Relightify: Relightable 3D Faces from a Single Image via Diffusion Models (Supplementary Material)

Foivos Paraperas Papantoniou[1,2] Alexandros Lattas[1,2] Stylianos Moschoglou[1,2] Stefanos Zafeiriou[1,2]
[1]Imperial College London [2]Huawei Noah's Ark Lab
{f.paraperas,a.lattas,s.moschoglou,s.zafeiriou}@imperial.ac.uk

1. Additional Qualitative Results

As discussed in the main paper, our method relies on a pre-trained diffusion model which serves as prior for guiding the inpainting procedure. To illustrate the generation capabilities of this model, we show an extension of Fig. 3 of the main paper in Fig. 1, with a few more unconditionally generated (*i.e.* based on the standard diffusion sampling process) samples. Besides the synthesized UV maps, we further apply the reflectance components on top of random 3D shapes drawn from the employed 3DMM and provide corresponding 3D renderings in different scenes. Finally, in Fig. 2, we show additional 3D reconstructions and renderings from monocular images with our proposed method.

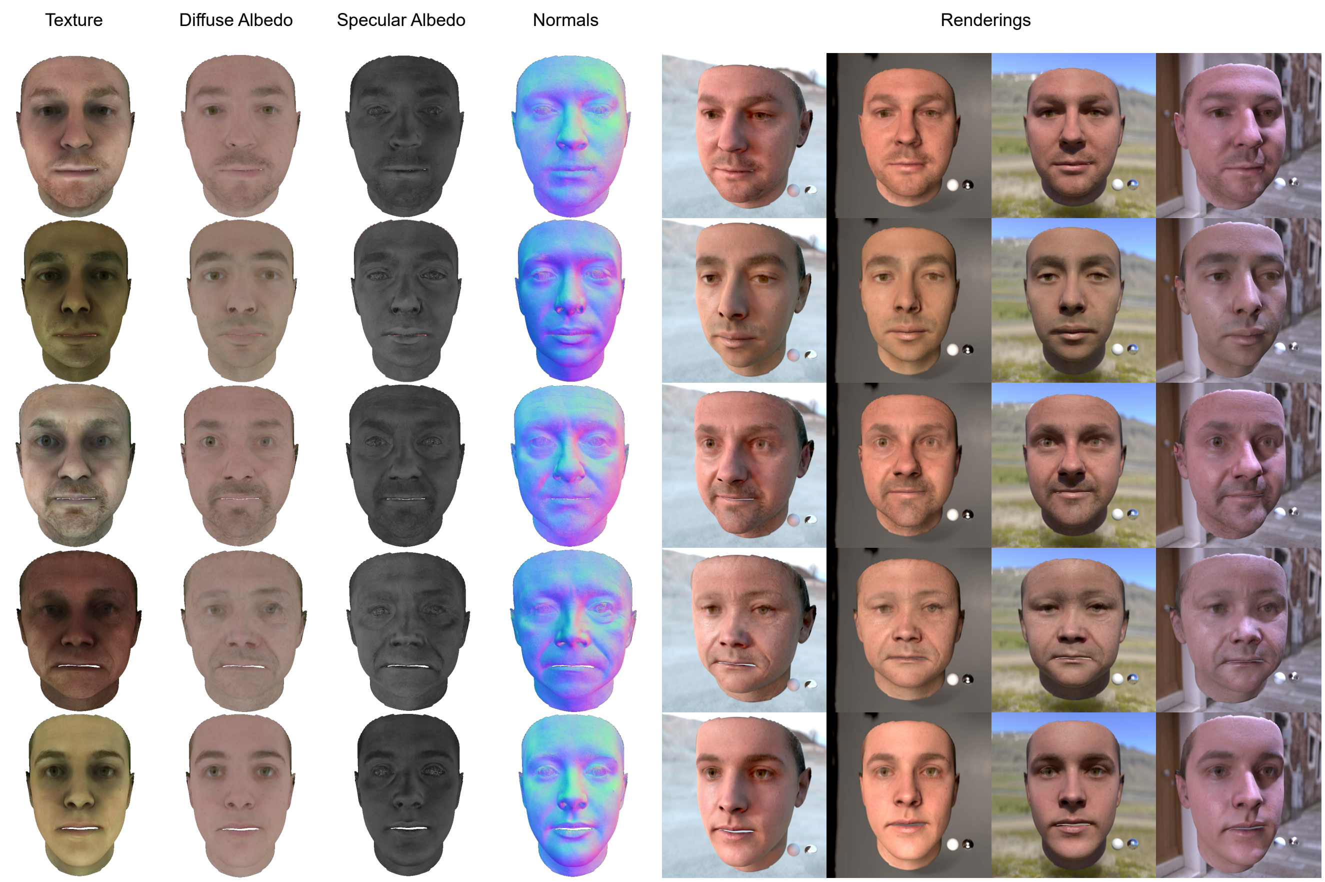

Figure 1. Texture and reflectance maps sampled from our diffusion model. Left: We visualize the UV maps on top of arbitrary 3D shapes from a 3DMM [1]. Right: Realistic renderings of the sampled reflectances on a commercial renderer.

