ScribbleLight: Single Image Indoor Relighting with Scribbles

Supplementary Material

Along with this supplementary Material, we provide additional visual materials (e.g., images and videos) in an websit, accessible via https://chedgekorea.github.io/ScribbleLight.

A. Implementation Details

We fine-tune a pre-trained Stable Diffusion v2 model [60] for Albedo-conditioned Image Diffusion and ScribbleLight ControlNet. To reconstruct the monochromatic shading map $\mathbf{S}_{mono}$ and the normal map $\mathbf{N}$ from the lighting feature map f, we utilize a control decoder $\mathcal{D}^C$. This control decoder $\mathcal{D}^C$ is structured similarly to the control encoder $\mathcal{E}^C$, consisting of 4 residual blocks, but with a transposed architecture. For training, we use a batch size of 16 and the AdamW optimizer with a learning rate of 1e–5. All inputs are resized to 512×512. Training each model takes approximately 48 hours on 4 A6000 GPUs. We employ the DDPM noise scheduler with 1000 diffusion steps during training. For inference, we apply the DDIM scheduler and sample only 20 steps.

B. Evaluation with Monochromatic Shading

This section is analogous to Section 4, but instead of using scribbles, we evaluate indoor scene relighting performance using the monochromatic shading map $\mathbf{S}_{mono}$.

Dataset and Baselines. As detailed in Section 4.1, we trained both LightIt* [38] and our method using the LSUN bedrooms dataset [76]. Instead of using auto-generated scribbles as the ControlNet input, however, here we utilize the monochromatic shading map instead. For our second baseline, we employed RGB↔X [80] without retraining: intrinsic components (normal, albedo, roughness, and metallicity) were extracted from the source image, and the irradiance field was derived from the target image using RGB→X. The source image was then relit by recomposing it with its intrinsic components and the target image's irradiance field through X→RGB. Since IIDiffusion [39] employs spherical Gaussians as its lighting representation, it was not feasible to perform a comparison based on scribbles. However, we extend the comparison with IIDiffusion by extracting intrinsic components from the source image and the spherical Gaussians from the target image, and recomposing the source image under the target spherical Gaussians, similar to RGB↔X.

Evaluation. We quantitatively compare the relighting quality using the monochromatic shading $\mathbf{S}_{mono}$ over the test set. The results, summarized in Table 5, show that ScribbleLight outperforms the baseline methods across all metrics.

	RMSE ↓	PSNR ↑	SSIM ↑	LPIPS ↓
IIDiffusion	0.137	18.02	0.690	0.367
LightIt*	0.227	13.17	0.390	0.447
RGB↔X	0.261	13.36	0.570	0.364
Ours	**0.132**	**18.22**	**0.697**	**0.275**

Table 5. Quantitative comparison of relighting accuracy between our ScribbleLight, RGB↔X [80], LightIt* [38], and IIDiffusion [39] with monochromatic shading map created from the target image. We compute the errors with respect to a target image.

Additionally, we present the qualitative results in website (see the 'Monochromatic Shading Map' section).

C. Comparison with IC-Light

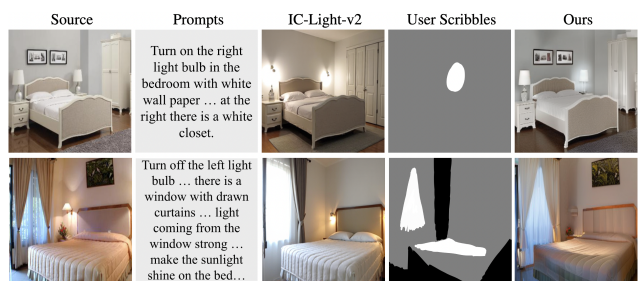

Figure 8. IC-Light uses prompt-driven relighting but lacks precise control, e.g., turn on only 1 light source (row 1) or add direct lighting on the bed (row 2). IC-Light also often changes the composition of the scene by removing or adding any objects, e.g., removes picture frames on the wall (row 1 & 2) and reshapes the curtains (row 2). In contrast, ScribbleLight provides precise lighting control while preserving the composition of the scene.

