Practical Measurements of Translucent Materials with Inter-Pixel Translucency Prior—Supplemental Materials

Zhenyu Chen[1], Jie Guo[1*], Shuichang Lai[1], Ruoyu Fu[1], Mengxun Kong[1], Chen Wang[1]
Hongyu Sun[2], Zhebin Zhang[2], Chen Li[2], Yanwen Guo[1]
[1]Nanjing University, Nanjing, China
[2]OPPO, Seattle, USA
{chenzy, sclai, 502022330022, chenwang}@smail.nju.edu.cn
{hongyu.sun, zhebin.zhang, chen.li}@oppo.com
{guojie,fry,ywguo}@nju.edu.cn

1. Network Details

In this section, we provide the implementation details of the encoders in out IPTNet. All the encoders are designed as described in Fig. 1. Each convolutional block is followed by LeakyReLU activations. For MLPs $\mathcal{M}_1$ and $\mathcal{M}_2$, we use a 128-channel, four-layer, fully connected network with ReLU activations as their shared architecture, and the last layer of each MLP applies Sigmoid activations.
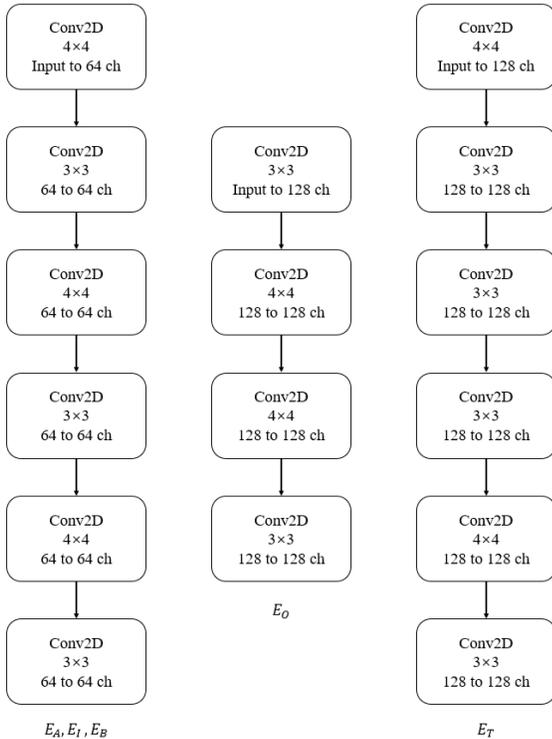
Figure 1. Flow charts of multiple encoders.

2. Impact of Lighting and Shadows
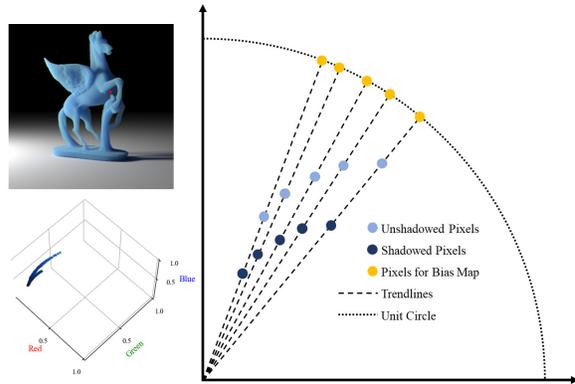
Figure 2. An illustration of a translucent curve showing the impact of shadows. Shadows are generated by applying a scalar to the pixel value. The effect of the scalar will be eliminated by the vector normalization operator.

Our experimental setting assumes that the object is lit by a pure white light source. If that is not the case for real-world images, we apply the color constancy algorithm in [5] to normalize the input images. After that, we have a pure white illumination color.

Shadows generally have little impact on the pseudo-albedo maps and bias maps if the shadow regions are not very dark. For optically thick parts, as pixel values are plotted under a linear distribution, shadows cause the corresponding pixel values to move along straight lines. For optically thin parts, shadows shift the translucent curves to different extent of nonlinearity, and later the effect is balanced by the regression process. The impact of shadows on bias maps is shown in Fig. 2. As bias maps are calculated by the mean Euclidean distance between two normalized

pixel values, both unshadowed and shadowed pixel values map to the same position on the unit circle, where the pixel on the bias map lies. However, for regions with very dark shadows, the normalization operation will introduce large errors due to the numerical instability.

