Supplementary Material for
LightFormer: Light-Oriented Global Neural Rendering in Dynamic Scene

HAOCHENG REN, State Key Lab of CAD&CG, Zhejiang University

YUCHI HUO, Zhejiang Lab and State Key Lab of CAD&CG, Zhejiang University

YIFAN PENG, University of Hong Kong

HONGTAO SHENG, WEIDONG XUE, HONGXIANG HUANG, JINGZHEN LAN, and RUI WANG, State Key Lab of CAD&CG, Zhejiang University

HUJUN BAO, Zhejiang Lab and State Key Lab of CAD&CG, Zhejiang University

1 NETWORK DETAILS

Figure 3 illustrates the detailed network architectures of each module in our framework. Specifically, there are *five* encoder networks and *two* decoder networks. The shadow encoder employs the UNet architecture, while the others are compact MLP networks. The direct encoder and indirect encoder incorporate a skip connection following AE [Diolatzis et al. 2022] and NeuralGI [Gao et al. 2022]. Additionally, we provide detailed descriptions of the inputs and outputs of each module in Table 2.

2 ADDITIONAL RESULTS

Since some results are cropped for better visual comparison in the main paper, we provide results with full resolution version and error maps in the supplementary material, which are presented in a web-based comparison tool following the denoising works [Áfra 2024]. For more dynamic results and generalization results, please refer to our supplementary video.

To validate our model with challenging scenes featuring multiple freely movable objects, we constructed a scene named *Interior Design*. This scene also demonstrates its modeling-while-rendering application. It consists of 5 pieces of freely movable furniture and 3 area lights. In addition to transformation objects, materials such as

color and roughness of sofa and wall, as well as the emission of light source, can also be edited. We show quantitative results in Table 1 and visual comparison in Figure 1. Similarly, our method can outperform previous neural rendering methods. However, path tracing methods produce more accurate results given the high number of samples per pixel. Dynamic results as well as inter-reflection effects are showcased in the video.

3 DISCUSSION WITH CLASSIC VPL METHOD

Though our method draws inspiration from classic VPL rendering framework, the results are quite different. Thanks to our neural shader, certain artifacts commonly observed in classic VPL method (e.g., instant radiosity [Keller 1997]) are not shown in our results. A visual comparison in our *Chess-Luminaire* scene is illustrated in Figure 2. We utilized the implementation of VPL rendering in the Mitsuba renderer [Jakob 2014], with timing measured on RTX 4090.

As can be seen, Instant radiosity (IR) introduces visual artifacts such as bright blotches. The application of clamping can mitigate these artifacts, but introducing significant energy loss. This is evident in the dark areas, especially on the surface of the luminary. As shown in the second row, a large number of VPLs are required to achieve visually acceptable results. leading to high timing costs. In contrast, our method do not suffer from these artifacts and can achieve real-time performance.

Authors' addresses: Haocheng Ren, State Key Lab of CAD&CG, Zhejiang University, swordigo@zju.edu.cn; Yuchi Huo, Zhejiang Lab and State Key Lab of CAD&CG, Zhejiang University, huo.yuchi.sc@gmail.com; Yifan Peng, University of Hong Kong, evanpeng@hku.edu.hk; Hongtao Sheng, 22251006@zju.edu.cn; Weidong Xue, 22251027@zju.edu.cn; Hongxiang Huang, crisprhhx@zju.edu.cn; Jingzhen Lan, lanjz@zju.edu.cn; Rui Wang, State Key Lab of CAD&CG, Zhejiang University, ruiwang@zju.edu.cn; Hujun Bao, Zhejiang Lab and State Key Lab of CAD&CG, Zhejiang University, bao@cad.zju.edu.cn.

Table 1. Quantitative results of *Interior Design* scene. Details of comparison methods can be found in the Table. 1 of the main paper.

Methods	*Interior Design* Time (ms) ↓	LPIPS ↓	RSE ↓
CNSR [1]	155.74	0.1325	0.0807
AE [2]	120.93	0.0687	0.0303
Ours	107.72	0.0583	0.0070
ONND [3]	38.09 (66spp)	0.0470	0.0124
OIDN [4]	37.99 (57spp)	0.0440	0.0100
Ours (TRT)	38.19	0.0583	0.0070

