NLOST: Non-Line-of-Sight Imaging with Transformer Supplementary Material

Yue Li Jiayong Peng Juntian Ye Yueyi Zhang Feihu Xu Zhiwei Xiong
University of Science and Technology of China

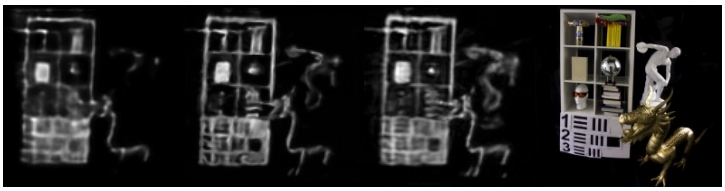
In this supplementary material, we provide more experimental results. Besides, we provide a video showing the reconstructed 3d scenes from the real-world transient measurements captured by [1] and our NLOS imaging system (https://github.com/Depth2World/NLOST).

Training time, inference time, inference memory, and parameters. The detailed information of the deep models is listed below, measured on a workstation with 4 A100 GPUs. Note that NeTF is trained on a single GPU for direct test rendering, and LFE is trained with multi-view supervision.

Method	UNet	LFE	NeTF	Ours
Training / Test	23h / 161ms	10h / 26ms	50h	24h / 61ms
Memory / Params	6G / 36M	5G / 0.03M	23G / 0.04M	20G / 1M

Table 1. Training time, inference time, inference memory, and total parameters of the deep models.

Ablation Study. We conduct the ablation study on shallow-deep feature fusion. The qualitative results under different feature fusion manners tested on real-world data are shown in Fig. 1.



Glo LocGlo Ours Scene

Figure 1. Reconstructed intensity images from a real-world scene under different feature fusion manners. Glo denotes that the network only uses the deep global features and LocGlo denotes that the network uses both deep local and global features.

Real-world Results. More results on the real-world scenarios provided by [1] are shown in Fig. 2. More results on the real-world transient measurements captured by our NLOS imaging system are shown in Fig. 3.

Simulated Results. More reconstructed intensity images and depth error maps from the simulated **Seen** test set are shown in Fig. 4 and Fig. 5. More reconstructed intensity images and depth error maps from the simulated **Unseen** test set are shown in Fig. 6 and Fig. 7.

References

[1] David B Lindell, Gordon Wetzstein, and Matthew O'Toole. Wave-based non-line-of-sight imaging using fast fk migration. *ACM Transactions on Graphics*, 38(4):1–13, 2019. 1, 2

