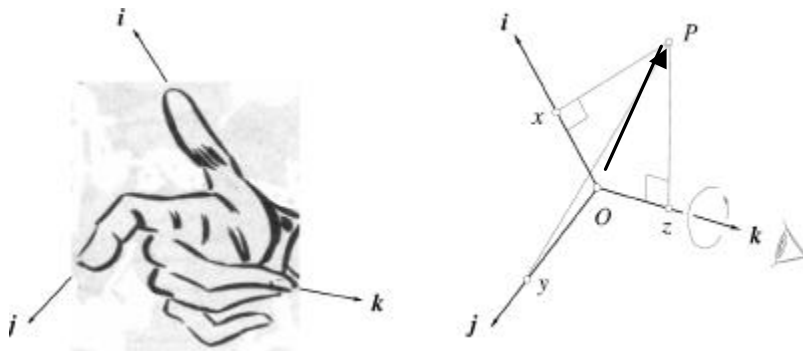


What can we actually measure:

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

CS 534 – Ahmed Elgammal 17

Euclidean Coordinate Systems

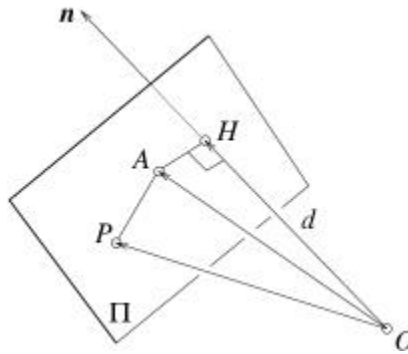


$$\begin{cases} x = \overline{OP} \cdot \mathbf{i} \\ y = \overline{OP} \cdot \mathbf{j} \\ z = \overline{OP} \cdot \mathbf{k} \end{cases} \Leftrightarrow \overline{OP} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \Leftrightarrow \mathbf{P} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

CS 534 – Ahmed Elgammal 18

Planes

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



$$\overrightarrow{AP} \cdot \mathbf{n} = 0 \Leftrightarrow OP \cdot \mathbf{n} - OA \cdot \mathbf{n} = 0 \Leftrightarrow ax + by + cz - d = 0 \Leftrightarrow ? \cdot \mathbf{P} = 0$$

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

CS 534 – Ahmed Elgammal 19

Homogenous Coordinates

- From Projective Geometry (will be revisited later)
- 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:

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

CS 534 – Ahmed Elgammal 20

Homogenous Coordinates

- Spheres

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

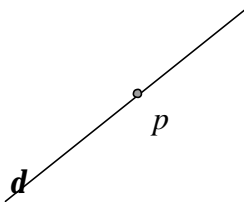
$$(x \ y \ z \ 1) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & R^2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = 0$$

CS 534 – Ahmed Elgammal 21

Homogenous Coordinates

- On 2D plane

$$ax + by - d = 0 \Leftrightarrow \mathbf{d}^T \cdot \mathbf{p} = 0$$



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

CS 534 – Ahmed Elgammal 22

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 20th 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.

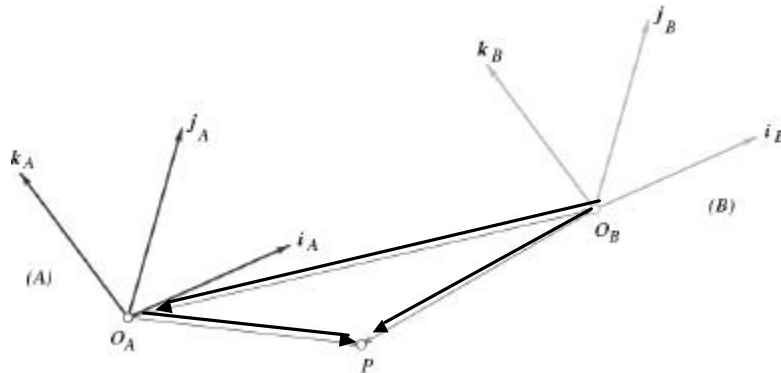
CS 534 – Ahmed Elgammal 23

Projective Geometry and Homogeneous coordinates

- If (x,y) are the rectangular coordinates of a point, P, and if (x_1, x_2, x_3) are any three real numbers such that:
 - $x_1/x_3 = x$
 - $x_2/x_3 = y$then (x_1, x_2, x_3) 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, (x_1, x_2, x_3) , the rectangular coordinates can be recovered.
- But (x,y) has an infinite number of homogeneous coordinate representations, because if (x_1, x_2, x_3) are homogeneous coordinates of (x,y) , then so are (kx_1, kx_2, kx_3) for any $k \neq 0$.

