Diving into Haze-Lines: Color Restoration of Underwater Images

Dana Berman[1]
www.eng.tau.ac.il/~berman
Tali Treibitz[2]
marsci.haifa.ac.il/profiles/tTreibitz
Shai Avidan[1]
www.eng.tau.ac.il/~avidan

[1] Tel Aviv University
Israel
[2] University of Haifa
Israel

Abstract

Underwater images suffer from color distortion and low contrast, because light is attenuated as it propagates through water. The attenuation varies with wavelength and depends both on the properties of the water body in which the image was taken and the 3D structure of the scene, making it difficult to restore the colors. Existing single underwater image enhancement techniques either ignore the wavelength dependency of the attenuation, or assume a specific spectral profile. We propose a new method that takes into account multiple spectral profiles of different water types, and restores underwater scenes from a single image.

We show that by estimating just two additional global parameters - the attenuation ratios of the blue-red and blue-green color channels - the problem of underwater image restoration can be reduced to single image dehazing, where all color channels have the same attenuation coefficients. Since we do not know the water type ahead of time, we try different parameter sets out of an existing library of water types. Each set leads to a different restored image and the one that best satisfies the Gray-World assumption is chosen. The proposed single underwater image restoration method is fully automatic and is based on a more comprehensive physical image formation model than previously used.

We collected a dataset of real images taken in different locations with varying water properties and placed color charts in the scenes. Moreover, to obtain ground truth, the 3D structure of the scene was calculated based on stereo imaging. This dataset enables a quantitative evaluation of restoration algorithms on natural images and shows the advantage of the proposed method.

1 Introduction

Underwater images often lack contrast and depict inaccurate colors due to scattering and absorption of light while it propagates in the water. Yet color and contrast are extremely important for visual surveys in the ocean. For example, enhanced images can improve automatic segmentation, increase the accuracy of feature matching between images taken from multiple viewpoints, and aid in navigation.

The attenuation of light depends both on the light's wavelength and the distance it travels [2, 23]. The wavelength-dependent attenuation causes color distortions that increase with an object's distance. In addition, the scattering induces a distance-dependent additive

component on the scene which reduces contrast. These phenomena cannot be globally corrected since the color degradation depends on the distance of the object from the camera. Moreover, the attenuation parameters are affected by seasonal, geographical, and climate variations. These variations were categorized into 10 water types by Jerlov [19].

Unfortunately, existing single underwater image enhancement techniques under-perform as they do not take into account the diverse spectral properties of water. In addition, their evaluation is generally based on a handful of images, and is mostly qualitative. A few methods have been evaluated using no-reference image quality metrics, which operate only on the luminance and cannot measure the color correction.

In this paper we suggest a method to recover the distance maps and object colors in scenes photographed under ambient illumination in water, using just a single image as input. Our recovery takes into account the different optical water types and is based on a more comprehensive physical image formation model than the one previously used. Restoration from a single image is desirable because water properties temporally change, sometimes within minutes. In addition, no additional equipment, such as a tripod or filters, is required.

A variety of methods have been developed for the closely related single image dehazing problem, in which images are degraded by weather conditions such as haze or fog. Under the assumption of wavelength-independent attenuation, single image dehazing is an ill-posed problem with three measurements per pixel (the R,G,B values of the input image) and four unknowns: the true R,G,B values of the object and its distance from the camera. The transmission is the fraction of the scene's radiance that reaches the camera, and is related to the distance via the attenuation coefficient.

Under water, where the assumption of wavelength-independent attenuation does not hold, there are theoretically three unknown transmission values per pixel, one per channel, yielding six unknowns with only three measurements. However, the color-dependent transmission is related to the distance via the attenuation coefficients. Based on this relation we reduce the problem to estimation of four unknowns per pixel as before, with two new *global* parameters, the ratios between the attenuation coefficients of the color channels. We estimate these parameters using an existing library of water types, and based on the color distribution of the image after correction, utilizing the fact that using a wrong water type results in distorted colors. Our results demonstrate a successful single image restoration of underwater scenes using a comprehensive physical image formation model. Thus, we are able to recover more complex 3D scenes than previous methods and, in addition, estimate the optical water type. Fig. 1 depicts the proposed method.

Since public ground truth data is not available, we took multiple stereo images that contain color charts. We used the stereo images to recover the true distance from the camera. We then conducted a thorough quantitative analysis, comparing the results of multiple algorithms to the ground truth data. Our algorithm outperforms other state-of-the-art methods.

