

Aleth-NeRF: Low-light Condition View Synthesis with Concealing Fields

Ziteng Cui^{1,2*} Lin Gu^{3,1} Xiao Sun² Yu Qiao² Tatsuya Harada^{1,3}

¹The University of Tokyo ²Shanghai AI Laboratory ³RIKEN AIP

Abstract

Common capture low-light scenes are challenging for most computer vision techniques, including Neural Radiance Fields (NeRF). Vanilla NeRF is viewer-centred that simplifies the rendering process only as light emission from 3D locations in the viewing direction, thus failing to model the low-illumination induced darkness. Inspired by emission theory of ancient Greek that visual perception is accomplished by rays casting from eyes, we make slight modifications on vanilla NeRF to train on multiple views of low-light scene, we can thus render out the well-lit scene in an **unsupervised** manner. We introduce a surrogate concept, **Concealing Fields**, that reduce the transport of light during the volume rendering stage. Specifically, our proposed method, **Aleth-NeRF**, directly learns from the dark image to understand volumetric object representation and concealing field under priors. By simply eliminating Concealing Fields, we can render a single or multi-view well-lit image(s) and gain superior performance over other 2D low light enhancement methods. Additionally, we collect the first paired **LOw-light and normal-light Multi-view (LOM)** datasets for future research. Our code and dataset will be released soon.

1. Introduction

Neural Radiance Field (NeRF) [34] has been demonstrated to be effective in understanding 3D scenes from 2D images and generate novel views. However, similar to most computer vision algorithms such as semantic segmentation [59, 24], object detection [54, 6], and etc., NeRF often fails in sub-optimal lighting scenes captured under insufficient illumination or limited exposure time [33].

This is because vanilla NeRF is *viewer-centred* which models the amount of light emission from a location to the viewer without counting the interaction between illumination and scenes. The emitted light results from the reflection of the light in the environment on the scene. The re-

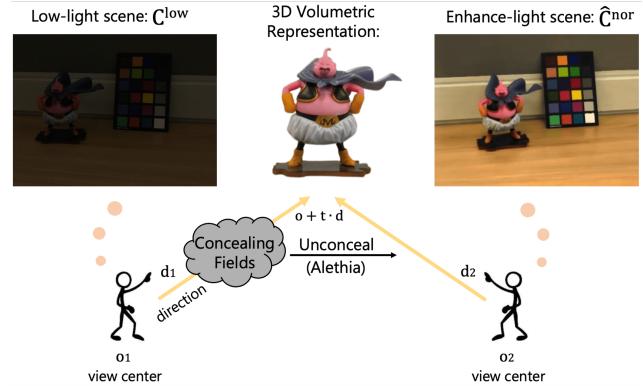


Figure 1. We assume objects are naturally visible. However, the Concealing Field attenuates the light in the viewing direction, making the left user see a low-light scene. Aleth-NeRF takes a low-light image as input and unsupervisly learns the distribution of the Concealing Field. Then, we unconceal (alethia) the Concealing field to render the enhanced image. This scene is taken from LOM dataset.

flected light will be further refracted, absorbed, and thus attenuated in the environment again [48]. As a result, the NeRF algorithm interprets a dark scene as resulting from insufficient radiation from the 3D particles representing objects in a scene. When training NeRF on dark images with a relatively high zero-mean noise signal [57], the algorithm may struggle to maintain multi-view projection consistency, leading to poor reconstruction results, shown in Fig. 2(a).

One solution is extending NeRF to *object-centred* rendering. That is, the observed lightness of the image is not due to the radiation intensity of the 3D particles representing the object, but rather to the attenuation of the light caused by other physical factors in the environment. However, this solution [48, 67] usually requires known target lighting conditions and additional parametric modeling of the lighting.

Leveraging mature 2D low-light image enhancement (LLIE) methods appear to be another solution. However, Fig. 2(b) shows that directly enhancing the dark images by 2D enhancement methods does not guarantee accurate NeRF estimation, since the independent and inconsistent enhancement of 2D images in multi-views may lead to the

*This work was done during his internship at Shanghai Artificial Intelligence Laboratory. [✉] means corresponding author.

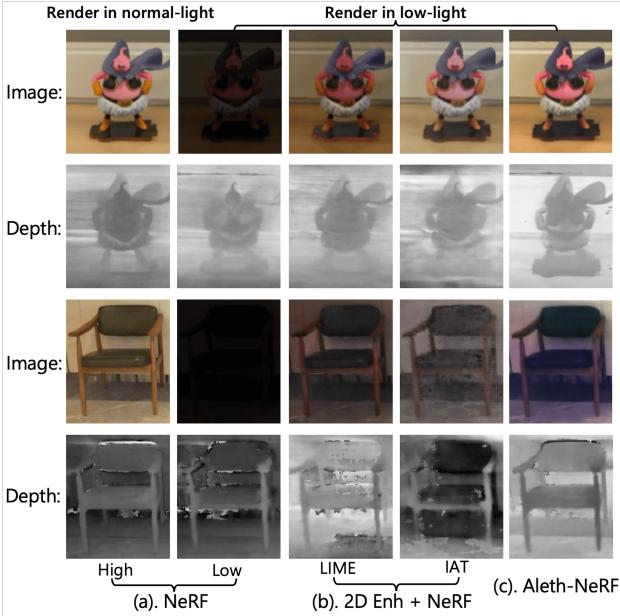


Figure 2. (a). NeRF rendering results in normal-light scene and low-light scene. (b). NeRF rendering on enhanced scene by 2D image enhancement methods LIME [12] and IAT [5]. (c) Aleth-NeRF rendering results in low-light scene.

destruction of 3D geometric consistency.

To explore the NeRF application in these commonly captured but often precluded low-light scenes, we aim to propose a framework that directly supervises on low-light sRGB images. The rendering process in NeRF is similar to the emission theory held by ancient Greek. Emission theory ignores the incident light but postulates visual rays emitted from the eye travel in straight lines and interacts with objects to form the visual perception. Therefore, the darkness of an entity is solely caused by the particles between the object and the eye. In other words, all objects are visible by default unless they are concealed. Inspired by this worldview, we assume a simple but NeRF-friendly model that it is the concealing fields in viewing direction makes the viewer see a low-light scene C^{low} shown in Fig.1. Aletheia ($\alpha\lambda\gamma\theta\epsilon\alpha$), normally translated as "unconcealedness", "disclosure" or "revealing" [14], removes the concealing fields to let right side viewer catch sight of normal-light scene C^{nor} .

Here, we slightly modify the NeRF and propose our Aleth-NeRF for low-light conditions, which takes multi-view dark images as input to train the model and learn the volumetric representation jointly with Concealing Fields. As shown in Fig.1, we model this Concealing Field between object and viewer to naturally extend the transmittance function in NeRF [34]. In training stage, we aim to estimate not only the Radiance Fields representing the consistent objects, but also the Concealing Fields explaining the image transition from normal-light to low-light. In the testing stage, we could directly render the normal-lit scenes by

simply eliminating Concealing Field. In addition, We collect the first paired low-light and normal-light multi-view dataset, **LOM**, to facilitate the understanding of 3D scenes under low-light conditions.

