# Self-calibrating Deep Photometric Stereo Networks


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



CVPR LONG BEACH CALIFORNIA June 16-20, 2019

Code & Model: http://guanyingc.github.io/SDPS-Net/


## Problem Definition and Contribution

**Goal:** Estimating the normal map of a non-Lambertian surface with unknown isotropic reflectances observed under unknown varying light directions.

**Motivations:**

- Traditional methods for uncalibrated photometric stereo often assume a simplified reflectance model.
- Recent learning based methods for photometric stereo often assume known light directions.
- The performance of the existing learning based method for uncalibrated photometric stereo (*i.e.*, the single-stage method UPS-FCN [1]) lags far behind the calibrated methods.

## Problem Formulation

**Image Formation Model:** Consider a non-Lambertian surface whose appearance is described by a general isotropic BRDF $\rho$. Given a surface point with normal $\boldsymbol{n} \in \mathbb{R}^3$ being illuminated by the $j$-th incoming lighting with direction $\boldsymbol{l}_j \in \mathbb{R}^3$ and intensity $e_j \in \mathbb{R}$, the image formation model can be expressed as

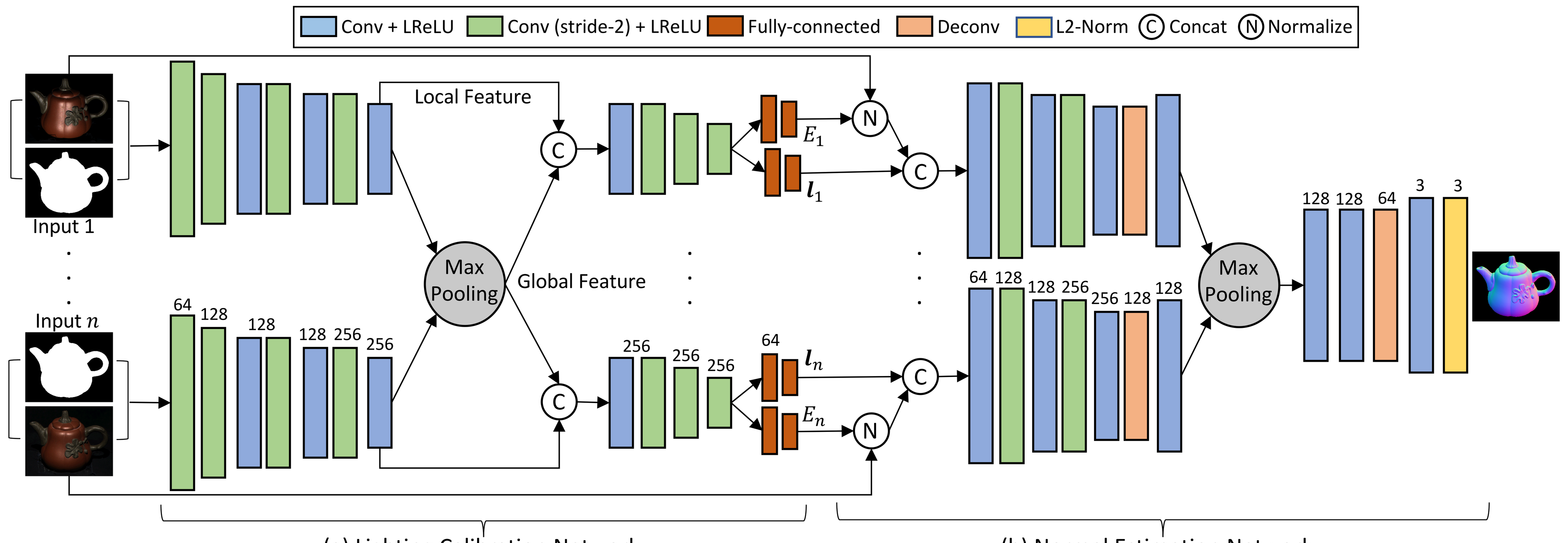
$$m_j = e_j \rho(\boldsymbol{n}, \boldsymbol{l}_j) \max(\boldsymbol{n}^\top \boldsymbol{l}_j, 0) + \epsilon_j,$$

where $m$ represents the measured intensity, $\max(:, 0)$ accounts for attached shadows, and $\epsilon$ accounts for the global illumination effects and noise.

**Main Idea:** We first estimate lightings from the measured intensities, and then solve for the surface normals using the estimated lightings and measured intensities.

