



A Rapid Scene Depth Estimation Model Based on Underwater Light Attenuation Prior for Underwater Image Restoration

Wei Song¹, Yan Wang¹, Dongmei Huang^{1(✉)},
and Dian Tjondronegoro^{2(✉)}

¹ College of Information Technology, Shanghai Ocean University,
Shanghai, China

{wsong, dmhuang}@shou.edu.cn, yanwang9310@163.com

² Southern Cross University, Gold Coast, Australia

Dian.tjondronegoro@scu.edu.au

Abstract. Underwater images present blur and color cast, caused by light absorption and scattering in water medium. To restore underwater images through image formation model (IFM), the scene depth map is very important for the estimation of the transmission map and background light intensity. In this paper, we propose a rapid and effective scene depth estimation model based on underwater light attenuation prior (ULAP) for underwater images and train the model coefficients with learning-based supervised linear regression. With the correct depth map, the background light (BL) and transmission maps (TMs) for R-G-B light are easily estimated to recover the true scene radiance under the water. In order to evaluate the superiority of underwater image restoration using our estimated depth map, three assessment metrics demonstrate that our proposed method can enhance perceptual effect with less running time, compared to four state-of-the-art image restoration methods.

Keywords: Underwater image restoration · Underwater light attenuation prior
Scene depth estimation · Background light estimation
Transmission map estimation

1 Introduction

Underwater image restoration is challenging due to complex underwater environment where images are degraded by the influence of water turbidity and light attenuation [1]. Compared with green (G) and blue (B) lights, red (R) light with the longer wavelength is the most affected, thus underwater images appear blue-greenish tone. In our previous work [2], the model of underwater image optical imaging and the light selective attenuation are presented in detail.

An image restoration method recovers underwater images by considering the basic physics of light propagation in the water medium. The purpose of restoration is to deduce the parameters of the physical model and then recover the underwater images by reserved compensation processing. A simplified image formation model (IFM) [3–6] is

often used to approximate the propagation equation of underwater Atmospheric scattering, can be shown as:

$$I_\lambda(x) = J_\lambda(x)t_\lambda(x) + (1 - t_\lambda(x))B_\lambda \quad (1)$$

where x is a point in the image, λ represents RGB lights in this paper, $I_\lambda(x)$ and $J_\lambda(x)$ are the hazed image and the restored image, respectively, B_λ is regarded as the background light (BL), $t_\lambda(x)$ is the transmission map (TM), which is a function of both λ and the scene-camera distance $d(x)$ and can be expressed as:

$$t_\lambda(x) = e^{-\beta(x)d(x)} = Nrer(\lambda)^{d(x)} \quad (2)$$

where $e^{-\beta(x)}$ can be represented as the normalized residual energy ratio $Nrer(\lambda)$, which depends on the wavelength of one channel and the water type in reference to [7].

In the IFM-based image restoration methods, a proper scene depth map is the key for background light and transmission map estimation. Since the hazing effect of underwater images caused by light scattering and color change is similar to the fog effect in the air, He's Dark Channel Prior (DCP) [3] and its variants [4, 8, 9] were used for depth estimation and restoration. Li et al. [10] estimated the background light by the map of the maximum intensity prior (MIP) to dehaze G-B channel and used Gray-World (G-W) theory to correct the R channel. However, Peng et al. [6] found the previous background light estimation methods were not robust for various underwater images, and they proposed a method based on the image blurriness and light absorption to estimate more accurate background light and scene depth to restore color underwater image precisely.

Learning-based methods for underwater image enhancement have been taken into consideration in recent years. Liu et al. [11] proposed the deep sparse non-negative matrix factorization (DSNMF) to estimate the image illumination to achieve image color constancy. Ding et al. [12] estimated the depth map using the Convolutional Neural Network (CNN) based on the balanced images that were produced by adaptive color correction. Although the above methods can obtain the scene depth map and enhance the underwater image, the deep learning is time consuming. A linear model to predict the scene depth of hazy images based on color attenuation prior is proposed by Zhu et al. [13]. This model is trained by supervised learning method and its expression is as follows:

$$d(x) = \theta_0 + \theta_1 v(x) + \theta_2 s(x) + \epsilon(x) \quad (3)$$

where $d(x)$ is the scene depth, $v(x)$ and $s(x)$ are the brightness and saturation components, respectively, $\epsilon(x)$ is a Gaussian function with zero mean and the standard deviation value σ . The model can successfully recover the scene depth map of outdoor hazed images. Unfortunately, it is not suitable for underwater images. Figure 1 shows an example image with the false depth map and invalid recovered image.