2 Related Works

The participating medium causes color shifts and reduces contrast in images. The degradation depends both on the water properties and the 3D structure, which can be estimated by analyzing *multiple* images that satisfy certain conditions. E.g., Schechner and Karpel [28, 29] take two images with orthogonal polarizer angles and utilize the partial polarization of light to restore the visibility. Multiple images of the same object taken from different *known* viewpoints [8, 9, 34] are used to estimate attenuation coefficients and recover the scene. These methods have limited applicability since they impose constraints on the imaging conditions.

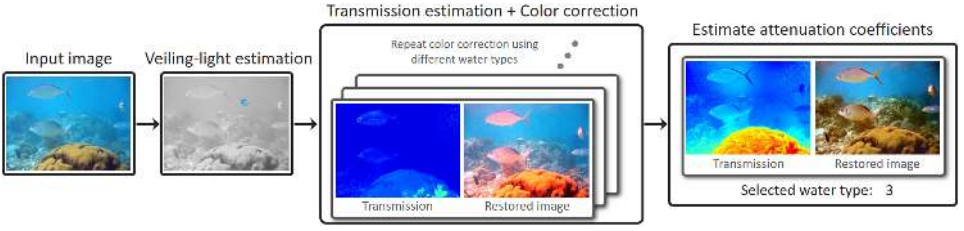

Figure 1: The proposed color restoration and transmission estimation method: First, the veiling-light (the ambient light that is scattered into the line of sight) is estimated. The pixels whose color is the veiling light are shown in color. Then, the transmission estimation and color restoration are repeated for multiple water types that have different optical characteristics, as described in Section 3.2. Finally, the best result is selected automatically based on the gray-world assumption. Photo from [11].

Scattering media limits visibility in terrestrial images as well, when imaging in fog, haze, or turbulence [30]. *Single* image dehazing methods, *e.g.* [7, 15, 17, 25, 33], take as input a single hazy image and estimate the unknown distance map and the clear image of the scene. Different priors are used in order to solve the ill-posed problem, under the assumption of color-independent transmission, which is reasonable for terrestrial images but is violated under water. The dark channel prior (DCP) [17] assumes that within small image patches, at least one pixel has a low value in some color channel, and uses the minimal value to estimate the distance. The Haze-Lines prior [7] is based on the observation that the number of colors in natural images is small, and that similar colors appear all over the image plane.

Some variations of DCP were proposed for the underwater domain [11, 12, 14, 16, 22]. Carlevaris-Bianco et al. [11] assume a color-independent transmission, and propose a variant of DCP based on the difference between the maximum of the red channel and the maximum of the blue and green channels in each patch. They claim this value is inversely related to transmission, since red is attenuated at a higher rate than blue or green. Drews et al. [14] apply DCP to the blue and green channels only, since the red channel is often attenuated rapidly and cannot be used to estimate the transmission. They get improved transmission maps, but still assume a uniform transmission across channels for recovery. Chiang and Chen [12] recover the transmission using standard DCP. They assume the recovered transmission is the transmission of the red channel, as under water this channel has the lowest transmission. They use fixed attenuation coefficients measured for open ocean waters to recover the image based on the estimated red transmission. Lu et al. [22] estimate the transmission using the darker of the blue and red channels. They use the same fixed water attenuation as in [12] to recover the scene. Galdran *et al*. [16] suggest the Red-Saturation prior, incorporating information from the inverse of the red channel and the saturation in order to estimate the transmission. They use a spectrally uniform transmission since the water type is not easy to determine. Instead, they add an additive veiling-light term to the color restoration.

Despite the abundance of DCP based methods, the underlying assumption does not hold in many underwater scenarios: bright foreground sand has high values in all color channels and might be mistaken to have a low transmission despite being close to the camera. Moreover, the background water has a dominant color (hence at least one color channel is low), and many of the mentioned methods inaccurately estimate the transmission there to be high.

Peng et al. [27] estimate the scene distance via image blurriness, which grows with the distance due to scattering. They disregard the spectral dependence of the transmission. Peng and Cosman [26] combine the blurriness prior with [11] and assume open ocean waters.

While this prior is physically valid, it has limited efficiency in textureless areas.

The above mentioned restoration methods aim for a physics-based recovery of the scenes' colors, while estimating its 3D structure in the process. Other methods aim for a visually pleasing result, *e.g.* [3, 4, 21], but have not shown color consistency that is required for scientific measurements.

The wavelength dependent absorption properties of water are used in [5] to recover depth. However this method requires an active illumination system and is valid only for short distances, since near-infrared light is rapidly attenuated in water. Known geometry is used in [31] to plan the imaging viewpoints under water that result in high contrast.

Morimoto et al. [24] estimate the optical properties of layered surfaces based on non-linear curves in RGB space, which is a similar physical model. Nonetheless, user scribbles are required in order to detect the curves.

