Deep Photometric Stereo for Non-Lambertian Surfaces -Supplementary Material-

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. More Results for Calibrated Photometric Stereo

1.1. Quantitative Results of PS-FCN and PS-FCN^{+N} on Different Materials

Figure 1 compares PS-FCN, PS-FCN^{+N}, and L2 Baseline [7] on SPHERE that was rendered with 100 different BRDFs. It can be seen that PS-FCN and PS-FCN^{+N} achieved comparable results on different materials, which indicates that training with data normalization strategy will not worsen the results on homogeneous surfaces. Besides, both models significantly outperformed the L2 Baseline [7].

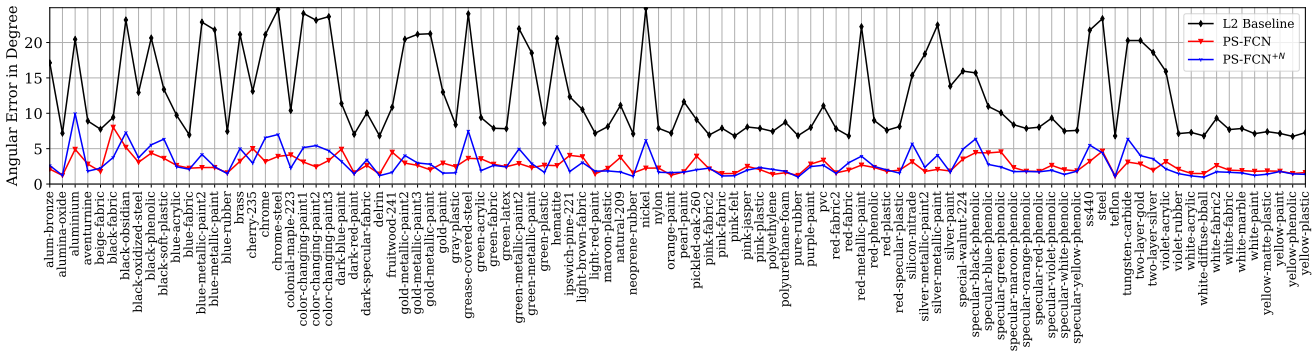


Figure 1. Quantitative results on SPHERE rendered with 100 MERL BRDFs. The average MAE for L2 Baseline [7], PS-FCN, and PS-FCN^{+N} are 12.59, 2.66 and 2.91, respectively.

1.2. Qualitative Results on the DiLiGenT Benchmark

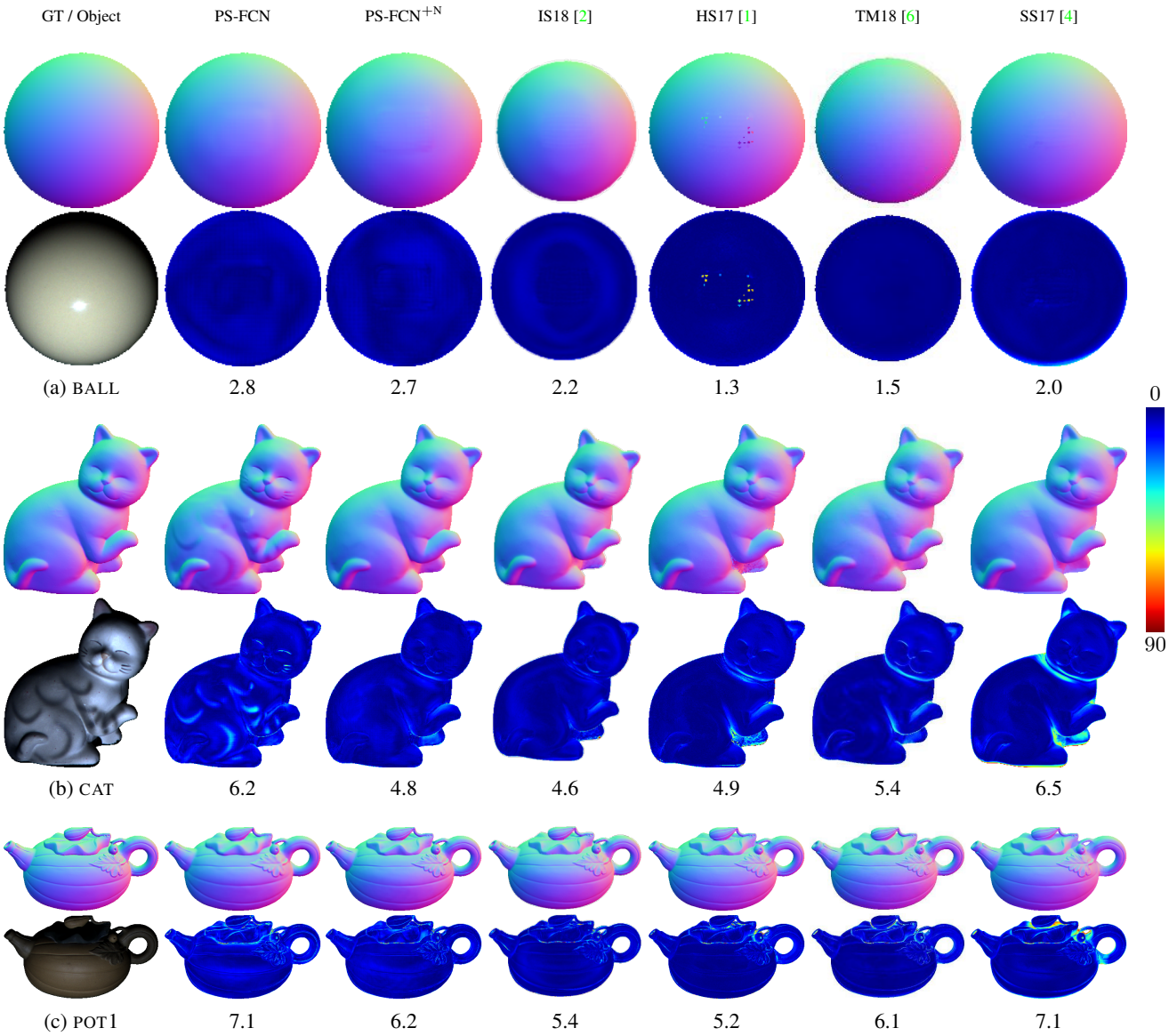


Figure 2. Results of calibrated photometric stereo for BALL, CAT and POT1 in the DiLiGenT benchmark.

