Formulation of Lane Lines 2D to 3D Inverse Perspective Mapping (IPM)

Smart Automobile, Inc.

June 18, 2022

1 Problem

Figure 1 illustrates three coordinate systems: ego frame (i.e., the world coordinate system $\mathcal{F}_w$), camera frame $\mathcal{F}_w$ and image pixel coordinate system $\mathcal{F}_i$. When lane lines are detected in 2D images via deep learning models, the most challenging task is to inverse project the points of lane lines into world coordinate system $\mathcal{F}_w$.

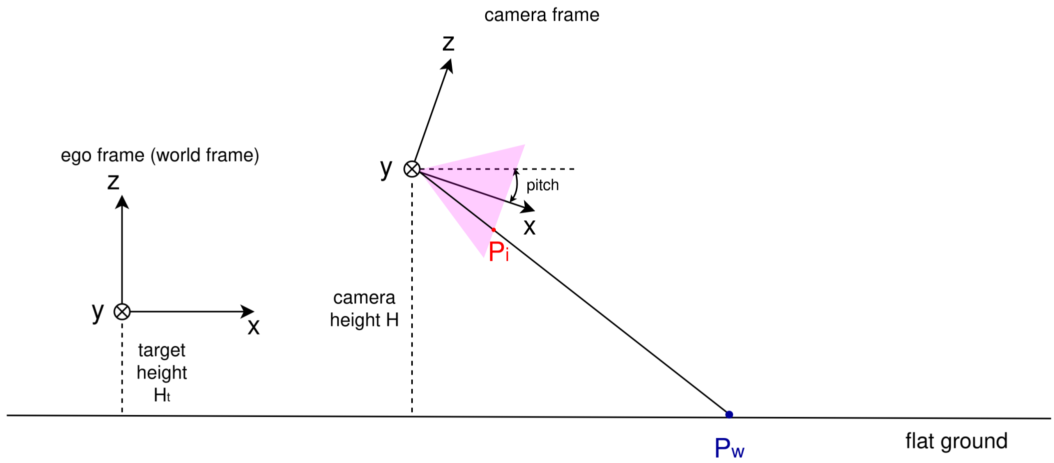

Figure 1: Illustration of ego frame, camera frame and camera pose.

2 Notation

Consider a pixel point of lane lines in image plane $P_i = (u, v, 1)^T$, its corresponding 3D point in camera frame is noted as $P_c = (x_c, y_c, z_c)^T$ and its representation in ego frame is $P_w = (x_w, z_w, z_w)^T$.

The rotation matrix from world frame to camera frame is represented as $\boldsymbol{R}_{3\times3}$.
The position of camera frame's origin point in world frame is represented as $\boldsymbol{T}_{3\times1}$.
The camera intrinsic matrix is represented as $\boldsymbol{I}_{3\times3}$.

3 Formualation

Follow the camera extrinisc and intrinsic model, the relationship between P_w and P_c, P_c and P_i are formulated as:

$$P_c = \boldsymbol{R}_{3\times3} P_w - \boldsymbol{T}_{3\times1} \tag{1}$$

$$P_i = \frac{1}{z_c} \boldsymbol{I}_{3\times3} P_c \tag{2}$$

respectively. Combine Equation 1 and 2, we can obtain:

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \boldsymbol{I}_{3\times3}\boldsymbol{R}_{3\times3} \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} - \boldsymbol{I}_{3\times3}\boldsymbol{T}_{3\times1} \tag{3}$$

Therefore,

$$\boldsymbol{R}_{3\times3}^{-1}\boldsymbol{I}_{3\times3}^{-1} \left(z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} + \boldsymbol{I}_{3\times3}\boldsymbol{T}_{3\times1} \right) = \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} \tag{4}$$

In the equation above, it is worth mentioning that the extrinsic matrix $\boldsymbol{R}_{3\times3}$, $\boldsymbol{T}_{3\times1}$ and the camera intrisic matrix $\boldsymbol{I}_{3\times3}$ are calibirated in advance and already known. However, the depth z_c is unknown and it is crital for calculating P_w.

We introduce three assumptions:
1. the ground plane is extremely flat;
2. the ego frame's $XY - plane$ is parallel with ground;
3. the origin point of ego frame is with height H_t with respect to ground.

Then, we could config z_w in Equation 4 to $-H_t$, we represent $\boldsymbol{R}_{3\times3}^{-1}\boldsymbol{I}_{3\times3}^{-1}$ and $\boldsymbol{I}_{3\times3}\boldsymbol{T}_{3\times1}$ as below:

$$\boldsymbol{R}_{3\times3}^{-1}\boldsymbol{I}_{3\times3}^{-1} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \boldsymbol{I}_{3\times3}\boldsymbol{T}_{3\times1} = \begin{bmatrix} b1 \\ b2 \\ b3 \end{bmatrix} \tag{5}$$

Combine Equation 5 and Equation 4 with condition $z_w = -H_t$, we can obtain:

$$z_c(a_{31}u + a_{32}v + a_{33}) = -H_t - b_3 \tag{6}$$

i.e, the depth z_c is:

$$z_c = \frac{-H_t - b_3}{a_{31}u + a_{32}v + a_{33}} \tag{7}$$

After z_c is calculated, then the corresponding point P_w could be easily obtained with Equation 4.