Self-calibrating Deep Photometric Stereo Networks
Supplementary Materials

Guanying Chen[1] Kai Han[2] Boxin Shi[3,4] Yasuyuki Matsushita[5] Kwan-Yee K. Wong[1]
[1]The University of Hong Kong [2]University of Oxford
[3]Peking University [4]Peng Cheng Laboratory [5]Osaka University

Contents

1. Regression Based Lighting Estimation Model LCNet$_{\text{reg}}$

Given multiple input images, one straightforward idea for lighting estimation is to regress the exact light direction vectors and intensity values. We examined a regression based counterpart of our LCNet, denoted as LCNet$_{\text{reg}}$, which shares the same architecture with LCNet, except that LCNet$_{\text{reg}}$ estimates a 3-vector for light direction and a scalar value for light intensity, rather than the softmax probability vectors. Given q images, the loss function for the lighting regression is

$$\mathcal{L}_{\text{Reg}} = \lambda_l \frac{1}{q} \sum_{i}^{q} (1 - \boldsymbol{l}_i^\top \tilde{\boldsymbol{l}}_i) + \lambda_e \frac{1}{q} \sum_{i}^{q} (e_i - \tilde{e}_i)^2, \tag{1}$$

where λ_l and λ_e are the weighting factors for the loss terms, $\boldsymbol{l}_i$ (e_i) and $\tilde{\boldsymbol{l}}_i$ ($\tilde{e}_i$) denote the predicted light direction (intensity) and the ground truth, respectively, for image i. During training, λ_l and λ_e are set to 1 (we found that using other weighting factors have similar results).

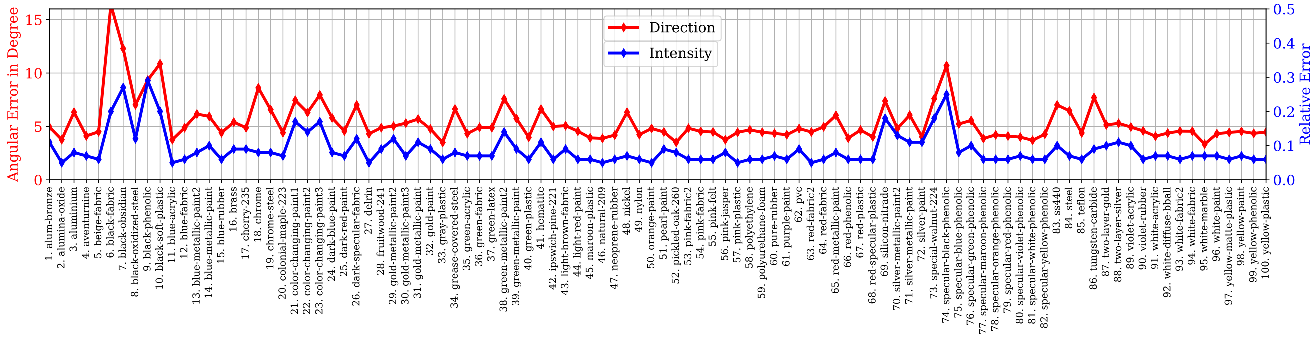
An alternative way is to regress a single intensity-scaled light direction vector for each image and use mean square error for training, but we experimentally found that such a coupled lighting representation decreased the performance on surfaces with complexed geometry (*e.g.*, BUNNY), as shown in Table 1 (ID 0 & 1), where this model is denoted as LCNet$_{\text{reg-coupled}}$.

To investigate the effect of different parameterization of light direction on lighting estimation, we also trained a regression based model, denoted as LCNet$_{\text{reg-}\phi\theta}$, to regress the azimuth and elevation angle (i.e., ϕ and θ) instead of a 3-vector light direction. Table 1 (ID 0 & 2) show that directly regressing (ϕ, θ) decreased the performance of light direction estimation.