3 Background

3.1 Image Formation Model

We follow the model developed in [29]. In each color channel $c \in \{R,G,B\}$, the image intensity at each pixel is composed of two components, attenuated signal and veiling-light:

$$I_c(\boldsymbol{x}) = t_c(\boldsymbol{x})J_c(\boldsymbol{x}) + (1 - t_c(\boldsymbol{x})) \cdot A_c \ , \tag{1}$$

where bold denotes vectors, $\boldsymbol{x}$ is the pixel coordinate, I_c is the acquired image value in color channel c, t_c is the transmission of that color channel, and J_c is the object radiance that we wish to restore. The global veiling-light component A_c is the scene value in areas with no objects ($t_c = 0, \forall c \in \{R,G,B\}$). Eq. (1) applies to linear captured data, prior to in-camera processing such as color-space conversion, tone mapping, and compression. Therefore, $\boldsymbol{I}$ refers to the image obtained from the raw file after minimal processing such as demosaicing and black current subtraction [1, 32].

The transmission depends on object's distance $z(\boldsymbol{x})$ and the water attenuation coefficient for each channel β_c:

$$t_c(\boldsymbol{x}) = \exp(-\beta_c z(\boldsymbol{x})) \ . \tag{2}$$

In the ocean, the attenuation of red colors can be an order of magnitude larger than the attenuation of blue and green [23]. Therefore, as opposed to the common assumption in single image dehazing, the transmission $\boldsymbol{t} = (t_R, t_G, t_B)$ is wavelength-dependent.

3.2 Water Attenuation

The attenuation of light under water is not constant and varies with geography, seasons, and climate related events. In clear open waters, visible light is absorbed at the longest wavelengths first, appearing deep-blue to the eye. In near-shore waters, sea water contains more suspended particles than the central ocean waters, which scatter light and make coastal waters less clear than open waters. In addition, the absorption of the shortest wavelengths is stronger, thus the green wavelength reaches deeper than other wavelengths.

Jerlov [19] developed a frequently used classification scheme for oceanic waters, based on water clarity. The Jerlov water types are I, IA, IB, II and III for open ocean waters, and 1 through 9 for coastal waters. Type I is the clearest and type III is the most turbid open ocean water. Likewise, for coastal waters, type 1 is clearest and type 9 is most turbid. Fig. 2(Left) depicts the attenuation coefficients' dependence on wavelength, while Fig. 2(Right) shows an RGB simulation of the appearance of a perfect white surface viewed at different depths in different water types. The common notion that red colors are attenuated faster than blue/green only holds for oceanic water types.

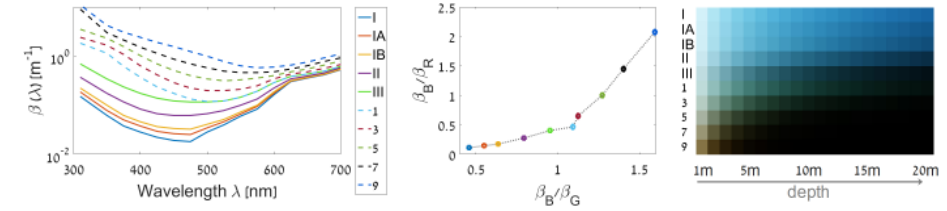

Figure 2: Left: Attenuation coefficients (β) of Jerlov water types, figure adapted from data in [23], based on measurements in [6]. Solid lines mark open ocean water types while dashed lines mark coastal water types. Middle: β ratios of R,G,B channels for each water type, based the wavelengths: 475nm, 525nm, and 600nm, respectively. Right: simulation of the appearance of a perfect white surface viewed at depth of 1-20m in different water types (reproduced from [2]).

When capturing an image using a commercial camera, three color channels R,G,B are obtained. Thus, we are interested in three attenuation coefficients: $(\beta_R,\beta_G,\beta_B)$ in order to correct the image. We use the Jerlov water types to constrain the space of attenuation coefficients in the RGB domain. We show in Sec. 4.2 that the three attenuation coefficient themselves are not required for transmission estimation, but rather their ratios (two variables). Fig. 2(Middle) shows the ratios of the attenuation coefficients: β_B/β_R vs. β_B/β_G of Jerlov water types for wavelengths of peak camera sensitivity according to [20] (475nm, 525nm, and 600nm for B,G,R, respectively). Essentially, we approximate the cameras' spectral response as a Dirac delta function, similarly to [24, 26]. This approximation is camera-agnostic, which is advantageous since the true response functions are rarely published.

