# Supplemental Material for High-Fidelity Facial Reflectance and Geometry Inference From an Unconstrained Image

## 1 ADDITIONAL RESULTS

We provide additional results with pictures from the Chicago Face Dataset [Ma et al. 2015] to further demonstrate the fidelity of our reflectance and geometry inference for a wide range of ethnicities. In Fig. 3 on the left side, we visualize the input (upper left), inferred diffuse albedo (lower left), specular albedo (lower right), and meso-scale geometry (upper right). On the right, we provide a 3D rendering under a novel lighting environment (the environment from www.debevec.org is shown in the inset of the first subject). As seen in the figure, our method captures person-specific skin tones, dynamic wrinkles (e.g., first column, third subject), and distinguishing features such as stubble hair (e.g., first column, sixth subject), while illumination-dependent shading in the input image is successfully removed in the resulting reflectance maps, making it possible to insert the digital faces into arbitrary virtual environments.

*Rendering.* Our final rendering is produced with a layered surface and subsurface skin material, as shown in Fig. 1, using brute-force path tracing and the image-based lighting in Solid Angle's Arnold renderer [Solid Angle 2018]. The output displacement is applied on the base surface mesh to simulate the mesoscopic geometric deviations through geometry tessellation. The resulting specular albedo and diffuse albedo are respectively used to influence the intensities of the specular BRDF and deep scattering components of the skin, as is done in the open source Digital Human project [The Digital Human League 2015].

*Ablation Study.* Fig. 2 highlights the importance of each output from the proposed network. While rendering with the diffuse reflectance layer only (second column) provides person-specific skin tones, it lacks the visual cues for the overall surface geometry provided by the specular reflection layer (third and fifth column). Using the specular albedo map in addition to the diffuse albedo map (third column) provides local specular variations in the surface reflection, but omits the subtle surface indentations provided by the displacement map (fourth column). Using all the maps provides the most realistic result (fifth column).

## 2 ADDITIONAL EVALUATIONS

We validate our results using ground truth measurements from the Light Stage with varying conditions. For a fair evaluation of the inferred textures, we employ the same base mesh obtained from the ground truth capture for each comparison. Fig.4 shows the robustness of our method using input images captured with varying camera angles. Despite the large missing regions in the input (first column), our results (shown in the second and the third rows) demonstrate a consistent estimation of the reflectance and geometry data. Rendering under a novel illumination (last column) produces results that are comparable to the ground truth (first row).

Fig.5 shows how our method performs under different facial expressions. While distinct, person-specific details such as freckles are maintained across the expressions, the shading introduced due to the expression around the nasolabial fold and the compressed forehead is successfully removed in the resulting reflectance layer. With the captured dynamic geometry, the final rendering exhibits a similar magnitude of the dynamic folds through simulated shading.

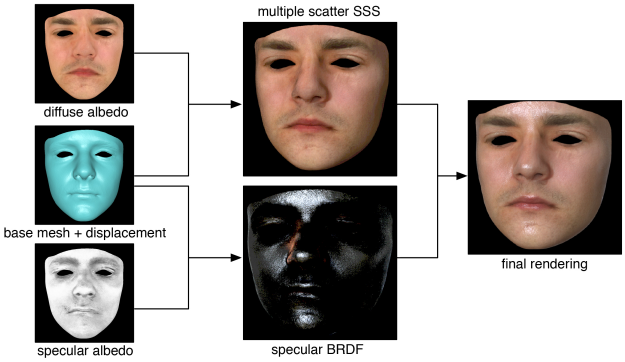


Fig. 1. Our rendering pipeline.