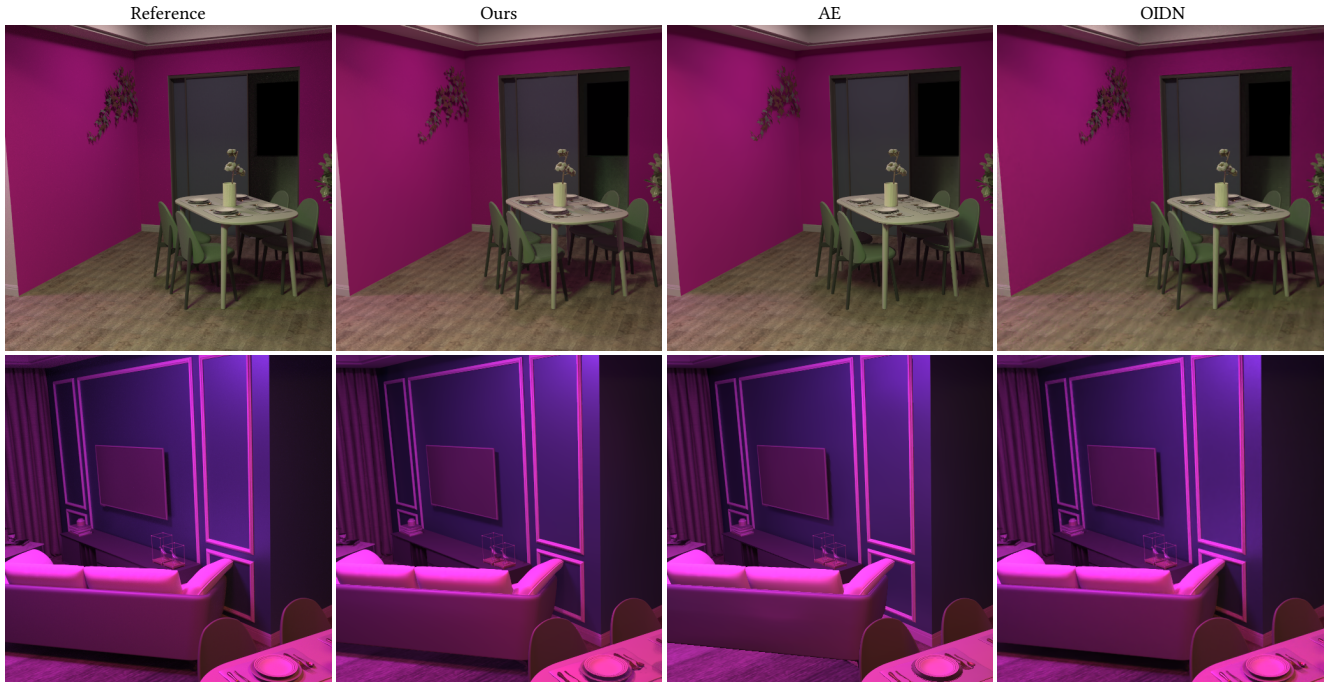| Reference	Ours	AE	OIDN

Fig. 1. Visual comparison results of *Interior Design* scene. Our method can generate visually plausible global illumination results, even with multiple freely movable objects. The high-frequency shading details are hard to be captured in previous neural rendering methods.
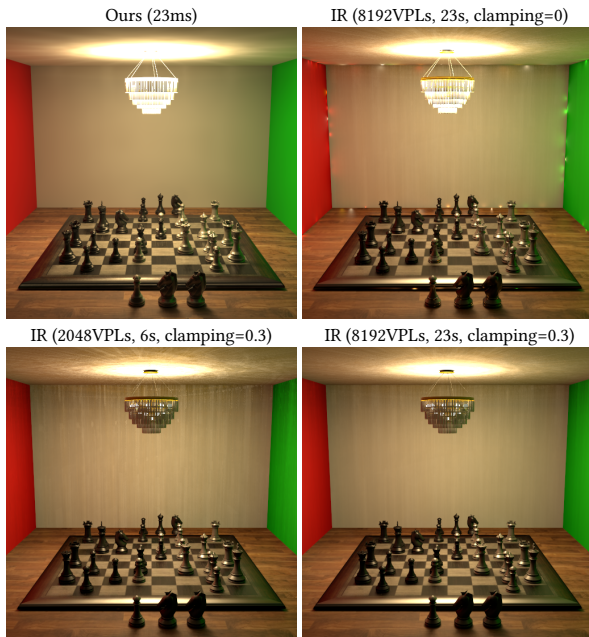

Fig. 2. Visual comparison with classic VPL rendering method (Instant Radiosity) and our method. Our result do not suffer from light leak (bright blotches) and the visible energy loss (dark areas) caused by clamping.

Stavros Diolatzis, Julien Philip, and George Drettakis. 2022. Active exploration for neural global illumination of variable scenes. *ACM Transactions on Graphics (TOG)* 41, 5 (2022), 1–18.

Duan Gao, Haoyuan Mu, and Kun Xu. 2022. Neural Global Illumination: Interactive Indirect Illumination Prediction under Dynamic Area Lights. *IEEE Transactions on Visualization and Computer Graphics* (2022).

Wenzel Jakob. 2014. Mitsuba documentation. *Version 0.5. 0* (2014).

Alexander Keller. 1997. Instant radiosity. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*. 49–56.

