2D-Driven 3D Object Detection in RGB-D Images

Jean Lahoud, Bernard Ghanem
King Abdullah University of Science and Technology (KAUST)
Thuwal, Saudi Arabia
{jean.lahoud, bernard.ghanem}@kaust.edu.sa

1. Implementation Details

Bounding Box Regression

Given the point density along one of the estimated directions as shown in Figure 1, we use a multilayer perceptron (MLP) network that learns to regress the boundaries of the bounding box (shown in red). This is equivalent to finding the length, width, and height along a fixed orientation. In our experiments, we use a network with 5 hidden layers of sigmoid neurons and an output layer with linear neurons. During training, we use the point densities along every dimension from the training set. The point densities are obtained by computing the histogram of the point coordinates along every direction. We use histogram bins with a constant size ($0.1m$) and fix the total number of bins to preserve the same input size. We also translate the groundtruth boxes along the given direction and append the translations to the training set. The translations extend from $-0.5m$ to $0.5m$ with a step size of $0.1m$. As for the training algorithm, we use Levenberg-Marquardt backpropagation.

Refinement Based on Context Information

- SVM classifier is learned on 10 classes plus background class.
- 3D bounding box features consist of deep features and geometric features. Using PCA, we reduce the feature dimension to 30.
- Binary term probability: $p_b = p_o^{\alpha} p_d^{1-\alpha}$ with $\alpha = 0.1$
- Once the new labels are obtained from the LP-MAP assignment, the scores of the boxes that retain their original label are increased by $0.2 \times posterior$ of that box having that particular label. In case the label was different, the score is decreased by 0.2.

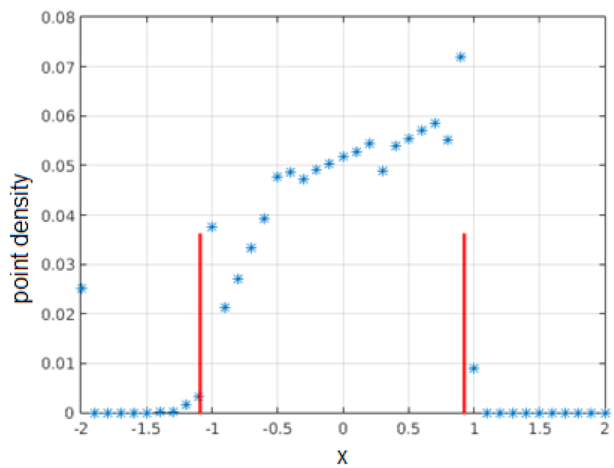

Figure 1. Given the point density along one of the estimated directions (x in this example), we use a multilayer perceptron (MLP) network that learns to regress the extents of the bounding box of the object (shown in red)