Improving Illumination Estimation using Shading Residue

Bhushan Sonawane
Stony Brook University
bsonawane@cs.stonybrook.edu

Abstract

SfSNet[1] estimates Albedo, Normal and Spherical Harmonics attributes for given face image. Due to lack of Shading loss, we see some of the residue pushed down into Albedo. We worked on improving albedo in SfSNet setting by computing residue difference and adding into shading (shading is computed using normal and spherical harmonics). We propose that learned shading difference will lead to removal of albedo residue and lead to updated correct shading. With the ground truth albedo, we can achieve this with shading loss similar to normal and albedo loss. But, in case of lack of ground truth, we propose to achieve the same with only re-construction loss and using generative adversarial networks[2] to remove residue from albedo- we are using Cycle-GAN[3] based training paradigm for albedo generation. We are using pretrained model provided by SfSNet. We also tried supervised approach for albedo with very less weight to albedo loss and experimented with shading residual net, such as using latent lighting representation, generating shading from normal and latent lighting representation.

1 Introduction

Learning face attributes such as albedo, normal and spherical harmonics is useful for face editing application. Different approaches have been tried out for the same. Deforming Autoencoders[4], SfSNet, LDAN[5] are some of the successful attempts. SfSNet have achieved highest results but faces issues with Albedo generation. Due to lack of shading loss, reconstruction loss pushes shading residual into albedo which leads to not so accurate albedo.

We started off with implementing SfSNet in PyTorch and Skip-Connection based network to generate CelebA data for pseudo-supervision. Then we built our experiments on top of the base model by adding network to generate shading residue and different approach to compute shading using latent lighting representation instead of Spherical Harmonics.

Later, we moved towards using pre-trained model provided of SfSNet and build experiments on it's top to remove residue from albedo and add into shading instead using reconstruction loss and optionally with lesser albedo loss. Pre-trained model does not generate albedo accurately and hence, we thought of using generative approach to generate albedo in synthetic albedo space which is known true ground truth. Here, we sample synthetic albedo as real albedo and generated albedo as a fake albedo to ensure accurate albedo generation. We experimented with both no albedo loss and albedo loss with less weight approach. For GAN training, we adopted CycleGAN.

2 Connection with work from previous semester

Last semester, we worked on Label Denoising Auto-Encoder[4] and focus was on domain adaptation. Problem we solved was to understand and learn using synthetic data where we lack ground truth. This work is important if our current work as well. Right now, we had performed experiments with multi-pie [5] and CelebA [6] dataset and have focused only on designing robust network to work with real images with ground truth. Due to lack of large variation in number of unique faces, we faced overfitting on the faces being generated from our system. Domain adaptation is promising approach to tackle this problem where we can use earlier method to generate more training instances and perform the training and then use domain adaptation to transfer the real images into same space to get desired lighting direction in input source image.

Now, we are combining our experience from last semester's use of GANs for domain transfer to generate CelebA albedo in synthetic domain space with less residue as one of the promising method.

3 Spherical Harmonics Lighting

Spherical harmonics(SH)[7] is well studied and commonly used rendering method for generating realistic lighting and shading. Lighting can be represented as a spherical function and used along with face normal to generate shading for the faces [8]. Figure 1 shows face normal and spherical lighting based lighting as discussed in SIRFS[7].

Face lighting transfer can be modeled with the use of face normal and spherical harmonics base lighting as we studied in Spring 2018, but use for face normal and lighting leads to use of many latent representations and is really difficult to model as un-supervised approach. We need supervision for face normal as well as spherical harmonics to ensure that we are indeed learning the representation in terms of face normals and SHs.

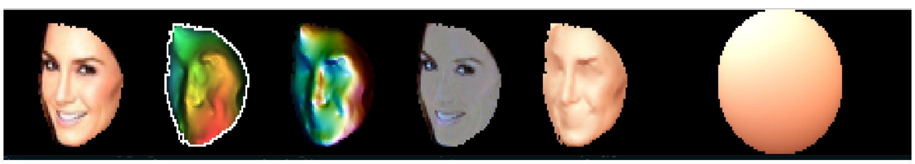

Figure 1: Input image, Shape, Normal, Shading, Reflectance and SH based lighting respectively

4 Baseline

We implemented base SfSNet and Skip connection based model for baseline and ground up for our work. Figure 2 demonstrates SfSNet pipeline and figure 3 demonstrates Skip connection based model.

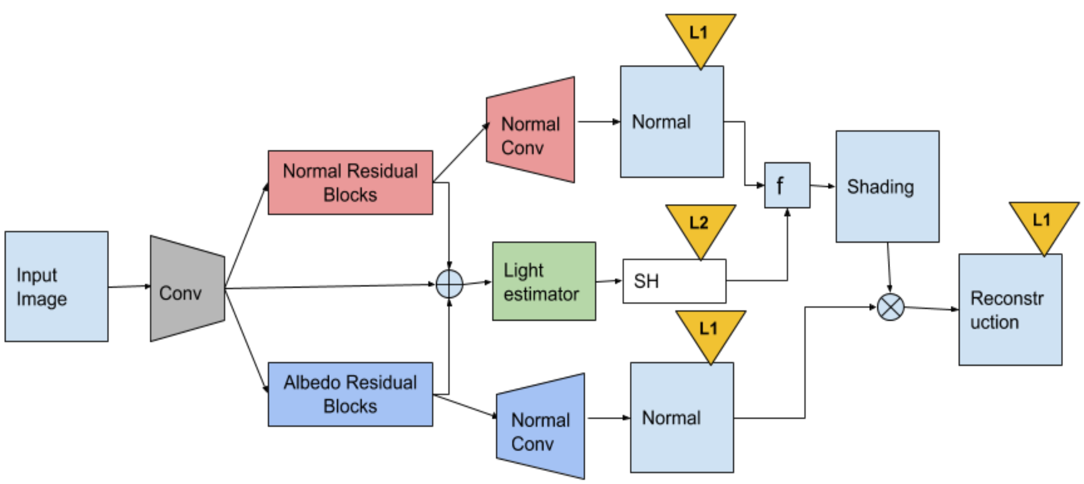

Figure 2: SfSNet Model

Training procedure is as follow:

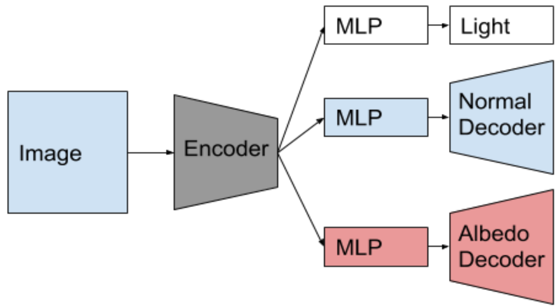

Figure 3: SkipNet

1. Train Skip-Connection based model on synthetic data
2. Generate Albedo, Normal and SH for CelebA data using trained skip-connection based model
3. Use Synthetic data with true ground truth and generated CelebA data as pseudo-supervision
4. Train SfSNet using dataset described in step 3

4.1 Skip-Connection based dataset

Skip connection network is same as used in SfSNet and Neural face editing[6] We trained Skip-Connection based network for 20 epochs with learning rate of 0.0002 with Adam optimizer and no weight decay.

CelebA generated data is not very good and does have some artifacts. Few experiments with this dataset leads to inaccurate albedo generation and hence, we initialized sfsnet with pretrained weights, fixed all the weights except albedo residual, generation and new residue generation network. Now, task is to learn albedo and shading residue. Now, problem is with pretrained albedo which does have residue and we need to un-learn it. High level idea is to learn residue and remove residue from albedo generation(i.e. albedo residual block and albedo generation net).

5 Experiments

High level idea is to generate shading residue and then add the residue into shading, then we use this updated shading for reconstruction. We hope to capture residue and successfully eliminate residue from albedo by training only for albedo and residue network.

Following are few approaches we experimented with for residue learning.

