Supplementary material of DebSDF: Delving into the Details and Bias of Neural Indoor Scene Reconstruction

Yuting Xiao*, Jingwei Xu*, Zehao Yu, Shenghua Gao

1 Large Curvature Radius Situation

The SDF mapping function $y(t)$ can be simplified to $y(t) = s(t)/|\cos\theta(t)|$ when the absolute value of the curvature radius is a very large number. This corresponds to the situation that the ray intersects with the planar surface. This approximation comes from the fact that the limit of $y(t)$ is $s(t)/|\cos\theta(t)|$ as the radius of curvature $R(\mathbf{r}(t), \mathbf{v}) \to \infty$. We take this approximation when the absolute value of the curvature radius is large due to limitations in numerical precision.

Proof. Since $a(t) + s(t) = R(\mathbf{r}(t), \mathbf{v}) > 0$,

$$\begin{aligned}\lim_{R\to\infty} y &= \lim_{R\to\infty}\left[(a+s)|\cos\theta| - \sqrt{a^2-(a+s)^2\sin^2\theta}\right] \\ &= \lim_{R\to\infty}\left[R|\cos\theta| - \sqrt{a^2-R^2\sin^2\theta}\right] \\ &= \lim_{R\to\infty}\left[\sqrt{R^2\cos^2\theta} - \sqrt{a^2-R^2(1-\cos^2\theta)}\right] \\ &= \lim_{R\to\infty}\left[\sqrt{R^2\cos^2\theta} - \sqrt{R^2\cos^2\theta - R^2 + a^2}\right]\end{aligned}$$

Let $f(x) = \sqrt{x}$,

$$\begin{aligned}&\lim_{R\to\infty} y = \lim_{R\to\infty}\left[\sqrt{R^2\cos^2\theta} - \sqrt{R^2\cos^2\theta - R^2 + a^2}\right] \\ &= \lim_{R\to\infty}\left[f(R^2\cos^2\theta) - f(R^2\cos^2\theta - R^2 + (R-s)^2)\right]\end{aligned}$$

Without loss of generality, assuming that $s > 0$, we know $(R-s)^2 < R^2$. Then we get,

$$R^2\cos^2\theta > R^2\cos^2\theta - R^2 + (R-s)^2.$$

Assuming that there is a ξ suffice the inequality that

$$R^2\cos^2\theta - R^2 + (R-s)^2 < \xi < R^2\cos^2\theta$$

According to Lagrange's mean value theorem, we get

$$\begin{aligned}&f(R^2\cos^2\theta) - f(R^2\cos^2\theta - R^2 + (R-s)^2) \\ &= f'(\xi)(R^2 - (R-s)^2) \\ &= f'(\xi)(2Rs - s^2)\end{aligned}$$

Therefore we know that,

$$\begin{aligned}&\lim_{R\to\infty}\left[f(R^2\cos^2\theta) - f(R^2\cos^2\theta - R^2 + (R-s)^2)\right] \\ &= \lim_{R\to\infty} f'(\xi)(2Rs - s^2) = \frac{2Rs - s^2}{2\sqrt{\xi}}\end{aligned}$$

Since we know that

$$\begin{aligned}&R^2\cos^2\theta - R^2 + (R-s)^2 \\ &= R^2\cos^2\theta - 2Rs + s^2 < \xi < R^2\cos^2\theta\end{aligned}$$

Then ξ can be denoted as $R^2cos^2\theta + o(R^2)$. Here $o(R^2)$ is the lower order infinity of R^2.

