Zero-shot Single Image Restoration through Controlled Perturbation of Koschmieder's Model

Aupendu Kar*, Sobhan Kanti Dhara*, Debashis Sen, Prabir Kumar Biswas
Indian Institute of Technology Kharagpur, WB, India

Abstract

Real-world image degradation due to light scattering can be described based on the Koschmieder's model. Training deep models to restore such degraded images is challenging as real-world paired data is scarcely available and synthetic paired data may suffer from domain-shift issues. In this paper, a zero-shot single real-world image restoration model is proposed leveraging a theoretically deduced property of degradation through the Koschmieder's model. Our zero-shot network estimates the parameters of the Koschmieder's model, which describes the degradation in the input image, to perform image restoration. We show that a suitable degradation of the input image amounts to a controlled perturbation of the Koschmieder's model that describes the image's formation. The optimization of the zero-shot network is achieved by seeking to maintain the relation between its estimates of Koschmieder's model parameters before and after the controlled perturbation, along with the use of a few no-reference losses. Image dehazing and underwater image restoration are carried out using the proposed zero-shot framework, which in general outperforms the state-of-the-art quantitatively and subjectively on multiple standard real-world image datasets. Additionally, the application of our zero-shot framework for low-light image enhancement is also demonstrated.

1. Introduction

Due to the scattering of light in adverse environmental conditions, captured images suffer from poor visibility [48, 34]. Such conditions are often encountered during imaging in hazy and underwater environments. Due to the presence of suspended particles like dust, aerosol, and water droplets in air, light gets absorbed to produce hazy images having reduced visibility and contrast [38]. The presence of color particles renders a hazy image to be a color cast one as well [32, 14, 31]. Due to absorption of light in turbid water medium, and attenuation of light depending on optical wavelengths and water salinity, underwater images suffer from poor visibility and color cast [46]. The well-accepted Koschmieder's light scattering model [34, 48] with the transmission map and atmospheric light parameters describes such degradations in hazy images. A modified version of the model [30] does so for the underwater images, where the atmospheric light is replaced by the global background light and the transmission map is modified into a channel-wise component. The said degradations in the images lead to poor performance in computer vision applications such as surveillance and robotic navigation [37, 27, 29].

A plethora of techniques have been proposed for image dehazing [26, 10, 36, 52, 47, 51, 32, 15, 37, 22] and underwater image restoration [9, 50, 63, 12, 3, 42, 5, 2, 61, 39, 19, 58] to restore the images degraded due to haze and turbid medium, respectively. Most of the approaches either provide hand-crafted prior or deep learning based solutions for the restoration. Most of the deep learning based techniques rely on synthetically generated degradation for training due to lack of large-scale datasets of real-world degraded images with ground truth [36]. Such a training may not capture the features of the real-world degradation, and hence, suffer from domain-shift issues [36]. A few deep learning based solutions are unsupervised approaches which either use hand-crafted priors for restoration [22] or use generative adversarial networks (GAN) [52].

A zero-shot approach learns the task at hand from the input image alone, and therefore, is helpful in scenarios where real-world paired data for training is limited and its creation is labor-intensive [35, 57, 53, 36]. Hence, a few zero-shot approaches have been proposed for the said restoration based on the Koschmieder's model. However, as these zero-shot approaches do not have any other reference than the input image itself, they are required to estimate the parameters of the Koschmieder's model using loss functions or regularizers based on classical hand-crafted priors. For example, the zero-shot image dehazing approaches Double-DIP [21] and ZID [36] use the dark channel prior DCP [26]

*Aupendu Kar and Sobhan Kanti Dhara share equal contribution
Webpage and Code: aupendu.github.io/zero-restore
Correspondence to: {*mailtoaupendu, dhara.sk*}@*gmail.com*

to train their transmission map estimator and consider the atmospheric light estimated by [8]. As such zero-shot approaches depend on the use of hand-crafted priors to derive a result relevant to the problem at hand, they carry a risk of obtaining an output dominated by the characteristics of the prior.

This paper proposes a zero-shot single real-world image restoration approach that does not employ any hand-crafted prior for optimization. Our approach is based on the Koschmieder's light scattering model and is free to converge at a solution without any prior bias. To this end, we theoretically show that a suitable degradation of the input image amounts to a controlled perturbation of the Koschmeider's model parameters that define the relation between the input image and the original uncorrupted image. Invoking the above finding, an image-specific network model is designed to estimate the parameters of the Koschmieder's model, which are then used for the image restoration. The zero-shot learning of the network model is achieved through an optimization targeted to maintain the relation, given by the perturbation, between the two pairs of Koschmeider's model parameters estimated from the input image and its degraded form.

Our zero-shot approach can be used for the restoration of real-world images where the degradation is due to the scattering of light, and hence, can be formulated through the Koshmieder's model. Using the proposed approach, we restore images captured in hazy conditions and turbid underwater medium, which suffer from such degradations. In the zero-shot optimization, a few generic no-reference loss functions are also employed that are designed for color cast reduction and to avoid pixel value saturation. Subjective and quantitative evaluations on multiple standard real-world image datasets show that our approach mostly outperforms the state-of-the-art of the respective domains.

To highlight, our work contributes in the following ways:

- To the best of our knowledge, the proposed approach is the first that can be used for image restoration in all the application domains where the degradations can be formulated based on the Koschmieder's model.
- To the best of our knowledge, the proposed zero-shot learning approach for dehazing and underwater image restoration is first of a kind, where a prior based loss function or regularizer is not required. Further, our approach is probably the first zero-shot approach for underwater image restoration.
- Despite being a zero-shot approach, our approach outperforms or performs as good as the state-of-the-art in real-world image dehazing and underwater image restoration. We further demonstrate the use of our approach for low-light image enhancement, where the Koschmieder's model can be employed [68].

2. Related Work

2.1. Koschmieder's Model for Image Restoration

Koschmieder's model is one of the most popular models that formulates image degradation due to light scattering and absorption. Most hand-crafted prior-based single image dehazing techniques like [26, 10, 51, 32] estimate the transmission map and atmospheric light guided by the prior's characteristics, and then follow the Koschmieder's model for haze removal. Many deep learning based techniques like [52, 47, 32] either use the Koschmieder's model for generating synthetic hazy images for training their networks or estimate its components for dehazing. Similar to image dehazing, most underwater image restoration approaches like [50, 58, 9, 39, 42, 63] are based on modified versions of the Koschmieder's model taking into account underwater characteristics. The techniques estimate medium energy ratio /transmission map and global background light components of the model for the restoration.