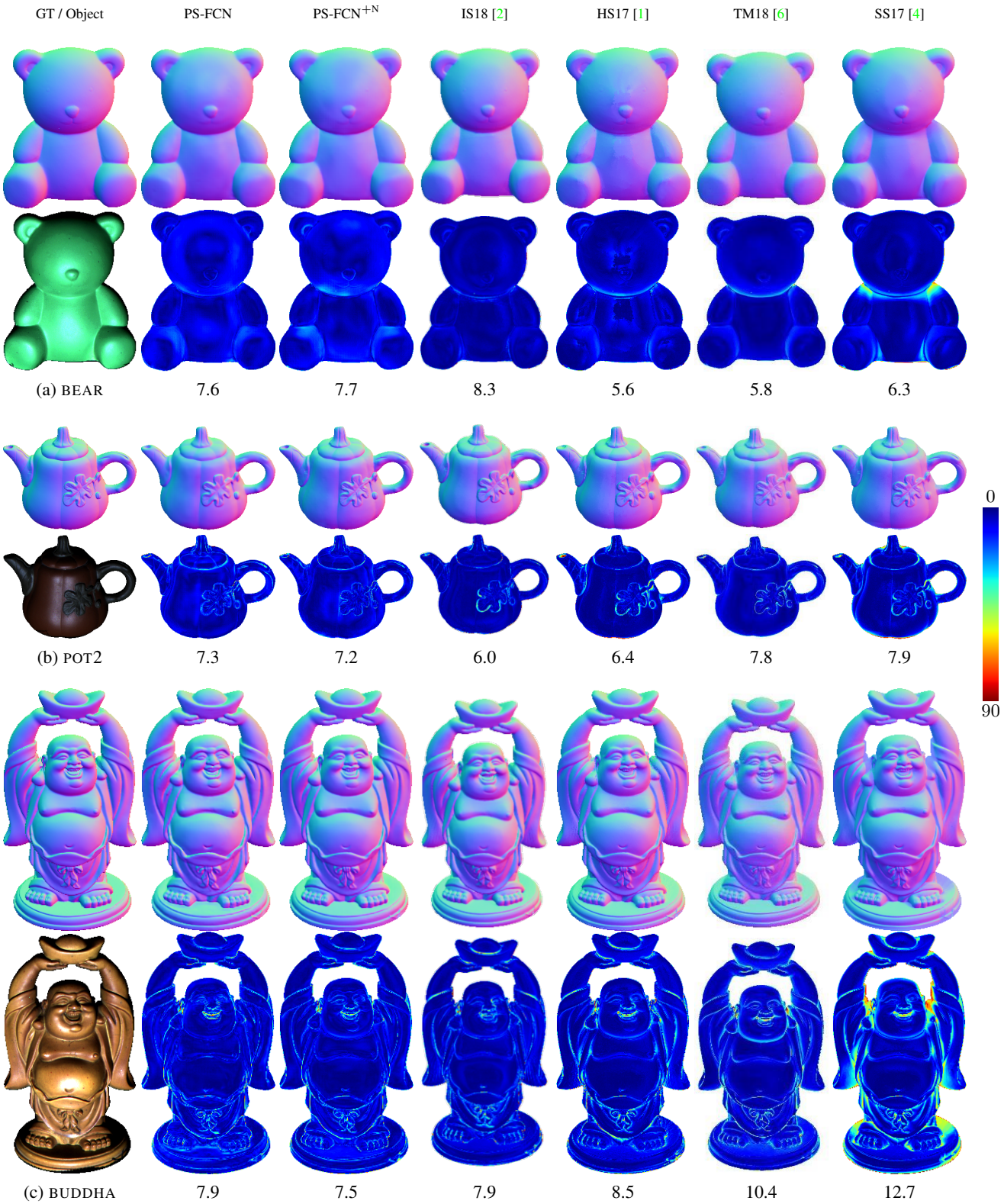


Figure 3. Results of calibrated photometric stereo for BEAR, POT2 and BUDDHA in the DiLiGenT benchmark.

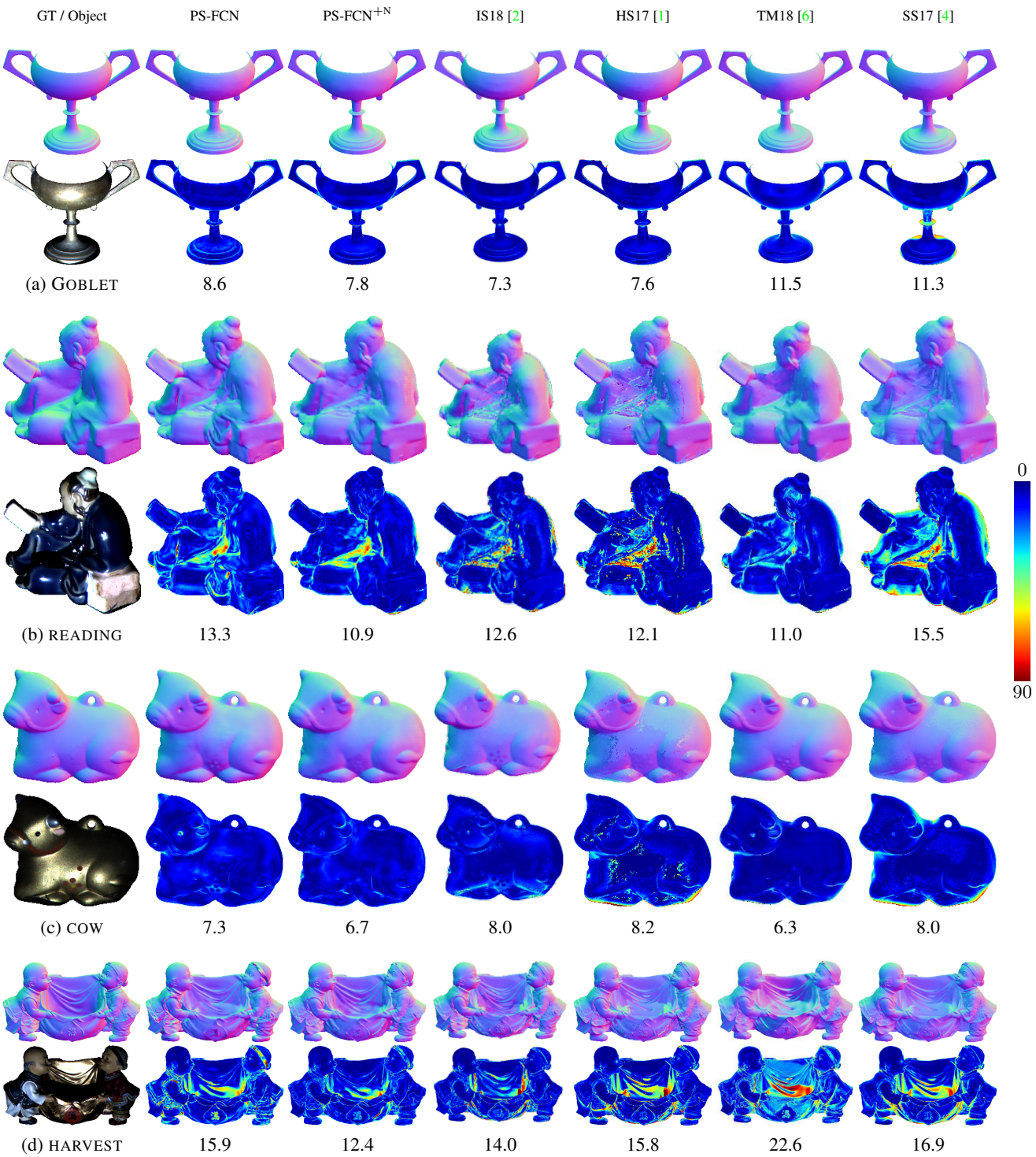


Figure 4. Results of calibrated photometric stereo for GOBLET, READING, COW and HARVEST in the DiLiGenT benchmark.