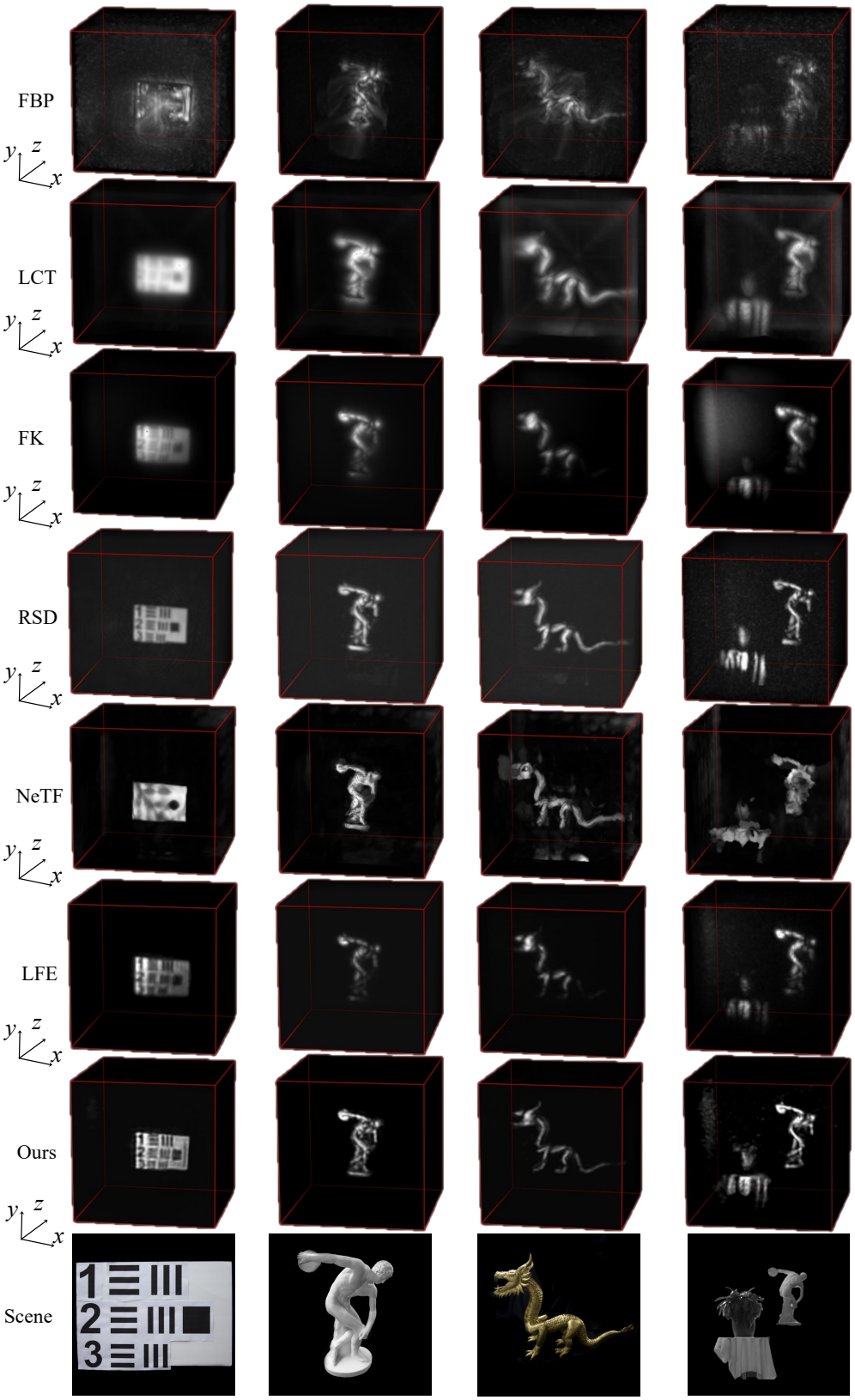


Figure 2. Reconstructed hidden scenes from the public real-world dataset in [1]. Zoom in for details.

