

Underwater Image Devignetting and Colour Correction

Yujie Li¹, Huimin Lu^{1,2,3(✉)}, and Seiichi Serikawa¹

¹ Kyushu Institute of Technology, Kitakyushu, Japan
luhuimin@ieee.org

² Shanghai Jiaotong University, Shanghai, China

³ Tongji University, Shanghai, China

Abstract. This paper describes a novel method to recover underwater images by devignetting and colour correction. Scattering and colour distortion are two major problems of degradation for underwater imaging. Scattering is caused by large suspended particles. Colour distortion corresponds to the varying degrees of attenuation encountered by light traveling in water with different wavelengths, rendering ambient underwater environments dominated by a bluish tone. To this end, we propose a novel underwater imaging model, which is much closer to the light propagation model in underwater environment. We remove the noise by dual-tree complex wavelet transform. Then, solve the non-uniform illumination of artificial lights by devignetting. Finally, we recover the image colour through camera spectral responses. The corrected images are characterized by reduced noised level, better exposedness of the dark regions, improved global contrast while the finest details and edges preserving.

Keywords: Underwater imaging · Inherent optical properties · Image devignetting · Colour correction

1 Introduction

Underwater exploration of the seafloor is used for various scientific reasons, such as assessing the biological environment, mineral exploration, and taking population census. The problem of mitigating sea mines is challenging and multifaceted. Ocean mines may be located on the seafloor in the water column. Recently, two deficiencies in mine counter measures are mentioned by researchers. First, the problem of mine hunting is one of the most important and difficult problems. Second, the key technologies for mine and mine-like objects recognition are not solving. To this end, in the last decade, most unmanned systems were developed for supporting underwater mining. The goal in developing automation for underwater mine detection is automatically determine the mine location and recognize the mine-like objects instead of the human. There are two stages of mine hunting operations. One is search-classify-map (SCM), which is intended to locate all sufficiently mine-like objects in a given operational area. The other is reacquire-and-identify (RI), which is to distinguish the mines or non-mines and prosecutes them accordingly [1].

Although the underwater objects detection technology makes a great progress, the recognition of underwater objects also remains a major issue in recent days. Challenges associated with obtaining visibility of objects have been difficult to overcome due to the physical properties of the medium. Different from the common images, underwater images suffer from poor visibility due to the medium scattering and light distortion. First of all, capture images underwater is difficult, mostly due to attenuation caused by light that is reflected from a surface and is deflected and scattered by particles, and absorption substantially reduces the light energy. The random attenuation of the light is mainly cause of the haze appearance while the fraction of the light scattered back from the water along the sight considerable degrades the scene contrast. In particular, the objects at a distance of more than 10 meters are almost indistinguishable while the colours are faded due to the characteristic wavelengths are cut according to the water depth . Moreover, as the artificial light is employed, there usually leave a distinctive footprint of the light beam on the seafloor.

There have been many techniques to restore and enhance the underwater images. Y. Y. Schechner et al. [2] exploited the polarization dehazing method to compensate for visibility degradation. Combining point spread function and a modulation transfer function to reduce the blurring effect by Hou et al. [3]. Although the aforementioned approaches can enhance the image contrast, these methods have demonstrated several drawbacks that reduce their practical applicability. First, the equipment of imaging is difficult in practice (e.g. range-gated laser imaging system, which is hardly applied in practice). Second, multiple input images are required, which is difficult to capture by hardware. Third, they cannot solve the colour distortion very well.

In this paper, we introduce a novel approach that is able to enhance underwater images based on single image to overcome the drawbacks of the above methods. We propose a new guided median filter instead of the soft matting to solve the alpha mattes more efficiently. In short summary, our technical contributions are in threefold: first, the proposed guided median filter can perform as an edge-preserving smoothing operator like the popular bilateral filter, but has better behavior near the edges. Second, the devinnetting method has a fast and non-approximate constant-time algorithm, whose computational complexity is independent of the filtering kernel size. Third, the proposed colour correction method is effectively in underwater image enhancement.

