Image Formation and Cameras (Part 2)

Introduction to Computer Vision
CSE 152
Lecture 5

Announcements

• Homework 1 is due Apr 24, 11:59 PM
• Wait list
• Reading:
  – Chapter 2 Image formation

The equation of projection

Cartesian coordinates:

\[ (x, y, z) \rightarrow \left( \frac{x}{f}, \frac{y}{f}, \frac{z}{f} \right) \]

Homogenous Coordinates and Camera matrix

\[
\begin{bmatrix}
  1 & 0 & 0 & 0 \\
  0 & 1 & 0 & 0 \\
  0 & 0 & \frac{1}{f} & 0 \\
  0 & 0 & 0 & 1
\end{bmatrix}
\]

What if camera coordinate system differs from object coordinate system

Euclidean Coordinate Systems

\[
\begin{align*}
  x &= \overrightarrow{OP}.i \\
  y &= \overrightarrow{OP}.j \\
  z &= \overrightarrow{OP}.k
\end{align*}
\]

\[ \overrightarrow{OP} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \]

\[ \mathbf{P} = 
\begin{bmatrix}
  x \\
  y \\
  z \\
  1
\end{bmatrix}
\]

Coordinate Changes: Pure Translations

\[
\begin{align*}
  \overrightarrow{O_bP} &= \overrightarrow{O_aP} + \overrightarrow{O_aO_b} \\
  \overrightarrow{O_bP} &= \overrightarrow{O_aP} + \mathbf{t}
\end{align*}
\]

Translation from coordinate frame A to coordinate frame B
Rotation Matrix

Dot Products between all pairs of coordinate axis of both systems

\[ R = \begin{bmatrix} i_A & i_B & k_A & k_B \\ i_A & i_B & k_A & k_B \\ j_A & j_B & k_A & k_B \end{bmatrix} = \begin{bmatrix} i_A & i_B & k_A & k_B \end{bmatrix} \begin{bmatrix} i_A & i_B & k_A & k_B \end{bmatrix} \]

Coordinate Changes: Pure Rotations

Rotation from coordinate frame A to coordinate frame B

\[ \overrightarrow{OP} = \begin{bmatrix} x_A \\ y_A \\ z_A \end{bmatrix} = \begin{bmatrix} i_A & j_A & k_A \end{bmatrix} \begin{bmatrix} x_B \\ y_B \\ z_B \end{bmatrix} \]

\[ \Rightarrow P_B = \begin{bmatrix} i_A & j_A & k_A \end{bmatrix} P_A = R P_A \]

Coordinate Changes: Euclidean Transformations

Euclidean transformation from coordinate frame A to coordinate frame B

\[ P_B = R P_A + t \]

3D Rotation Matrices

- \( R^T = R^{-1} \)
- \( R^T R = R R^T = I \)
- \( \det(R) = 1 \)
- \( R_{ij} \in [-1, +1] \)
- Rows (or columns) of \( R \) form a right-handed orthonormal coordinate system
- Even though a rotation matrix is 3x3 with nine numbers, it only has three degrees of freedom, it can be parameterized with three numbers. There are many parameterizations

Rotation: Homogenous Coordinates

- About z axis

\[ \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \]

- About x axis:

\[ \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \]

- About y axis:

\[ \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \]
Composition of Rotations

\[
(\mathbf{R}_j \cdot \mathbf{R}_k) \cdot \mathbf{R}_i = \mathbf{R}_i \cdot (\mathbf{R}_j \cdot \mathbf{R}_k)
\]

Roll-Pitch-Yaw

\[
R = \text{rot}(\mathbf{j}, \alpha) \cdot \text{rot}(\mathbf{j}, \beta) \cdot \text{rot}(\mathbf{k}, \phi)
\]

