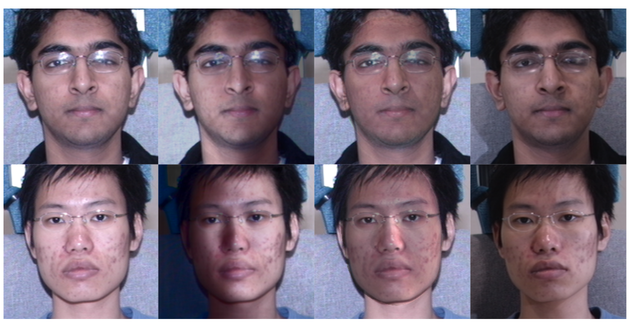
Face Relighting with Geometrically Consistent Shadows (Supplementary Materials)

Andrew Hou[*,1], Michel Sarkis[2], Ning Bi[2], Yiying Tong[1], Xiaoming Liu[1]
[1]Michigan State University, [2]Qualcomm Technologies Inc.
{houandr1, ytong, liuxm}@msu.edu, {msarkis, nbi}@qti.qualcomm.com
https://github.com/andrewhou1/GeomConsistentFR

Method	MSE	DSSIM	LPIPS
DPR [7]	0.0171	0.0796	0.1286
Proposed	**0.0080**	**0.0562**	**0.1268**

Table 1. **Quantitative Evaluation for General (Diffuse) Relighting** We outperform DPR quantitatively across all metrics in diffuse relighting on the Multi-PIE dataset.



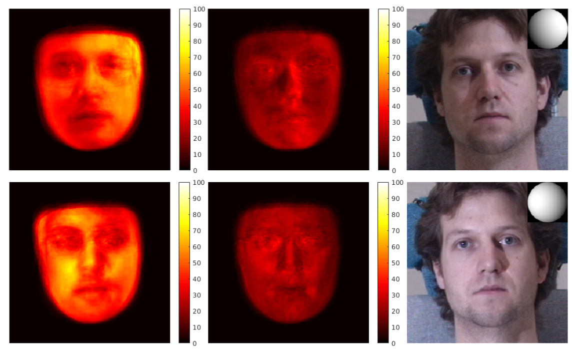
Input DPR [7] Proposed Target

Figure 1. **General (Diffuse) Relighting**. We outperform DPR qualitatively in diffuse relighting on the Multi-PIE dataset, where each input image is relit by averaging the predictions of 3 randomly selected target lightings. The groundtruth is the average of the 3 groundtruth Multi-PIE images.

1. Diffuse Relighting Evaluation

To compare with DPR [7] on more general, diffuse lightings, we follow their protocol and generate diffuse lighting groundtruth by averaging 3 random directional lighting images per Multi-PIE [1] subject. For both DPR and our method, we feed each of the 3 target lightings separately and average the predictions to generate the final relit image. We outperform DPR in general relighting both quantitatively and qualitatively (Tab. 1 and Fig. 1).

*All of the data mentioned in this paper was downloaded and used at Michigan State University.



Nestmeyer [5] Proposed Target Lighting

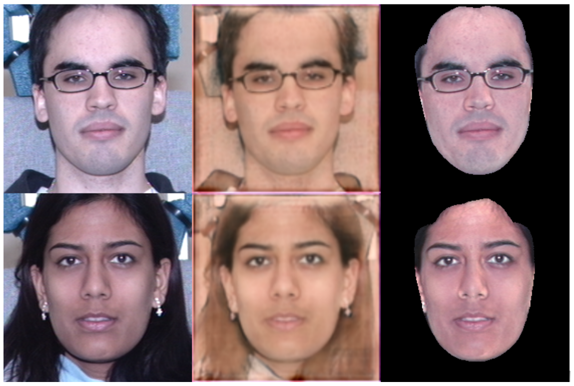
Figure 2. **Error Maps**. We visualize the average L_1 error for each Multi-PIE lighting's test subjects. Our method has significantly lower error around the hard shadow regions (nose and cheek) compared to Nestmeyer *et al.* [5], which demonstrates that our method produces more geometrically consistent hard shadows.

2. Geometric Consistency Comparison with Nestmeyer *et al.* [5]

To compare with the SoTA relighting method Hou *et al.* [2], we used the average L_1 error for each Multi-PIE lighting's test subjects to verify that our model was improving primarily around the hard shadow region in Fig. 6 of the main paper. We show the same error map to compare with Nestmeyer *et al.* [5] in Fig. 2. Our method has particularly low error in the hard shadow region (nose and cheek), whereas Nestmeyer *et al.* has high error in and around the shadow, especially for the first row's lighting. Our method thus produces more geometrically consistent hard shadows.

