SeaThru-NeRF: Neural Radiance Fields in Scattering Media

Deborah Levy[1], Amit Peleg[1], Naama Pearl[1], Dan Rosenbaum[1], Derya Akkaynak[1,2],
Simon Korman[1], Tali Treibitz[1]
[1]University of Haifa , [2]The Inter-University Institute for Marine Sciences in Eilat
sea-thru-nerf.github.io

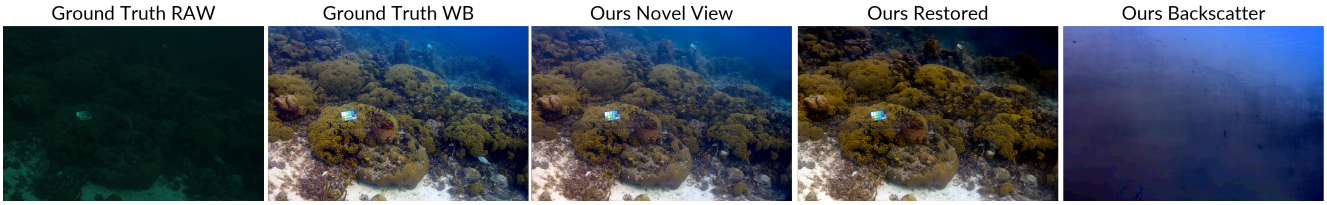

Figure 1. NeRFs have not yet tackled scenes in which the medium strongly influences the appearances of objects, as in the case of underwater imagery. By incorporating a scattering image formation model into the NeRF rendering equations, we are able to separate the scene into 'clean' and backscatter components. Consequently, we can render photorealistic novel-views with or without the participating medium, in the latter case recovering colors as if the image was taken in clear air. Results on the **Curaçao** scene: A RAW image (left) is brightened and white balanced (WB) for visualization, showing more detail, while areas further from the camera (top-right corner) are occluded and attenuated by severe backscatter - which is effectively removed in our restored image. Please zoom-in to observe the details.

Abstract

Research on neural radiance fields (NeRFs) for novel view generation is exploding with new models and extensions. However, a question that remains unanswered is what happens in underwater or foggy scenes where the medium strongly influences the appearance of objects. Thus far, NeRF and its variants have ignored these cases. However, since the NeRF framework is based on volumetric rendering, it has inherent capability to account for the medium's effects, once modeled appropriately. We develop a new rendering model for NeRFs in scattering media, which is based on the SeaThru image formation model, and suggest a suitable architecture for learning both scene information and medium parameters. We demonstrate the strength of our method using simulated and real-world scenes, correctly rendering novel photorealistic views underwater. Even more excitingly, we can render clear views of these scenes, removing the medium between the camera and the scene and reconstructing the appearance and depth of far objects, which are severely occluded by the medium. Our code and unique datasets are available on the project's website.

1. Introduction

The pioneering work of Mildenhall et al. [25] on Neural Radiance Fields (NeRFs) has tremendously advanced the field of Neural Rendering, due to its flexibility and unprecedented quality of synthesized images. Yet, the formulations of the original NeRF [25] and its followup variants assume that images were acquired in clear air, i.e., in a medium that does not scatter or absorb light in a significant manner and that the acquired image is composed solely of the object radiance. The NeRF formulation is based on volumetric rendering equations that take into account sampled points along 3D rays. Assuming a clear air environment, an implicit assumption, which is often enforced explicitly with dedicated loss components [5], is that a single opaque (high density) object is encountered per ray, with zero density between the camera and the object.

In stark contrast to clear air case, when the medium is *absorbing* and / or *scattering* (e.g., haze, fog, smog, and all aquatic habitats), the volume rendering equation has a true volumetric meaning, as the entire volume, and not only the object, contributes to image intensity. As the NeRF model estimates color and density at every point of a scene, it lends itself perfectly to general volumetric rendering, given that the appropriate rendering model is used. Here, we bridge this gap with *SeaThru-NeRF*, a framework that incorporates a rendering model that takes into account scattering media.

This is achieved by assigning separate color and density parameters to the object (scene) and the medium, within the NeRF framework. Our approach adopts the SeaThru underwater image formation model [1, 3] to account for scattering media. SeaThru is a generalization of the standard wavelength-independent attenuation (e.g., fog) image formation model, where *two* different wideband coefficients are used to represent the medium, which is more accurate when attenuation is wavelength-dependent (as in all wa-

ter bodies and under some atmospheric conditions). In our model, the medium parameters are separate per color channel, and are learned functions of the viewing angles, enforcing them to be constant only along 3D rays in the scene.

Attempting to optimize existing NeRFs on scenes with scattering medium results in cloud-like objects floating in space, while our formulation enables the network to learn the correct representation of the entire 3D volume, that consists of both the scene and the medium. Our experiments demonstrate that *SeaThru-NeRF* produces state-of-the-art photorealistic **novel view synthesis** on simulated and challenging real-world scenes (see Fig. 1) with scattering media, that include complex geometries and appearances. In addition, it enables:

1. **Color restoration** of the scenes as if they were not imaged through a medium, as our modeling allows full separation of object appearance from the medium effects.
2. **Estimation of 3D scene structure** which surpasses that of structure-from-motion (SFM) or current NeRFs, especially in far areas of bad visibility, as we jointly reconstruct and reason for the geometry and medium.
3. **Estimation of wideband medium parameters**, which are informative properties of the captured environment, and potentially allowing simulation under different conditions.

2. Related Work

Neural Radiance Fields (NeRFs): The original work on NeRFs [25] has paved the way to a large capacity of followup work, with rapid and significant progress in many related aspects. For brevity, we focus in the following only on works closely related to ours, and refer the reader to a comprehensive review of the field [43] prior to the introduction of NeRFs, and to [44] for the most recent.

NeRFs have been recently shown to be extremely powerful in multi-image settings that involve computational imaging tasks. These include HDR [50], de-blurring [20], super-resolution [49], low-light enhancement [23] and denoising [32]. The need to recover the clean 'medium-free' version of images degraded by scattering and attenuation effects, likewise, can benefit from the neural rendering approach. SeaThru-NeRF is designed to *model* the degradation, with the ability to recover its parameters, and reconstruct the clean underlying scene and novel-view images.

Our work is related to recent efforts to improve the quality and robustness of NeRFs in challenging environments (e.g. [14, 21, 29]). Nerf-W [21] learns a per-image latent embedding that can capture appearance variations in complex scenes. It decomposes the scene into image-dependent and shared components to disentangle transient elements from the static scene. NeRFReN [14] is designed for scenes with reflections. It separates the scene as a sum of transmitted and reflected components, which are modeled as separate NeRFs. Ref-NeRF [47] introduces a new parameterization and structuring of view-dependent appearance, which can represent scenes with specularities and reflections.

Our approach has much in common with some of these methods, most notably the way of reconstructing the target image as a composition of components: 'direct' and 'backscatter' in our case, 'transient' and 'static' in [21], or 'transmitted' and 'reflected' in [14]. However, as we demonstrate, these methods are limited on scenes in scattering media as they do not explicitly model the effects of the medium. Our method is different in that it can model a continuous medium component that is not an object.