1.3. Qualitative Results on the Light Stage Data Gallery

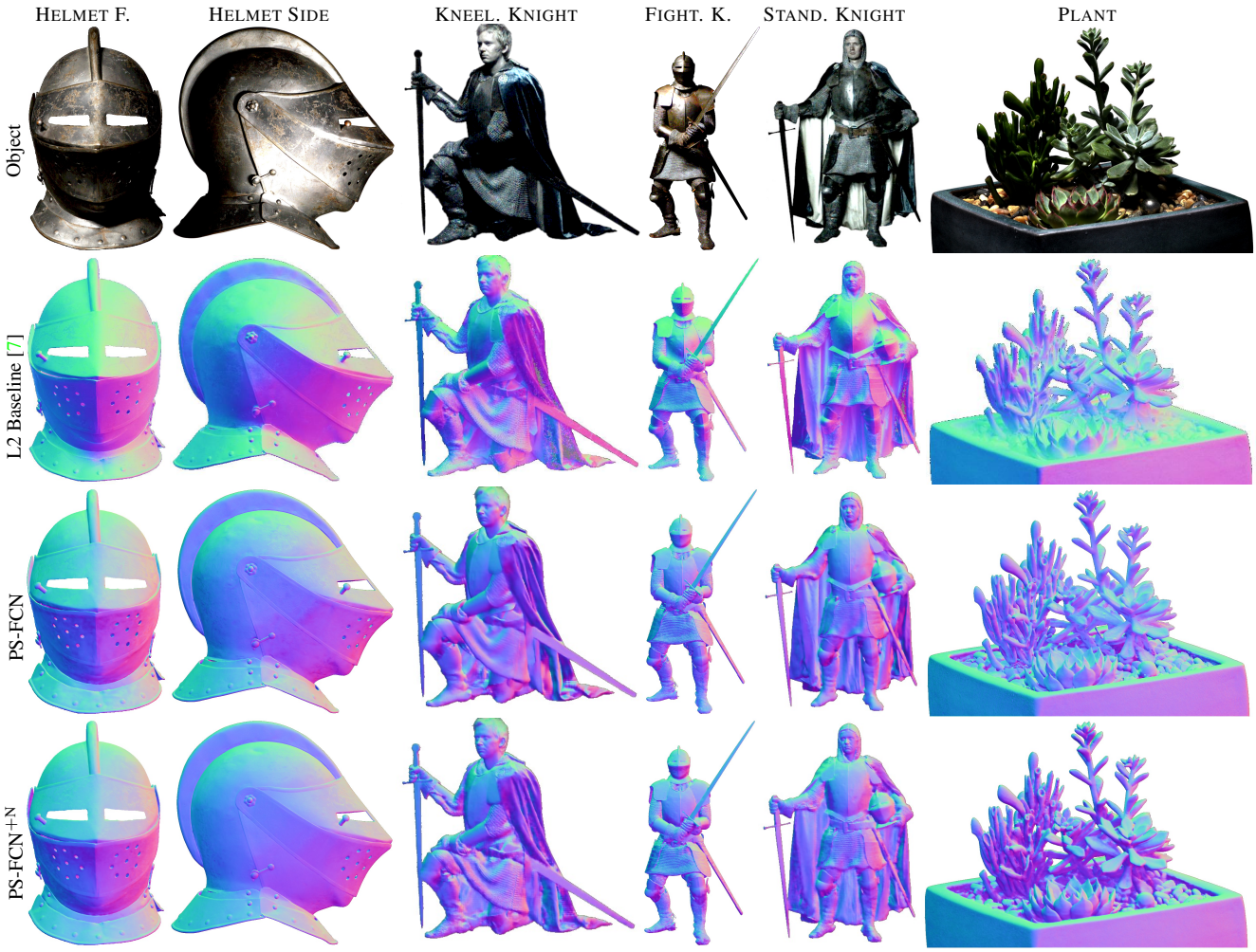
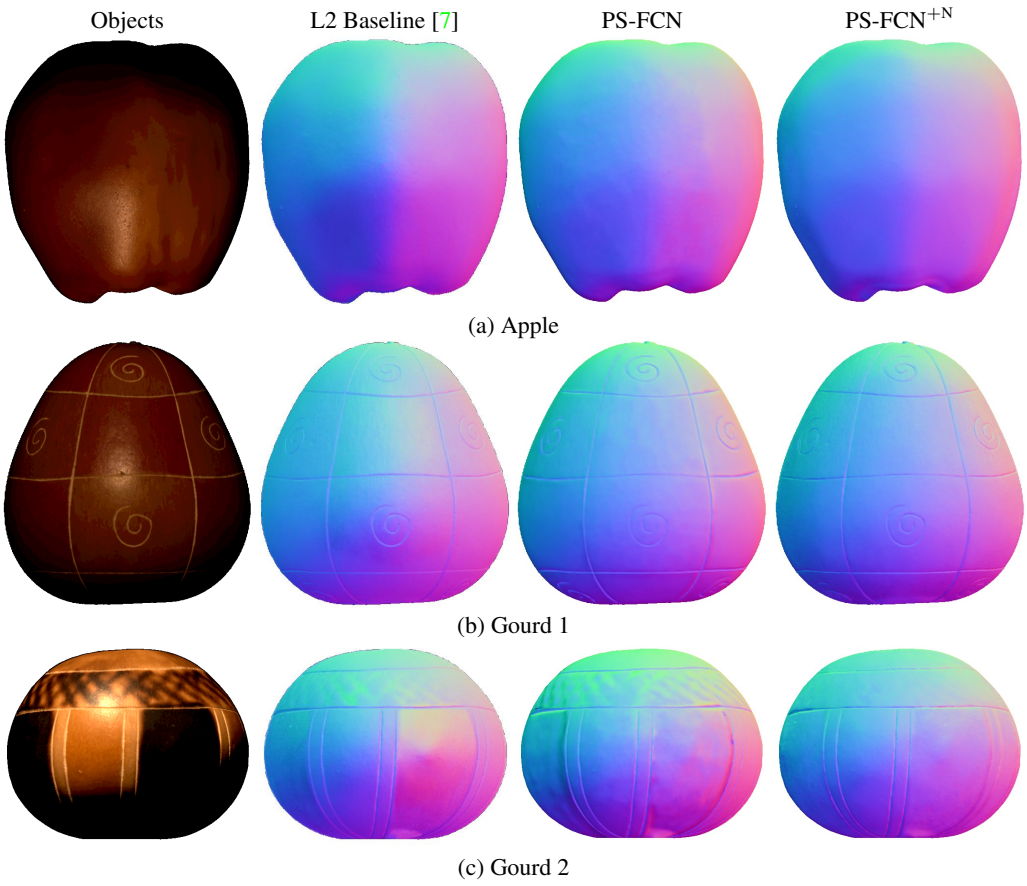


Figure 5. Qualitative results of calibrated photometric stereo on Light Stage Data Gallery.