The organization of this paper is as follows. In Sect. 2, underwater imaging model will be discussed. And we will demonstrate an image enhancement system on Sect. 3. We apply the enhancement model in underwater optical images in Sect. 4. Finally, a conclusion is presented in Sect. 5.

2 Underwater Imaging Model

In the optical model [4], the acquired image can be modelled as being composed of two components. One is the direct transmission of light from the object, and the other is the transmission due to scattering by the particles of the medium. Mathematically, it can be written as

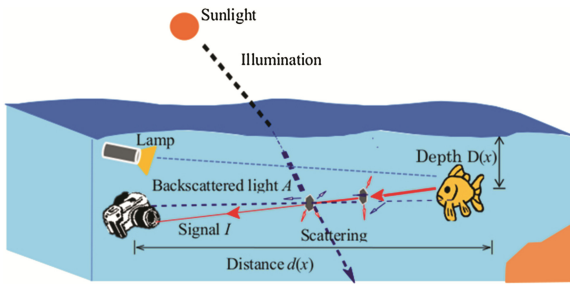
$$I(x) = J(x)t(x) + (1 - t(x))A \quad (1)$$

where I is the achieved image. J is the scene radiance or haze-free image, t is the transmission along the cone of vision, and $t(x) = e^{-\beta d(x)}$, β is the attenuation coefficient of the medium, $d(x)$ is the distance between the camera and the object, A is the veiling colour constant and $x = (x, y)$ is a pixel. The optical model assumes linear correlation between the reflected light and the distance between the object and observer.

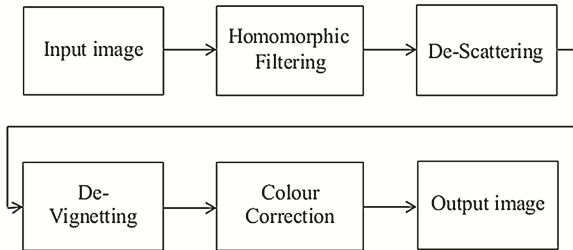
The light propagation model is slightly different underwater environment. In the underwater optical imaging model, absorption plays an important role in image degrading. Furthermore, unlike scattering, the absorption coefficient is different for each colour channel, being the highest for red and lowest for blue in seawater. These leads to achieve the following simplified hazy image formation model:

$$I(x) = J(x)e^{-(\beta_s + \beta_a)d(x)} + (1 - e^{-\beta_s d(x)})A \tag{2}$$

where β_s is the scattering coefficient and β_a is the absorption coefficient of light. The effects of haze are highly correlated with the range of the underwater scene. In this paper, we simplify the situation as at a certain water depth, the transmission t is defined only by the distance between camera and scene (see Fig. 1).



(a)



(b)

Fig. 1. Underwater optical imaging model; (a) underwater imaging model; (b) underwater imaging reconstruction pipeline.

3 Underwater Image Processing

Homomorphic Filtering. The homomorphic filter is used to correct non-uniform illumination and to enhance contrast of the image. Assume the captured image is a function of the product of the illumination and the reflectance as

$$f(x, y) = i(x, y)r(x, y) \quad (3)$$

where $f(x, y)$ is the captured image, $i(x, y)$ is the illumination multiplicative factor, and $r(x, y)$ is the reflectance function. Taking the logarithm to achieve,

$$\ln(f(x, y)) = \ln(i(x, y)) + \ln(r(x, y)) \quad (4)$$

Compute FFT of the Eq. (4),

$$F(w_x, w_y) = I(w_x, w_y) + R(w_x, w_y) \quad (5)$$

Then, utilize the High-pass Filters to the coefficients of FFT. And after inverse-FFT, we can get the filtered images. The processed images also contains some noise, in the next subsection, we use the DTC-wavelet transform based denoising method [5].

