Supplementary Material

Arsalan Mousavian[*]
George Mason University
amousavi@gmu.edu

Dragomir Anguelov
Zoox, Inc.
drago@zoox.com

John Flynn
Zoox, Inc.
john.flynn@zoox.com

Jana Košecká
George Mason University
kosecka@gmu.edu

1. Solving for 3D Box Translation

1.1. 2D Box Constraints

As mentioned in the paper, K denotes the known intrinsic parameters, R is the corresponding rotation matrix and D is the estimated dimensions of the box, both regressed using CNN. Let the vertical side of the 2D bounding box correspond to $\mathbf{X_o^i}$, the ith corner of the 3D box then the constraint for the correspondence is as follows:

$$\left(K \begin{bmatrix} I & R \times \mathbf{X_o^i} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} T_x \\ T_y \\ T_z \\ 1 \end{bmatrix}\right)_x = x_{vertical} \tag{1}$$

The only unknowns are the the translation $T = [T_x, T_y, T_z]$. Similarly for horizontal edges we have the following equation:

$$\left(K \begin{bmatrix} I & R \times \mathbf{X_o^j} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} T_x \\ T_y \\ T_z \\ 1 \end{bmatrix}\right)_y = y_{horizontal} \tag{2}$$

In the equations above I is a 3×3 identity matrix.

1.2. Solving the Linear System

There are 4 different equations corresponding to the 4 sides of the 2D box and we want to solve for the translation parameters T. The constraints of Eq (1) and (2) are rearranged to have the form of $Ax = 0$ and solution is found using SVD decomposition.

2. Equivalence between L2 Loss and Single Bin MultiBin Loss for Angle Regression

In this section, we demonstrate the equivalence between direct angle regression using L2 loss and using our MultiBin loss with a single bin. In our representation, the L2 loss for an angle is equal to the Euclidean distance between two vectors on a unit circle. As a result, the L2 loss for a ground truth angle θ^* and estimated angle θ is as follows:

$$L = \frac{1}{n} \sum_i \left(\cos(\theta_i^*) - \cos(\theta_i)\right)^2 + \left(\sin(\theta_i^*) - \sin(\theta_i)\right)^2\Big) \tag{3}$$

where n is the batch size. If we expand Eq. (3), we get the following equation:

$$\begin{aligned} L =& \frac{1}{n} \sum_i \left(\cos(\theta_i^*) - \cos(\theta_i)\right)^2 + \left(\sin(\theta_i^*) - \sin(\theta_i)\right)^2\Big) \\ =& \frac{1}{n} \sum_i \left(cos(\theta_i^*)^2 + cos(\theta_i)^2 - 2cos(\theta_i^*)cos(\theta_i)\right) + \\ & \frac{1}{n} \sum_i \left(sin(\theta_i^*)^2 + sin(\theta_i)^2 - 2sin(\theta_i^*)sin(\theta_i)\right) \\ =& 2 - \frac{2}{n} \sum_i cos(\theta_i^* - \theta_i). \end{aligned} \tag{4}$$

With one bin, the loss reduces to L_{loc} which only depends to the cosine of angle difference.

[*]Work done as an intern at Zoox, Inc.