1.4. Qualitative Results on the Gourd&Apple Dataset



(a) Apple

(b) Gourd 1

(c) Gourd 2

Figure 6. Qualitative results of calibrated photometric stereo on Gourd&Apple dataset.

2. More Results for Uncalibrated Photometric Stereo

2.1. 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

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

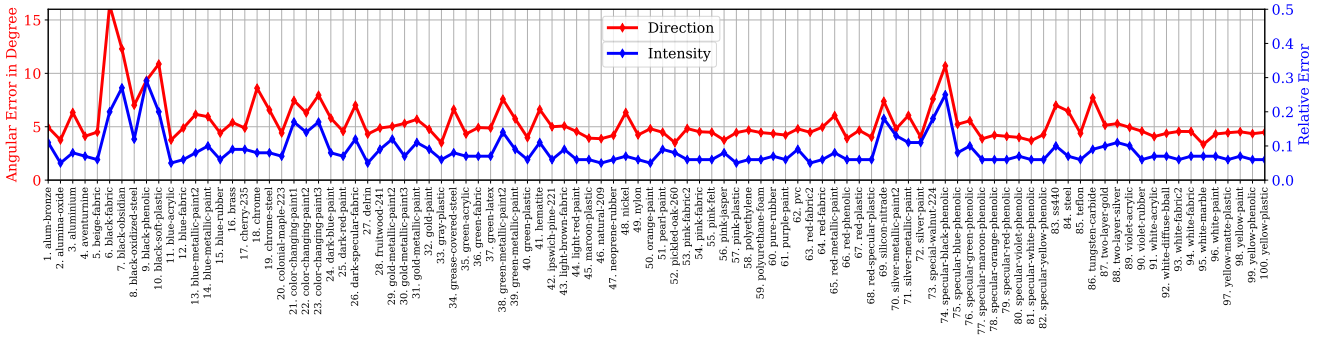
where λ_l and λ_e are the weighting factors for the loss terms, $\boldsymbol{l}_i$ (e_i) and $\tilde{\boldsymbol{l}}_i$ ($\tilde{e}_i$) denote the predicted light direction (intensity) and the ground truth, respectively, for image i. During training, λ_l and λ_e are set to 1 (we found that using other 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, as shown in Table 1, where this model is denoted as LCNet$_{\text{reg-coupled}}$.

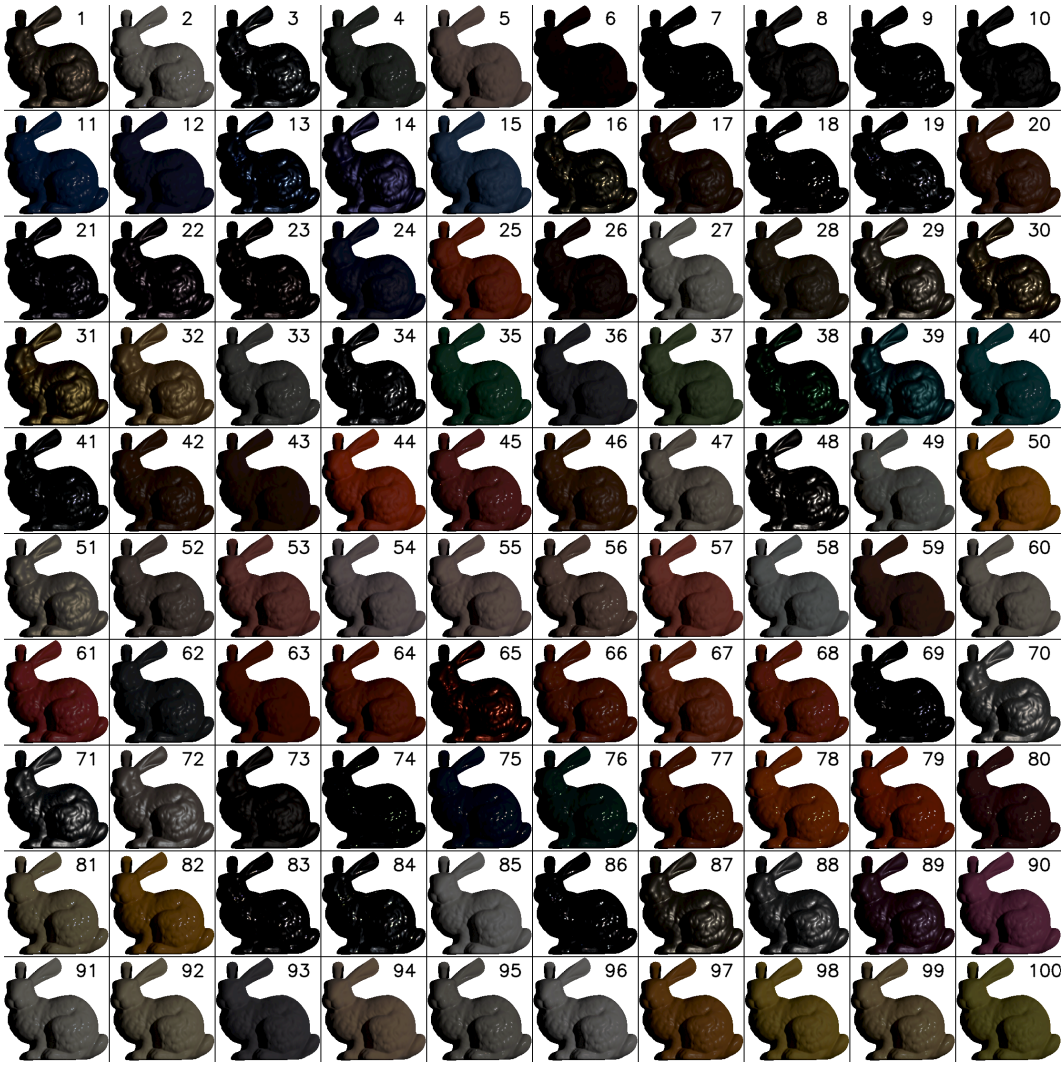
Table 1. Lighting estimation results of two regression based baseline models on the SynTest$^{\text{MERL}}$ dataset.

