CSE-483: Mobile Robotics **Monsoon 2019**

Linear Triangulation - Structure Computation

Prepared by: AadilMehdi Sanchawala, Rahul Sajnani

In this note, we discuss methods to compute the position of points in the 3D world given its images in two views and the camera matrices of those views. In the following we will look at methods we can use to estimate 3D points from known camera matrices and 2D correspondences between their respective image coordinates.

1 Recap

1.1 Camera Modeling

We model the camera to world projection using the camera intrinsic matrix $\boldsymbol{K}$ and the camera extrinsic matrix $[R|T]$. Here $\boldsymbol{K}$ gives the camera intrinsic parameters f_x, f_y, c_x, c_y, s.

1. f_x is the pixel focal length of the camera along x axis.
2. f_y is the pixel focal length of the camera along y axis.
3. c_x is the pixel optical center of the camera along x axis.
4. c_y is the pixel optical center of the camera along y axis.
5. s is the skew coefficient of the camera. Where s = $f_x tan(\alpha)$. Here α is the skew angle.

$$\boldsymbol{K} = \begin{bmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

Note: In the above camera intrinsic matrix we assume there is no distortion.

$\boldsymbol{R_c^w}$ is the rotation of camera with respect to world.
$\boldsymbol{T_c^w}$ is translation of camera with respect to world. The camera extrinsic matrix is composed of $[\boldsymbol{R_w^c}|\boldsymbol{T_w^c}]$. The extrinsic matrix projects the world points into the camera's frame. To project the 3D points from camera's frame to the image plane we pre-multiply the extrinsic matrix with the intrinsic matrix. Our camera matrix is $\boldsymbol{P}$. Where $\boldsymbol{P} = \boldsymbol{K}\boldsymbol{R_w^c}[\boldsymbol{I_3}|\boldsymbol{T_w^c}]$.

Note: $\boldsymbol{T_w^c = -X_0}$. Here $\boldsymbol{X_0}$ is the camera center in the world's frame.

2 Parameterization Model

We use the General Parameterization for dependent images model which uses a normalized direction vector $\boldsymbol{b}$ and a rotation matrix $\boldsymbol{R}$.
In this parameterization, the 1^{st} camera is placed in the origin of the reference frame. The 2^{nd} camera's position is given by $\mathbf{b}$ which tells us the direction of the 2^{nd} camera with respect to the 1^{st} camera and $\boldsymbol{R_2^1}$ which gives the orientation of the 2^{nd} camera with respect to the 1^{st} camera.

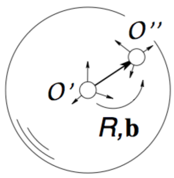

Figure 1: General Parameterization for dependent images

3 Motivation

Given the camera matrix and images from multiple views we would like to get the world coordinates corresponding to image coordinates.
Triangulation is used in visual odometry to get the correct R and T from essential matrix. It also has its applications in point cloud estimation and camera pose estimation.

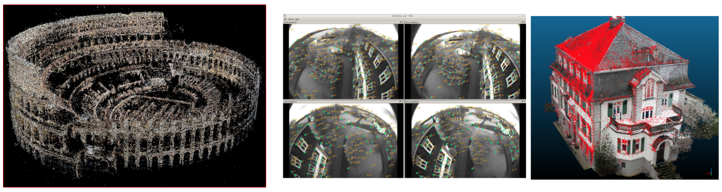

Figure 2: Triangulated 3D Point Cloud of the Roman Coloseum and Point Cloud estimation