1. **Shading Correcting Network**- In this approach, we update shading using latent lighting representation. We hypothesize that latent lighting representation (spherical harmonics 2.0) will captures details missed by spherical harmonics and fill in the residue for shading. Here, Shading correcting network is responsible for combine existing shading and residue.
2. **Latent Shading Generation**- In this approach, we generate shading without spherical harmonics and only use normal and latent lighting representation. Here, Shading generation network is responsible for generating correct shading and here we do not rely on residue as neural network is expected to learn the representation accurately.
3. **Shading Residue Network**- In this approach, we simply determine the residue and add into predicted shading. This simplifies overall learning and residue learning task.
4. **Two way Residue Network**- In this approach, we determine the residue similar to approach 3, but we not only add the residue in shading but also subtract the residue from albedo. Here, we have more strong supervision on residue. We can use analogy of take from albedo and give it to shading. Here, we fix albedo network as well and add supervision on updated albedo which is 'albedo - residue'.

5. **GAN based approach for Albedo generation for shading residue Network**- In all of the above approaches, we faced issue for generating ground truth albedo (which we only have for synthetic data). Hence, we use GANs to generate albedo in synthetic domain space. We additionally added supervision with very small weight to albedo with Smooth loss.

Below, we will first go over GAN based approach (i.e. approach 5) and then other approaches.

5.1 GAN based approach for Albedo generation for shading residual network

In this approach, we are using pre-trained model and hence fixing weights for normal and spherical harmonics estimation. We make a copy of albedo residual block and use albedo conv as generator in GAN training setting.

Figure 4 demonstrates GAN based model architecture. We keep rest of the flow similar to SfSNet model due to pretrained weights being used to ensure minimal changes and assess learning of residue and albedo.

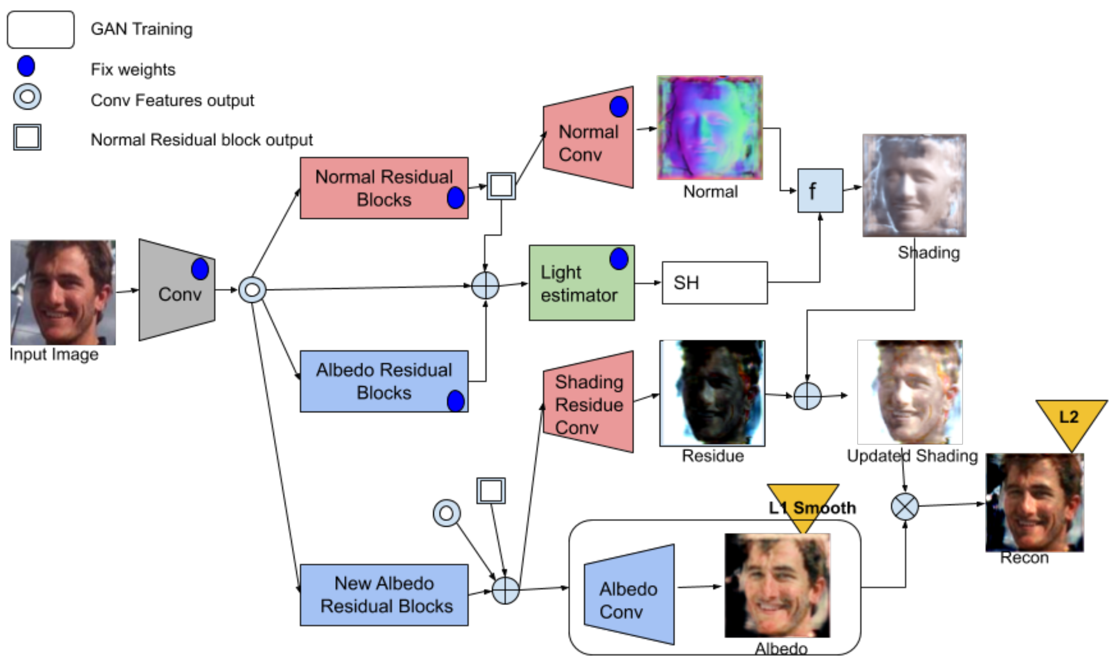

Figure 4: GAN based model

We use CycleGAN training paradigm for training albedo generation and traing Generator along with full network training. We experimented with and without L1 smoothing loss and observed that L1 smoothing loss with very small weight helps learning albedo generation. With only reconstruction loss, all the learning is pushed towards shading residue network and albedo generation is not learned properly. Note that, we are using small weights for gan loss and albedo loss (0.01 and 0.2 respectively).

We experimented with very small weight for gan loss i.e. 0.002, but that did not help and lead to incorrect albedo generation. Figure 11 and 12 shows interpolation and comparison for celeba and synthetic image.

Plot 5 shows training loss for gan based model. Plot 6 shows reconstruction and albedo loss for gan based model.

5.2 Shading Residue Network

Our claim is that, some of the shading residue is pushed down into albedo. Instead of learning normal, spherical harmonics from scratch, we propose to build on top current model by adding residue network.

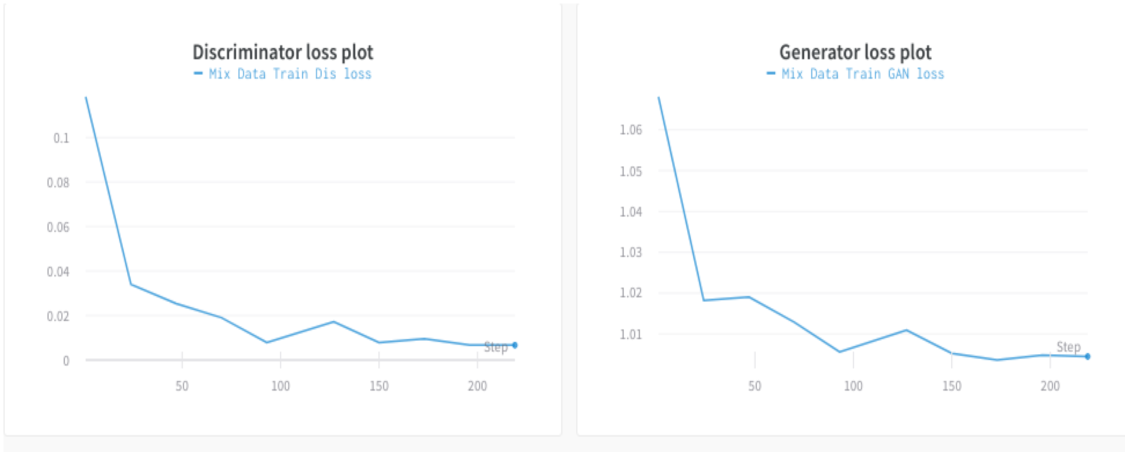

Figure 5: GAN model: GAN loss

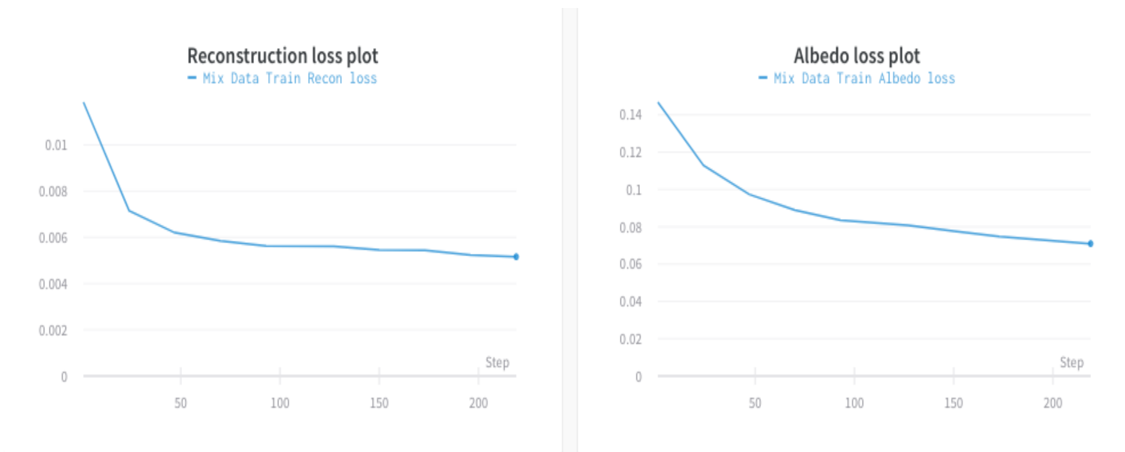

Figure 6: GAN Model: Reconstruction and albedo loss plot

Idea is to find the residue pushed down into albedo and add into shading. Network with this setting will be trained primarily with reconstruction loss. Use of shading loss will be more strong supervision and will be useful for finding better residue, but due to lack of ground truth normal, spherical harmonics, we cannot use the shading loss. Hence, we use lower weights for albedo and focus more on reconstruction loss and also don't comprise on albedo much.