D. Visual Materials

The structure of website is as follows. As demonstrated in Figure 1 and Figure 6, the top section presents a demo video and three iterative examples showcasing how ScribbleLight iteratively refines lighting effects. The 'User Scribble' section provides additional examples of user scribble comparisons, as illustrated in Figure 4. The 'Monochromatic Shading Map' section features qualitative comparisons referenced in Appendix B. We also present the target shading utilized by each method during relighting in Figure 9. Finally, the 'Turning On/Off the Light' section includes additional examples from Figure 5.

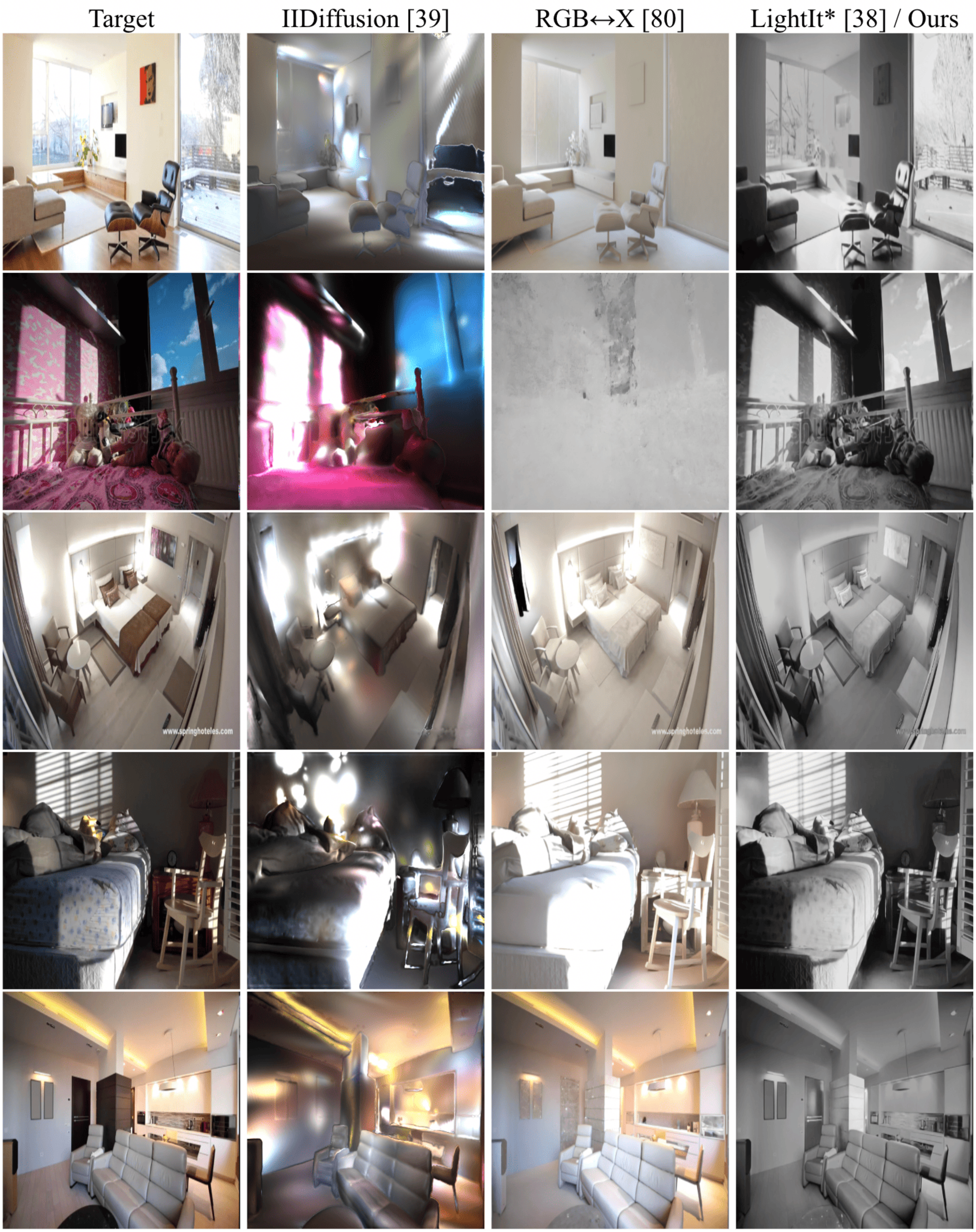

Figure 9. We demonstrate the target lighting representation used by each method—IIDiffusion [39], RGB↔X [80], LightIt* [38], and Ours—when performing relighting with a monochromatic shading map.