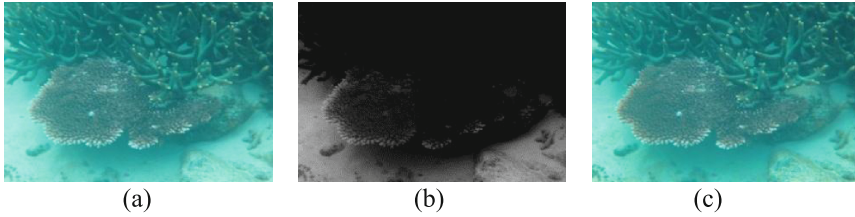


Fig. 1. An overview of Zhu et al.'s method. (a) Original underwater image; (b) Estimated depth map where the whiter the father; (c) Restored image.

In this paper, we reveal underwater light attenuation prior (ULAP) that the scene depth increases with the higher value of the difference between the maximum value of G and B lights and the value of the R light. On the basis of the ULAP and annotated scene depth data, we train a linear model of scene depth estimation. With the accurate scene depth map, BL and TM are easily estimated, and then underwater images are restored properly. The details of the scene depth learning model and the underwater image restoration process are presented in Sects. 2 and 3, respectively. In Sect. 4, we evaluate our results by comparing with four state-of-the-art image restoration methods from the perspectives of quantitative assessment and complexity. The conclusion is shown in Sect. 5.

2 Scene Depth Model Based on ULAP

2.1 Underwater Light Attenuation Prior

On account of little information about underwater scenes, restoring a hazed underwater image is a difficult task in computer vision. But, the human can quickly recognize the scene depth of the underwater image without any auxiliary information. When we explore a robust background light estimation, the farthest point in the depth map corresponding to the original underwater image is often considered as the background light candidate. With the light attenuation underwater, depending on the wavelength where the energy of red light is absorbed more than that of green and blue lights, the most intensity difference between R light and G-B light is used to estimate the background light [9, 10]. The rule motivates us to conduct experiments on different underwater images to discover an effective prior for single underwater image restoration.

After examining a large number of underwater images, we find the Underwater Light Attenuation Prior (ULAP), that is the difference between the maximum value of G-B intensity (simplified as MVGB) and the value of R intensity (simplified as VR) in one pixel of the underwater image is very strongly related to the change of the scene depth. Figure 2 gives an example with a typical underwater scene to show the MVGB, VR and the difference vary along with the change of the scene within different depths. As illustrated in Fig. 2(a), we select three regions as the test data from the close scene to the far scene and show the corresponding close-up patches in the right side. It is observed

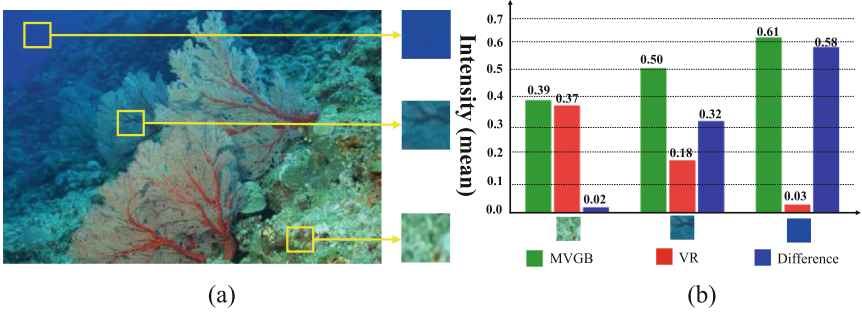


Fig. 2. The scene depth is positively correlated with the difference between MVGB and VR. (a) Original underwater image; (b) Three close-up patches of the close scene, moderately-close scene, relatively-far scene and their corresponding histograms, respectively.

in the left histogram of Fig. 2(b), in the close region, the MVGB and VR is relatively moderate and the difference is close to zero; In the middle histogram of Fig. 2(b), the MVGB in the moderately-close region increases while with the farther depth of the scene, the VR decreases at the same time, producing a higher value of the difference. Furthermore, in the utmost scene, the component of the red light remains nothing much due to the significant attenuation, the MVGB improves remarkably, and the difference is drastically higher than that in other patches in the right histogram of Fig. 2(b). In all, when the scene goes to a far region, the MVGB increase and the VR decreases, which leads to a positive correlation between the scene depth and the difference between MVGB and VR.