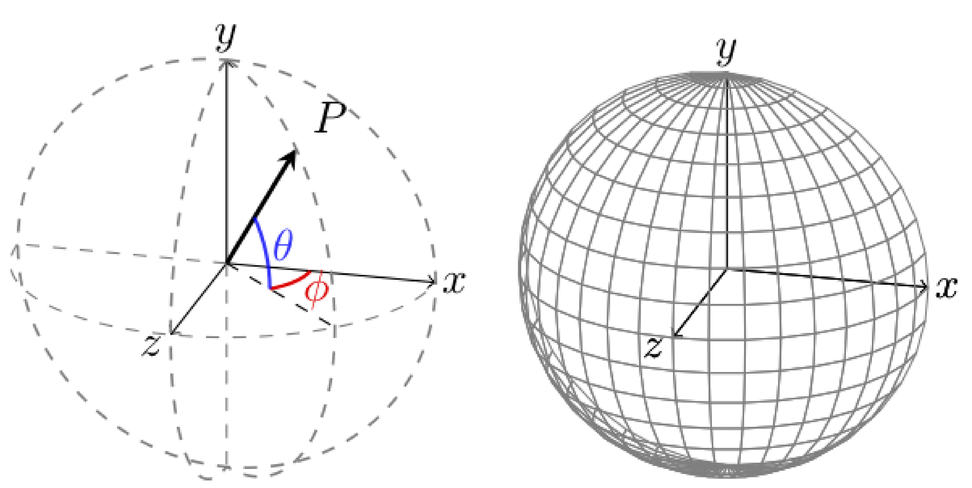
## Method

### Network Architecture:




(a) Lighting Calibration Network (b) Normal Estimation Network

### Discretization of Lighting Space:



- Illustration of the coordinate system (left)
- Example discretization (right)

### Loss Function for Lighting Estimation:

$$\mathcal{L}_{\text{Light}} = \lambda_{l_a}\mathcal{L}_{l_a} + \lambda_{l_e}\mathcal{L}_{l_e} + \lambda_e\mathcal{L}_e$$

- $\mathcal{L}_{l_a}$: azimuth classification loss
- $\mathcal{L}_{l_e}$: elevation classification loss
- $\mathcal{L}_e$: light intensity classification loss

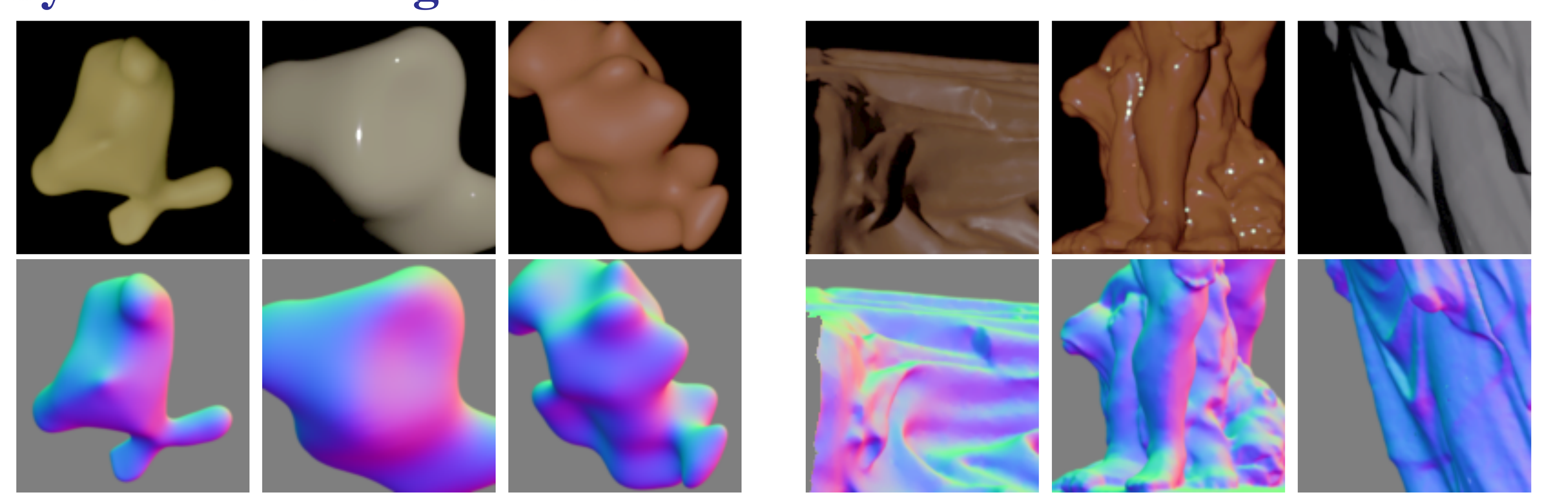
### Loss function for Normal Estimation:

$$\mathcal{L}_{\text{Normal}} = \frac{1}{hw}\sum_{i}^{hw}\left(1 - \boldsymbol{n}_i^\top \tilde{\boldsymbol{n}}_i\right)$$