3. Impact of Highlights and Transparency

Fig. 3 demonstrates some failure examples of our IPTNet. Since our goal is to predict the global information, not pixel-wise prediction, small regions of highlights are acceptable to our pipeline. These overexposed pixels will be considered as outliers and removed during regression. However, highlights covering a large area (*i.e.*, too many overexposed pixels) invalidate the calculation of the prior (the first row of Fig. 3), as a relatively large amount of pixel values are discarded, leading to large errors.

The second row of Fig. 3 shows that our IPTNet may also fail on near transparent materials. This is another limitation of our method.

4. Impact of Noise

Fig. 4 shows the effect of noise on the results by adding zero mean Gaussian noise with variable standard deviation σ to input images. A small amount of noise slightly alters the prediction of the network as well as the calculation of pseudo-albedo maps, because CNN effectively acts as a filter and these maps are obtained by regression. However, the prediction accuracy decrease when the noise becomes significant, since there will be large errors in the estimated pseudo-albedo maps and bias maps.

5. Details about Translucent Materials' Optical Parameters

The macroscopic bulk optical parameters of translucent materials include the extinction coefficient σ_t, the single-scattering albedo Λ and the phase function $p(\theta)$:

- the extinction coefficient σ_t represents the combined effect of absorption and out-scattering attenuation,
- the single-scattering albedo Λ represents the probability of scattering instead of absorption at a scattering event, and
- the phase function $p(\theta)$ describes the angular probability density of the scattered direction at a scattering point.

6. BSDF Layer Details

For some translucent materials with specular reflection, we optionally coat a dielectric BSDF layer during rendering. The reflective index is manually set according to the material categories (e.g., 1.5 for licite and 1.3 for water). The effect of coating is illustrated in Fig. 5.

7. More Comparison Results

In table 1 and Fig. 6, we have compared our method with an inverse-rendering based method [4], whose problem setting is similar with our work. Our method generally outperforms the method of Li *et al.* [4].

8. More Visualization Results

We show more visualization results about our pseudo-albedo maps and bias maps in Fig. 7. Additionally, more comparison results with ITN-based methods [1] and inverse rendering based methods [2, 3] are illustrated in Fig. 8 and Fig. 9. We also provide visualization results of ablation studies in Fig. 10. The noise in the rendered images is due to a low sampling rate of the renderer.

References

[1] Chengqian Che, Fujun Luan, Shuang Zhao, Kavita Bala, and Ioannis Gkioulekas. Towards learning-based inverse subsurface scattering. In *2020 IEEE International Conference on Computational Photography (ICCP)*, pages 1–12, 2020. 2, 7

[2] Wenzel Jakob, Sébastien Speierer, Nicolas Roussel, Merlin Nimier-David, Delio Vicini, Tizian Zeltner, Baptiste Nicolet, Miguel Crespo, Vincent Leroy, and Ziyi Zhang. Mitsuba 3 renderer, 2022. https://mitsuba-renderer.org. 2, 8

[3] Wenzel Jakob, Sébastien Speierer, Nicolas Roussel, and Delio Vicini. Dr.jit: A just-in-time compiler for differentiable rendering. *Transactions on Graphics (Proceedings of SIGGRAPH)*, 41(4), 2022. 2, 8

[4] Chenhao Li, Trung Thanh Ngo, and Hajime Nagahara. Inverse rendering of translucent objects using physical and neural renderers, 2023. 2, 3, 5