ID	Model	SPHERE		BUNNY		DRAGON		ARMADILLO	
		Dir.	Int.	Dir.	Int.	Dir.	Int.	Dir.	Int.
0	LCNet$_{\text{reg}}$	4.10	0.104	**5.46**	**0.094**	**8.27**	**0.125**	**8.31**	**0.104**
1	LCNet$_{\text{reg-coupled}}$	**4.03**	**0.103**	6.97	0.095	8.47	0.126	8.50	0.107

2.2. Detailed Lighting Estimation Results of LCNet on BUNNY



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



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

Figure 7. 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.

2.3. Detailed Normal Estimation Results of LCNet+PS-FCN† on BUNNY

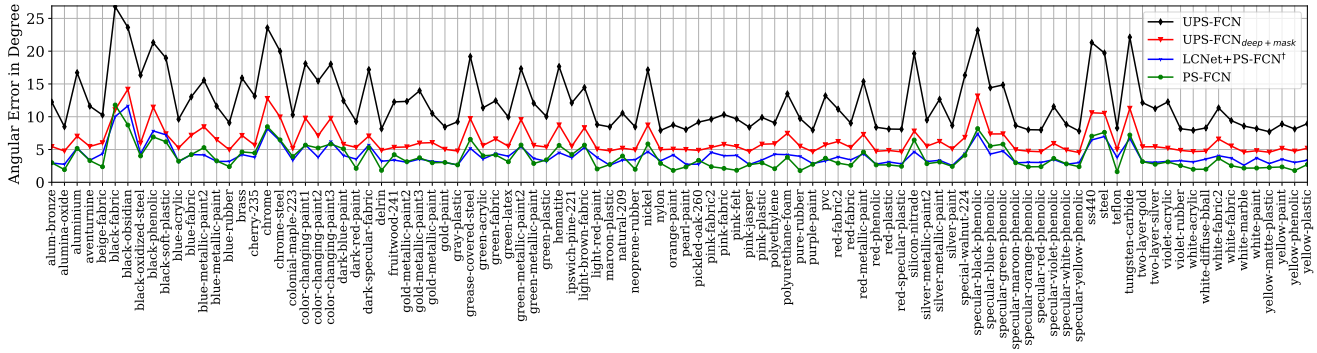


Figure 8. Quantitative comparison of normal estimation results among UPS-FCN, UPS-FCN$_{\text{deep+mask}}$, LCNet+PS-FCN†, and PS-FCN on BUNNY from SynTest$^{\text{MERL}}$.

2.4. Network Architectures for the Single-stage Model

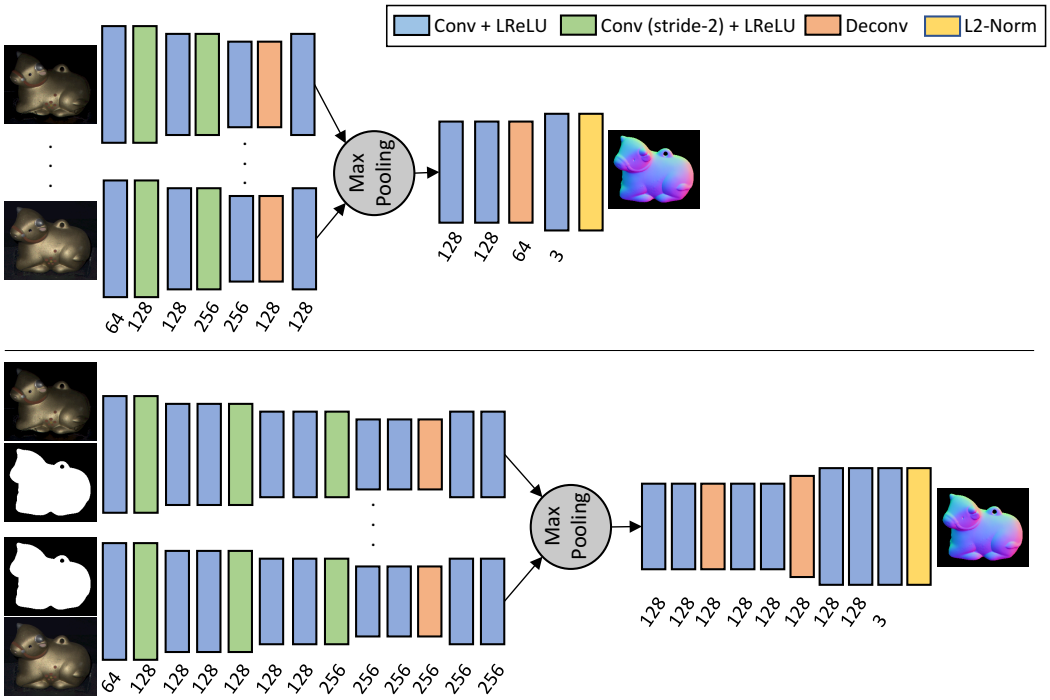


Figure 9. Network architectures for UPS-FCN (top) and UPS-FCN$_{\text{deep+mask}}$ (bottom).