Transmission Estimation. After denoising, we use the de-scattering method to remove the haze in turbidity particles. According to recent researches, we found that the red colour channel is attenuated at a much higher rate than the green or blue channel. We further assume that the transmission in the water is constant. We denote the patch's transmission as t_c . Take the maximum intensity of the red colour channel to compare with the maximum intensity of the green and blue colour channels. We define the dark channel $J_{\text{dark}}(x)$ for the underwater image $J(x)$ as,

$$J_{\text{dark}}(x) = \min_{c \in \{r, g, b\}} (\min_{y \in \Omega(x)} J_c(x)) \quad (6)$$

where $J_c(x)$ refers to a pixel x in colour channel $c \in \{r\}$ in the observed image, and Ω refers to a patch in the image. The dark channel is mainly caused by three factors, shadows, colourful objects or surfaces and dark objects or surfaces.

Here, take the min operation in the local patch on the haze imaging Eq. (1), we assume the transmission as:

$$\min_c(I_c(x)) = \tilde{t}_c(x) \min_{y \in \Omega(x)} (J_c(x)) + (1 - \tilde{t}_c(x))A_c \quad (7)$$

Since A_c is the homogeneous background light and the above equation perform one more min operation among all three colour channels as follows:

$$\min_c \min_{y \in \Omega(x)} \left(\frac{I_c(x)}{A_c} \right) = \tilde{t}_c(x) \min_c \left(\min_{y \in \Omega(x)} \left(\frac{J_c(x)}{A_c} \right) \right) + (1 - \tilde{t}_c(x)) \quad (8)$$

In Ref, let us set $V(x) = A_c(1-t(x))$ as the transmission veil, $W = \min_c(I_c(x))$ is the min colour components of $I(x)$. We have $0 \leq V(x) \leq W(x)$. For grayscale image, $W = I$. Utilize the guided trigonometric bilateral filter (GTBF), which will be discussed in the next subsection. We compute the $T(x) = \text{median}(x) - GTBF_{\Omega}(I W - \text{median}(x))$. And then, we can acquire the by $V(x) = \max\{\min[wT(x), W(x)], 0\}$, here w is the parameter in $(0,1)$. Finally, the transmission of each patch can be written as,

$$\tilde{t}(x) = 1 - \frac{V(x)}{A_c} \tag{9}$$

The background A_c is usually assumed to be the pixel intensity with the highest brightness value in an image. However, in practice, this simple assumption often renders erroneous results due to the presence of self-luminous organisms. So, in this paper, we compute the brightest pixel value among all local min corresponds to the background light A_c as follows:

$$A_c = \max_{x \in I} \min_{y \in \Omega(x)} I_c(y) \tag{10}$$

where $I_c(y)$ is the local colour components of $I(x)$ in each patch.

Guided Median Filtering Refinement. In the above subsection, we roughly estimated the camera-object distance $d(x)$. This distance depth contains mosaic effects and produces less accurately. Consequently, we need to use the proposed weighted guided median filter to reduce the mosaicking. In this section, we introduce our constant time algorithm for weighted guided median filter.

The traditional median filter [10] has been considered as an effective way of removing “outliers”. The traditional median filter usually leads to morphological artifacts like rounding sharp corners. To address this problem, the weighted median filter has been proposed. The weighted median filter is defined as

$$h(\mathbf{x}, i) = \sum_{\mathbf{y} \in N(\mathbf{x})} W(\mathbf{x}, \mathbf{y}) \delta(V(\mathbf{y}) - i) \tag{11}$$

where $W(\mathbf{x}, \mathbf{y})$ corresponds to the weight assigned to a pixel \mathbf{y} inside a local region centered at pixel x , the weight $W(\mathbf{x}, \mathbf{y})$ depends on the image d that can be different from V . $N(\mathbf{x})$ is a local window near pixel \mathbf{x} . i is the discrete bin index, and δ is the Kronecker delta function, δ is 1 when the argument is 0, and is 0 otherwise.