[5] T.T. Tan, K. Nishino, and K. Ikeuchi. Illumination chromaticity estimation using inverse-intensity chromaticity space. In *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.*, pages I–I, 2003. 1
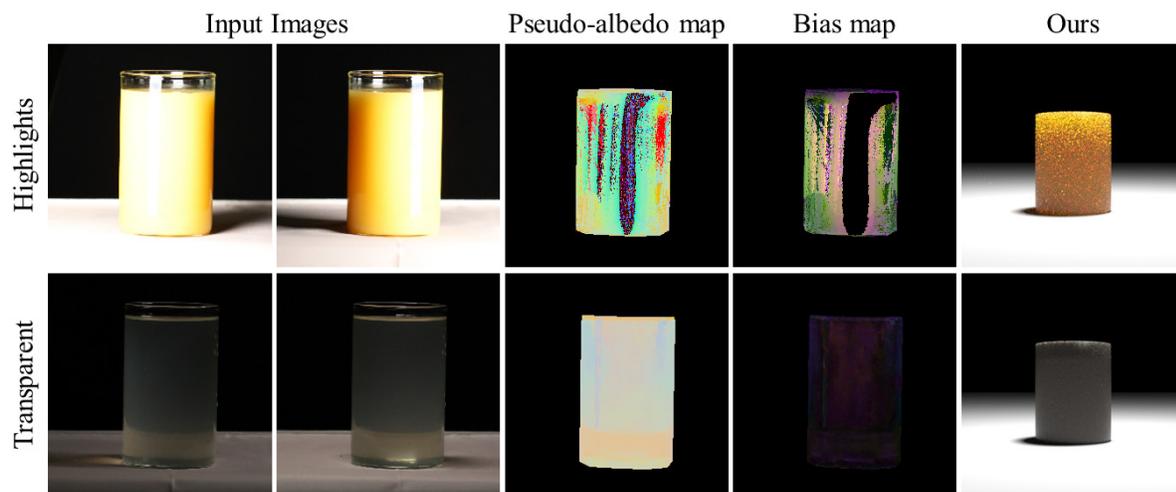
Figure 3. Two failure examples of our method.

Table 1. Quantitative comparison with Li *et al.* [4] on a test set with novel shapes and lighting conditions.