2.2. Zero-shot Learning

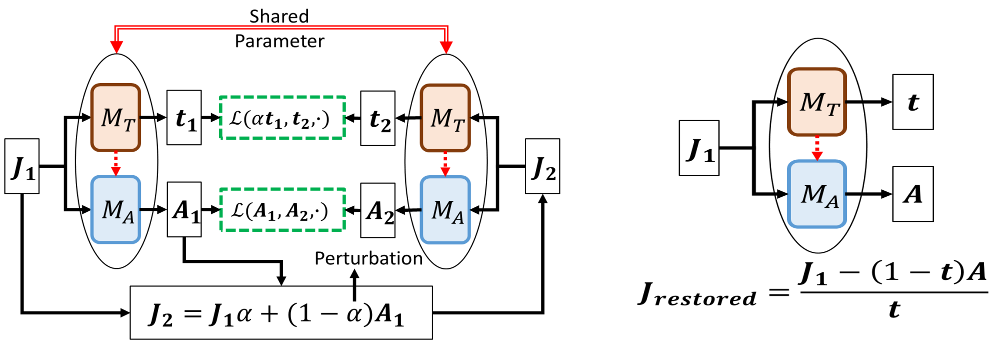
The zero-shot framework learns to perform a particular task from the input image itself. Such a learning is especially beneficial where the ground truth is unavailable or challenging to create, and synthetically generated ground truths suffer from domain-shift issues, such as in image restoration [36]. Due to the absence of any reference, zero-shot learning is one of the most challenging learning approaches. Recently, the approach has drawn significant attention due to its said significance, leading to proposals in the domains of denoising [35], dehazing [36], super-resolution [57], deblurring [53], back-lit image restoration [69] etc. Besides these domain-specific solutions by zero-shot learning, a few proposals like DIP [60] and Double-DIP [21] have been designed to learn low-level image statistics and features from the input image for performing different computer vision tasks.

3. Proposed Zero-Shot Methodology

3.1. Light-scattering Induced Real-world Image Degradation Model

As mentioned in Section 1, our approach is based on the Koschmieder's light scattering model, whose variants describe the degradations in hazy and underwater images. As per Beer–Lambert law [59], light propagation is associated with an attenuation factor $e^{-\beta d}$, where d is the distance from the source and $\beta(>0)$ is the extinction coefficient, which is higher when the density of scattering inducing particles is higher. Guided by the Beer–Lambert law, Koschmieder [34] formulated the effect of scattering of light at a distance d from the source as follows:

$$\mathcal{A}_{\mathcal{L}}(d) = (1 - e^{-\beta d})A \tag{1}$$



(a) Training Model (b) Testing Model

Figure 1. Proposed zero-shot learning framework. J_1 is the input corrupted image and J_2 is the image degraded from J_1 using the perturbation α and the estimated atmospheric /global background light in J_1. M_T is transmission map estimator and M_A is the atmospheric /global background light estimator. Model details are in Section 3.3 and in the supplementary. Red dotted line indicates feature sharing.

where the light intensity of the source is A. Later, the authors of [48] used the existing notion that the scene radiance (J) also gets attenuated by the factor $e^{-\beta d}$ to get:

$$\mathcal{J}_{\mathcal{A}}(d) = Je^{-\beta d} \tag{2}$$

Based on the above effects of light scattering, they formulated the related degradation in an image captured as the additive combination of $\mathcal{J}_{\mathcal{A}}$ and $\mathcal{A}_{\mathcal{L}}$ as follows:

$$\boldsymbol{I}(x) = \boldsymbol{J}(x)e^{-\beta d(x)} + (1 - e^{-\beta d(x)})\boldsymbol{A} \tag{3}$$

where $\boldsymbol{I}$ is the degraded image, $\boldsymbol{J}$ is the corresponding uncorrupted image /scene radiance, x is an image pixel, and $d(x)$ is essentially the scene depth at that pixel. In (3), $t(x) = e^{-\beta d(x)}$ is known as the transmission value /medium energy ratio /light-scattering attenuation rate. For light propagation through the atmosphere, the transmission value $t(x)$ in the degraded image is the same for all the primary color channels [26]. On the other hand, for underwater light propagation the $t(x)$ could be different for the different primary color channels due to the wavelength-selective characteristics of the attenuation ($\boldsymbol{t}$ with different β) [6, 30]. While $\boldsymbol{A}$ is referred to as the atmospheric light when the light propagates through the atmosphere, the attenuated quantity $\mathcal{A}_{\mathcal{L}}$ is called the airlight. On the other hand, $\boldsymbol{A}$ is called the global background light in the case of underwater light propagation, and the attenuated quantity $\mathcal{A}_{\mathcal{L}}$ is called the veiling light or spacelight. The model in (3) is commonly known as the Koschmieder's light scattering model [34], which explains hazy image formation and with certain modifications also explains underwater image formation [30].

3.2. Controlled Model Perturbation for Zero-shot Learning

As the model in (3) for the real-world image degradation is available, a straightforward approach of training deep networks for image restoration when ground truths are unavailable is to create synthetic hazy images using the model. However, it should be noted that such hazy image creation requires accurate knowledge of the scene depth at every image pixel, which if erroneous, could result in inappropriate synthetic hazy images. Alternatively, the model could be used to design a zero-shot learning network based on a judiciously deduced framework for its optimization, which we perform here.

Consider the Koschmieder's model in (3), which we rewrite here as follows:

$$J_1 = Jt_1 + (1 - t_1)A_1 \tag{4}$$

where (x) is dropped and a single color channel is considered for simplicity. In (4), let J_1 be a degraded input image pixel value corresponding to the original uncorrupted image pixel value J. Our objective is to restore the scene radiance J of the uncorrupted image pixel from J_1, for which we are required to estimate t_1 and A_1.

Let us begin by further degrading J_1 using the Koschmieder's model as follows:

$$J_2 = J_1t_2 + (1 - t_2)A_2 \tag{5}$$

Substituting J_1 from (4) in (5), we get:

$$J_2 = Jt_1t_2 + A_2 - t_1A_1t_2 - A_2t_2 + A_1t_2 \tag{6}$$

Now, if we impose a constraint on the degradation from J_1 to J_2 such that $A_1 = A_2 = A$, (6) becomes:

$$J_2 = Jt_1t_2 + (1 - t_1t_2)A \tag{7}$$

Interestingly, it is evident that (7) is the Koschmieder's model that explains the formation of the degraded image J_2 from the original uncorrupted image J, just like the one in (4) which explains the formation of J_1 from J. Moreover, as evident, (7) can be obtained from (4) by perturbing the transmission value t_1 using t_2 into the transmission

(a) Hazy (b) GridDN (c) MSBDN (d) HazeLine (e) DCPLoss (f) ALC (g) ZID (h) Ours

Figure 2. Subjective evaluation of the different dehazing methods on real-world hazy images. Effectiveness in terms of faithful color restoration, and haze and color cast reduction may be observed. Cropped regions are for detailed observation.