2.5. Qualitative Results on the DiLiGenT Benchmark

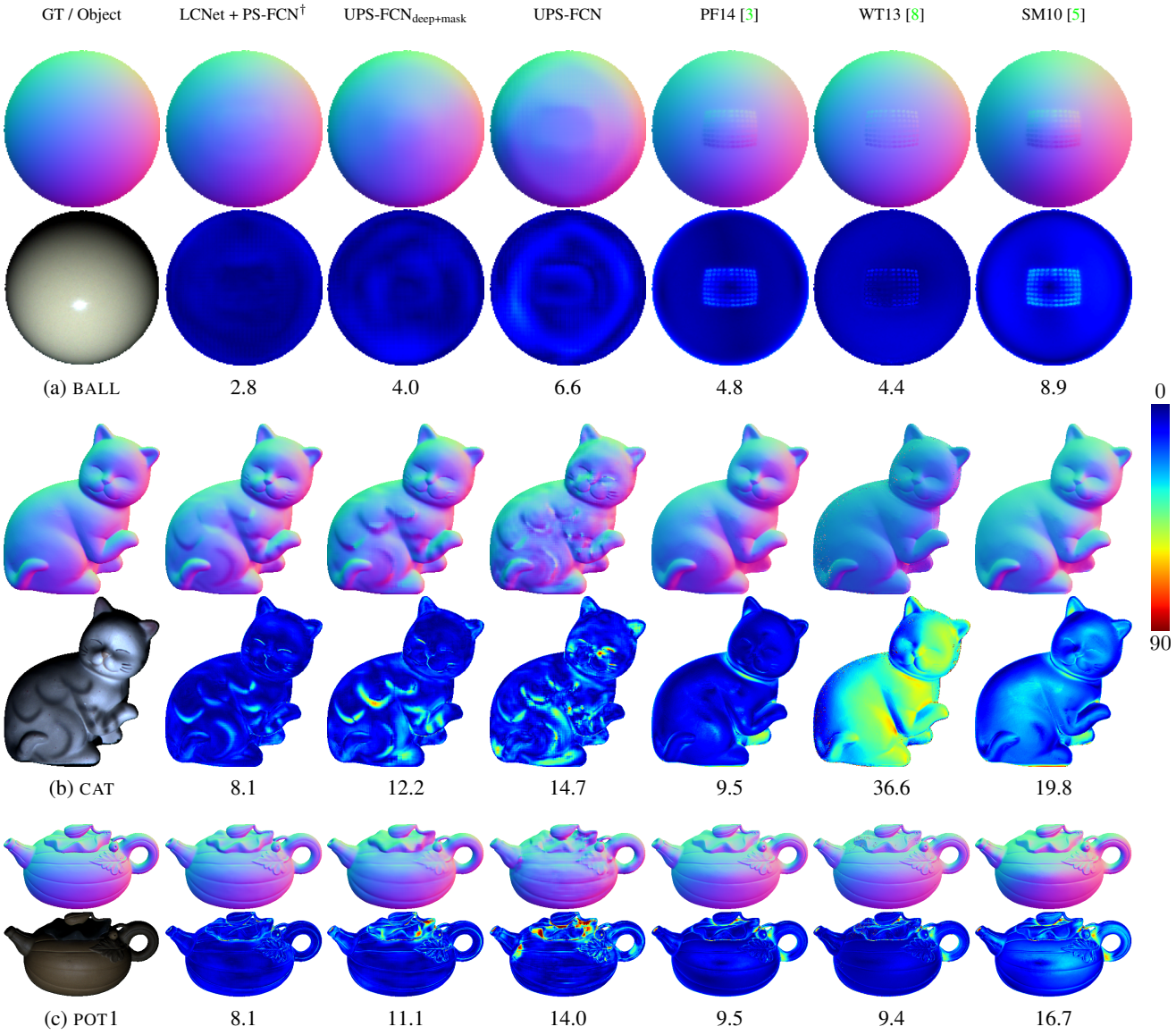


Figure 10. Results of uncalibrated photometric stereo for BALL, CAT and POT1 in the DiLiGenT benchmark.

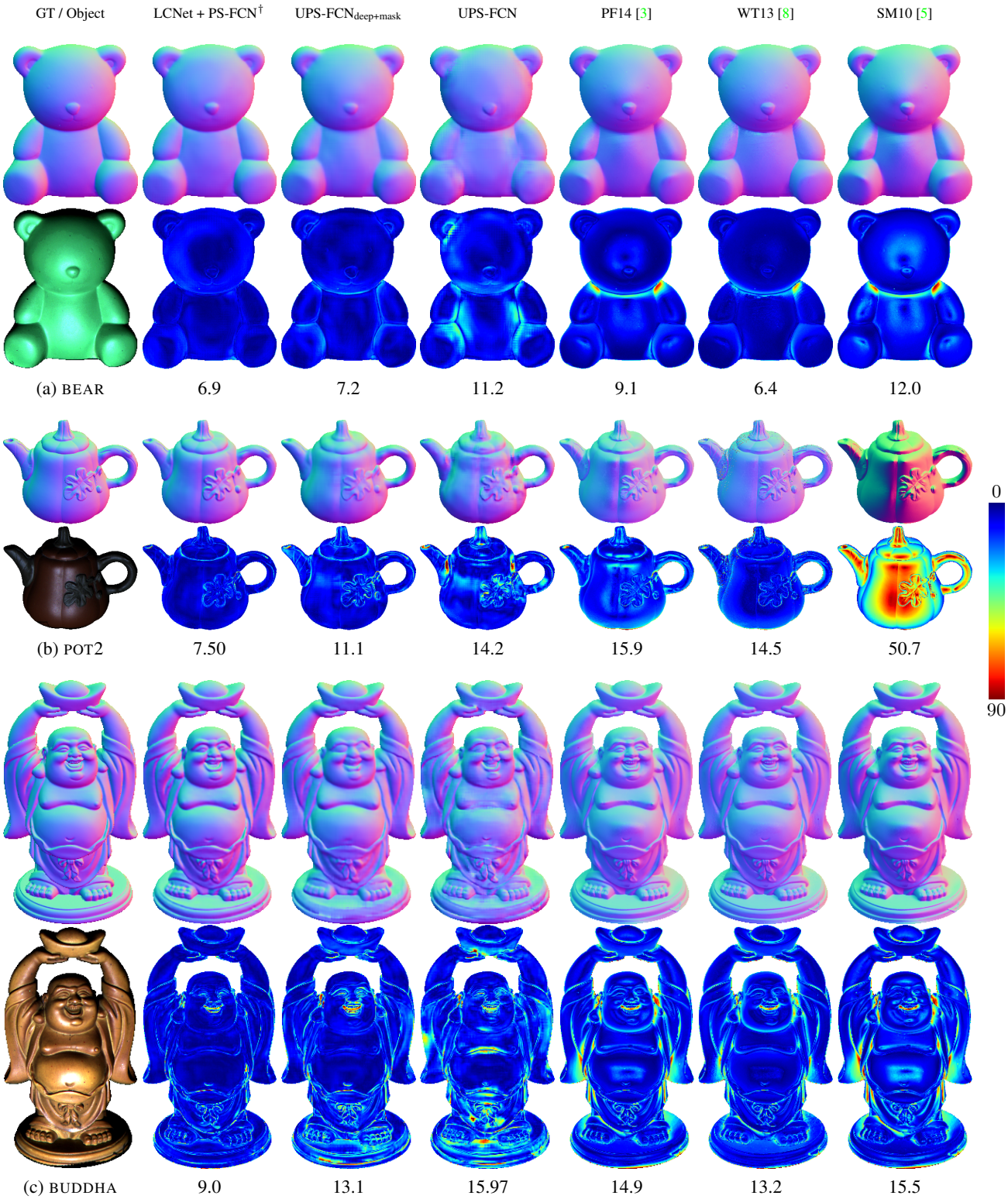


Figure 11. Results of uncalibrated photometric stereo for BEAR, POT2 and BUDDHA in the DiLiGenT benchmark.

