Variational Uncalibrated Photometric Stereo under General Lighting
Supplementary Material

Bjoern Haefner*, 1 Zhenzhang Ye*, 1 Maolin Gao[2] Tao Wu[1] Yvain Quéau[3] Daniel Cremers[1]
[1]Technical University of Munich [2]Artisense [3]GREYC, UMR CNRS 6072
{bjoern.haefner, zz.ye, tao.wu, cremers}@tum.de maolin@artisense.ai yvain.queau@ensicaen.fr

1. Further Details on Synthetic Experiments

To provide further insights on the synthetic experiments (in Section 6.1), we visualize the environment lightings ℓ^i, $i = 1 \ldots 25$, used to render each image. Figure 1 shows all 25 environment maps[1]. The impact of each incident lighting ℓ^i, $i = 1 \ldots 25$, is illustrated in Figure 2 showing the Joyful Yell with a White ($\rho \equiv 1$) albedo. Thus, color changes in the images are caused by lighting only, as depicted in model (1) and (7) in the main paper.

Table 1 shows the mean angular error (MAE) of each dataset on the state-of-the-art approaches [1, 2, 3] and our proposed methodology. It can be seen that our approach consistently overcomes [1, 2, 3] by a factor of 2–3. Only the Pattern albedo seems to bias the resulting depth negatively, yet even in this case our approach estimates the geometry more faithfully than the current state-of-the-art.

Two more qualitative results on synthetic data are shown in Figure 3. While [1] gives more meaningful results on Armadillo with Constant albedo, depth deteriorates strongly on Lucy with Hippie albedo. Methods of [2, 3] both result in rather flattened shapes (cf. Lucy). Most accurate results are achieved using the proposed method where fine geometric details, as well as non flattened depth estimates are shown.

Additional to the depth results, Figure 4 shows estimated lightings and albedos along with the ground truths. Although lighting estimates show less shadowed areas and seem brighter compared to ground truths, this does not seem to affect reflectance estimations much. The estimated albedos are satisfactory, although some shading information is slightly visible.

The initialization is indeed crucial for the whole algorithm. Here, we show two different non-trivial initializations for our algorithm in Table 1: 1) Hemisphere, we first compute the circumscribed sphere for the 3D points of ground truth. The projection of each point onto this sphere is considered as initialization; 2) Initialization by [2], we simply refine the result from [2] by our algorithm. In Figure 5, we show visualized results. In certain special cases, the initialization from [2] is slightly better. However, our minimal surface strategy is stable for all cases, and our algorithm improves the results from [2]) in most cases.

2. Further Details on Real-World Results

Supplementary to the real-world experiments (in Section 6.2), Figures 6 and 7 show alternative viewpoints of the real-world results. The estimated albedos, which are mapped onto the surfaces, appear satisfactory. Correspondingly, we also show the estimated albedos and lightings. In view of the multiplicative ambiguity between lightings and albedos, all visualized albedos are normalized to have maximum value 1.

References

[1] Paolo Favaro and Thoma Papadhimitri. A closed-form solution to uncalibrated photometric stereo via diffuse maxima. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 821–828, 2012. 1, 4, 5

[2] Zhipeng Mo, Boxin Shi, Feng Lu, Sai-Kit Yeung, and Yasuyuki Matsushita. Uncalibrated photometric stereo under natural illumination. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2936–2945, 2018. 1, 4, 5, 7

[3] Songyou Peng, Bjoern Haefner, Yvain Quéau, and Daniel Cremers. Depth super-resolution meets uncalibrated photometric stereo. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV) Workshops*, pages 2961–2968, 2017. 1, 4, 5

*Authors contributed equally.

[1]All environment maps were downloaded from http://www.hdrlabs.com/sibl/archive.html

Figure 1. All environment maps ℓ^i (360° view) used throughout the synthetic evaluation.

Figure 2. Illustration of the input data. The Joyful Yell dataset with White albedo to show the impact of the different environment maps used throughout the synthetic experimental validation.

