Deep Outdoor Illumination Estimation

Yannick Hold-Geoffroy, Kalyan Sunkavalli, Sunil Hadap, Emiliano Gambaretto, Jean-François Lalonde
Université Laval, Adobe
Capturing lighting conditions

[Debevec 1998]
Illumination capture and estimation

1. Specialized equipment
   [Stumpfel et al. '04]

2. Known object in scene
   [Lombardi et al. '15]

3. Handcrafted features
   [Lalonde et al. '12]
Our work

1. Specialized equipment
[Stumpfel et al. '04]

Conventional camera

2. Known object in scene
[Lombardi et al. '15]

Generic scene

3. Handcrafted features
[Lalonde et al. '12]

Learned features
Input → CNN → Light → Output
How do we train such an approach?
How to train?

Image

Lighting conditions

...
SUN360 dataset
[Xiao et al. CVPR’12]

- 45,000 outdoor panoramas
- Vast diversity of illumination conditions
Problem: SUN360 is Low Dynamic Range

LDR lighting

HDR lighting
Hošek-Wilkie Sky Model

\[ f^c_h(l; q_h) = \omega \left[ f^c_{sun}(l_{sun}, t) + f^c_{sky}(l_{sun}, t) \right] \]
Hošek-Wilkie Sky Model

\[
    f_h^c(l; q_h) = \omega \left[ f_{sun}^c(l_{sun}, t) + f_{sky}^c(l_{sun}, t) \right]
\]

\[
    q_h = [t, l_{sun}]
\]
Hošek-Wilkie Sky Model

\[ f_h^c(l; q_h) = \omega \left[ f_{sun}^c(l_{sun}, t) + f_{sky}^c(l_{sun}, t) \right] \]
Full panorama: millions of pixels

Sky model: 4 parameters
Framing the problem as one of end-to-end learning
# CNN Architecture and Training

<table>
<thead>
<tr>
<th>Layer</th>
<th>Stride</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td></td>
<td>320 × 240</td>
</tr>
<tr>
<td>conv7-64</td>
<td>2</td>
<td>160 × 120</td>
</tr>
<tr>
<td>conv5-128</td>
<td>2</td>
<td>80 × 60</td>
</tr>
<tr>
<td>conv3-256</td>
<td>2</td>
<td>40 × 30</td>
</tr>
<tr>
<td>conv3-256</td>
<td>1</td>
<td>40 × 30</td>
</tr>
<tr>
<td>conv3-256</td>
<td>2</td>
<td>20 × 15</td>
</tr>
<tr>
<td>conv3-256</td>
<td>1</td>
<td>20 × 15</td>
</tr>
<tr>
<td>conv3-256</td>
<td>2</td>
<td>10 × 8</td>
</tr>
<tr>
<td><strong>FC-2048</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>FC-160</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LogSoftMax</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>FC-5</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Output: sun position distribution

Output: turbidity and scale factor
Sun position results
Sun position result
Quantitative results — sun azimuth estimation

(a) Subset of SUN 360

Lalonde et al, 2012

SUN360 subset
Virtual object insertion
Insertion results on an HDR validation set

Network Output

Ground truth

Sun position estimation
Recovering camera parameters

\[ \mathbf{q}_h = [ t, \mathbf{l}_{\text{sun}}, \theta_{\text{cam}}, \alpha_{\text{cam}} ] \]

Camera pitch

Camera Field of View
Camera parameters estimation

Camera pitch: (est.) -2°
    (real) -1°
Field of View: (est.) 51°
    (real) 59°

Camera pitch: (est.) -16°
    (real) -16°
Field of View: (est.) 46°
    (real) 45°
Conclusion

SUN360 dataset

CNN

model parameters
Sun position
Atmospheric turbidity
Exposure
Camera pitch and FoV