2.2 Scene Depth Estimation

Based on the ULAP, we define a linear model of the MVGB and VR for the depth map estimation as follows:

$$d(x) = \mu_0 + \mu_1 m(x) + \mu_2 v(x) \tag{4}$$

Where x represents a pixel, $d(x)$ is the underwater scene depth at point x , $m(x)$ is the MVGB, $v(x)$ is the VR.

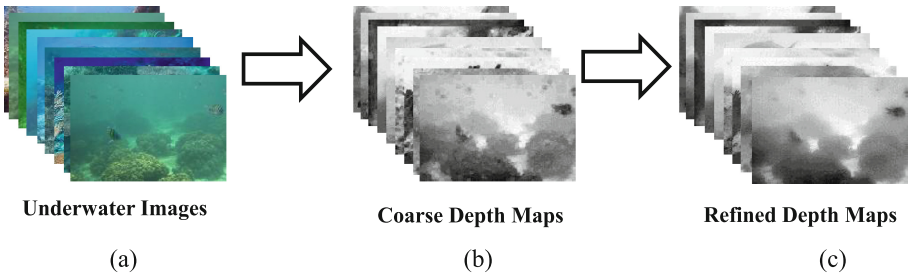


Fig. 3. The process of generating the training samples. (a) Original underwater images; (b) Coarse depth maps; (c) Refined depth maps by the guided filter [14].

The Training Data. In order to learn the coefficients μ_0 , μ_1 and μ_2 accurately, we need relatively correct training data. In reference with the depth map estimation proposed by Peng et al. [6], we computed the depth maps of different underwater images based on the light absorption and image blurriness. 500 depth maps of underwater images were gained following by [6], and some maps exist obvious estimation errors (e.g., a close-up fish is white in the depth) were discarded. Hence 100 fully-proper depth maps are well-chosen by the final manual selection. The process of generating the training samples is illustrated in Fig. 3. Firstly, for the 500 underwater images, the depth map estimation method is used to obtain the corresponding depth maps with the same size, and then 100 accurate depth maps from the above maps are ascertained as the final training data. Secondly, the guided filter [14] is used to refine the coarse depth maps and the radius of the guided filter is set as 15 in this paper. Finally, 100 accurate depth maps verified by manually-operated selection according to the perception of the scene depths. The prepared dataset has a total of 24 million points of depth information, which are called reference depth maps (RDMs).

Coefficients Learning. Based on the reference depth maps (RDMs), Pearson correlation analysis for the MVGB and VR is firstly run. The range of Pearson correction coefficient (PCC) is $[-1, 1]$, and the PCC close to 1 or -1 indicates a perfect linear relationship and the value closes to 0 demonstrates no relations. The PCC values of the MVGB and VR are 0.41257 and -0.67181 ($\alpha \leq 0.001$), respectively. We can readily see that there is a tight correlation between RDMs and the MVGB, and between RDMs and the VR. Therefore, the proposed model in (4) is reasonable.

To train the model, we take the ratio of training and testing dataset as 7:3 and use 10-fold cross validation. The best learning result is $\mu_0 = 0.53214829$, $\mu_1 = 0.51309827$ and $\mu_2 = -0.91066194$. Once the values of the coefficients have been determined, this model will be used to generate the scene depths of single underwater images under different scenarios.