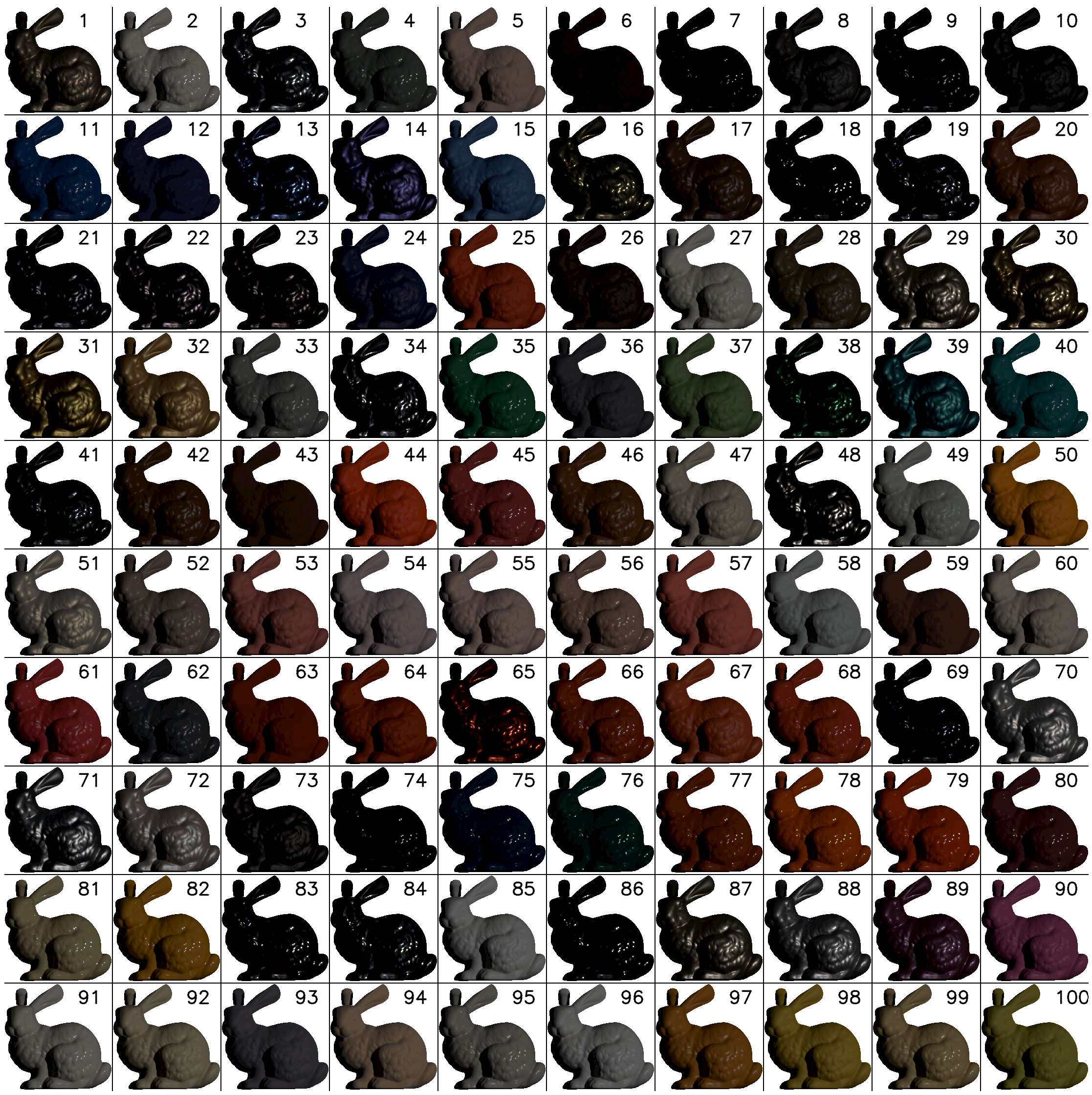
Table 1. Lighting estimation results of three regression based baseline models on the MERL$^{\text{Test}}$ dataset.

ID	Model	SPHERE Direction	SPHERE Intensity	BUNNY Direction	BUNNY Intensity
0	LCNet$_{\text{reg}}$	4.10	0.104	5.46	0.094
1	LCNet$_{\text{reg-coupled}}$	4.03	0.103	6.97	0.095
2	LCNet$_{\text{reg-}\phi\theta}$	4.57	0.083	5.87	0.091

2. Detailed Lighting Estimation Results of SDPS-Net on BUNNY



(a) Detailed lighting estimation results of LCNet on BUNNY from MERL$^{\text{Test}}$ dataset.



(b) Sample images for BUNNY rendered with 100 BRDFs.

Figure 1. Lighting estimation results of LCNet on 100 different BRDFs. We can see that LCNet can robustly estimate lighting conditions for different BRDFs. Note that the results on some of the dark materials are slightly worse (e.g., BRDFs with IDs 6-10 & 74.), which might be explained by the fact that images of dark material surfaces provide less information for feature extraction.

