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 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° of angle.
  – Appearance
    • a point on which three lines segments
      are concurrent. Total angle is 360°

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
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

\[ ax + by - d = 0 \iff \delta^T \cdot p = 0 \]

\[
\delta = \begin{pmatrix} a \\ b \\ -d \end{pmatrix} \quad p = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}
\]

Homogenous Coordinates

• 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: \( p_1 \times p_2 \)
• Three points on the same line: \( \text{det}[p_1 \ p_2 \ p_3] = 0 \)
• Three line intersecting in a point: \( \text{det}[l_1 \ l_2 \ l_3] = 0 \)
Table from S. Birchfield “An Introduction to Projective Geometry”

<table>
<thead>
<tr>
<th>point</th>
<th>$p = (X, Y, W)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>incidence</td>
<td>$p^T n = 0$</td>
</tr>
<tr>
<td>collinearity</td>
<td>$</td>
</tr>
<tr>
<td>join of 2 points</td>
<td>$n = p_1 \times p_2$</td>
</tr>
<tr>
<td>ideal points</td>
<td>$(X, Y, 0)$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>line</th>
<th>$n = (a, b, c)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>incidence</td>
<td>$p^T n = 0$</td>
</tr>
<tr>
<td>concurrence</td>
<td>$</td>
</tr>
<tr>
<td>intersection of 2 lines</td>
<td>$p = u_1 \times u_2$</td>
</tr>
<tr>
<td>ideal line</td>
<td>$(0, 0, c)$</td>
</tr>
</tbody>
</table>

Homogenous Coordinates
Planes

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

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

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

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:

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

Homogenous Coordinates

• Spheres

\[
x^2 + y^2 + z^2 = R^2
\]

\[
(x \ y \ z \ 1) \cdot \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & R^2
\end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = 0
\]
Euclidean Geometry
Coordinate Systems

\[ \begin{align*}
x &= \overrightarrow{OP}.i \\
y &= \overrightarrow{OP}.j \quad \iff \quad \overrightarrow{OP} = xi + yj + zk \quad \iff \quad P = \begin{bmatrix} x \\ y \\ z \end{bmatrix}
z &= \overrightarrow{OP}.k
\end{align*} \]

Coordinate System Change
Pure Translation

\[ \overrightarrow{OB}P = \overrightarrow{OB}O_A + \overrightarrow{O_A}P \quad , \quad \overrightarrow{BP} = \overrightarrow{AP} + \overrightarrow{BO_A} \]
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

\[ \overrightarrow{O_B P} = \overrightarrow{O_B O_A} + \overrightarrow{O_A P} \]
\[ \overrightarrow{B P} = \overrightarrow{A P} + \overrightarrow{B O_A} \]

Coordinate System Change

Pure Rotation

Multiply both sides by

\[ \overrightarrow{OP} = \begin{bmatrix} i_A & j_A & k_A \end{bmatrix} \begin{bmatrix} A_x \\ A_y \\ A_z \end{bmatrix} = \begin{bmatrix} i_B & j_B & k_B \end{bmatrix} \begin{bmatrix} B_x \\ B_y \\ B_z \end{bmatrix} \]

\[ \Rightarrow \overrightarrow{BP} = \overrightarrow{AP} = \overrightarrow{A R_A P} \]

\[ i_R = \begin{bmatrix} i_A & j_A & k_A \\ i_A & j_A & k_A \\ i_A & j_A & k_A \end{bmatrix} \]
Coordinate Changes: Pure Rotations

Rotation matrix describing the frame A in coordinate frame B

\[
^{B}_{A} R = \begin{bmatrix}
{i_A \cdot i_B} & {j_A \cdot i_B} & {k_A \cdot i_B} \\
{i_A \cdot j_B} & {j_A \cdot j_B} & {k_A \cdot j_B} \\
{i_A \cdot k_B} & {j_A \cdot k_B} & {k_A \cdot k_B}
\end{bmatrix}
= \begin{bmatrix}
{^A_0 i_B} \\
{^A_0 j_B} \\
{^A_0 k_B}
\end{bmatrix}
= \begin{bmatrix}
^{B}_0 A_i \\
^{B}_0 A_j \\
^{B}_0 A_k
\end{bmatrix}

