

# Supplementary Materials: Event-ID: Intrinsic Decomposition Using an Event Camera

Anonymous Authors

## 6 OVERVIEW

This supplemental material provides, 1). In Section 7, other details of our method in Section 3.3 of the main paper, including aspects of details of scene geometry, details of BRDF function, and implementation details; 2). In Section 8, additional experimental results, which include results on real scenes, and comparison with state-of-the-art (SOTA) methods. 3). In Section 9, one downstream application of our method: relighting.

We also provide a video that showcases our results on the test scenes and comparisons with baseline models. **We highly recommend readers to check out this video supplement.**

## 7 OTHER DETAILS OF OUR METHOD

### 7.1 Details of Scene Geometry

In this section, we provide a detailed description of SDF-based volume rendering. As detailed in [2, 9], the attributes of a 3D scene, consisting of signed distance and radiance, are parameterized by a geometry network  $f$  and a radiance network  $c$ ,

$$s = f(\mathbf{x}), \mathbf{c} = c(\mathbf{x}, \mathbf{d}), \quad (24)$$

where  $\mathbf{d}$  is view direction, the geometry network  $f$  maps a spatial position  $\mathbf{x}$  to its signed distance  $f(\mathbf{x})$  to the object. The SDF-based volume rendering can be written as,

$$\mathbf{R}(\mathbf{r}) = \sum_{i=1}^P T_i \alpha_i \mathbf{c}_i, \quad (25)$$

$$T_i = \exp\left(-\sum_{j=1}^{i-1} \alpha_j \delta_j\right), \quad (26)$$

$$\alpha_i = \max\left(\frac{\Theta_s(f(\mathbf{x}_i)) - \Theta_s(f(\mathbf{x}_{i+1}))}{\Theta_s(f(\mathbf{x}_i))}\right), \quad (27)$$

$$\Theta_s(x) = (1 + e^{-sx})^{-1}, \quad (28)$$

where  $P$  is the number of sampled points along a ray,  $\frac{1}{s}$  indicates the standard deviation of  $\Theta_s(x)$ . For detailed information, please refer to NeuS [9] and Ref-NeuS [2]

### 7.2 Details of BRDF Function

In this section, we describe detailed implementations of our normal distribution term  $D$ , Fresnel term  $F$ , and geometry term  $G$ . Like [10, 11], the normal distribution function  $D$  is approximated by the Spherical Gaussian function,

$$D(\mathbf{h}; r) = S(\mathbf{h}, \frac{1}{\pi r^2}, \mathbf{n}, \frac{2}{r^2}) = \frac{1}{\pi r^2} e^{\frac{2}{r^2}(\mathbf{h} \cdot \mathbf{n} - 1)}. \quad (29)$$

The Fresnel term is given as,

$$F(\omega_o, \mathbf{h}; \mathbf{b}, m) = F_0 + (1 - F_0)(1 - (\omega_o \cdot \mathbf{h})^5), \quad (30)$$

where  $F_0 = 0.04(1 - m) + \mathbf{b}m$ . The geometry term is approximated by the GGX function [8],

$$G(\omega_i, \omega_o, \mathbf{n}; r) = G_{\text{GGX}}(\omega_i \cdot \mathbf{n}) G_{\text{GGX}}(\omega_o \cdot \mathbf{n}), \quad (31)$$

$$G_{\text{GGX}}(z) = \frac{2z}{z + \sqrt{r^2 + (1 - r^2)z^2}}. \quad (32)$$

### 7.3 Implementation Details

We base our implementation on [10] and [9]. We set the threshold of event camera  $\Theta$  to 0.3. In the reconstruction phase, train for 200,000 epochs; in the decomposition phase, train for 30,000 epochs. During the reconstruction phase, we use  $\mathcal{L}_{\text{vol}}$  and  $\mathcal{L}_{\text{eik}}$  defined in NeuS [9]. In the decomposition phase, the weights for  $\mathcal{L}_{\text{surf}}$ ,  $\mathcal{L}_{\text{surf\_vol}}$ ,  $\mathcal{L}_{\text{base\_color}}$ , and  $\mathcal{L}_{\text{normal}}$  are 1, 1, 0.1, and 0.1, respectively.

## 8 ADDITIONAL EXPERIMENTAL RESULTS

### 8.1 Results on Real Scenes

We additionally demonstrate the experimental results on two real scenes, including a blurred specular bunny and an over-exposed diffuse bunny (shown in Figure 11). We can observe that our method is able to suppress the GS-IR [5] in both cases. As for albedo estimation, our Event-ID can produce smooth and clear albedo maps, but GS-IR estimates the unpleasing albedo maps that are related to over-exposure and blur conditions. For normal map estimation, the proposed Event-ID reconstructs a better geometry, while GS-IR fails and its estimated normal maps are highly affected by over-exposure and blur conditions. These results further validate the effectiveness of the proposed Event-ID.