Then the compute the refined depth map by weighted guided median filter is defined as:

$$I_x^{WG} = \frac{\sum_{y \in N(x)} f_S(x, y) f_R(I_x, I_y) I_y W_y}{\sum_{y \in N(x)} f_S(x, y) f_R(I_x, I_y) W_y} \tag{12}$$

where y is a pixel in the neighborhood $N(x)$ of pixel x . Note that kernels other than Gaussian kernels are not excluded.

$$f_S(x, y) = v(x - y) = \frac{1}{2} e^{-\frac{(x-y)(x+y)}{2\sigma_D^2}} \quad (13)$$

where x and y denote pixel spatial positions. The spatial scale is set by σ_D . The range filter weights pixels based on the photometric difference,

$$f_R(I_x, I_y) = w(f(x) - f(y)) = \frac{1}{2} e^{-\frac{(f(x)-f(y))(f(x)+f(y))}{2\sigma_R^2}} \quad (14)$$

where $f(\cdot)$ is image tonal values. The degree of tonal filter is set by σ_R . W_y is the weight map, which is defined as:

$$W_y = \sum_{y \in N(x)} f_s(y, q) f_R(y, q) e^{-\|I_y - I_q\|_2 / 2\sigma_R} \quad (15)$$

where q is the coordinate of support pixel centered around pixel y . The final refined depth map is produced by:

$$h(\tilde{d}(x), i) = \sum_{y \in N(x)} I_x^{WG}(d(x), x) \delta(V(x) - i) \quad (16)$$

This filters images, preserving edges and filters noise based on a dimensionality reduction strategy, having high quality results, while achieving significant speedups over existing techniques.

Recovering the Scene Radiance. With the transmission depth map, we can recover the scene radiance according to Eq. (1). We restrict the transmission $t(x)$ to a lower bound t_0 , which means that a small certain amount of haze are preserved in very dense haze regions. The final scene radiance $J(x)$ is written as,

$$J(x) = \frac{I_c(x) - A_c}{\max(t(x), t_0)} + A_c \quad (17)$$

Typically, we choose $t_0 = 0.1$. The recovered image may be too dark.

Deblurring. In the underwater environment, we must use artificial light for imaging. However, it will cause the vignetting effect. In Ref. [6], K. Sooknanan et al. proposed a multi-frame vignetting correction model for removing the vignetting phenomenon which involves estimating the light source footprint on the seafloor. This artificial light correction can well done, however, it cost large time for computing. So, in this paper, we intend to introduce a signal frame-based vignette removal method. Given the fact that we are interested in the overall effect of light attenuation through the system and not all of the image formation details, we have derived an effective degradation model, $Z(r, \theta)$ as follows,

$$Z(r, \theta) = O(r, \theta) V(r) \quad (18)$$

where Z is the image with vignetting, O is the vignetting-free image, and V is the vignetting function. Our goal is to find the optimal vignetting function V that minimizes asymmetry of the radial gradient distribution. By taking the log of Eq. (26), we get

$$\ln(Z(r, \theta)) = \ln O(r, \theta) + \ln V(r) \tag{19}$$

Let $\mathbf{Z} = \ln Z$, $\mathbf{O} = \ln O$, and $\mathbf{V} = \ln V$. We denote the radial gradients of \mathbf{Z} , \mathbf{O} , and \mathbf{V} for each pixel (r, θ) by $R_r^Z(r, \theta)$, $R_r^O(r, \theta)$, $R_r^V(r, \theta)$. Then,

$$R_r^Z(r, \theta) = R_r^O(r, \theta) + R_r^V(r, \theta) \tag{20}$$

Given an image \mathbf{Z} with vignetting, we find a maximum a posterior (MAP) solution to \mathbf{V} . Taking Bayes rule, we get,

$$\mathbf{V} = \arg \max_{\mathbf{V}} P(\mathbf{V}|\mathbf{Z}) \sim \arg \max_{\mathbf{V}} P(\mathbf{Z}|\mathbf{V})P(\mathbf{V}) \tag{21}$$