Figure 7 explains architecture used for this experiment. First, we compute residue and then simply add it into generated shading. Note that, we are not learning normal and spherical harmonics and hence, we ensure to only update albedo and shading using residue network and reconstruction loss.

5.3 Shading - Albedo Residue Network or Two-Way Residue Network

In Shading residue network, we added computed residue in generated shading and relied on network learning to remove residue from albedo. In this approach, we subtract residue from generated albedo and add same residue into shading.

In this architecture, we don't need to create new albedo generation network as we will not generate albedo from scratch but only update the existing one i.e. fixing and not generating. We can imagine residue network as Robinhood who takes from albedo and gives to shading.

Figure 8 demonstrates architecture being used.

Primary issue in this approach is again ground truth albedo for CelebA dataset. Hence, we again use smaller weight for albedo loss to focus on reconstruction loss.

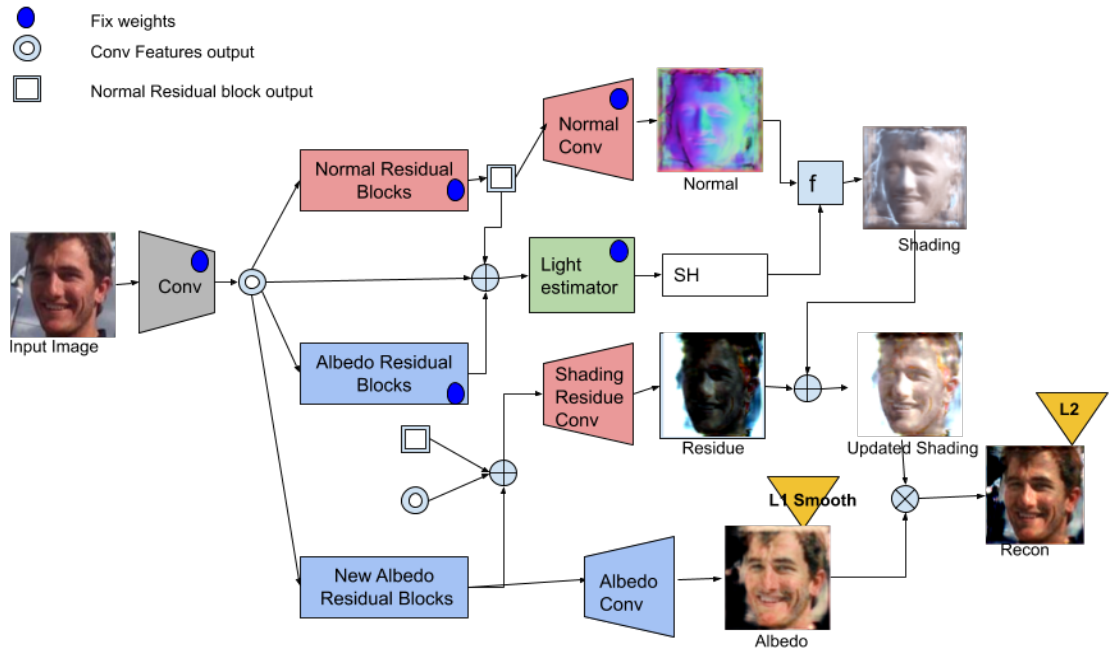

Figure 7: Shading Residual model

5.4 Shading Correcting Network

This was very first experimented we tried. In this approach, we use latent lighting representation, let's say Spherical Harmnics 2.0 to update or correct the generated shading.

We generate SH 2.0 using similar input as SH i.e. using conv features, normal residual block features and albedo residual block features. Later, we use this generated shading along with SH 2.0 to correct the shading. Shading correcting network is suppose to use SH 2.0 and add missing details in traditional shading.

Here, We believe that, SH 2.0 is nothing but add-on for spherical harmonics but do not have any method or representation to check what SH 2.0 is learning or representing.

Figure 9 shows the shading correcting model.

5.5 Shading Generation using SH 2.0

In Shading Correcting approach, we used two lighting representations. In this experiment we assessed need of spherical harmonics and used SH 2.0 along with normal to generate the shading.

Lighting representation is equivalent to image dimensions for each plane. We concatenate lighting representation along with normal and feed into shading generation network to generate shading using these two representation. We are fixing normal and hence, ensuring that latent lighting representation is being valid and representing lighting similar to spherical harmonics. Note that, we are not aware of function it is approximating to and expect it to be some what but not exactly similar to spherical harmonics.

Figure 10 explains the shading correcting model.

6 Results

We performed extensive experiments with above methods with different weights to albedo loss, adding shading loss and also removing albedo loss.

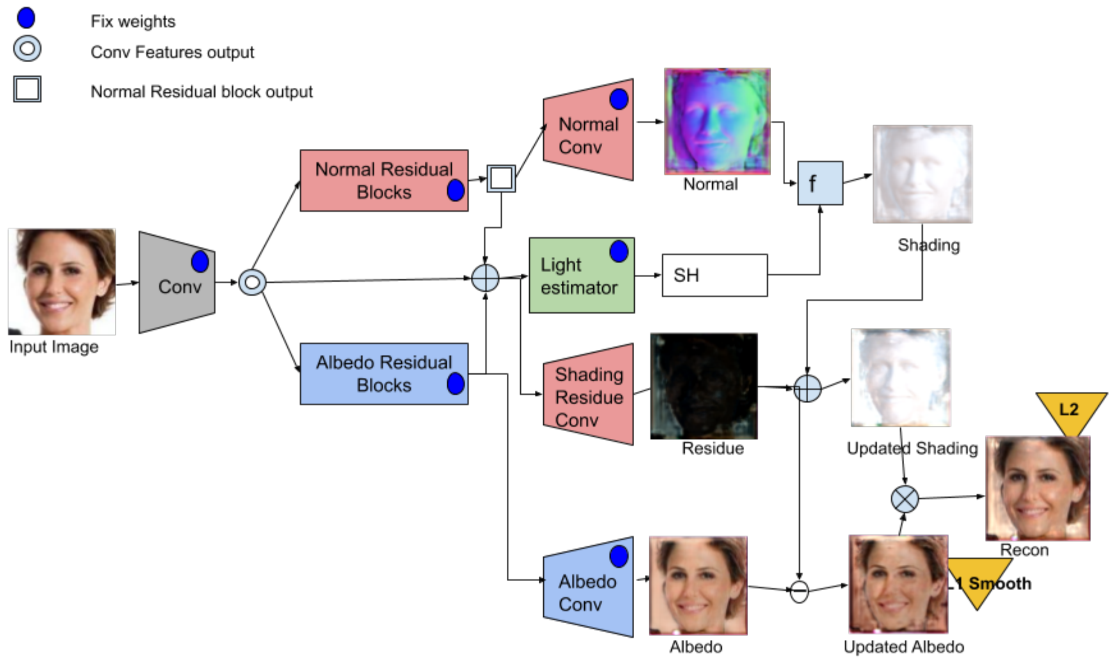

Figure 8: Shading-Albedo Residual model

Albedo loss: No including albedo loss leads to albedo gen network collapse and learning is pushed down into residual networks. Using albedo loss with less weight was viable choice as we indeed want to use albedo loss but give more weight on reconstruction instead of albedo for which ground truth albedo has residue in it.