4 Triangulation

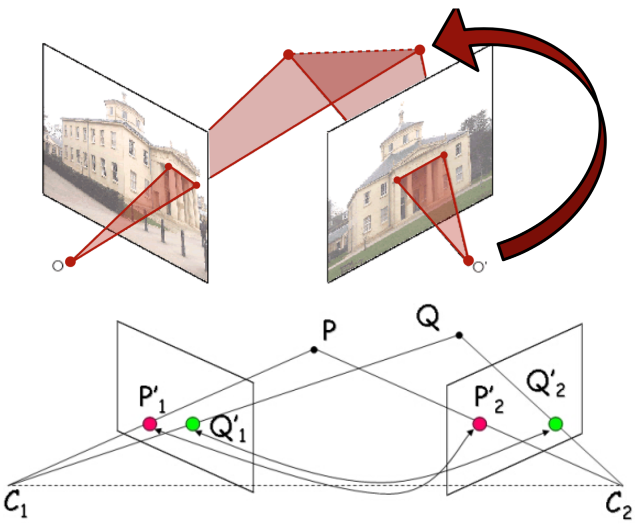

Figure 3: Triangulation

4.1 Linear Triangulation using DLT

In this section we describe a simple linear or algebraic triangulation method. This linear triangulation method is a direct analogue of DLT. In each image we have a measurement $\boldsymbol{x} = \boldsymbol{P}\boldsymbol{X}$ and $\boldsymbol{x'} = \boldsymbol{P'}\boldsymbol{X}$ which can be combined into a form $\boldsymbol{A}\boldsymbol{X} = \boldsymbol{0}$

In order to construct A matrix, we perform a cross product $\boldsymbol{x} \times \boldsymbol{PX}$ to give three equations for each image point. Let us take an example for the sake of clarity.

$$\boldsymbol{x} \times \begin{bmatrix} P_1^T \\ P_2^T \\ P_3^T \end{bmatrix} \boldsymbol{X} = 0 \tag{1}$$

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \times \begin{bmatrix} P_1^T \boldsymbol{X} \\ P_2^T \boldsymbol{X} \\ P_3^T \boldsymbol{X} \end{bmatrix} = 0 \tag{2}$$

where $\boldsymbol{x}$ is the homogeneous image coordinate, $\boldsymbol{P}$ is a 3×4 projection matrix and $\boldsymbol{X}$ is the homogeneous 3D world coordinate.

$$\begin{bmatrix} 0 & -1 & y \\ 1 & 0 & -x \\ -y & x & 0 \end{bmatrix} * \begin{bmatrix} P_1^T \boldsymbol{X} \\ P_2^T \boldsymbol{X} \\ P_3^T \boldsymbol{X} \end{bmatrix} = 0 \tag{3}$$

which yeild the following three equations

$$yP_3^T \boldsymbol{X} - P_2^T \boldsymbol{X} = 0 \tag{4}$$

$$xP_3^T \boldsymbol{X} - P_1^T \boldsymbol{X} = 0 \tag{5}$$

$$xP_2^T \boldsymbol{X} - yP_1^T \boldsymbol{X} = 0 \tag{6}$$

Out of these three equations, 2 of the equations are linearly independent. Moreover, these equations are linear in X.

Thus A matrix can be composed as follows

$$\begin{bmatrix} xP_3^T - P_1^T \\ yP_3^T - P_2^T \\ x^{'}P_3^{'T} - P_1^{'T} \\ y^{'}P_3^{'T} - P_2^{'T} \end{bmatrix} \tag{7}$$

Here we include two equations from each image giving a total of four equations in 4 homogeneous unknowns that is

$$\boldsymbol{X} = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \tag{8}$$

Thus, we obtain $\boldsymbol{X}$, the 3D triangulated world point.

4.2 Geometric Triangulation

In geometric triangulation we first find the ray from the camera to the 3D point in the world.

$$\boldsymbol{r} = K_1^{-1} \boldsymbol{x_1} \tag{9}$$

$$\boldsymbol{s} = R_2^1 K_2^{-1} \boldsymbol{x_2} \tag{10}$$

Here $\boldsymbol{x_1}$ is image pixel coordinate (homogeneous) in camera 1 and $\boldsymbol{x_2}$ is image pixel coordinate (homogeneous) in camera 2. $\boldsymbol{K_1}$ and $\boldsymbol{K_2}$ are intrinsic matrices for camera 1 and camera 2 respectively. $\boldsymbol{R_2^1}$ is the relative orientation of camera 2 with respect to camera 1.