Dataset	Techniques and Measures (PSNR/SSIM/CIEDE2000/VI/RI)								
	Supervised			Unsupervised			Zero-Shot		
	AODN [37]	GridDN [47]	MSBDN [15]	DCPLoss [22]	ALC [51]	Haze-Lines [10]	D-DIP [21]	ZID [36]	Ours
I-Haze	14.87/0.585/ **52.35**/0.888/ 0.973	12.24/0.473/ 63.56/0.869/ 0.920	**16.57**/**0.636**/ **49.38**/**0.912**/ **0.974**	14.91/0.585/ 54.09/0.889/ 0.974	14.34/0.553/ 57.01/0.906/ **0.972**	15.48/0.595/ 55.85/0.908/ 0.968	14.22/0.525/ 62.95/0.887/ 0.964	14.01/0.443/ 75.32/0.902/ 0.958	**16.79**/**0.605**/ 57.05/**0.915** /0.971
O-Haze	15.57/0.371/ 66.47/0.824/ 0.962	13.54/0.367/ 68.21/0.765/ 0.915	**16.83**/0.455/ 62.45/0.815/ 0.963	**16.92**/0.475/ **59.57**/0.823/ **0.965**	16.06/0.449/ 63.94/0.866/ 0.962	15.80/**0.522**/ 61.77/**0.880**/ 0.962	14.35/0.391/ 72.94/0.862/ 0.953	14.60/0.377/ 76.68/0.845/ 0.953	16.63/**0.601**/ **59.64**/**0.879**/ **0.968**

Table 1. Quantitative comparison of the different dehazing approaches on standard hazy image datasets. Higher PSNR, SSIM, VI and RI are better, lower CIEDE2000 is better. (Best: Red highlight, Second best: Blue highlight)

value $t_1 t_2$. This finding suggests that a constrained degradation of the input image using the Koschmieder's model amounts to a controlled perturbation of the Koschmieder's model parameters that resulted in the formation of the input image from the uncorrupted one. This property of degradation through the Koschmieder's model can be leveraged for zero-shot training as described below.

We are required to devise a zero-shot learning strategy, where $t' = t_1$ and $A = A_1$ are the parameters to be estimated for image restoration. For that, let $\alpha = t_2$ be a value that is applied as a controlled perturbation to get J_2 from J_1 leveraging the finding discussed earlier, where A of J_1 is used in J_2 as well. So, from (7), (5) and (4), we have:

$$\begin{aligned} J_2 &= J\alpha t' + (1-\alpha t')A \qquad (8)\\ &= J_1\alpha + (1-\alpha)A \\ J_1 &= Jt' + (1-t')A \qquad (9) \end{aligned}$$

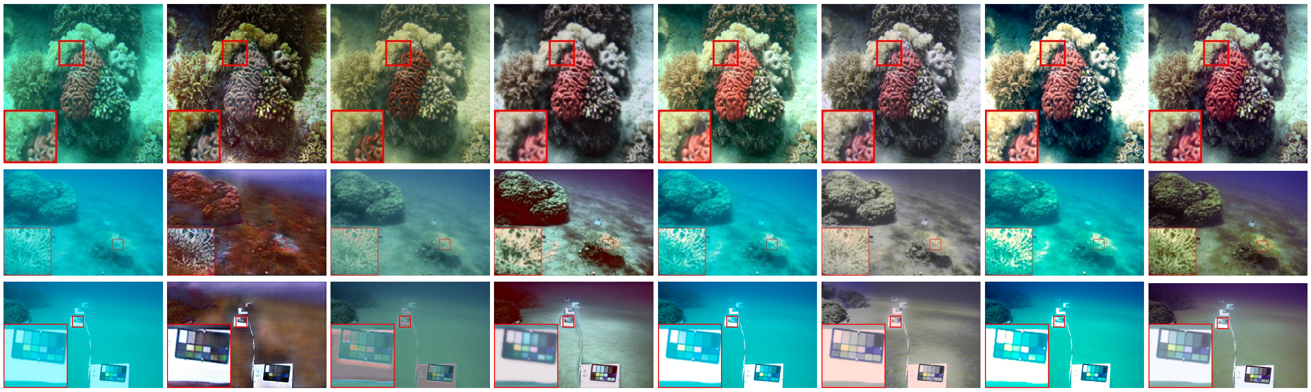
So, if an appropriate estimator of transmission value estimates $\hat{t'}$ as the transmission value from J_1, it must estimate a value very close to $\alpha\hat{t'}$ as the transmission value from J_2. Similarly, if an appropriate estimator of atmospheric light or global background light estimates $\hat{A}$ from J_1, it must estimate a value almost the same from J_2. These relations between the corresponding estimates from J_1 and J_2 are targeted to be maintained through suitable loss functions during the optimization of our zero-shot deep network $\mathfrak{F} = \{M_T, M_A\}$, which yields image-specific estimates of transmission map (by M_T) and atmospheric light /global background light (by M_A).

Let us hence consider all pixels and color channels together. If $(\hat{\boldsymbol{t'_{J_1}}}, \hat{\boldsymbol{A_{J_1}}}) = \mathfrak{F}(\boldsymbol{J_1})$ and $(\hat{\boldsymbol{t'_{J_2}}}, \hat{\boldsymbol{A_{J_2}}}) = \mathfrak{F}(\boldsymbol{J_2})$, then our optimization process minimizes the dissimilarity between $\hat{\boldsymbol{t'_{J_2}}}$ and $\alpha\hat{\boldsymbol{t'_{J_1}}}$, where α is known, along with the dissimilarity between $\hat{\boldsymbol{A_{J_1}}}$ and $\hat{\boldsymbol{A_{J_2}}}$. The optimized estimates $\hat{\boldsymbol{t'_{J_1}}}$ and $\hat{\boldsymbol{A_{J_1}}}$ obtained from the zero-shot deep network are then used to get the restored output $\hat{\boldsymbol{J}}$. Note that, the above zero-shot learning is possible as:

- The controlled perturbation does not add any distortion to the degraded image $\boldsymbol{J_2}$ other than the ones that can be explained by the Koschmieder's model (evident from (8)).
- The controlled perturbation in (8) forming $\boldsymbol{J_2}$ is related to the degradation of $\boldsymbol{J_1}$ to $\boldsymbol{J2}$ through the Koschmieder's model (evident from (4)-(7)).

3.3. The Zero-shot Network Architecture

Figure 1 shows our proposed zero-shot training and testing framework. $\boldsymbol{J_1}$ is the input corrupted image, M_T is the transmission map estimate model, and M_A is the atmospheric light /global background light estimation model. During training, at each iteration, we estimate transmission map $\boldsymbol{t_1}$ and atmospheric light /global background light $\boldsymbol{A_1}$ from image $\boldsymbol{J_1}$. Image $\boldsymbol{J_2}$ is then obtained from the input



(a) Input (b) All-in-One (c) UWCNN (d) Retinex (e) IBLA (f) ColorFusion (g) Statistical (h) Ours

Figure 3. Subjective evaluation of the different restoration methods on real-world underwater images. Effectiveness in terms of faithful color restoration, contrast improvement and color cast reduction may be observed. Cropped regions are for detailed observation.

