Adaptive Surface Normal Constraint for Depth Estimation Supplementary Materials

Xiaoxiao Long¹ Cheng Lin^{1,4} Lingjie Liu² Wei Li³ Christian Theobalt² Ruigang Yang³ Wenping Wang⁵ ¹The University of Hong Kong ²Max Planck Institute for Informatics ³Inceptio ⁴Digital Content Technology Center, CROS, Tencent Games ⁵Texas A&M University

1. More Results

In this section, we present more qualitative results for several SOTA methods that also use geometric constraints, including i) GeoNet [2] (least square normal); ii)VNL [3] (virtual normal constraint); iii) BTS [1] (predicting local plane equations instead of directly predicting depth).

Guidance feature In Fig. 1, we show more visualization results of the guidance feature. Compared with the edge maps detected by Canny operator, our guidance feature map determines the true geometric shape boundaries instead of image intensity edges. For example, in the second row of Fig. 1, the Canny operator detects all image intensity edges including the posters on the wall, while our guidance feature only perceives the true geometric boundaries.

Depth As shown in Fig. 2, we present more qualitative comparisons of the estimated depth maps. Compared with the other methods, our estimated depth is more accurate and can preserve more details. For example, in the second row of Fig. 2, our estimated depth even accurately predicts the depth of the human body, which demonstrates our capacity to identify various shape geometries. Moreover, in the fourth row of Fig. 2, all the other methods incorrectly perceive the carpets on the floor while our estimated depth faithfully recovers its true geometry.

Surface normal More qualitative results for the recovered surface normals are shown in Fig. 3. Compared with the other methods, the surface normals recovered from our estimated depth are more accurate and smooth, which indicates that our method faithfully captures the true 3D geometry of various shapes.

Point cloud We show more qualitative results for the reconstructed point clouds in Fig. 4. Our point clouds are less distorted and more accurate than the other methods,



Figure 1. More visualization results of guidance feature.

which thus can faithfully reconstruct sharp corners, sofas, and even texture-less walls. With the proposed Adaptive Surface Normal (ASN) constraint, our method preserves not only global structural information but also local geometric features.

References

- Jin Han Lee, Myung-Kyu Han, Dong Wook Ko, and Il Hong Suh. From big to small: Multi-scale local planar guidance for monocular depth estimation. *arXiv preprint arXiv:1907.10326*, 2019.
- [2] Xiaojuan Qi, Renjie Liao, Zhengzhe Liu, Raquel Urtasun, and Jiaya Jia. Geonet: Geometric neural network for joint depth and surface normal estimation. In *Proceedings of the*



Figure 2. More qualitative comparisons of estimated depth.



Figure 3. More qualitative comparisons of surface normal.

IEEE Conference on Computer Vision and Pattern Recognition, pages 283–291, 2018. 1

forcing geometric constraints of virtual normal for depth prediction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5684–5693, 2019. 1

[3] Wei Yin, Yifan Liu, Chunhua Shen, and Youliang Yan. En-



Figure 4. More qualitative comparisons of point cloud.