Our contribution could be summarized as follow:

- We propose **Aleth-NeRF**, by our knowledge, the first NeRF trains on dark multi-view sRGB images for unsupervised enhancement. Inspired by ancient Greek philosophy, we naturally extend the transmittance function in vanilla NeRF by modeling Concealing Fields between scene and viewer to interpret low light condition, endowing robustness to darkness with minimal modification on vanilla NeRF.
- We contribute the first low-light & normal-light paired multi-view real-world dataset **LOM**, along with comparisons with various 2D image enhancement methods, experiments show that our **Aleth-NeRF** achieves satisfactory performance in both enhancement quality and multi-view consistency.

2. Related Work

2.1. Low-light Image Enhancement

Low-light image enhancement (LLIE) is a classical image processing task aiming to recover an image taken under inadequate illumination to its counterpart taken under normal illumination. Traditional LLIE methods usually make pre-defined illumination assumptions and use hand-crafted features for image restoration. For example, the Retinex-based methods [26, 40, 12] assume that a low-light image is multiplied by illumination and reflection. Histogram Equalization (HE) based methods [10, 49, 41] perform LLIE by spreading out the most frequent intensity values.

With the fast growth of deep learning in recent years, deep neural networks (DNNs) based methods have become the mainstream solutions these years. A series of CNN or Transformer-based methods have been developed in this area [28, 56, 35, 55, 58, 4, 70, 72, 69, 25, 47, 5, 53, 52], which trained on paired low-light and normal-light images for supervision and gain satisfactory results.

Beyond supervised methods, unsupervised low-light image enhancement is a more general but challenging setup, which does not require ground truth normal-light images for training. Some works like [18, 19] leverage the statistical illumination distribution in the target image domain for unsupervised training. Other works complete this task with only low-light information [11, 30, 21, 65, 73, 37]. Like Zero-DCE [11] propose to train the network only with low-light images and regularize the network with modeled curves and some additional non-reference constraints (*i.e.* color constancy [3]). SCI [30] completes this task with a self-calibrated module and unsupervised loss.

However, current LLIE methods almost build on 2D image space operations, which often fail to exploit the 3D geometry of the scene and make it hard to deal with multi-view inputs. The proposed Aleth-NeRF belongs to the unsupervised family but is empowered with a much stronger and more reliable understanding of 3D space. We follow the volume rendering formulation in NeRF and trace this problem back to the phase of image rendering from 3D radiant particles, where additional light concealing fields are introduced and learned unsupervisedly for the LLIE problem, meanwhile maintaining the multi-view consistency.

2.2. Novel View synthesis with NeRF

NeRF [34] is proposed for novel view synthesis from a collection of posed input images. It gains substantial attention even though the quality of the rendered image is yet to be comparable to state-of-the-art GAN [23] or diffusion-based [15, 8] models. The unique advantage of NeRF models exists in preserving the 3D geometry consistency thanks to its physical volume rendering scheme.

In addition to general efforts to speed up and improve NeRF training [2, 44, 27, 63, 17, 42, 7, 36] or to use NeRF for scene understanding applications [71, 38, 62, 61, 9]. Many of the latter works focus on improving NeRF’s performance under various degradation conditions, such as blurry [29], noisy [39], reflection [13], glossy surfaces [50], or use NeRF to complete low-level tasks in 3D space, like super-resolution [51, 1] and HDR reconstruction [60, 20].

Another line of research extends NeRF for lightness editing in 3D space. Some work, like NeRF-W [31], focuses on rendering NeRF with uncontrolled in-the-wild images, other relighting works [48, 43, 67] rely on known illumination conditions and introduce additional physical elements (*i.e.* normal, light, albedo, etc.), along with complex parametric modeling of these elements. Meanwhile, these methods are not specifically designed for low-light scene enhancement under extreme low-light conditions.

Among these, RAW-NeRF [33] is most close to our work, which proposes to render a NeRF model in nighttime RAW domain with high-dynamic range, and then post-process the rendered scene with image signal processor (ISP) [22], RAW-NeRF has shown a preliminary ability to enhance the scene light but requires sufficient RAW data for training. We present the first work to render NeRF with low-light sRGB inputs and injection unsupervised LLIE into 3D space by an effective concealing fields manner.

3. Method

At first we briefly review core concepts of vanilla NeRF on the radiance field, volume rendering, and image reconstruction loss in Section. 3.1. Then we introduce the proposed Concealing Fields in Section. 3.2 and show how it is integrated into the NeRF framework to mimic the process

of low-light scene generation. Finally, in Section. 3.3, we show how to learn Concealing Field effectively in an unsupervised manner.

3.1. Neural Radiance Field Revisited

In the NeRF representation [34], a Radiance Field is defined as the density σ and RGB color value c of a 3D location \mathbf{x} under a 2D viewing direction \mathbf{d} . The density σ , on the one hand, represents the radiation capacity of the particle itself at \mathbf{x} , and on the other hand, controls how much radiance is absorbed when other lights pass through \mathbf{x} .

When rendering an image with a neural radiance field, a camera ray $\mathbf{r}(t) = \mathbf{o} + t \cdot \mathbf{d}$ ($\mathbf{r} \in \mathbf{R}$), casting from the given camera position \mathbf{o} towards direction \mathbf{d} , is used to accumulate all the radiance along the ray to render its corresponding pixel value $\mathbf{C}(\mathbf{r})$. This process is commonly implemented with the volume rendering function [32]. Formally,

$$\mathbf{C}(\mathbf{r}) = \int_{t_n}^{t_f} T(\mathbf{r}(t))\sigma(\mathbf{r}(t))c(\mathbf{r}(t), \mathbf{d})dt, \quad (1)$$

where

$$T(\mathbf{r}(t)) = \exp\left(-\int_{t_n}^t \sigma(\mathbf{r}(s))ds\right), \quad (2)$$

is known as the *accumulated transmittance*. It denotes the radiance decay rate of the particle at $\mathbf{r}(t)$ when it is occluded by particles closer to the camera (at $\mathbf{r}(s)$, $s < t$). The integrals are computed by a discrete approximation over sampled 3D points along the ray \mathbf{r} .

$$\mathbf{C}(\mathbf{r}) = \sum_{i=1}^N T(\mathbf{r}(i))(1 - \exp(-\sigma(\mathbf{r}(i)) \cdot \delta)) \cdot c(\mathbf{r}(i), \mathbf{d}) \quad (3)$$

$$T(\mathbf{r}(i)) = \exp\left(-\sum_{j=1}^{i-1} \sigma(\mathbf{r}(j)) \cdot \delta\right) \quad (4)$$

where δ is a constant distance value between adjacent sample points under uniform sampling.

Neural Network Implementation. NeRF learns two multilayer perceptron (MLP) networks ($F_{\Theta_\sigma}, F_{\Theta_c}$) to map the 3D location $\mathbf{r}(i)$ and 2D viewing direction \mathbf{d} to its density σ and colour c . Specifically,

$$F_{\Theta_\sigma}(\mathbf{r}(i)) \rightarrow \sigma(\mathbf{r}(i)), \mathbf{h} \quad (5)$$

$$F_{\Theta_c}(\mathbf{h}, \mathbf{d}) \rightarrow c(\mathbf{r}(i), \mathbf{d}) \quad (6)$$

where \mathbf{h} is a hidden feature vector, Θ_σ and Θ_c are learnable network parameters. Note that $c(\mathbf{r}(i), \mathbf{d})$ and $\sigma(\mathbf{r}(i))$ are further activated by Sigmoid and ReLU functions so that their value ranges are $[0, 1]$ and $[0, \infty)$, respectively.