Graphics and Vision in Scattering Media: Propagation of light in a medium is governed by the radiative transfer equation [10]. This equation describes light interaction per particle in the medium and requires extensive Monte Carlo simulations for complete solutions [12, 16, 30, 31], see [30] for an excellent review. In many cases simplifying assumptions can be made about the medium that ease rendering [22], where the major one is single-scattering [33, 42, 48]. Realistic rendering requires having the scattering properties of the medium, which can be estimated in the lab [13, 26], or from *in situ* images [6]. An MLP for rendering synthetic atmospheric clouds was suggested in [17].

In computer vision, image formation models under ambient illumination for bad weather [28] and underwater [37] share the same general structure for the case of horizontal viewing. Artificial illumination used underwater [46] and in fog has additional terms that account for the nonuniformity of the source. Underwater, the medium parameters exhibit strong wavelength dependency that was shown to affect model accuracy in wideband camera channels. The SeaThru model [1–3] was suggested to overcome this issue.

Scene reconstruction in scattering media is an ill-posed problem which was initially solved with multiple frames [27, 37, 38, 46]. Later on, single-image methods have proposed a multitude of image priors to overcome the ill-posed nature of the problem, see [19, 51] for comprehensive reviews on single-image dehazing and underwater reconstruction, respectively. The introduction of deep-learning has led to an explosion of single-image dehazing and underwater reconstruction networks. See [41] and [4] for recent reviews of deep-learning dehazing and underwater reconstruction works, respectively.

Underwater, it was suggested to solve for the 3D structure of the scene, prior to image restoration [9, 35] using the haze model and [2] using the revised one. In *SeaThru-NeRF*, we simultaneously reconstruct the scene and its 3D structure, yielding multiple advantages that we demonstrate empirically. WaterNeRF [40] suggests an underwater neural renderer that estimates the medium parameters only with respect to histogram-equalized images, separately from the rendering. Here, we show the huge benefit of modeling the scattering medium within the rendering equations, over a

variety of different scenes. Concurrently and independently to our work, [11] presented a NeRF for haze.

3. Scientific Background

3.1. Neural Radiance Fields (NeRFs)

The original NeRF formulation [25] implicitly represents a 3D scene by a trainable continuous function. It is typically parameterized by an MLP $f_\Theta : (\mathbf{x}, \mathbf{d}) \rightarrow (\mathbf{c}, \sigma)$ which encodes the density σ at the 3D point $\mathbf{x} = (x, y, z)$ and the color $\mathbf{c} = (r, g, b)$ emitted from this point in the viewing direction $\mathbf{d} = (\theta, \phi)$ (which is typically represented as a 2-element unit normalized 3D vector).[1]

This simple representation is used to simulate classical image-based rendering, by color accumulation along rays that are back-projected from a posed camera. If we parameterize points along a camera ray $\mathbf{r}$ by $\mathbf{r}(t) = \mathbf{o} + \mathbf{d}(t)$, where $\mathbf{o}$ is the camera center and $t \in \mathbb{R}_+$, the expected (image) color $C(\mathbf{r})$ along the ray can be written as:[2]

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

where: integration along the ray is limited to the range $t \in [t_n, t_f]$ (near and far bounds on the scene contents); $\sigma(t)$ and $\mathbf{c}(t)$ are shorthands for the *density* at the point $\mathbf{r}(t)$ and its emitted *color* towards the camera center; $T(t)$ denotes the accumulated *transmittance* along the ray from t_n to t (the probability that the ray travels from t_n to t without hitting other particles along the way), and is given by:

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

In practice, the rendered color $C(\mathbf{r})$ in Eq. (1) is approximated with the quadrature rule, by discretizing the range $[t_n, t_f]$ into a set of N intervals $I_i = [s_i, s_{i+1}]$ (where $t_n = s_0 < \cdots < s_N = t_f$), assuming that the density σ and color $\mathbf{c}$ are constant along each interval (e.g. by querying the model once per interval at its center-point).

In the discretized version of Eq. (1):

$$\hat{C}(\mathbf{r}) = \sum_{i=1}^{N} C_i(\mathbf{r}) \tag{3}$$

the contribution $C_i(\mathbf{r})$ of the interval I_i is given by:

$$C_i(\mathbf{r}) = \int_{s_i}^{s_{i+1}} T(t)\sigma_i\mathbf{c}_i dt = T(s_i)\left(1 - e^{-\sigma_i\delta_i}\right)\mathbf{c}_i \tag{4}$$

where σ_i and $\mathbf{c_i}$ are the (constant) density and color along the ith interval, whose length is $\delta_i = s_{i+1} - s_i$, and the transmittance $T(s_i)$ at the beginning of the interval is:

$$T(s_i) = \exp\left(-\sum_{j=0}^{i-1} \sigma_j\delta_j\right) \tag{5}$$

This fully differentiable NeRF model (with the discretized rendering scheme) is trained with a simple reconstruction loss

$$L = \sum_{\mathbf{r}\in R} \|\hat{C}(\mathbf{r}) - C(\mathbf{r})\|^2 \tag{6}$$

comparing each rendered training image pixel $\hat{C}(\mathbf{r})$ to its ground truth color $C(\mathbf{r})$. The model can then be used to synthesize photorealistic novel view images.

3.2. Image Formation in Scattering Media

Image formation in fog, haze, or underwater differs from image formation in clear air in two major aspects. First, the *direct* signal emanating from the object is attenuated as a function of distance and wavelength. Second, this signal is occluded by *backscatter* (termed also path-radiance or veiling-light) - radiance that is added due to the in-scattering from the particles along the line-of-sight (LOS), illustrated in Fig. 2. The intensity and color of the occluding backscatter layer are independent of the scene contents, and its intensity accumulates along the LOS, increasing with distance. As a result, the visibility and contrast of further objects is significantly reduced and their colors are distorted.

We adopt the revised model [1] as the general model for image formation in scattering media under ambient illumination. Image intensity (per pixel, per color channel) is given as:

$$I = \overbrace{\underbrace{J}_{\text{color}} \cdot \underbrace{\left(e^{-\beta^D(\mathbf{v}_D)\cdot z}\right)}_{\text{attenuation}}}^{\textit{direct}} + \overbrace{\underbrace{B^\infty}_{\text{color}} \cdot \underbrace{\left(1 - e^{-\beta^B(\mathbf{v}_B)\cdot z}\right)}_{\text{attenuation}}}^{\textit{backscatter}} \tag{7}$$

where I is the linear image captured by the camera of a scene with range z, J is the clear scene that would have been captured had there been no medium along z, and B^∞ is the backscatter water color at infinity, i.e., the backscatter at areas that contain no objects. Lastly, β^D and β^B are the attenuation and backscatter coefficients, respectively, the two parameters that describe the medium effects. The vectors $\mathbf{v}_D$ and $\mathbf{v}_B$ represent the dependencies of β^D and β^B on range, object reflectance, spectrum of ambient light, spectral response of the camera, and the physical scattering and beam attenuation coefficients of the water body, all of which are functions of wavelength. It was shown in [2] that β^B can be assumed constant in an image, while β^D mainly depends on the object distance and weakly on object reflectance, thus solving for the full model requires at least 6 unknowns. The value of B^∞ is usually assumed to be uniform in the scene

[1] Vectors (3D coordinates or color components) are denoted in **bold**.
[2] Each color channel is integrated separately.