Coordinate Changes: Rotations about the z Axis

\[
^{B}_{A} R = \begin{bmatrix}
{^A_0 i_B} & {^A_0 j_B} & {^A_0 k_B} \\
{^A_0 i_B} & {^A_0 j_B} & {^A_0 k_B} \\
{i_A \cdot k_B} & {j_A \cdot k_B} & {k_A \cdot k_B}
\end{bmatrix}
\]

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

\[ \underleftrightarrow{b} A R = \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \]

- 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

\[ B P = \underleftrightarrow{b} A R \underleftrightarrow{a} P + \underleftrightarrow{b} O_A \]
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 \(AB\) ?

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

\[
\begin{bmatrix}
B_P \\
1
\end{bmatrix} = \begin{bmatrix}
B_R A_P + B_O_A \\
A_T^T
\end{bmatrix} = \begin{bmatrix}
B_R & B_O_A \\
A_T^T & 1
\end{bmatrix} \begin{bmatrix}
A_P \\
1
\end{bmatrix}
\]

Homogeneous Representation of Rigid Transformations

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
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.

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

\[
\begin{align*}
\begin{cases}
  x' = f' \frac{x}{z} \\
y' = f' \frac{y}{z}
\end{cases}
\end{align*}
\]

NOTE: \( z \) is always negative.

Affine projection models: Weak perspective projection

\[
\begin{align*}
\begin{cases}
  x' = f' \frac{x}{z} \\
y' = f' \frac{y}{z}
\end{cases}
\end{align*}
\]

\[
\begin{align*}
x' = -mx \\
y' = -my
\end{align*}
\]

where \( m = -\frac{f'}{z_0} \) 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)
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?
• 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 $k, l$
  \[ u' = k \frac{f x}{z} = \alpha \frac{x}{z} \]
  \[ v' = l \frac{f y}{z} = \beta \frac{y}{z} \]

• 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
  \[ \theta \approx 90 \text{ but not exactly} \]

\[ u = \alpha \frac{x}{z} + u_o \]
\[ v = \beta \frac{y}{z} + v_o \]
\[ u = \alpha \frac{x}{z} + s \frac{y}{z} + u_o \]
\[ v = \beta \frac{y}{z} + v_o \]
\[ v = \frac{\beta}{\sin \theta} \frac{y}{z} + v_o \]

$\alpha$ is skew parameter
• Five intrinsic camera parameters:
  – Magnification \( \alpha, \beta \) \textit{(in pixels)}
  – Image center location \( u_o, v_o \) \textit{(in pixels)}
  – Skew measured as \( \theta \) or \( s \)
• Five intrinsic camera parameters:
  – Magnification $\alpha, \beta$ \textit{(in pixels)}
  – Image center location $u_o, v_o$ \textit{(in pixels)}
  – Skew measured as $\theta$ or $s$

\[
\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z} MP, \quad \text{where} \quad M \overset{\text{def}}{=} \begin{bmatrix} \alpha & 0 & u_o & 0 \\ 0 & \frac{\rho}{\sin \theta} & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad \text{OR} \quad \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z} \begin{bmatrix} \alpha & s & u_o & 0 \\ 0 & \beta & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}
\]

\[p = \frac{1}{z} M P,\]
The Intrinsic Parameters of a Camera

Calibration Matrix

\[ p = K \hat{p}, \quad \text{where} \quad p = \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \quad \text{and} \quad K \overset{\text{def}}{=} \begin{pmatrix} \alpha & -\alpha \cot \theta & u_0 \\ 0 & \beta \sin \theta & v_0 \\ 0 & 0 & 1 \end{pmatrix} \]

The Perspective Projection Equation

\[ p = \frac{1}{z} M P, \quad \text{where} \quad M \overset{\text{def}}{=} \begin{pmatrix} K & 0 \end{pmatrix} \]