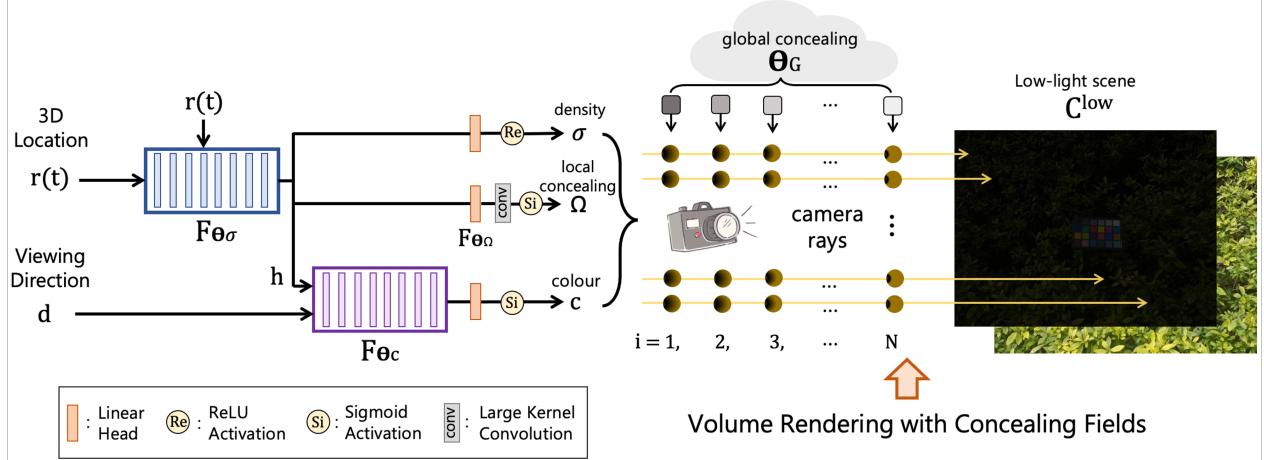


Figure 3. Overview of the Aleth-NeRF architecture. Local Ω and Global Θ_G Concealing Fields are additionally learned and integrated into the NeRF framework. We use a modified volume rendering function to render low-light scene taking the Concealing Fields into account.

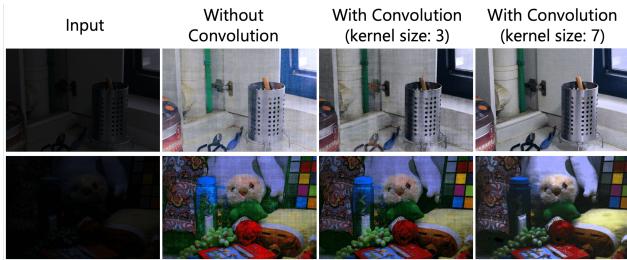


Figure 4. Ablation analyze on the LOL dataset [56], larger convolution size to generate local concealing field Ω would further improve enhanced scene $\hat{\mathbf{C}}^{nor}$'s smoothness.

Given ground truth rendered image \mathbf{C} , the network is optimized by minimizing the image reconstruction loss between predicted image $\hat{\mathbf{C}}$ and ground truth image \mathbf{C} :

$$\mathcal{L}_{nerf} = \sum_{\mathbf{r}}^R \|\hat{\mathbf{C}}(\mathbf{r}) - \mathbf{C}(\mathbf{r})\|_2. \quad (7)$$

We refer techniques such as positional encoding and hierarchical volume sampling to original paper [34] for more details.

3.2. Aleth-NeRF with Concealing Field Assumption

Given low-light scene $\{\mathbf{C}^{low}(\mathbf{r}), \mathbf{r} \in \mathbf{R}\}$ taken under poor illumination, the goal of LLIE is to recover its normal-lit correspondence $\{\mathbf{C}^{nor}(\mathbf{r}), \mathbf{r} \in \mathbf{R}\}$. The key idea of our model is that we assume \mathbf{C}^{low} and \mathbf{C}^{nor} are rendered under the same Radiance Field condition (namely, they share the same underlying densities σ and colors c at all 3D locations in the scene) but with or without the proposed Concealing Field assumption.

Global and Local Concealing Fields We design two types of Concealing Fields, namely the local Concealing Field Ω and global Concealing Field Θ_G for low-light scene generation. Specifically, Ω controls the light concealing at voxel level while Θ_G controls at scene level.

The local Concealing Field, denoted as $\Omega(\mathbf{r}(i))$, defines an extra light concealing capacity of a particle at 3D location $\mathbf{r}(i)$. As it shown in Fig. 3, $\Omega(\mathbf{r}(i))$ is learned for each 3D location, for implementation we add a linear layer head F_{Θ_Ω} upon the first MLP network F_{Θ_σ} , an additional convolution layer has been added after F_{Θ_Ω} to generate the local concealing $\Omega(\mathbf{r}(i))$, convolution process could build spatial relations between pixels and let Concealing Field contain more light information rather than structure information [64], make rendering results smooth (see Fig. 4).

$$F_{\Theta_\Omega}(F_{\Theta_\sigma}(\mathbf{r}(i))) \rightarrow \Omega(\mathbf{r}(i)) \quad (8)$$

The global-wise Concealing Field, denoted as $\Theta_G(i)$, is defined as a set of learnable parameters corresponding to the camera distance i for all camera rays in \mathbf{R} . $\Theta_G(i)$ is kept the same in a scene rendering and irrelevant to pixel level.

In the training stage, when we render and reconstruct a given low-light image \mathbf{C}^{low} with the reconstruction loss \mathcal{L}_{nerf} computed between predicted $\hat{\mathbf{C}}^{low}$ and ground truth \mathbf{C}^{low} images, the *accumulated transmittance* T in Eq. 4 is further modulated with the additional local Concealing Field Ω and global Concealing Field Θ_G , to mimic the process of light suppression in a normal scene:

$$T^{low}(\mathbf{r}(i)) = \exp \left(- \sum_{j=1}^{i-1} \sigma(\mathbf{r}(j)) \cdot \delta \right) \cdot \prod_{j=1}^{i-1} \Omega(\mathbf{r}(j)) \Theta_G(j) \quad (9)$$

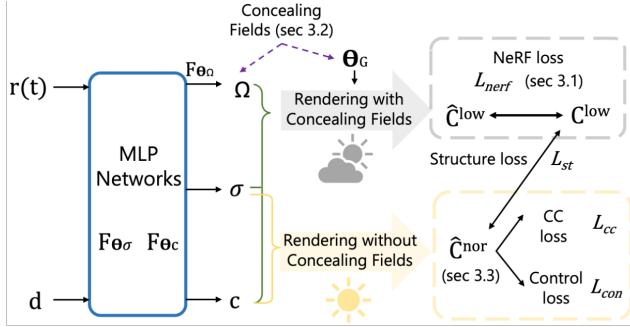


Figure 5. Aleth-NeRF uses reconstruction loss (\mathcal{L}_{nerf}) on the low-light scene \mathbf{C}^{low} , meanwhile additional constraints (control loss \mathcal{L}_{con} , structure loss \mathcal{L}_{st} and color constancy loss \mathcal{L}_{cc}) are added to regularize the predicted normal-light scene $\hat{\mathbf{C}}^{nor}$.