Finally, we prove the limit of y approaches $s/|cos\theta|$, which is the same as TUVR [1].

$$\begin{aligned}\lim_{R\to\infty} y &= \lim_{R\to\infty}\frac{2Rs - s^2}{2\sqrt{R^2\cos^2\theta + o(R^2)}} \\ &= \lim_{R\to\infty}\frac{Rs}{\sqrt{R^2\cos^2\theta + o(R^2)}} \\ &= \frac{s}{\sqrt{\cos^2\theta + \dfrac{o(R^2)}{R^2}}} = \frac{s}{|\cos\theta|}\end{aligned}$$

The proof process is very similar when $s < 0$ and y also converge to $s/|\cos\theta|$ in this situation. □

2 Progressively Warm-up of Bias-aware SDF to Density Transformation

As aforementioned before, we design the bias-aware SDF to density transformation in a progressive manner to ensure the stabilization of training. In this subsection, we conduct quantitative and qualitative experiments to validate the efficiency of this design. As shown in Table. 1, the "Non-Prog" indicates applying the bias-aware SDF to density

- *Yuting Xiao and Jingwei Xu contributed equally to this work;*
- *Corresponding Author: Shenghua Gao; E-mail: gaoshh@shanghaitech.edu.cn*
- *Yuting Xiao, Jingwei Xu and Shenghua Gao are with the School of Information Science and Technology, ShanghaiTech University, Shanghai 201210, China;*
- *Zehao Yu is with the Department of Computer Science, University of Tübingen, 72076, Germany;*

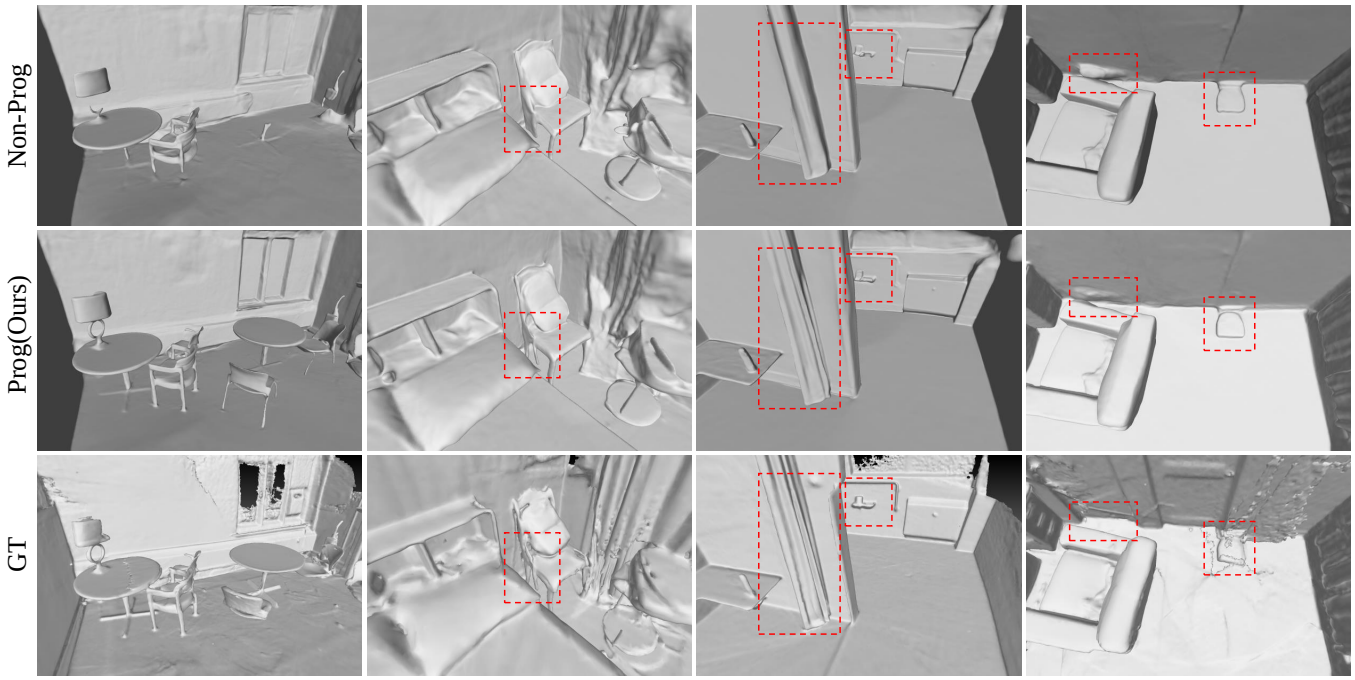

Fig. 1: The qualitative experiments to evaluate the efficiency of the progressive manner in our bias-aware SDF to density transformation (Sec. 3.5). The "Non-Prog" indicates applying the bias-aware SDF to density transformation without progressive growing. The progressive manner can make the optimization of thin and detailed structures more stable.