	$\sigma_t'\downarrow$	$\Lambda'\downarrow$	$g\downarrow$	PSNR↑	SSIM↑
Ours	0.023	**0.007**	**0.041**	**30.60**	**0.975**
Li *et al.* [4]	**0.011**	0.116	0.107	24.33	0.929
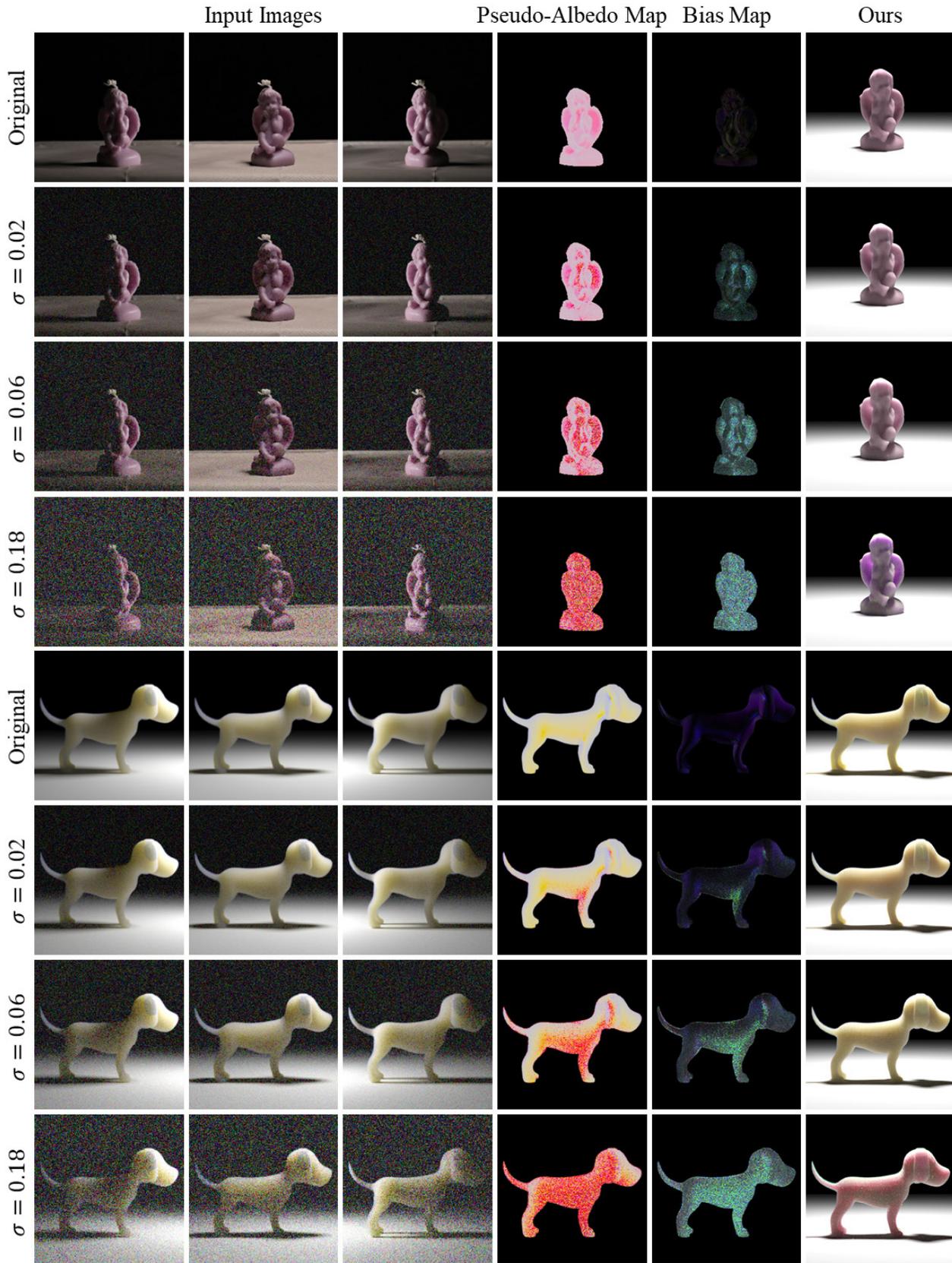
Figure 4. Impact of the noise on the estimated pseudo-albedo maps and bias maps, as well as the final prediction.
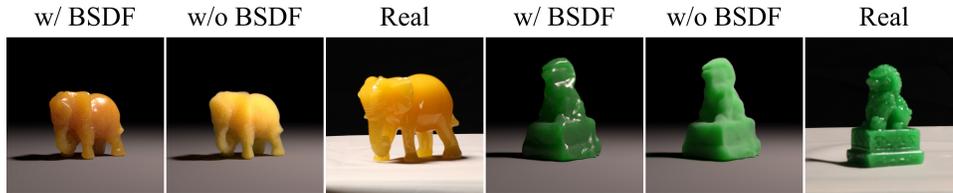
Figure 5. Rendering without/with BSDF coating.
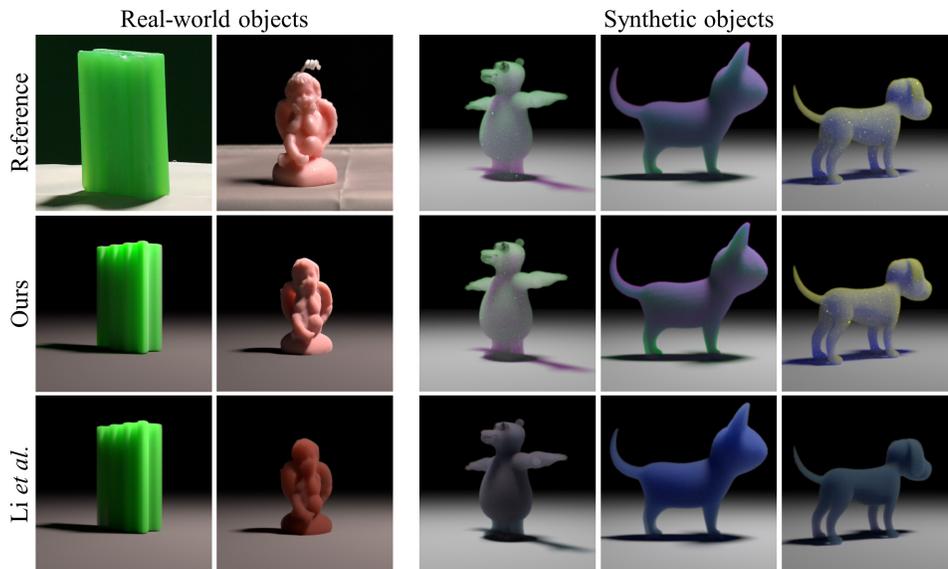
Figure 6. Visual comparison with Li *et al.* [4].

