Appendix (VIO Lecture): Triangulation

Scribe: Charbel Toumieh

1 Method using 2 lines

The objective is to triangulate a point seen by 2 images taken at different poses. You will need the focal length in pixels (computed below), and the position of the camera relative to the body frame: $(x = 0.03, y = 0, z = 0.01)$. The orientation of the camera is the same as the body frame in the sense that it is looking in the forward direction of the drone. **However, the frame orientation conventions are different and need to be adjusted with a rotation matrix such that:** $z_{\text{cam}} = x_{\text{drone}},\ x_{\text{cam}} = -y_{\text{drone}},\ y_{\text{cam}} = -z_{\text{drone}}$.

First, get the pixel coordinates of the point in each image (using computer vision techniques such as color masks ...). Then get the vector that goes from the first camera's optical center to the point, $\boldsymbol{v}$. To get this vector, use slides 13, 14, and 15 in the course: first move the origin of the image from the top left to the center of the image and compute the new pixel coordinates; v_x and v_y are the new pixel coordinates. v_z is simply the focal length in pixels which is:

$$v_z = f_{\text{pixels}} = \frac{W}{2\tan(\frac{\text{FOV}}{2})} = 161.013922282 \tag{1}$$

where FOV is the camera field of view (1.5 rad) and W is the width of the image (300 pixels). These coordinates are in the first camera's frame. Do the same for the second image to get the coordinates of the vector going from that camera's optical center to the pixel, which we will call $\boldsymbol{v}'$.

Then move these vectors from each camera's frame to the world/inertial frame:

$$\boldsymbol{r} = \boldsymbol{R}_{C_1 2W}\boldsymbol{v} \tag{2}$$

$$\boldsymbol{s} = \boldsymbol{R}_{C_2 2W}\boldsymbol{v}' \tag{3}$$

where $\boldsymbol{R}_{C_1 2W}$ takes a vector from the first camera frame to the world frame, and, $\boldsymbol{R}_{C_2 2W}$ takes a vector from the second camera frame to the world frame (**it is a composition of the rotation matrix that takes us from the camera frame to the body frame multiplied by the rotation matrix that takes us from the body frame to the inertial frame**).

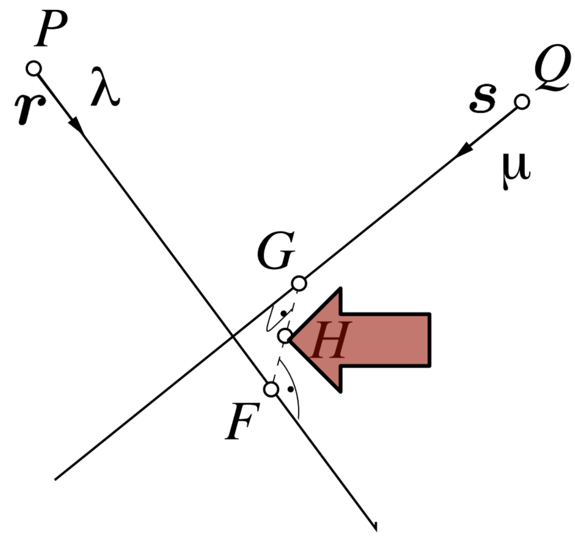

Figure 1: Lines going out of the cameras centers P and Q and through the point/pixel we want to triangulate. They do not meet, but there is a segment orthogonal to both that represents the shortest distance between any 2 points on these lines. The midpoint of this segment is a good approximation of the position of the 3D point we want to triangulate.

Now we can write the equations of the lines that go out from the optical center and go through the pixel (Fig. 1 - the equations are now in the world frame since all vectors/points are now in the world frame):

$$F = P + \lambda \boldsymbol{r} \tag{4}$$

$$G = Q + \mu \boldsymbol{s} \tag{5}$$

where P is the position of the first camera, Q the position of the second camera, and F and G points on each line (by varying λ and μ you move along the line). **Note that the positions of the camera are not the same as the drone but there is a translation vector/offset**. However, the lines may not necessarily intersect because of inaccuracies in the point detection (Fig. 1)

So we find the midpoint of the segment $F - G$ which represents the closest distance between the lines and is orthogonal to both. to do that we first write out the conditions that $F - G$ is orthogonal to both segments:

$$(F - G) \cdot \boldsymbol{r} = (P + \lambda \boldsymbol{r} - Q - \mu \boldsymbol{s}) \cdot \boldsymbol{r} = 0 \tag{6}$$

$$(F - G) \cdot \boldsymbol{s} = (P + \lambda \boldsymbol{r} - Q - \mu \boldsymbol{s}) \cdot \boldsymbol{s} = 0 \tag{7}$$

This gives us 2 linear equations with 2 unknowns λ and μ which we can solve by writing it in matrix form and inverting the matrix ($Ax = b$ then $x = A^{-1}b$). After we find λ and μ we also find F and G by replacing the values of λ and μ in the line equations. Then we find H which is the midpoint between F and G:

$$H = \frac{F + G}{2} \tag{8}$$

λ and μ can also be solved by getting the least squares solution of the equation $P + \lambda r = Q + \mu s$:

$$\begin{bmatrix} r_x & -s_x \\ r_y & -s_y \\ r_z & -s_z \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \end{bmatrix} = \begin{bmatrix} Q_x - P_x \\ Q_y - P_y \\ Q_z - P_z \end{bmatrix} \implies \begin{bmatrix} \lambda \\ \mu \end{bmatrix} = \text{pseudoinverse}(\begin{bmatrix} \boldsymbol{r} & -\boldsymbol{s} \end{bmatrix}) \cdot (Q - P) \tag{9}$$

We can also use more than 2 lines to triangulate the point and get more accuracy using.