TABLE 1: The quantitative results on the ScaNet dataset. The "Non-Prog" indicates applying the bias-aware SDF to density transformation without progressive warm-up, and the "Prog" indicates our proposed model.

Metrics	Chamfer↓	F-score↑	Normal C↑
Non-Prog	0.042	73.19	89.85
Prog	**0.038**	**78.54**	**90.21**

transformation without progressive warm-up. It can be observed that this progressive manner achieves significant improvement by 5.35 F-score.

The qualitative results are shown in Fig. 1. Though applying the bias-aware SDF to density transformation without the progressive manner can still reconstruct the textureless region well, the reconstruction of detailed and small structures seriously degraded, such as the tables, chairs, and the curtain. These experimental results prove that the designed progressive manner can make the training robust.

3 Evaluating Gradient Detach for Depth and Normal Prior

We conduct ablation studies on the masked depth loss and masked normal loss. The masked depth and normal loss can filter the predicted geometry prior, which makes the training phase stable since the monocular geometry priors are not absolutely accurate. So we investigate the effect of masked depth and normal loss in this section. We disable the adaptive gradient detach operation based on the uncertainty map to evaluate the efficiency of our proposed masked geometry loss function.

TABLE 2: The quantitative results show that the mask-guided gradient detaching design can slightly improve the performance because this design mainly influences the thin and detailed structure, whose proportion in each image is not significant.

Metrics	Chamfer↓	F-score↑	Normal C↑
W/o mask detach	0.0385	78.31	90.17
Ours	**0.0382**	**78.54**	**90.21**

The quantitative and qualitative results are shown in Table 2 and Fig. 2, respectively. It can be observed that, though applying the mask-guided gradient detaching loss can only achieve numerical improvement, the qualitative improvements are obvious, especially for those thin and small structures. The reason is that this design mainly aims to improve the reconstruction of detailed regions, whose uncertainty score is high. And only occupying a small area results in little impact on the numerical performance. The experimental results indicate the effectiveness of our proposed approach.

4 Time Consumption

The time consumed in the training phase of each method is shown in the Table. 3. Though the implementation of MonoSDF [2] and our method are both based on the VolSDF [3], our method consumes more time for training. The reason is that our model needs to compute the SDF gradient and uncertainty score of each point at the point sampling for the opacity estimation. Apply other point sampling algorithms, such as the sampling based on proposal MLP [4], may eliminate this negative impact.

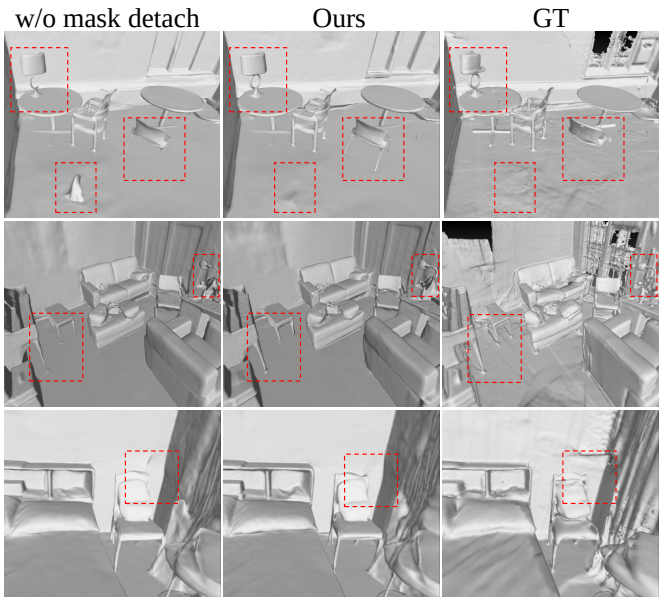

Fig. 2: The ablation studies for evaluating the effectiveness of detaching the gradients at regions with high uncertainty scores. It can be observed that detaching the gradients of geometry with high uncertainty can avoid losing and damaging some thin and detailed surfaces.