Dataset	Measures	Techniques							
			Supervised		Unsupervised				Zero-Shot
		Input	AllinOne [61]	UWCNN [39]	Retinex [19]	IBLA [50]	ColorFusion [6]	Statistical [58]	Ours
U45	UIQM/ UCIQE	2.35/ 17.3	**5.72**/ 27.5	4.30/ 19.7	4.94/ **27.6**	2.94/ 24.9	5.05/ 19.6	**5.75**/ 23.3	4.97/ **29.5**
Challenging-60	UIQM/ UCIQE	0.68/ 20.4	**3.64**/ 29.2	2.0/ 19.9	2.82/ 24.2	1.19/ **29.6**	2.73/ 20.8	1.68/ 27.7	**3.81**/ **29.3**
Stereo	UIQM/ UCIQE	-1.63/ 15.5	**4.73 28.7**	1.65/ 17.2	3.67/ 26.0	-1.14/ 21.8	**4.19**/ 17.0	-1.54/ 21.0	3.84/ **28.4**

Table 2. Quantitative comparison of the different underwater image restoration approaches on standard underwater image datasets. Higher UIQM and UCIQE are better. (Best: Red highlight, Second best: Blue highlight)

image $\boldsymbol{J_1}$ using a fixed transmission map α and the estimated $\boldsymbol{A_1}$. $\boldsymbol{J_2}$ is the degraded image after the model perturbation, from which we estimate the transmission map $\boldsymbol{t_2}$ and atmospheric light /global background light $\boldsymbol{A_2}$. The objective is to maintain the relation given by the perturbation between the $\boldsymbol{t_1}$ and $\boldsymbol{t_2}$ estimates along with the similarity between the $\boldsymbol{A_1}$ and $\boldsymbol{A_2}$ estimates. After training, we use the estimates $\boldsymbol{t}$ and $\boldsymbol{A}$ to compute the restored image $\boldsymbol{J_{\text{restored}}}$ as shown in Figure 1.

We use overall the same network topology across the different applications discussed in Sections 4 and 5 for the M_T and M_A estimators. The topology for M_T is based on color channel-wise multi-scale feature extraction and feature selection [17, 45, 52]. In the topology for M_A, intermediate features from M_T are used as multi-scale feature attention [28] in a convolution network. Due to space constraint, the detailed description of the networks are given in the supplementary.

3.4. Loss Functions

The formulation of our zero-shot approach is based on the imposition of the relation between the transmission maps and similarity between the atmospheric lights /global background lights estimated from $\boldsymbol{J_1}$ (9) and $\boldsymbol{J_2}$ (8). Therefore, transmission relation loss $\mathcal{L}_{TR}$ and light similarity loss $\mathcal{L}_{LS}$ are correspondingly employed. In addition to the above losses, a few other no-reference losses are used as described below, all of which used together additively.

Transmission Relation Loss: Transmission relation loss $\mathcal{L}_{TR}$ is one of the two primary loss functions proposed. It computes the extent to which the relation (as explained in Section 3.2) between the estimates of $\boldsymbol{t_1}$ and $\boldsymbol{t_2}$ are not maintained. As a transmission map contains pixel-level information, we compute a pixel-wise loss as follows:

$$\mathcal{L}_{TR} = \sum_{x} ||\alpha \hat{\boldsymbol{t_1}}(x) - \hat{\boldsymbol{t_2}}(x)||_2{}^2 \tag{10}$$

where x represents a pixel, $\hat{\boldsymbol{t_1}} = M_T(\boldsymbol{J_1})$, and $\hat{\boldsymbol{t_2}} = M_T(\boldsymbol{J_2})$ with M_T representing our transmission map estimator.

Light Similarity Loss: The other primary loss function, light similarity loss $\mathcal{L}_{LS}$ measures the dissimilarity between the estimates of $\boldsymbol{A_1}$ and $\boldsymbol{A_2}$, which is formulated as:

$$\mathcal{L}_{LS} = ||\hat{\boldsymbol{A_1}} - \hat{\boldsymbol{A_2}}||_2{}^2 \tag{11}$$

where $\hat{\boldsymbol{A_1}} = M_A(\boldsymbol{J_1})$ and $\hat{\boldsymbol{A_2}} = M_A(\boldsymbol{J_2})$ with M_A representing our atmospheric /global background light estimator.

The above two losses are at the heart of our zero-shot framework and directly deduced from its design.

Saturated Pixel Penalty: The third set of losses proposed is saturated pixel penalty, which is used to constrain the optimization to a subspace of the solution space avoiding overflow /underflow that results in pixel value saturation [25].

These losses are designed to avoid operations such as clipping so that there is no hindrance to gradient flow.

The no-reference pure white $\mathcal{L}_{SPW}$ and pure black $\mathcal{L}_{SPB}$ saturation penalties used are as follows:

$$\mathcal{L}_{SPW} = \sum_{x,c} \left(\max(\hat{J}^c(x), 1) + \max(\hat{J}_{\prime}^c(x), 1) \right) \quad (12)$$

$$\mathcal{L}_{SPB} = -\sum_{x,c} \left(\min(\hat{J}^c(x), 0) + \min(\hat{J}_{\prime}^c(x), 0) \right) \quad (13)$$

where x represents an image pixel, c represents a primary color channel, $\hat{\boldsymbol{J}}$ and $\hat{\boldsymbol{J}}_{\prime}$ are respectively the estimates of the original scene radiance from $\boldsymbol{J_1}$ (input image) and $\boldsymbol{J_2}$. With the range of $\hat{J}^c(x)$ and $\hat{J}_{\prime}^c(x)$ being $[0, 1]$, it can be seen that both the penalties are at the minimum for a pixel that is not saturated and are related linearly to the estimates otherwise.

Gray-world Assumption Loss: For imbibing the characteristics of non-cast natural images in the restored output, a fourth loss is proposed based on the Gray-world assumption [11] of natural image statistics. Use of this no-reference loss particularly facilitates diminishing of color cast in the restored image. This loss is computed as:

$$\mathcal{L}_{GW} = \sum_{(c_1,c_2)\in\Omega} |\mu(\hat{\boldsymbol{J}^{c_1}}) - \mu(\hat{\boldsymbol{J}^{c_2}})| \quad (14)$$

where $\Omega = \{(R, G), (R, B), (G, B)\}$ is a set of color pairs with red (R), Green (G) and Blue (B) colors and $\mu(\hat{\boldsymbol{J}}^{\boldsymbol{c}})$ represents the mean of the estimated uncorrupted image's color channel c. As can be seen the loss penalizes deviation from the Gray-world assumption, which is that the mean vector of a non-cast natural color images is achromatic.

