Differentiable Adaptive 4D Structured Illumination for Joint Capture of Shape and Reflectance

Supplementary Material

1. Calibrations

Below we describe a differentiable, end-to-end calibration approach for the spatial-angular structured light. Compared with previous work [44], our approach is fully end-to-end and does not rely on manual feature extractions. Our goal is to minimize the differences between image measurements and synthetic rendering results of a planar calibration board at different poses, lit with different light and mask patterns.

Initialization. With the ARTags [11] printed on the calibration board, we estimate the intrinsic and extrinsic parameters of the camera. We also use a ruler to roughly estimate the extrinsic parameters of the structured light, including the position/orientation of each LED and the pose of the LCD mask.

Acquisition. Next, we capture using 6 different sets of light and mask patterns (visualized in Fig. S1), with 12 different poses of the calibration board for each set. This amounts to a total of 72 HDR photographs. Please see Fig. S2 for examples. For each photograph, a region of interest is manually specified with the help of SAM [20]. Each pose of the calibration board can be determined by identifying the ARTags under environment illumination.

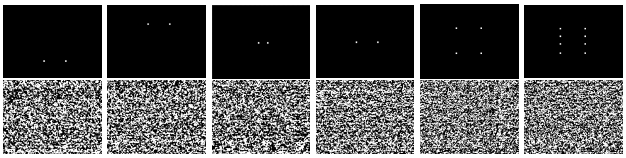

Figure S1. Visualization of the patterns used in calibration. Top row: light patterns; bottom row: mask patterns.

Optimization. Finally, we optimize various parameters of our system to minimize the differences between our rendering results and image measurements, similar to Sec. 6.2. We assume that the calibration board is diffuse with unknown, spatially-varying albedos. The position/orientation of each LED, the pose of the LCD mask, along with the spatially-varying albedos are jointly refined with the Adam optimizer. Please refer to Fig. S2 for examples.

2. Additional Results

We show additional relighting results for objects in Fig. 6 with novel lighting conditions in Fig. S3, S4, S5, S7 and S5. Quantitative errors are reported in the figures as the average SSIM/LPIPS/PSNR over all image pairs.

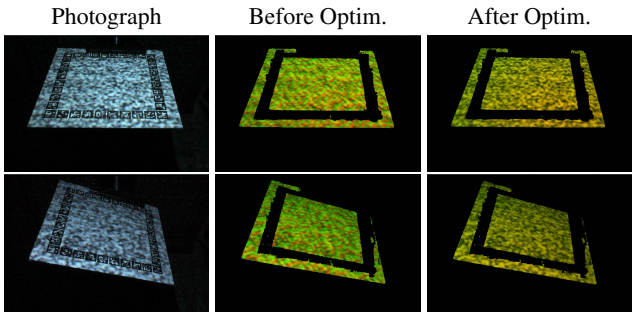

Figure S2. Visualization of the optimization process in our calibration. The first column shows captured photographs. The second and last columns are visualizations of our rendering result (in the red channel) in conjunction with a corresponding photograph (in the green channel), before and after our optimization, respectively. Note that when our rendering result closely matches the photograph, such a visualization appears yellow, instead of being red or green.