CS 534 – Ahmed Elgammal 24

Coordinate Changes: Pure Translations



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

CS 534 – Ahmed Elgammal 25

Coordinate Changes: Pure Translations

Notations:

Left Superscript : Coordinate Frame of Reference

Origin of coordinate frame A in
coordinate frame B

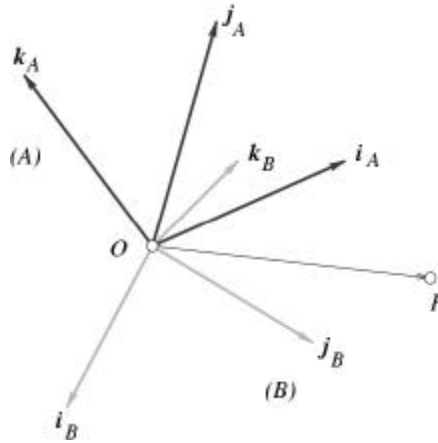
Point P in coordinate Frame A

Point P in coordinate Frame B

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

CS 534 – Ahmed Elgammal 26

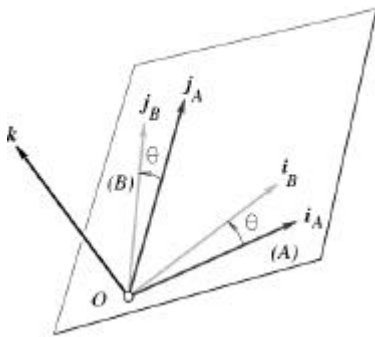
Coordinate Changes: Pure Rotations



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

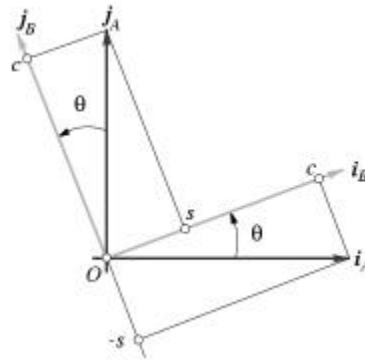
CS 534 – Ahmed Elgammal 27

Coordinate Changes: Rotations about the z Axis



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

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



CS 534 – Ahmed Elgammal 28

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

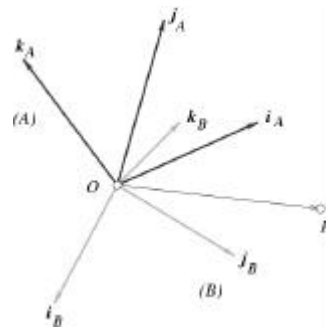
$${}^B_A R = \begin{bmatrix} \cos \mathbf{q} & \sin \mathbf{q} & 0 \\ -\sin \mathbf{q} & \cos \mathbf{q} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

CS 534 – Ahmed Elgammal 29

Coordinate Changes: Pure Rotations

Multiply both sides by

$$\begin{bmatrix} \mathbf{i}_B \\ \mathbf{j}_B \\ \mathbf{k}_B \end{bmatrix}$$



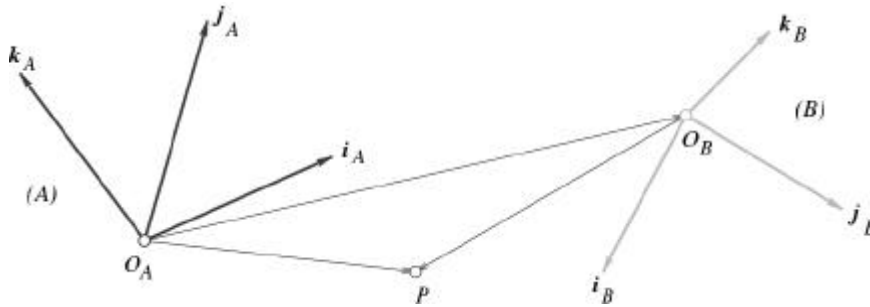
$$\overrightarrow{OP} = \begin{bmatrix} \mathbf{i}_A & \mathbf{j}_A & \mathbf{k}_A \end{bmatrix} \begin{bmatrix} {}^A x \\ {}^A y \\ {}^A z \end{bmatrix} = \begin{bmatrix} \mathbf{i}_B & \mathbf{j}_B & \mathbf{k}_B \end{bmatrix} \begin{bmatrix} {}^B x \\ {}^B y \\ {}^B z \end{bmatrix}$$

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

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

CS 534 – Ahmed Elgammal 30

Coordinate Changes: Rigid Transformations



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

CS 534 – Ahmed Elgammal 31

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 ?

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

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

Homogeneous Representation of Rigid Transformations

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

CS 534 – Ahmed Elgammal 32

Other transformations

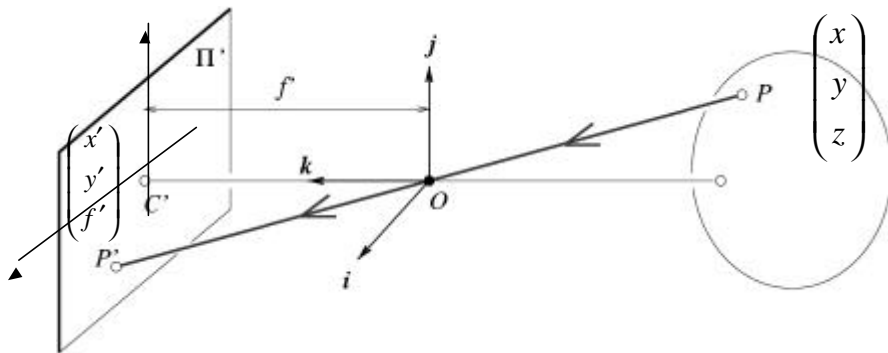
$$T = \begin{pmatrix} A & t \\ 0^T & 1 \end{pmatrix}$$