- Cosine similarity loss

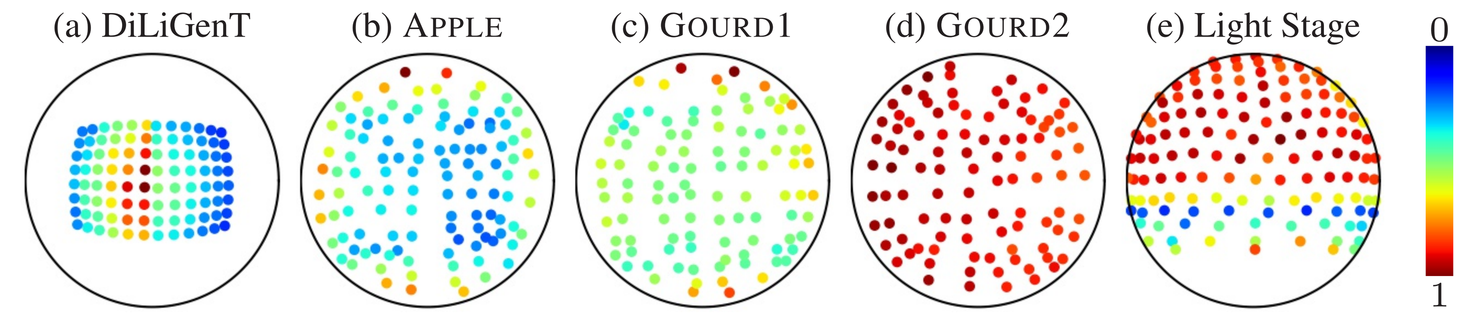
## Experiments & Results

### Synthetic Training Datasets:



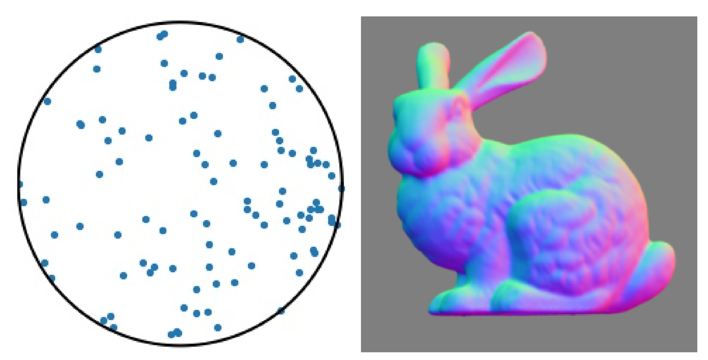
(a) Blobby dataset (b) Sculpture dataset