### 8.2 Comparison with State-of-the-Art Methods

We additionally provide a comparison between the Neif++ [11] and the proposed Event-ID, shown in Figure 12.

Besides, we also demonstrate a further comparison between the GS-IR [5] using event-enhanced data (with the output of LL23 [6]) and the proposed Event-ID, shown in Figure 13.

In both experiments, the proposed Event-ID produces better albedo maps and normal maps, validating its effectiveness.

**Over-fitting problem on implicit representation based methods.** We should indicate that some of the existing works fail to handle such severe scenes, as they are designed to process multi-view information under good conditions (especially those using implicit representation, e.g., NVDiffRec [7] (based on implicated mesh), TensoIR [4] (based on TensoRF [1])). Figure 15 shows the under-fitting problems of those works. We find methods based on SDF [11] and 3DGS [5] representation could tolerate such challenges. The reason might be that SDF/3DGS has a strong explicit restriction on geometry and provides a well-constructed geometry reconstruction (though the 3DGS is difficult to represent surface normal [3]), which is crucial to intrinsic decomposition.

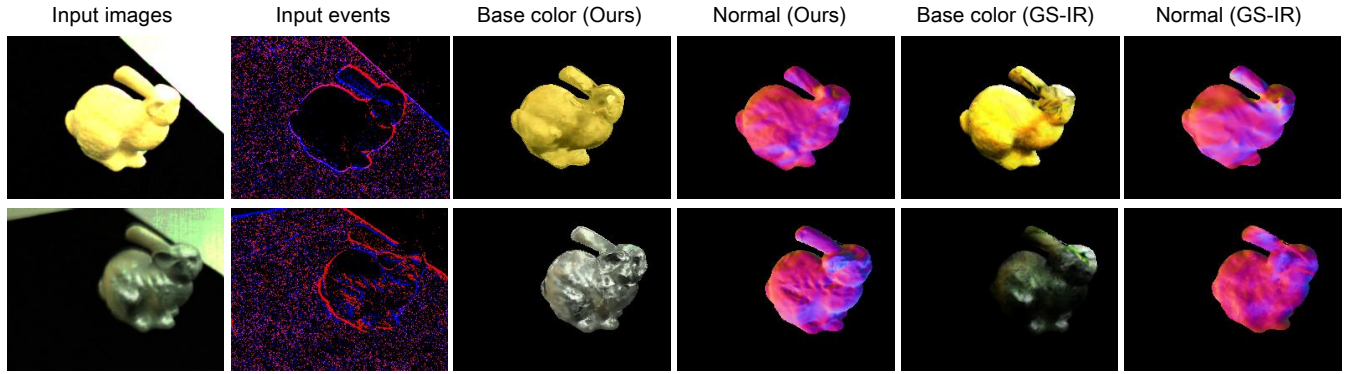


Figure 11: Qualitative comparison on real-world data. (Top) A diffuse bunny under the over-exposure condition. (Bottom) A specular bunny under the blur condition. We visualize the inputs, estimated base color, and estimated normal produced by our Event-ID and GS-IR [5].

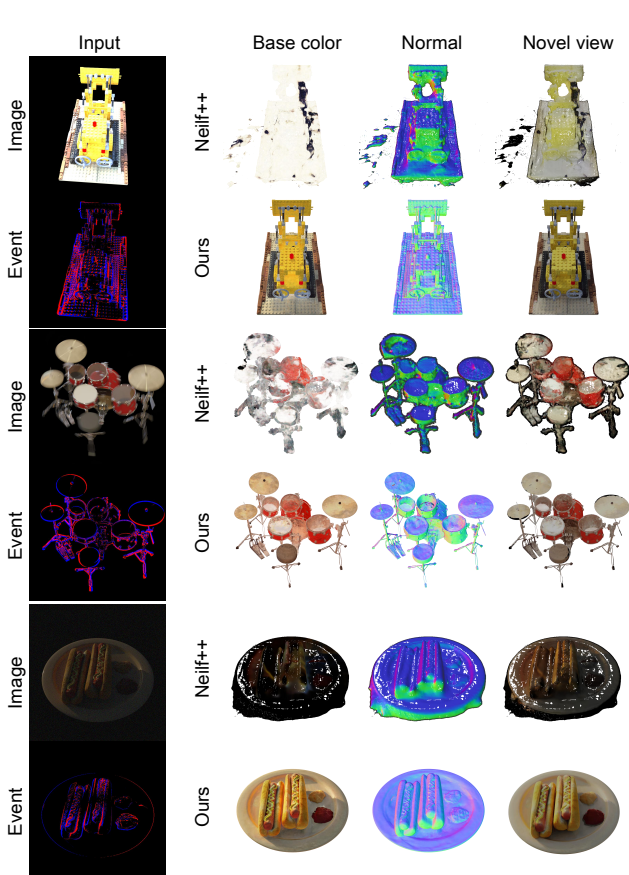


Figure 12: Qualitative comparison on our synthetic dataset. We visualize the estimated based color, normal, novel view render and input of our Event-ID and Neif++ [11]. (From top to bottom, they are overexposed, blurred, and low-light respectively)

