CS 217: Light and Geometry in Vision **Spring 2009**

Lecture 17 — May 8

Scribe: Jiang, Shan *Lecturer: Deva Ramanan*

Note: These lecture notes are still rough, and have only have been mildly proofread.

References:
3D - Chapter 2, 6
M.V. - "Classic in field"

17.1 Formal Definition of Homogeneous Points and Projective Space

- P^n is a set of 1-D subspaces(lines through the origin) in $\Re^{n+1}$
- Any p $\in P^n$ can be represented by any non-origin point in the subspace $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n+1} \end{bmatrix}$
  where at least one $x_i \neq 0$
- We call $\mathbf{x}$ a homogeneous representation of a p $\in P^n$
- Fact: $\mathbf{x}$, $\lambda\mathbf{x}$ correspond to the same point p $\in P^n$, $\lambda > 0$
  Notation: $\mathbf{x} \sim \lambda\mathbf{x}$, where $\sim$ means equivalence in homogeneous coordinates

- Points p $\in P^n$ that do not intersect z = 1 plane are called ideal points or points at ∞

  - Why are ideal points useful?
    Recall: a point $\mathbf{x}$ on a line $\mathbf{l}$ ax + by + c = 0 is given by
    $\begin{bmatrix} x, y, 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \Rightarrow \mathbf{x}\cdot\mathbf{l} = 0$
    let $\mathbf{l}_1$, $\mathbf{l}_2$ be two lines, $\mathbf{x} = \mathbf{l}_1 \times \mathbf{l}_2$, $\mathbf{x}$ is the point of intersection in homogeneous coordinates
    e.g. $\mathbf{l}_1 = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$, $\mathbf{l}_2 = \begin{bmatrix} a \\ b \\ c' \end{bmatrix} \Rightarrow \mathbf{x} = \mathbf{l}_1 \times \mathbf{l}_2 = \begin{bmatrix} b \\ -a \\ 0 \end{bmatrix}$
    where $\mathbf{x}$ is parallel to z = 1 plane, i.e. $\mathbf{x}$ is an ideal point

17.2 Coordinate Systems(Camera Projection)

17.2.1 Rigid-Body Transformation

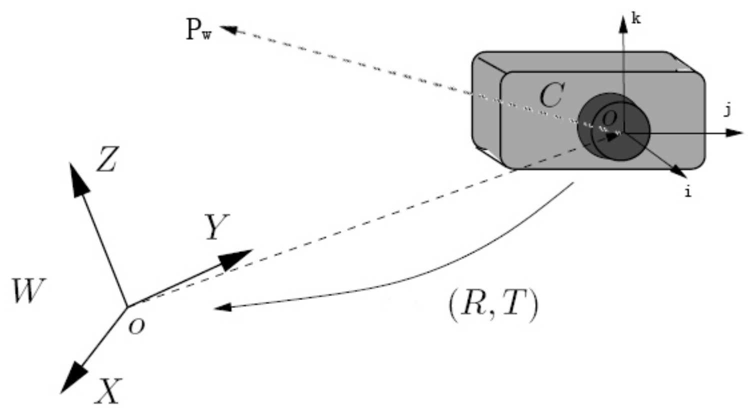

Figure 17.1. A rigid-body motion between a camera frame C: (x, y, z) and a world coordinate frame W: (X, Y, Z).

As shown in Figure 17.1, $\hat{i}$, $\hat{j}$, $\hat{k}$ are 3×1 unit vectors, such that R = $\begin{bmatrix} \hat{i}^T \\ \hat{j}^T \\ \hat{k}^T \end{bmatrix}$, $RR^T = I$ $= \begin{bmatrix} 1,0,0 \\ 0,1,0 \\ 0,0,1 \end{bmatrix}$; P_w is a point in the world coordinate system, where $P_w = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix}$.
The cooresponding coordinates of P_w in the camera coordinate system is

$$P_c = \begin{bmatrix} \hat{i}^T \\ \hat{j}^T \\ \hat{k}^T \end{bmatrix} \begin{bmatrix} P_w - C \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} R_{3\times3}, T_{3\times1} \\ \mathbf{0}^T, 1 \end{bmatrix}}_{g} \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \\ 1 \end{bmatrix},$$

where $T = -RC$. Matrix g is called euclidean transformation, also called rigid-body.