Shading loss: Addition of shading loss is important as it helps pushing the residue into shading network with valid shading ground truth. Due to lack of which, we added shading loss optionally with small weight. But, no significant benefit was seen with small shading loss due to incorrect ground truth available.

Following are the results of each method.

6.1 GAN based albedo generation in synthetic domain space

Figure 21 shows predicted albedo, shading, residue, updated shading and reconstruction with this method.

Figure 11 and 12 shows results on GAN based method on unseen samples on celeba and synthetic images.

6.2 Shading residue network

Figure 13 and 14 shows output of residual based network. We have added more examples of sample output in appendix.

We can see that residue indeed captures the missing shading and shading is updated partially. But albedo was not completely accurate and some artifact was observed in it.

6.3 Shading-Albedo Residue network

Figure 15 and 16 shows output of residual based network. We have added more examples of sample output in appendix.

This approach did not perform better than only shading residue. We believe this is due to shading again being pushed into albedo due to lack of supervision on albedo.

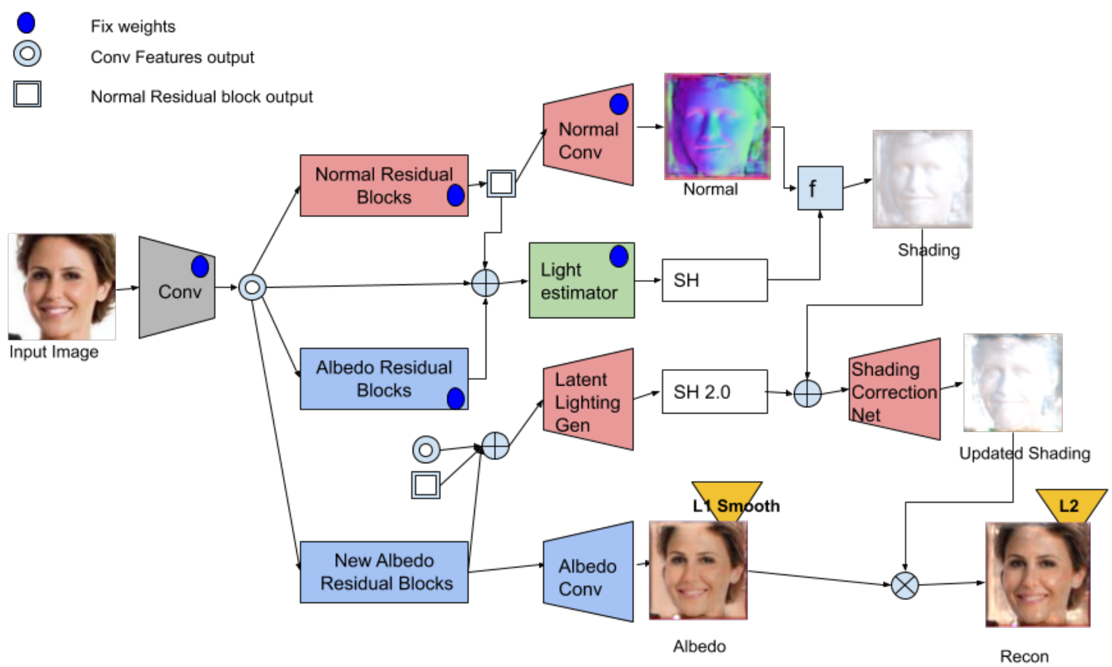

Figure 9: Shading Correcting model

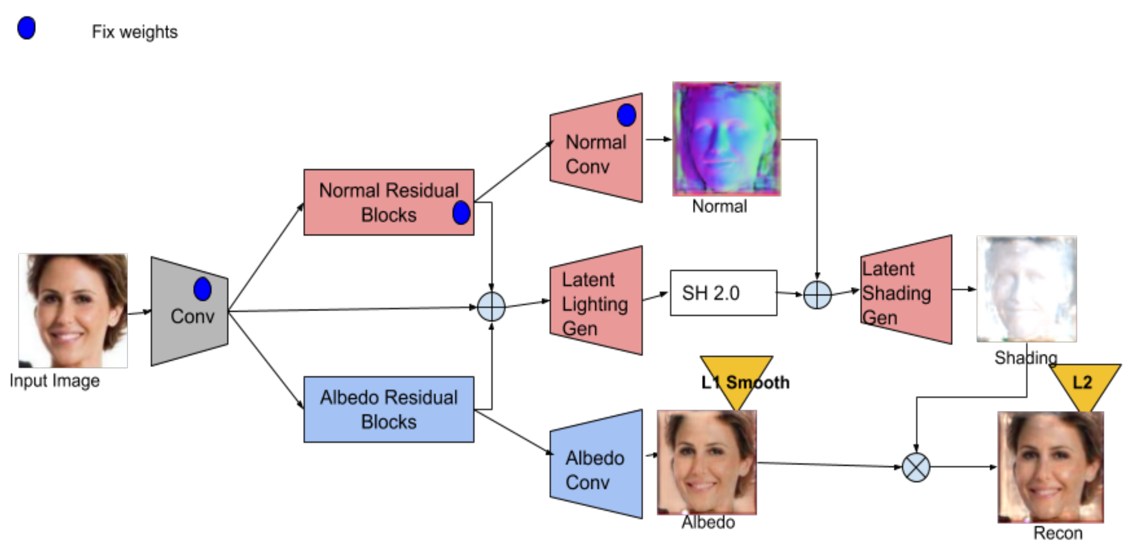

Figure 10: Shading generation using SH 2.0

6.4 Shading Correction Network

Figure 17 and 18 shows output of shading correcting network on celeba and synthetic data respectively. We have added more examples of sample output in appendix.

This approach has many downsides-

1. lots of parameters
2. Albedo is not albedo exactly
3. Due to shading correcting network, image features are pushed into updated shading

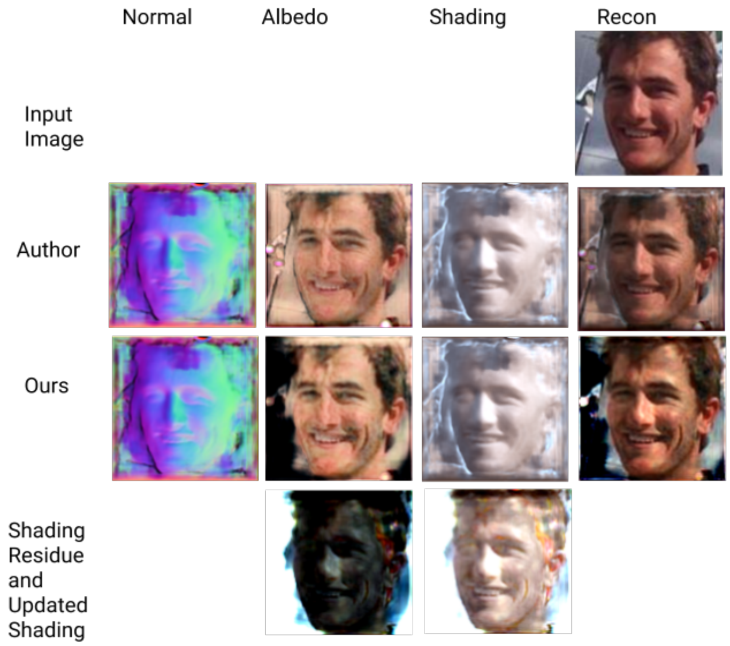

Figure 11: GAN based model result for CelebA

6.5 Shading Generation using SH 2.0

Figure 19 and 20 shows output of shading correcting network on celeba and synthetic data respectively. We have added more examples of sample output in appendix.