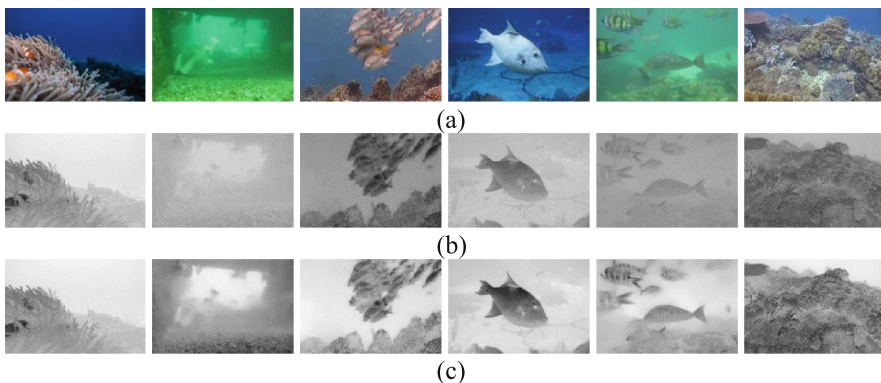


Fig. 4. The process of generating the training samples. (a) Original underwater images; (b) Coarse depth maps based on our method; (c) Refined depth maps.

Estimation of the Depth Map. As the relationship among the scene depth map $d(x)$, the MVGB and VR has been established and the coefficients have been learned, the depth maps of any underwater images can be obtained by Eq. (4). In order to check the validity of the assumption, we collected a large database of underwater images from several well-known photo websites, e.g., Google Images, Filck.com, and computed the underwater scene depth maps of each test images. Some of the results are shown in the Fig. 4. For different types of input underwater images, the corresponding estimated coarse depth maps and the refined depth maps are shown in the Fig. 4(a)–(c), respectively. As can be seen, the estimated depth maps have brighter color in the regions within the farther depth while having lighter color in the closer region as expected. After obtaining the correct depth map, the background light estimation and the transmission maps for RGB lights are rather simple and rapid.

3 Underwater Image Restoration

3.1 Background Light Estimation

The background light BL in the Eq. (1) is often estimated as the brightest pixel in an underwater image. However, the assumption is not correct in some situations, e.g., the foreground objects are brighter than the background light. The background light is selected from the farthest point of the input underwater image, i.e., the position of the maximum value in the refined depth map corresponding the input underwater image is the background light candidate value. But directly select the farthest point as the final background light, some suspended particles can interrupt the valid estimation result. After generating an accurate depth map, we firstly remove the effects of suspended particles via selecting the 0.1% farthest point, and then select the pixel with the highest intensity in the original underwater image. An example to illustrate the background light estimation method is shown in Fig. 5.

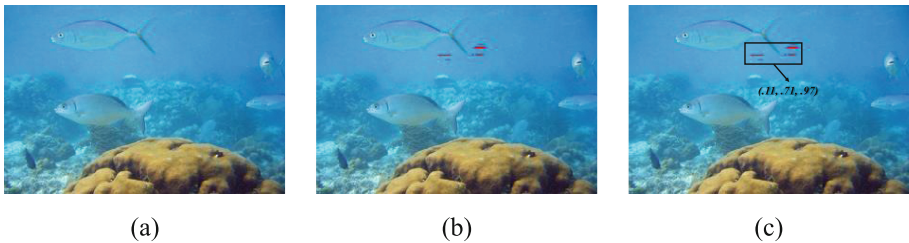


Fig. 5. An example to illustrate the global background light estimation algorithm. (a) Original image; (b) The result of searching for the 0.1% farthest pixels in the refined depth map; (c) The brightest intensity of the 0.1% farthest pixels corresponding to the original image.

3.2 Transmission Map Estimation for Respective R-G-B Channel

The relative depth map cannot be directly used to estimate the final TMs for R-G-B channel. To measure the distance from the camera to each scene point, the actual scene depth map d_a is defined as follows:

$$d_a(x) = D_\infty \times d(x) \quad (5)$$

where D_∞ is a scaling constant for transforming the relative distance to the real distance, and in this paper, the D_∞ is set as 10. With the estimated d_a , we can calculate the TM for the R-G-B channel as:

$$t_\lambda(x) = Nrer(\lambda)^{d_a(x)} \quad (6)$$

In approximately 98% of the world's clear oceanic or coastal water (ocean type I), the accredited ranges of $Nrer(\lambda)$ in red, green and blue lights are 80%–85%, 93%–97%, and 95%–99%, respectively [7]. In this paper, we set the $Nrer(\lambda)$ for R-G-B light is 0.83, 0.95 and 0.97, respectively. Figure 6(c)–(e) gives an example of TMs for the RGB channels of a blue-greenish underwater image based on Eqs. (5)–(6). Figure 6(c)–(e) gives an example of TMs for the RGB channels of a blue-greenish underwater image based on Eqs. (5)–(6).