### Lighting Distribution of Real Datasets:

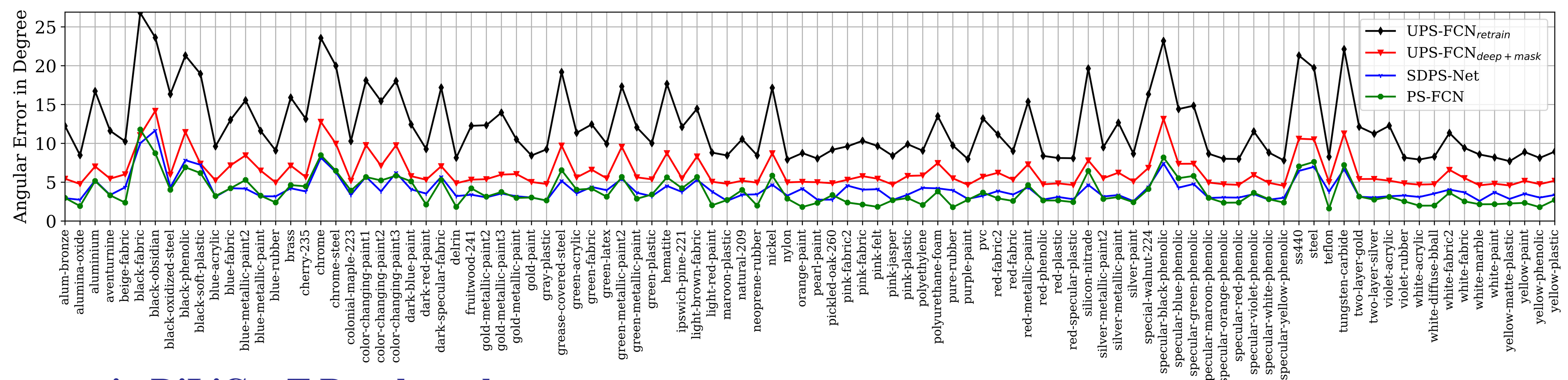



(a) DiLiGenT (b) APPLE (c) GOURD1 (d) GOURD2 (e) Light Stage

### Quantitative Results on BUNNY Rendered with 100 MERL BRDFs:



(a) Light source (b) BUNNY




### Results on DiLiGenT Benchmark [2]:

(a) Results on light direction estimation.

| Method | BALL | CAT | POT1 | BEAR | POT2 | BUDDHA | GOBLET | READING | COW | HARVEST | Avg. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\text{LCNet}_{\text{reg}}$ | 4.94 | 5.82 | 5.62 | 7.19 | 4.82 | **3.90** | 12.89 | 7.90 | **4.19** | 9.50 | 6.68 |
| LCNet | **3.27** | **4.08** | **5.44** | **3.47** | **2.87** | 4.34 | **10.36** | **4.50** | 4.52 | **6.32** | **4.92** |

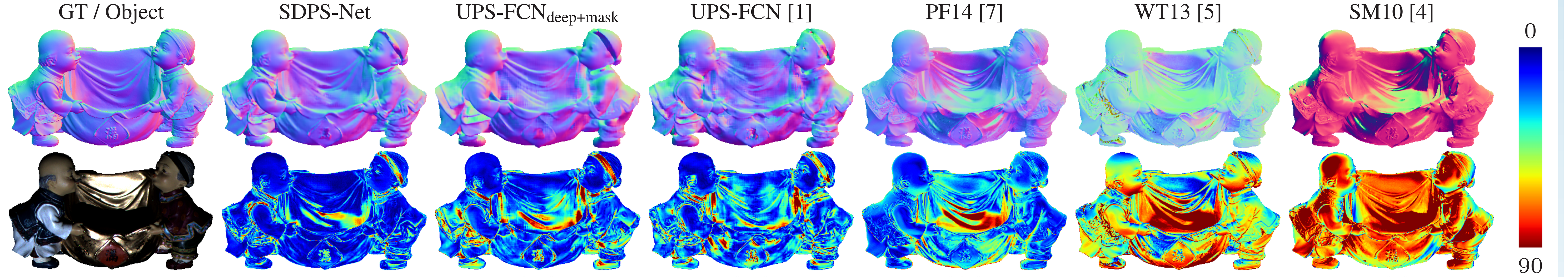
(b) Results on light intensity estimation.

| Method | BALL | CAT | POT1 | BEAR | POT2 | BUDDHA | GOBLET | READING | COW | HARVEST | Avg. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\text{LCNet}_{\text{reg}}$ | **0.032** | **0.051** | **0.048** | 0.167 | 0.074 | 0.080 | 0.075 | 0.141 | **0.044** | 0.085 | 0.080 |
| LCNet | 0.039 | 0.095 | 0.058 | **0.061** | **0.048** | **0.048** | **0.067** | **0.105** | 0.073 | **0.082** | **0.068** |