A is 3x3 rotation matrix :
Rigid transformation
 Lengths and angles 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

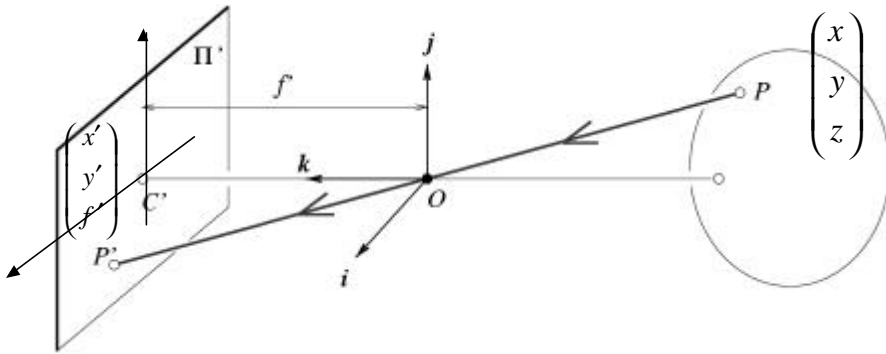
Pinhole Perspective Equation



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

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

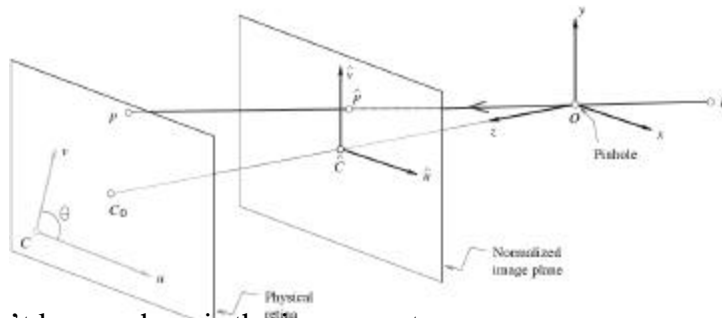
Pinhole Perspective Equation



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

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

CS 534 – Ahmed Elgammal 37



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

CS 534 – Ahmed Elgammal 38

- Pixels are rectangular with scale parameters k, l

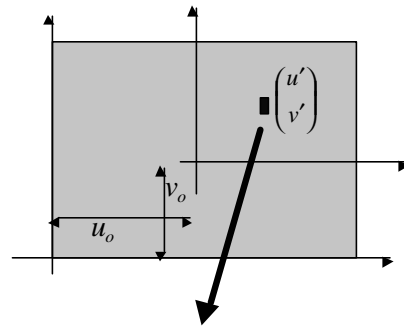
$$u' = kf \frac{x}{z} = \mathbf{a} \frac{x}{z}$$

$$v' = lf \frac{y}{z} = \mathbf{b} \frac{y}{z}$$

- Move coordinate system to the corner

$$u = \mathbf{a} \frac{x}{z} + u_o$$

$$v = \mathbf{b} \frac{y}{z} + v_o$$



$1/l$ meter/pixel

$1/k$ meter/pixel

CS 534 – Ahmed Elgammal 39

- Pixel grid may not be exactly orthogonal
- $\mathbf{q} @ 90$ but not exactly

$$u = \mathbf{a} \frac{x}{z} + u_o$$

$$v = \mathbf{b} \frac{y}{z} + v_o$$



$$u = \mathbf{a} \frac{x}{z} - \mathbf{a} \cot \mathbf{q} \frac{y}{z} + u_o$$

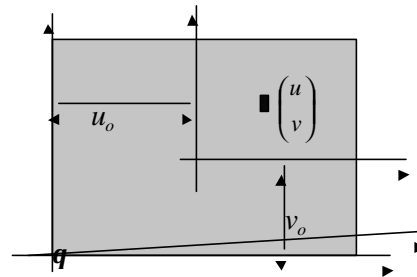
$$v = \frac{\mathbf{b}}{\sin \mathbf{q}} \frac{y}{z} + v_o$$



Approximation
 s is skew parameter

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

$$v = \mathbf{b} \frac{y}{z} + v_o$$



CS 534 – Ahmed Elgammal 40

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

$$\longrightarrow \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z} \begin{bmatrix} \mathbf{a} & 0 & u_o & 0 \\ 0 & \mathbf{b} & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \longrightarrow$$

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z} \begin{bmatrix} \mathbf{a} & -\mathbf{a} \cot \mathbf{q} & u_o & 0 \\ 0 & \frac{\mathbf{b}}{\sin \mathbf{q}} & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad \text{OR} \quad \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z} \begin{bmatrix} \mathbf{a} & s & u_o & 0 \\ 0 & \mathbf{b} & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

CS 534 – Ahmed Elgammal 41

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z} \begin{bmatrix} \mathbf{a} & -\mathbf{a} \cot \mathbf{q} & u_o & 0 \\ 0 & \frac{\mathbf{b}}{\sin \mathbf{q}} & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad \text{OR} \quad \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z} \begin{bmatrix} \mathbf{a} & s & u_o & 0 \\ 0 & \mathbf{b} & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

