Supplementary Material for ReLeaPS : Reinforcement Learning-based Illumination Planning for Generalized Photometric Stereo

Jun Hoong Chan[1,2,†] Bohan Yu[1,2,†] Heng Guo[3,4] Jieji Ren[5] Zongqing Lu[1,2] Boxin Shi[1,2,‡]

[1] National Key Laboratory for Multimedia Information Processing, School of Computer Science, Peking University
[2] National Engineering Research Center of Visual Technology, School of Computer Science, Peking University
[3] School of Artificial Intelligence, Beijing University of Posts and Telecommunications [4] Osaka University
[5] School of Mechanical Engineering, Shanghai Jiao Tong University
junhoong95@stu.pku.edu.cn, {ybh1998, zongqing.lu, shiboxin}@pku.edu.cn,
heng.guo@ist.osaka-u.ac.jp, jiejiren@sjtu.edu.cn

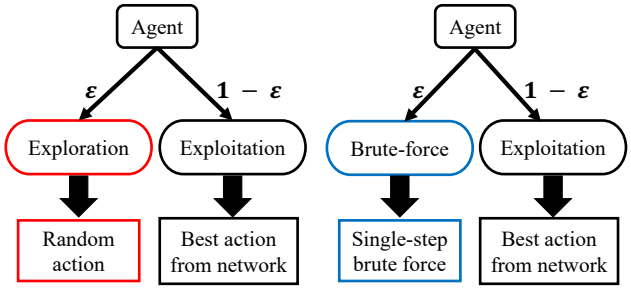
The following materials are provided in this supplementary file:

- We first present our proposed strategy to facilitate exploration and assist in network training in Section 4.2 of the main paper.
- We describe more details about synthetic rendering and real-world equipment setup in Section 4.3 of the main paper.
- We provide real-world experimental results in addition to Section 5.3 of the main paper.
- We elaborate more details about the ablation studies in Section 5.4 of the main paper.
- We conduct an experiment to compare ReLeaPS with other illumination planning approaches using 10 lights with CNN-PS backbone.
- We discuss other properties of ReLeaPS, including time efficiency, limited number of lights, reward design for external factors, and adaptation to different PS backbones.
- We capture a short video clip that illustrates the online illumination planning process.

7. Appendix

7.1. Epsilon One-Step Brute-force Exploration

Epsilon greedy exploration strategy is the common exploration strategy in reinforcement learning (RL). As shown in Fig. 6, the agent explores the environment with a random action of probability epsilon ϵ and the best-known action with probability $1-\epsilon$. However, this random exploration can be inefficient and time-consuming which requires a lot of exploration samplings to obtain a plausible result.



(a) Epsilon greedy exploration (b) Epsilon brute-force exploration

Figure 6. Differences between (a) Epsilon-greedy and (b) Epsilon Brute-Force strategy based on the prior from the photometric stereo [9].

We propose an alternative approach to replace the random exploration with a one-step brute-force method. Specifically, given the state $\mathcal{S}_t = \{\mathcal{I}_t, \boldsymbol{L}_t\}$ and F light directions for the candidate action $\mathcal{A}_{t+1}$, we brute-forcely try one out of F candidate lights and the corresponding image observation as $\{\boldsymbol{l}', \boldsymbol{I}'\}$ to build a new state $\mathcal{S}_{t+1} = \{\mathcal{I}_{t+1}, \boldsymbol{L}_{t+1}\}$. Under this trial, we apply a Lambertian photometric stereo (*i.e.*, the least square) method [9] as the function $\mathrm{PS}(\mathcal{I}_{t+1}, \boldsymbol{L}_{t+1})$ for computing the surface normal $\boldsymbol{N}$ efficiently and compare with ground truth normal $\tilde{\boldsymbol{N}}$. If the angular error by adding $\{\boldsymbol{l}', \boldsymbol{I}'\}$ is the smallest over all of the F brute-force trials, we set $\boldsymbol{l}'$ as the optimal RL action in the next single step.

