Self-calibrating Deep Photometric Stereo Networks
Supplementary Materials

Guanying Chen¹  Kai Han²  Boxin Shi³,⁴  Yasuyuki Matsushita⁵  Kwan-Yee K. Wong¹
¹The University of Hong Kong  ²University of Oxford
³Peking University  ⁴Peng Cheng Laboratory  ⁵Osaka University

Contents

1. Regression Based Lighting Estimation Model LCNet_{reg}  2
2. Detailed Lighting Estimation Results of SDPS-Net on BUNNY  3
3. Detailed Normal Estimation Results of SDPS-Net on BUNNY  4
4. Different Network Architectures for the Single-stage Model  4
5. Qualitative Results on the DiLiGenT Dataset  5
6. Qualitative Results on the Light Stage Data Gallery  8
7. Qualitative Results on the Gourd&Apple Dataset  10
1. Regression Based Lighting Estimation Model LCNet_{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_{reg}, which shares the same architecture with LCNet, except that LCNet_{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

$$L_{\text{Reg}} = \lambda_l \frac{1}{q} \sum_{i} (1 - \hat{t}_i^\top \tilde{t}_i) + \lambda_e \frac{1}{q} \sum_{i} (e_i - \tilde{e}_i)^2,$$

(1)

where $\lambda_l$ and $\lambda_e$ are the weighting factors for the loss terms, $\hat{t}_i$ ($e_i$) and $\tilde{t}_i$ ($\tilde{e}_i$) denote the predicted light direction (intensity) and the ground truth, respectively, for image $i$. During training, $\lambda_l$ and $\lambda_e$ 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_{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_{reg-ϕθ}, to regress the azimuth and elevation angle (i.e., $\phi$ and $\theta$) instead of a 3-vector light direction. Table 1 (ID 0 & 2) show that directly regressing ($\phi$, $\theta$) decreased the performance of light direction estimation.

<table>
<thead>
<tr>
<th>ID</th>
<th>Model</th>
<th>SPHERE</th>
<th></th>
<th>BUNNY</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>LCNet_{reg}</td>
<td>4.10</td>
<td>0.104</td>
<td>5.46</td>
<td>0.094</td>
</tr>
<tr>
<td>1</td>
<td>LCNet_{reg-coupled}</td>
<td>4.03</td>
<td>0.103</td>
<td>6.97</td>
<td>0.095</td>
</tr>
<tr>
<td>2</td>
<td>LCNet_{reg-ϕθ}</td>
<td>4.57</td>
<td>0.083</td>
<td>5.87</td>
<td>0.091</td>
</tr>
</tbody>
</table>
2. Detailed Lighting Estimation Results of SDPS-Net on **BUNNY**

(a) Detailed lighting estimation results of LCNet on **BUNNY** from MERL 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\textsubscript{retrain}, UPS-FCN\textsubscript{deep+mask}, SDPS-Net, and PS-FCN on BUNNY from MERL\textsuperscript{Tm} (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\textsubscript{est\_light} to handle inputs with varying scales.
5. Qualitative Results on the DiLiGenT Dataset

![Qualitative Results for BALL, CAT, and BEAR](image)

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 rightmost column visualizes the error distributions of the light direction estimation.
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

![Qualitative results on Gourd&Apple dataset](Image)

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

References