The testing stage performs aletheia or unconcealedness to remove Concealing Fields Ω and Θ_G to render the predicted normal-light image $\hat{\mathbf{C}}^{nor}$ directly. As a result, we transform the unsupervised LLIE problem into unsupervised learning of Concealing Fields in the low-light scene. By adding or removing the Concealing Fields, Aleth-NeRF can render images under low-light or normal-light conditions of the same scene easily.

3.3. Effective Priors for Unsupervised Training

In this section, we introduce the priors and constraints that promote unsupervised learning of Concealing Fields. Training strategy overview is shown in Fig. 5.

Value Range Prior. Value range of concealing fields would determine the enhancement scene $\hat{\mathbf{C}}^{nor}$'s lightness, generally the concealing fields' range should be less than 1, for local concealing field Ω , we add an sigmoid activation after the convolution process (see Fig. 3), the activated Ω would lie in the range of (0, 1). Meanwhile, the global concealing field $\Theta_G(i)$ is a set of learnable parameters in the network, the initial value of $\Theta_G(i)$ is all set to 0.3 along the camera ray.

To control the degree of image enhancement, we add a control loss \mathcal{L}_{con} on the local concealing field Ω , where we introduce a hyper-parameter: η , representing the degree of Aleth-NeRF's concealing ability. To calculate \mathcal{L}_{con} , we apply an average pooling (stride 64) to local concealing field Ω , then minimize loss \mathcal{L}_{con} between the pooled Ω 's mean and the concealing degree η :

$$\mathcal{L}_{con} = \|\text{avgpool}(\Omega(\mathbf{r}(i))) - \eta\|^2, \quad (10)$$

Here, the concealing degree η should be set larger than 0. In our work, we set η to 0.05 in extremely dark low-light conditions and to 0.1 in conditions not so dark.

Structure Similarity Prior. To maintain the structure information of the predicted normal light images $\hat{\mathbf{C}}^{nor}$ consistent with the original low light image \mathbf{C}^{low} , we addition-

ally add a structure loss between them, to maintain the detail information and increase the contrast. Specifically, for each rendered pixel $\hat{\mathbf{C}}^{nor}(\mathbf{r})$, we keep a structure consistency with neighbored pixel $\hat{\mathbf{C}}^{nor}(\mathbf{r}-1)$ and $\hat{\mathbf{C}}^{nor}(\mathbf{r}+1)$ along the ray as follows:

$$\begin{aligned} \mathcal{L}_{st} = & \sum_{k \in \{+1, -1\}} (\hat{\mathbf{C}}^{nor}(\mathbf{r}) - \hat{\mathbf{C}}^{nor}(\mathbf{r}+k)) - \\ & \frac{0.5}{\eta} (\mathbf{C}^{low}(\mathbf{r}) - \mathbf{C}^{low}(\mathbf{r}+k)), \end{aligned} \quad (11)$$

here $\frac{0.5}{\eta}$ determines generated scene $\hat{\mathbf{C}}^{nor}$'s contrast degree, which is inversely proportional to conceal degree η . More ablation analyses on conceal degree and contrast degree are in our supplementary.

Color Constancy Prior. Pixels taken under low-light conditions would lose some color information, and direct take off the concealing fields would easily cause color imbalance. To regularize the color of the predicted normal-light images $\hat{\mathbf{C}}^{nor}$, we add an extra color constancy loss \mathcal{L}_{cc} on $\hat{\mathbf{C}}^{nor}$. Here we assume that $\hat{\mathbf{C}}^{nor}$ obey the gray-world assumption [3, 11, 19], as follows:

$$\mathcal{L}_{cc} = \sum_{p,q} (\hat{\mathbf{C}}^{nor}(\mathbf{r})^p - \hat{\mathbf{C}}^{nor}(\mathbf{r})^q)^2, \quad (12)$$

where $(p, q) \in \{(R, G), (G, B), (B, R)\}$ represents any pair of color channels.

Above all, the total loss function \mathcal{L}_{total} of Aleth-NeRF include 4 parts, low-light scene rendering loss \mathcal{L}_{nerf} and additional constrain losses \mathcal{L}_{con} , \mathcal{L}_{st} , \mathcal{L}_{cc} , the total loss used for training can be represented as:

$$\mathcal{L}_{total} = \mathcal{L}_{nerf} + \lambda_1 \cdot \mathcal{L}_{con} + \lambda_2 \cdot \mathcal{L}_{st} + \lambda_3 \cdot \mathcal{L}_{cc}, \quad (13)$$

where λ_1 , λ_2 and λ_3 are three non-negative parameters to balance total loss weights, which we set to $1e^{-4}$, $1e^{-3}$ and $1e^{-4}$ respectively.

Discussion Although our method was originally designed for multi-view conditions, surprisingly, it achieves the same excellent low-light enhancement result for single-view images (see experiments in Sec. 4.1). Despite the ambiguity of depth estimation and the lack of synthesis capability for novel views, we believe that the success of single-view low-light enhancement greatly benefits from the effective priors we present in this section.

4. Experiments

In this section, we first benchmark on single image low-light enhancement dataset LOL [56] in Sec. 4.1. Then we would introduce our collected **LOM** dataset with paired



Figure 6. Example of enhancement results on LOL [56] evaluation set, (a) is input image, then are the enhancements results by (b). Zero-DCE [11], (c). SCI [30], (d). IAT [5] and (e) our **Aleth-NeRF**, (f) is the ground truth image.

multi-view low-light and normal-light images in Sec. 4.2. Multi-view experimental results and analyses are shown in Sec. 4.3.

Our framework is base on NeRF-Factory [16] toolbox. We adopt the Adam optimizer with the initial learning rate set to $5e^{-4}$. Besides cosine learning rate decay strategy are set at every 2500 iters. For the single image experiments on LOL [56] dataset, the camera position \mathbf{o} and viewing direction \mathbf{d} are fixed, and batch size is set to 8192 for 20000 iters training. For multi-view experiments on **LOM** dataset, batch size is set to 4096 for 62500 iters training.

4.1. Generation Quality Assessment

We first conduct the experiments on benchmark single image low-light enhancement dataset **LOL** [56]. **LOL** consists of 500 paired low-light images \mathbf{C}^{low} and normal-light images \mathbf{C}^{nor} , 485 pairs are split into a train set, and the other 15 pairs are split into evaluation set.

During training stage, we only learn the 15 low-light images \mathbf{C}^{low} in the evaluation set. Without either the reference normal-light images \mathbf{C}^{nor} or the other 485 low-light images in the train set. The quantitative comparison with various low-light enhancement methods [10, 56, 28, 68, 11, 12, 18, 21, 30, 73] is shown in Table. 1, we report three image quality metrics: PSNR, SSIM and LPIPS [66]. Table. 1 shows that our methods gain satisfactory results among the unsupervised methods, ensuring the **generation quality** of Aleth-NeRF. We show 4 examples in Fig. 6, including other enhancement results by current SOTA meth-

Table 1. Quantitative comparison on LOL [56] dataset, here * denotes the zero-shot methods.