Dataset		[1]	[3]	[2]	Our approach with different initializations		
Shape	Albedo				Hemisphere	Using [2]	Minimal surface (Sec. 5.1)
Armadillo	Bars	26.22	27.84	36.91	79.54	20.08	**16.78**
	Constant	25.84	26.64	36.87	83.01	18.81	**13.97**
	Ebsd	25.34	26.88	27.80	82.53	15.99	**14.26**
	Hippie	28.21	27.30	25.82	79.12	**12.56**	14.52
	Lena	27.07	27.33	28.36	84.24	17.79	**14.78**
	Pattern	45.87	26.82	24.01	82.59	19.39	**19.06**
	Rectcircle	26.97	26.71	36.23	80.68	19.64	**14.06**
	Voronoi	25.62	26.91	50.70	79.65	55.29	**14.07**
	White	26.19	26.64	52.04	83.04	56.74	**14.13**
Joyful Yell	Bars	21.84	16.26	31.80	21.21	28.82	**8.69**
	Constant	23.95	14.93	33.47	16.85	29.31	**5.96**
	Ebsd	26.08	15.63	15.91	17.63	7.49	**7.28**
	Hippie	28.67	16.23	22.96	17.68	**7.47**	7.49
	Lena	21.33	16.33	19.70	20.11	13.16	**9.21**
	Pattern	26.07	18.76	26.67	18.76	21.03	**16.97**
	Rectcircle	35.27	15.19	52.41	16.27	61.77	**7.34**
	Voronoi	22.27	16.42	45.74	18.62	54.78	**6.57**
	White	27.12	14.32	33.06	17.70	28.99	**6.20**
Lucy	Bars	49.13	21.90	36.51	40.55	26.15	**8.16**
	Constant	54.98	19.89	36.57	41.00	25.74	**8.71**
	Ebsd	62.33	20.81	23.56	40.80	13.36	**9.61**
	Hippie	58.61	21.29	32.38	39.93	8.10	**7.87**
	Lena	64.01	22.24	30.93	40.16	19.14	**9.56**
	Pattern	48.83	22.25	32.68	40.11	20.56	**17.78**
	Rectcircle	24.68	20.99	43.13	41.17	10.01	**8.98**
	Voronoi	61.53	22.10	48.14	40.39	71.32	**7.59**
	White	64.43	19.33	44.76	41.54	72.45	**8.76**
Thai Statue	Bars	25.53	21.91	66.17	78.72	8.94	**8.55**
	Constant	27.20	18.91	38.47	81.14	24.26	**9.58**
	Ebsd	27.85	20.22	34.11	79.58	19.23	**9.47**
	Hippie	21.91	21.86	30.62	77.27	12.78	**8.83**
	Lena	33.53	19.66	34.00	79.43	19.55	**9.19**
	Pattern	26.77	22.06	28.81	83.92	16.69	**15.27**
	Rectcircle	29.36	19.92	43.86	81.88	79.88	**8.84**
	Voronoi	30.65	21.56	36.58	78.92	25.21	**8.69**
	White	28.02	18.64	37.31	81.54	24.94	**9.16**
Median		27.16	21.14	34.06	59.41	19.86	**9.17**
Mean		34.15	21.18	35.53	55.20	27.43	**10.72**

Table 1. Quantitative comparison between our method and other state-of-the-art methods on challenging synthetic datasets. The last three columns refer to the results with different initializations for our approach.