3. Detailed Normal Estimation Results of SDPS-Net on BUNNY

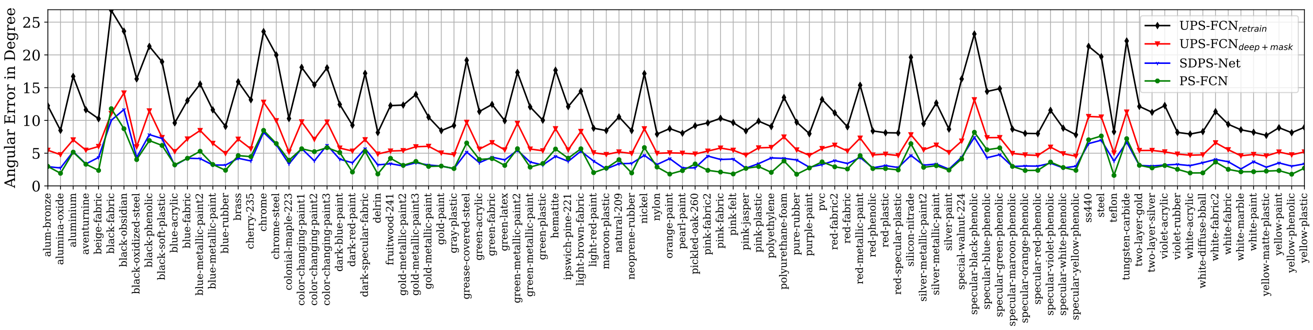


Figure 2. Quantitative comparison of normal estimation results among UPS-FCN$_{\text{retrain}}$, UPS-FCN$_{\text{deep+mask}}$, SDPS-Net, and PS-FCN on BUNNY from MERL$^{\text{Test}}$ (note that PS-FCN is a fully calibrated method).

4. Different Network Architectures for the Single-stage Model

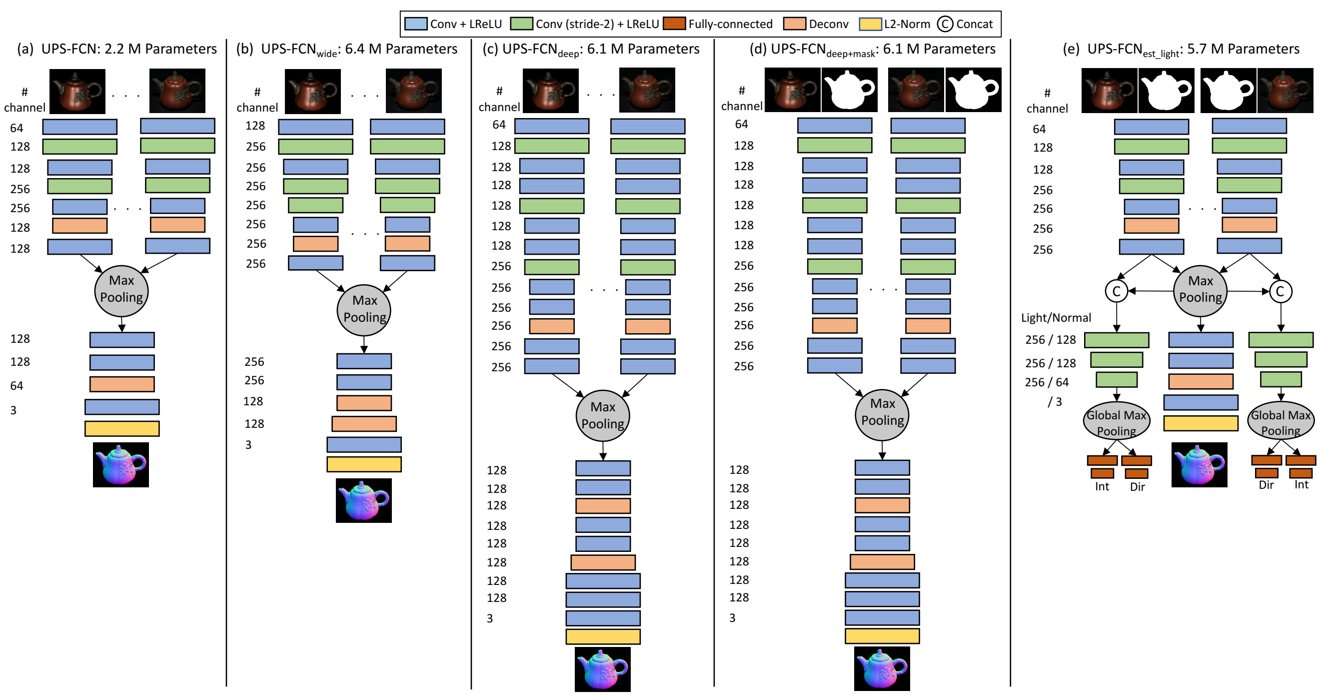


Figure 3. Different single-stage network architectures for normal estimation. Note that a global max-pooling layer is used in the lighting estimation sub-network of UPS-FCN$_{\text{est_light}}$ to handle inputs with varying scales.

