Camera Calibration and Depth Evaluation

Angelo Nutu

Driverless Project 4

This document will try not to be overly verbose, hopefully the whole process is clear enough.

First step of the process is converting the given pixel coordinates $(u, v) = (795, 467)[px]$ to image coordinates. The concept behind this choice is to shift the point of reference from the principal point along the principal axis, and get a new tuple centered on the pinhole or lens.

$$u' = u - c_x = 159px$$

$$v' = v - c_y = -81px$$

Next step is to normalize the image coordinates in order to truncate the dependency from the focal lengths, and get a general scale of the coordinates in a standardized range. From there we can directly re-map the object back into its 3D position based on the camera coordinate frame.

$$X_c = \frac{u'}{f_x} \cdot d = 1,782m$$

$$Y_c = \frac{v'}{f_y} \cdot d = -0.919m$$

$$Z_c = d = 2.7m$$

From there, we need to find and apply the rotation matrix to account for the angle of the camera, the steps are rather bothersome but we can do this.

$$R_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 100^\circ & -\sin 100^\circ \\ 0 & \sin 100^\circ & \cos 100^\circ \end{bmatrix}$$

$$R_y = \begin{bmatrix} \cos 0^\circ & 0 & \sin 0^\circ \\ 0 & 1 & 0 \\ -\sin 0^\circ & 0 & \cos 0^\circ \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R_z = \begin{bmatrix} \cos 90^\circ & -\sin 90^\circ & 0 \\ \sin 90^\circ & \cos 90^\circ & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R = R_z \cdot R_y \cdot R_x = \begin{bmatrix} \mathbf{0} & \mathbf{-\cos(100^\circ)} & \mathbf{\sin(100^\circ)} \\ \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\sin(100^\circ)} & \mathbf{\cos(100^\circ)} \end{bmatrix}$$

Last step is to apply the rotation matrix, and since we're working with a camera reference system already, we might want to add a traslation matrix to the gist as to find the position of the object with reference to the vehicle. Thankfully, that's easy enough since the task provides us with the position of the camera in the vehicle frame.

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \left[\begin{array}{c|c} R & t \\ 0 & 1 \end{array} \right] \cdot \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \left[\begin{array}{ccc|c} 0 & -\cos 100^\circ & \sin 100^\circ & 0.5 \\ 1 & 0 & 0 & 0.16 \\ 0 & \sin 100^\circ & \cos 100^\circ & 1.14 \\ 0 & 0 & 0 & 1 \end{array} \right] \cdot \begin{bmatrix} 1.782 \\ -0.919 \\ 2.7 \\ 1 \end{bmatrix}$$

$$\approx \begin{bmatrix} \mathbf{-0.075} \\ \mathbf{1.942} \\ \mathbf{3.934} \\ \mathbf{1} \end{bmatrix}$$

In conclusion, we have walked through the steps of a simple instance of depth evaluation. We started by converting pixel coordinates to image coordinates, factoring in the principal point and normalizing the values. With the camera projection equation and rotation matrices, we nailed down the object's 3D position in the camera coordinate frame. By adding the translation matrix, we transformed the coordinates to match the vehicle's frame of reference. These findings provide crucial insights into the object's exact location relative to the vehicle.