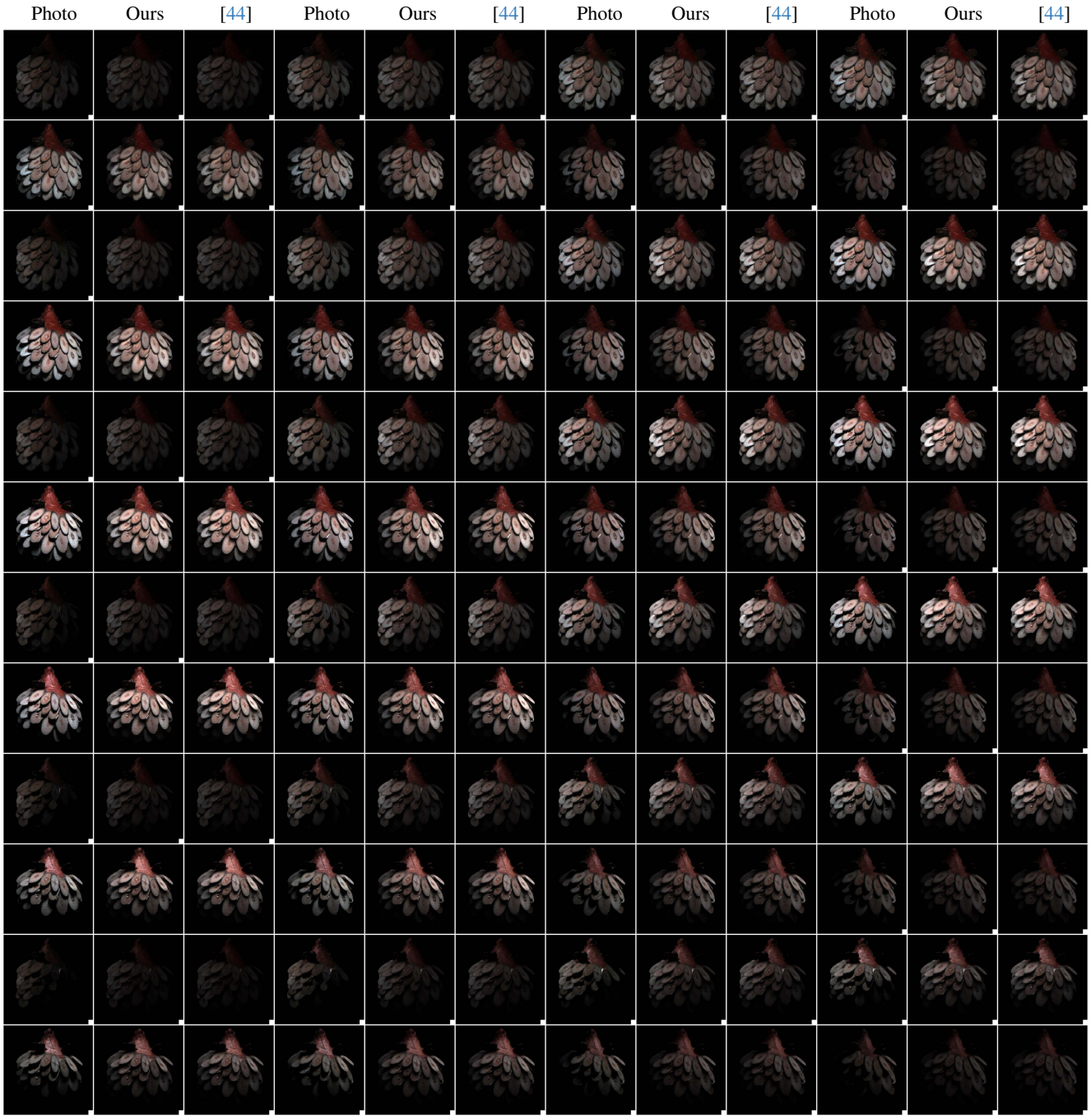

Figure S3. Relighting results of HEDGEHOG under 48 novel lighting conditions. For each triplet of images, from the left to right, a photograph, the relighting result of our approach and [44]. The average SSIM/LPIPS/PSNR are 0.89/0.080/31.03 for our approach, and 0.89/0.080/31.03 for [44]. Note that 24 out of 48 lighting conditions are unseen ones, whose LEDs are never used in any of our acquisition patterns. We mark such cases with a missing bottom-right corner.

Figure S4. Relighting results of MARUKO under 48 novel lighting conditions. For each triplet of images, from the left to right, a photograph, the relighting result of our approach and [44]. The average SSIM/LPIPS/PSNR are 0.96/0.036/34.16 for our approach, and 0.96/0.035/34.17 for [44]. Note that 24 out of 48 lighting conditions are unseen ones, whose LEDs are never used in any of our acquisition patterns. We mark such cases with a missing bottom-right corner.