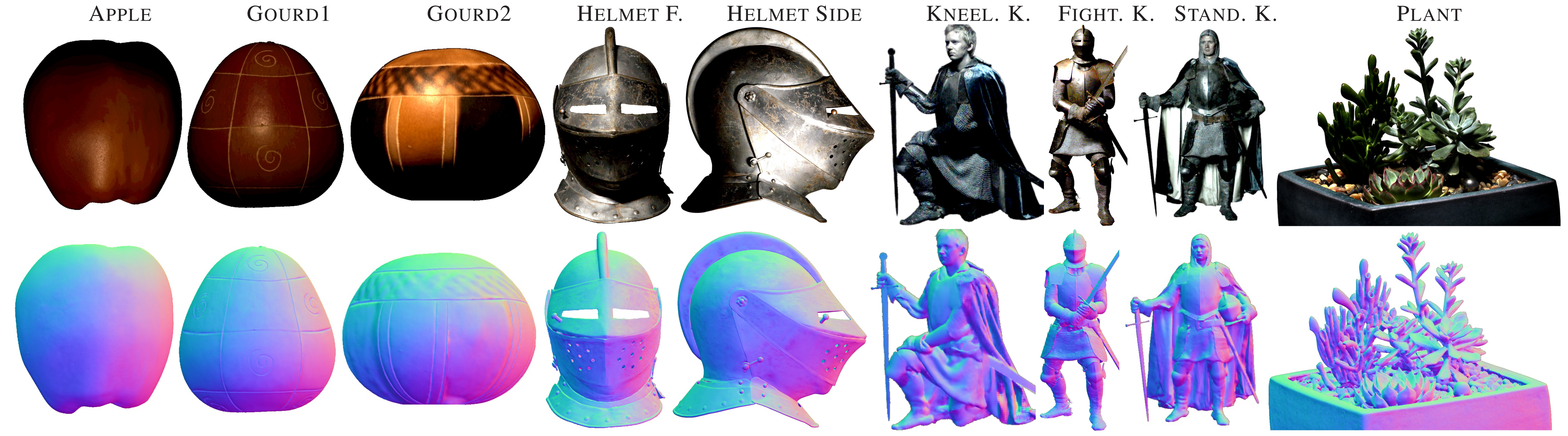
(c) Results on normal estimation.

| Method | BALL | CAT | POT1 | BEAR | POT2 | BUDD. | GOBL. | READ. | COW | HARV. | Avg. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| AM07 [3] | 7.3 | 31.5 | 18.4 | 16.8 | 49.2 | 32.8 | 46.5 | 53.7 | 54.7 | 61.7 | 37.3 |
| SM10 [4] | 8.9 | 19.8 | 16.7 | 12.0 | 50.7 | 15.5 | 48.8 | 26.9 | 22.7 | 73.9 | 29.6 |
| WT13 [5] | 4.4 | 36.6 | 9.4 | **6.4** | 14.5 | 13.2 | 20.6 | 59.0 | 19.8 | 55.5 | 23.9 |
| LM13 [6] | 22.4 | 25.0 | 32.8 | 15.4 | 20.6 | 25.8 | 29.2 | 48.2 | 22.5 | 34.5 | 27.6 |
| PF14 [7] | 4.8 | 9.5 | 9.5 | 9.1 | 15.9 | 14.9 | 29.9 | 24.2 | 19.5 | 29.2 | 16.7 |
| LC18 [8] | 9.3 | 12.6 | 12.4 | 10.9 | 15.7 | 19.0 | 18.3 | 22.3 | 15.0 | 28.0 | 16.3 |
| UPS-FCN [1] | 6.6 | 14.7 | 14.0 | 11.2 | 14.2 | 15.9 | 20.7 | 23.3 | 11.9 | 27.8 | 16.0 |
| LCNet + L2 [9] | 4.90 | 11.12 | 9.72 | 9.35 | 14.70 | 14.86 | 18.29 | 20.11 | 25.08 | 29.17 | 15.73 |
| LCNet + IS18 [10] | 6.37 | 15.64 | 10.58 | 8.48 | 12.24 | 13.94 | 18.54 | 23.78 | 29.31 | 25.69 | 16.46 |
| SDPS-Net | **2.8** | **8.1** | **8.1** | 6.9 | **7.5** | **9.00** | **11.9** | **14.9** | **8.5** | **17.4** | **9.5** |

### Quantitaive Comparison on HARVEST in DiLiGenT Benchmark:




### Qualitative Results on Gourd&Apple Dataset [11] and Light Stage Data Gallery [12]:




### References:


[1] UPS-FCN [Chen *et al.*, ECCV8]
[2] DiLiGenT [Shi *et al.*, TPAMI19]
[3] AM07 [Alldrin *et al.*, ICCV07]
[4] SM10 [Shi *et al.*, CVPR10]
[5] WT13 [Wu and Tan, CVPR13]
[6] LM13 [Lu *et al.*, CVPR13]
[7] PF14 [Papadhimitri and Favaro, IJCV14]
[8] LC18 [Lu *et al.*, TPAMI18]
[9] L2 [Woodham, OE1980]
[10] IS18 [Ikehata, ECCV18]
[11] Gourd&Apple [Alldrin *et al.*, CVPR08]
[12] Light Stage [Einarsson *et al.*, EGSR06]