- Five intrinsic camera parameters:
 - Magnification \mathbf{a}, \mathbf{b} (in pixels)
 - Image center location u_o, v_o (in pixels)
 - Skew measured as \mathbf{q} or s

CS 534 – Ahmed Elgammal 42

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z} \begin{bmatrix} \mathbf{a} & -\mathbf{a} \cot \mathbf{q} & u_o & 0 \\ 0 & \frac{\mathbf{b}}{\sin \mathbf{q}} & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad \text{OR} \quad \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z} \begin{bmatrix} \mathbf{a} & s & u_o & 0 \\ 0 & \mathbf{b} & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

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

- Five intrinsic camera parameters:
 - Magnification \mathbf{a}, \mathbf{b} (in pixels)
 - Image center location u_o, v_o (in pixels)
 - Skew measured as \mathbf{q} or s

CS 534 – Ahmed Elgammal 43

The Intrinsic Parameters of a Camera

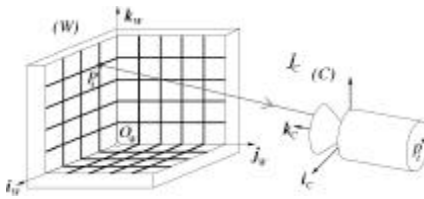
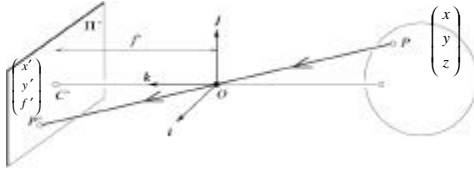
Calibration Matrix

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

The Perspective Projection Equation $\mathbf{p} = \frac{1}{z} \mathbf{M} \mathbf{P}$, where $\mathbf{M} \stackrel{\text{def}}{=} (\mathcal{K} \ \mathbf{0})$

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]



$$\begin{pmatrix} {}^C P \\ 1 \end{pmatrix} = \begin{pmatrix} {}^C R_{3 \times 3} & {}^C O_{W 3 \times 1} \\ 0^T_{1 \times 3} & 1 \end{pmatrix}_{4 \times 4} \begin{pmatrix} {}^W P \\ 1 \end{pmatrix}$$

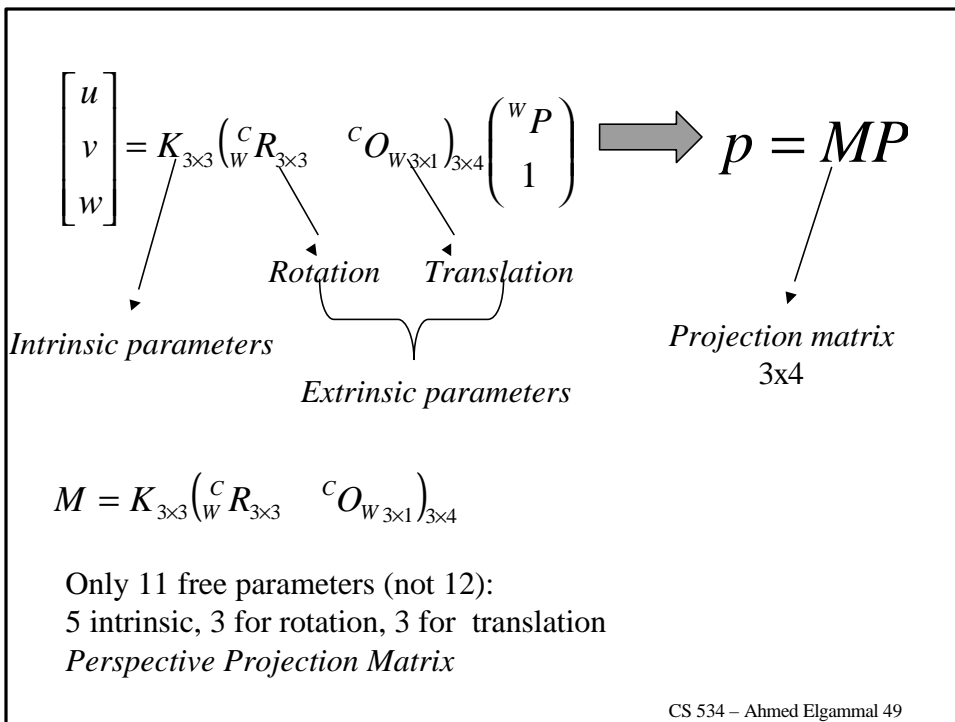
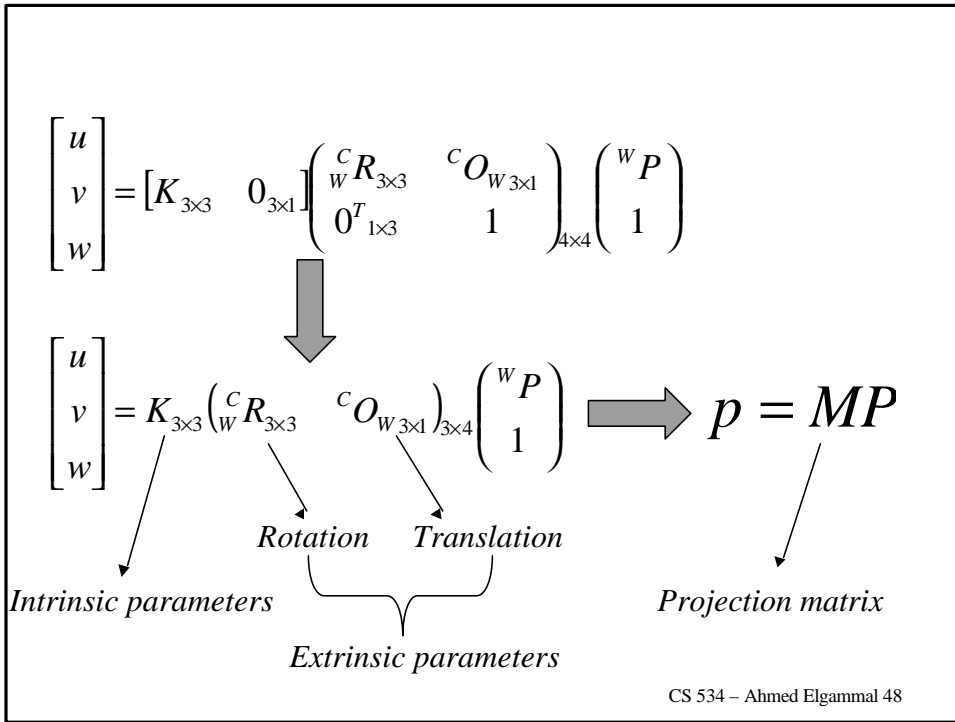
CS 534 – Ahmed Elgammal 46

