Towards High-quality Intrinsic Images in the Wild

Gang Fu\textsuperscript{1}, Qing Zhang\textsuperscript{2}, and Chunxia Xiao\textsuperscript{1}

\textsuperscript{1}Wuhan University, \textsuperscript{2}Sun Yat-sen University

Wuhan University

July 7, 2019
Intrinsic images

- Decompose an input image into an albedo image and a shading image

(a) Input  (b) Albedo  (c) Shading
Our key observations

Our method is built upon two observations on real-world images:

- Reflectance is generally sparse and there are limited number of reflectance values in an image
- Shading usually has locally smooth transition
- Shading also tends to be towards a constant
Intrinsic Decomposition Model

We formulate the intrinsic image decomposition problem as the minimization of the following objective function including the four terms:

$$\arg\min_{R,S} \|I - R \times S\|^2_2 + \lambda_r \|\nabla R\|_0 + \lambda_s \|\nabla S\|^2 + \lambda_a \|S - \overline{S}\|^2_2$$

- Date term
- Reflectance sparseness term
- Shading smoothness term
- Absolute scale term
Model Solver

We adopt the alternating direction method of multipliers (ADMM) technique to solve the optimization problem model.

\[
\arg \min_{R,S,G,X} \| I - R \times S \|^2 + \lambda_s \| \nabla S \|^2 + \lambda_a \| S - \overline{S} \|^2 \\
+ \lambda_r \{ \| G \|_0 + \mu \| \nabla R - G + X \|^2 \}, \quad \text{s.t.} \quad I \leq S
\]

Figure: Results from different iterations and the convergence curves for reflectance (top) and shading (bottom).
Visual comparison

- Our method can effectively separate the albedo and shading from the input image.
- Other methods either fail to distinguish shading variations from the albedo variations, or damage the structures in the albedo image and make the intrinsic images unrealistic.

Figure: Comparison between our proposed method with the state-of-the-arts.
Visual comparison with more methods

Please also refer to our supplemental material for more results
Ablation study

- Reflectance sparseness term using different norm

(a) Input  (b) $L_1$ norm  (c) Ours ($L_0$ norm)

- With and without shading smoothness term

(a) Input  (b) Without $f_s$  (c) With $f_s$
## Ablation study

Table: Quantitative comparison of different algorithm variants of our method on the IIW and SAW datasets in terms of WHDR and SAW, respectively.

<table>
<thead>
<tr>
<th>Algorithm variant</th>
<th>WHDR (IIW)</th>
<th>AP (SAW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without ( f_r ) (L₀)</td>
<td>23.54%</td>
<td>95.23%</td>
</tr>
<tr>
<td>( f_r ) (L₀) using L₁ instead</td>
<td>27.43%</td>
<td>92.23%</td>
</tr>
<tr>
<td>( f_r ) (L₀) using L₂ instead</td>
<td>38.23%</td>
<td>91.55%</td>
</tr>
<tr>
<td>Without ( f_s )</td>
<td>24.32%</td>
<td>93.12%</td>
</tr>
<tr>
<td>Without ( f_a )</td>
<td>29.34%</td>
<td>94.52%</td>
</tr>
<tr>
<td>Ours (full method)</td>
<td><strong>19.20%</strong></td>
<td><strong>98.80%</strong></td>
</tr>
</tbody>
</table>
Quantitative evaluation

Table: Quantitative comparison between our method and the state-of-the-art methods on the IIW and SAW datasets.

<table>
<thead>
<tr>
<th>Method</th>
<th>WHDR (IIW)</th>
<th>AP (SAW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retinex-color 2009</td>
<td>26.89%</td>
<td>91.93%</td>
</tr>
<tr>
<td>Shen et al. 2011</td>
<td>36.90%</td>
<td>87.23%</td>
</tr>
<tr>
<td>Garces et al. 2012</td>
<td>25.46%</td>
<td>96.89%</td>
</tr>
<tr>
<td>Zhao et al. 2012</td>
<td>23.20%</td>
<td>97.11%</td>
</tr>
<tr>
<td>Bell et al. 2017</td>
<td>20.64%</td>
<td>97.37%</td>
</tr>
<tr>
<td>Fu et al. 2016</td>
<td>36.43%</td>
<td>91.87%</td>
</tr>
<tr>
<td>Li and Snavely 2018</td>
<td>20.30%</td>
<td>97.90%</td>
</tr>
<tr>
<td>Ours</td>
<td><strong>19.20%</strong></td>
<td><strong>98.80%</strong></td>
</tr>
</tbody>
</table>
Application to low-light image enhancement

Figure: Low-light image enhancement comparison. **Left:** Input. **Middle:** LIME. **Right:** Ours. We can see that, the LIME [17] overexposes the leaves and the sky, while our method produces more appealing result with moderate brightness, clear details, distinct contrasts and vivid colors.
Limitation

- It may wrongly assign hard shadows that violate the Lambertian assumption to the reflectance layer
- Our current implementation based on CPU is not fast enough for real-time performance

Figure: Failure case of our method. **Left:** Input. **Middle:** Albedo. **Right:** Shading. We can see that, the cast shadow led by shading discontinuities also exists in the reflectance layer.
Conclusion and future work

• We have presented a novel method for estimating high-quality intrinsic images for real-world scenes
• An $L_0$ norm is utilized to encourage the reflectance sparseness while preserving the prominent structures
• A total variation based shading smoothness term is adopted to avoid the texture-copy problem
• Future work
  • We would like to adopt the scene semantic and the depth information to further refine our model
  • We will improve the performance of our method by GPU implementation
End

- The paper, supplementary material, poster, presentation and code will be available in our lab homepage graphvision.whu.edu.cn.
End

Q & A