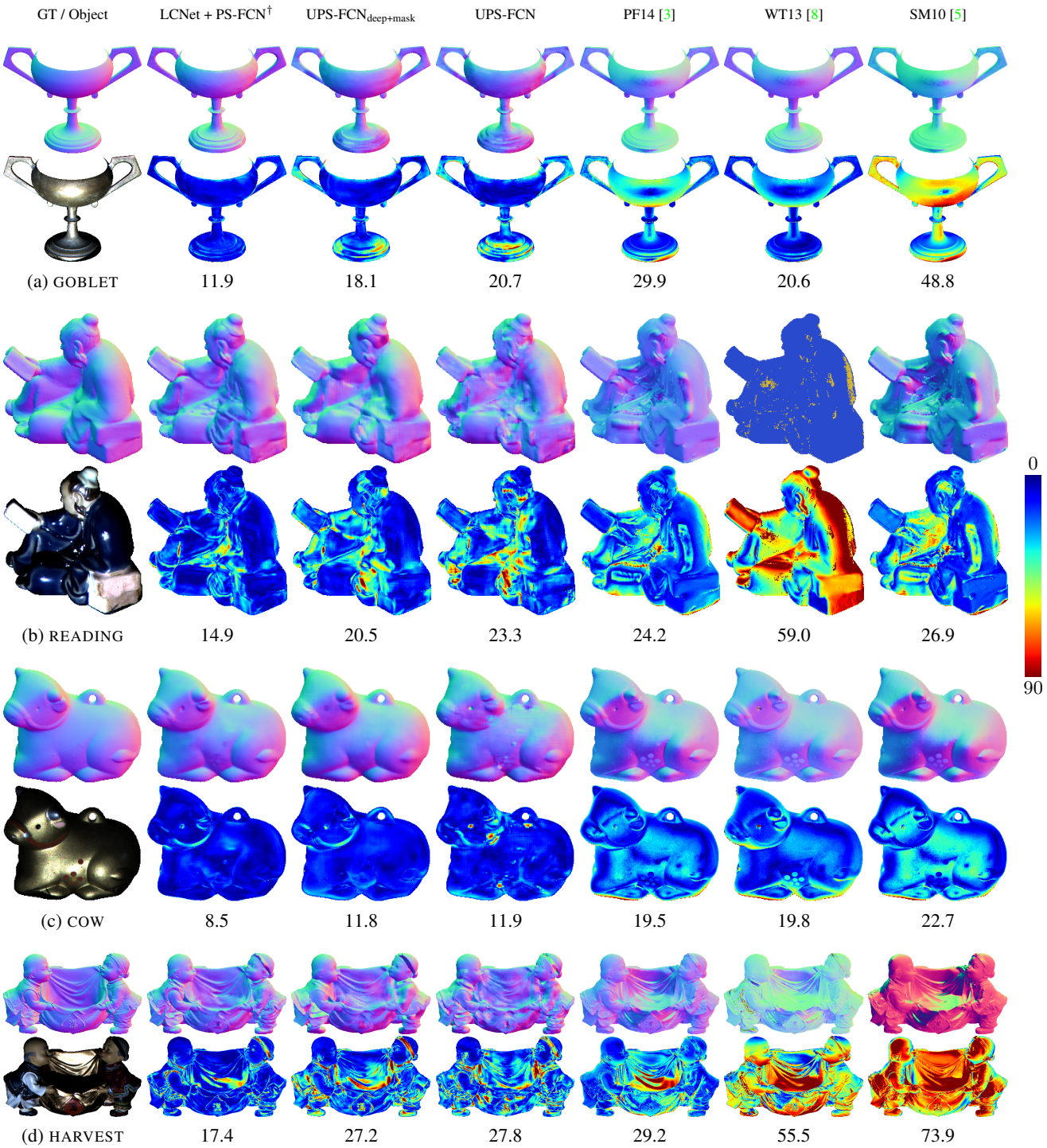


Figure 12. Results of uncalibrated photometric stereo for GOBLET, READING, COW and HARVEST in the DiLiGenT benchmark.

2.6. Qualitative Results on the Light Stage Data Gallery

Table 2. Lighting estimation results of LCNet on the Light Stage Data Gallery.

	HELMET SIDE	PLANT	FIGHTING KNIGHT	KNEELING KNIGHT	STANDING KNIGHT	HELMET FRONT	Avg.
Direction	6.57	16.06	15.95	19.84	11.60	11.62	13.61
Intensity	0.212	0.170	0.214	0.199	0.286	0.248	0.221