- How many degree of freedom in T? 3.
- How many degree of freedom in R? 3.
- There are lots of ways to represent R.

17.2.2 Camera Projection

- $x = f\frac{X}{Z}$, $y = f\frac{Y}{Z} \Rightarrow$
  $$Z \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f,0,0 \\ 0,f,0 \\ 0,0,1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \Rightarrow$$

$\mathbf{x} \sim K\mathbf{X}$

- If f = 1, $\mathbf{x} \sim \mathbf{X}$, and $\lambda\mathbf{x} \sim \mathbf{X}$ for some $\lambda \neq 0$.

- Camera projection:

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} f,0,0 \\ 0,f,0 \\ 0,0,1 \end{bmatrix}}_{K} \underbrace{\begin{bmatrix} 1,0,0,0 \\ 0,1,0,0 \\ 0,0,1,0 \end{bmatrix}}_{\Pi_0} \underbrace{\begin{bmatrix} R,T \\ \mathbf{0}^T,1 \end{bmatrix}}_{g} \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \\ 1 \end{bmatrix} \tag{17.1}$$

Where K is "intrinsic parameters", g is "extrinsic parameters".

- Formular 17.1 is the one used by 3D, and V.G. uses $\widetilde{K} = K\Pi_0$.

- K can be more complex like $\begin{bmatrix} f_{sx}, f_{s\theta}, O_x \\ 0, f_{sy}, O_y \\ 0,0,1 \end{bmatrix}$, where f_{sx} and f_{sy} define the scaled focus length, $f_{s\theta}$ defines the shear parameter, and O_x and O_y define the translation of the center of the image plane.

  However, typically, $K \approx \begin{bmatrix} f,0,0 \\ 0,f,0 \\ 0,0,1 \end{bmatrix}$

17.3 Two View Calibrated Geometry

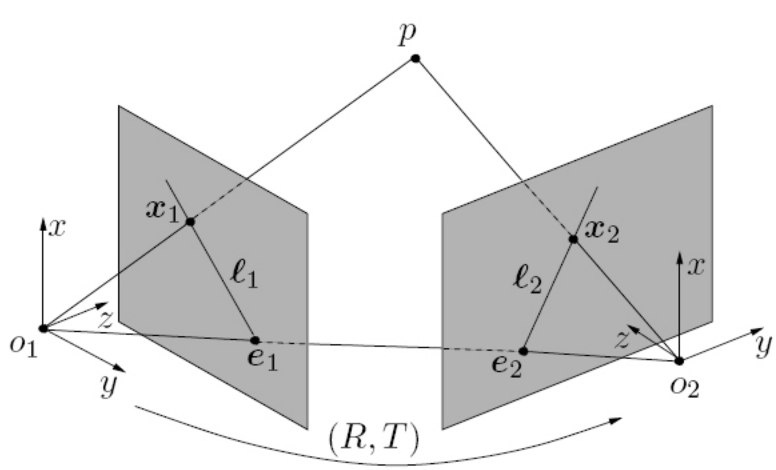

Figure 17.2. Two calibrated cameras viewing at the same point in 3D space.

- Assumption:
  Let $K_1 = I$ and $K_2 = I$ (Images generated by $f = 1$).

- Terms:
  $\mathbf{b} = \mathbf{o}_2 - \mathbf{o}_1$ is the baseline;
  Plane formed by $\mathbf{o}_1$, $\mathbf{o}_2$, p is called the epipolar plane;
  Intersections of $\mathbf{b}$ with image planes yield $\mathbf{e}_1$, $\mathbf{e}_2$, which are called epipoles;
  Intersections of epipolar plane with images yield $\mathbf{l}_1$, $\mathbf{l}_2$, which are called epipolar lines.