4 Color Restoration

We show in the following that if we were given the two *global* attenuation ratios, then the problem can be reduced to single image dehazing (i.e., single attenuation coefficient across all channels) for which good algorithms exist. The question now becomes how to estimate these two parameters? Our solution is to evaluate every possible water type, as defined by Jerlov, and pick the best one. Each water type defines a known, and fixed, pair of attenuation ratios that we can use. Once we have evaluated all possible water types (10 in total) we use a standard natural image statistics prior to pick the best result. In particular, we found the Gray-World assumption prior to work well for our needs. The method is illustrated in Fig. 1 and summarized in Alg. 1.

4.1 Veiling-Light Estimation

First we describe how we estimate the veiling-light, which is required in every dehazing algorithm. We assume an area without objects is visible in the image, in which the pixels' color is determined by the veiling-light alone. Such an area should be smooth and not have texture. This assumption often holds when the line of sight is horizontal. It does not hold when photographing a reef wall up close, or when the camera is pointed downwards. However, in these cases, the distance of objects from the camera usually varies less then in horizontal scenes, and a simple contrast stretch is likely to be sufficient.

In order to detect the pixels that belong to the veiling-light, we generate an edge map of the scene using the Structured Edge Detection Toolbox [13] with pre-trained model and default parameters. We then threshold the edge map and look for the largest connected component. The pixels belonging to the largest component are classified as veiling-light pixels ($\boldsymbol{x} \in VL$). The veiling-light $\boldsymbol{A}$ is the average color of those pixels. This is demonstrated on the veiling-light estimation step of Fig. 1, where only the pixels $\boldsymbol{x} \in VL$ are shown in color.

4.2 Transmission Estimation

Combining and rearranging Eqs. (1,2) yields for the different channels:

$$A_R - I_R = e^{-\beta_R Z} \cdot (A_R - J_R) \ , \quad A_G - I_G = e^{-\beta_G Z} \cdot (A_G - J_G) \ , \quad A_B - I_B = e^{-\beta_B Z} \cdot (A_B - J_B) \ . \tag{3}$$

Raising Eq. (3) of the red channel to the power of $\frac{\beta_B}{\beta_R}$ yields:

$$(A_R - I_R)^{\frac{\beta_B}{\beta_R}} = e^{-\beta_R z \cdot \frac{\beta_B}{\beta_R}} \cdot (A_R - J_R)^{\frac{\beta_B}{\beta_R}} = t_B \cdot (A_R - J_R)^{\frac{\beta_B}{\beta_R}} \ . \tag{4}$$

Denote the ratios between the attenuation coefficients:

$$\beta_{BR} = \beta_B / \beta_R \ , \ \beta_{BG} = \beta_B / \beta_G \ . \tag{5}$$

Then, in this medium-compensated space we achieve a form similar to Eq. (1), with one unknown transmission per-pixel, common to all color channels:

$$\begin{bmatrix} (I_R(\boldsymbol{x}) - A_R)^{\beta_{BR}} \\ (I_G(\boldsymbol{x}) - A_G)^{\beta_{BG}} \\ (I_B(\boldsymbol{x}) - A_B) \end{bmatrix} = t_B(\boldsymbol{x}) \begin{bmatrix} (J_R(\boldsymbol{x}) - A_R)^{\beta_{BR}} \\ (J_G(\boldsymbol{x}) - A_G)^{\beta_{BG}} \\ (J_B(\boldsymbol{x}) - A_B) \end{bmatrix} \ . \tag{6}$$

This form is similar to the Haze-Lines [7] formulation. Therefore, we similarly cluster the pixels to Haze-Lines and obtain an initial estimation of the transmission of the blue channel $\tilde{t}_B$. Since the value $(I_c - A_c)$ might be negative, we avoid numerical issues when raising to the power β_{Bc} by raising the absolute value and keeping the sign.

In [7] it was assumed that there is a haze-free pixel in each Haze-Line. However, the attenuation coefficients measures by Jerlov indicate that even scene points that are located at a distance of only one meter from the camera have a t_B of about 0.9, depending on water type. Thus, we multiply the initial transmission estimation by 0.9 (similarly to [27]).