We conduct the one-step brute-force exploration with

Project page: https://jhchan0805.github.io/ReLeaPS
† Equal contributions. ‡ Corresponding author.

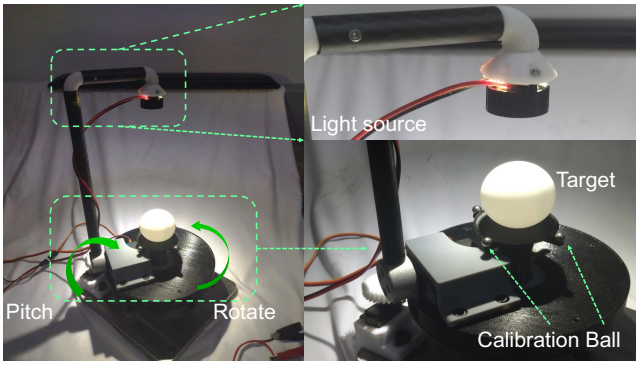


Figure 7. Our illumination equipment. It has two axes to control the pitch and axial rotation angularly. A high-power LED is mounted on the robot arm, which can move into any location of the upper hemisphere, and a convex lens is placed in front of the LED to provide distant illumination. Four shiny balls are placed around the target for calibrating the light directions.

probability epsilon ϵ. As the training step increases, we gradually reduce the ϵ probability so that the agent chooses the light direction learned from the network. Compared to random exploration, this approach provides a more efficient exploration strategy by using the photometric stereo solution [9].

7.2. Detail of Rendering the Synthetic Dataset

In this paper, our training involves examining the complex light transport phenomena that occur in real-world scenarios, including inter-reflection, specular highlights, and cast shadows. To create a training dataset that accurately reflects these phenomena, we employ a physically-based ray-tracing renderer to generate a synthetic dataset with the Cook-Torrance [1] reflectance model while also considering global illumination effects, such as cast shadows and inter-reflections. Through the use of this physically-based synthetic dataset, we are able to comprehensively address generalized photometric stereo and develop an improved illumination planning strategy during the implementation of RL techniques.

7.3. Equipment Setup for Real-World Experiment

As shown in Fig. 7, we build an illumination equipment with a concentric equidistant design that allows for omnidirectional illumination and a moveable light source to evaluate the effectiveness of ReLeaPS. The concentric design ensures the light source is located on top of a hemisphere surface, whereas the light is placed opposite to the camera's field of view (*i.e.*, negative Z-axis in the general case).

Our illumination equipment incorporates two rotation axes, enabling pitch and axial rotations of the light source to emit light toward the target from any direction sampled on the visible hemisphere with high precision control of angular rotation. To approximate distant light in photometric stereo, we attach a focus lens in front of the LED light source and put it 50 cm far from the target object.

Figure 8. Objects captured by our real-world equipment: (a) LIONHEAD, (b) MONSTER, and (c) TWOFISHES.

The DaHeng Image MER-503-36U3C camera with a 50 mm lens and an original resolution of 2448×2048 is utilized to capture observed images of the object, which are saved in a raw format (linear radiometric response) with a cropped resolution of 1001×1001. The camera is placed 50 cm far from the target object. The camera's exposure time is adjusted from 1 ms to 4 ms depending on the material's reflectivity to obtain a high dynamic range (HDR) image observation under each light direction.

7.4. Qualitative Comparison on Real Dataset

Based on the setup introduced in Section 7.3, we capture three objects, LIONHEAD, MONSTER, and TWOFISHES as shown in Fig. 8. In Fig. 9, we present the estimated surface normals and angular error maps for comparing our proposed method with other illumination planning methods [2, 7].

As described in Section 5.3 using the LIONHEAD data in the main paper, our method yields superior results compared to other illumination planning methods for generalized photometric stereo. In this supplementary file, we further compare the performance of these methods on the object MONSTER and TWOFISHES, which exhibit gaps on their surfaces. From Fig. 9, it is evident that DC05 [2] fails to properly recover the gaps on the object surfaces, resulting in a large angular error. Although TK22 [7] can recover the gaps on the surface, our RL-based illumination planning approach achieves a lower mean angular error than TK22 [7].