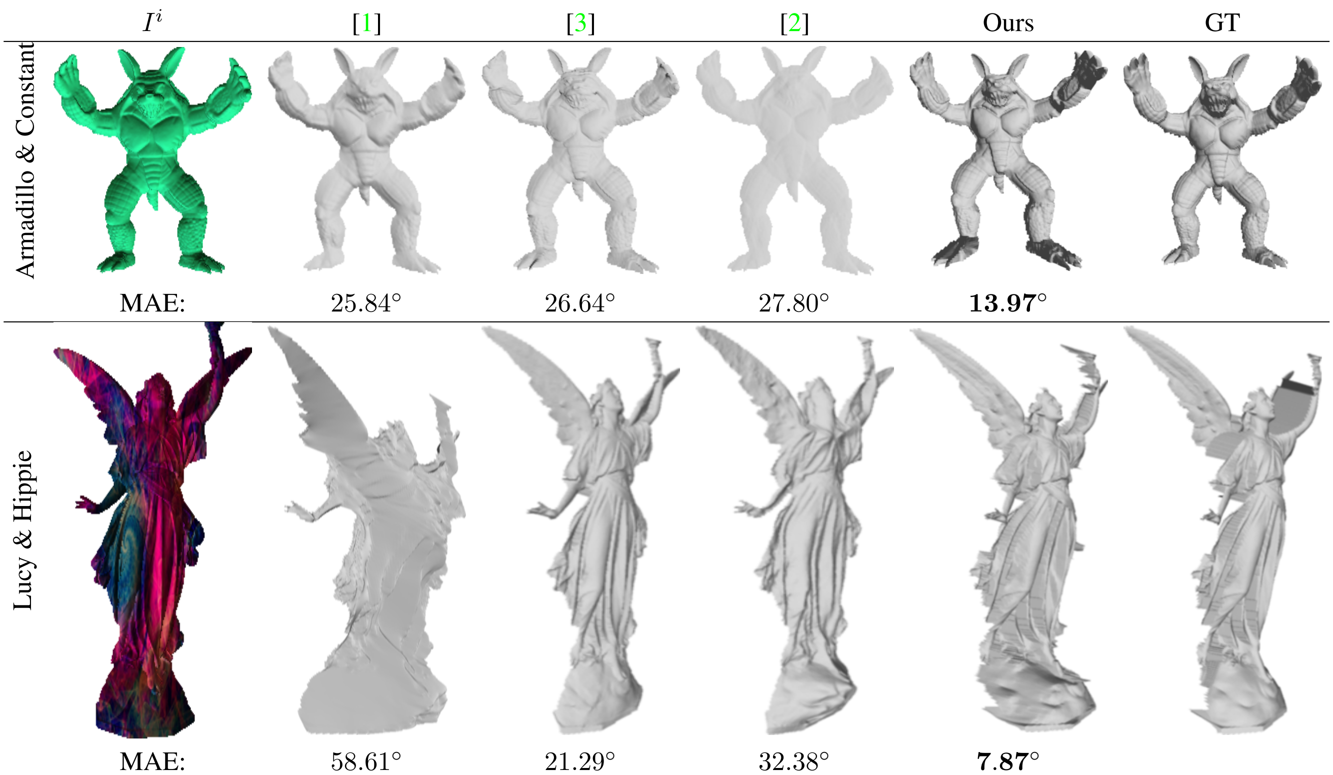

Figure 3. Results of state-of-the-art approaches and our approach on two out of the 36 synthetic datasets. Numbers show the mean angular error (MAE) in degrees.