Total Variation Loss: Finally, total variation [55] loss $\mathcal{L}_{TV}$ is proposed to invoke neighborhood consistency in the restored image facilitating reduction of spurious local changes. This no-reference loss is defined as

$$\mathcal{L}_{TV} = \sum_{x,c} \left(|\nabla_{\mathcal{H}} \hat{J}^c(x)|^2 + |\nabla_{\mathcal{V}} \hat{J}^c(x)|^2 \right) \quad (15)$$

where x represents an image pixel, c represents a primary color channel, $\nabla_{\mathcal{H}}$ and $\nabla_{\mathcal{V}}$ represent the horizontal and vertical gradients, and $\hat{\boldsymbol{J}}$ is the estimate of the original uncorrupted image.

It should be noted, that the above proposed no-reference losses are generic in nature and not domain /application-specific, and they facilitate convergence of the proposed approach to produce real-world-like restored images.

3.5. Training Procedure

The loss function used in our training is $\mathcal{L} = \omega_1\mathcal{L}_{TR} + \omega_2\mathcal{L}_{LS} + \omega_3\mathcal{L}_{SPB} + \omega_4\mathcal{L}_{SPW} + \omega_5\mathcal{L}_{GW} + \omega_6\mathcal{L}_{TV}$. The weight values are discussed in the supplementary. As J_2 is obtained by further degrading J_1, appropriate range of the perturbation α is $(0, 1)$. From Section 3.2, it is evident α can take any value in the mentioned range. We fix $\alpha = 0.9$ across all applications. Data augmentation is considered by transforming the input image using 8 geometrical transformations through 4 rotations by $90°$ combined with vertical and horizontal mirror reflection, which has been found useful for unsupervised internal learning [57]. Model weight values are drawn from normal distribution with zero mean and standard deviation is set to 0.001. We do not use bias in any layer of the model. The optimization process is done using the ADAM optimizer [33] with default parameters in PyTorch environment and learning rate is set to 10^{-3}. The model is trained for 10000 iterations on an NVIDIA 2080Ti GPU. The time analysis is discussed in the supplementary.

4. Single Image Dehazing

We perform Image dehazing using the proposed zero-shot framework. Haze in an image $\boldsymbol{I}$ is described by the Koschmieder's model (3), where the pixel-wise transmission map $\boldsymbol{t}$ is same for all the primary color channels and the atmospheric light $\boldsymbol{A}$ is a three element vector representing a color. Chromatic shift in the image $\boldsymbol{I}$ due to $\boldsymbol{A}$ is responsible for color cast in hazy images [26]. The dehazed image $\boldsymbol{J}_{\textbf{restored}}$ is obtained as described in Section 3.3, where the architecture of our dehazing model is also mentioned, which is further elaborated in the supplementary along with the hyperparameters used. Among the loss functions used (see Section 3.4), $\mathcal{L}_{GW}$ particularly contributes to color cast reduction during the dehazing.

Table 1 presents the quantitative performance comparison of single image dehazing techniques on two standard hazy image datasets, I-Haze [4] for indoor images and O-Haze [7] for outdoor images. Visibility Index (VI) and Realness Index (RI) that are specifically designed to evaluate dehazing [70], CIEDE2000 [56] that measures faithful color restoration, and the popular PSNR and SSIM [65] measures are listed for our approach and the techniques AODN [37], GriDN [47], MSBDN [15], DCPLoss [22], ALC [51], Haze-Lines [10], D-DIP [21], and ZID [36], which are categorized into supervised, unsupervised and zero-shot approaches. Higher the values of VI, RI, PSNR, SSIM, and lower the value of CIEDE2000, better is the performance. The results shows that our approach outperforms the other zero-shot techniques. Despite being a zero-shot approach, our technique either outperforms or produces results comparable to the best performing techniques in both the supervised and unsupervised categories. Figure 2 presents the subjective comparison of the above mentioned techniques, where we find our approach outperforms the rest in terms of dehazing, faithful color restoration and effective color cast reduction.

(a) Input (b) Using $\mathcal{L}$ (c) w/o $\mathcal{L}_{GW}$ (d) w/o $\mathcal{L}_{TR}$&$\mathcal{L}_{LS}$ (e) w/o $\mathcal{L}_{SPW}$&$\mathcal{L}_{SPB}$

Figure 4. A study of importance of the different loss functions in our zero-shot learning framework. Haze reduction, color cast reduction and pixel saturation prevention may be noted.

5. Underwater Image Restoration

We also perform underwater image restoration using the proposed zero-shot framework. Underwater images with reduced contrast and color cast, is degraded due to wavelength-selective light scattering [5, 1, 20] that can be described by a modified Koschmieder's model [30]. In the model of (3), due to the wavelength-selectivity, the transmission map can be different for the different primary color channels, and the global background light replaces the atmospheric light. In an image captured underwater, it is well-known that the intensity of red color channel is severely diminished due to depth-dependent attenuation of higher wavelength light [5, 20]. Additionally, the green channel is known to reliably contain details about the scene content [5]. Due to these properties, as suggested in [5, 20], a red channel compensation is required in underwater images before the use of the Gray world assumption for restoration, which can be achieved using the green channel. As proposed in [5], the compensation factor to be applied on the red channel of the underwater image I before restoration is given by:

$$CF(x) = \left(\mu(\boldsymbol{I^G}) - \mu(\boldsymbol{I^R})\right) \bar{I}^R(x) I^G(x) \qquad (16)$$

where x represents an image pixel, $\mu(\boldsymbol{I^G})$ and $\mu(\boldsymbol{I^R})$ respectively represent means of the red and green channels of the degraded image and $\bar{\boldsymbol{I}}^{\boldsymbol{R}}$ represent the negative of the red channel. The compensated red channel is then obtained as $\boldsymbol{I^R} + \boldsymbol{CF}$ which then replaces $\boldsymbol{I^R}$ of $\boldsymbol{I}$. The modified $\boldsymbol{I}$ is used as $\boldsymbol{J_1}$ in our zero-shot framework described in Section 3 whose loss functions are given in Section 3.4. The architecture of our model is also mentioned in Section 3.3 and elaborated in the supplementary along with the hyper-parameters used.

Table 2 presents quantitative performance comparison of underwater image restoration techniques on three standard real-world underwater image datasets, U45 [43], Challenging-60 [40] and Stereo [9]. Underwater image specific quality measures UIQM [49] and UCIQE [67] are listed for our approach and the techniques All-in-One [61], UWCNN [39], Retinex [19], IBLA [50], ColorFusion [6], Statistical [58], which are classified into supervised, unsupervised and zero-shot categories. To the best of our knowledge, our proposal is the first in the zero-shot category. Higher the values of UIQM and UCIQE, better is the performance. As can be seen, our approach in spite of being a zero-shot network performs better or as good as the best performing supervised and unsupervised techniques. Figure 3 presents the subjective comparison of the aforesaid techniques where it is seen that our approach outperforms the rest in terms of distortion-free natural color restoration, contrast improvement and color cast reduction.

