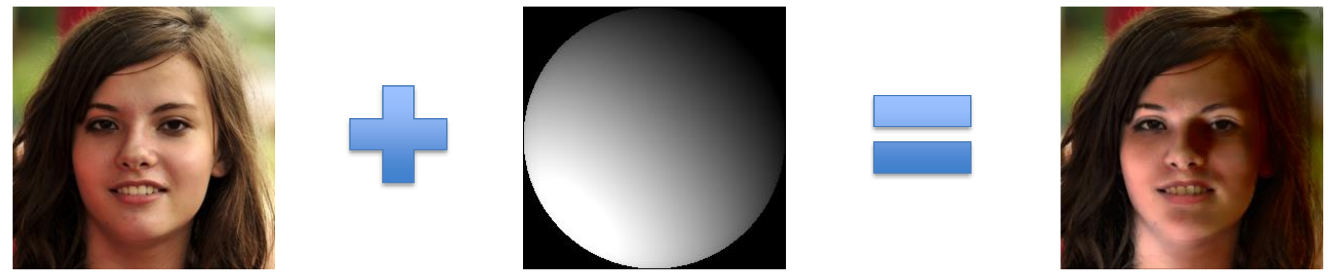
Towards High Fidelity Face Relighting with Realistic Shadows

Andrew Hou[1], Ze Zhang[1], Michel Sarkis[2], Ning Bi[2], Yiying Tong[1], Xiaoming Liu[1]
Michigan State University[1], Qualcomm Technologies Inc.[2]

CVPR VIRTUAL JUNE 19-25

Problem Overview



Input Image | Target Lighting | Relit Image

- Face relighting is the problem of turning a source image of a human face into a new image of the same face under a desired illumination different from the original lighting.

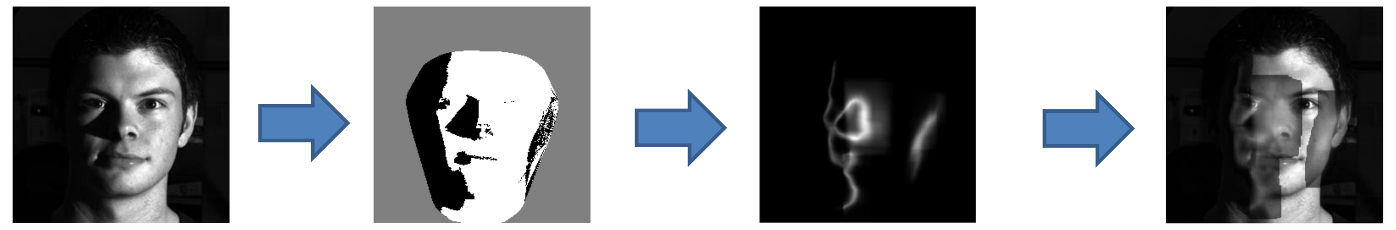
Related Work

- Prior face relighting methods can handle diffuse lightings, assumes light is evenly scattered across the face, produces soft shadows.
- Cannot handle strong directional light, which produces hard shadows such as cast shadows around the nose.
- Existing methods can also struggle to preserve the subject's facial details during relighting.

Contributions

- We propose a novel face relighting method that models both high-contrast cast shadows and soft shadows, while preserving the local facial details.
- Our technical approach involves single image based ratio image estimation to better preserve local details, shadow border reweighting to handle hard shadows, and ambient light compensation to account for dataset differences.
- Our approach achieves the state-of-the-art relighting results on two benchmarks quantitatively and qualitatively.