3. Albedo Comparison

Our albedo supervision from SfSNet [6] is far from perfect, as shown in Fig. 3, which is why we define the albedo loss in grayscale and not RGB. We adopt this supervision primarily because albedo supervision has limited options for single image in-the-wild datasets besides PCA,



Input Image SfSNet [6] Proposed

Figure 3. **Albedo Comparison**. Our method is able to produce high quality albedo despite the imperfect supervision from SfSNet [6] by keeping the albedo loss $\mathcal{L}_{albedo}$ in grayscale, which gives our model more freedom in the RGB space.

which often does not preserve facial details well. However, our model's estimated albedo clearly improves over SfSNet.

4. Comprehensive FFHQ Relighting Results

We strongly believe in diversity and the representation of all groups in the computer vision community. We therefore show a wide variety of relighting results with diversity and inclusion in mind. Our results cover as many racial groups as possible, as well as other factors such as different ages, genders, poses, expressions, subjects with facial hair, and the presence of glasses (See Fig. 4). We also increased the lighting diversity to demonstrate that our model can handle many different desired illuminations.

5. FFHQ Relighting Video

We include a video with 4 FFHQ [4] subjects where we rotate the light around the face, move the light horizontally, and move the light vertically. From left to right, we visualize the target lighting, the relighting results of Hou *et al.* [2], and our proposed method's relighting results. Our video demonstrates our high relighting quality as well as the geometric consistency of our shadows across many lightings. Compared to [2], it is clear that the shape of our shadows is superior, especially when comparing the first subject. We also modify the tone of the image significantly less, while [2] seems to frequently produce overly dark shadows. The video can be viewed here.

6. Licenses for Face Related Datasets

Although we don't collect any face data ourselves in this work, we do make use of existing face datasets, including Multi-PIE [1], FFHQ [4], and CelebA-HQ [3]. The Multi-PIE database was collected at Carnegie Mellon University, where all subjects agreed that their data would be used for research purposes. We only use the database internally for our work and primarily for evaluation. FFHQ consists of images published on Flickr, which are all under multiple licenses that allow free use, adaptation, and redistribution for noncommercial purposes. The creators also provide a way to remove an individual's photo from the dataset if they so desire. CelebA-HQ consists entirely of images collected from the internet. Although there is no associated IRB approval, the authors assert in the dataset agreement that the dataset is only to be used for noncommercial research purposes, which we strictly adhere to. Users must also agree not to sell, reproduce, or exploit any of the data and can only make copies of the data within their own organization, which we also adhere to.

References

[1] Ralph Gross, Iain Matthews, Jeffrey Cohn, Takeo Kanade, and Simon Baker. Multi-PIE. *Image and Vision Computing*, 2010. 1, 2

[2] Andrew Hou, Ze Zhang, Michel Sarkis, Ning Bi, Yiying Tong, and Xiaoming Liu. Towards high fidelity face relighting with realistic shadows. In *CVPR*, 2021. 1, 2

[3] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. Progressive growing of GANs for improved quality, stability, and variation. In *ICLR*, 2018. 2

[4] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In *CVPR*, 2019. 2

[5] Thomas Nestmeyer, Jean-Francois Lalonde, Iain Matthews, and Andreas Lehrmann. Learning physics-guided face relighting under directional light. In *CVPR*, 2020. 1

[6] Soumyadip Sengupta, Angjoo Kanazawa, Carlos D. Castillo, and David W. Jacobs. SfSNet: Learning shape, refectance and illuminance of faces in the wild. In *CVPR*, 2018. 1, 2

[7] Hao Zhou, Sunil Hadap, Kalyan Sunkavalli, and David W Jacobs. Deep single-image portrait relighting. In *ICCV*, 2019. 1

Figure 4. **Comprehensive and Diverse Relighting Performance on FFHQ**. Every two rows (*e.g.* c, d) shows the input image in the first row and our relighting results in the second row. We demonstrate our relighting performance on a wide variety of racial groups, genders, ages, expressions, and poses and also include subjects with facial hair and glasses. We find that our model is able to generalize to a wide range of subjects across many different lightings. Best viewed if enlarged.