$$\begin{pmatrix} {}^C P \\ 1 \end{pmatrix} = \begin{pmatrix} {}^C R_{3 \times 3} & {}^C O_{W 3 \times 1} \\ 0^T_{1 \times 3} & 1 \end{pmatrix}_{4 \times 4} \begin{pmatrix} {}^W P \\ 1 \end{pmatrix}$$

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} \mathbf{a} & -\mathbf{a} \cot \mathbf{q} & u_o \\ 0 & \frac{b}{\sin \mathbf{q}} & v_o \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} \mathbf{K}_{3 \times 3} & \mathbf{0}_{3 \times 1} \end{bmatrix} \begin{pmatrix} {}^C R_{3 \times 3} & {}^C O_{W 3 \times 1} \\ 0^T_{1 \times 3} & 1 \end{pmatrix}_{4 \times 4} \begin{pmatrix} {}^W P \\ 1 \end{pmatrix}$$

CS 534 – Ahmed Elgammal 47



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

$$p = MP = \begin{pmatrix} m_1^T \\ m_2^T \\ m_3^T \end{pmatrix} P \quad \longrightarrow \quad \begin{cases} u = \frac{m_1^T \cdot P}{m_3^T \cdot P} \\ v = \frac{m_2^T \cdot P}{m_3^T \cdot P} \end{cases}$$

Replacing M by IM doesn't change u or v

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

CS 534 – Ahmed Elgammal 50

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,

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

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

$$\mathcal{M} = \begin{pmatrix} m_1^T \\ m_2^T \\ m_3^T \end{pmatrix} \implies z = m_3 \cdot P, \quad \text{or} \quad \begin{cases} u = \frac{m_1 \cdot P}{m_3 \cdot P}, \\ v = \frac{m_2 \cdot P}{m_3 \cdot P}. \end{cases}$$

mmal 51

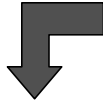
Explicit Form of the Projection Matrix

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

Note: If $\mathcal{M} = (\mathcal{A} \ \mathbf{b})$ 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 .

\mathcal{M} is only defined up to scale in this setting!!

CS 534 – Ahmed Elgammal 52

Theorem (Faugeras, 1993)

Let $\mathcal{M} = (\mathcal{A} \ \mathbf{b})$ be a 3×4 matrix and let \mathbf{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 $\text{Det}(\mathcal{A}) \neq 0$.
- A necessary and sufficient condition for \mathcal{M} to be a zero-skew perspective projection matrix is that $\text{Det}(\mathcal{A}) \neq 0$ and

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

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

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

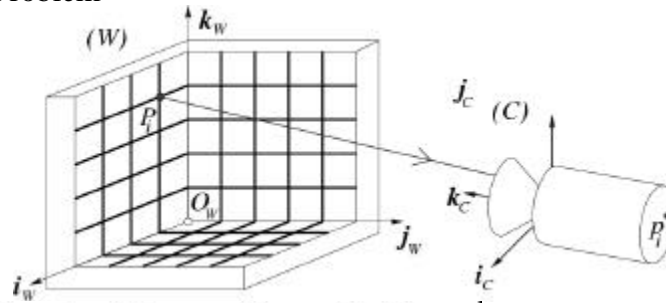
CS 534 – Ahmed Elgammal 53

Camera Calibration

- Find the intrinsic and extrinsic parameters of a camera
 - Extrinsic parameters: the camera's location and orientation in the world.
 - Intrinsic parameters: the relationships between pixel coordinates and camera coordinates.
- VERY large literature on the subject
- Work of Roger Tsai influential
- Good calibration is important when we need to:
 - Reconstruct a world model.
 - Interact with the world: Robot, hand-eye coordination
- Basic idea:
 - Given a set of world points P_i and their image coordinates (u_i, v_i)
 - find the projection matrix M
 - and then find intrinsic and extrinsic parameters.