Considering the vignetting function at discrete, evenly sampled radii: $(\mathbf{V}(r_t), r_t \in S_r)$, where $S_r = \{r_0, r_1, \dots, r_{n-1}\}$. Each pixel (r, θ) is associated with the sector in it resides, and sector width is δr . The vignetting function is in general smooth, therefore, we obtain,

$$P(\mathbf{V}) = e^{-\lambda_s \sum_{r_t \in S_r} V''(r_t)^2} \tag{22}$$

where λ_s is chosen to compensate for the noise level in the image, and $\mathbf{V}''(r_t)$ is approximated as

$$V''(r_t) = \frac{V(r_{t-1}) - 2V(r_t) + V(r_{t+1}))}{(\delta r)^2} \tag{23}$$

Using the sparsity prior method on the vignetting-free image \mathbf{O} ,

$$P(\mathbf{Z}|\mathbf{V}) = P(R_r^O) = e^{-|R_r^O|^\alpha}, < 1 \tag{24}$$

Substituting Eq. (32) and Eq. (28), we have

$$P(\mathbf{Z}|\mathbf{V}) = e^{-\sum_{(r, \theta)} |R_r^Z(r, \theta) - R_r^V(r)|^\alpha} \tag{25}$$

The overall energy function $P(\mathbf{Z}|\mathbf{V})P(\mathbf{V})$ can be written as

$$E = \sum_{(r, \theta)} |R_r^Z(r, \theta) - R_r^V(r)|^\alpha + \lambda_s \sum_{r_t \in S_r} V''(r_t)^2 \tag{26}$$

Through minimize E , we can estimate the $\mathbf{V}(r_t)$. Then, we use the IRLS technique for estimating the vignetting function.

Spectral Properties-Based Colour Correction. We take the chromatic transfer function τ for weighting the light from the surface to a given depth of objects as

$$\tau_\lambda = \frac{E_\lambda^{surface}}{E_\lambda^{object}} \quad (27)$$

where the transfer function τ at wavelength λ is derived from the irradiance of the surface $E_\lambda^{surface}$ by the irradiance of the object E_λ^{object} . According to the spectral response of RGB camera, we convert the transfer function to RGB domain:

$$\tau_{RGB} = \sum^k \tau_\lambda \cdot C_b(\lambda) \quad (28)$$

where the weighted RGB transfer function is τ_{RGB} , $C_b(\lambda)$ is the underwater spectral characteristic function of colour band b , $b \in \{r, g, b\}$. k is the number of discrete bands of the camera spectral characteristic function.

Finally, the corrected image is gathered from the weighted RGB transfer function by

$$J_\lambda(x) = \hat{J}_\lambda(x) \cdot \tau_{RGB} \quad (29)$$

where $J_\lambda(x)$ and $\hat{J}_\lambda(x)$ are respectively the colour corrected and uncorrected images.

4 Experimental Results and Discussions

The performance of the proposed algorithm is evaluated both analytically and experimentally by utilizing ground-truth colour patches. Both results demonstrate that the proposed algorithm has superior haze removal effects and colour balancing capabilities of the proposed method over the others. In the experiment, we compare our method with Fattal's model [7], He's model [8], and Xiao's model [9]. Here, we select the best parameters for each model. The computer used is equipped with Windows XP and an Intel Core 2 (2.0 GHz) with 1 GB RAM. The size of the images is 345×292 pixels.

In the first experiment, we simulate the mine detection system in the darkroom of our laboratory (see Fig. 2). We take OLYMPUS STYLUS TG-2 15 m/50ft underwater camera for capture images [13]. The distance between light-camera and the objects is 3 meters. The size of the images is 640×480 pixels. We take the artificial light as an auxiliary light source. As a fixed light source, it caused uneven distribution of light. Because the light is absorbed in water, the imaging array of the camera captured a distorted video frame, see Fig. 2(a). Figure 2(b) shows the denoised image, electrical noise and additional noise are removed. After estimation, we use single frame vignetting method to remove artificial light. And the dehazing method is proposed to eliminate the haze in the image. After that, the contrast of Fig. 2(d) is obviously than Fig. 2(c). The obtained image is also too dark. So, α ACE is used to enhancement the image. And finally, the sharp-based recognition method is used to distinguish the objects. Fig. 3 shows the results of devignetting by different methods.