This approach is simplified version of shading correcting network. This generates shading directly using latent lighting and normal. But faces similar issue of image features being pushed into shading.

7 Conclusion

Latent ligthing representations were not useful due to lack of shading loss and more importantly lack of ground truth for pseudo-supervision. Shading correcting and latent shading generation approach are not promising as of now.

GAN based approach produced convincing results with use of low weights to albedo and gan loss. This approch shows that we can work with noisy and spurious data with the use of synthetic dataset and shifting generated output in synthetic domain space.

Shading Residue network is also promising approach to pursue further. Unlike GAN based approach, it is solely depended on supervision. In this setting, we relied on reconstruction error and rely on noise albedo loss in small proportion. We can investigate this furhter once we generate or get pseudo supervision data without albedo residue.

Shading-Albedo residue network was suppose to perform better due to more manual residue addition and subtraction and hence should be looked into more details.

8 Next Steps

1. GAN based approach: Try different or more stable GAN setting e.g. Wassertian-GAN gradient penalty[8]
2. Fine-tune skip-connection network to generate stable and artifact free pseudo-supervision dataset
3. Once, accurate pseudo-supervision dataset is available

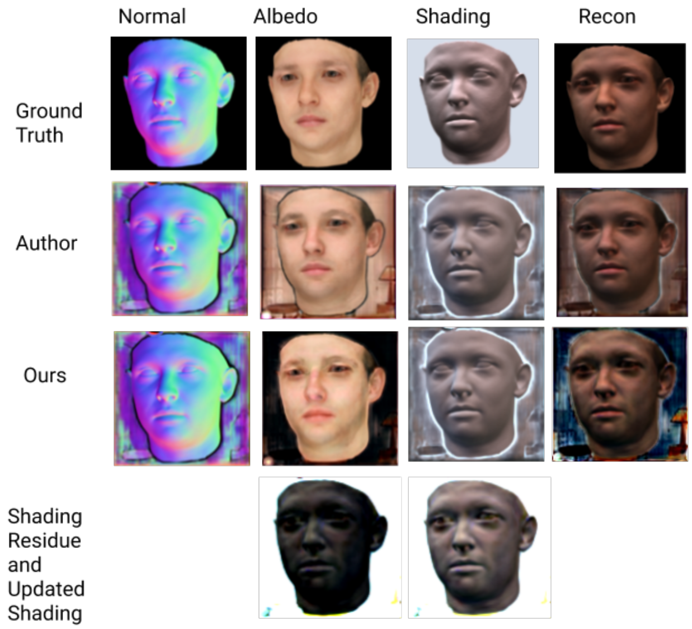

Figure 12: GAN based model result for Synthetic dataset

(a) Adding shading loss in base sfsnet
(b) Adding shading loss in residue networks
(c) Add L1 smoothing loss with more weight

Reference

1. Soumyadip Sengupta, Angjoo Kanazawa, Carlos D. Castillo, David Jacobs. SfSNet: Learning Shape, Reflectance and Illuminance of Faces in the Wild Soumyadip Sengupta, Angjoo Kanazawa, Carlos D. Castillo, David Jacobs
2. Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza. Generative Adversarial Networks
3. Jun-Yan Zhu, Taesung Park, Phillip Isola, Alexei A. Efros. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks
4. Zhixin Shu, Mihir Sahasrabudhe, Alp Guler, Dimitris Samaras, Nikos Paragios, Iasonas Kokkinos. Deforming Autoencoders: Unsupervised Disentangling of Shape and Appearance
5. Hao Zhau, Jin Sun. Label Denoising Adversarial Network (LDAN) for Inverse Lighting of Face Images
6. Z. Shu, E. Yumer, S. Hadap, K. Sunkavalli, E. Shechtman, and D. Samaras. Neural face editing with intrinsic image disentangling
7. Jonathan T. Barron, Jitendra Malik , SIRFS
8. Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky , Vincent Dumoulin , Aaron Courville. Improved Training of Wasserstein GANs.

9 Appendix

9.1 More examples of sample output

Following, we go over more sample output of each of the model described above. Figures 22, 23, 24, 25, 21 shows sample output for shading correcting, latent shading generation, shading residue, shading-albedo residue and gan based approach respectively.

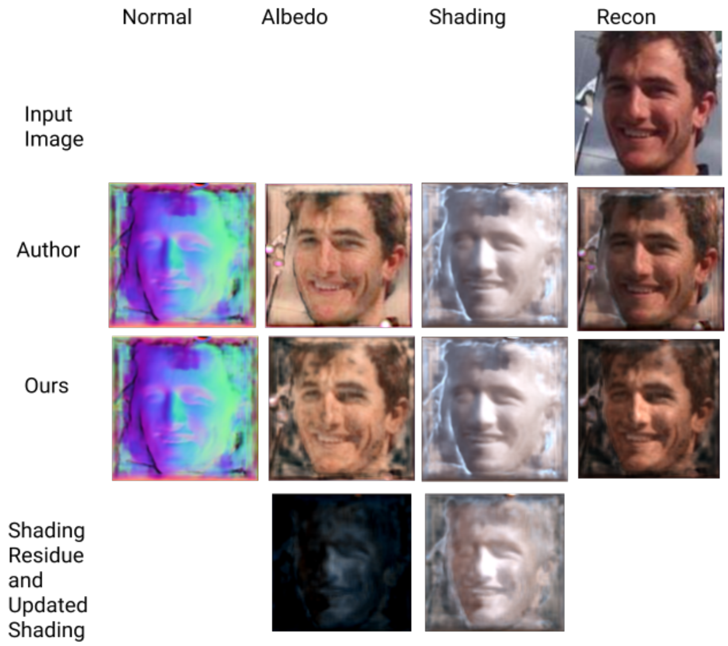

Figure 13: Shading Residual based model result for CelebA

9.2 Source code

1. Baseline SFSNet and SkipNet https://github.com/bhushan23/SfSNet-PyTorch
2. All experiments- https://github.com/bhushan23/SC-Net