7.5. Ablation Studies

This section provides ablation studies to analyze the individual contributions of these components (*i.e.*, non-Lambertian reflectance, sparse-to-dense reward strategy, and dualling DQN network design in our online illumination planning method) to the overall performance of ReLeaPS for illumination planning in generalized photometric stereo. The evaluation settings and testing set are consistent with the ablation analysis in Section 4 of the main paper.

ReLeaPS w/ Lambert. The conventional illumination planning methods [2, 7] in photometric stereo is limited to Lambertian surfaces and rely on simplistic analytical assumptions. In contrast, our proposed approach provides a

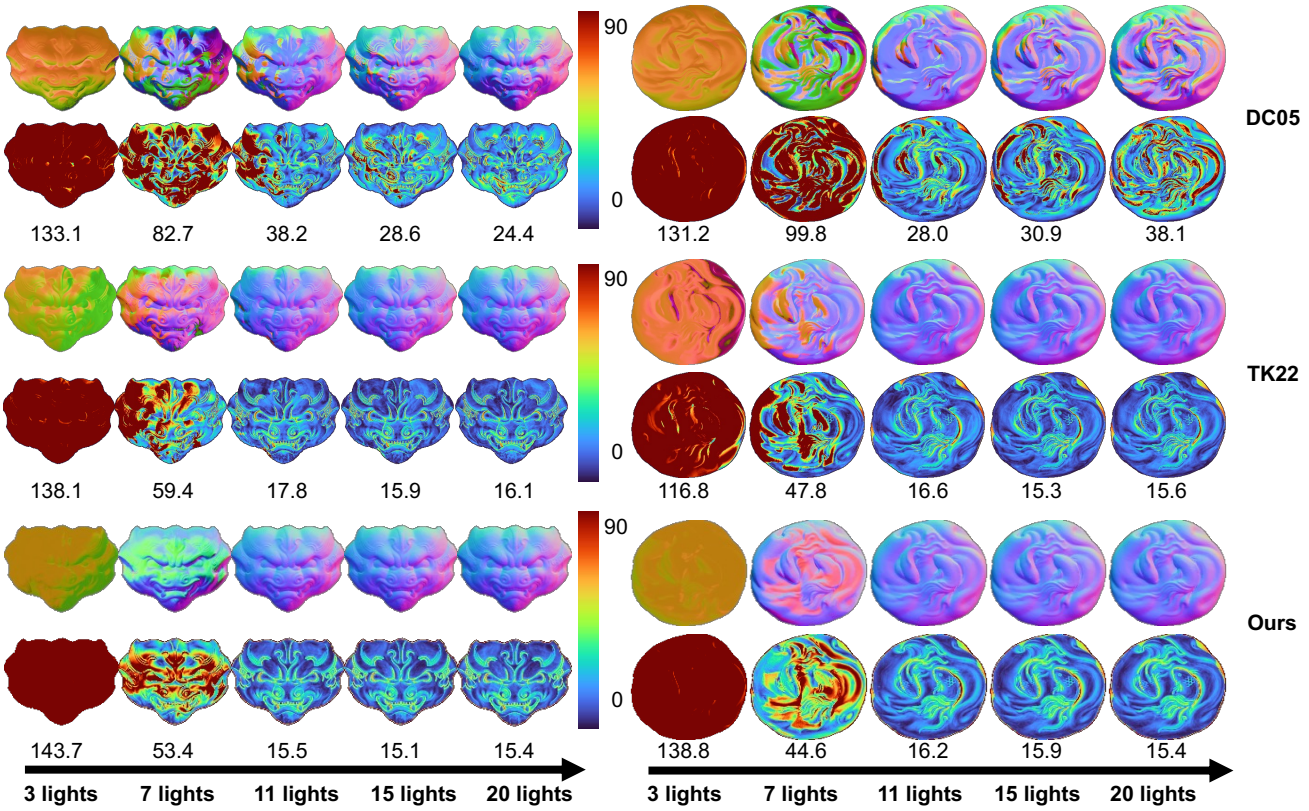


Figure 9. Qualitative comparison of recovered surface normals and error maps for **(left)** MONSTER and **(right)** TWOFISHES captured using our equipment under different illumination planning methods (*i.e.*, DC05 [2], TK22 [7], and Ours) with increasing light directions (3, 7, 11, 15, 20 lights) in CNN-PS [3] backbone. For each error map, we show the mean angular errors in degrees. The red circle indicates a region with shadows that cannot be effectively recovered by CNN-PS [3], resulting in a large angular error.

novel solution to more complex challenges, such as non-Lambertian surfaces and global illumination effects with complex light transport phenomena. To demonstrate the effectiveness of ReLeaPS trained with generalized image formation under such conditions, we compare it with a version of ReLeaPS trained with Lambertian properties, which is denoted as 'ReLeaPS w/ Lambert'.

ReLeaPS w/o reward. RL is a learning approach that relies on trial and error and is guided by a reward signal. The reward function we use for illumination planning in photometric stereo is based on the mean angular error at the final step of each episode. However, it is important to acknowledge that this particular reward function is sparse in nature. To enhance the network training process, we devise a sparse-to-dense reward design and apply reward-shaping methods as described in Section 4.1 of the main paper. However, it is possible that the RL agent may achieve satisfactory performance without the need for a dense reward design. To verify the effectiveness of this reward design, we conduct the ablation study and denote the method with sparse reward as 'ReLeaPS w/o reward'.

