Reflectance Adaptive Filtering Improves Intrinsic Image Estimation

Thomas Nestmeyer[1,3], Peter V. Gehler[1,2,3]
[1]University of Würzburg; [2]BCCN, Tübingen; [3]MPI for Intelligent Systems, Tübingen

The Intrinsic Images Problem

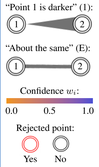
Decompose single image into its reflectance and shading layers

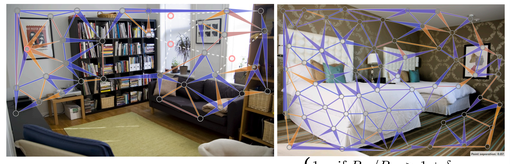
Input Reflectance Shading

Reflectance: Physical property of objects, invariant under different lighting conditions.
Shading: Separates scene illumination: number, location, and color of the light sources, light occlusion by geometry, etc.

Intrinsic Images in the Wild (IIW)

5230 Flickr images having each about 100 pairwise relative reflectance judgments from humans [1].



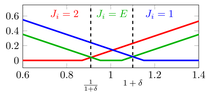


$$\mathrm{WHDR}_\delta(J,R) = \frac{\sum_i w_i \cdot \mathbb{1}\left(J_i \neq \hat{J}_\delta(R,i)\right)}{\sum_i w_i} \qquad \hat{J}_\delta(R,i) = \begin{cases} 1 & \text{if } R_{q_i}/R_{p_i} > 1+\delta \\ 2 & \text{if } R_{p_i}/R_{q_i} > 1+\delta \\ E & \text{else} \end{cases}$$

Direct CNN Prediction

Novel loss function for training, the *WHDR-Hinge* loss

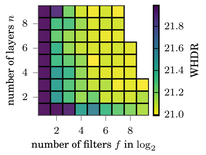
$$\ell_{\delta,\xi}(J,R,i) = \begin{cases} \max\left(0,\ \frac{R_{p_i}}{R_{q_i}} - \frac{1}{1+\delta+\xi}\right) & \text{if } J_i = 1 \\ \max\left(0, \left\{\frac{1}{1+\delta-\xi} - \frac{R_{p_i}}{R_{q_i}}, \frac{R_{p_i}}{R_{q_i}} - (1+\delta-\xi)\right\}\right) & \text{if } J_i = E \\ \max\left(0,\ 1+\delta+\xi - \frac{R_{p_i}}{R_{q_i}}\right) & \text{if } J_i = 2, \end{cases}$$

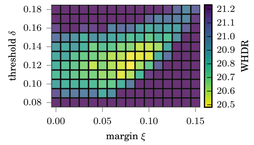


Visualization of WHDR-Hinge

Full resolution CNN utilizing ReLUs and skip connections.

Pixelwise Reflectance Prediction



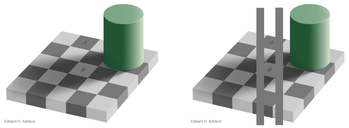


Insight: 1x1 convolutions work just as well as bigger kernels

Input Scalar reflectance prediction by CNN Reconstructed Reflectance

Predictions for IIW ID 97794:

Input Scalar reflectance Reflectance Shading



- While small receptive fields might work to infer the *true* reflectance (Retinex theory), estimating it from a single input pixel is not possible.
- Still: Performance on par with state-of-the-art!
- Very fast reflectance predictions by construction
- Many small variations in reflectance prediction`

Take home message

- CNN method allowing *end-to-end training on sparse relative reflectance judgments*.
- Learned *pixelwise* non-linear *reflectance prediction* with *competitive WHDR*.
- Allows *real time intrinsic video* (~180 fps on GPU).
- Use *reflectance adaptive filtering* to encourage piecewise constant reflectance assumption.
- Filtered results outperform the current state-of-the-art by far.

Reflectance Adaptive Filtering

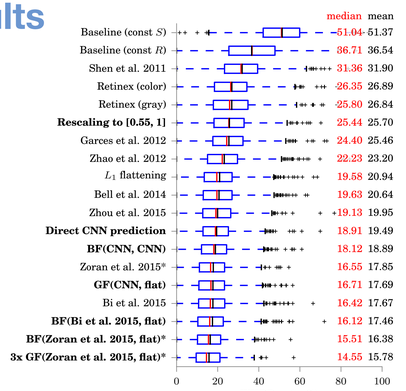
New idea: Encourage piecewise constant reflectance by filtering reflectance predictions.
Bi et al. [2] formulated an optimization problem inferring an L1-flattened image (grouping pixels into regions of similar reflectance) as part of their pipeline. We call this the 'flat guidance image'.
We can filter any reflectance estimate using this guidance.

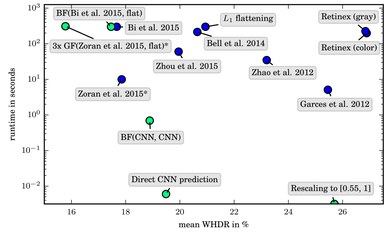
Input Reflectance of [3] Our filtered reflectance

Flat guidance image [2] Shading of [3] Our shading

Guided filtering [5] is very fast, but a reflectance prediction as input and a guidance image is needed. We are working on speeding the most expensive latter part up.

Results





References:
1. Sean Bell et al. "Intrinsic Images in the Wild" (*SIGGRAPH 2014*)
2. Bi et al. "An L1 image transform for edge-preserving smoothing and scene-level intrinsic decomposition" (SIGGRAPH 2015)
3. Zoran et al. "Learning ordinal relationships for mid-level vision" (ICCV 2015)
4. Petschnigg et al. "Digital photography with flash and no-flash image pairs" (SIGGRAPH 2004)
5. He et al. "Guided image filtering" (ECCV 2010)

{thomas.nestmeyer, peter.gehler}@tuebingen.mpg.de Perceiving Systems https://ps.is.tue.mpg.de/research_projects/reflectance-filtering