5. Qualitative Results on the DiLiGenT Dataset

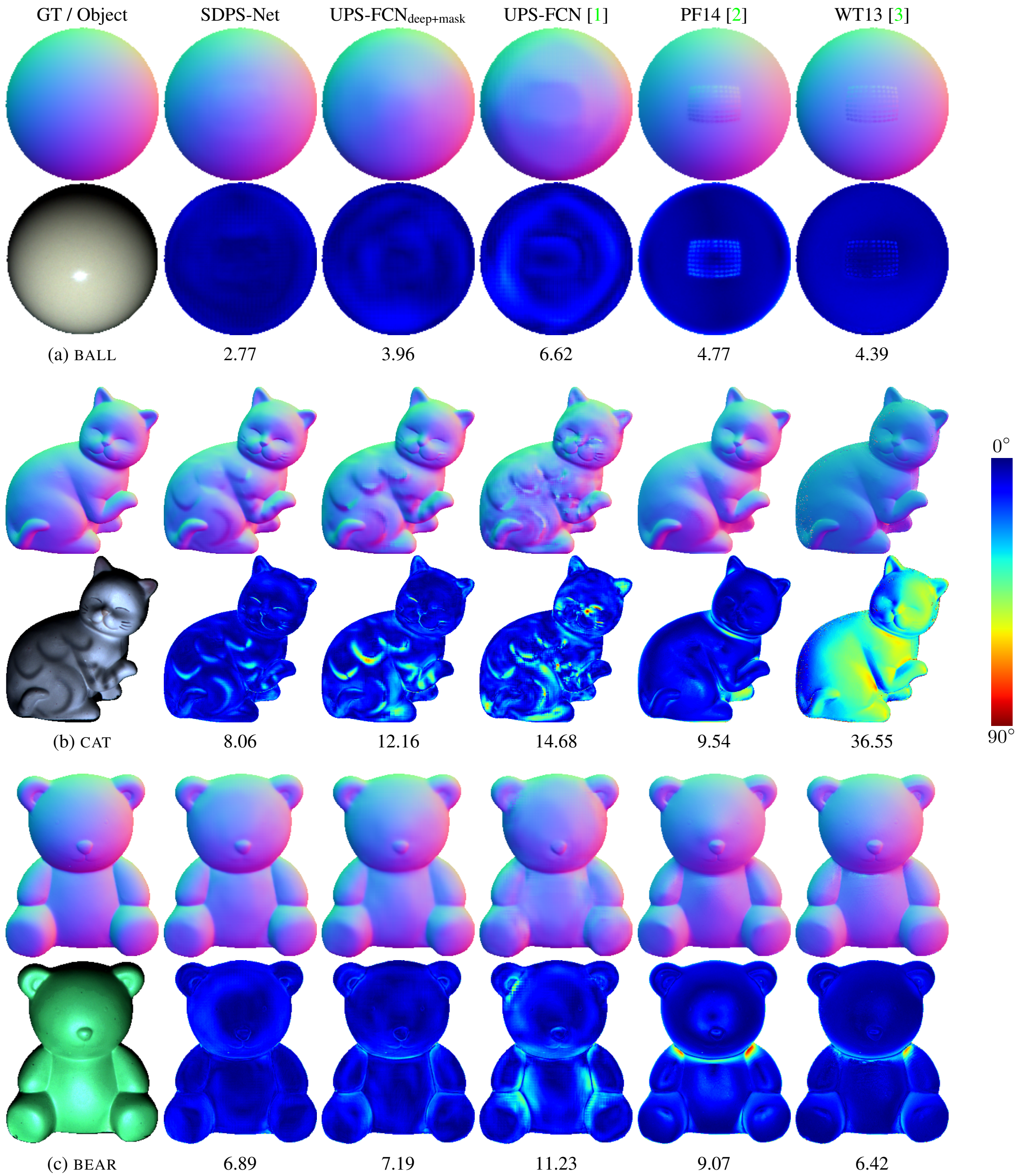


Figure 4. Qualitative results for BALL, CAT and BEAR in the DiLiGenT dataset.

Figure 5. Qualitative results for POT2, BUDDHA and GOBLET in the DiLiGenT dataset.