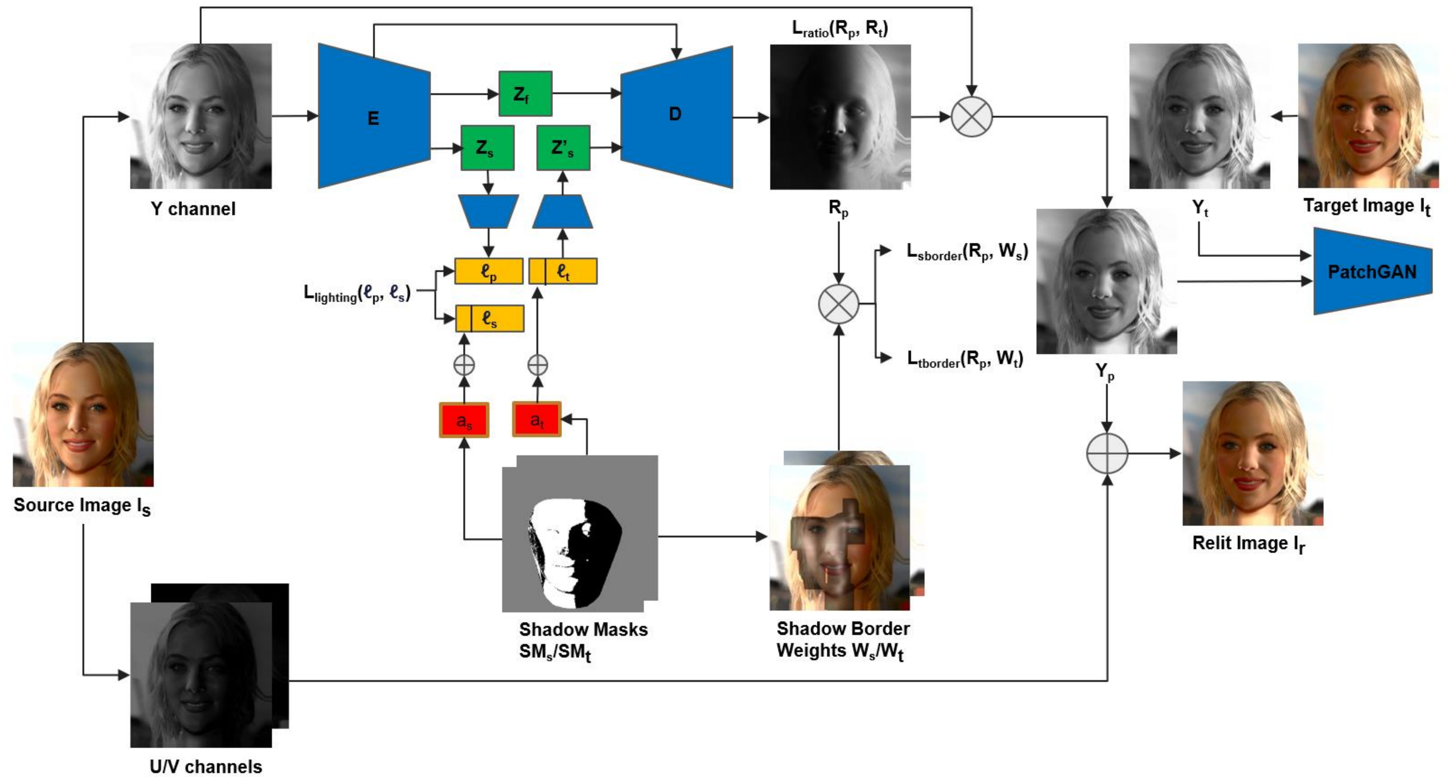
Shadow Mask Estimation and Shadow Border Weights



Input Image | Shadow Mask | Shadow Border Weights | Shadow Border Weights Overlaid on Image