CS 534 – Ahmed Elgammal 54

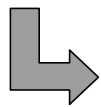
Calibration Problem



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

Find \mathbf{i} and \mathbf{e} such that

$$\begin{cases} u_i = \frac{\mathbf{m}_1(\mathbf{i}, \mathbf{e}) \cdot \mathbf{P}_i}{\mathbf{m}_3(\mathbf{i}, \mathbf{e}) \cdot \mathbf{P}_i} \\ v_i = \frac{\mathbf{m}_2(\mathbf{i}, \mathbf{e}) \cdot \mathbf{P}_i}{\mathbf{m}_3(\mathbf{i}, \mathbf{e}) \cdot \mathbf{P}_i} \end{cases} \quad \text{for } i = 1, \dots, n$$



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

CS 534 – Ahmed Elgammal 55

Camera Calibration and least-squares

- Camera Calibration is least-squares parameter estimation problem.
- Estimate the intrinsic and extrinsic parameters that minimize the mean-squared deviation between predicted and observed image features.
- Least-squares parameter estimation is a fundamental technique that is used extensively in computer vision.
- You can formulate many problems as error minimization between observed and predicted values.
 - Linear Least-Squares
 - Nonlinear Least-Squares

CS 534 – Ahmed Elgammal 56

Linear Least-Squares

- Linear system of equations

$$\begin{aligned}a_{11}x_1 + a_{12}x_2 + \dots + a_{1q}x_q &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2q}x_q &= b_2 \\ \dots & \\ a_{p1}x_1 + a_{p2}x_2 + \dots + a_{pq}x_q &= b_p\end{aligned}$$

$$\Leftrightarrow Ax = b$$

$$A_{p \times q} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1q} \\ a_{21} & a_{22} & \dots & a_{2q} \\ \dots & \dots & \dots & \dots \\ a_{p1} & a_{p2} & \dots & a_{pq} \end{pmatrix} \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_q \end{pmatrix} \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{pmatrix}$$

CS 534 – Ahmed Elgammal 57

How do you solve overconstrained linear equations ??

- Define $E = \|e\|^2 = e \cdot e$ with

$$e = Ax - b = \begin{bmatrix} | & | & | & | \\ c_1 & c_2 & \dots & c_n \\ | & | & | & | \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} - b$$

$$= x_1 c_1 + x_2 c_2 + \dots + x_n c_n - b$$

$$\frac{\partial e}{\partial x_i} = c_i$$

- At a minimum,

$$\frac{\partial E}{\partial x_i} = \frac{\partial e}{\partial x_i} \cdot e + e \cdot \frac{\partial e}{\partial x_i} = 2 \frac{\partial e}{\partial x_i} \cdot e$$

$$= 2 \frac{\partial}{\partial x_i} (x_1 c_1 + \dots + x_n c_n - b) \cdot e = 2 c_i \cdot e$$

$$= 2 c_i^T (Ax - b) = 0$$

- or

$$0 = \begin{bmatrix} c_1^T \\ \vdots \\ c_n^T \end{bmatrix} (Ax - b) = A^T (Ax - b) \Rightarrow A^T Ax = A^T b$$

where $x = A^\dagger b$ and $A^\dagger = (A^T A)^{-1} A^T$ is the *pseudoinverse* of A !

CS 534 – Ahmed Elgammal 60

Singular Value Decomposition

- Chief tool for dealing with m by n systems and singular systems.
- SVD: If A is a real m by n matrix then there exist orthogonal matrices U ($m \times m$) and V ($n \times n$) such that

$$U^T A V = S = \text{diag}(s_1, s_2, \dots, s_p) \quad p = \min\{m, n\}$$

$$U^T A V = S \quad A = U S V^T$$

- Singular values:** Non negative square roots of the eigenvalues of $A^T A$. Denoted $s_i, i=1, \dots, p$
- $A^T A$ is symmetric \Rightarrow eigenvalues and singular values are real.
- Singular values arranged in decreasing order.

$$A^T A = (U \Sigma V^T)^T (U \Sigma V^T) = V \Sigma^T U^T U \Sigma V^T = V \Sigma^T \Sigma V^T = V \Sigma^2 V^{-1}$$

$$(A^T A) V = V \Sigma^2$$

$$(A^T A) v = v \lambda$$

$$\begin{array}{c} \boxed{A} \\ m \times n \end{array} = \begin{array}{c} \boxed{U} \\ m \times m \end{array} \begin{array}{c} \boxed{\Sigma} \\ m \times n \end{array} \begin{array}{c} \boxed{V^T} \\ n \times n \end{array}$$

CS 534 – Ahmed Elgammal 61

- For over-constrained system $Ax=b$:
- Use SVD to decompose $A=USV^T$

$$Ax = b$$

$$U \Sigma V^T x = b$$

$$(V \Sigma^{-1} U^T) U \Sigma V^T x = (V \Sigma^{-1} U^T) b$$

$$x = V \Sigma^{-1} U^T b$$

$$\Sigma^{-1} = \text{diag}(1/s_1, 1/s_2, \dots, 1/s_n)$$

CS 534 – Ahmed Elgammal 62

Homogeneous Linear Systems

$$\begin{array}{|c|} \hline A \\ \hline \end{array} \begin{array}{|c|} \hline x \\ \hline \end{array} = \begin{array}{|c|} \hline 0 \\ \hline \end{array}$$