Rendered images under varying lighting   Pseudo-albedo map   Bias map
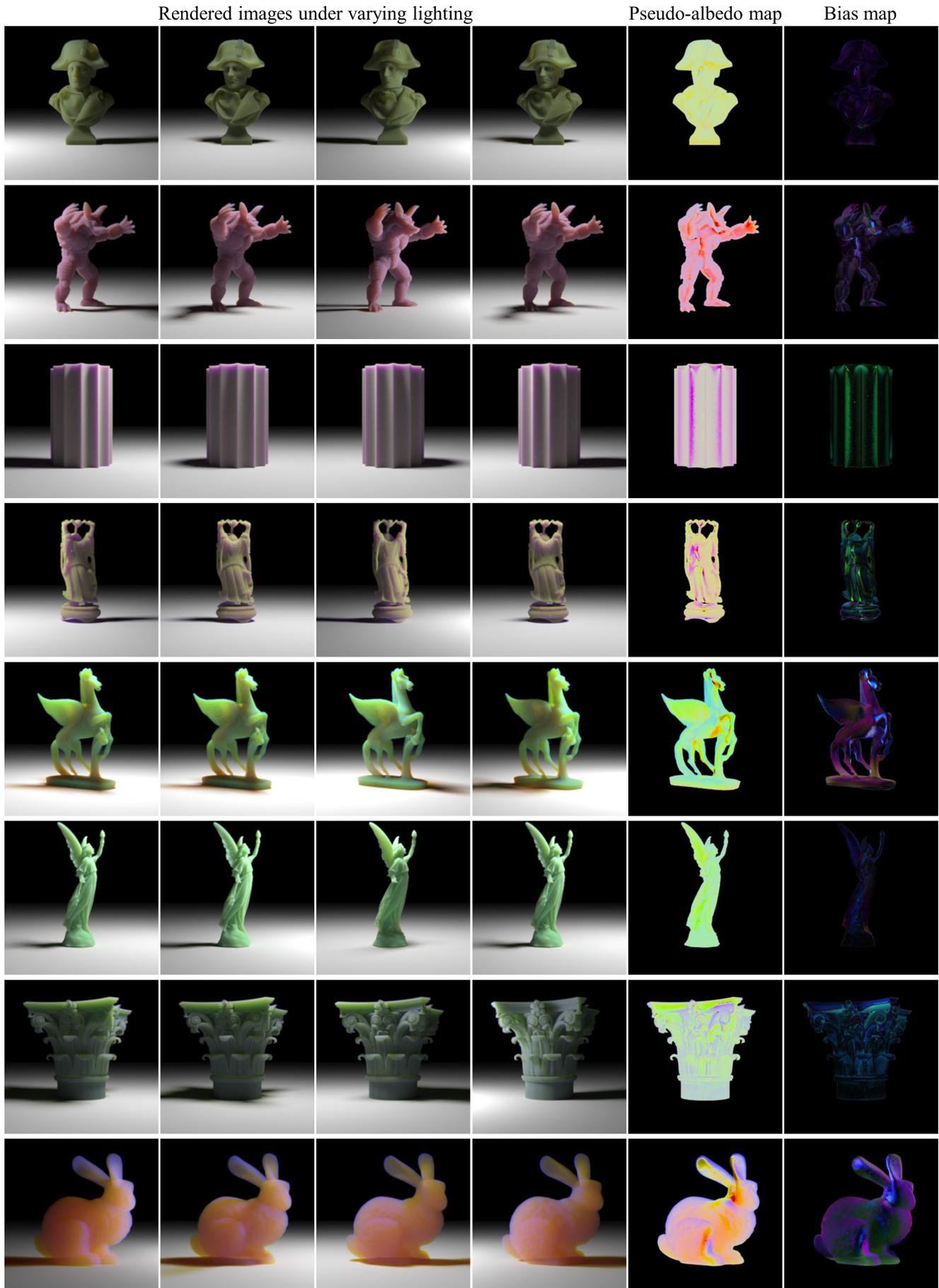
Figure 7. Demonstration of rendered images, pseudo-albedo maps and bias maps.
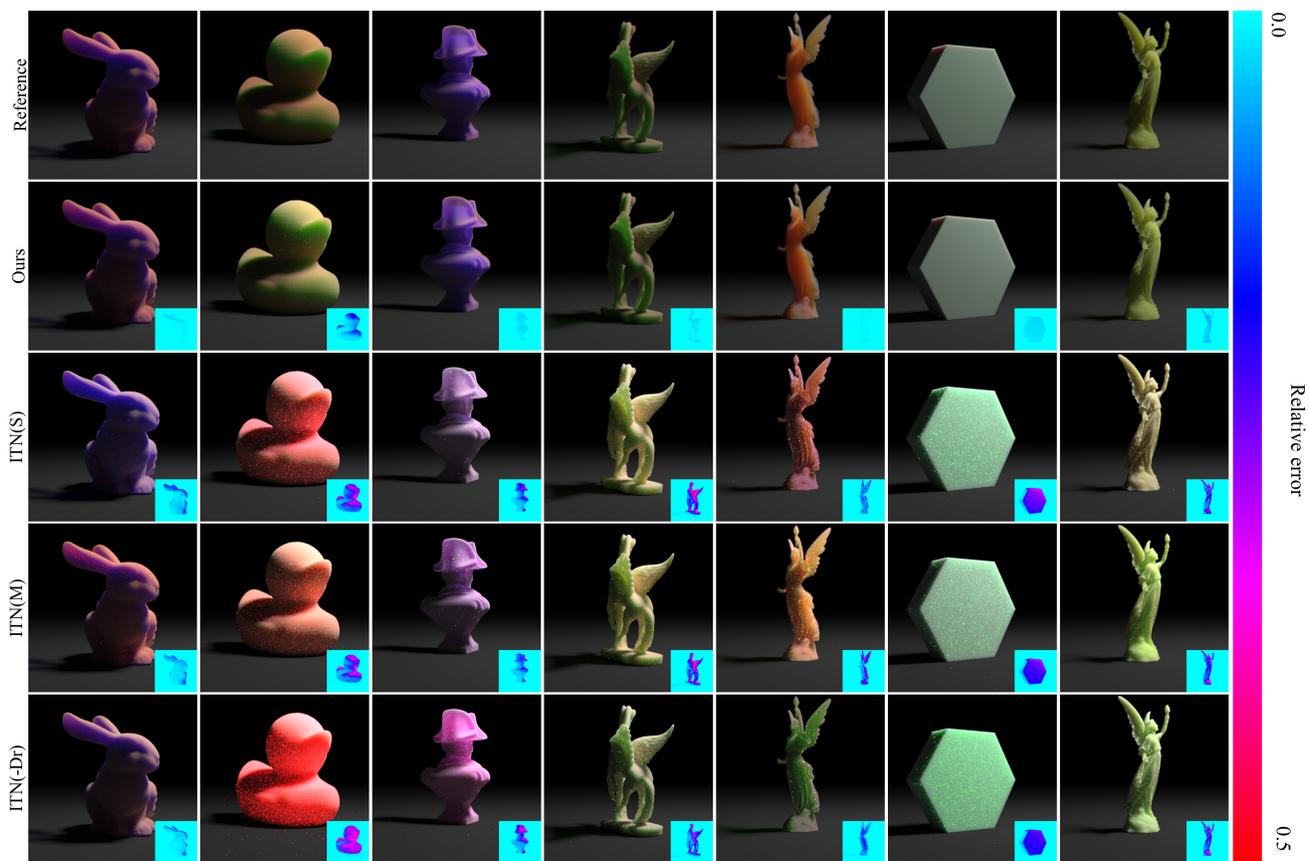
Figure 8. More comparison results with different variants of ITN [1].