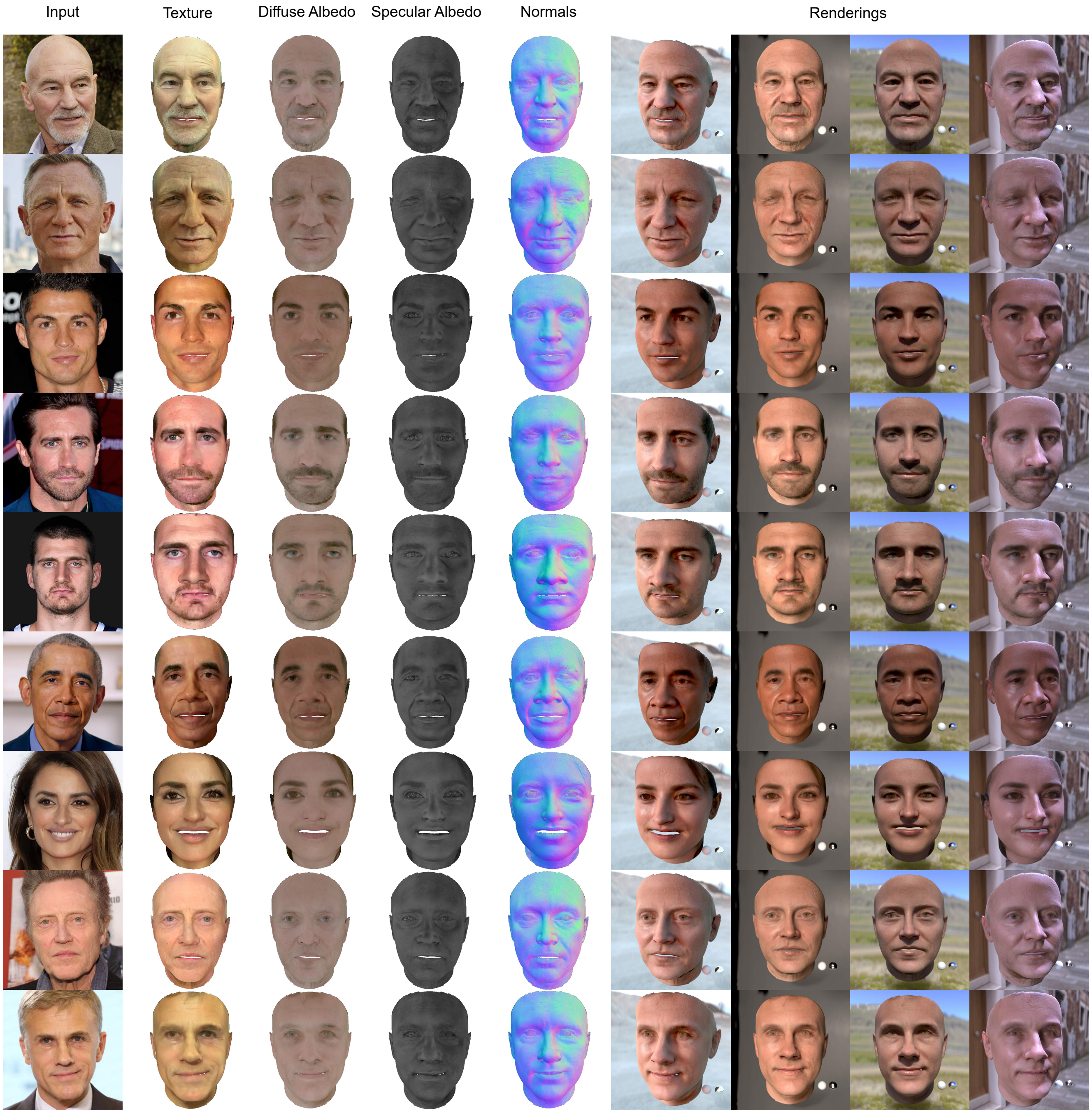

Figure 2. 3D reconstructions with recovered texture/reflectance maps by our method and renderings in different environments.