Euclidean Transformations, Homogeneous Coordinates

\[
\begin{bmatrix}
P_x \\
P_y \\
P_z \\
1
\end{bmatrix} = \begin{bmatrix}
R & t \\
0 & 1
\end{bmatrix}
\begin{bmatrix}
P_x' \\
P_y' \\
P_z' \\
1
\end{bmatrix}
\]

where

\[
c = \cos \theta \quad \text{and} \quad s = \sin \theta
\]

What if camera coordinate system differs from object coordinate system

\[
E_{cw} = \begin{bmatrix}
R & t \\
0^T & 1
\end{bmatrix}
\]

Intrinsic parameters

- 3x3 homogenous matrix
- Focal length:
- Principal Point: C'
- Units (e.g. pixels)
- Orientation and position of image coordinate system
- Pixel Aspect ratio
Camera parameters

- Extrinsic Parameters: Since camera may not be at the origin, there is a rigid transformation between the world coordinates and the camera coordinates.
- Intrinsic parameters: Since scene units (e.g., cm) differ image units (e.g., pixels) and coordinate system may not be centered in image, we capture that with a 3x3 transformation comprised of focal length, principal point, pixel aspect ratio, angle between axes, etc.

\[
\begin{bmatrix}
U \\
V \\
W
\end{bmatrix}
= \begin{bmatrix}
1 & 0 & 0 & X \\
0 & 1 & 0 & Y \\
0 & 0 & 1 & Z
\end{bmatrix}
\begin{bmatrix}
E \\
F \\
G
\end{bmatrix}
\begin{bmatrix}
X \\
Y \\
Z
\end{bmatrix}
\]

Extrinsic parameters: 4 x 4

Intrinsic parameters: 3 x 3

Camera Calibration

Given \( n \) points \( P_1, \ldots, P_n \) with known positions and their images \( p_1, \ldots, p_n \), estimate intrinsic and extrinsic camera parameters.

- See Text book for how to do it.
- Camera Calibration Toolbox for Matlab (Bouguet)
  - http://www.vision.caltech.edu/bouguetj/calib_doc/

Beyond the pinhole Camera

Getting more light – Bigger Aperture

The reason for lenses

We need light, but big pinholes cause blur.

Thin Lens

- Rotationally symmetric about optical axis.
- Spherical interfaces.
Thin Lens: Center

- All rays that enter lens along line pointing at \( O \) emerge in same direction.

Thin Lens: Focus

Parallel lines pass through the focus, \( F \).

Thin Lens: Image of Point

- All rays passing through lens and starting at \( P \) converge upon \( P' \)
- So light gather capability of lens is given the area of the lens and all the rays focus on \( P' \) instead of become blurred like a pinhole

\[
\frac{1}{z'} + \frac{1}{z} = \frac{1}{f}
\]

Relation between depth of Point (\( Z \)) and the depth where it focuses (\( Z' \))

Thin Lens: Image Plane

A price: Whereas the image of \( P \) is in focus, the image of \( Q \) isn’t.

Thin Lens: Aperture

- Smaller Aperture -> Less Blur
- Pinhole -> No Blur
Field of view is a function of \( f \) and size of image plane.

Deviations from this ideal are **aberrations**

**Two types**

1. geometrical
   - spherical aberration
   - astigmatism
   - distortion
   - coma

2. chromatic
   Aberrations are reduced by combining lenses
   - Compound lenses

**Spherical aberration**

Rays parallel to the axis do not converge

Outer portions of the lens yield smaller focal lengths

**Astigmatism**

An optical system with astigmatism is one where rays that propagate in two perpendicular planes have different focus. If an optical system with astigmatism is used to form an image of a cross, the vertical and horizontal lines will be in sharp focus at two different distances.

**Distortion**

Magnification/focal length different for different angles of inclination

Can be corrected! (if parameters are known)

**Chromatic aberration**

(great for prisms, bad for lenses)
Chromatic aberration

rays of different wavelengths focused in different planes
cannot be removed completely

Vignetting

– Only part of the light reaches the sensor
– Periphery of the image is dimmer