6. Additional Experiments

(a) Input ($\boldsymbol{I}$) (b) T-map ($\boldsymbol{t}$) (c) A/GB-light($\boldsymbol{A}$) (d) Output ($\boldsymbol{J}$)

Figure 5. Visualization of our approach's estimates of $\boldsymbol{t}$ and $\boldsymbol{A}$ (See expression (3)), which are seen to have expected physical characteristics.

6.1. Ablation Studies

Using an example of dehazing in Figure 4, we study the importance of using the different of loss functions in our model. Transmission relation $\mathcal{L}_{TR}$ and light similarity $\mathcal{L}_{LS}$ losses mainly drive our zero-shot framework, and as can been seen from Figure 4(d), haze does not reduce without them. Diminishing of color cast is due to Gray-world assumption loss $\mathcal{L}_{GW}$ and as evident Figure 4(c), cast is not reduced without it. Saturated pixel penalties $\mathcal{L}_{SPW}$ and $\mathcal{L}_{SPB}$ are important losses that ensures that the dehazed image pixels are not saturated, which is clearly seen in Figure 4(e) for which it is not used.

6.2. The Estimates in Our Restoration Approach

Our zero-shot learning approach is aimed at estimating $\boldsymbol{t}$ and $\boldsymbol{A}$ from the input $\boldsymbol{I}$ to obtain $\boldsymbol{J}$ as given in the Koschmieder's model of (3). Figure 5 shows a couple of examples of different kinds of images and their estimated $\boldsymbol{t}$, $\boldsymbol{A}$, and corresponding restored $\boldsymbol{J}$. While $\boldsymbol{t}$ is depth dependent in hazy images, it is dependent on both the depth and color (wavelength) [30]. The atmospheric light or global

(a) Input (b) DRD (c) LightenNet (d) LIME (e) SRIE (f) RRM (g) ZeroDCE (h) Ours

Figure 6. Subjective evaluation of the different low-light image enhancement methods on real-world low-light images.

Dataset	Techniques and Measures (PSNR/SSIM/CIEDE2000)									
	Supervised			Unsupervised						Zero-Shot
	DRD [66]	LightenNet [41]	PR [62]	LIME [24]	SRIE [18]	RRM [44]	LECARM [54]	ALSM [64]	ZeroDCE [23]	Ours
LOL	16.81/0.560/ 52.23	10.28/0.361/ 81.80	**18.80**/**0.721**/ –	17.23/0.635/ 53.95	11.87/0.498/ 69.35	13.88/0.657/ **51.81**	14.43/0.569/ 62.76	17.19/0.568/ 58.82	14.87/0.585/ 60.42	**17.50**/**0.695**/ **46.87**

Table 3. Quantitative comparison of the different low-light image enhancement approaches on a standard low-light image dataset. Higher PSNR, SSIM are better, lower CIEDE2000 is better. (Best: Red highlight, Second best: Blue highlight)

background light $\boldsymbol{A}$ is expected to correspond to the brightest light in the scene [26, 42]. As can be seen from the figure, the parameters $\boldsymbol{t}$ and $\boldsymbol{A}$ correspond very well with the expected physical characteristics mentioned above.

6.3. Low-light Image Enhancement

Here, we demonstrated the use of our zero-shot framework for low-light image enhancement. Koschmieder's model has been employed a few times in literature [16, 68] for low-light image enhancement. In a significant change to the model, the atmospheric light /global background light component of the model in (3) is replaced by a pixel-wise map for all the three primary color channels [68]. The transmission map in the model is an achromatic pixel-wise map [68] similar to that used for dehazing. Therefore, when our zero-shot framework of Section 3.2 is employed for low-light image enhancement, the estimated A_1 and A_2 needs to be pixel-wise three-channel quantity. Among the losses used in our framework (see Section 3.4), $\mathcal{L}_{GW}$, which facilitates color cast reduction as low-light image seldom contains color cast. Further, in $\mathcal{L}_{TR}$, L_1 loss is considered in place of L_2 loss as transmission values associated with low-light images are near zero in low-light areas, where L_1 loss provides larger gradients for learning than L_2 loss. A study related to this is given in the supplementary. The architecture of our network model for the enhancement is similar to that mentioned in Section 3.3, except that both the Koschmieder's model parameters are estimated pixel-wise, which is elaborated in the supplementary along with the hyperparameters used. As low-light images are prone to noise, like many other approaches [24], we use BM3D [13] to reduce noise in the enhanced image.

Table 3 presents quantitative performance comparison of low-light image enhancement on the standard LOL dataset [66] using our approach, and the techniques DRD [66], LightenNet [41], PR [62], LIME [24], SRIE [18], RRM [44], LECARM [54], ALSM [64], and ZeroDCE [23], which are segregated into supervised, unsupervised and zero-shot categories. PSNR and SSIM [65] and CIEDE2000 [56] are used for the evaluation. As can be seen from the table, our approach outperforms all the approaches except the supervised approach PR. Figure 6 presents the subjective comparison where our approach is found to perform as good as any other.

7. Conclusion

This paper contributes a zero-shot framework for the restoration of images whose degradation is described by the Koschmieder's light scattering model. The framework is designed based on the theoretical finding that a further degradation of the corrupted input image amounts only to a perturbation in the Koschmieder's model that explains the corruption in the input image, and hence, keeps its functional form intact. The proposed approach is used for single image dehazing and underwater image restoration, where it outperforms or performs as good as the state-of-the-art in the respective domains. The potential use of our framework for low-light image enhancement is also demonstrated. The success of our zero-shot framework suggests that the preservation of a model's functional form through successive transformations can be leveraged for zero-shot training. This paves the way for further investigation and design of similar frameworks for other applications, where such a preservation of functional forms is evident.