Method		PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow
supervised	RetiNexNet [56]	16.67	0.562	0.474
	MBLLEN [28]	17.96	0.756	0.386
	KIND [68]	20.86	0.790	0.448
unsupervised	HE* [10]	14.54	0.376	0.466
	LIME* [12]	14.01	0.513	0.391
	Zero-DCE [11]	14.63	0.558	0.385
	Zero-Restore* [21]	14.03	0.460	0.451
	RRDNet* [73]	11.39	0.468	0.451
	En-GAN [18]	12.23	0.596	0.498
	SCI [30]	15.65	0.510	0.419
	Aleth-NeRF*	17.14	0.653	0.445

ods Zero-DCE [11], SCI [30] and IAT [5]. Fig. 6 shows that our Aleth-NeRF could gain high-quality and vivid results compared with various 2D enhancement methods.

4.2. Multi-view Paired Dataset in Real-world

In this section, we introduce our collected paired low-light and normal-light multi-view dataset, namely **LOM** dataset. To the best of our knowledge, this is the first dataset containing paired multi-view low-light and normal-light images. Previous work [33] also includes some low-light scenes. However, their dataset concentrates on RAW denoising rather than sRGB low-light enhancement. Besides [33] does not include normal-light sRGB counterpart,

Table 2. LOM [56] dataset results, we evaluate PSNR \uparrow (higher the better), SSIM \uparrow (higher the better) and LPIPS \downarrow (lower the better).

method	“buu”	“chair”	“sofa”	“bike”	“shrub”	mean
	PSNR/ SSIM/ LPIPS	PSNR/ SSIM/ LPIPS	PSNR/ SSIM/ LPIPS	PSNR/ SSIM/ LPIPS	PSNR/ SSIM/ LPIPS	PSNR/ SSIM/ LPIPS
NeRF [34]	7.53/ 0.313/ 0.414	6.01/ 0.142/ 0.599	6.26/ 0.206/ 0.566	6.32/ 0.067/ 0.625	8.00/ 0.027/ 0.684	5.24/ 0.151/ 0.578
NeRF + HE [10]	14.90/ 0.678/ 0.578	15.11/ 0.615/ 0.658	17.28/ 0.713/ 0.624	14.26/ 0.547/ 0.582	12.40/ 0.384/ 0.647	14.79/ 0.587/ 0.618
NeRF + LIME [12]	13.67/ 0.745/ 0.375	11.11/ 0.637/ 0.600	12.28/ 0.723/ 0.562	10.82/ 0.504/ 0.557	14.12/ 0.360/ 0.540	12.40/ 0.593/ 0.526
NeRF + RetiNexNet [56]	16.18/ 0.754/ 0.385	16.74/ 0.731/ 0.531	15.90/ 0.826/ 0.512	17.79/ 0.654/ 0.552	15.24/ 0.283/ 0.577	16.37/ 0.657/ 0.495
NeRF + SCI [30]	13.66/ 0.769/ 0.433	18.42/ 0.736/ 0.545	19.82/ 0.813/ 0.569	12.61/ 0.560/ 0.550	16.87/ 0.411/ 0.555	16.27/ 0.658/ 0.530
NeRF + IAT [5]	14.00/ 0.652/ 0.421	19.53/ 0.806/ 0.562	10.58/ 0.532/ 0.668	16.30/ 0.660/ 0.532	10.03/ 0.149/ 0.577	14.09/ 0.560/ 0.552
HE [10] + NeRF	15.54/ 0.755/ 0.473	15.46/ 0.756/ 0.492	17.89/ 0.801/ 0.457	15.47/ 0.699/ 0.455	14.97/ 0.414/ 0.511	15.87/ 0.685/ 0.470
LIME [12] + NeRF	13.81/ 0.783/ 0.253	11.20/ 0.694/ 0.498	12.02/ 0.747/ 0.420	11.35/ 0.586/ 0.448	14.11/ 0.426/ 0.473	12.50/ 0.647/ 0.418
RetiNexNet [56] + NeRF	16.16/ 0.777/ 0.339	16.82/ 0.773/ 0.438	16.87/ 0.806/ 0.548	18.00/ 0.717 / 0.448	14.65/ 0.269/ 0.513	16.50/ 0.668/ 0.457
SCI [30] + NeRF	13.62/ 0.821/ 0.309	11.75/ 0.756/ 0.526	10.10/ 0.750/ 0.505	19.10 / 0.644/ 0.458	18.13/ 0.510/ 0.469	14.54/ 0.696/ 0.453
IAT [5] + NeRF	14.33/ 0.697/ 0.311	18.68/ 0.789/ 0.563	17.77/ 0.811/ 0.519	13.68/ 0.621/ 0.501	13.84/ 0.322/ 0.536	15.66/ 0.648/ 0.486
Aleth-NeRF	18.93/ 0.825/ 0.251	19.71/ 0.812/ 0.476	19.98/ 0.851/ 0.411	15.53/ 0.691/ 0.444	18.31/ 0.511/ 0.506	18.49/ 0.738/ 0.417

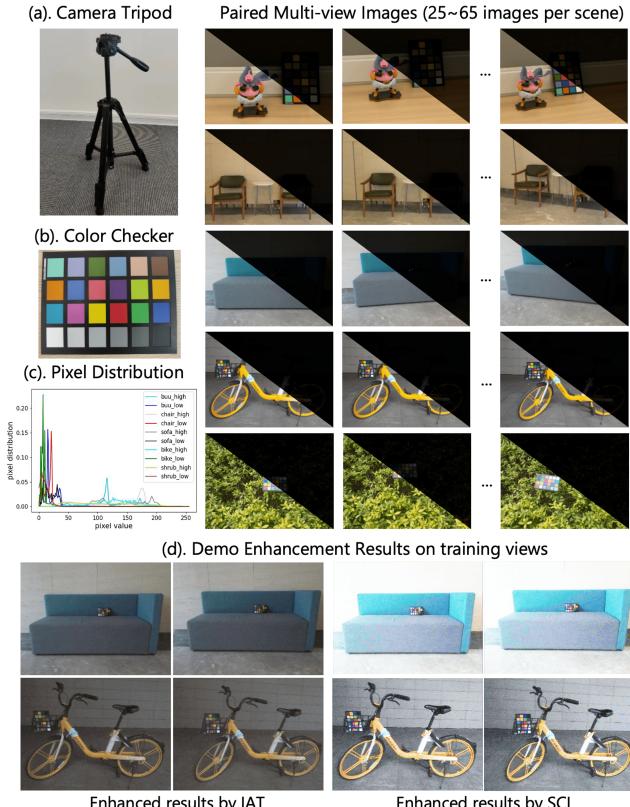


Figure 7. Collection detail of the LOM dataset.

making it hard to evaluate the model performance of multi-view low-light enhancement in real-world.

LOM dataset has 5 real-world scenes (“**buu**”, “**chair**”, “**sofa**”, “**bike**”, “**shrub**”). Each scene includes 25 ~ 65 low-light and normal-light pairs. We collect the scenes with DJI Osmo Action 3 camera, then generate pair low-light and normal-light images by adjusting exposure time and ISO while other configurations of the camera are fixed. We also

take a camera tripod to prevent camera shake (Fig. 7(a)), we capture multi-view images by moving and rotating the tripod. Additionally, for each scene, we add a DSLR color checker (Fig. 7(b)) to help us determine the color and better evaluate generated images. Images are collected with resolution 3000×4000 . We down-sample the original resolution with ratio 8 to 375×500 for convenience, and generate the ground truth view and angle information by adopting COLMAP [45, 46] on the normal-light scenes. For dataset split, in each scene, we choose 3 ~ 5 images as the testing set, 1 image as the validation set, and other images to be the training set. The Y-channel pixel distribution of each scene has been shown in Fig. 7(c). Exemplary low-light scene enhancement results by SOTA 2D enhancement methods IAT [5] and SCI [30] are shown in Fig. 7(d).