Table 4. Ablation studies on ReLeaPS with 20 lights on Blobby [4], Sculpture [8], DiLiGenT [6], and DiLiGenT10^2 [5] datasets using the LS [9] backbone: including generalized image formation, reward design, and network design (simplified to 'ReLeaPS w/ Lambert', 'ReLeaPS w/o reward', and 'ReLeaPS w/ MLP' respectively).

Dataset	ReLeaPS w/ Lambert	ReLeaPS w/o reward	ReLeaPS w/ MLP	Ours
Blobby [4]	3.7	3.5	3.3	**2.9**
Sculpture [8]	12.9	11.9	10.6	**9.7**
DiLiGenT [6]	14.8	13.9	13.9	**13.8**
DiLiGenT10^2[5]	24.2	24.1	24.5	**23.5**

ReLeaPS w/ MLP. In order to address the challenges posed by non-Lambertian reflectances and global illumination effects, *i.e.*, shadows and inter-reflections in illumination planning, we develop a specialized dense network architecture incorporating dueling deep Q-network (DQN) as our agent network to predict the optimal light direction based on highest Q-value. Nevertheless, other neural network architectures may also be effective in handling this generalized photometric stereo. To validate the effectiveness of our dense dueling DQN, we compare its performance with a simple multi-layer perceptron network,

Table 5. Evaluation of ReLeaPS in DiLiGenT [6] dataset with 10 lights using the CNNPS [3] backbone.

Dataset	Rnd.	DC05	TK22	Ours	All
Blobby	33.0	37.4	16.0	**8.2**	2.3
Sculpture	40.6	48.9	27.0	**12.4**	5.2
DiLiGenT	15.3	14.9	17.8	**13.0**	7.4
DiLiGenT10^2	33.3	32.7	22.9	**21.0**	16.4

referred to as 'ReLeaPS w/ MLP'. The network initially concatenates a sequence of $T - t$ zero images to the input. This approach ensures that a fixed number of T input images are available for all steps, which keeps the input shape consistent. Next, two fully connected layers are employed, each consisting of 256 hidden units. To avoid overparameterization, the input images and the output Q-value map are downsampled by a factor of 4.

To summarize, Table 4 shows that our proposed ReLeaPS architecture demonstrates superior performance compared to 'ReLeaPS w/ Lambert', 'ReLeaPS w/o reward', and 'ReLeaPS w/ MLP' in addressing the challenges posed by non-Lambertian surfaces and global illumination effects in illumination planning. This highlights the effectiveness of our specially designed network architecture, which incorporates a self-designed reward function integrated with an RL framework for illumination planning in generalized photometric stereo.