Our equations of lines are as follows,

$$\boldsymbol{f} = \boldsymbol{P} + \lambda \boldsymbol{r} \tag{11}$$

$$\boldsymbol{g} = \boldsymbol{Q} + \mu \boldsymbol{s} \tag{12}$$

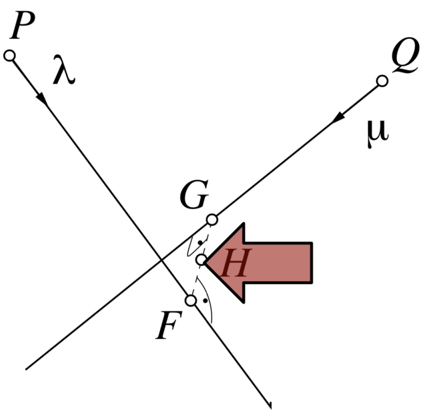

Figure 4: Geometric Triangulation

Here $\boldsymbol{P}$ and $\boldsymbol{Q}$ are centers of camera 1 and 2 in the world respectively. Next, we have to find the points on the line where the distance between the is minimum as the points may not intersect always. For this we have to find μ and λ. Considering the world points are F and G. For the distance to be minimum we need the line (F - G) to be perpendicular to both lines **r** and **s**.

We have the following equations:

$$(\boldsymbol{F} - \boldsymbol{G}) \cdot \boldsymbol{r} = 0 \tag{13}$$

$$(\boldsymbol{F} - \boldsymbol{G}) \cdot \boldsymbol{s} = 0 \tag{14}$$

$$(\boldsymbol{P} + \lambda\boldsymbol{r} - (\boldsymbol{Q} + \mu\boldsymbol{s})) \cdot \boldsymbol{r} = 0 \tag{15}$$

$$(\boldsymbol{P} + \lambda\boldsymbol{r} - (\boldsymbol{Q} + \mu\boldsymbol{s})) \cdot \boldsymbol{s} = 0 \tag{16}$$

We have two equations and two unknowns. To solve for λ and μ we use the following equations:

$$(\boldsymbol{P} + \lambda\boldsymbol{r} - (\boldsymbol{Q} + \mu\boldsymbol{s})) \cdot \boldsymbol{r} = 0 \tag{17}$$

$$(\boldsymbol{P} + \lambda\boldsymbol{r} - (\boldsymbol{Q} + \mu\boldsymbol{s})) \cdot \boldsymbol{s} = 0 \tag{18}$$

which can be written as.

$$\begin{bmatrix} r \cdot r - s \cdot r \\ r \cdot s - s \cdot s \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \end{bmatrix} = \begin{bmatrix} (\boldsymbol{Q} - \boldsymbol{P}) \cdot r \\ (\boldsymbol{Q} - \boldsymbol{P}) \cdot s \end{bmatrix} \tag{19}$$

$$\begin{bmatrix} \lambda \\ \mu \end{bmatrix} = \begin{bmatrix} r \cdot r - s \cdot r \\ r \cdot s - s \cdot s \end{bmatrix}^{-1} \begin{bmatrix} (\boldsymbol{Q} - \boldsymbol{P}) \cdot r \\ (\boldsymbol{Q} - \boldsymbol{P}) \cdot s \end{bmatrix} \tag{20}$$

Once we have found λ and μ, we can obtain $\boldsymbol{F}$ and $\boldsymbol{G}$. We take the mid-point of this line segment $\boldsymbol{H} = \boldsymbol{F} - \boldsymbol{G}$ as our estimate for the 3D triangulated world point.

5 Summary

In these set of notes we,

1. revisited the concepts of camera modeling
2. described the parameterization convention used
3. discussed the motivation and application of triangulation
4. discussed the linear / algebraic method od triangulation
5. discussed the geometric method of triangulation

References

[1] Class Lectures

[2] Cyrill Stachniss course on photogrammetry

[3] Multi-View Geometry - Hartley and Zisserman