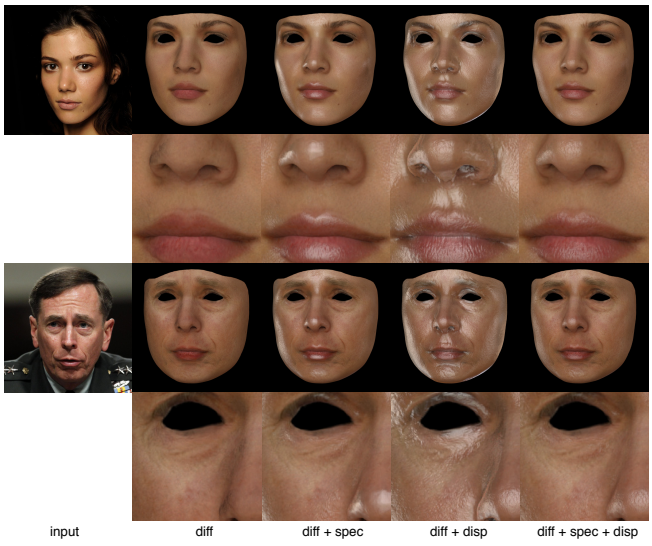


Fig. 2. Ablation study demonstrating the importance of each of the outputs of the proposed network. The second column shows the rendering of the diffuse reflectance using the diffuse albedo map, which lacks the surface geometry cues provided by the specular reflection. With the specular reflection added (third and fifth column), the view-dependent nature of the surface reflection enhances the sense of the object's 3D shape. The specular albedo map simulates the specular occlusion and provides local variations in the specular intensities (third column). Adding the displacement map provides more lifelike skin textures, but the rendering looks too shiny without the specular albedo map (fourth column). Combining all the maps provides the best rendering result (fifth column).