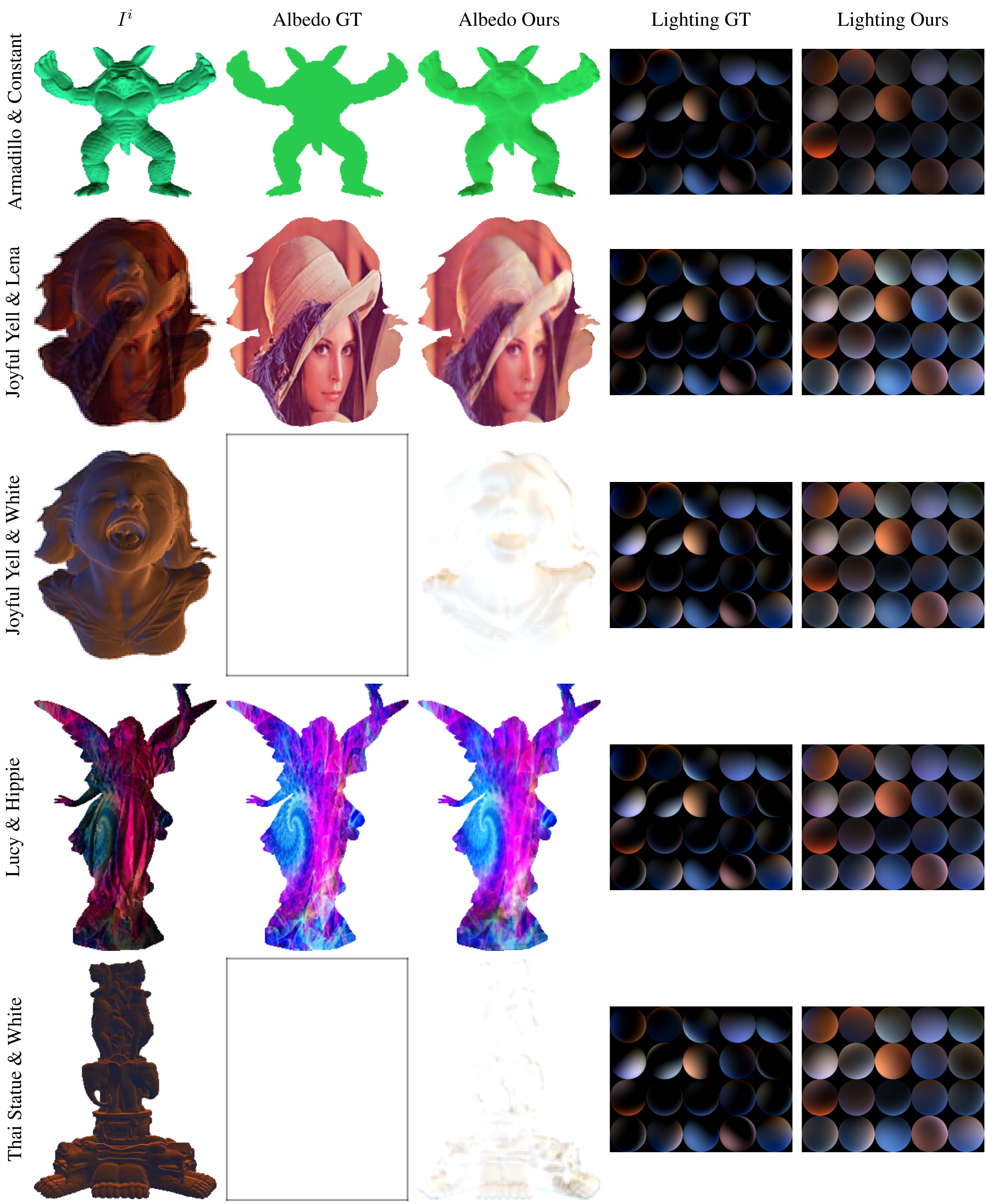

Figure 4. Our estimated albedos and lighting next to the ground truth. Lighting estimates show less shadowed areas and seem brighter compared to ground truth, yet this does not seem to affect reflectance and geometry estimation much, cf. Figure 7 in main paper and Figure 3 in the supplementary material. The estimated albedos are satisfactory, although some shading information is slightly visible.