Now that we have the BL_λ and $t_\lambda(x)$ for the R-G-B channel, we can restore the underwater scene radiance J_λ with the Eq. (7). A lower bound and an upper bound for $t_\lambda(x)$ empirically set to 0.1 and 0.9, respectively.

$$J_\lambda(x) = \frac{1}{\min(\max(t_\lambda(x), 0.1), 0.9)} (I_\lambda(x) - BL_\lambda) + BL_\lambda \quad (7)$$

Figures 6(f) and 7(f) show some final results recovered by our proposed method.

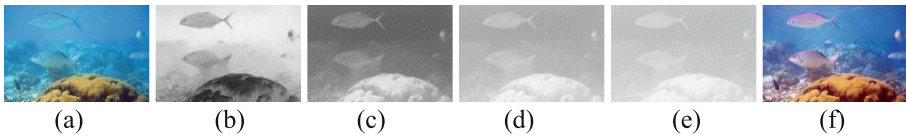


Fig. 6. The processing of underwater image restoration (a) Original image; (b) The refined depth map; (c) The estimated red transmission map; (d) The estimated green transmission map; (e) The estimated blue transmission map; (f) The restored image. (Color figure online)

4 Results and Discussion

In this part, we compare our proposed underwater image restoration method based on underwater light attenuation prior against with the typical image dehazing method by He et al. [3], the variant of the DCP (UDCP) by Drew et al. [8], the image restoration based on the dehazing of G-B channel and the correction of R channel [10], and Peng et al.'s method [6] based on both the image blurriness and light absorption. In order to

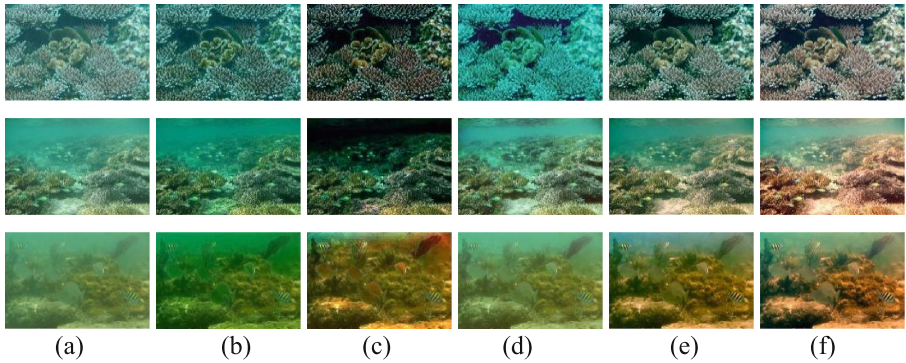


Fig. 7. (a) Original images with a size of 600×400 pixels; (b) He et al.'s results; (c) Drew et al.'s results; (d) Li et al.'s results; (e) Peng et al.'s results; (f) Our results.

demonstrate the outstanding performance of our proposed method, we present some examples of the restored images and introduce quantitative assessments and complexity comparison in this section.

Figure 7(a) shows four raw underwater images with different underwater characteristics in terms of color tones and scenes from our underwater images datasets. In Fig. 7(b), the DCP has no or little work on the test images due to incorrect depth map estimation. This indicates the direct application of the outdoor image dehazing is not suitable to the underwater image restoration. Figure 7(c) shows that the UDCP fails to recover the scene of the underwater images and even bring color distortion and error restoration. As shown in the Fig. 7(d)–(f), although all the methods can remove the haze of the input images, the color and contrast of Fig. 7(d–e) are not as good as those of Fig. 7(f) because the underwater light selection attenuation is ignored when estimating transmission maps or background light. As shown in Fig. 7(f), our image restoration method can effectively descatter and dehaze different underwater images, improves details and colorfulness of the input images and finally produce a natural underwater images.