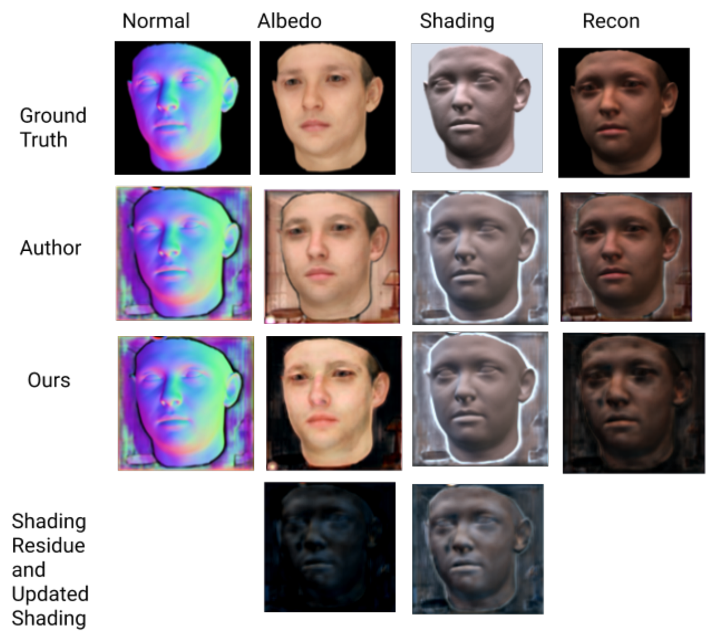

Figure 14: Shading Residual based model result for Synthetic dataset

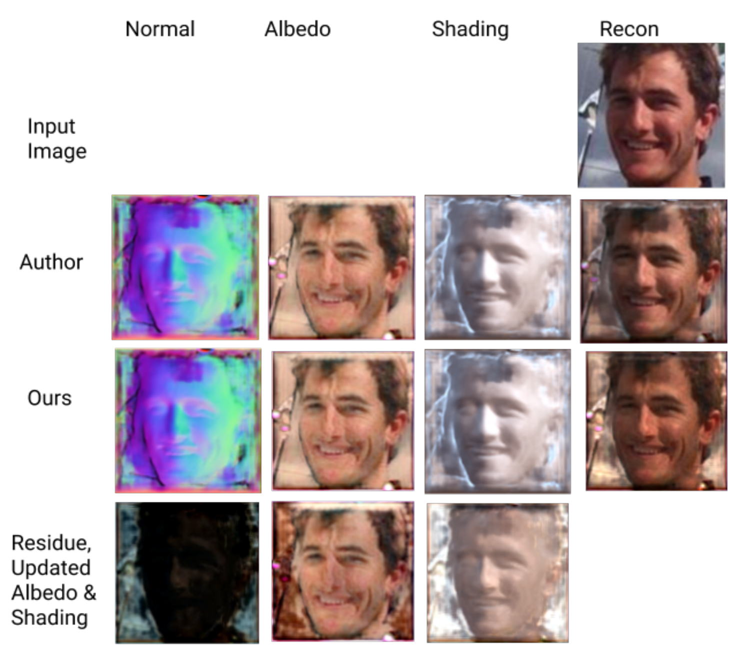

Figure 15: Shading-Albedo Residual based model result for CelebA

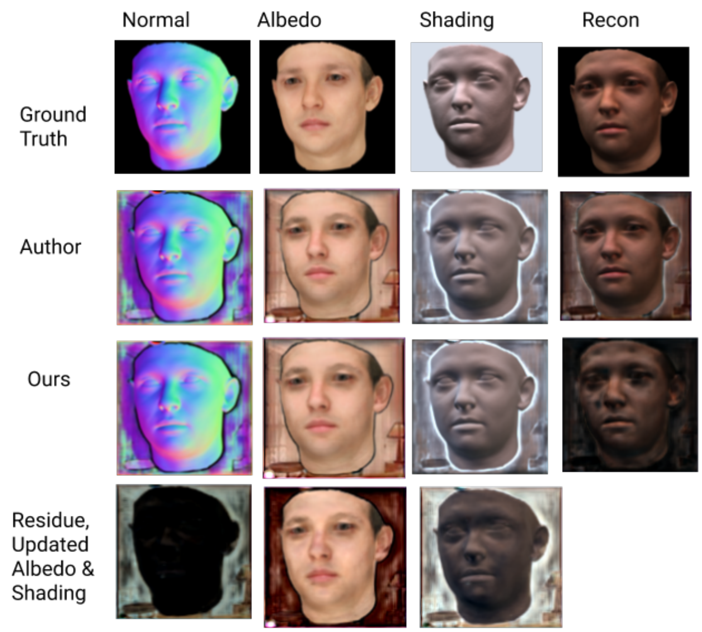

Figure 16: Shading-Albedo Residual based model result for Synthetic dataset

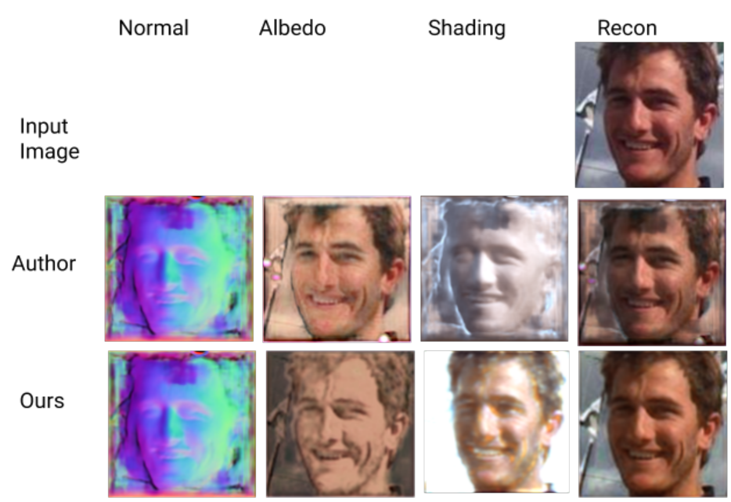