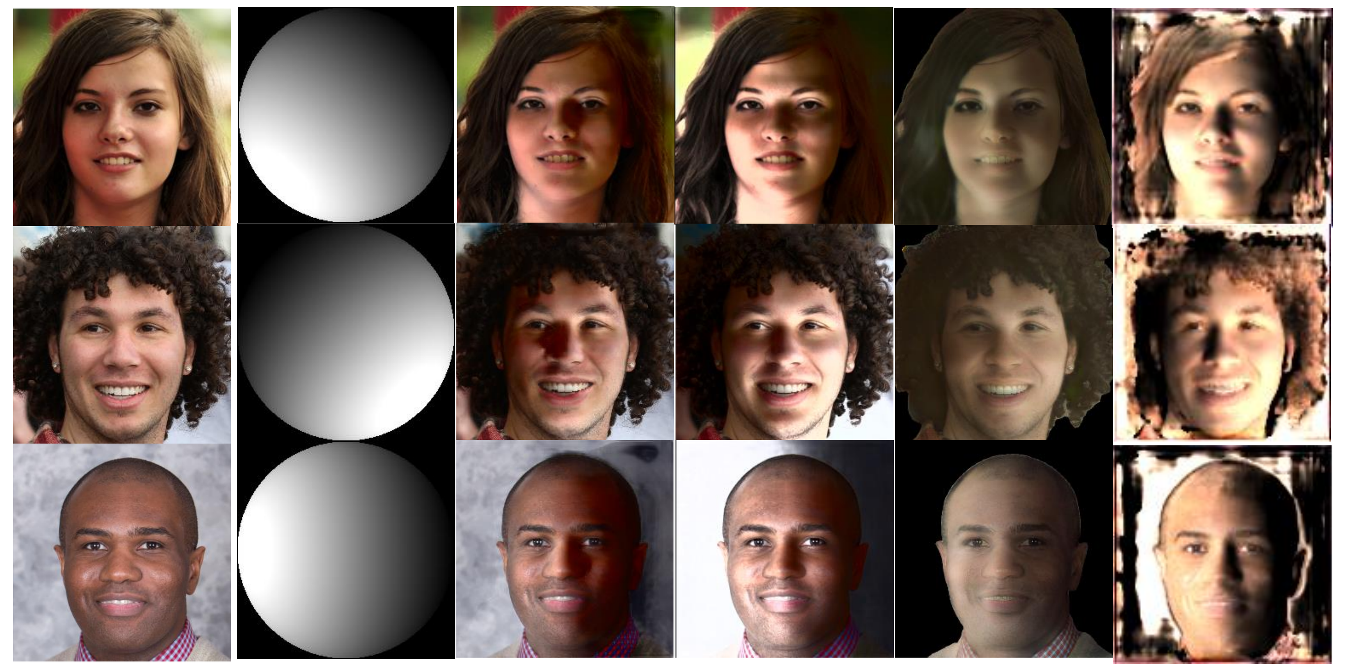
- We estimate the shadow mask through ray tracing using estimated 3D face geometry from 3DMM fitting and target light direction.
- Shadow mask shows where hard shadows lie.
- From the shadow mask, we compute where shadow borders lie.
- We apply higher weights during training on the pixels near the shadow borders.
- **Key insight**: gradient near the borders of hard shadows is often very steep and requires high weights to properly handle.

Model Architecture



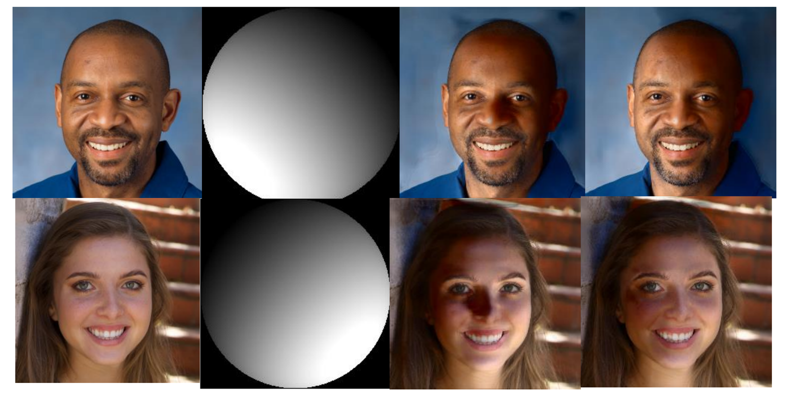
- Input image converted to YUV space.
- Input: Y channel and target lighting ℓ_t , Output: ratio image R_p and source lighting ℓ_p .
- Higher loss weights near shadow borders to properly handle the steep gradients.

FFHQ Relighting Performance (In-the-Wild)