A bound on the transmission arises from the fact that $J_c \geq 0, \forall c \in \{R,G,B\}$. We substitute this bound in Eq. (1) and obtain a lower bound t^{LB} on the transmission of the blue channel, t_B, taking into account the different attenuation coefficients of the different color channels:

$$t^{LB} := max \left\{ 1 - \frac{I_B}{A_B}, \left(1 - \frac{I_G}{A_G}\right)^{\beta_{BG}}, \left(1 - \frac{I_R}{A_R}\right)^{\beta_{BR}} \right\} . \tag{7}$$

We detect pixels $\boldsymbol{x}$ with no scene objects based on their Mahalanobis distance from the distribution of the veiling-light pixels: $D_M(\boldsymbol{I}(\boldsymbol{x}))$ (from here on, when referring to the Mahalanobis distance, it is with respect to the distribution of intensities of veiling-light pixels). We set the transmission of such pixels to be the lower bound t^{LB} calculated in Eq. (7). However, a binary classification of the pixels to $VL, \overline{VL}$ often results in abrupt discontinuities in the transmission map, which are not necessarily distance discontinuities. Therefore, we use a soft-matting and calculate the transmission as follows:

$$t_B(\boldsymbol{x}) = \begin{cases} t^{LB}(\boldsymbol{x}) & D_M(\boldsymbol{I}(\boldsymbol{x})) \leq \overline{D_M} + \sigma_M \\ \tilde{t}_B(\boldsymbol{x}) & D_M(\boldsymbol{I}(\boldsymbol{x})) \geq D_M^{max} + \sigma_M \\ \alpha(\boldsymbol{x}) \cdot t^{LB}(\boldsymbol{x}) + (1 - \alpha(\boldsymbol{x})) \cdot \tilde{t}(\boldsymbol{x}) & otherwise \end{cases} \tag{8}$$

where $\overline{D_M} = \frac{1}{|VL|} \sum_{\boldsymbol{x} \in VL} D_M(\boldsymbol{I}(\boldsymbol{x}))$, is the average Mahalanobis distance of the veiling-light pixels, $D_M^{max} = \max_{\boldsymbol{x} \in VL} \{D_M(\boldsymbol{I}(\boldsymbol{x}))\}$ is their maximal Mahalanobis distance and σ_M is the standard deviation. $\alpha(\boldsymbol{x})$ is the matting coefficient for pixels that cannot be classified to object/ water with high probability: $\alpha(\boldsymbol{x}) = \frac{D_M(\boldsymbol{I}(\boldsymbol{x})) - \overline{D_M} - \sigma_M}{D_M^{max} - \overline{D_M}}$, yielding a relatively steep transition between VL and $\overline{VL}$.

Finally, we regularize the transmission using Guided Image Filter [18], with a contrast-enhanced input image as guidance.

Algorithm 1 Underwater image restoration

Input: $\boldsymbol{I}(\boldsymbol{x})$
Output: $\hat{\boldsymbol{J}}(\boldsymbol{x}), \hat{\boldsymbol{t}}(\boldsymbol{x})$

Detect veiling light pixels using structured edges (Sec. 4.1) and calculate $\boldsymbol{A}$
for each (β_{BR}, β_{BG}) values of water types (Fig. 2 middle) **do**
  for each $c \in \{R,G,B\}$ **do**
    $\tilde{I}_c(\boldsymbol{x}) = \text{sign}\left(I_c(\boldsymbol{x}) - A_c\right) \cdot \text{abs}\left(I_c(\boldsymbol{x}) - A_c\right)^{\beta_{Bc}}$ (β_{Bc} is defined in Eq. 5)
  Cluster pixels to 500 Haze-Lines as in [7] and estimate an initial transmission $\tilde{t}_B$
  Apply soft matting to $\tilde{t}_B$ with the lower bound (Eq. 8)
  Regularization using guided image filter, with a contrast-enhanced input as guidance
  Calculate the restored image using Eq. 9
  Perform a global WB on the restored image
Return the image that best adheres to the Gray-World assumption on pixels $\boldsymbol{x} \notin VL$

4.3 Scene Recovery

Once t_B is estimated, we can compensate for the color attenuation using the following:

$$J_c = A_c + \frac{I_c - A_c}{e^{-\beta_c Z}} = A_c + \frac{I_c - A_c}{t_B^{\beta_c/\beta_B}} \quad , \quad \forall c \in \{R,G,B\} \ . \tag{9}$$

Eq. (9) compensates for the intensity changes that happen in the path between the object and the camera. In addition, the ambient illumination is attenuated by the water column from the surface to the imaging depth, resulting in a colored global illumination. We are interested in restoring the colors as if they were viewed under white light, without a color cast. Since this effect is global in the scene, we correct it by performing white balance on the result. This global operator works well only because the distance-dependent attenuation has already been compensated for.

Finally, since Eq. (1) applies to the linear captured data, we convert the linear image to sRGB using a standard image processing pipeline, including color-space conversion and gamma curve for tone mapping [32].