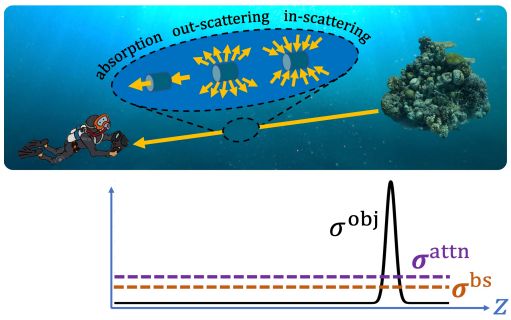

Figure 2. An illustration of our ray model. We assume a scene with at most a single opaque object at every ray. The medium is semi-transparent with constant density *per-ray*. The densities governing the backscatter and the object in water are not the same [1].

but as discussed in [6] it is almost never really uniform, because of the directionality of the sun, among other factors.

This model is also applicable to haze and fog, where attenuation has very little dependence on wavelength. Then, image formation is greatly simplified because it can be assumed that there is only one medium parameter that is *uniform* across the color channels, represented by a scalar ($\beta^D = \beta^B = \beta$) [38]. Underwater, often the simplifying (less accurate) assumption that $\beta^D = \beta^B$ is made, reducing to 3 unknowns [6, 7, 37].

4. *SeaThru*-NeRF

4.1. Basic Model Derivation

Here we consider a more general setting than the original NeRF [25], in which light travels through a scattering *medium* rather than free-space, resulting in a significant impact on the captured color. Following [22] we suggest adding the medium to Eqs. (1) and (2):

$$C(\mathbf{r}) = \int_{t_n}^{t_f} T(t)\Big(\sigma^{\text{obj}}(t)\mathbf{c}^{\text{obj}}(t) + \sigma^{\text{med}}(t)\mathbf{c}^{\text{med}}(t)\Big)dt \tag{8}$$

where:

$$T(t) = \exp\left(-\int_{t_n}^{t}\Big(\sigma^{\text{obj}}(s) + \sigma^{\text{med}}(s)\Big)ds\right) \quad , \tag{9}$$

using separate color and density parameters for the *object* and *medium*. Notice that these equations reduce to Eqs. (1) and (2) by simply considering the case where the medium density σ^{med} is zero.

Moving to the discretized version, we similarly get[3] the following generalizations of Eqs. (4) and (5):

$$C_i(\mathbf{r}) = T(s_i)\left(1 - e^{-(\sigma_i^{\text{obj}}+\sigma_i^{\text{med}})\delta_i}\right)\frac{\sigma_i^{\text{obj}}\mathbf{c}_i^{\text{obj}} + \sigma_i^{\text{med}}\mathbf{c}_i^{\text{med}}}{\sigma_i^{\text{obj}} + \sigma_i^{\text{med}}} \tag{10}$$

[3]Full derivations of Eqs. (10, 11) is provided in the appendix.

with transmittance

$$T(s_i) = \exp\left(-\sum_{j=0}^{i-1}(\sigma_j^{\text{obj}} + \sigma_j^{\text{med}})\delta_j\right) \quad . \tag{11}$$

Splitting the discretized color rendering equations (3 ,10 and 11) into the 'object' and 'medium' components, we get:

$$\hat{C}(\mathbf{r}) = \sum_{i=1}^{N}\hat{C}_i^{\text{obj}}(\mathbf{r}) + \sum_{i=1}^{N}\hat{C}_i^{\text{med}}(\mathbf{r}) \tag{12}$$

$$\hat{C}_i^{\text{obj}}(\mathbf{r}) = T(s_i)\left(1 - e^{-(\sigma_i^{\text{obj}}+\sigma_i^{\text{med}})\delta_i}\right)\frac{\sigma_i^{\text{obj}}\mathbf{c}_i^{\text{obj}}}{\sigma_i^{\text{obj}} + \sigma_i^{\text{med}}} \tag{13}$$

$$\hat{C}_i^{\text{med}}(\mathbf{r}) = T(s_i)\left(1 - e^{-(\sigma_i^{\text{obj}}+\sigma_i^{\text{med}})\delta_i}\right)\frac{\sigma_i^{\text{med}}\mathbf{c}_i^{\text{med}}}{\sigma_i^{\text{obj}} + \sigma_i^{\text{med}}} \tag{14}$$

As a first step towards constraining and simplifying our model, we constrain the medium parameters to be constant along 3D viewing rays. We drop the respective interval indices and remain with $\boldsymbol{\sigma}^{\text{med}}$ and $\mathbf{c}^{\text{med}}$ that depend only on the ray $\mathbf{r}$. The $\mathbf{c}^{\text{med}}$ uniformity stems from the common assumption of a uniform phase function (the dependence of scattered radiance on scattering angle) along the LOS [22, 30]. Regarding the density $\boldsymbol{\sigma}^{\text{med}}$, we take it to be separate per color channel, but constant per ray. This is far less restrictive compared to models that assume constancy per image or even per scene [2]. These constraints will be enforced by respective structural choices in the network.

Furthermore, we assume that objects in the scene are opaque, hence the object density along the ray is close to zero except for a high peak in the object location. On the other hand, the medium is semi-transparent, characterized by a low non-zero density. This implies that $\boldsymbol{\sigma}^{\text{med}} \gg \sigma^{\text{obj}}$ before the object and $\boldsymbol{\sigma}^{\text{med}} \ll \sigma^{\text{obj}}$ at the object, as illustrated in Fig. 2. Therefore, Eqs. (13-14) reduce to

$$\hat{C}_i^{\text{obj}}(\mathbf{r}) = T_i \cdot \left(1 - e^{-\sigma_i^{\text{obj}}\delta_i}\right)\cdot \mathbf{c}_i^{\text{obj}} \tag{15}$$

$$\hat{C}_i^{\text{med}}(\mathbf{r}) = T_i \cdot \left(1 - e^{-\boldsymbol{\sigma}^{\text{med}}\delta_i}\right)\cdot \mathbf{c}^{\text{med}} \tag{16}$$

$$T_i = \exp\left(-\sum_{j=0}^{i-1}\sigma_j^{\text{obj}}\delta_j\right)\cdot\exp\left(-\boldsymbol{\sigma}^{\text{med}}s_i\right) \tag{17}$$

4.2. Relation to Underwater Image Formation

In this section we show that our model can be reduced to the image formation model described in Sec. 3.2, which is not based on volumetric rendering. To this end we consider sampling the ray along intervals with constant size δ, and the appearance of the opaque object at a depth z, at the beginning of interval $I_k = [s_k, s_{k+1}]$, for some integer k (that is: $s_i = i \cdot \delta$ for every i and $z = k \cdot \delta$ in particular).

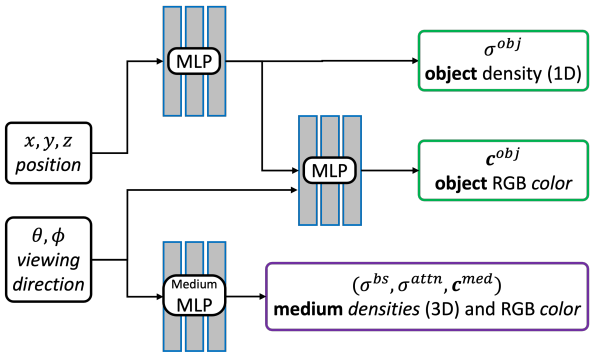

Figure 3. **SeaThru-NeRF architecture.** The computation of the 'object' outputs (density and color, in green), follows the standard NeRF architecture, while the 'medium' components (in purple) are computed once per ray by a separate subnet (the 'medium MLP') that depends only on the viewing direction.