References

[1] Derya Akkaynak and Tali Treibitz. A revised underwater image formation model. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pages 6723–6732, 2018. 7

[2] Derya Akkaynak and Tali Treibitz. Sea-thru: A method for removing water from underwater images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1682–1691, 2019. 1

[3] Cosmin Ancuti, Codruta Orniana Ancuti, Tom Haber, and Philippe Bekaert. Enhancing underwater images and videos by fusion. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 81–88. IEEE, 2012. 1

[4] Cosmin Ancuti, Codruta O Ancuti, Radu Timofte, and Christophe De Vleeschouwer. I-haze: a dehazing benchmark with real hazy and haze-free indoor images. In *International Conference on Advanced Concepts for Intelligent Vision Systems*, pages 620–631. Springer, 2018. 6

[5] Codruta O Ancuti, Cosmin Ancuti, Christophe De Vleeschouwer, and Philippe Bekaert. Color balance and fusion for underwater image enhancement. *IEEE Transactions on Image Processing*, 27(1):379–393, 2018. 1, 7

[6] Codruta O Ancuti, Cosmin Ancuti, Christophe De Vleeschouwer, and Philippe Bekaert. Color balance and fusion for underwater image enhancement. *IEEE Transactions on image processing*, 27(1):379–393, 2018. 3, 5, 7

[7] Codruta O Ancuti, Cosmin Ancuti, Radu Timofte, and Christophe De Vleeschouwer. O-haze: a dehazing benchmark with real hazy and haze-free outdoor images. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 754–762, 2018. 6

[8] Yuval Bahat and Michal Irani. Blind dehazing using internal patch recurrence. In *2016 IEEE International Conference on Computational Photography (ICCP)*, pages 1–9. IEEE, 2016. 2

[9] Dana Berman, Deborah Levy, Shai Avidan, and Tali Treibitz. Underwater single image color restoration using haze-lines and a new quantitative dataset. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020. 1, 2, 7

[10] Dana Berman, Tali Treibitz, and Shai Avidan. Single image dehazing using haze-lines. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(3):720–734, 2020. 1, 2, 4, 6

[11] Gershon Buchsbaum. A spatial processor model for object colour perception. *Journal of the Franklin institute*, 310(1):1–26, 1980. 6

[12] John Y Chiang and Ying-Ching Chen. Underwater image enhancement by wavelength compensation and dehazing. *IEEE Transactions on Image Processing*, 21(4):1756–1769, 2011. 1

[13] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Color image denoising via sparse 3d collaborative filtering with grouping constraint in luminance-chrominance space. In *2007 IEEE International Conference on Image Processing*, volume 1, pages I–313. IEEE, 2007. 8

[14] Sobhan Kanti Dhara, Mayukh Roy, Debashis Sen, and Prabir Kumar Biswas. Color cast dependent image dehazing via adaptive airlight refinement and non-linear color balancing. *IEEE Transactions on Circuits and Systems for Video Technology*, 2020. 1

[15] Hang Dong, Jinshan Pan, Lei Xiang, Zhe Hu, Xinyi Zhang, Fei Wang, and Ming-Hsuan Yang. Multi-scale boosted dehazing network with dense feature fusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, June 2020. 1, 4, 6

[16] Xuan Dong, Guan Wang, Yi Pang, Weixin Li, Jiangtao Wen, Wei Meng, and Yao Lu. Fast efficient algorithm for enhancement of low lighting video. In *2011 IEEE International Conference on Multimedia and Expo*, pages 1–6. IEEE, 2011. 8

[17] Akshay Dudhane, Kuldeep M. Biradar, Prashant W. Patil, Praful Hambarde, and Subrahmanyam Murala. Varicolored image de-hazing. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020. 5

[18] Xueyang Fu, Delu Zeng, Yue Huang, Xiao-Ping Zhang, and Xinghao Ding. A weighted variational model for simultaneous reflectance and illumination estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2782–2790, 2016. 8

[19] Xueyang Fu, Peixian Zhuang, Yue Huang, Yinghao Liao, Xiao-Ping Zhang, and Xinghao Ding. A retinex-based enhancing approach for single underwater image. In *2014 IEEE International Conference on Image Processing (ICIP)*, pages 4572–4576. IEEE, 2014. 1, 5, 7

[20] Adrian Galdran, David Pardo, Artzai Picón, and Aitor Alvarez-Gila. Automatic red-channel underwater image restoration. *Journal of Visual Communication and Image Representation*, 26:132–145, 2015. 7

[21] Yossi Gandelsman, Assaf Shocher, and Michal Irani. "double-dip": Unsupervised image decomposition via coupled deep-image-priors. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 6, page 2, 2019. 1, 2, 4, 6

[22] Alona Golts, Daniel Freedman, and Michael Elad. Unsupervised single image dehazing using dark channel prior loss. *IEEE Transactions on Image Processing*, 29:2692–2701, 2020. 1, 4, 6

[23] Chunle 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/CVF Conference on Computer Vision and Pattern Recognition*, pages 1780–1789, 2020. 8

[24] 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, 2016. 8

[25] Nicolas Hautiere, Jean-Philippe Tarel, Didier Aubert, and Eric Dumont. Blind contrast enhancement assessment by gradient ratioing at visible edges. *Image Analysis & Stereology*, 27(2):87–95, 2008. 5

[26] Kaiming He, Jian Sun, and Xiaoou Tang. Single image haze removal using dark channel prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(12):2341–2353, 2010. 1, 2, 3, 6, 8

[27] Ming Hong, Yuan Xie, Cuihua Li, and Yanyun Qu. Distilling image dehazing with heterogeneous task imitation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3462–3471, 2020. 1

[28] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7132–7141, 2018. 5

[29] Md Jahidul Islam, Youya Xia, and Junaed Sattar. Fast underwater image enhancement for improved visual perception. *IEEE Robotics and Automation Letters*, 5(2):3227–3234, 2020. 1

[30] Jules S Jaffe. Computer modeling and the design of optimal underwater imaging systems. *IEEE Journal of Oceanic Engineering*, 15(2):101–111, 1990. 1, 3, 7

[31] Aupendu Kar, Sobhan Kanti Dhara, Debashis Sen, and Prabir Kumar Biswas. Transmission map and atmospheric light guided iterative updater network for single image dehazing. *arXiv preprint arXiv:2008.01701*, 2020. 1

[32] Se Eun Kim, Tae Hee Park, and Il Kyu Eom. Fast single image dehazing using saturation based transmission map estimation. *IEEE Transactions on Image Processing*, 29:1985–1998, 2020. 1, 2

[33] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 6

[34] Harald Koschmieder. Theorie der horizontalen sichtweite. *Beitrage zur Physik der freien Atmosphare*, pages 33–53, 1924. 1, 2, 3

[35] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila. Noise2noise: Learning image restoration without clean data. In *International Conference on Machine Learning*, pages 2965–2974, 2018. 1, 2

[36] Boyun Li, Yuanbiao Gou, Jerry Zitao Liu, Hongyuan Zhu, Joey Tianyi Zhou, and Xi Peng. Zero-shot image dehazing. *IEEE Transactions on Image Processing*, 29:8457–8466, 2020. 1, 2, 4, 6

[37] Boyi Li, Xiulian Peng, Zhangyang Wang, Jizheng Xu, and Dan Feng. Aod-net: All-in-one dehazing network. In *Proceedings of the IEEE international conference on computer vision*, pages 4770–4778, 2017. 1, 4, 6

