The supplementary materials provide more experimental details and more visualization results.

1. Detailed Metrics

In our paper, we adopt eight standard metrics for evaluation, including RMSE, MAE, iRMSE, iMAE, REL, $\delta_i (i = 1.25, 1.25^2, 1.25^3)$. Formally, the metrics are defined as follows:

- Root mean square error (RMSE):
  $$\sqrt{\frac{1}{|V|} \sum_{x \in V} (\hat{D}(x) - D(x))^2}$$

- Mean absolute error (MAE):
  $$\frac{1}{|V|} \sum_{x \in V} |\hat{D}(x) - D(x)|$$

- Root mean square error of the inverse depth (iRMSE):
  $$\sqrt{\frac{1}{|V|} \sum_{x \in V} (\frac{1}{\hat{D}(x)} - \frac{1}{D(x)})^2}$$

- Mean absolute error of the inverse depth (iMAE):
  $$\frac{1}{|V|} \sum_{x \in V} \left| \frac{1}{\hat{D}(x)} - \frac{1}{D(x)} \right|$$

- Mean relative error (REL):
  $$\frac{1}{|V|} \sum_{x \in V} \left| \frac{\hat{D}(x) - D(x)}{D(x)} \right|$$

- Thresholded accuracy ($\delta_i (i = 1.25, 1.25^2, 1.25^3)$)
  $$\max \left( \frac{D(x)}{\hat{D}(x)}, \frac{\hat{D}(x)}{D(x)} \right) < 1.25^i$$

where $D$ and $\hat{D}$ refer to ground truth depth map and predicted depth map, respectively; $x$ denotes the indexes of depth map; $V$ represents the set of valid pixels in $D$; $|V|$ is the number of valid pixels.

2. Camera Models

Let denote the image coordinate of an pixel as $C = (u_i, v_i)$, whose corresponding depth value is $d_i$, and the coordinate in the 3D world system as $P = (x_i, y_i, z_i)$. Denoting the camera intrinsic matrix as $K \in \mathbb{R}^{4 \times 4}$, then $C$ can be transformed to $P$ as follows

$$[x_i, y_i, z_i, 1]^T = K^{-1} \cdot [u_i \times d_i, v_i \times d_i, d_i, 1]^T.$$  

3. More Qualitative Comparison Results

In this section, we provide more visualization results for qualitative comparison analysis.

References

Figure 1. Qualitative depth completion results on the NYU Depth V2 dataset [6] with 500 sampling points. (a) Color image, (b) Sparse depth, (c) CSPN [1], (d) NLSPN [5], (e) GraphCSPN [4], (f) CostDCNet [3], (e) PointDC (ours), (h) Ground truth.

Figure 2. Qualitative depth completion results on the NYU Depth V2 dataset [6] with 300 sampling points. (a) Color image, (b) Sparse depth, (c) CSPN [1], (d) NLSPN [5], (e) GraphCSPN [4], (f) CostDCNet [3], (e) PointDC (ours), (h) Ground truth.


Figure 3. Qualitative depth completion results on the NYU Depth V2 dataset [6] with 100 sampling points. (a) Color image, (b) Sparse depth, (c) CSPN [1], (d) NLSPN [5], (e) GraphCSPN [4], (f) CostDCNet [3], (e) PointDC (ours), (h) Ground truth.

Figure 4. Qualitative depth completion results on the KITTI DC Dataset [8] with sampling ratio as 1. (a) Color image, (b) GuideNet [7], (d) NLSPN [5], (d) PENet [2], (e) ACMNet [9], (f) PointDC (ours).

Figure 5. Qualitative depth completion results on the KITTI DC Dataset [8] with sampling ratio as 0.4. (a) Color image, (b) GuideNet [7], (d) NLSPN [5], (d) PENet [2], (e) ACMNet [9], (f) PointDC (ours).

Figure 6. Qualitative depth completion results on the KITTI DC Dataset [8] with sampling ratio as 0.1. (a) Color image, (b) GuideNet [7], (d) NLSPN [5], (d) PENet [2], (e) ACMNet [9], (f) PointDC (ours).
Figure 7. More qualitative depth completion results on the NYU Depth V2 dataset [6]. (a) Color image, (b) Sparse depth, (c) CSPN [1], (d) NLSPN [5], (e) GraphCSPN [4], (f) CostDCNet [3], (g) PointDC (ours), (h) Ground truth.

Figure 8. More qualitative depth completion results on the KITTI DC Dataset [8]. (a) Color image, (b) GuideNet [7], (d) NLSPN [5], (d) PENet [2], (e) ACMNet [9], (f) PointDC (ours).