In this case, $\sigma_i^{\text{obj}} \approx 0$ for all $i < k$, and $\sigma_k^{\text{obj}} \gg \boldsymbol{\sigma}^{\text{med}}$. This implies that $C_k^{\text{med}}(\mathbf{r}) \ll C_k^{\text{obj}}(\mathbf{r})$ and $C_i^{\text{obj}}(\mathbf{r}) \approx 0$, therefore we can write the rendering Eq. (12) as

$$\hat{C}(\mathbf{r}) \approx \hat{C}_k^{\text{obj}}(\mathbf{r}) + \sum_{i=0}^{k-1} \hat{C}_i^{\text{med}}(\mathbf{r}) \ . \tag{18}$$

Furthermore, since the object is opaque, we can assume that the density σ_k^{obj} is large enough such that the transmittance at the end of the kth interval $T(s_{k+1})$ drops practically to zero, that is: $e^{-\sigma_k^{\text{obj}}\delta} \approx 0$. Therefore

$$\begin{aligned} \hat{C}_k^{\text{obj}}(\mathbf{r}) &= T_i \cdot \left(1 - e^{-\sigma_i^{\text{obj}}\delta}\right) \cdot \mathbf{c}_i^{\text{obj}} \\ &\approx e^{-k\boldsymbol{\sigma}^{\text{med}}\delta}\left(1 - e^{-\sigma_k^{\text{obj}}\delta}\right)\mathbf{c}_k^{\text{obj}} \approx e^{-\boldsymbol{\sigma}^{\text{med}}\delta\cdot k}\mathbf{c}_k^{\text{obj}} \end{aligned} \tag{19}$$

and

$$\begin{aligned} \sum_{i=0}^{k-1} \hat{C}_i^{\text{med}}(\mathbf{r}) &= \sum_{i=0}^{k-1} T_i \cdot \left(1 - e^{-\boldsymbol{\sigma}^{\text{med}}\delta}\right) \cdot \mathbf{c}^{\text{med}} \\ &\approx \sum_{i=0}^{k-1} e^{-i\boldsymbol{\sigma}^{\text{med}}\delta} \cdot \left(1 - e^{-\boldsymbol{\sigma}^{\text{med}}\delta}\right) \cdot \mathbf{c}^{\text{med}} \\ &= \left(1 - e^{-\boldsymbol{\sigma}^{\text{med}}\delta\cdot k}\right) \cdot \mathbf{c}^{\text{med}} \end{aligned} \tag{20}$$

Substituting Eqs. (19) and (20) into Eq. (18) (recalling that the depth z equals $k \cdot \delta$) yields:

$$\hat{C}(\mathbf{r}) \approx e^{-\boldsymbol{\sigma}^{\text{med}} z} \cdot \mathbf{c}_k^{\text{obj}} + \left(1 - e^{-\boldsymbol{\sigma}^{\text{med}} z}\right) \cdot \mathbf{c}^{\text{med}} \ , \tag{21}$$

resulting precisely in the commonly used scattering media image formation model [8, 37]. This can be seen by comparing to Eq. (7) where $\boldsymbol{\sigma}^{\text{med}} = \beta^D = \beta^B$ plays the role of the attenuation coefficient (equal for direct and backscatter), $\mathbf{c}^{\text{med}}$ is the veiling light B^∞ (backscatter color at infinity) and $\mathbf{c}_k^{\text{obj}}$ the clear image color J.

4.3. Final Model

We make several refinements with respect to the basic model presented in Section 4.1. In Sec. 4.2, we showed that our rendering equations lead to an image formation model with identical attenuation coefficients for the direct and backscatter components. Following the discussion in Sec. 3.2, and as has been shown in [1], when using such equations with a camera with wideband color channels, the effective $\boldsymbol{\sigma}^{\text{med}}$ that is experienced by the camera in $C_i^{\text{obj}}(\mathbf{r})$ is different than the one experienced in $C_i^{\text{med}}(\mathbf{r})$. Therefore, in our final model we use different parameters for $\boldsymbol{\sigma}^{\text{med}}$ in each component and term them $\boldsymbol{\sigma}^{\text{attn}}$ and $\boldsymbol{\sigma}^{\text{bs}}$ for $C_i^{\text{obj}}(\mathbf{r})$ and $C_i^{\text{med}}(\mathbf{r})$ respectively. Our final equations are:

$$\hat{C}_i^{\text{obj}}(\mathbf{r}) = T_i^{\text{obj}} \cdot \exp\left(-\boldsymbol{\sigma}^{\text{attn}} s_i\right) \cdot \left(1 - \exp(-\sigma_i^{\text{obj}}\delta_i)\right) \cdot \mathbf{c}_i^{\text{obj}}$$

$$\hat{C}_i^{\text{med}}(\mathbf{r}) = T_i^{\text{obj}} \cdot \exp\left(-\boldsymbol{\sigma}^{\text{bs}} s_i\right) \cdot \left(1 - \exp(-\boldsymbol{\sigma}^{\text{bs}}\delta_i)\right) \cdot \mathbf{c}^{\text{med}}$$

$$T_i^{\text{obj}} = \exp\left(-\sum_{j=0}^{i-1} \sigma_j^{\text{obj}}\delta_j\right) \tag{22}$$

Based on our derivations we suggest the following architecture (Fig. 3). As in [25], the object properties are computed by a pair of MLPs, such that the density σ^{obj} is a function of the position (x, y, z) only, while color $\mathbf{c}^{\text{obj}}$ is determined by the viewing direction (θ, ϕ) as well. The medium parameters $\mathbf{c}^{\text{med}}$, $\boldsymbol{\sigma}^{\text{bs}}$, $\boldsymbol{\sigma}^{\text{attn}}$ are handled by a separate MLP with only the direction input, following our decision to constrain them to be constant per viewing direction.

4.4. Loss Function

Let us denote the sequence of samples along a ray by $\mathbf{s} = \{s_i\}_{i=0}^N$, the object 'weight' at the ith segment $[s_i, s_{i+1}]$,

$$w_i^{\text{obj}} = T_i^{\text{obj}} \cdot \left(1 - \exp(-\sigma_i^{\text{obj}}\delta_i)\right) \ , \tag{23}$$

the sequence of these weights by $\mathbf{w} = \{w_i^{\text{obj}}\}_{i=0}^{N-1}$ and the ground-truth (supervised) pixel color by C^*.

Our loss is the following combination:

$$\mathcal{L} = \mathcal{L}_{\text{recon}}(\hat{C}, C^*) + \mathcal{L}_{\text{prop}}(\mathbf{s}, \mathbf{w}) + \lambda\mathcal{L}_{\text{acc}}(\mathbf{w}) \ , \tag{24}$$

where $\lambda = 0.0001$ was chosen through cross-validation. Since we work on linear images, we adopt the reconstruction loss of RawNeRF [23]:

$$\mathcal{L}_{\text{recon}}(\hat{C}, C^*) = \left(\frac{\hat{C} - C^*}{\text{sg}(\hat{C}) + \epsilon}\right)^2 \tag{25}$$

where $\text{sg}(\cdot)$ stands for stop-gradient and $\epsilon = 10^{-3}$. The 'proposal' loss $\mathcal{L}_{\text{prop}}(\mathbf{s}, \mathbf{w})$ is an inherent part of Mip-NeRF-360 [5]. It penalizes for the discrepancy between the

distributions of object weights at 'original' and 'proposed' samplings, where only the latter is used for rendering [5].