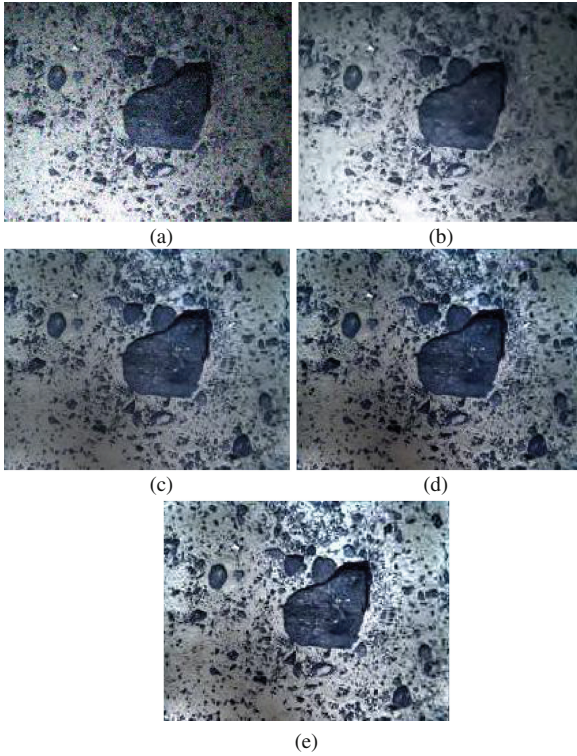


Fig. 2. Simulation of objects detection in water tank; (a) Captured video frame; (b) denoised image; (c) de-vignetting; (d) GMF dehazing; (e) colour correction.

In addition to the visual analysis mentioned above, we conducted quantitative analysis, mainly from the perspective of statistics and the statistical parameters for the images (see Table 1). This analysis includes Peak Signal to Noise Ratio (PSNR) [11], Quality mean-opinion-score (Q-MOS) [6], and Structural Similarity (SSIM) [12].

Table 1. Comparative analysis of different de-scattering methods.

Methods	PSNR	Q-MOS	SSIM	CPU time [s]
Fattal	11.21	46.11	0.4225	20.61
He	16.94	88.45	0.6992	10.66
Xiao	24.63	92.38	0.8030	5.21
Ours	29.87	94.58	0.8963	5.46

Let x_i and y_i be the i -th pixel in the original image A and the distorted image B , respectively. The MSE and $PSNR$ between the two images are given by

$$MSE = \frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2, \quad (30)$$

$$PSNR = 10 \log_{10} \left(\frac{L^2}{MSE} \right) \quad (31)$$



(a)



(b)



(c)

Fig. 3. Results of different devisnetting methods; (a) input image; (b) histogram equation; (c) our method.

where, PSNR means the peak signal to noise ratio (values are over 0, the higher the best).

Reference proposed a multi-scale SSIM method for image quality assessment. Input image A and B , let μ_A , μ_B , and σ_{AB} respectively as the mean of A , the mean of B , the covariance of image A and image B . The parameters of relative importance α , β , γ are equal to 1. The SSIM is given as follow:

$$SSIM(x, y) = \frac{(2\mu_A\mu_B + C_1)(2\sigma_{AB} + C_2)}{(\mu_A^2 + \mu_B^2 + C_1)(\mu_A^2 + \mu_B^2 + C_2)} \quad (32)$$

where C_1, C_2 are the small constants. SSIM is named as structural similarity (values are between 0 (worst) to 1 (best)).