REFERENCES

Attila T. Áfra. 2024. Intel® Open Image Denoise. https://www.openimagedenoise.org.
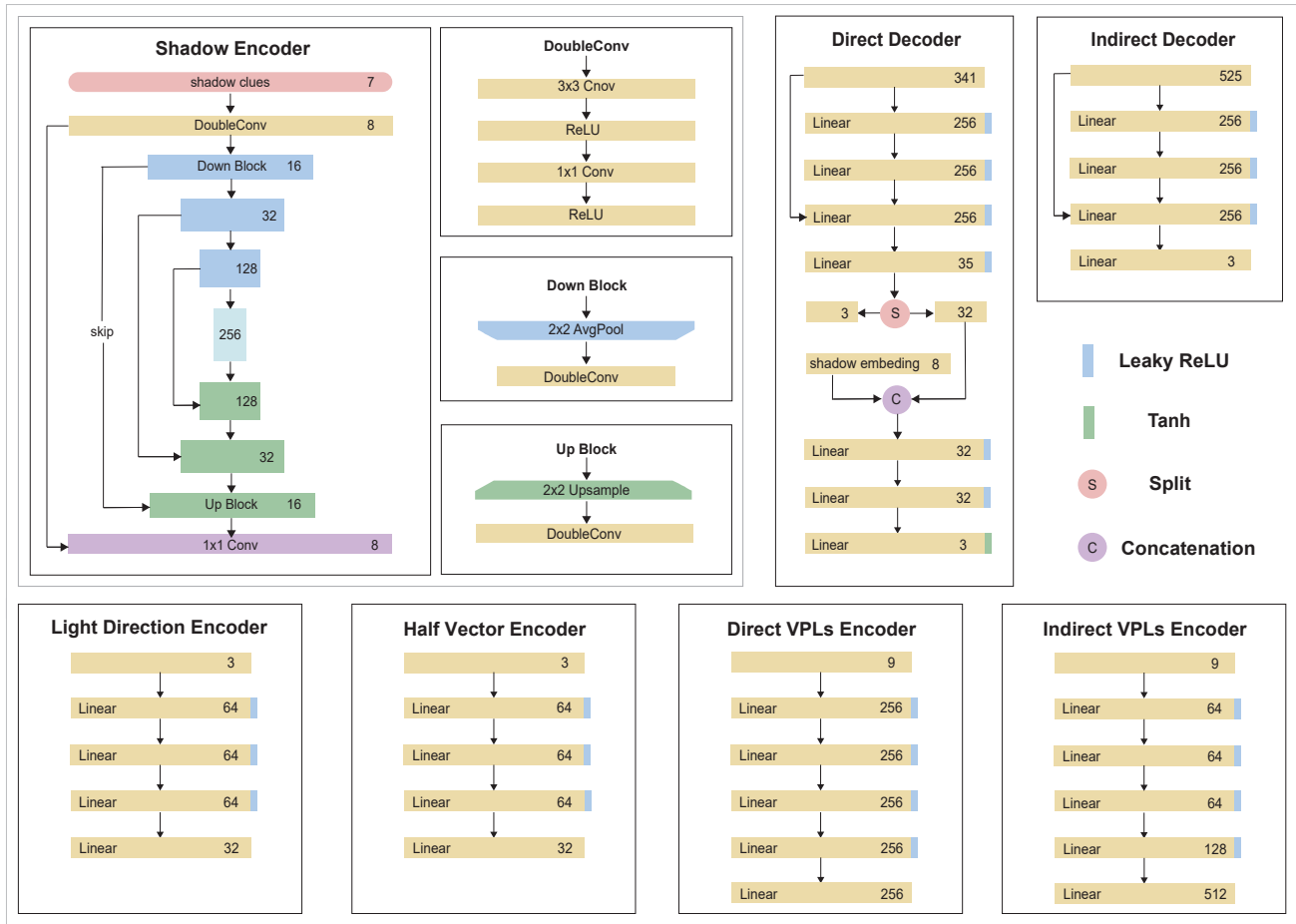
Fig. 3. The architectures of encoder and decoder networks of *LightFormer*.

Table 2. Detailed inputs and outputs of our networks. The number in the bracket indicates image channels or the length of feature vector.

Network	Input	Output
Direct VPLs Encoder	position(3) + normal(3) + power(3)	direct light embedding(256)
Indirect VPLs Encoder	position(3) + normal(3) + flux(3)	indirect light embedding(512)
Light Direction Encoder	light direction(3)	light direction embedding(32)
Half Vector Encoder	half vector(3)	half vector embedding(32)
Shadow Encoder	shadow clues($(d - d_f)$ (1), (d / d_f) (1), c_e(1), c_c(1), position(3))	shadow embedding(8)
Direct Decoder	GBuffers(position(3) + normal(3) + albedo(3) + specular(3) + roughness(1)) + composed direct light embedding(256) + composed half vector embedding(32) + composed light direction embedding(32) + composed shadow embedding(8)	direct shading(3) , direct shadow(3)
Indirect Decoder	GBuffers(position(3) + normal(3) + albedo(3) + specular(3) + roughness(1)) + composed indirect light embedding(512)	indirect shading(3)