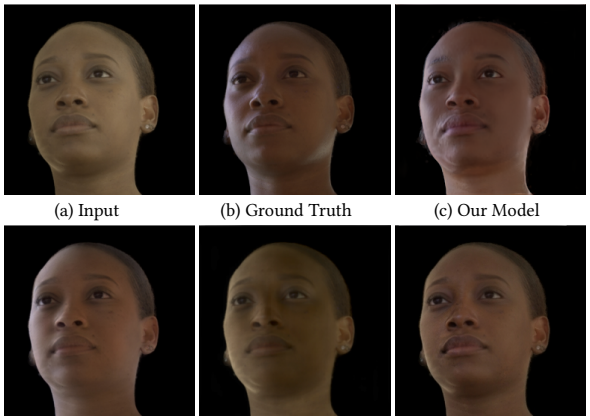
Supplemental material: Single Image Portrait Relighting via Explicit Multiple Reflectance Channel Modeling

ZHIBO WANG, BNRist and school of software, Tsinghua University, China
XIN YU, University of Technology Sydney, Australia
MING LU, Intel Labs, China
QUAN WANG, SenseTime, China
CHEN QIAN, SenseTime, China
FENG XU, BNRist and school of software, Tsinghua University, China

ACM Reference Format:
Zhibo Wang, Xin Yu, Ming Lu, Quan Wang, Chen Qian, and Feng Xu. 2020. Supplemental material: Single Image Portrait Relighting via Explicit Multiple Reflectance Channel Modeling. *ACM Trans. Graph.* 39, 6, Article 220 (December 2020), 4 pages. https://doi.org/10.1145/3414685.3417824



(a) Input (b) Ground Truth (c) Our Model
(d) [Sun et al. 2019] (e) [Zhou et al. 2019] (f) [Shih et al. 2014]

Fig. 1. Comparisons of portrait relighting on paired images.

Fig. 2. Illustration of normal estimation and relit results. In each row, we show the source image (left), the target lighting (left bottom), the normal estimation (middle) and the relit result (right).

1 ADDITIONAL RESULTS

Although there might be some errors in the normal estimation, such as the areas of glasses and hats, our model generates convincing relighting effects on these images, as shown in Figure 2. To better evaluate the relighting quality, we demonstrate the relit results of images provided by the paper [Sun et al. 2019], which are generated from "one-light-at-a-time" (OLAT) images. As shown in Figure 1, the specular effects are not reconstructed by [Shih et al. 2014; Zhou et al. 2019]. The specular in [Sun et al. 2019] is overly-smoothed, while our method achieves realistic specular effects. More relit results of our method are shown in Figure 3. Moreover, to demonstrate the high-quality of our captured face scans, we present some rendered albedo and normal images in Figure 4. Real-world images used in this document are from FFHQ dataset [Karras et al. 2019].

2 BLURRINESS

In the main paper, our model is trained on images synthesized by 16 samples per pixel (spp). Considering the trade-off between time cost and the rendering quality, we choose 16 spp at first to synthesize our

Authors' addresses: Zhibo Wang, BNRist and school of software, Tsinghua University, 100091, Beijing, China, wzb17@mails.tsinghua.edu.cn; Xin Yu, University of Technology Sydney, Australia, xin.yu@uts.edu.au; Ming Lu, Intel Labs, China, lu199192@gmail.com; Quan Wang, SenseTime, China, wangquan@sensetime.com; Chen Qian, SenseTime, China, qianchen@sensetime.com; Feng Xu, BNRist and school of software, Tsinghua University, 100091, Beijing, China, xufeng2003@gmail.com.

0730-0301/2020/12-ART220 $15.00
https://doi.org/10.1145/3414685.3417824

relighting dataset. As the configuration of low spp may lead to noise in the rendered results, we apply a denoising filter to the rendered images. However, this denoising step will result in image blurriness in our dataset, as shown in Figure 5. Therefore, the model trained on the 16 spp dataset may suffer from some blurriness in the relit results. This can be alleviated by using images rendered by higher spp for training. Hence, we also construct a dataset synthesized by 64 spp. As shown in Figure 6, our model trained on images with 64 spp achieves richer facial details in the relit results. Images rendered by 16 and 64 spp will be released.

REFERENCES

Tero Karras, Samuli Laine, and Timo Aila. 2019. A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 4401–4410.

YiChang Shih, Sylvain Paris, Connelly Barnes, William T Freeman, and Frédo Durand. 2014. Style transfer for headshot portraits. *ACM Transactions on Graphics (TOG)* 33, 4 (2014), 148.

Tiancheng Sun, Jonathan T Barron, Yun-Ta Tsai, Zexiang Xu, Xueming Yu, Graham Fyffe, Christoph Rhemann, Jay Busch, Paul Debevec, and Ravi Ramamoorthi. 2019. Single image portrait relighting. *ACM Transactions on Graphics (Proceedings SIGGRAPH)* (2019).

Hao Zhou, Sunil Hadap, Kalyan Sunkavalli, and David W Jacobs. 2019. Deep Single-Image Portrait Relighting. In *Proceedings of the IEEE International Conference on Computer Vision*. 7194–7202.

Fig. 3. Our relit results. In each row, we show the source image (left), the target lighting (right bottom) and our relit result (right).

Fig. 4. Examples of our captured albedo and normal maps.

(a) spp=16 (b) spp=64 (c) spp=8196

Fig. 5. Synthetic images rendered by different spp. Images rendered by 8196 spp are regarded as the ground-truth. Images rendered by 16 spp suffer from blurriness and missing details. Those artifacts will be alleviated by using 64 spp.

(a) Input (b) spp=16 (c) spp=64

Fig. 6. Relit results of our models trained on images rendered by different spp. (a) Input images. (b) Results of our model trained on images rendered by 16 spp. (c) Results of our model trained on images rendered by 64 spp. The model trained on images rendered by 64 spp significantly achieves much richer facial details.