Square system:

- unique solution: 0
- unless $\text{Det}(A)=0$

Rectangular system ??

$$\begin{array}{|c|} \hline \\ \hline A \\ \hline \\ \hline \end{array} \begin{array}{|c|} \hline x \\ \hline \end{array} = \begin{array}{|c|} \hline \\ \hline 0 \\ \hline \\ \hline \end{array}$$



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

x is the unit singular vector of A corresponding to the smallest singular value (the last column of V , where $A = U S V^T$ is the SVD of A)

CS 534 – Ahmed Elgammal 63

Linear Camera Calibration



- Calibration target: 2 planes at right angle with checkerboard patterns (Tsai grid)
- We know positions of pattern corners only with respect to a coordinate system of the target
- We position camera in front of target and find images of corners
- We obtain equations that describe imaging and contain internal parameters of camera
- As a side benefit, we find position and orientation of camera with respect to target (camera *pose*)

CS 534 – Ahmed Elgammal 66

Linear Camera Calibration

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

$$\Rightarrow \begin{pmatrix} u_i \\ v_i \end{pmatrix} = \begin{pmatrix} m_1 \cdot P_i \\ m_3 \cdot P_i \\ m_2 \cdot P_i \\ m_3 \cdot P_i \end{pmatrix} \iff \begin{pmatrix} m_1 - u_i m_3 \\ m_2 - v_i m_3 \end{pmatrix} P_i = 0$$

$$\begin{matrix} m_1 X_i + m_2 Y_i + m_3 Z_i + m_4 & -m_3 u_i X_i - m_3 v_i Y_i - m_3 u_i Z_i - m_3 u_i = 0 \\ m_2 X_i + m_2 Y_i + m_2 Z_i + m_2 & -m_3 v_i X_i - m_3 v_i Y_i - m_3 v_i Z_i - m_3 v_i = 0 \end{matrix}$$

$$\Rightarrow \mathcal{P}m = 0 \text{ with } \mathcal{P} \stackrel{\text{def}}{=} \begin{pmatrix} P_1^T & 0^T & -u_1 P_1^T \\ 0^T & P_1^T & -v_1 P_1^T \\ \dots & \dots & \dots \\ P_n^T & 0^T & -u_n P_n^T \\ 0^T & P_n^T & -v_n P_n^T \end{pmatrix}_{n \times 12} \text{ and } m \stackrel{\text{def}}{=} \begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix}_{12 \times 1} = 0$$

nal 67

$$\mathcal{P}\mathbf{m} = 0 \text{ with } \mathcal{P} \stackrel{\text{def}}{=} \begin{pmatrix} \mathbf{P}_1^T & \mathbf{0}^T & -u_1\mathbf{P}_1^T \\ \mathbf{0}^T & \mathbf{P}_1^T & -v_1\mathbf{P}_1^T \\ \dots & \dots & \dots \\ \mathbf{P}_n^T & \mathbf{0}^T & -u_n\mathbf{P}_n^T \\ \mathbf{0}^T & \mathbf{P}_n^T & -v_n\mathbf{P}_n^T \end{pmatrix}_{n \times 12} \text{ and } \mathbf{m} \stackrel{\text{def}}{=} \begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix}_{12 \times 1} = 0$$

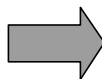
- When $n \geq 6 \Rightarrow$ homogenous linear least-squares can be used to compute unite vector $\mathbf{m} \Rightarrow M$
- Degenerate case: All points lie on the same plane.
Points should not be on the same plane
- When possible, have at least 5 times as many equations as unknowns (28 points)

CS 534 – Ahmed Elgammal 68

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

This is a decomposition problem, not an estimation problem.

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



- Intrinsic parameters
- Extrinsic parameters

CS 534 – Ahmed Elgammal 69

Another approach to decomposition

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = K_{3 \times 3} \begin{pmatrix} {}^C R_{3 \times 3} & {}^C O_{W 3 \times 1} \end{pmatrix}_{3 \times 4} \begin{pmatrix} {}^W P \\ 1 \end{pmatrix}$$

The translation vector can be written as:

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = K_{3 \times 3} \begin{pmatrix} {}^C R_{3 \times 3} & -{}^C R_{3 \times 3} {}^W O_{C 3 \times 1} \end{pmatrix}_{3 \times 4} \begin{pmatrix} {}^W P \\ 1 \end{pmatrix}$$

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = K_{3 \times 3} {}^C R_{3 \times 3} \begin{pmatrix} I_{3 \times 3} & -{}^W O_{C 3 \times 1} \end{pmatrix}_{3 \times 4} \begin{pmatrix} {}^W P \\ 1 \end{pmatrix}$$