To enforce binary separation between points in space that contain objects and those that contain solely a medium - we add a prior on the transmittance T_i^{obj} of each point on the ray to be either 0 or 1, not allowing semi-transparent objects. This is modelled as a mixture of two Laplacian distributions with modes at 0 and 1 (following [34]):

$$\mathbb{P}(x) \propto e^{-\frac{|x|}{0.1}} + e^{-\frac{|1-x|}{0.1}} \tag{26}$$

and use the negative log likelihood loss:

$$\mathcal{L}_{\text{acc}}(\mathbf{w}) = -\log \mathbb{P}(T_i^{\text{obj}}) \ . \tag{27}$$

4.5. Implementation and Optimization

Our implementation is based on the code released in Mip-NeRF-360 [5], choosing the best performing baseline on our scenes which was the forward-looking configuration with normalized device coordinates (NDC). For the mediumMLP , we use 6 linear layers with 256 features and a softplus activation, followed by 3 branches of dense layers and a sigmoid activation for predicting $\boldsymbol{c}^{\text{med}}$ and softplus activations for $\boldsymbol{\sigma}^{\text{attn}}$ and $\boldsymbol{\sigma}^{\text{bs}}$.

In the rendering scheme, [5] initialized the farthest δ_i with infinity, enabling the network to predict a background color for rays that do not intersect with any object. We disabled this addition as it prevents our method from explaining the medium that contributes to the rendered color along the ray. We keep the learning rate and optimization parameters the same as in [5]. The network is trained for 250,000 iterations with a batch size of 16384 rays, taking around 10 hours on an Nvidia A100 GPU. The loss function and metrics are calculated on the output before any post-processing.

5. Experiments and Results

5.1. Experiments

Real world scenes. We acquired multi-image underwater scenes by diving in three different seas: the Red Sea (Eilat, Israel), the Caribbean Sea (Curaçao) and the Pacific Sea (Panama) with a total of 20, 20 and 18 images respectively, from which three are set aside for validation in each set. This data encapsulates a diverse set of water conditions and imaging conditions. The images were acquired as RAW images with a Nikon D850 SLR camera in a Nauticam underwater housing with a dome port to avoid refractions that jeopardize the pinhole model [45], and downsampled to average size 900×1400. The input linear images were white-balanced before processing with 0.5% clipping per channel to remove extreme noisy pixels. Finally, COLMAP [39] is used to extract camera poses.

Simulated Scene. We constructed a simulation using the *Fern* scene of the LLFF dataset [24]. We ran MIP-NeRF-360 [5] and used the predicted depth maps to simulate an underwater and a foggy scene. Water was added according to (7) with parameter values: $\beta^D = [1.3, 1.2, 0.9]$, $\beta^B = [0.95, 0.85, 0.7]$, $B^\infty = [0.07, 0.2, 0.39]$. Fog was simulated based on [36] with $\beta = 1.2$.

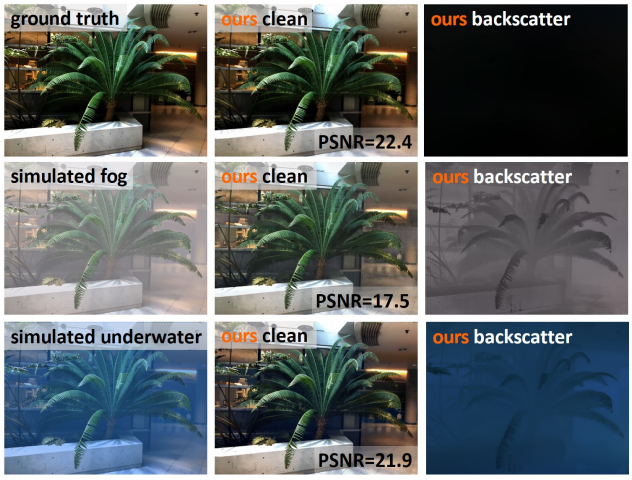

Figure 4. **Synthetic experiment on *Fern* scene of the LLFF dataset [24].** ***Top***: Our method can handle 'in-air' scenes, with a uniform zeros backscatter image and a reconstruction PSNR that is close to that of the baseline [5] which was 23.73. ***Middle*** and ***bottom***: Our method's separation between clean and backscatter components, for simulations of fog and underwater effects. The reconstruction quality degrades rather gracefully.

Baseline methods. The input to all methods is the same set of white-balanced linear images. In the task of **rendering scenes in the medium** we compare with the following NeRF methods: MIP-NeRF-360, forward-looking with NDC (MIP360) [5], NeRF-W [21] and NeRFReN [14]. In the task of **reconstructing clean medium-free images** we compare with NeRF-W, where the 'transient' component can be viewed as the clean image and the 'static' image as backscatter, and to NeRFReN with the 'reflected' and 'transmitted' components as clean and backscatter images. We additionally compare to single-image scene reconstruction methods: plain white balance, SeaThru [2], the leading "classical" methods of Bekerman et al [6], and the recent deep-learning based [52] (as well as [15, 52] whose results are provided in the project website). SeaThru requires depth maps, that we generate using SFM (with Agisoft Metashape) as originally suggested. These depth-maps are used to compare with those of the different methods.

Photo-finishing: As the input and output images are linear, we apply photofinishing on all linear reconstructed scenes to enhance scene contrast and appearance, using the digital camera pipeline from [18]. This is done to improve visualization for easier qualitative comparisons, while PSNR is calculated on the original non-photofinished linear images.