TABLE 3: The time-consuming of different methods. All methods are implemented on NVIDIA Geforce RTX 2080ti. The * indicates applying the hash encoding feature grids.

Methods	NeuRIS	MonoSDF	Ours	MonoSDF*	Ours*
Times (h)	11	26	31	17	21

5 Failure Case

Only applying the reduction of bias and prior filtering still cannot solve the reconstruction defect at reflective regions. We demonstrate the failure case on DTU dataset of our method without geometry prior in Fig. 3.

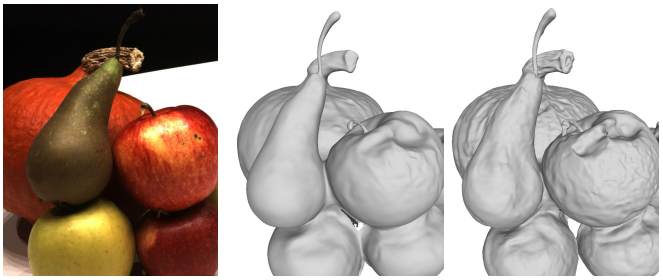

Fig. 3: The visualization of a failure case on the DTU dataset. It demonstrates the reflective regions cannot be reconstructed perfectly.

6 License

We include another 7 scenes published by artists on BlendSwap in numerical evaluation of the ICL-NUIM dataset [5]. The license information is shown in Table. 4. We utilize the synthetic mesh and rendered images of each scene.

Scene	URL	License
Breakfast room	https://blendswap.com/blend/13363	CC-BY
Green room	https://blendswap.com/blend/8381	CC-BY
Grey-white room	https://blendswap.com/blend/13552	CC-BY
ICL living room	https://www.doc.ic.ac.uk/~ahanda/VaFRIC/iclnuim.html	CC-BY
Morning apart.	https://blendswap.com/blend/10350	CC-BY
Kitchen 1	https://blendswap.com/blend/5156	CC-0
Kitchen 2	https://blendswap.com/blend/11801	CC-0
White room	https://blendswap.com/blend/5014	CC-BY

TABLE 4: The license information of each scene in the numerical evaluation of the ICL-NUIM dataset [5] and scenes published by artists.

References

[1] Y. Zhang, Z. Hu, H. Wu, M. Zhao, L. Li, Z. Zou, and C. Fan, "Towards unbiased volume rendering of neural implicit surfaces with geometry priors," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 4359–4368. 1

[2] Z. Yu, S. Peng, M. Niemeyer, T. Sattler, and A. Geiger, "Monosdf: Exploring monocular geometric cues for neural implicit surface reconstruction," in *Advances in Neural Information Processing Systems*. 2

[3] L. Yariv, J. Gu, Y. Kasten, and Y. Lipman, "Volume rendering of neural implicit surfaces," *Advances in Neural Information Processing Systems*, vol. 34, pp. 4805–4815, 2021. 2

[4] J. T. Barron, B. Mildenhall, D. Verbin, P. P. Srinivasan, and P. Hedman, "Mip-nerf 360: Unbounded anti-aliased neural radiance fields," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 5470–5479. 2

[5] A. Handa, T. Whelan, J. McDonald, and A. J. Davison, "A benchmark for rgb-d visual odometry, 3d reconstruction and slam," in *2014 IEEE international conference on Robotics and automation (ICRA)*. IEEE, 2014, pp. 1524–1531. 3