4.1 Quantitative Assessment

Considering the fact that the clear underwater image presents better color, contrast and visual effect, we rely on three non-reference quantitative metrics: ENTROPY, underwater image quality measure (UIQM) [15] and the Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) [16] to assess the restored image quality. Table 1 lists the average scores of the three metrics on the recovered 100 low-quality underwater images and notes that the best results are in bold. Entropy represents the abundance of information. An image with higher entropy value prevents more valuable information. The UIQM is a linear combination of colorfulness, sharpness and contrast and a larger value represents higher image quality. The highest values of both ENTROPY and UIQM values mean that our proposed method can recover high-quality underwater images and reserve a lot of image information. The BRISQUE quantifies possible

losses of naturalness in an image due to the presence of distortions. The BRISQUE value indicates the image quality from 0 (best) to 100 (worst). The lowest value of the BRISQUE obtained by our proposed method indicates the restored underwater image achieves natural appearance.

Table 1. Quantitative analysis in terms of ENTROPY, UIQM and BRISQUE.

Method	ENTROPY	UIQM	BRISQUE
He et al. [9]	6.1252	1.9296	33.5317
Drew et al. [8]	6.3714	2.6425	32.9704
Li et al. [10]	6.8031	3.3072	34.2450
Peng et al. [6]	6.7529	3.4514	30.1825
Ours	7.2398	3.9884	28.9147

4.2 Complexity Comparison

In this part, we compare running time of our proposed method with other methods (maps refined by the guided filter [14]), including He et al. [3], Drew et al. [8], Li et al. [10] and Peng et al. [6], processed on a Windows 7 PC with Intel(R) Core(TM) i7-4790U CPU@3.60 GHz, 8.00 GB Memory, running on Python3.6.5. When an underwater image with the size $m \times n$ is refined by the guided filter with the radius r , the complexity of our proposed underwater image restoration method is $O(m \times n \times r)$ after the linear model is used to estimate the depth map. In the Fig. 8, the running time

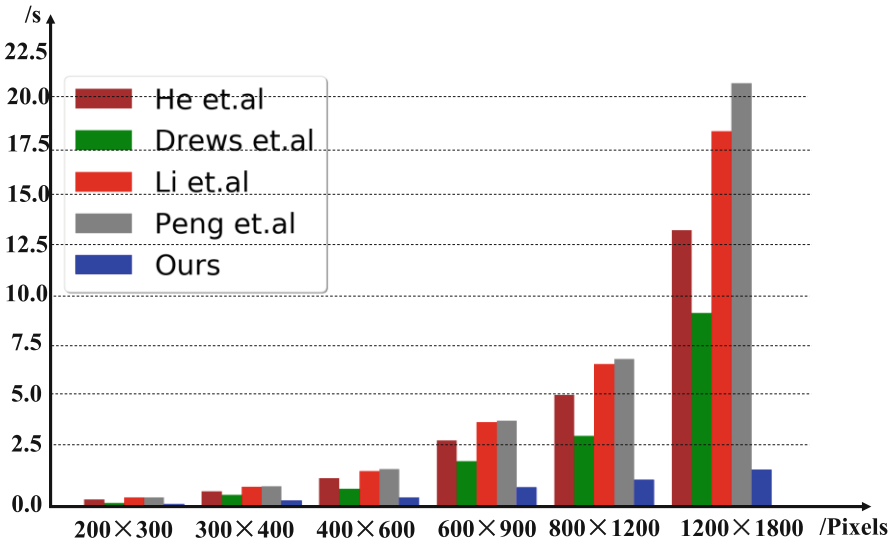


Fig. 8. The processing time (/s) of our proposed method (blue bar) with the different sizes of the input underwater images comparison with He et al. (brown bar), Drew et al. (green bar), Li et al. (red bar), Peng et al. (gray bar). (Color figure online)

(RT) is the average time (/s) of 50 underwater images with different sizes (/pixels). The RT of other compared methods improves significantly as the size test image becomes larger. Even though an underwater image is the size of 1200×1800 pixels, the RT of our method is lower than 2 s. In our method, the most time is used to obtain three transmission maps for respective RGB lights.