4.3. Multi-view Enhancement Results

In this section, we show multi-view rendering results on low-light scenes in LOM dataset. We design multiple comparison experiments to evaluate the generation quality and multi-view consistency. We first make the comparison with the vanilla NeRF [34], which has the same setting as Aleth-NeRF that only trains with low-light scene images \mathbf{C}^{low} . Then we compare with five low-light enhancement methods: histogram equalization (HE) [10], LIME [12], RetiNexNet [56], IAT [5] and SCI [30], here HE and LIME are two traditional enhancement method, IAT and SCI are two very recent SOTA network-based 2D enhancement methods. As shown in Table 2, we design two settings for comparison, (1). Rendering low-light scenes by NeRF and then using 2D enhancement methods to post-process these low-light novel views, name as “NeRF + *”. (2). Using 2D enhancement methods to pre-process training set and rendering NeRF on these enhanced views, name as “* + NeRF”. All the comparison experiments take the same training epochs, batch size, learning rate, and strategies to

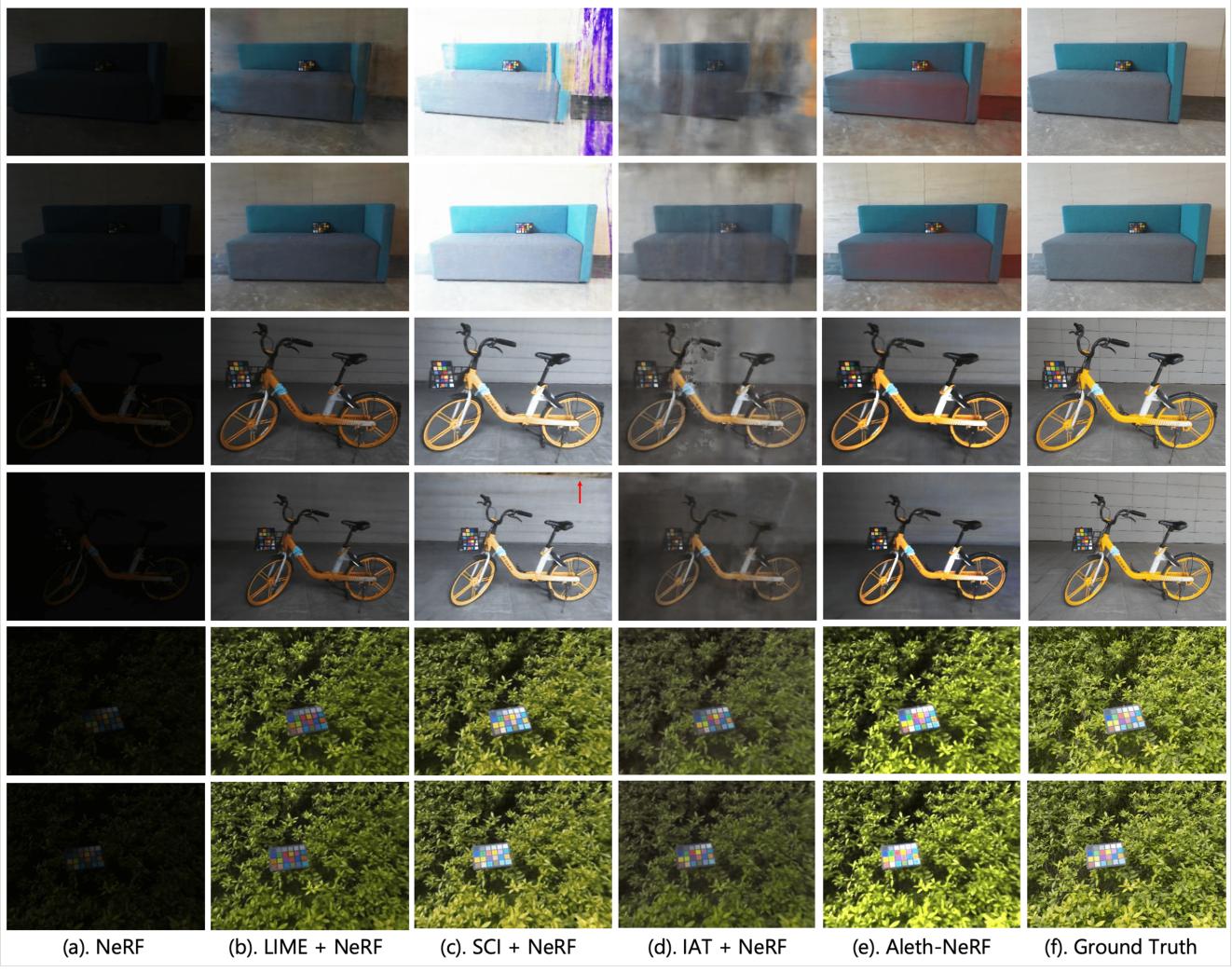


Figure 8. Multi-view rendering results on **LOM** dataset’s “*sofa*”, “*bike*” and “*shrub*” scenes.

ensure fairness.

The comparison results are shown in Table. 2. All rendering results would compare with LOM’s normal-lit view counterpart. We report to three image quality metrics: SSIM, PSNR and LPIPS [66]. The results of each scene are then averaged to generate the *mean* results (last column in Table. 2). From the results, we could find that “NeRF + ***” series methods almost perform worse than “*** + NeRF” series methods. This is probably due to the poor image quality NeRF produces in dark scenes, making enhancement methods easily fail on the generated dark scenes by NeRF. Meanwhile, for the “*** + NeRF” series methods, although the enhancement quality performs well on the training set (see Fig. 7(d)), 2D enhancement methods lack multi-view consistency and often fail on novel view generation (see Fig. 8), make it easier to generate artifacts and noise. Overall, our Aleth-NeRF is an end-to-end method and gain the best performance on most scenes and the *mean* results, the gener-

ated novel views both maintain multi-view consistency and image generation quality. Please refer to our supplementary for more visualization results and ablation analyze.

5. Conclusion and Discussion

We propose a novel unsupervised method to handle multi-view synthesis in low-light condition, which directly take low-light scene as input and render out normal-light scene. Inspired by the wisdom of the ancient Greeks, we introduce a concept: Concealing Fields. Experiments demonstrate our superior performance in both image quality and 3D multi-view consistency.

One limitation is that Aleth-NeRF should be specifically trained for each scene, which is the same as original NeRF [34]. Besides, Aleth-NeRF may fail in scenes with non-uniform lighting conditions or shadow conditions. We’ll solve these problems in the future.

6. Acknowledgments

This work was partially supported by JST Moonshot R&D Grant Number JPMJPS2011, CREST Grant Number JPMJCR2015 and Basic Research Grant (Super AI) of Institute for AI and Beyond of the University of Tokyo. Also this work is partially supported by the National Key R&D Program of China(NO.2022ZD0160100),and in part by Shanghai Committee of Science and Technology (Grant No. 21DZ1100100).