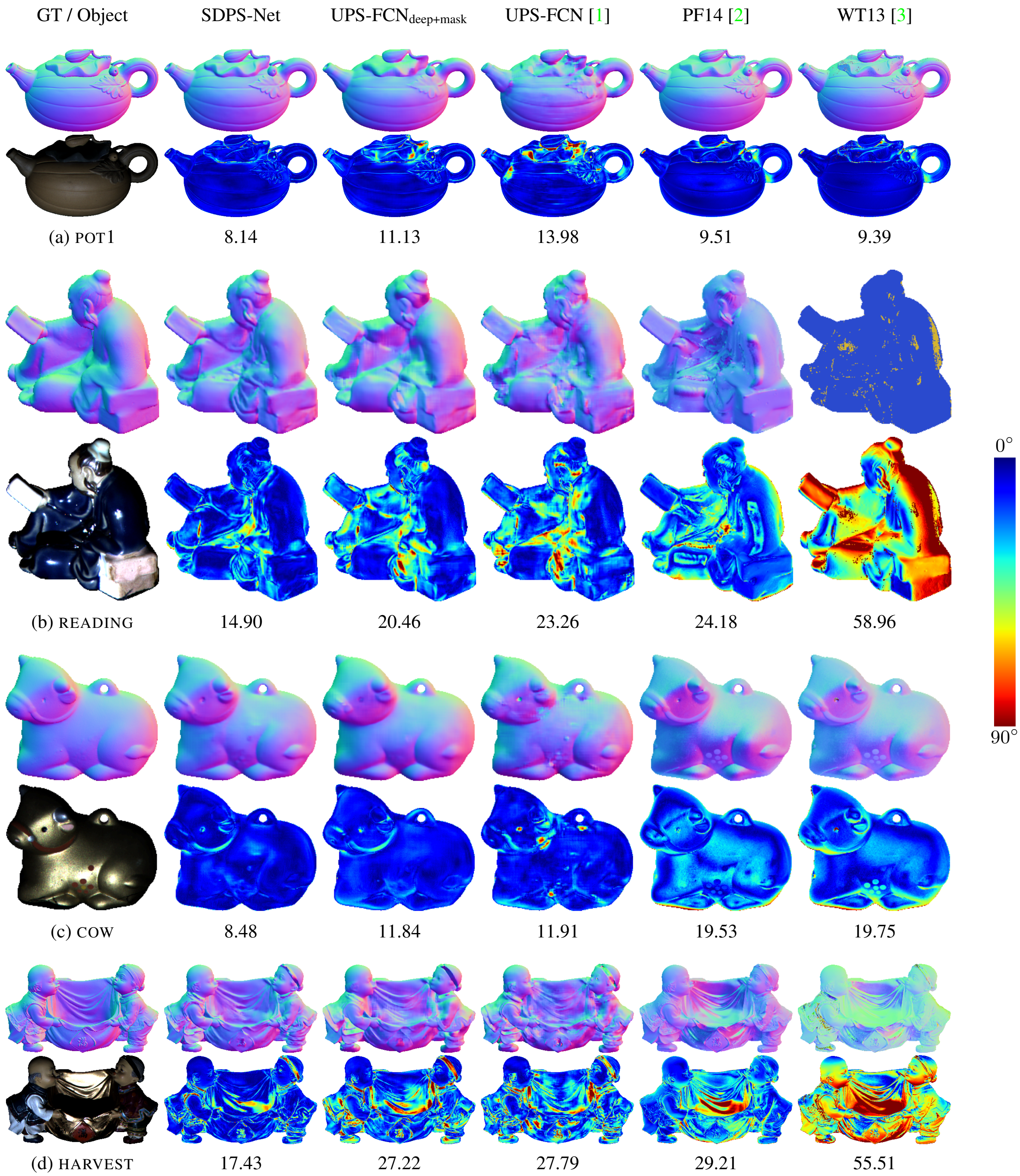


Figure 6. Qualitative results for POT1, READING, COW and HARVEST in the DiLiGenT dataset.