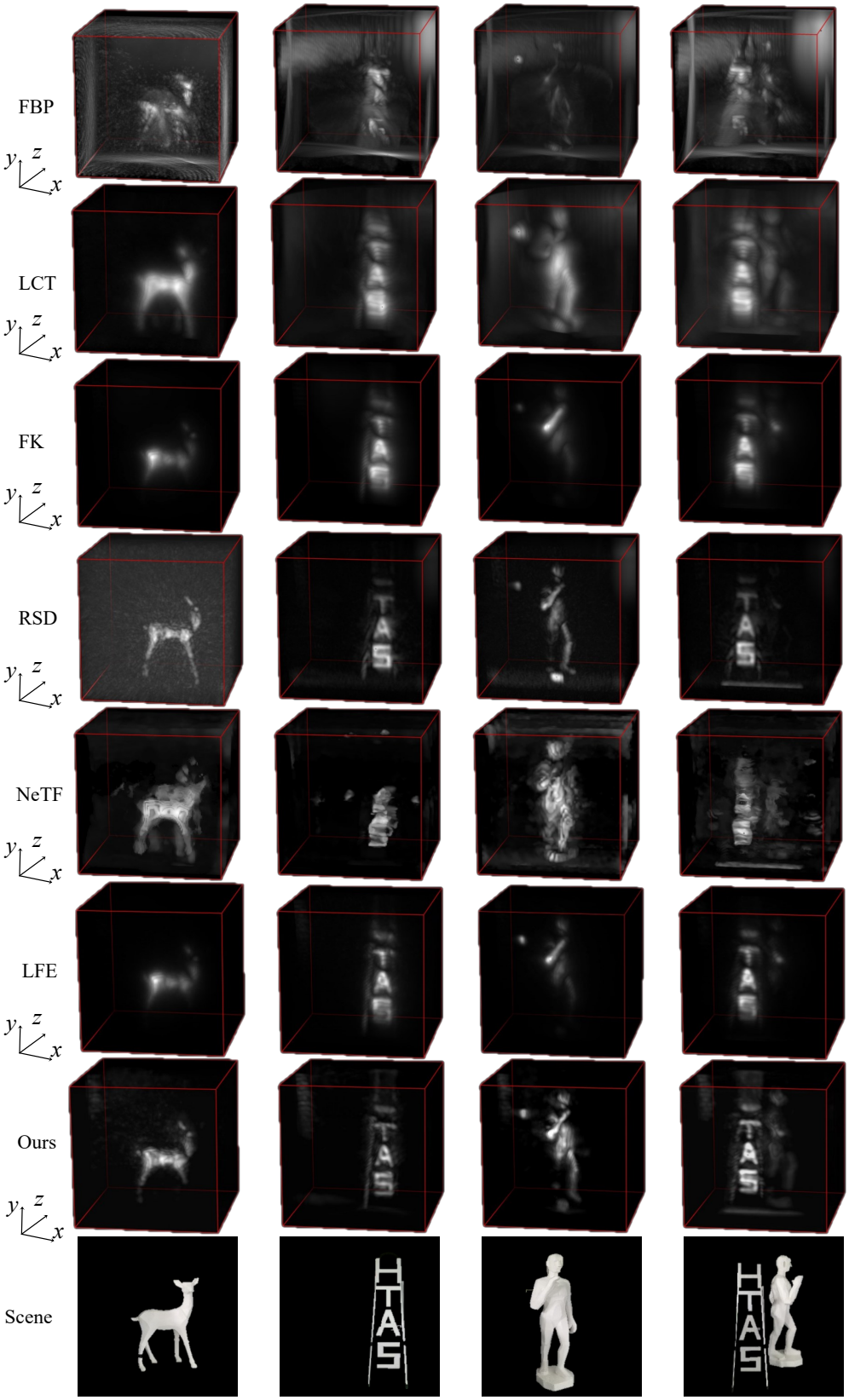


Figure 3. Reconstructed hidden scenes from real-world measurements captured by our NLOS imaging system. Zoom in for details.

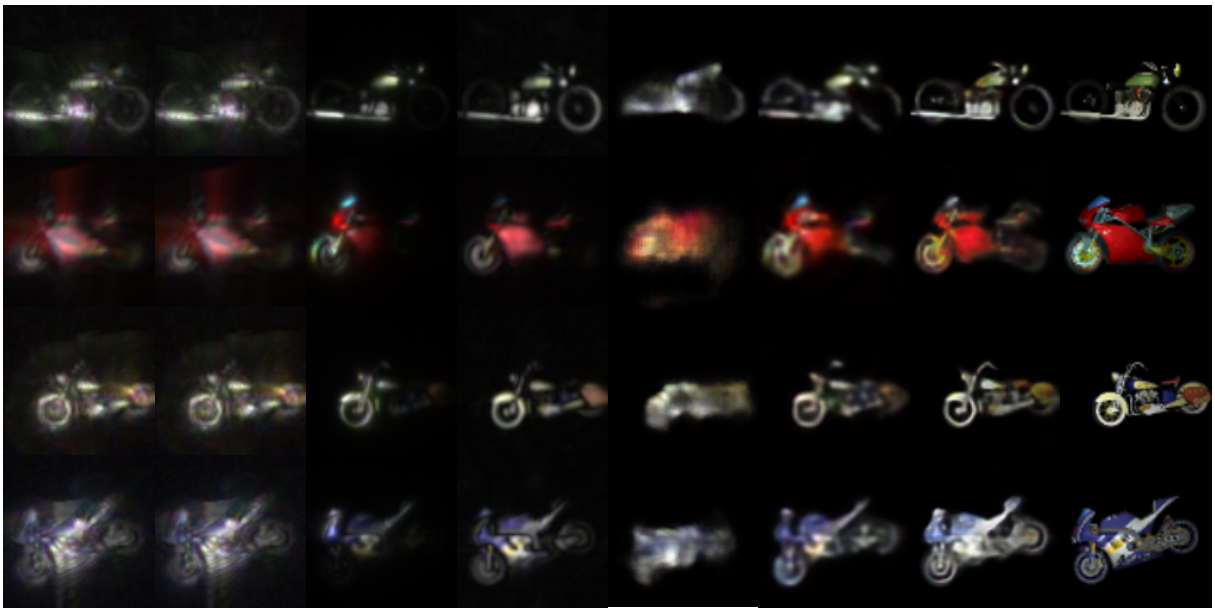


Figure 4. Reconstructed intensity results from **Seen** test set. GT denotes the intensity ground truth.