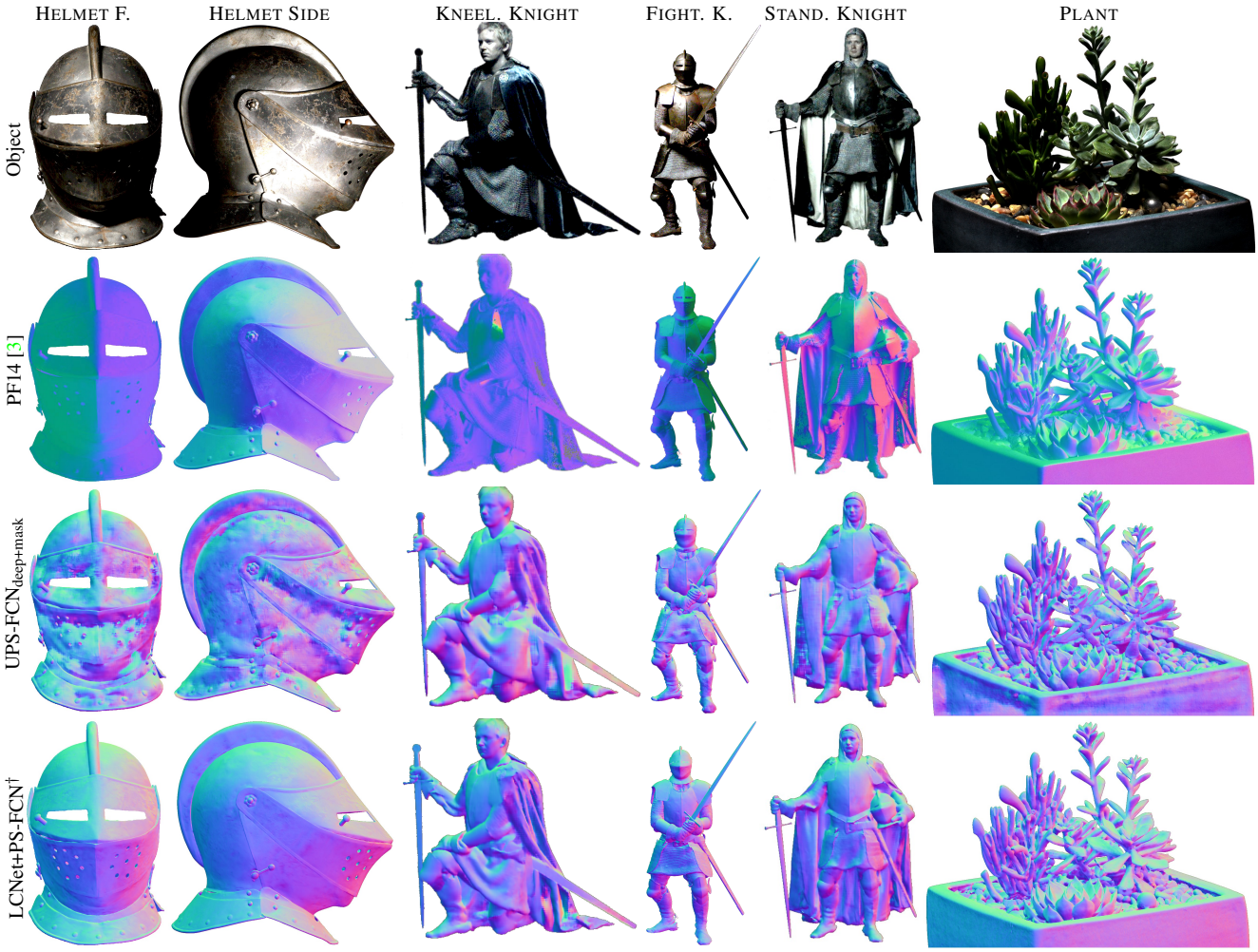
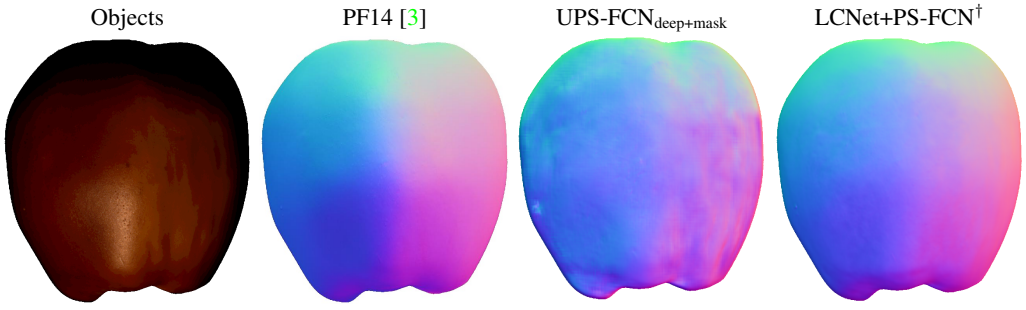


Figure 13. Qualitative results of uncalibrated photometric stereo on Light Stage Data Gallery. Note that although UPS-FCN$_{deep+mask}$ produced good results for the DiLiGenT benchmark, it had difficulty in handling the objects in Light Stage Data Gallery and Gourd&Apple Dataset. This is because the image sizes of these two datasets are considerably larger (e.g., PLANT: 1200×850, APPLE: 650×700), compared with the DiLiGenT benchmark (e.g., READING: 200×220). Single-stage methods estimate surface normals for each patch based solely on the observations of that patch and thus tend to produce noisy estimation when the image size is large. In contrast, the proposed two-stage method first estimates the global directional lightings using the whole image and then estimate the surface normals based on the lighting and observations. This result reiterates the advantage of our two-stage method.

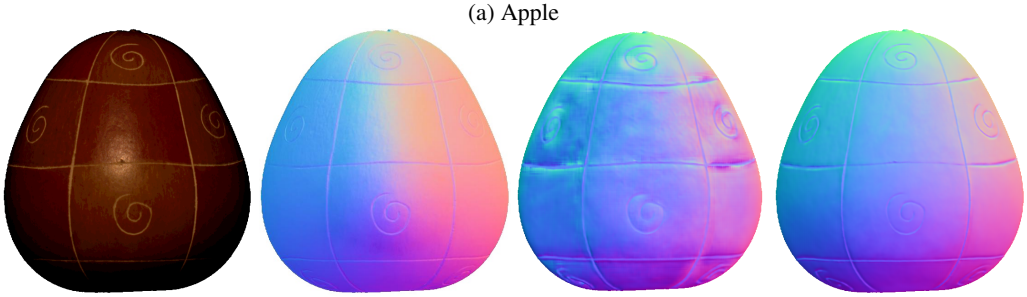
2.7. Qualitative Results on the Gourd&Apple Dataset

Table 3. Lighting estimation results of LCNet on the Gourd&Apple dataset.

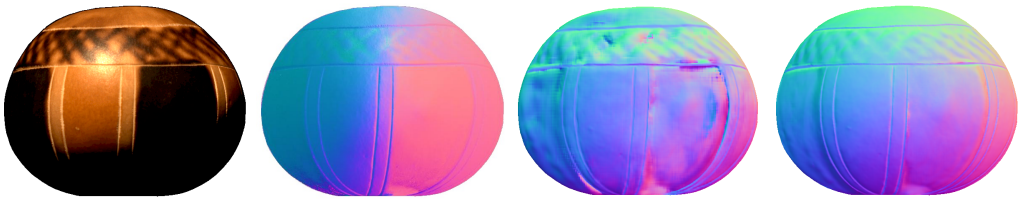
	APPLE	GOURD1	GOURD2	Avg.
Direction	9.31	4.07	7.11	6.83
Intensity	0.106	0.048	0.186	0.113



(a) Apple



(b) Gourd 1



(c) Gourd 2

Figure 14. Qualitative results of uncalibrated photometric stereo on Gourd&Apple dataset.

References

[1] Z. Hui and A. C. Sankaranarayanan. Shape and spatially-varying reflectance estimation from virtual exemplars. *IEEE TPAMI*, 2017. 3, 4, 5

[2] S. Ikehata. CNN-PS: CNN-based photometric stereo for general non-convex surfaces. In *ECCV*, 2018. 3, 4, 5

[3] T. Papadhimitri and P. Favaro. A closed-form, consistent and robust solution to uncalibrated photometric stereo via local diffuse reflectance maxima. *IJCV*, 2014. 11, 12, 13, 14, 15

[4] H. Santo, M. Samejima, Y. Sugano, B. Shi, and Y. Matsushita. Deep photometric stereo network. In *ICCV Workshops*, 2017. 3, 4, 5

[5] B. Shi, Y. Matsushita, Y. Wei, C. Xu, and P. Tan. Self-calibrating photometric stereo. In *CVPR*, 2010. 11, 12, 13

[6] T. Taniai and T. Maehara. Neural inverse rendering for general reflectance photometric stereo. In *ICML*, 2018. 3, 4, 5

[7] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical engineering*, 1980. 2, 6, 7

[8] Z. Wu and P. Tan. Calibrating photometric stereo by holistic reflectance symmetry analysis. In *CVPR*, 2013. 11, 12, 13