Multi-view Multi-illuminant Intrinsic Dataset

Shida Beigpour
shida@mpi-inf.mpg.de
Mai Lan Ha
hamailan@informatik.uni-siegen.de
Sven Kunz
sven.kunz83@gmail.com
Andreas Kolb
andreas.kolb@uni-siegen.de
Volker Blanz
blanz@informatik.uni-siegen.de

Institute for Vision and Graphics
University of Siegen
Siegen, Germany

Decomposing an image into its intrinsic components (e.g. reflectance and shading) is a fundamental concept in computer vision. New trends in computer vision such as fusion of colour and depth (RGB-D) as well as reconstructing 3D scenes from photo collections strongly benefit from a correct reflectance estimation in multi-view scenes. Producing the true colour and texture of the surface (reflectance) is highly challenging due to the presence of shadows and illumination colours which result in strong artefacts unless a correct reflectance estimation is performed. In these topics, multi-view intrinsic image estimation is desired.

This paper proposes a novel high-resolution multi-view dataset of complex multi-illuminant scenes with precise reflectance and shading ground-truth as well as raw depth and 3D point cloud. Our dataset challenges the intrinsic image methods by providing complex coloured cast shadows, highly textured and colourful surfaces, and specularity. This is the first publicly available multi-view real-photo dataset at such complexity with pixel-wise intrinsic ground-truth. Our work improves over the state-of-the-art intrinsic datasets [1, 2, 3]. In the effort to help evaluating different intrinsic image methods, we propose a new perception-inspired metric that is based on the reflectance consistency. We provide the evaluation of three intrinsic image methods using our dataset and metric.

Fig. 1 demonstrates an example of a scene captured with six different cameras under different illumination conditions along with its raw depth, point cloud, and a rough surface reconstruction. Here the advantage of using reflectance instead of the captured pixel colour for the 3D surface colour is evident. In total the dataset consists of 20 illumination conditions, 5 scenes, and 6 cameras. Fig. 2 shows the five scenes as seen by one of the cameras. Our complete dataset consists of 600 high-resolution (5208×3476) images along with their ground-truth and is publicly available online at:

http://www.cg.informatik.uni-siegen.de/data/iccv2015/intrinsic/

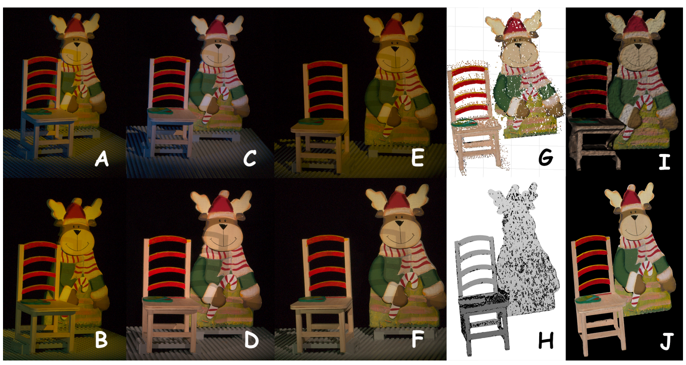

Figure 1: Examples of different views and illuminations (A-F), rough point cloud (G), raw depth (H), 3D surface (I), and ground-truth reflectance (J).

In addition, our proposed Point-wise Consistency Metric (PCM) evaluates the estimated intrinsic image objectively using the perceptual colour distance CIE DE2000 [4]. Given the ground-truth reflectance image G and the estimated reflectance image T, the PCM algorithm samples a set of point pairs $\mathcal{S} = \{(p,q)\}$ in G that are *perceptually similar* using CIE DE2000. The *point-wise consistency error* $PCE(G,T)$ is computed as follows:

$$PCE(G,T) = \frac{1}{|\mathcal{S}|} \sum_{(p,q)\in\mathcal{S}} f(pce_{G,T}(p,q)), \quad (1)$$

$$f(pce_{G,T}(p,q)) = \begin{cases} 1, & pce_{G,T}(p,q) > \sigma \\ \frac{pce_{G,T}(p,q)}{\sigma}, & otherwise \end{cases}, \quad (2)$$

$$pce_{G,T}(p,q) = \triangle E_{00}(G(p) - G(q), T(p) - T(q)), \quad (3)$$

where $pce_{G,T}(p,q)$ is the difference between the similarity of the points in the ground-truth and the estimated reflectances for a pair of points (p,q). $\triangle E_{00}$ denotes CIE DE2000 colour distance. f is a linear function to normalize $pce_{G,T}(p,q)$ to $[0,1]$ with a cut-off threshold σ.

Figure 2: The five scenes captured by one of the six cameras.

Using our dataset and metric, we evaluate 3 state-of-the-art intrinsic methods and compare against the Linear Mean Square Error (LMSE) [3] metric. Fig. 3 shows that PCM delivers more consistent results across 20 illuminations, whereas LMSE exhibits large variations.

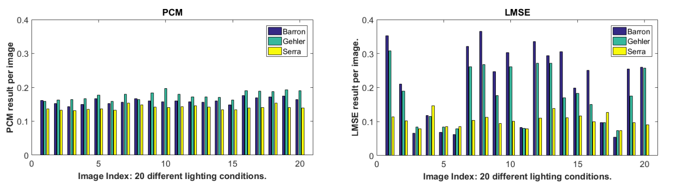

Figure 3: Qualitative results comparison between PCM and LMSE over 20 different illumination conditions.

In summary, we contribute to intrinsic image research a challenging multi-view dataset with multi-coloured illuminations, rich texture, high resolution ground-truth, depth, and dense point cloud. We further introduce a perception-inspired metric that objectively evaluates the estimated reflectance against the ground-truth. We believe that our dataset and metric can help in improving the quality of intrinsic image methods in complex scenes and lighting conditions. Please refer to our paper for more information.

[1] Shida Beigpour, Andreas Kolb, and Sven Kunz. A comprehensive multi-illuminant dataset for benchmarking of the intrinsic image algorithms. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 172–180, 2015.

[2] Sean Bell, Kavita Bala, and Noah Snavely. Intrinsic images in the wild. *ACM Transactions on Graphics (TOG)*, 33(4):159, 2014.

[3] Roger Grosse, Micah K Johnson, Edward H Adelson, and William T Freeman. Ground truth dataset and baseline evaluations for intrinsic image algorithms. In *2009 IEEE 12th International Conference on Computer Vision*, pages 2335–2342. IEEE, 2009.

[4] M Ronnier Luo, Guihua Cui, and B Rigg. The development of the CIE 2000 colour-difference formula: CIEDE2000. *Color Research & Application*, 26 (5), 2001.