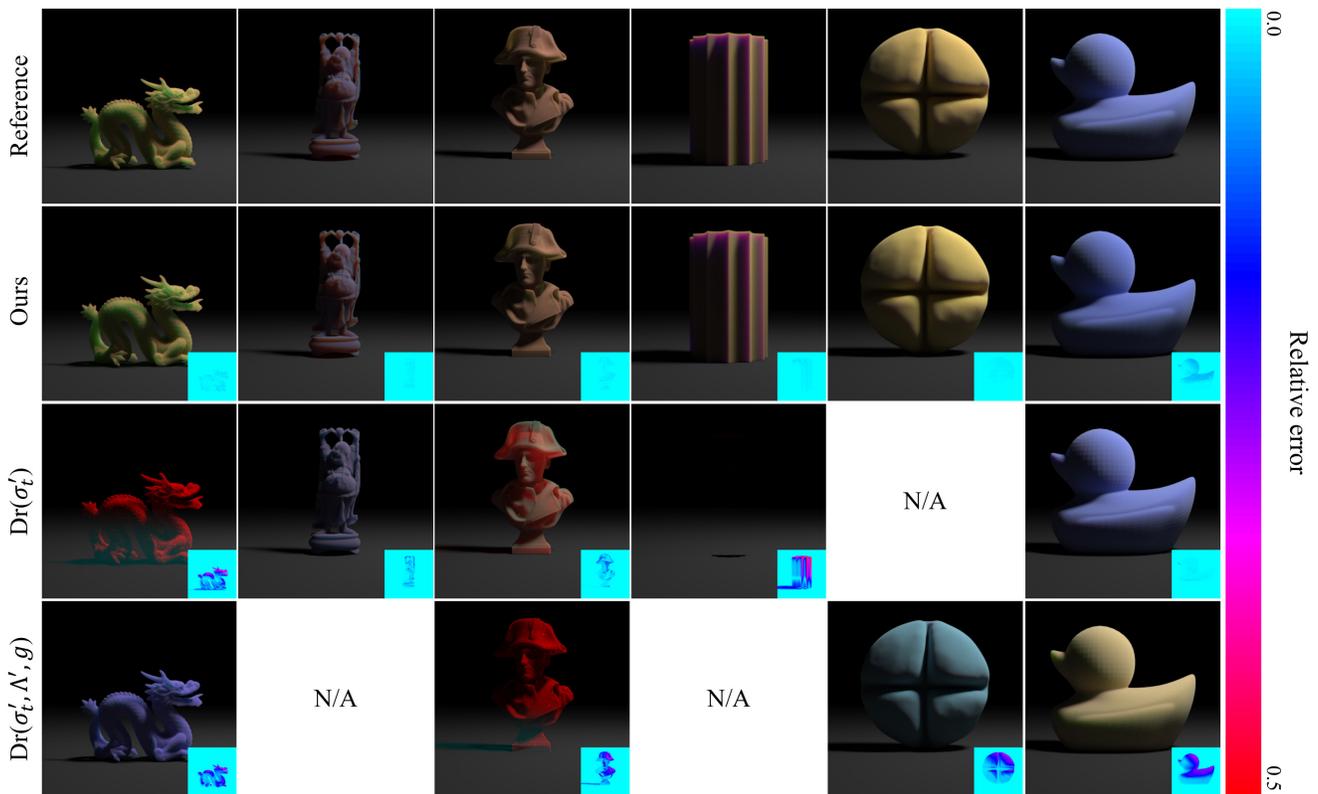
Figure 9. More comparison results with inverse rendering based methods [2, 3]. N/A indicates that the optimized scattering parameters are invalid and can not be rendered. Taking scene lighting