6. Qualitative Results on the Light Stage Data Gallery

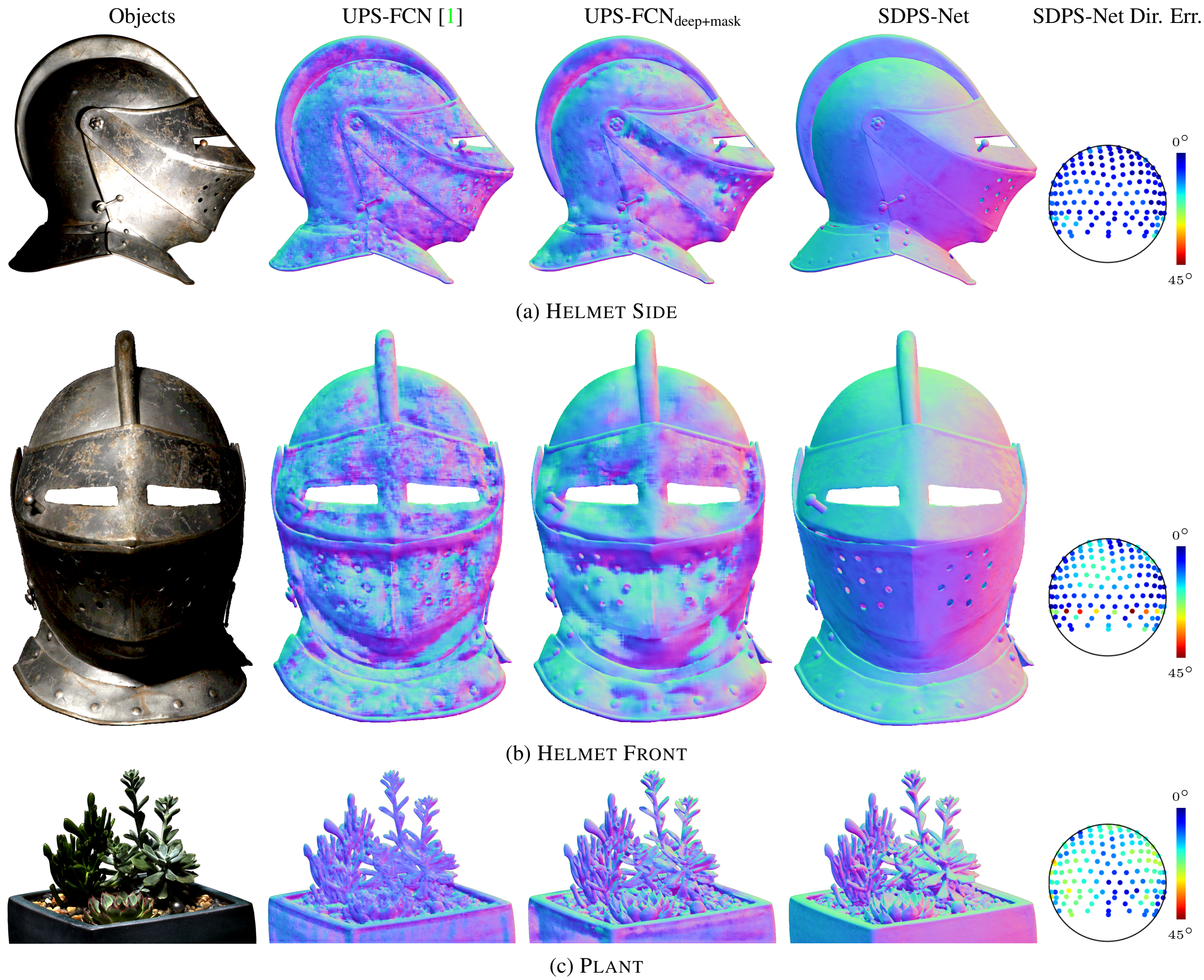
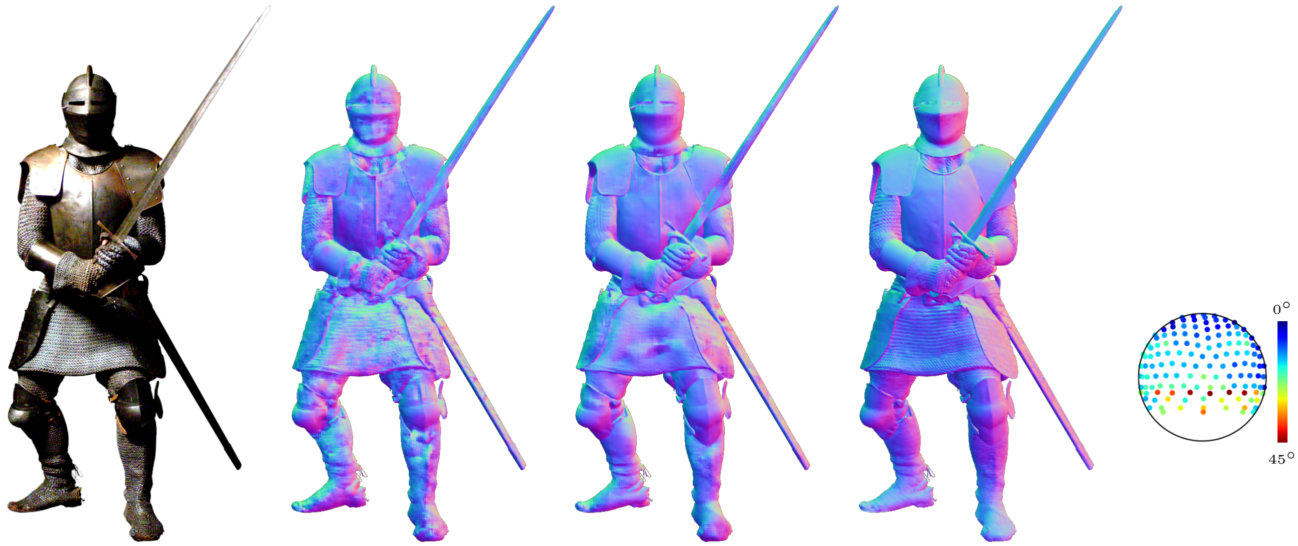
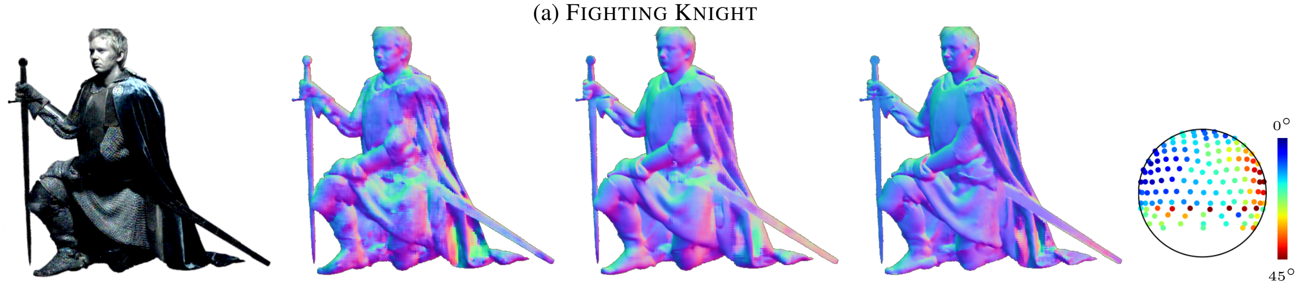