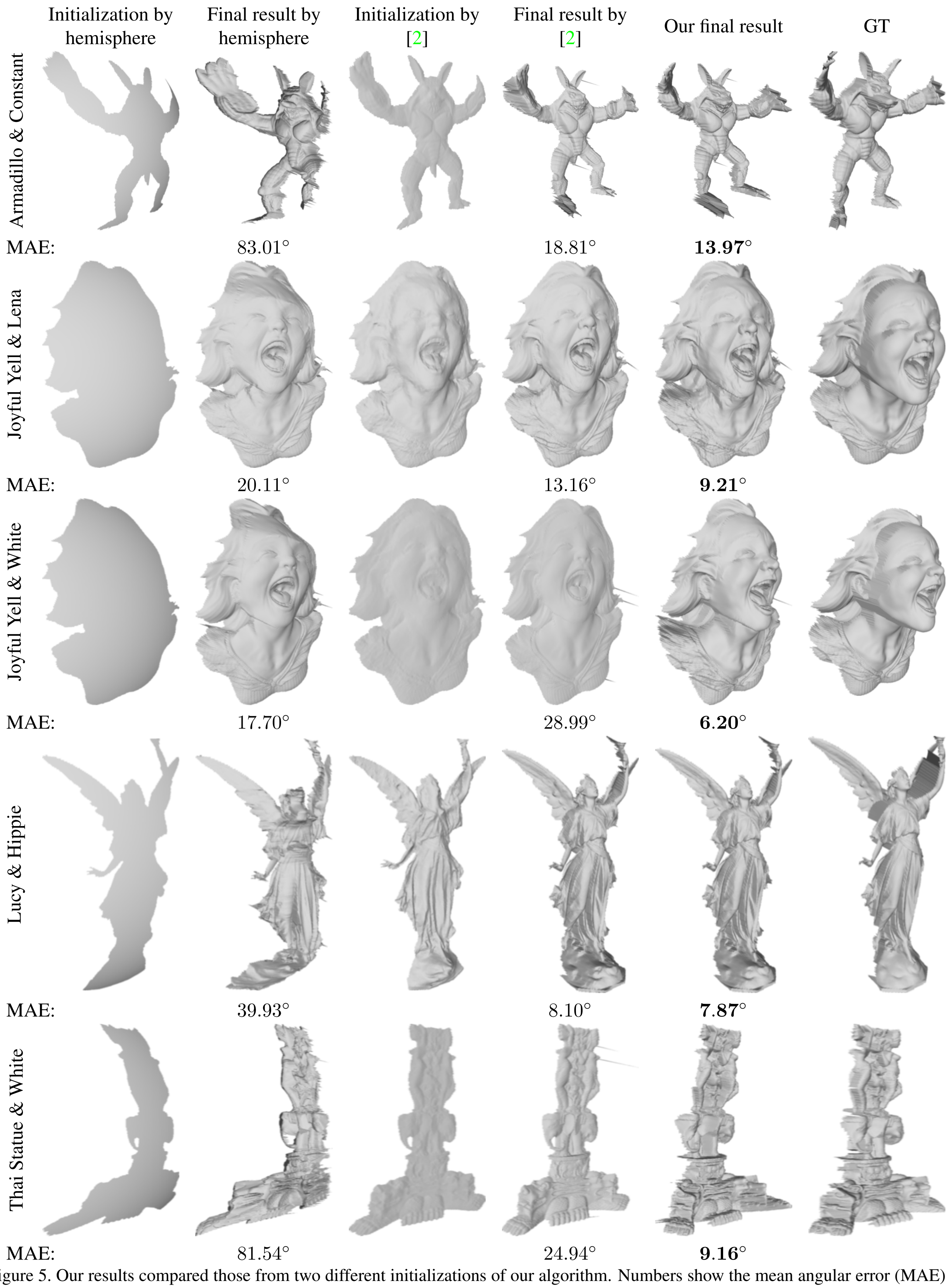

Figure 5. Our results compared those from two different initializations of our algorithm. Numbers show the mean angular error (MAE) in degrees. Though the initialization by [2] achieves comparable result to ground truth on "Lucy & Hippie" dataset, its performance is not stable across different datasets.