5 Conclusion

In this paper, we have explored the rapid and effective model of depth map estimation based on the underwater light attenuation prior within underwater images. After the linear model is created, the BL and the TMs for R-G-B light are smoothly estimated to restore the scene radiance simply. Our proposed method can achieve better quality of the restored underwater images, meanwhile it can save a mass of consumption time due to the depth map estimation using the linear model based on the ULAP, which further simplifies the deduction of the background light and transmission maps for R-G-B channel. The experimental results prove that our method can be well-suitable for underwater image restoration under different scenarios, faster and more effective to improve the quality of underwater images, according to the best objective evaluations and the lowest running time.

Acknowledgment. This work was supported by the National Natural Science Foundation of China (NSFC) Grant 61702323 and the Program for Professor of Special Appointment at Shanghai Institutions of Higher Learning (TP2016038). The first two authors contributed equally to this work.

References

1. Zhao, X., Jin, T., Qu, S.: Deriving inherent optical properties from background color and underwater image enhancement. *Ocean Eng.* **94**, 163–172 (2015)
2. Huang, D., Wang, Y., Song, W., Sequeira, J., Mavromatis, S.: Shallow-water image enhancement using relative global histogram stretching based on adaptive parameter acquisition. In: Schoeffmann, K., et al. (eds.) *MMM 2018*. LNCS, vol. 10704, pp. 453–465. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-73603-7_37
3. He, K.M., Sun, J., Tang, X.O.: Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **33**(12), 2341–2353 (2011)
4. Galdran, A., Pardo, D., Picón, A., Alvarez-Gila, A.: Automatic Red-Channel underwater image restoration. *J. Vis. Commun. Image Represent.* **26**, 132–145 (2015)
5. Wang, J.B., He, N., Zhang, L.L., Lu, K.: Single image dehazing with a physical model and dark channel prior. *Neurocomputing* **149**, 718–724 (2015)
6. Peng, Y.T., Cosman, P.C.: Underwater image restoration based on image blurriness and light absorption. *IEEE Trans. Image Process.* **26**(4), 1579–1594 (2017)
7. Chiang, J.Y., Chen, Y.C.: Underwater image enhancement by wavelength compensation and dehazing. *IEEE Trans. Image Process.* **21**(4), 1756–1769 (2012)
8. Drews, P., Nascimento, E., Moraes, F., Botelho, S., Campos, M.: Transmission Estimation in Underwater Single Images. In: *Proceedings of IEEE ICCVW 2013*, pp. 825–830 (2013)

9. Wen, H.C., Tian, Y.H., Huang, T.J., Guo, W.: Single underwater image enhancement with a new optical model. In: Proceedings of IEEE ISCAS 2013, pp. 753–756 (2013)
10. Li, C.Y., Quo, J.C., Pang, Y.W., Chen, S.J., Wang, J.: Single underwater image restoration by blue-green channels dehazing and red channel correction. In: ICASSP 2016, pp. 1731–1735 (2016)
11. Liu, X.P., Zhong, G.Q., Cong, L., Dong, J.Y.: Underwater image colour constancy based on DSNMF. IET. Image Process. **11**(1), 38–43 (2017)
12. Ding, X.Y., Wang, Y.F., Zhang, J., Fu, X.P.: Underwater image dehaze using scene depth estimation with adaptive color correction. In: OCEAN 2017, pp. 1–5 (2017)
13. Zhu, Q.S., Mai, J.M., Shao, L.: A fast single image haze removal algorithm using color attenuation prior. IEEE Trans. Image Process. **24**(11), 3522–3533 (2015)
14. He, K.M., Sun, J., Tang, X.O.: Guided image filtering. IEEE Trans. Pattern Anal. Mach. Intell. **35**(6), 1397–1409 (2013)
15. Panetta, K., Gao, C., Agaian, S.: Human-visual-system-inspired underwater image quality measures. IEEE J. Ocean. Eng. **41**(3), 541–551 (2016)
16. Mittal, A., Moorthy, A.K., Bovik, A.C.: No-reference image quality assessment in the spatial domain. IEEE Trans. Image Process. **21**(12), 4695–4708 (2012)