Figure 7. Qualitative results for HELMET SIDE, HELMET FRONT and PLANT in Light Stage Data Gallery. The right most column visualizes the error distributions of the light direction estimation.

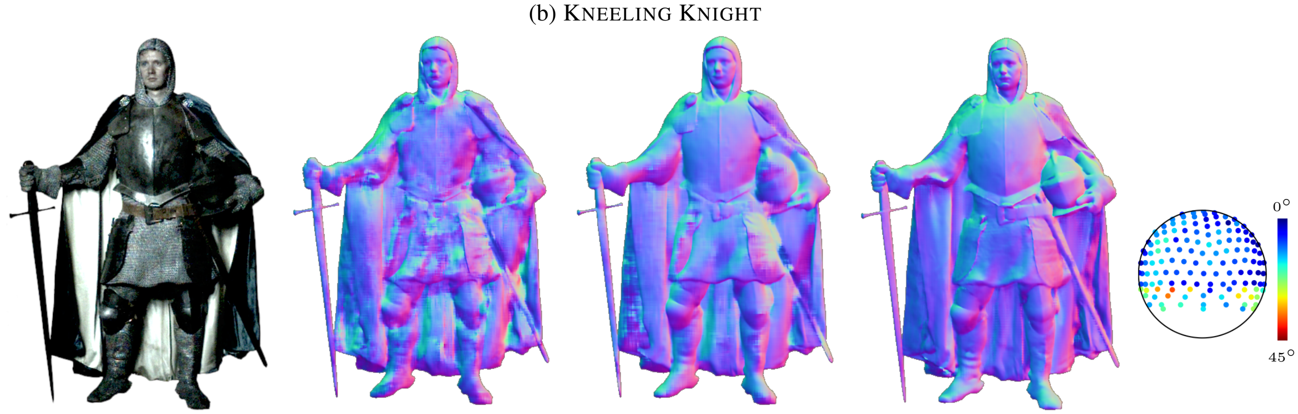
Objects UPS-FCN [1] UPS-FCN$_{\text{deep+mask}}$ SDPS-Net SDPS-Net Dir. Err.