7.6. Experiment with Ten Lights

From Fig. 4 in the main paper, the performance of ReLeaPS on DiLiGenT decreases compared to other illumination planning methods, especially as the number of lights increases or when using more lights. We suspect that this is due to the narrow distribution of light candidates in the DiLiGenT [6] dataset, which diminishes the advantage of ReLeaPS. We further evaluate Table 1 in the main paper with 10 lights to demonstrate the effectiveness of ReLeaPS compared to different illumination planning approaches with fewer lights. As shown in Table 5, ReLeaPS with CNN-PS [3] as a baseline brings a more significant performance boost over other approaches.

7.7. Discussion

This section mainly discusses other properties of ReLeaPS, including the time efficiency, the limited number of lights, reward design for external factors, and the adaptation to different PS backbones.

Time efficiency. ReLeaPS predicts an optimal light direction around 0.6s and the time required to optimize and capture 20 lights using ReLeaPS is shorter than capturing 100 lights, where the time spent on data actuation is much longer than the light prediction.

Limited light numbers in ReLeaPS. Although ReLeaPS restricts the light selection to 100 discrete candidates, it can easily adopt more light candidates to approximate a continuous light space and integrate with our setup for high precision control (*e.g.*, the number of light candidates can be increased by fourfold by adding another convolution layer, resulting in an approximate 5% and 0.1% increase in inference time and memory usage, respectively).

Reward design for external factor. Reward design of ReLeaPS is based on the mean angular error, which takes synthetic images as input while taking into account various distortions encountered in real observations, such as image noise, tone-mapping, saturation, and 8-bit quantization. Hence, ReLeaPS is intrinsically trained to be robust against these external influences.

Adaptation to different PS backbone. ReLeaPS is adaptable to different PS backbones; however, retraining on a new PS backbone is necessary due to varying PS models and assumptions. This results in adaptive changes in optimized light directions. As a trade-off, ReLeaPS may not be suitable for time-consuming PS approaches.

References

[1] Robert L Cook and Kenneth E. Torrance. A reflectance model for computer graphics. *ACM Transactions on Graphics (ToG)*, 1(1):7–24, 1982. 2

[2] Ondrej Drbohlav and Mike Chantler. On optimal light configurations in photometric stereo. In *Proc. of International Conference on Computer Vision*, 2005. 2, 3

[3] Satoshi Ikehata. CNN-PS: CNN-based photometric stereo for general non-convex surfaces. In *Proc. of European Conference on Computer Vision*, 2018. 3, 4

[4] Micah K Johnson and Edward H Adelson. Shape estimation in natural illumination. In *Proc. of Computer Vision and Pattern Recognition*, 2011. 3

[5] Jieji Ren, Feishi Wang, Jiahao Zhang, Qian Zheng, Mingjun Ren, and Boxin Shi. *DiLiGenT*10^2: A photometric stereo benchmark dataset with controlled shape and material variation. In *Proc. of Computer Vision and Pattern Recognition*, 2022. 3

[6] Boxin Shi, Zhe Wu, Zhipeng Mo, Dinglong Duan, Sai-Kit Yeung, and Ping Tan. A benchmark dataset and evaluation for non-lambertian and uncalibrated photometric stereo. In *Proc. of Computer Vision and Pattern Recognition*, 2016. 3, 4

[7] Hirochika Tanikawa, Ryo Kawahara, and Takahiro Okabe. Online illumination planning for shadow-robust photometric stereo. In *International Workshop on Frontiers of Computer Vision*, 2022. 2, 3

[8] Olivia Wiles and Andrew Zisserman. SILNet: Single-and multi-view reconstruction by learning from silhouettes. *arXiv preprint arXiv:1711.07888*, 2017. 3

[9] Robert J Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19:139–144, 01 1980. 1, 2, 3