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]

\[
\begin{pmatrix} ^{c}P \\ 1 \end{pmatrix} = \begin{pmatrix} ^{c}R_{3 \times 3} & ^{c}O_{3 \times 1} \\ 0_{1 \times 3} \end{pmatrix} \begin{pmatrix} ^{w}P \\ 1 \end{pmatrix}
\]
\[
\begin{pmatrix}
  \frac{C}{P} \\
  1
\end{pmatrix} =
\begin{pmatrix}
  C_W R_{3\times3} & C_O W_{3\times1} \\
  0^T_{1\times3} & 1
\end{pmatrix}_{4\times4}
\begin{pmatrix}
  \frac{W}{P} \\
  1
\end{pmatrix}
\]

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

\[
\begin{bmatrix}
  u' \\
  v' \\
  w'
\end{bmatrix} =
\begin{bmatrix}
  K_{3\times3} & 0_{3\times1} \\
  0_{3\times3} & 0^T_{1\times3}
\end{bmatrix}
\begin{pmatrix}
  C_W R_{3\times3} & C_O W_{3\times1} \\
  0^T_{1\times3} & 1
\end{pmatrix}_{4\times4}
\begin{pmatrix}
  \frac{W}{P} \\
  1
\end{pmatrix}
\]

\[
\begin{bmatrix}
  u' \\
  v' \\
  w'
\end{bmatrix} =
K_{3\times3}
\begin{pmatrix}
  C_W R_{3\times3} & C_O W_{3\times1}
\end{pmatrix}_{3\times4}
\begin{pmatrix}
  \frac{W}{P} \\
  1
\end{pmatrix} \Rightarrow p = MP
\]

**Intrinsic parameters**

**Extrinsic parameters**

**Projection matrix**
Intrinsic parameters
Rotation
Translation
Extrinsic parameters
Projection matrix
3x4

\[
\begin{bmatrix}
  u' \\
  v' \\
  w'
\end{bmatrix} = K_{3\times3}
\begin{bmatrix}
  c R_{3\times3} & c O_{w_{3\times1}}
\end{bmatrix}_{3\times4}
\begin{bmatrix}
  w P \\
  1
\end{bmatrix} 
\rightarrow p = MP
\]

\[
M = K_{3\times3}
\begin{bmatrix}
  c R_{3\times3} & c O_{w_{3\times1}}
\end{bmatrix}_{3\times4}
\]

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

Perspective Projection Matrix

Replacing \( M \) by \( \lambda M \) in

\[
\begin{align*}
  u &= \frac{m_1 \cdot P}{m_3 \cdot P} \\
  v &= \frac{m_2 \cdot P}{m_3 \cdot P}
\end{align*}
\]

does not change \( u \) and \( v \).

\( M \) is only defined up to scale in this setting!!
Explicit Form of the Projection Matrix

\[ \mathcal{M} = \begin{pmatrix}
\alpha r_1^T - \alpha \cot \theta r_2^T + u_0 r_3^T \\
\beta \frac{r_2^T}{\sin \theta} + v_0 r_3^T \\
r_3^T
\end{pmatrix} \begin{pmatrix}
\alpha t_x - \alpha \cot \theta t_y + u_0 t_2 \\
\beta \frac{t_y}{\sin \theta} + v_0 t_3 \\
t_3
\end{pmatrix} \]

Note: If \( \mathcal{M} = (A \ b) \) then \( |a_3| = 1 \).

\[ p = MP = \begin{pmatrix}
m_1^T \\
m_2^T \\
m_3^T
\end{pmatrix} P \]

\[ u = \frac{m_1^T \cdot P}{m_3^T \cdot P} \]

\[ v = \frac{m_2^T \cdot P}{m_3^T \cdot P} \]

Replacing \( M \) by \( \lambda M \) doesn’t change \( u \) or \( v \)

\[ M \text{ is only defined up to scale in this setting!!} \]

Theorem (Faugeras, 1993)

Let \( \mathcal{M} = (A \ b) \) be a \( 3 \times 4 \) matrix and let \( a_i^T \ (i = 1, 2, 3) \) denote the rows of the matrix \( A \) 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 \( \det(A) \neq 0 \).

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

\[ (a_1 \times a_3) \cdot (a_2 \times a_3) = 0. \]

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

\[ \begin{cases}
(a_1 \times a_3) \cdot (a_2 \times a_3) = 0, \\
(a_1 \times a_3) \cdot (a_1 \times a_3) = (a_2 \times a_3) \cdot (a_2 \times a_3). 
\end{cases} \]
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