Figure 17: Shading correcting model result for CelebA

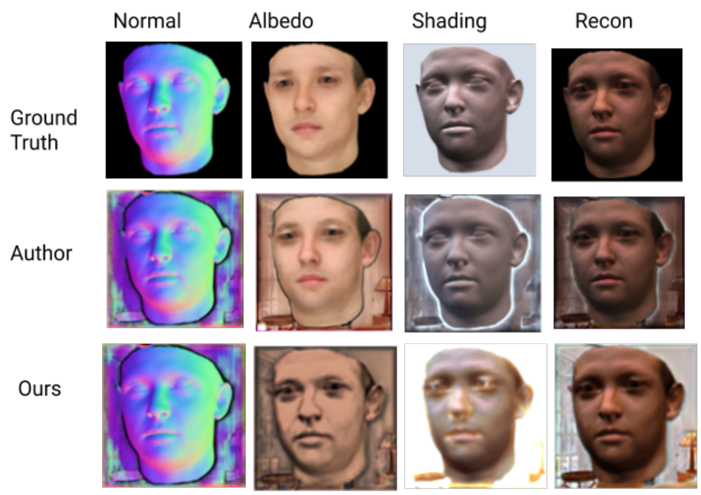

Figure 18: Shading correcting model result for Synthetic dataset

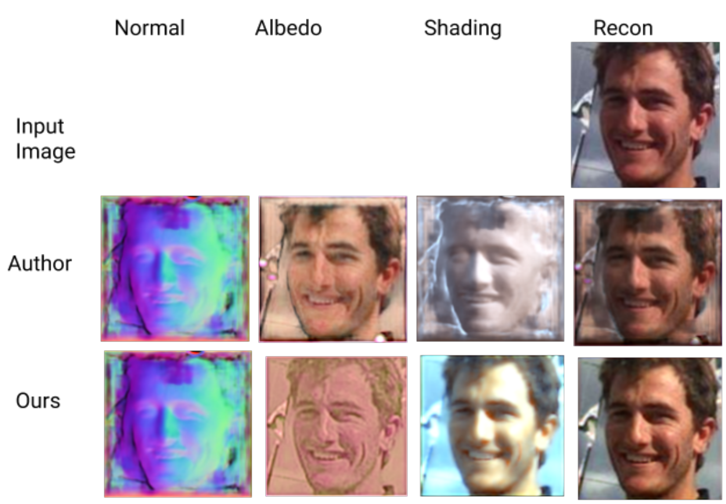

Figure 19: Shading generation using SH 2.0 result for CelebA

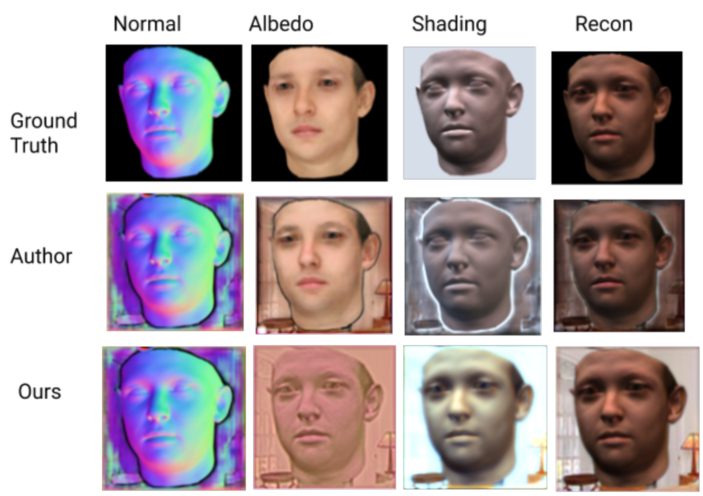

Figure 20: Shading generation using SH 2.0 result for Synthetic dataset

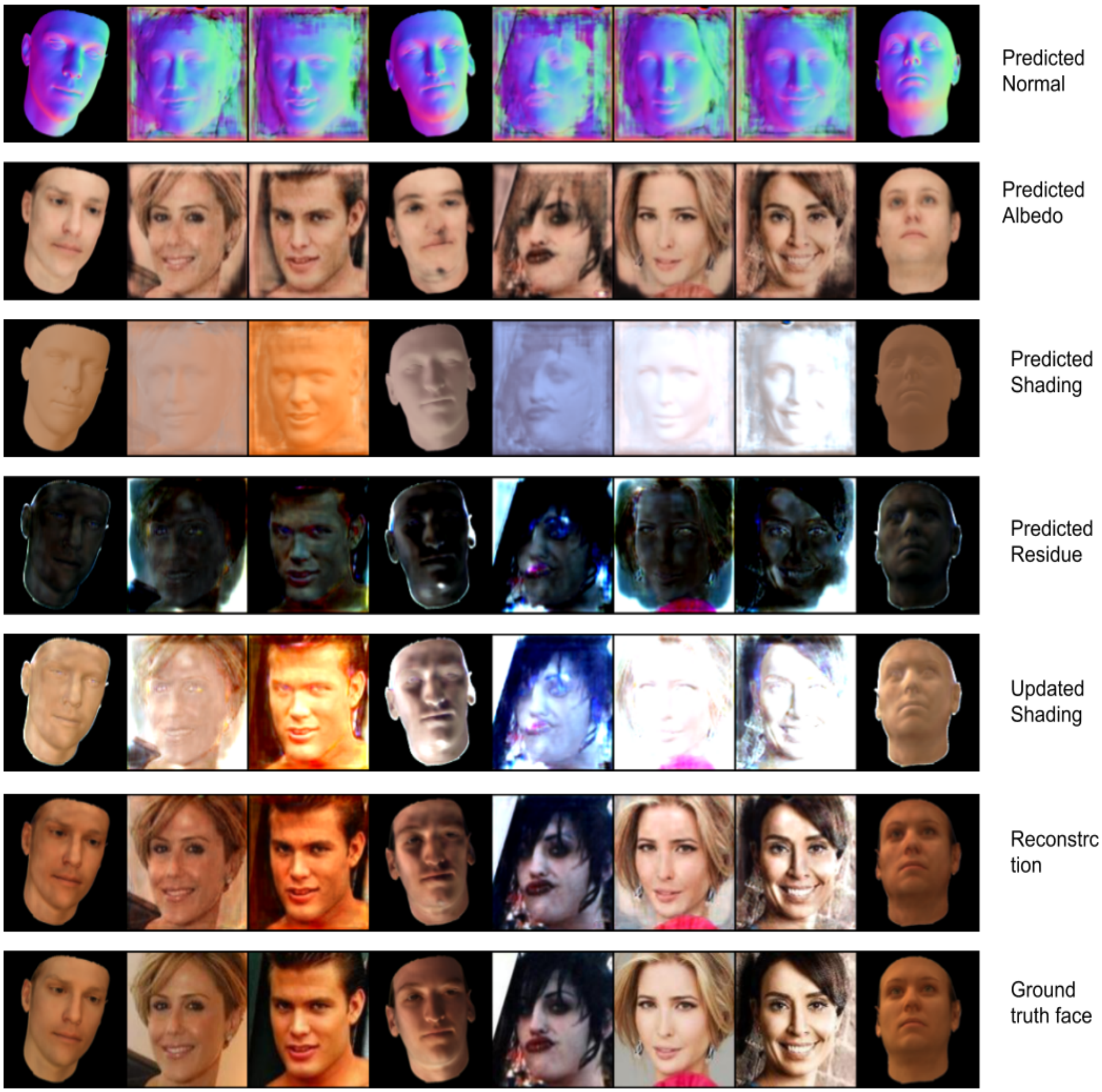