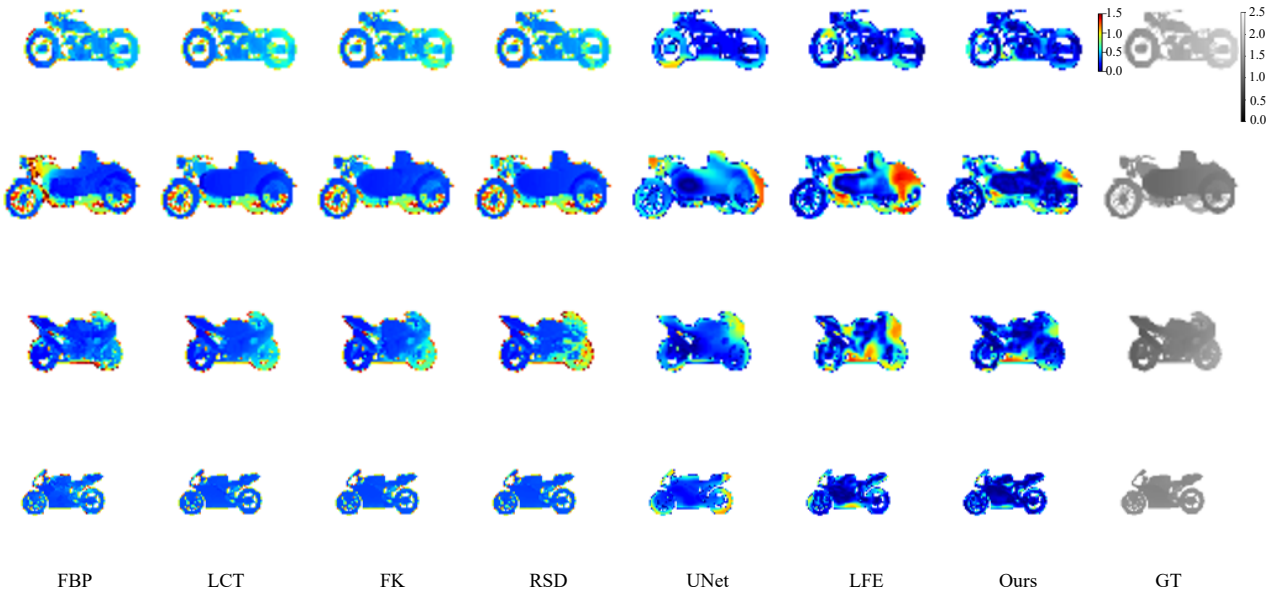


Figure 5. Reconstructed depth error maps from **Seen** test set. GT denotes the depth ground truth. The color bars show the value of depth and the error map.