References

- [1] Yuval Bahat, Yuxuan Zhang, Hendrik Sommerhoff, Andreas Kolb, and Felix Heide. Neural volume super-resolution, 2022.
- [2] Jonathan T. Barron, Ben Mildenhall, Matthew Tancik, Peter Hedman, Ricardo Martin-Brualla, and Pratul P. Srinivasan. Mip-nerf: A multiscale representation for anti-aliasing neural radiance fields, 2021.
- [3] G. Buchsbaum. A spatial processor model for object colour perception. *Journal of the Franklin Institute*, 310(1):1–26, 1980.
- [4] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. Learning to see in the dark. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.
- [5] Ziteng Cui, Kunchang Li, Lin Gu, Shenghan Su, Peng Gao, ZhengKai Jiang, Yu Qiao, and Tatsuya Harada. You only need 90k parameters to adapt light: a light weight transformer for image enhancement and exposure correction. In *33rd British Machine Vision Conference 2022, BMVC 2022, London, UK, November 21-24, 2022*. BMVA Press, 2022.
- [6] Ziteng Cui, Guo-Jun Qi, Lin Gu, Shaodi You, Zenghui Zhang, and Tatsuya Harada. Multitask aet with orthogonal tangent regularity for dark object detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 2553–2562, October 2021.
- [7] Kangle Deng, Andrew Liu, Jun-Yan Zhu, and Deva Ramanan. Depth-supervised NeRF: Fewer views and faster training for free. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2022.
- [8] Prafulla Dhariwal and Alexander Nichol. Diffusion models beat gans on image synthesis. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, and J. Wortman Vaughan, editors, *Advances in Neural Information Processing Systems*, volume 34, pages 8780–8794. Curran Associates, Inc., 2021.
- [9] Zhiwen Fan, Peihao Wang, Yifan Jiang, Xinyu Gong, Dejia Xu, and Zhangyang Wang. Nerf-sos: Any-view self-supervised object segmentation on complex scenes, 2022.
- [10] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing (3rd Edition)*. Prentice-Hall, Inc., USA, 2006.
- [11] Chunle Guo Guo, Chongyi Li, Jichang Guo, Chen Change Loy, Junhui Hou, Sam Kwong, and Runmin Cong. Zero-reference deep curve estimation for low-light image enhancement. In *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, pages 1780–1789, June 2020.
- [12] Xiaojie Guo, Yu Li, and Haibin Ling. Lime: Low-light image enhancement via illumination map estimation. *IEEE Transactions on Image Processing*, 26(2):982–993, 2017.
- [13] Yuan-Chen Guo, Di Kang, Linchao Bao, Yu He, and Song-Hai Zhang. Nerfren: Neural radiance fields with reflections. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 18409–18418, June 2022.
- [14] M. Heidegger, J. Stambaugh, and D.J. Schmidt. *Being and Time*. SUNY Series in Contemporary Co. State University of New York Press, 2010.
- [15] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 6840–6851. Curran Associates, Inc., 2020.
- [16] <https://github.com/kakaobrain/nerf factory>. Github, 2022.
- [17] Ajay Jain, Matthew Tancik, and Pieter Abbeel. Putting nerf on a diet: Semantically consistent few-shot view synthesis. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 5885–5894, October 2021.
- [18] Yifan Jiang, Xinyu Gong, Ding Liu, Yu Cheng, Chen Fang, Xiaohui Shen, Jianchao Yang, Pan Zhou, and Zhangyang Wang. Enlightengan: Deep light enhancement without paired supervision. *IEEE Transactions on Image Processing*, 30:2340–2349, 2021.
- [19] Yeying Jin, Wenhan Yang, and Robby T Tan. Unsupervised night image enhancement: When layer decomposition meets light-effects suppression. *arXiv preprint arXiv:2207.10564*, 2022.
- [20] Kim Jun-Seong, Kim Yu-Ji, Moon Ye-Bin, and Tae-Hyun Oh. Hdr-plenoxels: Self-calibrating high dynamic range radiance fields. In *ECCV*, 2022.
- [21] Aupendu Kar, Sobhan Kanti Dhara, Debasish Sen, and Prabir Kumar Biswas. Zero-shot single image restoration through controlled perturbation of koschmieder’s model. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 16205–16215, June 2021.
- [22] Hakkı Can Karaimer and Michael S. Brown. A software platform for manipulating the camera imaging pipeline. In *European Conference on Computer Vision (ECCV)*, 2016.
- [23] Tero Karras, Miika Aittala, Samuli Laine, Erik Hätkönen, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Alias-free generative adversarial networks. In *Proc. NeurIPS*, 2021.
- [24] Abdulrahman Kerim, Felipe Chamone, Washington LS Ramos, Leandro Soriano Marcolino, Erickson R Nascimento, and Richard Jiang. Semantic segmentation under adverse conditions: A weather and nighttime-aware synthetic data-based approach. In *33rd British Machine Vision Conference 2022, BMVC 2022*, 2022.
- [25] Mohit Lamba and Kaushik Mitra. Restoring extremely dark images in real time. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3487–3497, 2021.