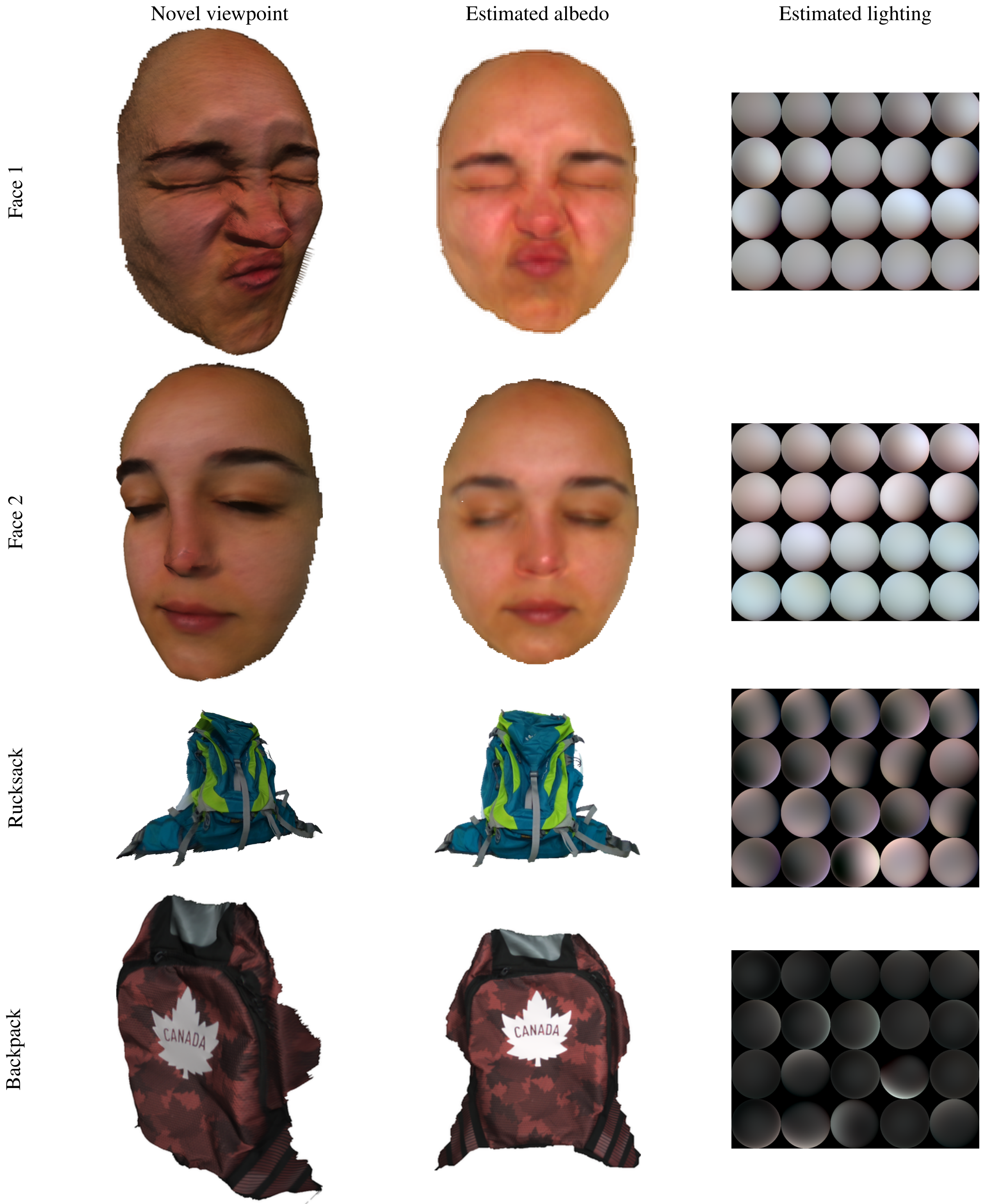

Figure 6. Real-world results: (left) estimated albedos mapped onto estimated surfaces rendered under a novel viewpoint, (middle) estimated albedos, (right) estimated lightings for all $M = 20$ input images.