The objective quality predictions do not map directly to the subjective mean opinion scores (MOS). There is a non-linear mapping function between subjective and objective predictions. In a novel logistic function to account for such a mapping is proposed,

$$Q - MOS = \frac{100}{1 + \exp(q_1(Q + q_2))} \quad (33)$$

where, Q is the multi-band pooling that produced the strongest correction with the LIVE database. q_1 and q_2 represent the different observers. The Q-MOS value is between 0 (worst) to 100(best). Table 1 displays the numerical results of Q-MOS, PSNR and SSIM measured on several images. In this paper, we first transfer the RGB image to gray image, and then take the mathematical indexes for evaluating the different methods. The results indicate that our approach works well for haze removal.

5 Conclusions

This work has shown that it is possible to enhance degraded video sequences from seabed observation systems using the image processing ideas. The proposed algorithm is automatic and requires little parameters adjustment. Total computing time of our system is about 10 s. We proposed a simple prior based on the difference in attenuation among the different colour channels, which inspire us to estimate the transmission map. Another contribution is to compensate the transmission by guided median filters, which not only has the benefits of edge-preserving and noise removing, but also speed up the computational cost. Meanwhile, the proposed underwater image colourization method can recover the underwater distorted images well than the state-of-the-art methods. The artificial light correction method can eliminate the non-uniform illumination very well.

Acknowledgments. This work was supported by Grant in Aid for Research Fellows of Japan Society for the Promotion of Science (No.13J10713), Grant in Aid for Foreigner Research Fellows of Japan Society for the Promotion of Science (No.15F15077), Open Research Fund of State Key Laboratory of Marine Geology in Tongji University (MGK1407), and Open Research Fund of State Key Laboratory of Ocean Engineering in Shanghai Jiaotong University (OEK1315).

References

1. Stack, J.: Automation for underwater mine recognition: current trends and future strategy. Proc. SPIE **80170K**, 1–21 (2011)
2. Schechner, Y.Y., Averbuch, Y.: Regularized image recovery in scattering media. IEEE Trans. Pattern Anal. **29**(9), 1655–1660 (2007)

3. Hou, W., Gray, D.J., Weidemann, A.D., Fournier, G.R., Forand, J.L.: Automated underwater image restoration and retrieval of related optical properties. In: Proceedings of International Symposium. of Geoscience and Remote Sensing, pp. 1889–1892 (2007)
4. Narasimhan, S.G., Nayar, S.K.: Vision and the atmosphere. *Int. J. Comput. Vis.* **48**(3), 233–254 (2002)
5. Selesnick, I.W., Baraniuk, R.G., Kingsbury, N.G.: The dual-tree complex wavelet transform. *IEEE Signal Process. Mag.* **22**(6), 123–151 (2005)
6. Sooknanan, K., Kokaram, A., Corrigan, D., et al.: Improving underwater visibility using vignetting correction. *Proc. SPIE* **8305**, 83050M-1–83050M-8 (2012)
7. Fattal, R.: Single image dehazing. In: SIGGRAPH, pp. 1–9 (2008)
8. He, K., Sun, J., Tang, X.: Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal.* **33**(12), 2341–2353 (2011)
9. Xiao, C., Gan, J.: Fast image dehazing using guided joint bilateral filter. *Visual Comput.* **28**(6–8), 713–721 (2012)
10. Lu, H., Li, Y., Serikawa, S.: Underwater image enhancement using guided trigonometric bilateral filter and fast automation color correction. In: Proceedings of International Conference on Image Processing, pp. 3412–3416 (2013)
11. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**(4), 600–612 (2004)
12. Zheng, Y., Lin, S., Kang, S.B., Xiao, R., Gee, J.C., Kambhamettu, C.: Single-image vignetting correction from gradient distribution symmetries. *IEEE Trans. Pattern Anal.* **35**(6), 1480–1494 (2013)
13. Lu, H., Li, Y., Zhang, L., Serikawa, S.: Contrast enhancement for images in turbid water. *J. Opt. Soc. Am.* **32**(5), 886–893 (2015)