2. Visual Comparison with Reflectance Reconstruction Methods

In Fig. 5 of the main paper, we have shown examples of rendered reconstructions by our method, as well as two related methods, namely AlbedoMM [3] and AvatarMe++ [2], both of which recover the facial reflectance from singe images, similarly to our method. Here, we extend this comparison by adding the diffuse and specular albedos produced by these methods in Fig. 3. Note that although AvatarMe++ [2] generates UV assets of higher resolution (4K by 6K), our method shows better detail preservation properties by directly capitalizing on the observed texture. Also, for the case of AlbedoMM [3], which uses per-vertex albedo values, we transform them into a flat plane UV parameterization for visualization purposes.

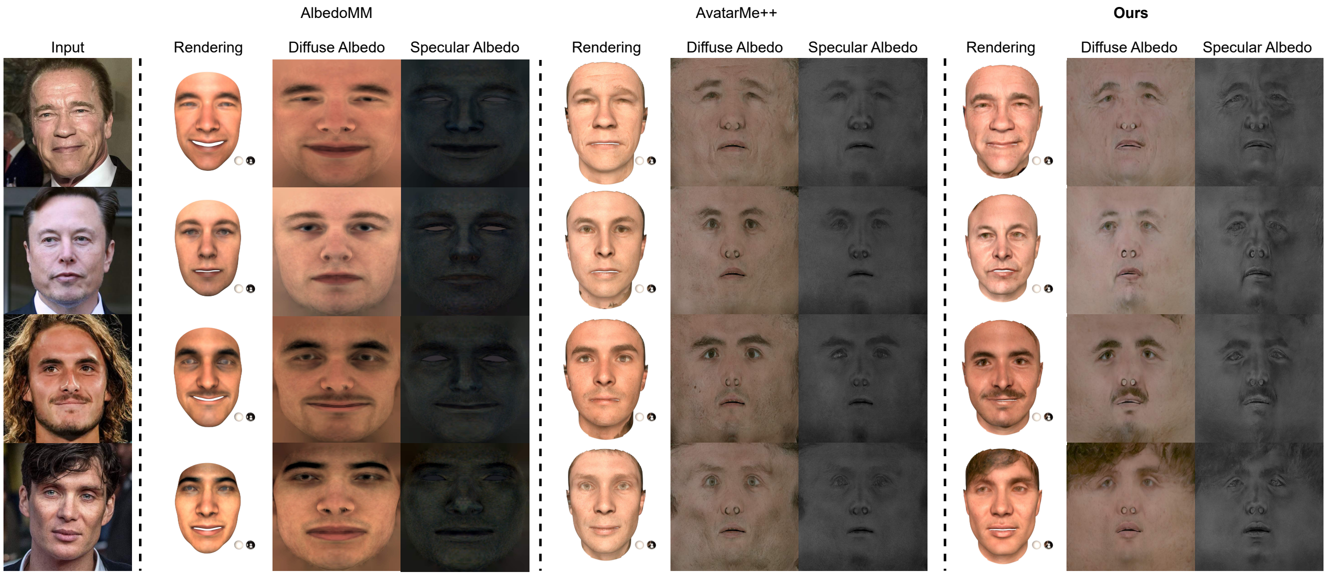

Figure 3. Comparison with AvatarMe++ [2] and AlbedoMM [3]. For a fair comparison, we sample the per-vertex albedo values of AlbedoMM to a flat plane UV parameterization, and crop AvatarMe++ and Our results to the central facial area, since AlbedoMM is using a tigher facial crop.

3. Ablation Study on Texture Augmentation

As mentioned in the paper, in order to enhance our model's generalization ability, we create arbitrary textures during training by re-rendering their delighted versions using random illumination environments. Preliminary experiments showed that this approach improves performance compared to using the original MimicMe textures. Here, we provide an evaluation of the model trained on the original textures, following the setting of Tab. 2 of the paper.

	PSNR (diffuse albedo)	PSNR (specular albedo)	PSNR (normals)
Ours (w/o aug.)	22.07	26.98	26.63
Ours (w/ aug.)	**22.47**	**27.17**	**26.69**

Table 1. Ablation study for UV texture augmentation (as in Tab. 2).

References

[1] James Booth, Anastasios Roussos, Stefanos Zafeiriou, Allan Ponniah, and David Dunaway. A 3d morphable model learnt from 10,000 faces. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[2] Alexandros Lattas, Stylianos Moschoglou, Stylianos Ploumpis, Baris Gecer, Abhijeet Ghosh, and Stefanos P Zafeiriou. Avatarme++: Facial shape and brdf inference with photorealistic rendering-aware gans. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.

[3] William AP Smith, Alassane Seck, Hannah Dee, Bernard Tiddeman, Joshua B Tenenbaum, and Bernhard Egger. A morphable face albedo model. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5011–5020, 2020.