(a) FIGHTING KNIGHT



(b) KNEELING KNIGHT



(c) STANDING KNIGHT

Figure 8. Qualitative results FIGHTING KNIGHT, KNEELING KNIGHT and STANDING KNIGHT in Light Stage Data Gallery. The right most column visualizes the error distributions of the light direction estimation.

7. Qualitative Results on the Gourd&Apple Dataset

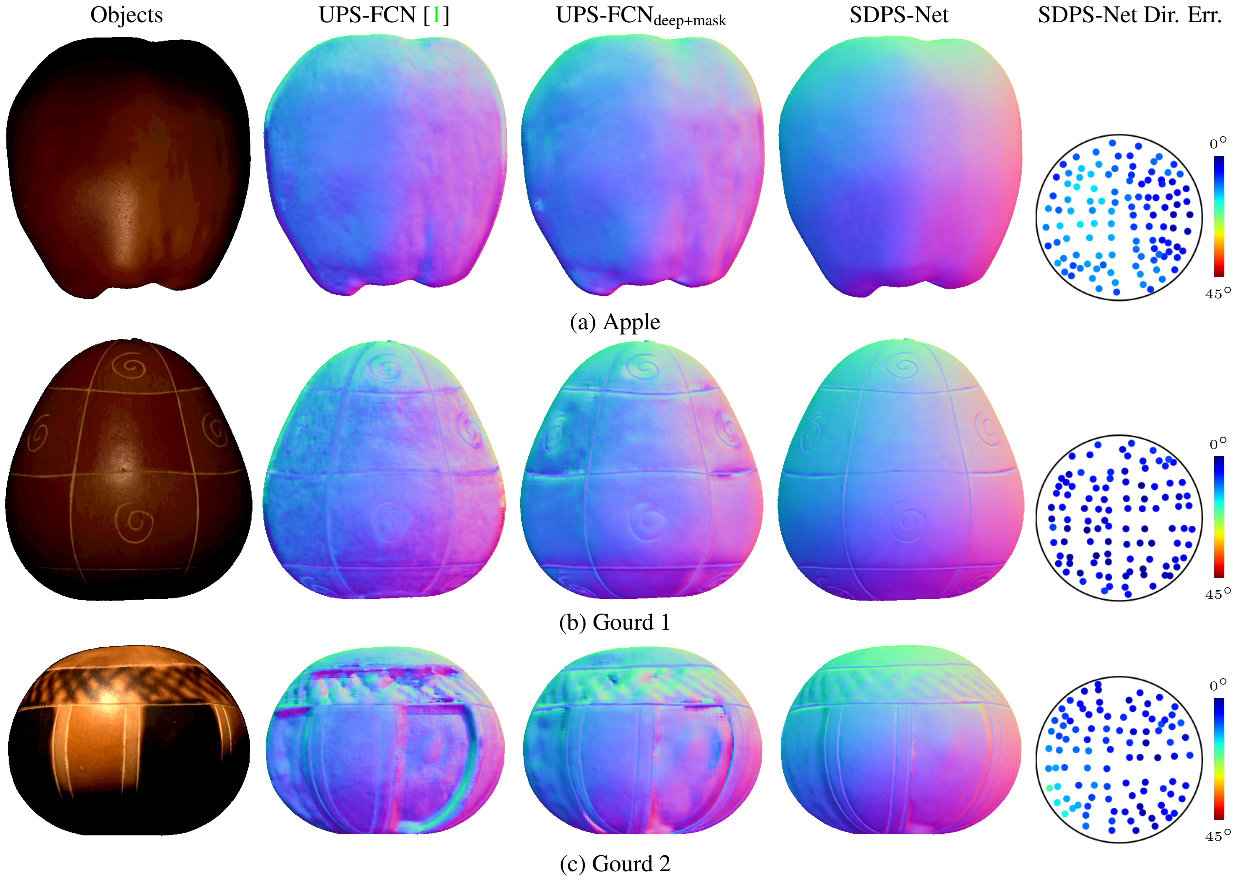


Figure 9. Qualitative results on Gourd&Apple dataset. The right most column visualizes the error distributions of the light direction estimation.

References

[1] G. Chen, K. Han, and K.-Y. K. Wong. PS-FCN: A flexible learning framework for photometric stereo. In *ECCV*, 2018. 5, 6, 7, 8, 9, 10

[2] T. Papadhimitri and P. Favaro. A closed-form, consistent and robust solution to uncalibrated photometric stereo via local diffuse reflectance maxima. *IJCV*, 2014. 5, 6, 7

[3] Z. Wu and P. Tan. Calibrating photometric stereo by holistic reflectance symmetry analysis. In *CVPR*, 2013. 5, 6, 7