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

$${}^W O_W = {}^W_C R {}^C O_W + {}^W O_C$$

$$0 = {}^W_C R {}^C O_W + {}^W O_C$$

$${}^W_C R {}^C O_W = -{}^W O_C$$

$${}^C O_W = -{}^C_W R {}^W O_C$$

CS 534 – Ahmed Elgammal 70

Getting Camera Translation

$$M_{3 \times 4} = K_{3 \times 3} {}^C R_{3 \times 3} \begin{pmatrix} I_{3 \times 3} & -{}^W O_{C 3 \times 1} \end{pmatrix}_{3 \times 4}$$

Upper triangle matrix

Orthogonal matrix

Translation Vector:

Where in the world is the camera center

Note that:

$$\begin{pmatrix} 1 & 0 & 0 & -X_c \\ 0 & 1 & 0 & -Y_c \\ 0 & 0 & 1 & -Z_c \end{pmatrix} \begin{pmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{pmatrix} = 0$$

If we solve the system $M \cdot t = 0$ we can get the translation vector! (t is the null vector for M)

Find null vector t of M using SVD

- t is the unit singular vector of M corresponding to the smallest singular value (the last column of V , where $M = U S V^T$ is the SVD of M)

CS 534 – Ahmed Elgammal 71

Getting Camera Orientation

$$M_{3 \times 4} = K_{3 \times 3}^c R_{3 \times 3} (I_{3 \times 3} - {}^w O_{C_{3 \times 1}})_{3 \times 4}$$

$K_{3 \times 3}^c$ → Upper triangle matrix
 $R_{3 \times 3}$ → Orthogonal matrix
 $(I_{3 \times 3} - {}^w O_{C_{3 \times 1}})_{3 \times 4}$ → Translation Vector:
 Where in the world is the camera center

- Left 3x3 submatrix M' of M is of form $M' = KR$
- K is an upper triangular matrix
- R is an orthogonal matrix
- Any non-singular square matrix M' can be decomposed into the product of an upper-triangular matrix K and an orthogonal matrix R using the RQ factorization
- Similar to QR factorization but order of 2 matrices is reversed $A=Q.R$

CS 534 – Ahmed Elgammal 72

Rotation matrices

- 3-D rotation is result of three consecutive rotations around the coordinate axes $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$.
 - the angles of the rotations are the parameters of the rotation

$$R_1(\mathbf{a}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \mathbf{a} & -\sin \mathbf{a} \\ 0 & \sin \mathbf{a} & \cos \mathbf{a} \end{bmatrix} \quad R_2(\mathbf{b}) = \begin{bmatrix} \cos \mathbf{b} & 0 & \sin \mathbf{b} \\ 0 & 1 & 0 \\ -\sin \mathbf{b} & 0 & \cos \mathbf{b} \end{bmatrix} \quad R_3(\mathbf{g}) = \begin{bmatrix} \cos \mathbf{g} & -\sin \mathbf{g} & 0 \\ \sin \mathbf{g} & \cos \mathbf{g} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R = R_1 R_2 R_3 = \begin{bmatrix} \cos \mathbf{b} \cos \mathbf{g} & -\cos \mathbf{b} \sin \mathbf{g} & \sin \mathbf{b} \\ \sin \mathbf{a} \sin \mathbf{b} \cos \mathbf{g} + \cos \mathbf{a} \sin \mathbf{g} & -\sin \mathbf{a} \sin \mathbf{b} \sin \mathbf{g} + \cos \mathbf{a} \cos \mathbf{g} & -\sin \mathbf{a} \cos \mathbf{b} \\ -\cos \mathbf{a} \sin \mathbf{b} \cos \mathbf{g} - \sin \mathbf{a} \sin \mathbf{g} & \cos \mathbf{a} \sin \mathbf{b} \sin \mathbf{g} + \sin \mathbf{a} \cos \mathbf{g} & \cos \mathbf{a} \cos \mathbf{b} \end{bmatrix}$$

- The order in which we perform the multiplications matters!
 - Six ways to represent a rotation matrix

CS 534 – Ahmed Elgammal 73

Rotation matrices

$$R = R_1 R_2 R_3 = \begin{bmatrix} \cos b \cos g & -\cos b \sin g & \sin b \\ \sin a \sin b \cos g + \cos a \sin g & -\sin a \sin b \sin g + \cos a \cos g & -\sin a \cos b \\ -\cos a \sin b \cos g - \sin a \sin g & \cos a \sin b \sin g + \sin a \cos g & \cos a \cos g \end{bmatrix}$$

- The order in which we perform the multiplications matters!
 - Six ways to represent a rotation matrix
- It is easy to recover the rotation angles (α, β, γ) from R
 - $R_{1,3}$ gives β
 - Then γ can be recovered from $R_{1,1}$, and α from $R_{2,3}$
- R is orthonormal $RR^T = R^T R = I$
- So, even though there are 9 entries in R , there are still only the 3 parameters
- This makes estimating R difficult in practice.

CS 534 – Ahmed Elgammal 74

Sources

- Forsyth and Ponce, Computer Vision a Modern approach: 1.1, 2.1, 2.2, 3.1, 3.2
- Fougeras, Three-dimensional Computer Vision
- Numerical Recipes
- Slides by J. Ponce @ UIUC
- Slides by L. S. Davis and R. Duraiswami @ UMD

CS 534 – Ahmed Elgammal 75