Figure 21: GAN based model Result samples

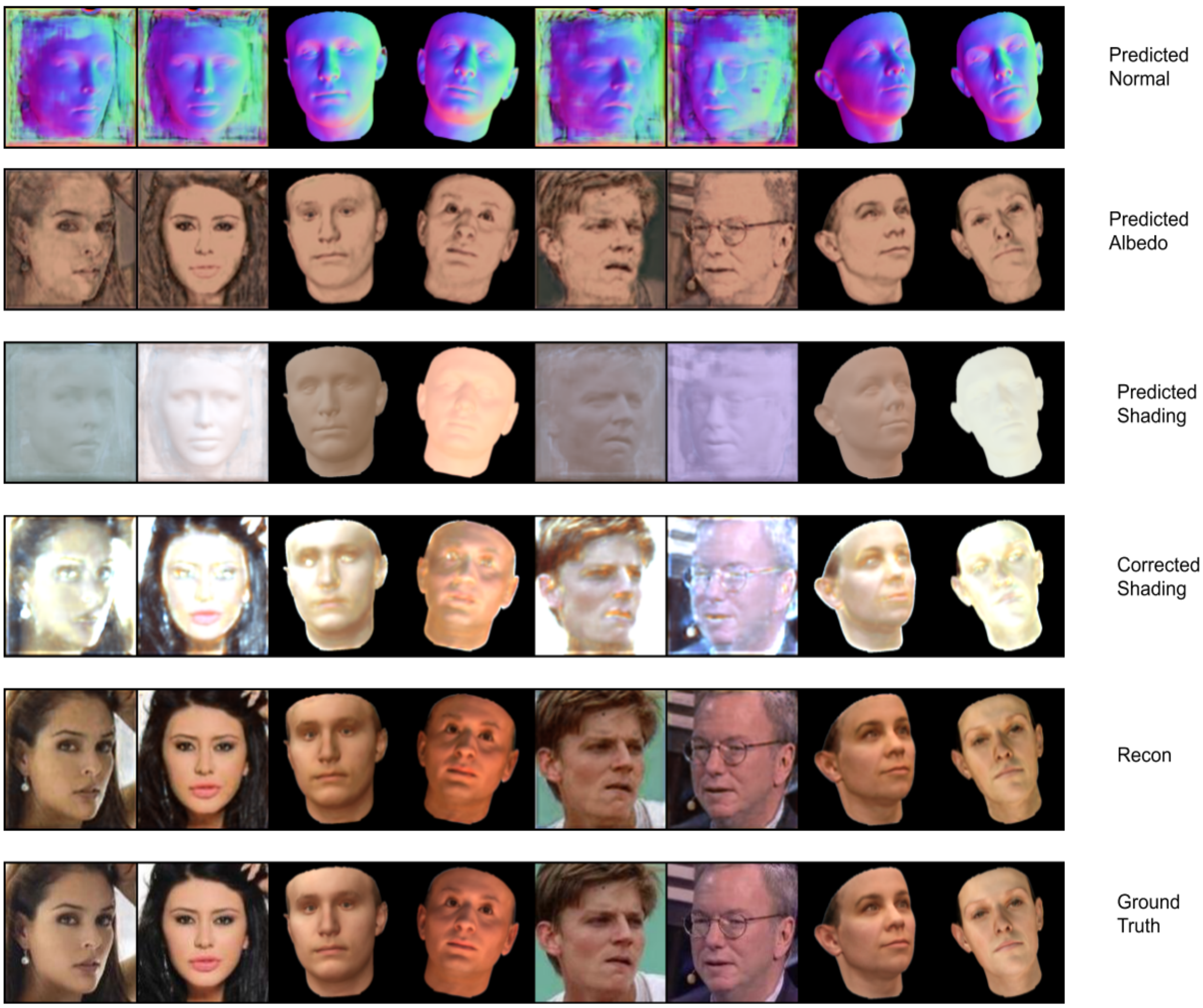

Figure 22: Shading Correcting model Result samples

Figure 23: Latent shading generation model Result samples

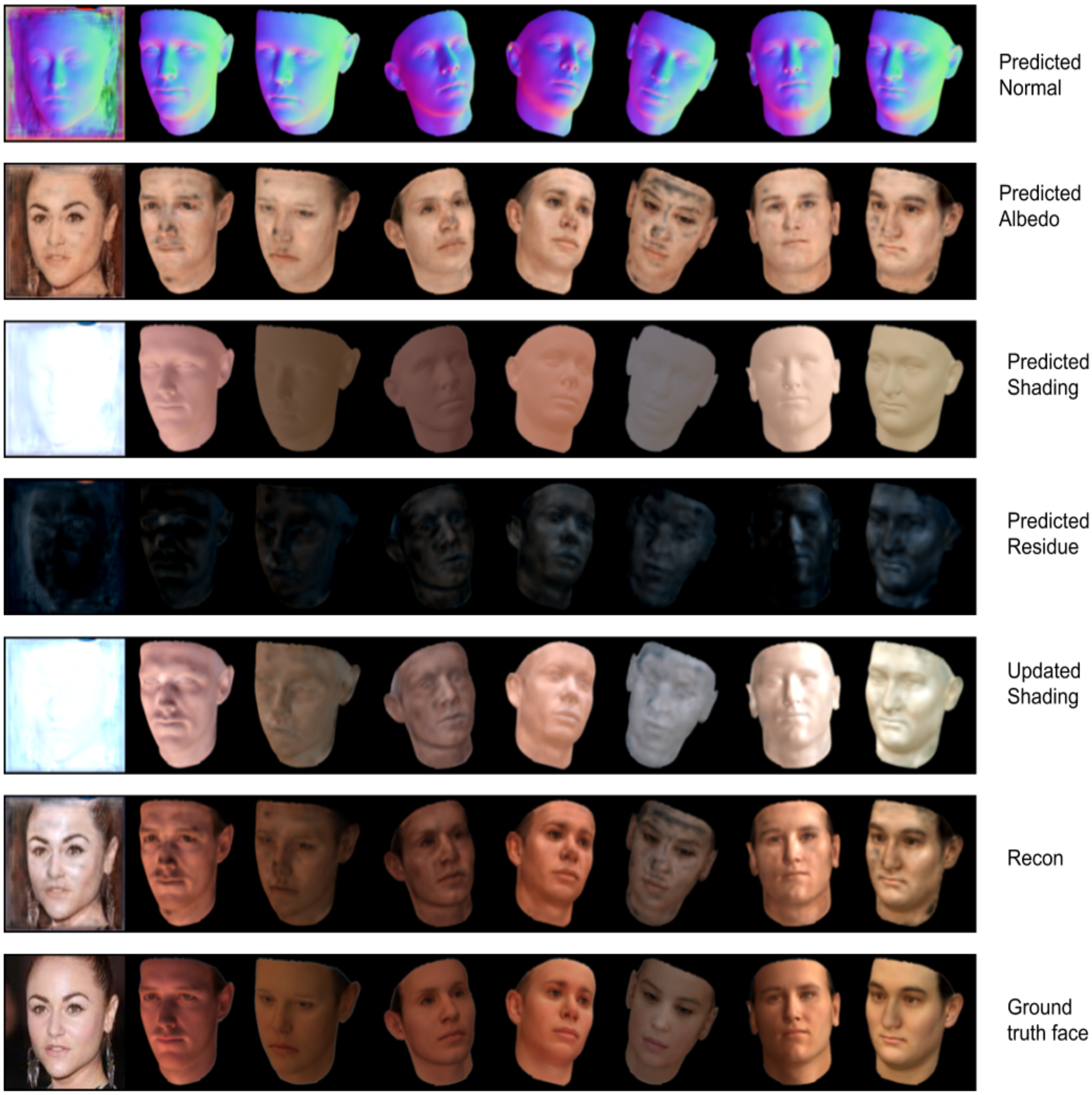

Figure 24: Shading Residue model Result samples

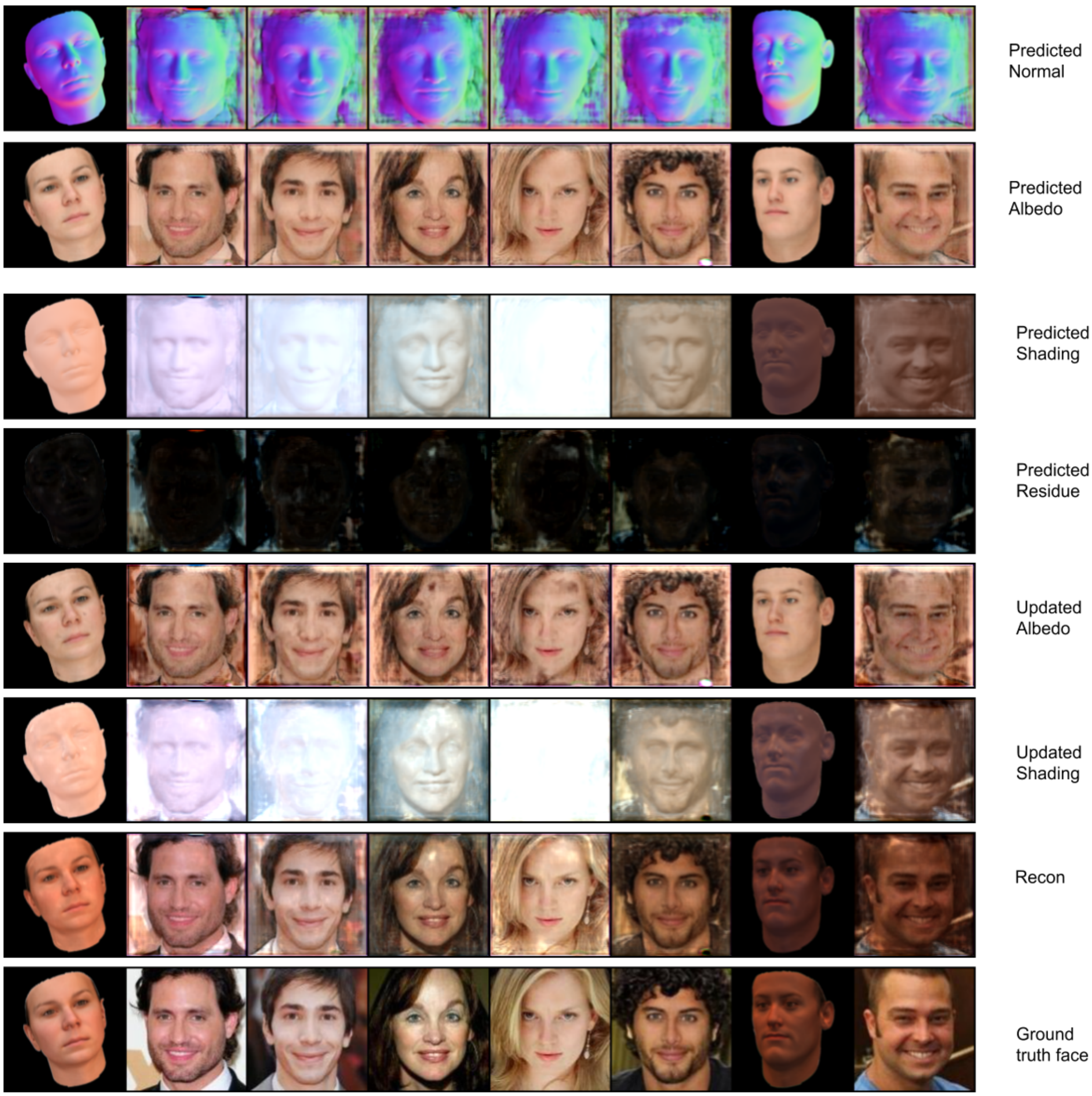

Figure 25: Shading-Albedo Residue model Result samples