4.4 Estimation of Attenuation Coefficient Ratios

Using the wrong coefficients (β_{BR}, β_{BG}) leads to restorations with skewed colors and wrong transmission maps. We use this insight to determine the most appropriate water type. We perform the restoration multiple times using different attenuation coefficients' ratios, corresponding to different Jerlov water types, and choose the best result automatically based on the Gray-World assumption. The attenuation coefficients' ratios are shown in Fig. 2(middle).

According to the Gray-World assumption [10], the average reflectance of surfaces in the world is achromatic. It has been used in the past for estimating attenuation coefficients underwater using known distances [8]. However, a significant portion of images taken under water often contain water without any objects. The Gray-World assumption obviously does not hold there. Therefore, we apply the Gray-World assumption only at image regions that contain objects, i.e., those that were not identified as veiling-light pixels. Thus, among all results using different water types, we choose the image with the smallest difference between the average values of the red, green and blue channels.

We considered several other measures (e.g. maximal contrast [33]), but found that the simple Gray-World assumption gave the best results and therefore we focus on this measure.

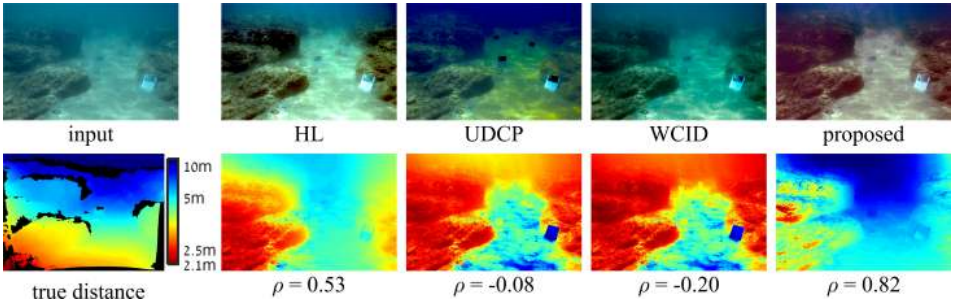

Figure 3: The left column contains the input image (top) and the true distance (bottom) which was calculated using stereo. The other columns show both the enhanced image and the estimated transmission of the following methods: HL [7], UDCP [14], WCID [12] and our proposed method (t_R is shown out of the three-channel transmission). The true distance and the transmission are shown in different scales, since the distance is given in meters while the transmission is in the range $[0,1]$. The quality of the 3D estimation is measured by the Pearson correlation coefficient ρ, which is calculated between the negative logarithm of the estimated transmission and the true distance. Best viewed in color.

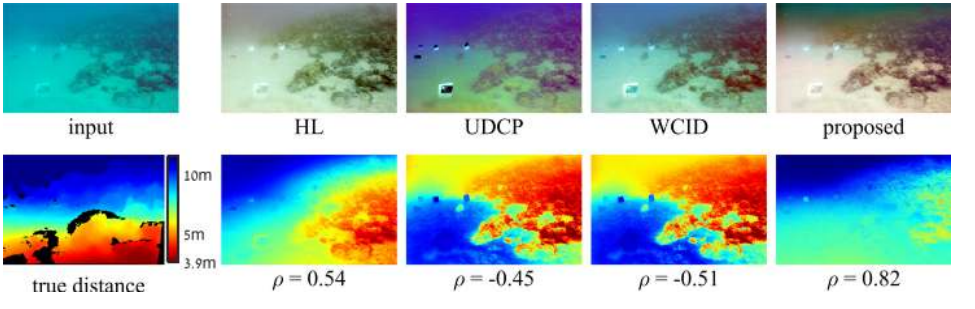

Figure 4: Similar to Fig. 3, different water type. The reader is encouraged to zoom-in.

5 Experimental Results

We use raw images taken with a Nikon D810 in two different locations, in tropical waters and in murkier coastal waters[1]. Using raw images is crucial in order to process the linear intensity, and since the signal of highly attenuated areas is stored in the least significant bits of each pixel. In order to quantitatively evaluate the color restoration and 3D reconstruction, we collected measurable data using color charts and stereo imaging. First, we placed in the scene five identical waterproof color charts (from DGK Color Tools), at different distances from the camera. The color charts are used only for validation and are masked when fed to our algorithm. The transmission values of the masked pixels are determined based on neighboring values at the bottom of the chart. In addition, we scuba dived with two synchronized cameras and calculated the ground-truth distances of the objects in the scene from the camera (only for pixels which are visible on both images). The stereo pair was *not* used for the color restoration, as it was only used to generate the ground-truth distance map. This type of data enables us to quantitatively evaluate the accuracy of the estimated transmission map.

Figs. 3 and 4 show a qualitative comparison between WCID [12][2], UDCP [14], and the

[1] High resolution images can be viewed in the supplementary material.