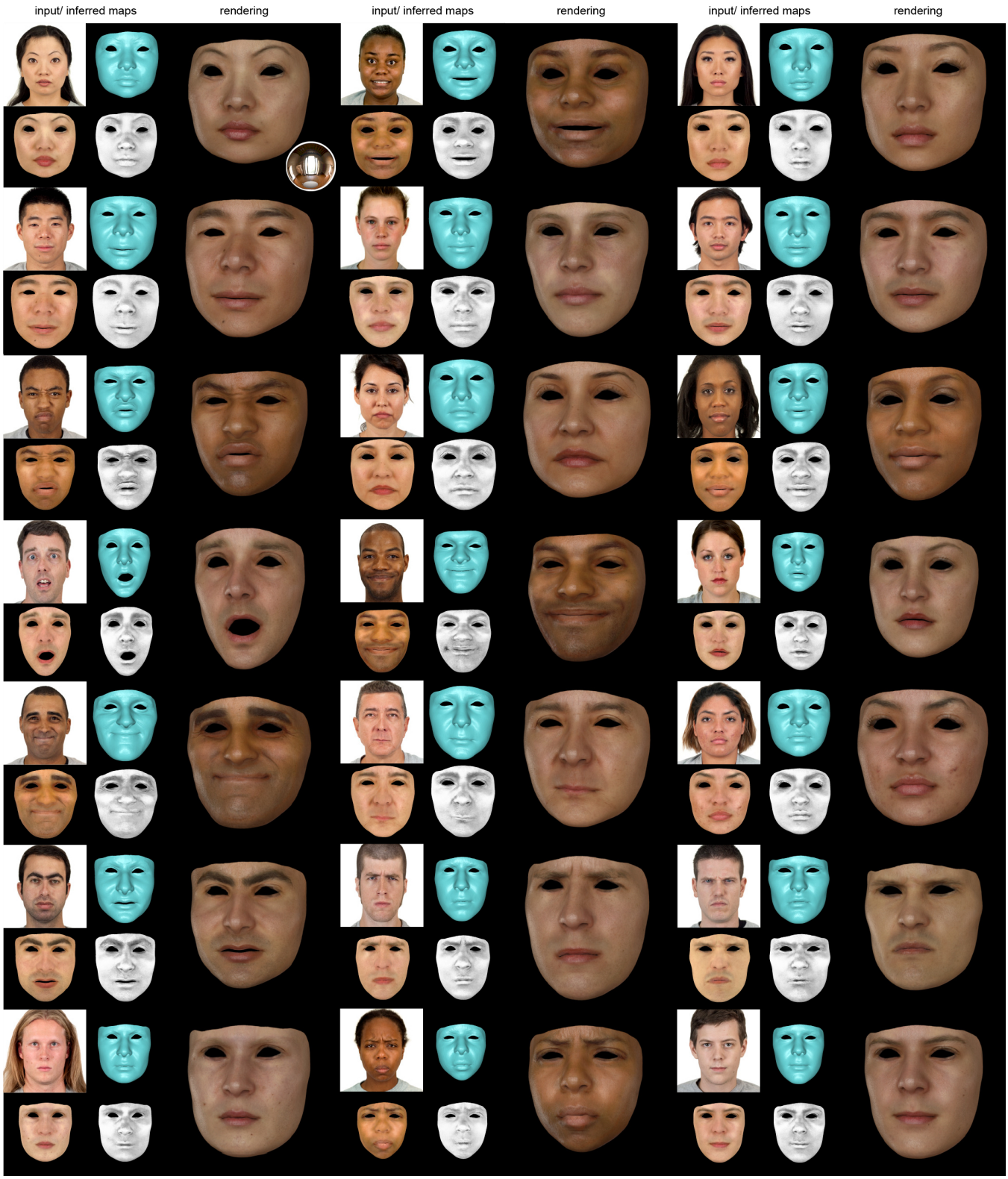


Fig. 3. Additional results with images from the Chicago Face Dataset [Ma et al. 2015]

Fig. 6 compares the results obtained under different lighting environments. Despite the strong contrast in the lighting and the strong Fresnel reflection, the inferred results exhibit a consistent skin tone and quality across different inputs. We also provide numerical evaluations in Table 1 that correspond to Figs. 4, 5, and 6, measured in the peak signal-to-noise ratio (PSNR) and the root-mean-square error (RMSE). As shown in the table, our method exhibits consistent numerical results despite the varying lighting conditions, indicating its robustness.

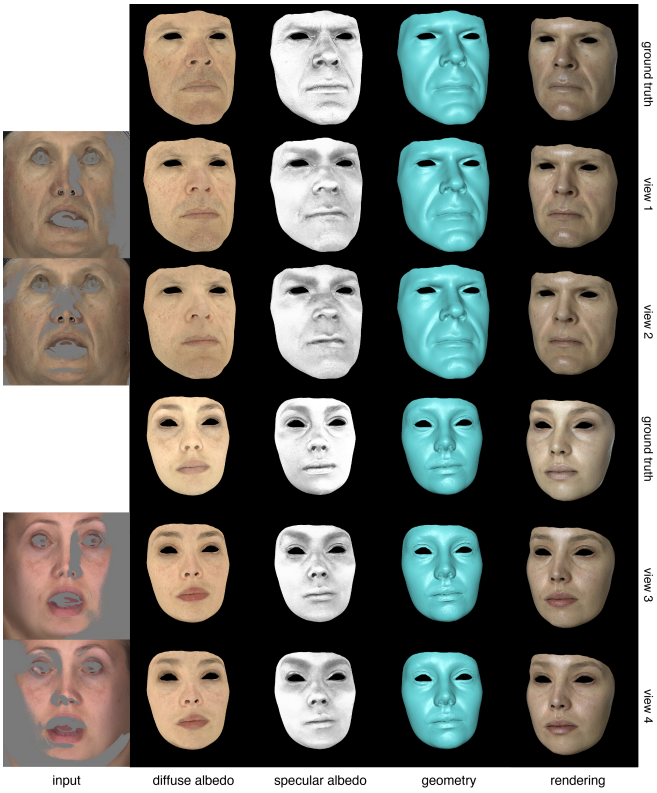


Fig. 4. Evaluation on view angles.

| | | diffuse | specular | disp |
|---|---|---|---|---|
| PSNR | cameras | 20.5 | 17.0 | 17.7 |
| | expressions | 20.4 | 17.0 | 18.4 |
| | illuminations | 20.4 | 17.1 | 18.6 |
| RMSE | cameras | 0.095 | 0.142 | 0.135 |
| | expressions | 0.096 | 0.141 | 0.132 |
| | illuminations | 0.096 | 0.139 | 0.124 |

Table 1. Evaluations for each condition


Fig. 5. Evaluation on expressions.


Solid Angle. 2018. (2018). http://www.solidangle.com/arnold/.

The Digital Human League. 2015. Digital Emily 2.0. (2015). http://gl.ict.usc.edu/Research/DigitalEmily2/.


## REFERENCES


D. S. Ma, J. Correll, and B. Wittenbrink. 2015. The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods* 47, 4 (2015), 1122–1135. DOI: https://doi.org/10.3758/s13428-014-0532-5


Fig. 6. Evaluation on various illuminations.