conditions, Λ_t' and g as its input, $\mathrm{Dr}(\sigma_t')$ may outperform our IPTNet (the sixth column of the image) when it converges to correct parameters.
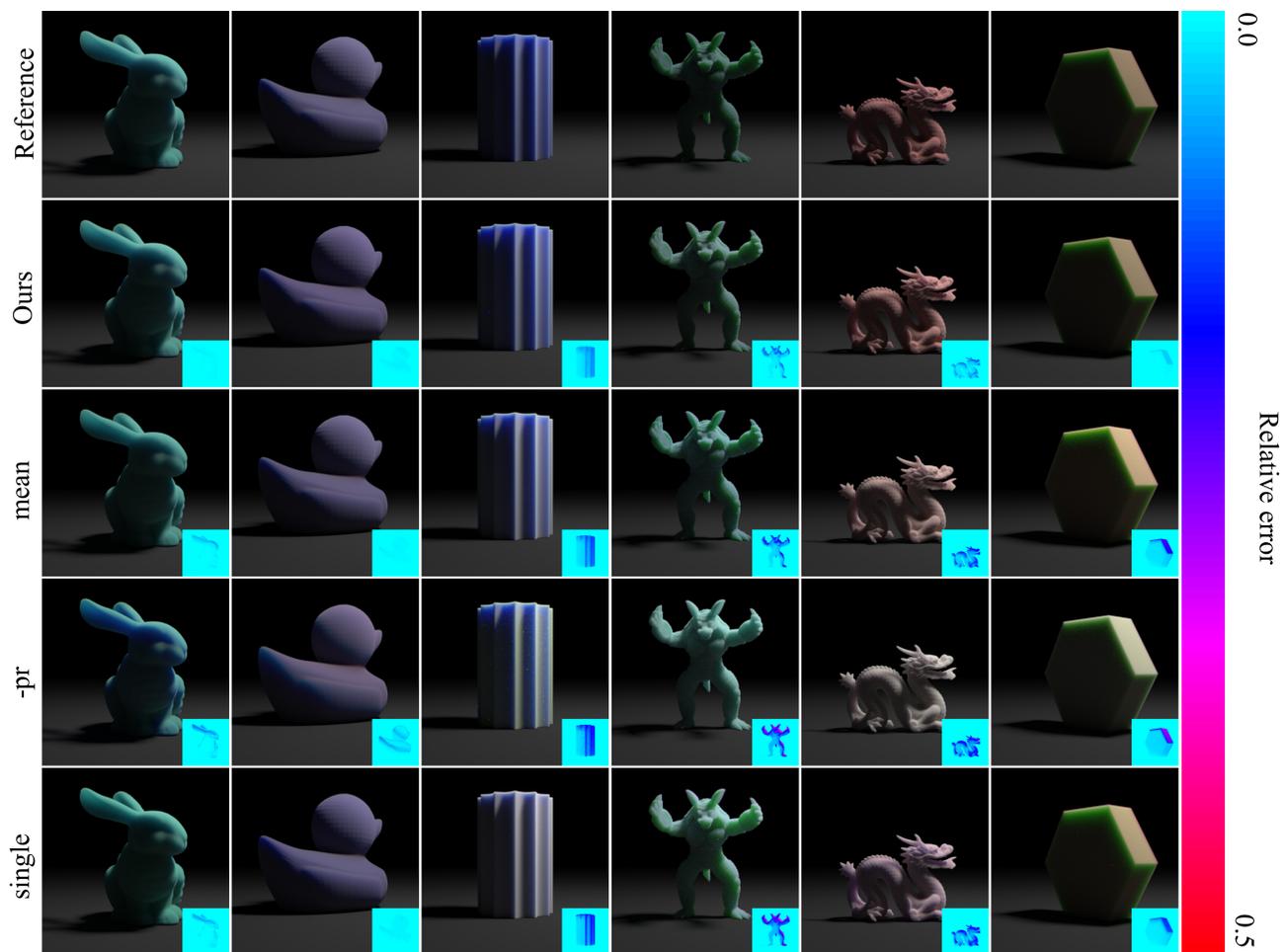
Figure 10. Visual comparison results of different variants of our model.