[38] Boyi Li, Wenqi Ren, Dengpan Fu, Dacheng Tao, Dan Feng, Wenjun Zeng, and Zhangyang Wang. Benchmarking single-image dehazing and beyond. *IEEE Transactions on Image Processing*, 28(1):492–505, 2019. 1

[39] Chongyi Li, Saeed Anwar, and Fatih Porikli. Underwater scene prior inspired deep underwater image and video enhancement. *Pattern Recognition*, 98:107038, 2020. 1, 2, 5, 7

[40] Chongyi Li, Chunle Guo, Wenqi Ren, Runmin Cong, Junhui Hou, Sam Kwong, and Dacheng Tao. An underwater image enhancement benchmark dataset and beyond. *IEEE Transactions on Image Processing*, 29:4376–4389, 2019. 7

[41] Chongyi Li, Jichang Guo, Fatih Porikli, and Yanwei Pang. Lightennet: a convolutional neural network for weakly illuminated image enhancement. *Pattern Recognition Letters*, 104:15–22, 2018. 8

[42] Chong-Yi Li, Ji-Chang Guo, Run-Min Cong, Yan-Wei Pang, and Bo Wang. Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior. *IEEE Transactions on Image Processing*, 25(12):5664–5677, 2016. 1, 2, 8

[43] Hanyu Li, Jingjing Li, and Wei Wang. A fusion adversarial underwater image enhancement network with a public test dataset. *arXiv preprint arXiv:1906.06819*, 2019. 7

[44] Mading Li, Jiaying Liu, Wenhan Yang, Xiaoyan Sun, and Zongming Guo. Structure-revealing low-light image enhancement via robust retinex model. *IEEE Transactions on Image Processing*, 27(6):2828–2841, 2018. 8

[45] Xiang Li, Wenhai Wang, Xiaolin Hu, and Jian Yang. Selective kernel networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 510–519, 2019. 5

[46] Risheng Liu, Xin Fan, Ming Zhu, Minjun Hou, and Zhongxuan Luo. Real-world underwater enhancement: Challenges, benchmarks, and solutions under natural light. *IEEE Transactions on Circuits and Systems for Video Technology*, 2020. 1

[47] Xiaohong Liu, Yongrui Ma, Zhihao Shi, and Jun Chen. Griddehazenet: Attention-based multi-scale network for image dehazing. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 7314–7323, 2019. 1, 2, 4, 6

[48] Srinivasa G Narasimhan and Shree K Nayar. Contrast restoration of weather degraded images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(6):713–724, 2003. 1, 3

[49] Karen Panetta, Chen Gao, and Sos Agaian. Human-visual-system-inspired underwater image quality measures. *IEEE Journal of Oceanic Engineering*, 41(3):541–551, 2015. 7

[50] Yan-Tsung Peng and Pamela C Cosman. Underwater image restoration based on image blurriness and light absorption. *IEEE Transactions on Image Processing*, 26(4):1579–1594, 2017. 1, 2, 5, 7

[51] Yan-Tsung Peng, Zhihui Lu, Fan-Chieh Cheng, Yalun Zheng, and Shih-Chia Huang. Image haze removal using airlight white correction, local light filter, and aerial perspective prior. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(5):1385–1395, 2020. 1, 2, 4, 6

[52] Yanyun Qu, Yizi Chen, Jingying Huang, and Yuan Xie. Enhanced pix2pix dehazing network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8160–8168, 2019. 1, 2, 5

[53] Dongwei Ren, Kai Zhang, Qilong Wang, Qinghua Hu, and Wangmeng Zuo. Neural blind deconvolution using deep priors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3341–3350, 2020. 1, 2

[54] Yurui Ren, Zhenqiang Ying, Thomas H Li, and Ge Li. Lecarm: Low-light image enhancement using the camera response model. *IEEE Transactions on Circuits and Systems for Video Technology*, 29(4):968–981, 2019. 8

[55] Paul Rodríguez. Total variation regularization algorithms for images corrupted with different noise models: a review. *Journal of Electrical and Computer Engineering*, 2013. 6

[56] Gaurav Sharma, Wencheng Wu, and Edul N Dalal. The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Research & Application*, 30(1):21–30, 2005. 6, 8

[57] Assaf Shocher, Nadav Cohen, and Michal Irani. “zero-shot” super-resolution using deep internal learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3118–3126, 2018. 1, 2, 6

[58] Wei Song, Yan Wang, Dongmei Huang, Antonio Liotta, and Cristian Perra. Enhancement of underwater images with statistical model of background light and optimization of transmission map. *IEEE Transactions on Broadcasting*, 66(1):153–169, 2020. 1, 2, 5, 7

[59] Donald F Swinehart. The beer-lambert law. *Journal of Chemical Education*, 39(7):333, 1962. 2

[60] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Deep image prior. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 9446–9454, 2018. 2

[61] Pritish M Uplavikar, Zhenyu Wu, and Zhangyang Wang. All-in-one underwater image enhancement using domain-adversarial learning. In *CVPR Workshops*, pages 1–8, 2019. 1, 5, 7

[62] Yang Wang, Yang Cao, Zheng-Jun Zha, Jing Zhang, Zhiwei Xiong, Wei Zhang, and Feng Wu. Progressive retinex: Mutually reinforced illumination-noise perception network for low-light image enhancement. In *Proceedings of the 27th ACM International Conference on Multimedia*, pages 2015–2023, 2019. 8

[63] Yi Wang, Hui Liu, and Lap-Pui Chau. Single underwater image restoration using adaptive attenuation-curve prior. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 65(3):992–1002, 2017. 1, 2

[64] Yun-Fei Wang, He-Ming Liu, and Zhao-Wang Fu. Low-light image enhancement via the absorption light scattering model. *IEEE Transactions on Image Processing*, 28(11):5679–5690, 2019. 8

[65] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004. 6, 8

[66] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep retinex decomposition for low-light enhancement. In *British Machine Vision Conference*, 2018. 8

[67] Miao Yang and Arcot Sowmya. An underwater color image quality evaluation metric. *IEEE Transactions on Image Processing*, 24(12):6062–6071, 2015. 7

[68] Shun-Yuan Yu and Hong Zhu. Low-illumination image enhancement algorithm based on a physical lighting model. *IEEE Transactions on Circuits and Systems for Video Technology*, 29(1):28–37, 2019. 2, 8

[69] Lin Zhang, Lijun Zhang, Xiao Liu, Ying Shen, Shaoming Zhang, and Shengjie Zhao. Zero-shot restoration of back-lit images using deep internal learning. In *Proceedings of the 27th ACM International Conference on Multimedia*, pages 1623–1631, 2019. 2

[70] Shiyu Zhao, Lin Zhang, Shuaiyi Huang, Ying Shen, and Shengjie Zhao. Dehazing evaluation: Real-world benchmark datasets, criteria and baselines. *IEEE Transactions on Image Processing*, 2020. 6