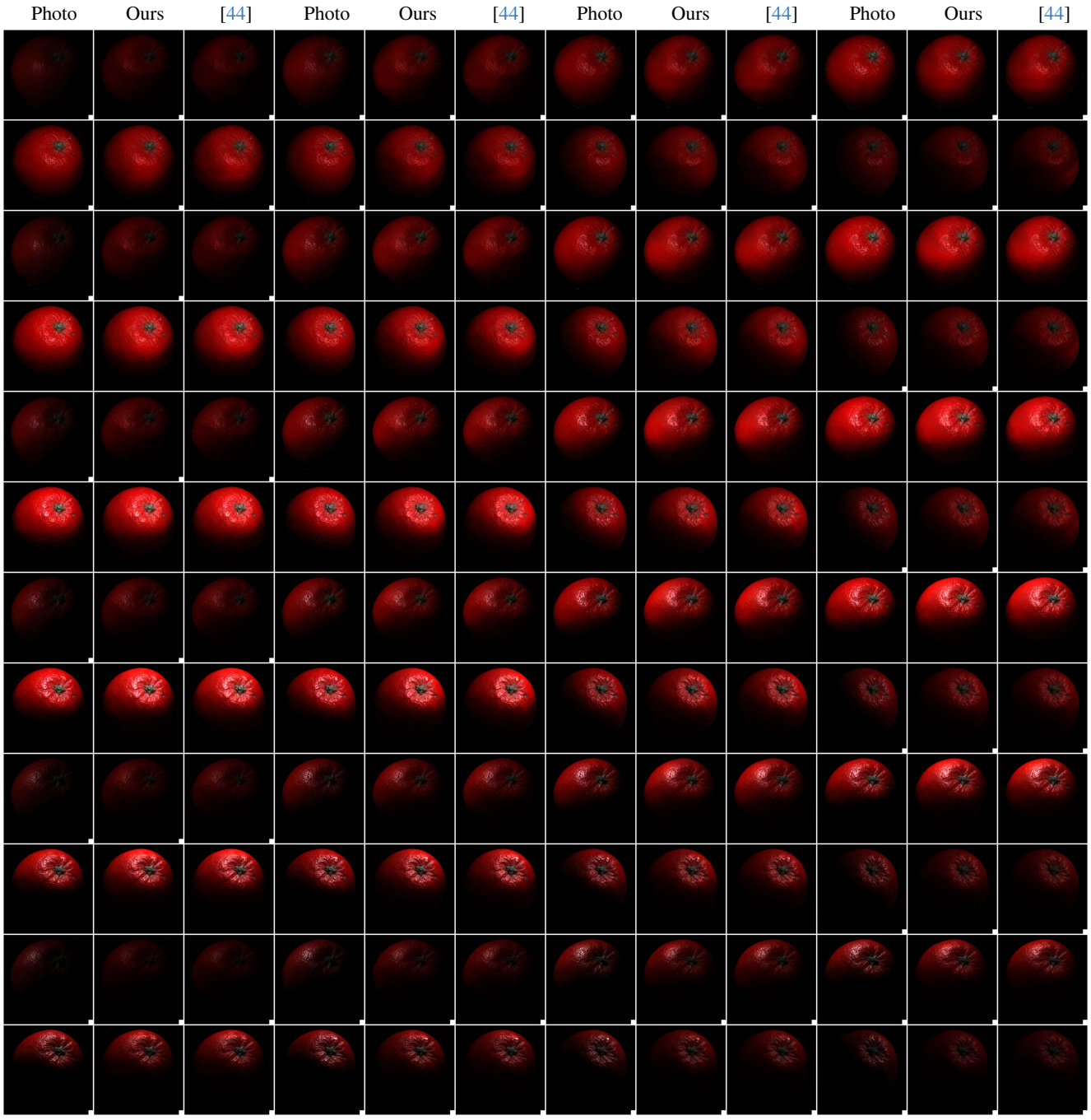

Figure S5. Relighting results of ORANGE under 48 novel lighting conditions. For each triplet of images, from the left to right, a photograph, the relighting result of our approach and [44]. The average SSIM/LPIPS/PSNR are 0.90/0.075/32.84 for our approach, and 0.90/0.063/32.89 for [44]. Note that 24 out of 48 lighting conditions are unseen ones, whose LEDs are never used in any of our acquisition patterns. We mark such cases with a missing bottom-right corner.