[2] The code was not available so we implemented it ourselves, except for the White-Balance section of the

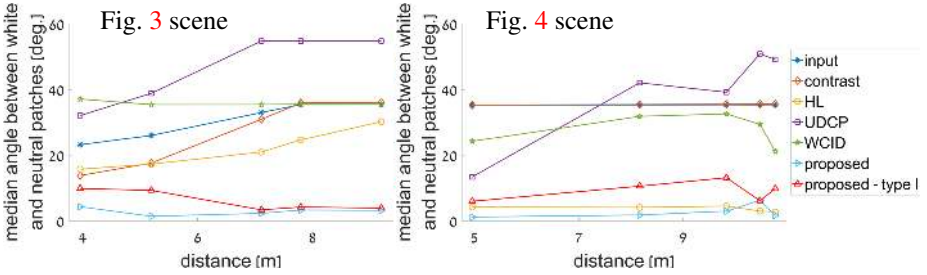

Figure 5: The median angle in RGB space between the neutral patches and a pure gray color is displayed for each color chart, as a function of distance. Lower is better.

proposed method. In addition, we include HL [7] as a baseline to emphasize the importance of handling spectral non-uniformity. Both UDCP and WCID are based on the dark channel prior, whose underlying assumption does not hold in those scenes due to two phenomena. First, the top of the scene has no objects and therefore its transmission should tend to zero, but it is relatively dark and according to the prior contains no haze. Second, the bright sand in the foreground has a significant value in all color channels and therefore is estimated to contain veiling-light by the prior. In contrast, the proposed method is better able to estimate the 3D structure of the scene and as a result, to better restore the colors.

The transmission maps, which were estimated from a single image, are evaluated quantitatively based on the true distance map of each scene (shown in Figs. 3, 4). Since the transmission is a function of the distance ($t_c = e^{-\beta_c z}$), we measure the correlation between the true distance and the negative logarithm of the estimated transmission ($-log(t_c) = \beta_c z$). The Pearson correlation coefficient ρ is noted below each transmission map, and shows that the proposed method estimates a much more accurate transmission.

Fig. 5 compares the colors of the different color charts in the restored images shown in Figs. 3, 4. We calculate the median angle between the gray-scale patches and a pure gray color, in RGB space. By calculating the angle we eliminate the influence of the global illumination and are able to calculate a robust measure on the neutral patches. We include two additional restoration techniques beyond those shown in Figs. 3, 4. First, the result of a naïve contrast stretch demonstrates that a global operation cannot compensate for the distance-dependent degradation. Second is the restoration result using the proposed method but assuming a specific water type - open ocean waters (type I, marked in red) - as assumed by WCID and UDCP. This restoration is less accurate and demonstrates the importance of using different spectral profiles.

The supplementary material includes more results and further comparisons to additional prominent single underwater image restoration methods by Carlevaris-Bianco *et al.* [11], Peng *et al.* [27], Lu *et al.* [22], Galdran *et al.* [16] and Ancuti *et al.* [4].

6 Conclusions

We demonstrate a physics-based method for the recovery of underwater scenes using just a single input image. As opposed to previous methods we consider wavelength-dependent attenuation and take into account various water types. We show that recovering transmission for each color channel separately adds just two global parameters to the problem. By considering a comprehensive physical model we are able to reconstruct scenes with a complex 3D structure and accurately correct the colors of objects that are further away.

pipeline, since we applied the same contrast stretch and White-Balance to all of the results.

7 Acknowledgments

This research was supported by ISF grant 1917/2015 and an IUI research scholarship. DB was partially supported by Apple Graduate Fellowship. TT was supported by the The Leona M. and Harry B. Helmsley Charitable Trust, The Maurice Hatter Foundation, and Ministry of Science, Technology and Space grant #3-12487.

References

[1] Derya Akkaynak, Tali Treibitz, Bei Xiao, Umut A. Gürkan, Justine J. Allen, Utkan Demirci, and Roger T. Hanlon. Use of commercial off-the-shelf digital cameras for scientific data acquisition and scene-specific color calibration. *JOSA A*, 31(2):312–321, 2014.

[2] 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 *Proc. IEEE CVPR*, 2017.

[3] Cosmin Ancuti, Codruta Orniana Ancuti, Tom Haber, and Philippe Bekaert. Enhancing underwater images and videos by fusion. In *Proc. IEEE CVPR*, pages 81–88, 2012.

[4] Cosmin Ancuti, Codruta Orniana Ancuti, Christophe De Vleeschouwer, Rafael Garcia, and Alan C Bovik. Multi-scale underwater descattering. In *Proc. ICPR*, 2016.