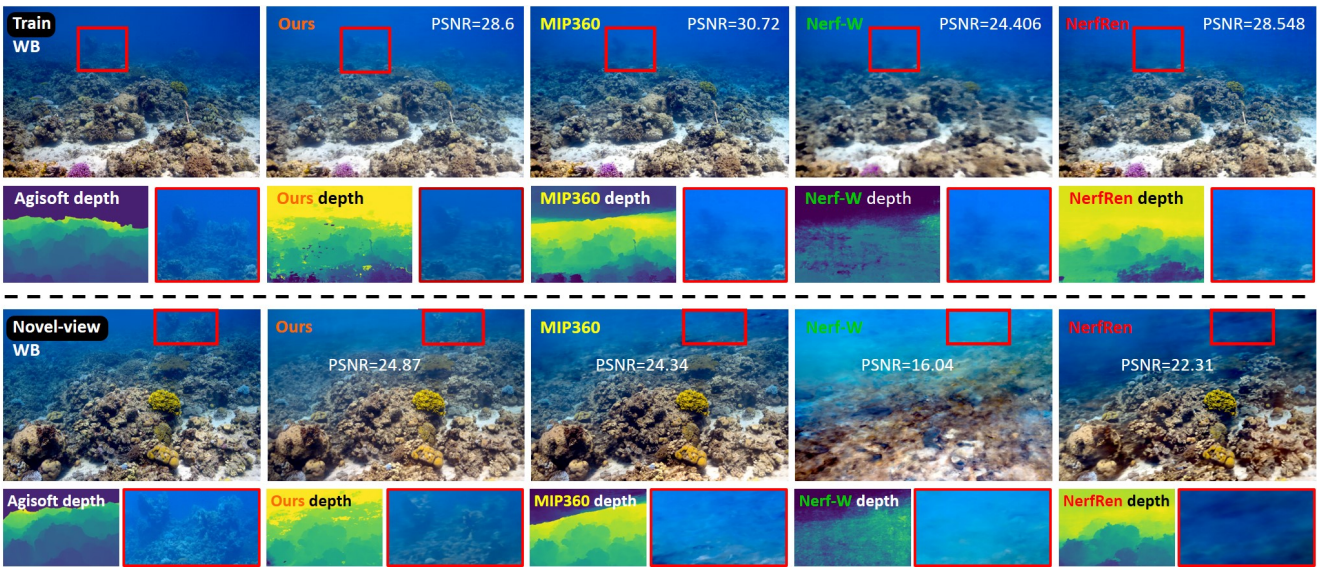

Figure 5. **Scene rendering in the medium, on 'Red Sea'**. Left to right: white-balanced input image, our result, MIP360 [23], Nerf-W [21], NeRFReN [14]. **[Top 2 rows]** Train image and zoom-ins. Even in the task of overfitting to the training images, there are noticeable differences between the methods, and as can be seen in the zoom-ins (red squares) we are able to reconstruct fine details in further areas (albeit a lower PSNR). **[Bottom 2 rows]** Novel view and zoom-ins. Our method achieves the highest PSNR (see tables 1, 2), and provides much better details in further areas (see red square zoom-ins). Our depth reconstructions provide more detail in further areas.

5.2. Results

We start with a **sanity check** on a clean image. Our method correctly estimates zero backscatter, and does not force itself to estimate a medium (Fig. 4 Top). On our **simulated** fog and underwater scenes (Fig. 4 middle and bottom), our method separates the scenes' components very well, with only slight reduction in PSNR in the underwater scene.

Image rendering and novel-view synthesis (in medium). Qualitative results are summarized in Figs. 1, 5. Table 1 summarizes the average PSNR on the validation set of the Red Sea scene, in which our method achieves the highest PSNR. Note that both NeRF-W [21] and NeRFReN [14] are based on the NeRF [25] model while our code is based on Mip-NeRF-360 [5] that in general improves reconstruction details in further areas. So our real comparison is with [5], which we improve by $\sim 0.8dB$ on average over the validation set. Table 2 compares PSNR, SSIM and LPIPS with those of [5] on validation sets of all real-world and simulation scenes. Our method is better in the majority of cases, and is *especially good on the further areas (red squares)*.

On the train set [5] achieves a reasonable rendering of the scene in terms of PSNR by **wrongly modeling the water as a nearby blue object**, as can be seen in the depth map, indicating a close object in the top area. Even then, in our results further objects are reconstructed with more detail. The NeRFReN [14] depth map flattens at the mid-range, while ours is more informative further on. SFM does not estimate depth at these areas because of lack of features.

Scene restoration (Fig. 6). In comparison with NeRF methods for scene separation it is clear that our separation handles the medium the best. In the result of NeRF-W [21] the scene is somewhat separated but blurry, while in that of NeRFReN [14] the closer part of the scene is relatively well reconstructed but the further areas are missing. In both, the scene's objects are visible in the backscatter image, which is clearly wrong and reduces relevant signal from the object image. In our results, due to the adequate modeling, the backscatter is not polluted with objects. Compared to scene restoration methods, SeaThru [2] is the best performing one but is limited to areas where the depth map is available from SFM. The single image methods are not on par, especially from the mid-ranges and on. In **novel-view synthesis of clean scenes** we are able to predict further the entire scene.

Ours	I	II	III	[5]	[21]	[14]
21.83	21.76	21.43	21.72	21.05	16.52	21.05

Table 1. **Average PSNRs of in-medium rendering on Red Sea set**: ours, 3 ablation variations (description in the text: I– 1 parameter, II– 3 parameters, III– Eqs. 13, 14) and other methods.

Scene	**MIP360 [5]**	**Ours**
Red Sea	21.05 / 0.75 / 0.29	**21.83 / 0.77 / 0.25**
Red Sea red square	29.66 / 0.84 / 0.43	**33.80 / 0.90 / 0.23**
Curaçao	26.54 / 0.81 / 0.33	**30.48 / 0.87 / 0.20**
Curaçao red square	27.04 / 0.84 / 0.45	**33.20 / 0.88 / 0.08**
Panama	27.43 / 0.82 / 0.23	**27.89 / 0.83 / 0.22**
Fern fog	30.23 / **0.88 / 0.15**	**30.75** / 0.87 / **0.16**
Fern underwater	29.62 / **0.87** / 0.26	**29.76** / 0.86 / **0.15**

Table 2. **Comparison to baseline**: PSNR↑ / SSIM↑ / LPIPS↓ .