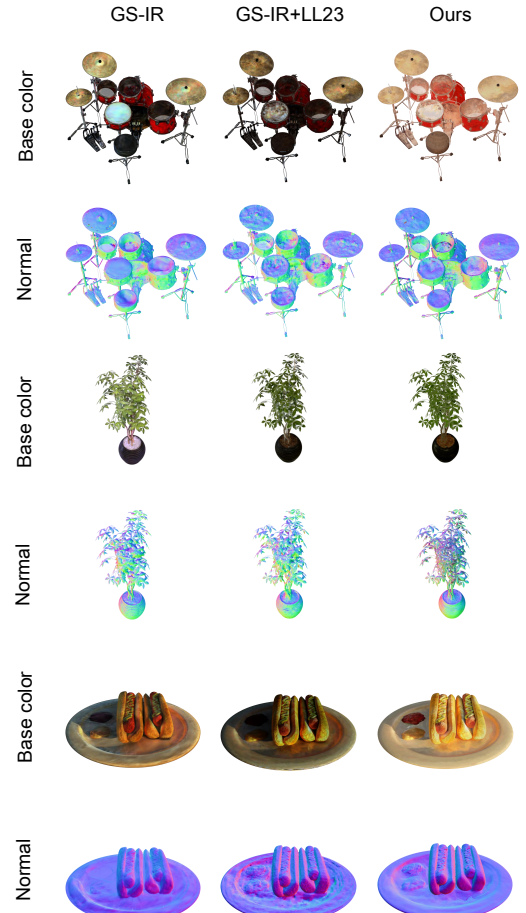


Figure 13: Qualitative comparison on our synthetic dataset. We visualize the estimated based color and normal of our Event-ID and GS-IR [5], GS-IR [5]+LL23 [6]. (From top to bottom, they are overexposed, blurred, and low-light respectively)

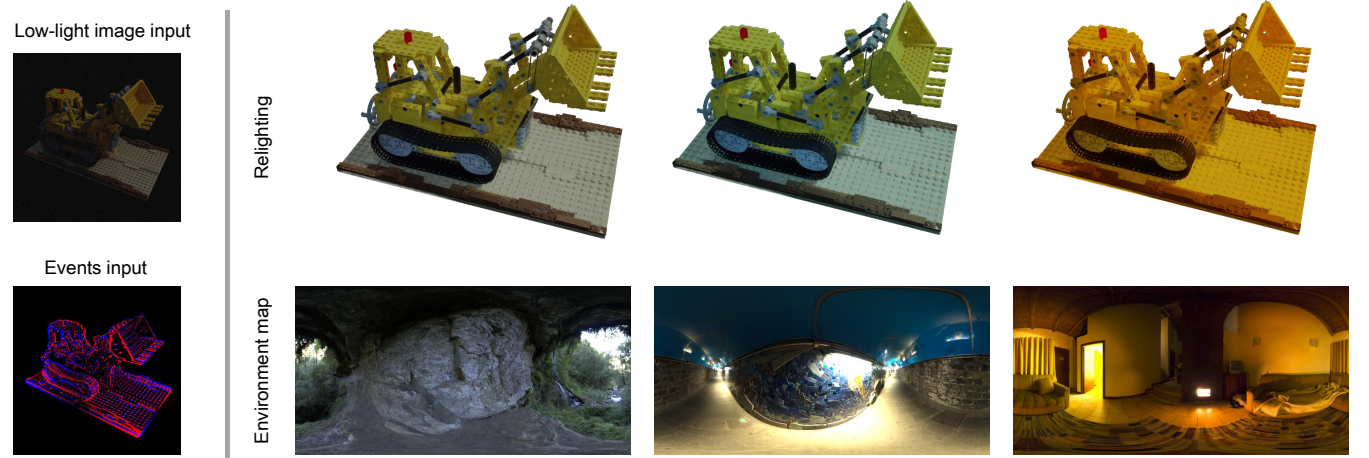


Figure 14: Relighting visualization with synthetic data. (Left) the input images and events. (Right) the results of relighting with different environment maps as inputs.