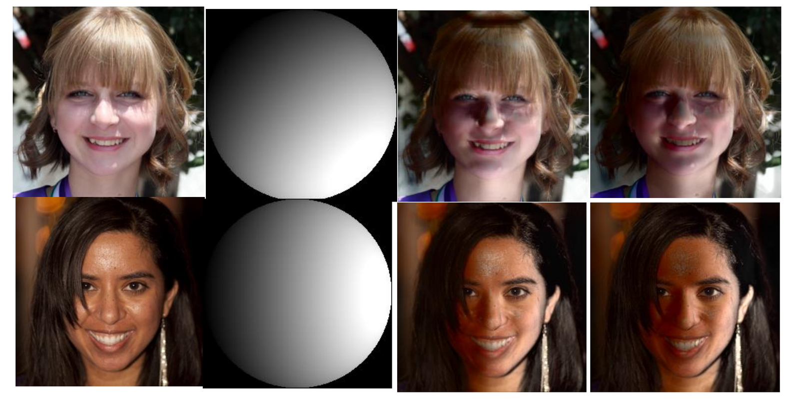


Input Image | Target Lighting | Our Method | DPR | SIPR | SfSNet

Ablation Studies



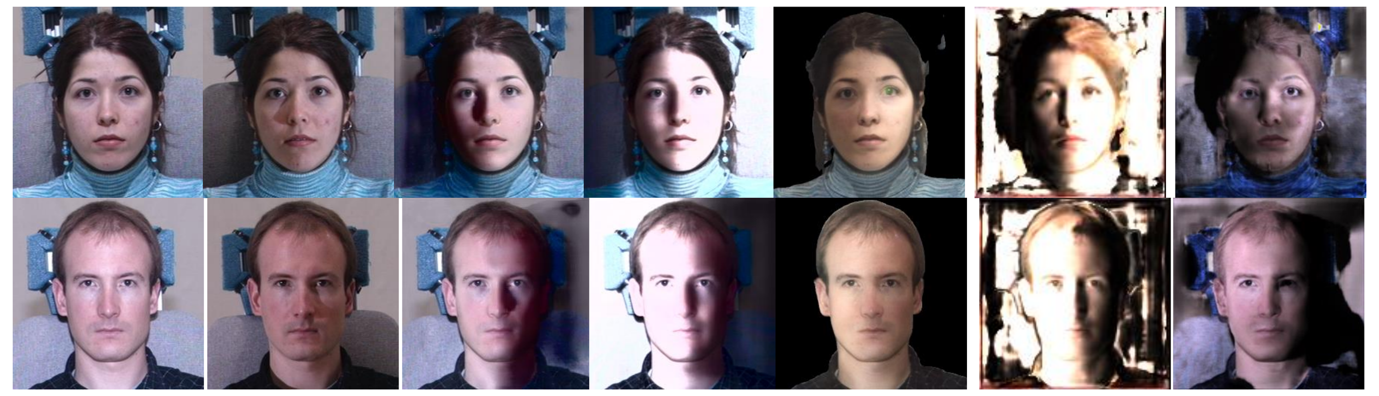
Input Image | Target Lighting | Proposed | w/o Shadow Border Weights



Input Image | Target Lighting | Proposed | w/o Ratio Image

Multi-PIE Experiments

Multi-PIE Evaluation Using Target Lightings



Input Image | Target Image | Our Model | DPR | SIPR | SfSNet | Nestmeyer

Method	Si-MSE	MSE	DSSIM
SfSNet [24]	0.0545 ± 0.0201	0.1330 ± 0.0531	0.3151 ± 0.0523
DPR [41]	0.0282 ± 0.0125	0.0702 ± 0.0361	0.1818 ± 0.0490
Nestmeyer [20]	0.0484 ± 0.0318	0.0584 ± 0.0335	0.2722 ± 0.0950
Our model	$\mathbf{0.0220 \pm 0.0078}$	$\mathbf{0.0292 \pm 0.0148}$	$\mathbf{0.1605 \pm 0.0487}$

Target Lighting Results (mean ± standard deviation)

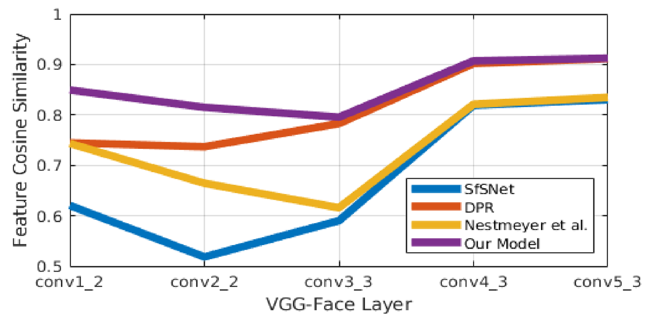
Multi-PIE Evaluation Using Lighting Transfer

Input Image | Reference Image | Target Image | Our Model | Shu *et al.* | Shih *et al.*

Method	Si-MSE	MSE	DSSIM
Shih [27]	0.0374 ± 0.0190	0.0455 ± 0.0203	0.2260 ± 0.0443
Shu [28]	0.0162 ± 0.0102	0.0243 ± 0.0170	0.1383 ± 0.0499
Our model	$\mathbf{0.0148 \pm 0.0096}$	$\mathbf{0.0204 \pm 0.0153}$	$\mathbf{0.1150 \pm 0.0404}$

Lighting Transfer Results (mean ± standard deviation)

Facial Detail Preservation Evaluation (VGG-Face)



- Cosine similarity between VGG-Face features of relit and groundtruth Multi-PIE images.
- Our method's cosine similarity is noticeably higher for earlier layers of VGG-Face, which indicates we better preserve local facial details.

Takeaways

- An effective way to handle hard shadows is to focus on the high-contrast shadow borders.
- Estimating the ratio image better preserves the facial details of the subject during relighting.