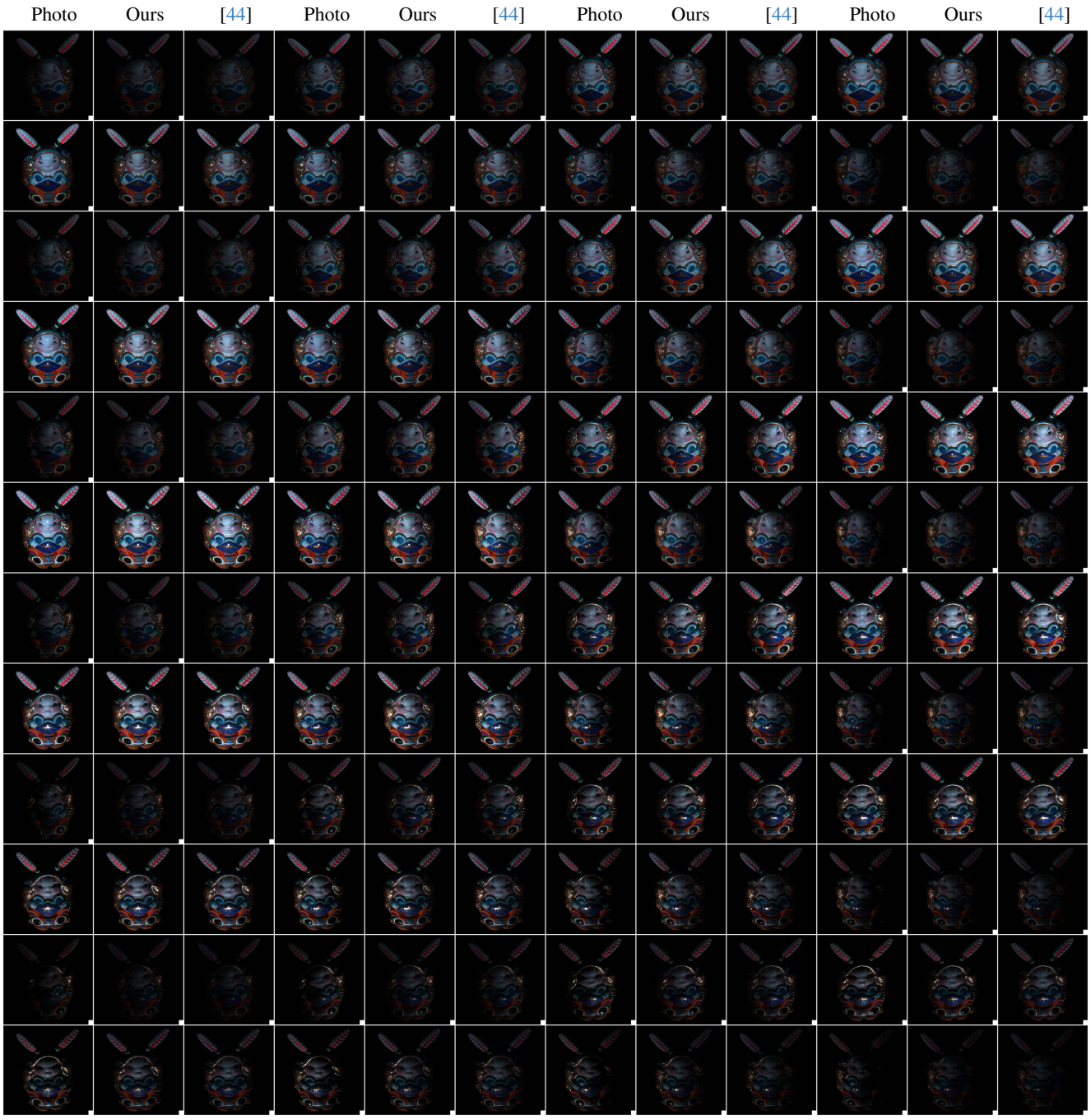

Figure S6. Relighting results of RABBIT under 48 novel lighting conditions. For each triplet of images, from the left to right, a photograph, the relighting result of our approach and [44]. The average SSIM/LPIPS/PSNR are 0.94/0.047/32.78 for our approach, and 0.93/0.051/31.41 for [44]. Note that 24 out of 48 lighting conditions are unseen ones, whose LEDs are never used in any of our acquisition patterns. We mark such cases with a missing bottom-right corner.