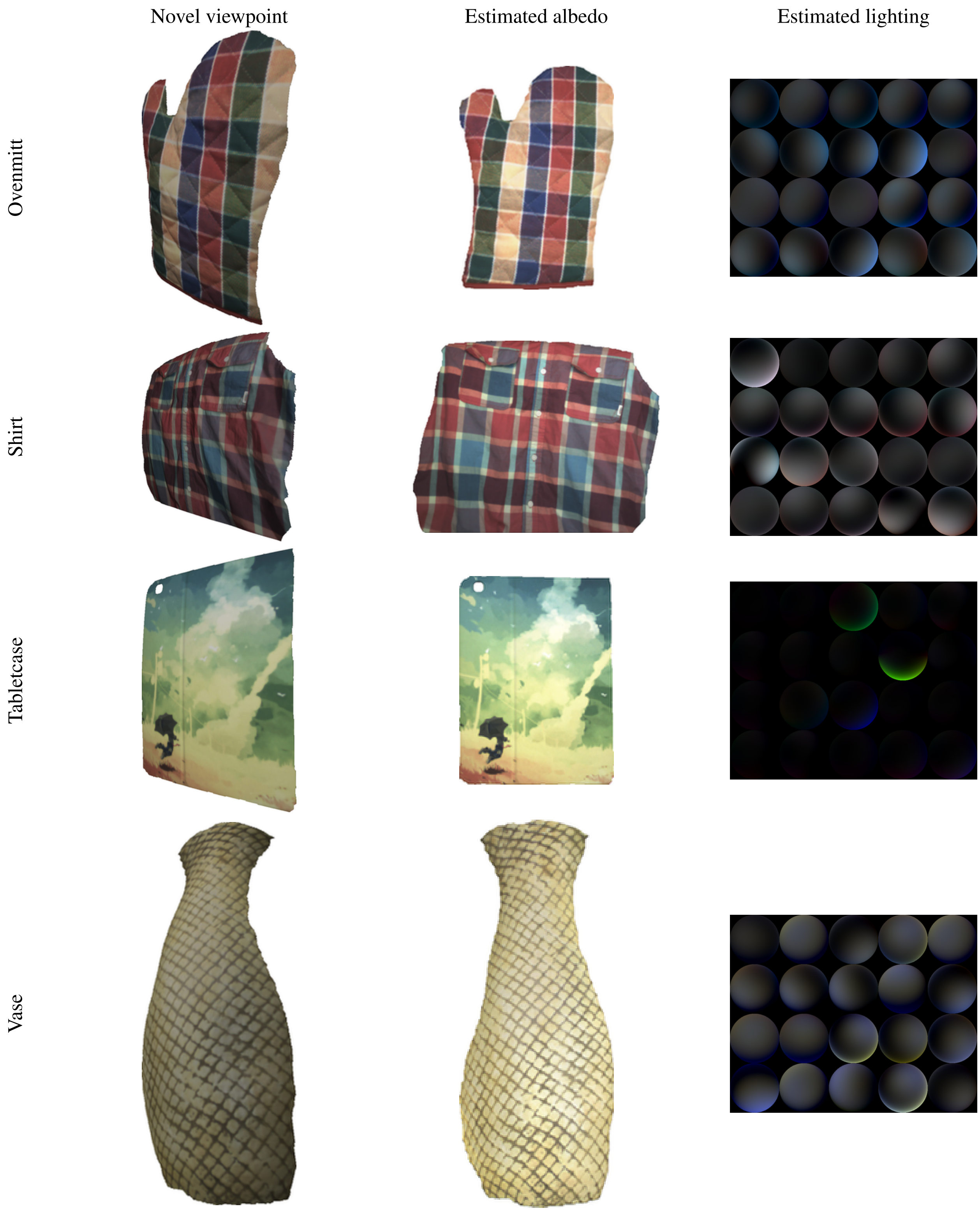

Figure 7. More real-world results: (left) estimated albedos mapped onto estimated surfaces rendered under a novel viewpoint, (middle) estimated albedos, (right) estimated lightings for all $M = 20$ input images.