[5] Yuta Asano, Yinqiang Zheng, Ko Nishino, and Imari Sato. Shape from water: Bispectral light absorption for depth recovery. In *Proc. ECCV*, 2016.

[6] Roswell W. Austin and Theodore J. Petzold. Spectral dependence of the diffuse attenuation coefficient of light in ocean waters. *Optical Engineering*, 25(3):253471–253471, 1986.

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

[8] Mitch Bryson, Matthew Johnson-Roberson, Oscar Pizarro, and Stefan B. Williams. Colour-consistent structure-from-motion models using underwater imagery. In *Robotics: Science and Systems*, 2012.

[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, 2015.

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

[11] Nicholas Carlevaris-Bianco, Anush Mohan, and Ryan M Eustice. Initial results in underwater single image dehazing. In *Proc. IEEE/MTS Oceans*, 2010.

[12] John Y. Chiang and Ying-Ching Chen. Underwater image enhancement by wavelength compensation and dehazing. *IEEE Trans. Image Processing*, 21(4):1756–1769, 2012.

[13] Piotr Dollár and C. Lawrence Zitnick. Structured forests for fast edge detection. In *Proc. IEEE ICCV*, 2013.

[14] Paul Drews, E Nascimento, Filipe Moraes, Silvia Botelho, and Mario Campos. Transmission estimation in underwater single images. In *Proc. IEEE ICCV Underwater Vision Workshop*, pages 825–830, 2013.

[15] Raanan Fattal. Dehazing using color-lines. *ACM Trans. Graph.*, 34(1):13, 2014.

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

[17] Kaiming He, Jian Sun, and Xiaoou Tang. Single image haze removal using dark channel prior. In *Proc. IEEE CVPR*, 2009.

[18] Kaiming He, Jian Sun, and Xiaoou Tang. Guided image filtering. *IEEE Trans. PAMI*, 35(6):1397–1409, 2013.

[19] Nils Gunnar Jerlov. *Marine optics*, volume 14. Elsevier, 1976.

[20] Jun Jiang, Dengyu Liu, Jinwei Gu, and Sabine Susstrunk. What is the space of spectral sensitivity functions for digital color cameras? In *Proc. IEEE Workshop Applications of Computer Vision (WACV)*, pages 168–179, 2013.

[21] Huimin Lu, Yujie Li, and Seiichi Serikawa. Underwater image enhancement using guided trigonometric bilateral filter and fast automatic color correction. In *Proc. IEEE ICIP*, 2013.

[22] Huimin Lu, Yujie Li, Lifeng Zhang, and Seiichi Serikawa. Contrast enhancement for images in turbid water. *JOSA A*, 32(5):886–893, 2015.

[23] Curtis D Mobley. *Light and water: radiative transfer in natural waters*. Academic press, 1994.

[24] Tetsuro Morimoto, Robby T. Tan, Rei Kawakami, and Katsushi Ikeuchi. Estimating optical properties of layered surfaces using the spider model. In *Proc. IEEE CVPR*, 2010.

[25] Ko Nishino, Louis Kratz, and Stephen Lombardi. Bayesian defogging. *International Journal of Computer Vision*, 98(3):263–278, 2012.

[26] 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, April 2017.

[27] Yan-Tsung Peng, Xiangyun Zhao, and Pamela C. Cosman. Single underwater image enhancement using depth estimation based on blurriness. In *Proc. IEEE ICIP*, 2015.

[28] Yoav Y. Schechner and Nir Karpel. Clear underwater vision. In *Proc. IEEE CVPR*, 2004.

[29] Yoav Y. Schechner and Nir Karpel. Recovery of underwater visibility and structure by polarization analysis. *IEEE J. Oceanic Engineering*, 30(3):570–587, 2005.

[30] Armin Schwartzman, Marina Alterman, Rotem Zamir, and Yoav Y. Schechner. Turbulence-induced 2D correlated image distortion. In *Proc. IEEE ICCP*, 2017.

[31] Mark Sheinin and Yoav Y. Schechner. The next best underwater view. In *Proc. IEEE CVPR*, pages 3764–3773, 2016.

[32] Robert Sumner. Processing raw images in matlab. https://users.soe.ucsc.edu/~rcsumner/rawguide/RAWguide.pdf, 2014.

[33] Robby T. Tan. Visibility in bad weather from a single image. In *Proc. IEEE CVPR*, 2008.

[34] Atsushi Yamashita, Megumi Fujii, and Toru Kaneko. Color registration of underwater images for underwater sensing with consideration of light attenuation. In *IEEE. Conf. Robotics and Automation*, pages 4570–4575, 2007.