	Undiluted	Diluted 4x	Diluted 16x
Reference			
Ours			
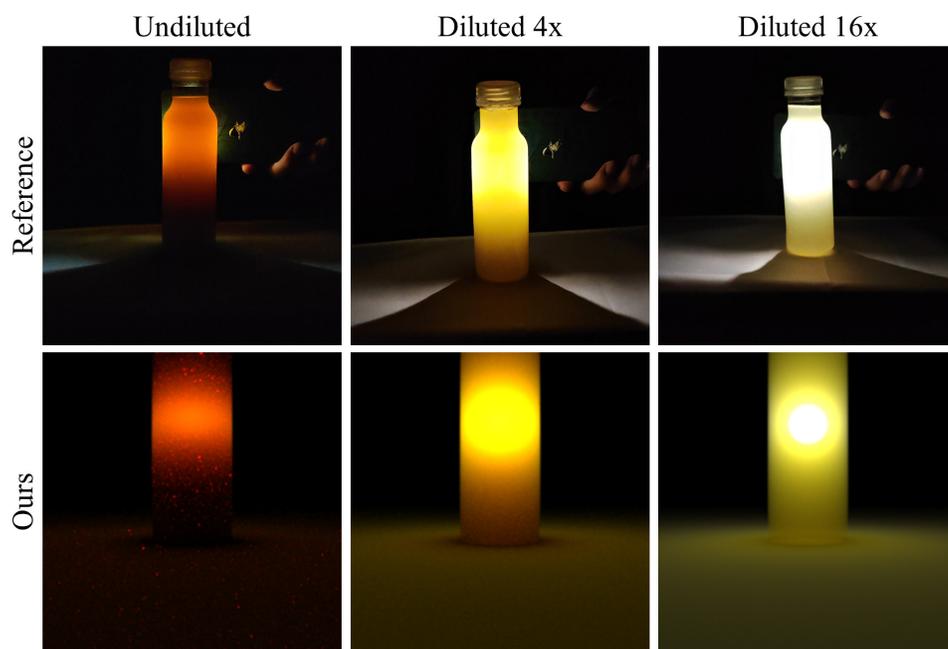
Figure 11. Test on orange juice at various concentrations.