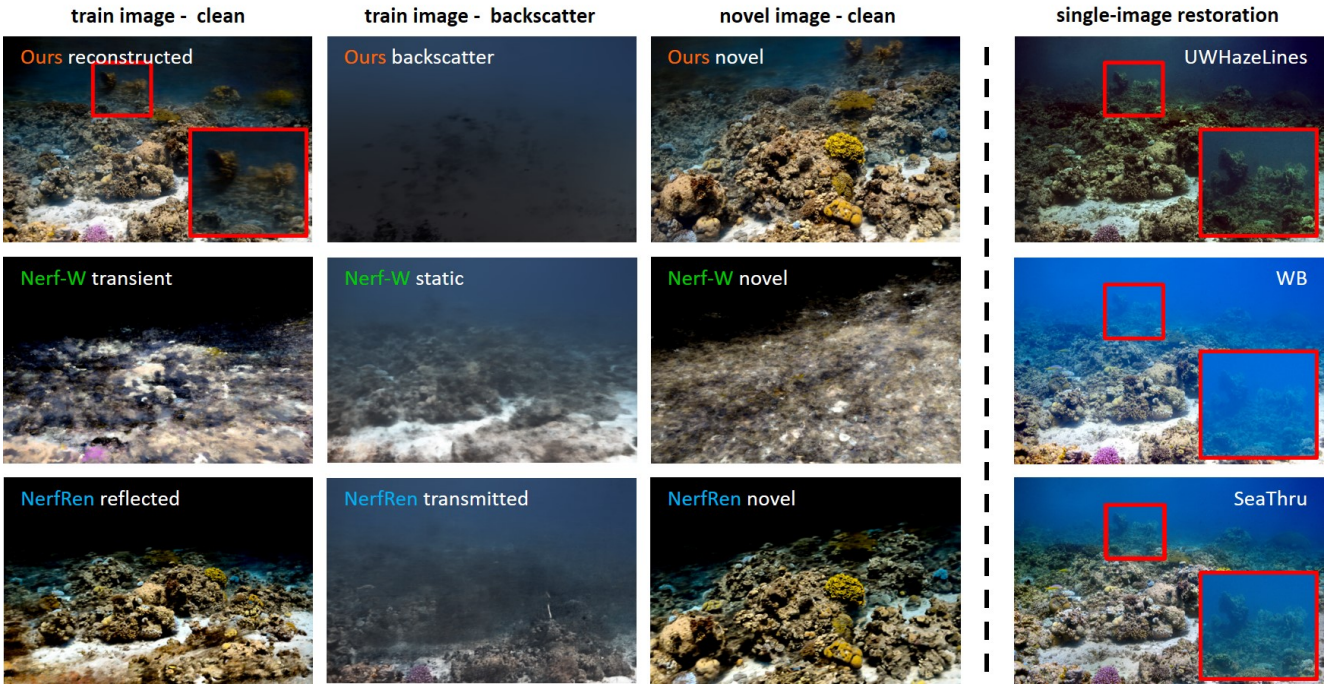

Figure 6. **Scene reconstruction on 'Red Sea'**. ***Columns 1-3***: Comparing our method (top row) and 2 other NeRFs - Nerf-W [21] (middle row) and NeRFReN [14] (bottom row). Even though not developed for scattering media, they manage to separate a training image into two components (columns 1, 2), resembling the clean scene and the backscatter. However, their backscatter components contains details of the objects, while ours correctly displays only the signal stemming from the medium. Similarly (column 3), their 'clean' rendering of a novel view is inferior to ours, especially in the far distances. ***Column 4***: Among the single-image methods, SeaThru [2] is the best performing (but requires the depth map as input). Nevertheless, color artifacts are visible in the far parts due to inaccurate depth information.

Ablation Study. (i) *Number of medium parameters:* following the discussion in Sec. 3.2, the elaborate version of Eq. (7) uses 6 parameters for the medium coefficients, while simplified versions use 1 (marked I in Table 1) or 3 (marked II). This is ablated in Table 1, in terms of PSNR, where the elaborate model with 6 parameters achieves a higher score. (ii) *Rendering equations.* We compare our basic rendering model (Eqs. 13, 14), marked III in Table 1 to the final one marked 'ours' (Eq. 22). This option achieves lower PSNR.

6. Discussion

We provide an important extension of NeRFs that enables rendering scenes acquired in scattering media such as haze, fog and underwater. So far, NeRF provided a framework for volumetric rendering, but without considering the nature of the medium, resulting in a binary 'occupancy' volume. Our formulation enables opaque objects to exist in a semi-transparent medium that is both scattering and absorbing in a wavelength-dependent manner. We demonstrate it on challenging real-world scenes with a complex 3D structure. Our scenes are forward-facing and contain areas with no objects at all, which our method is able to explain.

Water effects in further regions are very strong, hampering single image methods and feature matching in multi-view methods. Our method incorporates information from all images at once, learning the scene in a holistic way. This enables better scene reconstruction (in medium and clean), and estimating the depth and the medium's parameters.

Our method has several limitations. While it is based on the current state-of-the-art image formation model, that model does not account for multiple scattering or for artificial illumination. As common in NeRFs, it requires camera poses that are extracted beforehand, which can be challenging in bad visibility. Lastly, the medium's parameters are better learned in sets where there is enough variation in the scene range between the viewpoints. The formulation takes its strength from the modeling of the medium. Thus it struggles in scenes that do not adhere to the model's assumptions, e.g., underwater scenes with significant flickering. In the future we plan to add components that will explain transient effects such as flickering, and continue to explore estimation of more diverse scenes and medium parameters.

Acknowledgements. The research was funded by Israel Science Foundation grant #680/18, Israeli Ministry of Science and Technology, European Union's Horizon 2020 research and innovation programme GA 101094924 (ANERIS), the Leona M. and Harry B. Helmsley Charitable Trust, the Maurice Hatter Foundation, and Schmidt Marine Technology Partners. We thank Matan Yuval for substantial data contribution, Opher Bar-Nathan and Yuval Goldfracht for help with experiments.

References

[1] Derya Akkaynak and Tali Treibitz. A revised underwater image formation model. In *CVPR*, pages 6723–6732, 2018. 1, 2, 3, 4, 5

[2] Derya Akkaynak and Tali Treibitz. Sea-thru: A method for removing water from underwater images. In *CVPR*, pages 1682–1691, 2019. 2, 3, 4, 6, 7, 8

[3] Derya Akkaynak, Tali Treibitz, Tom Shlesinger, Raz Tamir, Yossi Loya, and David Iluz. What is the space of attenuation coefficients in underwater computer vision? In *CVPR*, 2017. 1, 2

[4] Saeed Anwar and Chongyi Li. Diving deeper into underwater image enhancement: A survey. *Signal Processing: Image Communication*, 89:115978, 2020. 2

[5] Jonathan T Barron, Ben Mildenhall, Dor Verbin, Pratul P Srinivasan, and Peter Hedman. Mip-nerf 360: Unbounded anti-aliased neural radiance fields. In *CVPR*, pages 5470–5479, 2022. 1, 5, 6, 7

[6] Yael Bekerman, Shai Avidan, and Tali Treibitz. Unveiling optical properties in underwater images. In *Proc. IEEE ICCP*, 2020. 2, 4, 6

[7] Dana Berman, Shai Avidan, and Tali Treibitz. Non-local image dehazing. In *CVPR*, 2016. 4

[8] Dana Berman, Deborah Levy, Shai Avidan, and Tali Treibitz. Underwater single image color restoration using haze-lines and a new quantitative dataset. *IEEE TPAMI*, 43(8):2822–2837, 2020. 5

[9] Mitch Bryson, Matthew Johnson-Roberson, Oscar Pizarro, and Stefan B Williams. True color correction of autonomous underwater vehicle imagery. *J. of Field Robotics*, 33(6):853–874, 2016. 2

[10] Subrahmanyan Chandrasekhar. *Radiative transfer*. Courier Corporation, 2013. 2

[11] Wei-Ting Chen, Wang Yifan, Sy-Yen Kuo, and Gordon Wetzstein. Dehazenerf: Multiple image haze removal and 3d shape reconstruction using neural radiance fields. *arXiv preprint arXiv:2303.11364*, 2023. 3

[12] Iliyan Georgiev, Jaroslav Krivanek, Toshiya Hachisuka, Derek Nowrouzezahrai, and Wojciech Jarosz. Joint importance sampling of low-order volumetric scattering. *ACM TOG*, 32(6), 2013. 2

[13] Ioannis Gkioulekas, Shuang Zhao, Kavita Bala, Todd Zickler, and Anat Levin. Inverse volume rendering with material dictionaries. *ACM TOG*, 32(6), 2013. 2

[14] Yuan-Chen Guo, Di Kang, Linchao Bao, Yu He, and Song-Hai Zhang. Nerfren: Neural radiance fields with reflections. In *CVPR*, pages 18409–18418, 2022. 2, 6, 7, 8

[15] Junlin Han, Mehrdad Shoeiby, Tim Malthus, Elizabeth Botha, Janet Anstee, Saeed Anwar, Ran Wei, Mohammad Ali Armin, Hongdong Li, and Lars Petersson. Underwater image restoration via contrastive learning and a real-world dataset. *Remote Sensing*, 14(17):4297, 2022. 6

[16] Henrik Wann Jensen and Per H Christensen. Efficient simulation of light transport in scenes with participating media using photon maps. In *Proc. of the 25th annual conference on Computer graphics and interactive techniques*, pages 311–320, 1998. 2