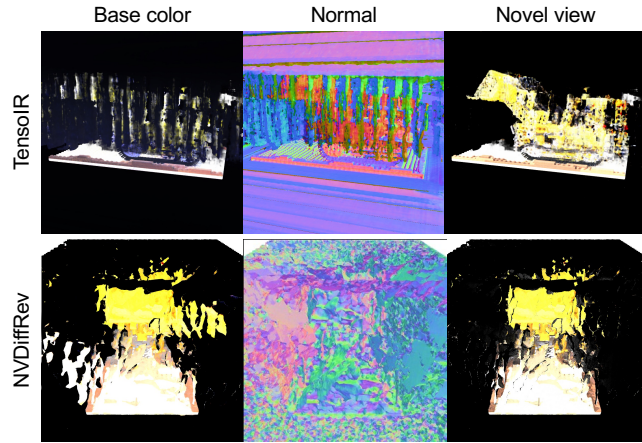


Figure 15: Under-fitting of several implicit representations on over-exposure scene.

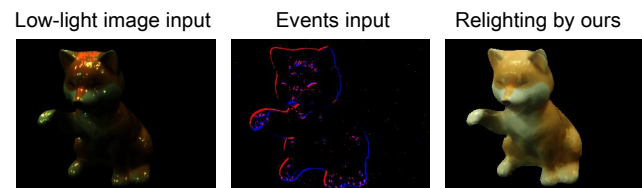


Figure 16: Relighting visualization in real-world scenarios.

## 9 APPLICATION

**Relighting.** In this section, we show the relighting result of our method in low-light environments. We replace the output of the lighting MLP  $\mathcal{L}$  directly with an environment map, which can be directly used in rendering pipelines for relighting. Figure 16 and Figure 14 display the relighting results in real and synthetic scenes,

respectively. It can be seen, our method is capable of performing Intrinsic Decomposition and relighting effectively, even in low-light conditions. Please refer to our supplementary video for convincing results.

## REFERENCES

- [1] Anpei Chen, Zexiang Xu, Andreas Geiger, Jingyi Yu, and Hao Su. 2022. Tensoror: Tensorial radiance fields. In *European Conference on Computer Vision*. Springer, 333–350.
- [2] Wenhao Ge, Tao Hu, Haoyu Zhao, Shu Liu, and Ying-Cong Chen. 2023. Ref-neus: Ambiguity-reduced neural implicit surface learning for multi-view reconstruction with reflection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 4251–4260.
- [3] Binbin Huang, Zehao Yu, Anpei Chen, Andreas Geiger, and Shenghua Gao. 2024. 2D Gaussian Splatting for Geometrically Accurate Radiance Fields. *arXiv preprint arXiv:2403.17888* (2024).
- [4] Haian Jin, Isabella Liu, Peijia Xu, Xiaoshuai Zhang, Songfang Han, Sai Bi, Xiaowei Zhou, Zexiang Xu, and Hao Su. 2023. Tensoror: Tensorial inverse rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 165–174.
- [5] Zhihao Liang, Qi Zhang, Ying Feng, Ying Shan, and Kui Jia. 2023. Gs-ir: 3d gaussian splatting for inverse rendering. *arXiv preprint arXiv:2311.16473* (2023).
- [6] Weng Fei Low and Gim Hee Lee. 2023. Robust e-NeRF: NeRF from Sparse & Noisy Events under Non-Uniform Motion. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 18335–18346.
- [7] Jacob Munkberg, Jon Hasselgren, Tianchang Shen, Jun Gao, Wenzheng Chen, Alex Evans, Thomas Müller, and Sanja Fidler. 2022. Extracting Triangular 3D Models, Materials, and Lighting From Images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 8280–8290.
- [8] Bruce Walter, Stephen R Marschner, Hongsong Li, and Kenneth E Torrance. 2007. Microfacet models for refraction through rough surfaces. In *Proceedings of the 18th Eurographics conference on Rendering Techniques*, 195–206.
- [9] Peng Wang, Lingjie Liu, Yuan Liu, Christian Theobalt, Taku Komura, and Wenping Wang. 2021. NeuS: Learning Neural Implicit Surfaces by Volume Rendering for Multi-view Reconstruction. *Advances in Neural Information Processing Systems* 34 (2021), 27171–27183.
- [10] Yao Yao, Jingyang Zhang, Jingbo Liu, Yihang Qu, Tian Fang, David McKinnon, Yanghai Tsin, and Long Quan. 2022. Neif: Neural incident light field for physically-based material estimation. In *European Conference on Computer Vision*. Springer, 700–716.
- [11] Jingyang Zhang, Yao Yao, Shiwei Li, Jingbo Liu, Tian Fang, David McKinnon, Yanghai Tsin, and Long Quan. 2023. Neif+: Inter-reflectable light fields for geometry and material estimation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 3601–3610.