- [26] Edwin H. Land. An alternative technique for the computation of the designator in the retinex theory of color vision. *Proceedings of the National Academy of Sciences of the United States of America*, 1986.
- [27] David B. Lindell, Julien N. P. Martel, and Gordon Wetzstein. AutoInt: Automatic integration for fast neural volume rendering. In *Proc. CVPR*, 2021.
- [28] Feifan Lv, Feng Lu, Jianhua Wu, and Chongsoon Lim. Mbllen: Low-light image/video enhancement using cnns. In *British Machine Vision Conference*, 2018.
- [29] Li Ma, Xiaoyu Li, Jing Liao, Qi Zhang, Xuan Wang, Jue Wang, and Pedro V. Sander. Deblur-nerf: Neural radiance fields from blurry images. *arXiv preprint arXiv:2111.14292*, 2021.
- [30] Long Ma, Tengyu Ma, Risheng Liu, Xin Fan, and Zhongxuan Luo. Toward fast, flexible, and robust low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5637–5646, 2022.
- [31] Ricardo Martin-Brualla, Noha Radwan, Mehdi S. M. Sajjadi, Jonathan T. Barron, Alexey Dosovitskiy, and Daniel Duckworth. NeRF in the Wild: Neural Radiance Fields for Unconstrained Photo Collections. In *CVPR*, 2021.
- [32] N. Max. Optical models for direct volume rendering. *IEEE Transactions on Visualization and Computer Graphics*, 1(2):99–108, 1995.
- [33] Ben Mildenhall, Peter Hedman, Ricardo Martin-Brualla, Pratul P. Srinivasan, and Jonathan T. Barron. NeRF in the dark: High dynamic range view synthesis from noisy raw images. *arXiv*, 2021.
- [34] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. In *ECCV*, 2020.
- [35] Sean Moran, Pierre Marza, Steven McDonagh, Sarah Parisot, and Gregory Slabaugh. DeepLpf: Deep local parametric filters for image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020.
- [36] Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Trans. Graph.*, 41(4):102:1–102:15, July 2022.
- [37] Hue Nguyen, Diep Tran, Khoi Nguyen, and Rang Nguyen. Psenet: Progressive self-enhancement network for unsupervised extreme-light image enhancement. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2023.
- [38] Atsuhiro Noguchi, Xiao Sun, Stephen Lin, and Tatsuya Harada. Neural articulated radiance field. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 5762–5772, October 2021.
- [39] Naama Pearl, Tali Treibitz, and Simon Korman. Nan: Noise-aware nerfs for burst-denoising. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 12672–12681, June 2022.
- [40] Zia-ur Rahman, Daniel J Jobson, and Glenn A Woodell. Multi-scale retinex for color image enhancement. In *Proceedings of 3rd IEEE international conference on image processing*, volume 3, pages 1003–1006. IEEE, 1996.
- [41] Ali M Reza. Realization of the contrast limited adaptive histogram equalization (clahe) for real-time image enhancement. *Journal of VLSI signal processing systems for signal, image and video technology*, 38(1):35–44, 2004.
- [42] Barbara Roessle, Jonathan T. Barron, Ben Mildenhall, Pratul P. Srinivasan, and Matthias Nießner. Dense depth priors for neural radiance fields from sparse input views. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2022.
- [43] Viktor Rudnev, Mohamed Elgharib, William Smith, Lingjie Liu, Vladislav Golyanik, and Christian Theobalt. Nerf for outdoor scene relighting. In *European Conference on Computer Vision (ECCV)*, 2022.
- [44] Sara Fridovich-Keil and Alex Yu, Matthew Tancik, Qinhong Chen, Benjamin Recht, and Angjoo Kanazawa. Plenoxels: Radiance fields without neural networks. In *CVPR*, 2022.
- [45] Johannes Lutz Schönberger and Jan-Michael Frahm. Structure-from-motion revisited. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
- [46] Johannes Lutz Schönberger, Enliang Zheng, Marc Pollefeys, and Jan-Michael Frahm. Pixelwise view selection for unstructured multi-view stereo. In *European Conference on Computer Vision (ECCV)*, 2016.
- [47] Yuda Song, Hui Qian, and Xin Du. Starenhancer: Learning real-time and style-aware image enhancement. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 4126–4135, 2021.
- [48] Pratul P Srinivasan, Boyang Deng, Xiuming Zhang, Matthew Tancik, Ben Mildenhall, and Jonathan T Barron. Nerv: Neural reflectance and visibility fields for relighting and view synthesis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7495–7504, 2021.
- [49] Carlo Tomasi and Roberto Manduchi. Bilateral filtering for gray and color images. In *Sixth international conference on computer vision (IEEE Cat. No. 98CH36271)*, 1998.
- [50] Dor Verbin, Peter Hedman, Ben Mildenhall, Todd Zickler, Jonathan T. Barron, and Pratul P. Srinivasan. Ref-NeRF: Structured view-dependent appearance for neural radiance fields. *CVPR*, 2022.
- [51] Chen Wang, Xian Wu, Yuan-Chen Guo, Song-Hai Zhang, Yu-Wing Tai, and Shi-Min Hu. Nerf-sr: High-quality neural radiance fields using supersampling. *arXiv*, 2021.
- [52] Haoyuan Wang, Ke Xu, and Rynson W.H. Lau. Local color distributions prior for image enhancement. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2022.
- [53] Tao Wang, Kaihao Zhang, Tianrun Shen, Wenhan Luo, Bjorn Stenger, and Tong Lu. Ultra-high-definition low-light image enhancement: A benchmark and transformer-based method. *arXiv preprint arXiv:2212.11548*, 2022.
- [54] Wenjing Wang, Wenhan Yang, and Jiaying Liu. Hla-face: Joint high-low adaptation for low light face detection. In

- Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2021.
- [55] Yufei Wang, Renjie Wan, Wenhan Yang, Haoliang Li, Lap-Pui Chau, and Alex C Kot. Low-light image enhancement with normalizing flow. *arXiv preprint arXiv:2109.05923*, 2021.
 - [56] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep retinex decomposition for low-light enhancement. In *British Machine Vision Conference*, 2018.
 - [57] Kaixuan Wei, Ying Fu, Jiaolong Yang, and Hua Huang. A physics-based noise formation model for extreme low-light raw denoising. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2020.
 - [58] Wenhui Wu, Jian Weng, Pingping Zhang, Xu Wang, Wenhan Yang, and Jianmin Jiang. Uretinex-net: Retinex-based deep unfolding network for low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5901–5910, June 2022.
 - [59] Xinyi Wu, Zhenyao Wu, Hao Guo, Lili Ju, and Song Wang. Dannet: A one-stage domain adaptation network for unsupervised nighttime semantic segmentation. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2021.
 - [60] Huang Xin, Zhang Qi, Feng Ying, Li Hongdong, Wang Xuan, and Wang Qing. Hdr-nerf: High dynamic range neural radiance fields. *arXiv preprint arXiv:2111.14451*, November 2021.
 - [61] Dejia Xu, Yifan Jiang, Peihao Wang, Zhiwen Fan, Humphrey Shi, and Zhangyang Wang. Sinnerf: Training neural radiance fields on complex scenes from a single image. 2022.
 - [62] Lin Yen-Chen, Pete Florence, Jonathan T. Barron, Alberto Rodriguez, Phillip Isola, and Tsung-Yi Lin. iNeRF: Inverting neural radiance fields for pose estimation. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2021.
 - [63] Alex Yu, Rui long Li, Matthew Tancik, Hao Li, Ren Ng, and Angjoo Kanazawa. PlenOctrees for real-time rendering of neural radiance fields. In *ICCV*, 2021.
 - [64] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Cycleisp: Real image restoration via improved data synthesis. In *CVPR*, 2020.
 - [65] Rongkai Zhang, Lanqing Guo, Siyu Huang, and Bihan Wen. Rellie: Deep reinforcement learning for customized low-light image enhancement. *MM '21*, page 2429–2437, New York, NY, USA, 2021. Association for Computing Machinery.
 - [66] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *CVPR*, 2018.
 - [67] Xiuming Zhang, Pratul P. Srinivasan, Boyang Deng, Paul Debevec, William T. Freeman, and Jonathan T. Barron. Nerfactor: Neural factorization of shape and reflectance under an unknown illumination. *ACM Trans. Graph.*, 40(6), dec 2021.
 - [68] Yonghua Zhang, Jiawan Zhang, and Xiaojie Guo. Kindling the darkness: A practical low-light image enhancer. In *Proceedings of the 27th ACM international conference on multimedia*, 2019.
 - [69] Zhao Zhang, Huan Zheng, Richang Hong, Mingliang Xu, Shuicheng Yan, and Meng Wang. Deep color consistent network for low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1899–1908, 2022.
 - [70] Chuanjun Zheng, Daming Shi, and Wentian Shi. Adaptive unfolding total variation network for low-light image enhancement. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 4439–4448, October 2021.
 - [71] Shuaifeng Zhi, Tristan Laidlow, Stefan Leutenegger, and Andrew J. Davison. In-place scene labelling and understanding with implicit scene representation. In *ICCV*, 2021.
 - [72] Shangchen Zhou, Chongyi Li, and Chen Change Loy. Lednet: Joint low-light enhancement and deblurring in the dark. In *ECCV*, 2022.
 - [73] Anqi Zhu, Lin Zhang, Ying Shen, Yong Ma, Shengjie Zhao, and Yicong Zhou. Zero-shot restoration of underexposed images via robust retinex decomposition. In *2020 IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6, 2020.