[17] Simon Kallweit, Thomas Müller, Brian Mcwilliams, Markus Gross, and Jan Novák. Deep scattering: Rendering atmospheric clouds with radiance-predicting neural networks. *ACM TOG*, 36(6), 2017. 2

[18] Hakki Can Karaimer and Michael S Brown. A software platform for manipulating the camera imaging pipeline. In *ECCV*, pages 429–444, 2016. 6

[19] Yu Li, Shaodi You, Michael S Brown, and Robby T Tan. Haze visibility enhancement: A survey and quantitative benchmarking. *Computer Vision and Image Understanding*, 165, 2017. 2

[20] Li Ma, Xiaoyu Li, Jing Liao, Qi Zhang, Xuan Wang, Jue Wang, and Pedro V Sander. Deblur-nerf: Neural radiance fields from blurry images. In *CVPR*, pages 12861–12870, 2022. 2

[21] Ricardo Martin-Brualla, Noha Radwan, Mehdi SM Sajjadi, Jonathan T Barron, Alexey Dosovitskiy, and Daniel Duckworth. Nerf in the wild: Neural radiance fields for unconstrained photo collections. In *CVPR*, pages 7210–7219, 2021. 2, 6, 7, 8

[22] Nelson Max and Min Chen. Local and global illumination in the volume rendering integral. Technical report, Lawrence Livermore National Lab (LLNL), Livermore, CA (United States), 2005. 2, 4

[23] 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. In *CVPR*, pages 16190–16199, 2022. 2, 5, 7

[24] Ben Mildenhall, Pratul P. Srinivasan, Rodrigo Ortiz-Cayon, Nima Khademi Kalantari, Ravi Ramamoorthi, Ren Ng, and Abhishek Kar. Local light field fusion: Practical view synthesis with prescriptive sampling guidelines. *TOG*, 2019. 6

[25] 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*, pages 405–421, 2020. 1, 2, 3, 4, 5, 7

[26] Srinivasa G Narasimhan, Mohit Gupta, Craig Donner, Ravi Ramamoorthi, Shree K Nayar, and Henrik Wann Jensen. Acquiring scattering properties of participating media by dilution. *ACM TOG*, pages 1003–1012, 2006. 2

[27] Srinivasa G Narasimhan and Shree K Nayar. Vision and the atmosphere. *IJCV*, 48(3):233–254, 2002. 2

[28] Shree K Nayar and Srinivasa G Narasimhan. Vision in bad weather. In *CVPR*, pages 820–827, 1999. 2

[29] Michael Niemeyer and Andreas Geiger. Giraffe: Representing scenes as compositional generative neural feature fields. In *CVPR*, 2021. 2

[30] Jan Novák, Iliyan Georgiev, Johannes Hanika, and Wojciech Jarosz. Monte carlo methods for volumetric light transport simulation. In *Computer Graphics Forum*, volume 37, pages 551–576. Wiley Online Library, 2018. 2, 4

[31] Mark Pauly, Thomas Kollig, and Alexander Keller. Metropolis light transport for participating media. In *Eurographics Workshop on Rendering Techniques*, pages 11–22. Springer, 2000. 2

[32] Naama Pearl, Tali Treibitz, and Simon Korman. Nan: Noise-aware nerfs for burst-denoising. In *CVPR*, pages 12672–12681, 2022. 2

[33] Vincent Pegoraro and Steven G Parker. An analytical solution to single scattering in homogeneous participating media. In *Computer Graphics Forum*, volume 28, pages 329–335. Wiley Online Library, 2009. 2

[34] Daniel Rebain, Mark Matthews, Kwang Moo Yi, Dmitry Lagun, and Andrea Tagliasacchi. Lolnerf: Learn from one look. In *CVPR*, 2022. 6

[35] Miran Roser, Matthew Dunbabin, and Andreas Geiger. Simultaneous underwater visibility assessment, enhancement and improved stereo. In *IEEE Conf. Robotics and Automation*, pages 3840–3847, 2014. 2

[36] Christos Sakaridis, Dengxin Dai, and Luc Van Gool. Guided curriculum model adaptation and uncertainty-aware evaluation for semantic nighttime image segmentation. In *ICCV*, 2019. 6

[37] Yoav Y Schechner and Nir Karpel. Recovery of underwater visibility and structure by polarization analysis. *IEEE J. oceanic engineering*, 30(3):570–587, 2005. 2, 4, 5

[38] Yoav Y Schechner, Srinivasa G Narasimhan, and Shree K Nayar. Instant dehazing of images using polarization. In *CVPR*, 2001. 2, 4

[39] Johannes L Schonberger and Jan-Michael Frahm. Structure-from-motion revisited. In *CVPR*, 2016. 6

[40] Advaith Venkatramanan Sethuraman, Manikandasriram Srinivasan Ramanagopal, and Katherine A Skinner. Waternerf: Neural radiance fields for underwater scenes. *arXiv preprint arXiv:2209.13091*, 2022. 2

[41] Neeraj Sharma, Vijay Kumar, and Sunil Kumar Singla. Single image defogging using deep learning techniques: past, present and future. *Archives of Computational Methods in Engineering*, 28(7):4449–4469, 2021. 2

[42] Bo Sun, Ravi Ramamoorthi, Srinivasa G Narasimhan, and Shree K Nayar. A practical analytic single scattering model for real time rendering. *ACM TOG*, 24(3):1040–1049, 2005. 2

[43] Ayush Tewari, Ohad Fried, Justus Thies, Vincent Sitzmann, Stephen Lombardi, Kalyan Sunkavalli, Ricardo Martin-Brualla, Tomas Simon, Jason Saragih, Matthias Nießner, et al. State of the art on neural rendering. In *Computer Graphics Forum*, volume 39. Wiley Online Library, 2020. 2

[44] Ayush Tewari, Justus Thies, Ben Mildenhall, Pratul Srinivasan, Edgar Tretschk, W Yifan, Christoph Lassner, Vincent Sitzmann, Ricardo Martin-Brualla, Stephen Lombardi, et al. Advances in neural rendering. In *Computer Graphics Forum*, volume 41, pages 703–735. Wiley Online Library, 2022. 2

[45] Tali Treibitz, Yoav Schechner, Clayton Kunz, and Hanumant Singh. Flat refractive geometry. *PAMI*, 34(1):51–65, 2011. 6

[46] Tali Treibitz and Yoav Y Schechner. Active polarization descattering. *IEEE TPAMI*, 31(3):385–399, 2009. 2

[47] 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. In *CVPR*, pages 5481–5490, 2022. 2

[48] Bruce Walter, Shuang Zhao, Nicolas Holzschuch, and Kavita Bala. Single scattering in refractive media with triangle mesh boundaries. *ACM TOG*, 2009. 2

[49] 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. In *Proc. the 30th ACM Int. Conf. on Multimedia*, pages 6445–6454, 2022. 2

[50] 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. 2

[51] Miao Yang, Jintong Hu, Chongyi Li, Gustavo Rohde, Yixiang Du, and Ke Hu. An in-depth survey of underwater image enhancement and restoration. *IEEE Access*, 7:123638–123657, 2019. 2

[52] Weidong Zhang, Peixian Zhuang, Hai-Han Sun, Guohou Li, Sam Kwong, and Chongyi Li. Underwater image enhancement via minimal color loss and locally adaptive contrast enhancement. *IEEE TIP*, 31:3997–4010, 2022. 6