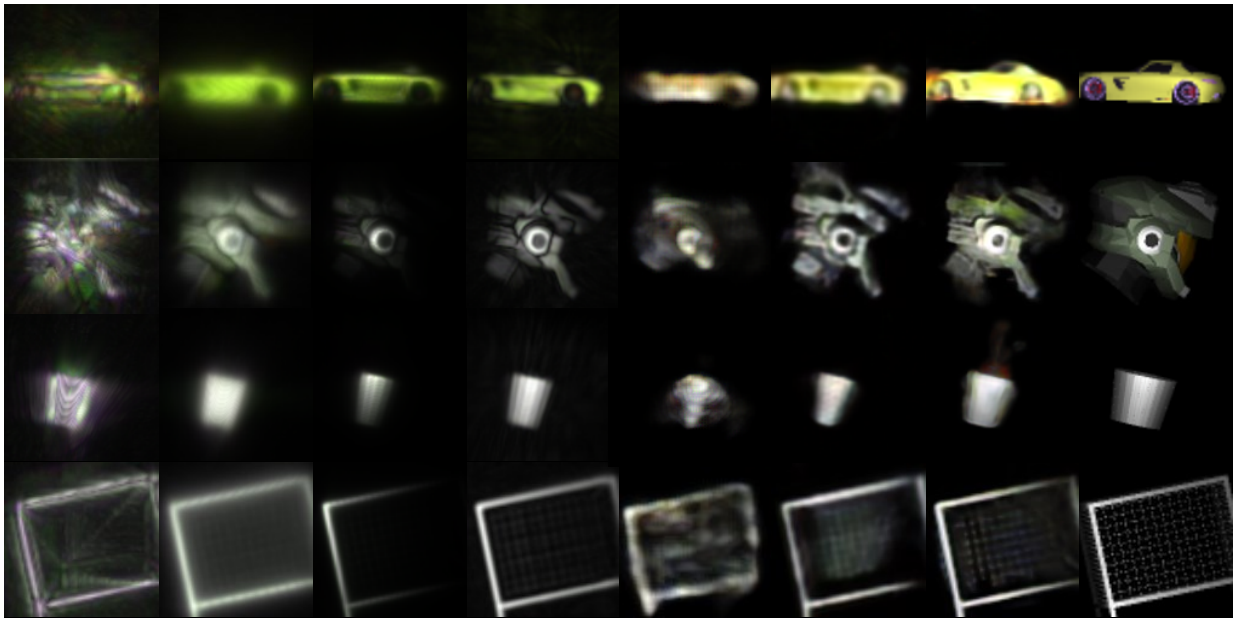


Figure 6. Reconstructed intensity results from **Unseen** test set. GT denotes the intensity ground truth.

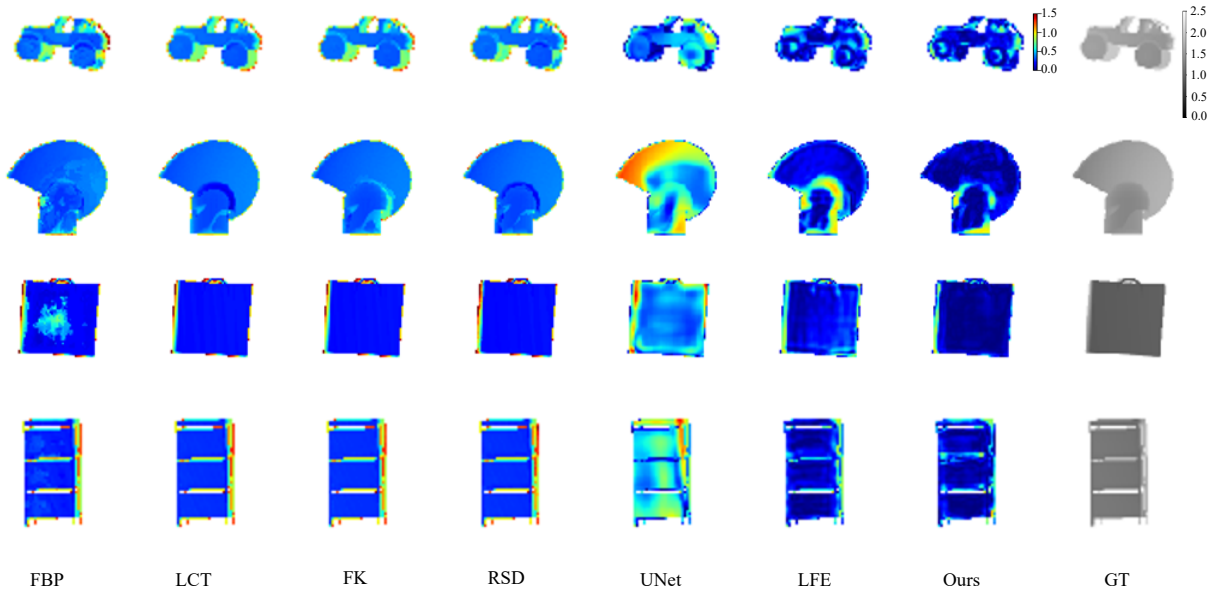


Figure 7. Reconstructed depth error maps from **Unseen** test set. GT denotes the depth ground truth. The color bars show the value of depth and the error map.