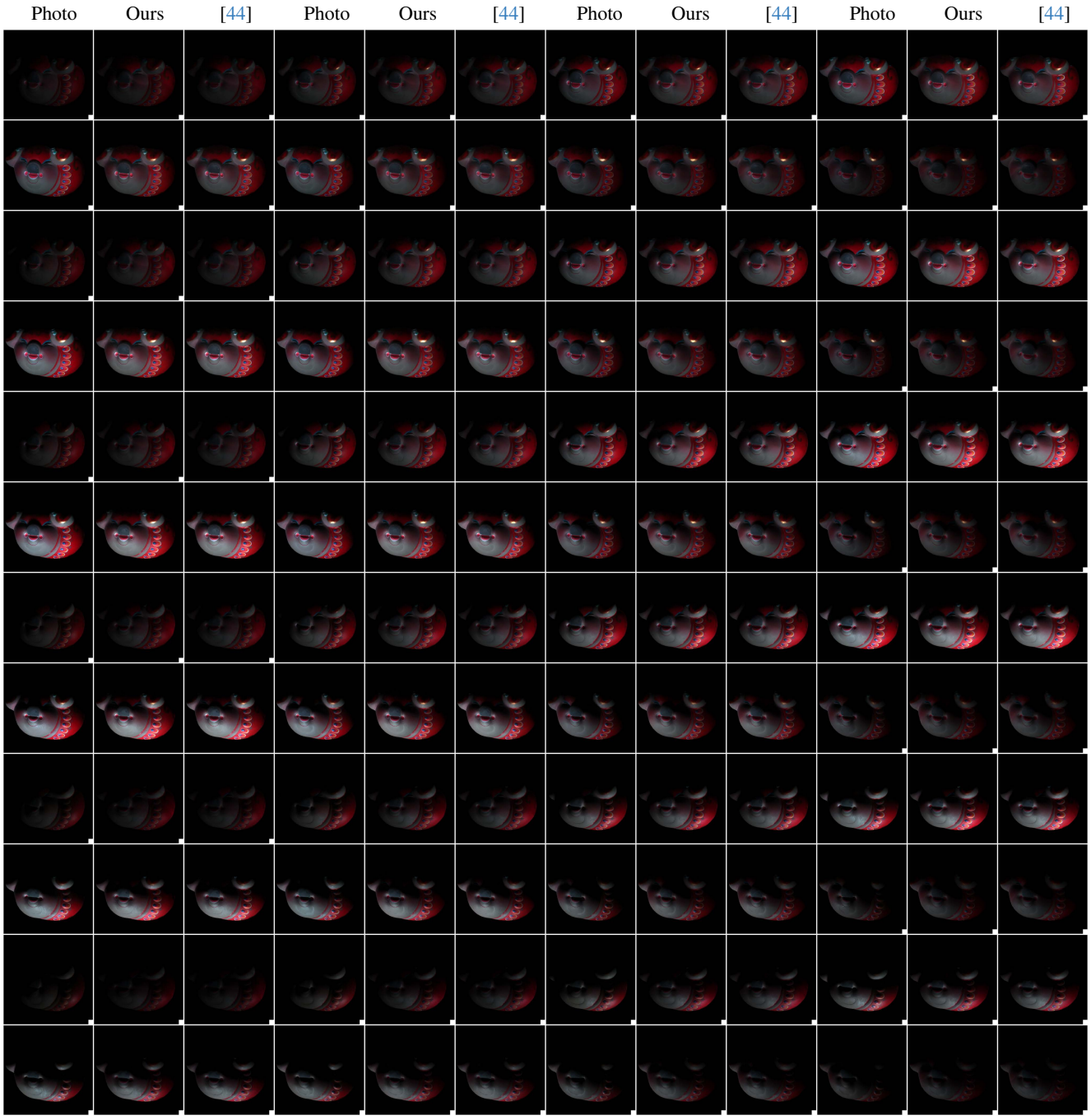

Figure S7. Relighting results of PIG under 48 novel lighting conditions. For each triplet of images, from the left to right, a photograph, the relighting result of our approach and [44]. The average SSIM/LPIPS/PSNR are 0.95/0.037/36.60 for our approach, and 0.96/0.033/36.76 for [44]. Note that 24 out of 48 lighting conditions are unseen ones, whose LEDs are never used in any of our acquisition patterns. We mark such cases with a missing bottom-right corner.