An Approach to Synthesize Diverse Underwater Image Dataset

Xiaodong LIU[1], Ben M. CHEN[1, 2]

(1. *National University of Singapore*, *Singapore*, 119077;

2. *Chinese University of Hong Kong*, *Hong Kong*, 999077)

Abstract: Images that are taken underwater mostly present color shift with hazy effects due to the special property of water. Underwater image enhancement methods are proposed to handle this issue. However, their enhancement results are only evaluated on a small number of underwater images. The lack of a sufficiently large and diverse dataset for efficient evaluation of underwater image enhancement methods provokes the present paper. The present paper proposes an organized method to synthesize diverse underwater images, which can function as a benchmark dataset. The present synthesis is based on the underwater image formation model, which describes the physical degradation process. The indoor RGB-D image dataset is used as the seed for underwater style image generation. The ambient light is simulated based on the statistical mean value of real-world underwater images. Attenuation coefficients for diverse water types are carefully selected. Finally, in total 14490 underwater images of 10 water types are synthesized. Based on the synthesized database, state-of-the-art image enhancement methods are appropriately evaluated. Besides, the large diverse underwater image database is beneficial in the development of learning-based methods.

Key words: Image Processing; Underwater Image Enhancement; Underwater Image Synthesis

1 Introduction

Many underwater tasks are beneficial in both scientific research and civil activities such as underwater pipeline inspection [1], coral reef monitoring [2], ship hull inspection [3], and the search or study of the underwater objects [4].

All the vision-related underwater tasks rely on the clear underwater images. However, unlike optical imaging in air, capturing images underwater poses unique difficulties due to the light absorption and light scattering underwater. Consequently, images that are taken underwater mostly present color shift with decreased visibility. The degraded quality of the underwater image may hamper and limit the performance of those tasks mentioned above. Two sample underwater images are given in Fig.1.

Fig. 1 Two samples of undersea images [12].

Underwater image enhancement methods are developed to cope with the above-mentioned quality degradation. Two major quality degradations are caused by light attenuation and light backscatter, which will result in image color cast with hazy effects.

Many pioneering approaches seek to handle this problem. In the early stage, underwater image enhancement methods are mainly hardware-based. For example, the backscatter was attempted to be removed by polarization or the precise control of the shutter gate of the camear [5-7].

Besides, since the image degradation underwater is similar to that in hazy scenes in part, many approaches are tailored based on the haze removal methods, for example, the influential dark channel prior [8], and its modifications [9-11].

Alternatively, deep learning has achieved great success in many high-level computer vision tasks. However, there are limited deep learning-based methods for underwater image correction, comparatively. This is partly because of the lack of sufficient data for network training [13].

Besides, there also lacks a fair benchmark to evaluate the image enhancement methods, where most methods only manage to test their performance on certain monotonous images. In this case, the generalization ability could not be verified.

Thus, the motivation of the present paper for proposing a benchmark underwater image dataset is two-fold:

1) It can be used as the benchmark to make comparisons with different underwater image enhancement methods sufficiently and fairly;

2) The benchmark can be utilized to train the deep convolutional neural network, thus pushes the research in this direction.

The present paper proposes a systematic and general scheme to synthesize diverse underwater images. The underwater image simulation imitates the image degradation process in water, which is based on an underwater image formation model. Therefore, in the present paper, first the image formation process is studied, and the parameters used for image generation are calculated and collected. Based on the synthesized underwater database, the paper also evaluates the state of the art methods for underwater image enhancement.

2 Related Work

2.1 Image formation model

Many researchers have made attempts to model the physical process of image formation. One of the most widely used models in this community is the Jaffe-McGlamery model [14-15], and the captured image (total radiance received by the observer E_T) consists of the direct illumination E_d, the forward scattering E_f, and the backscatter E_b:

$$E_T = E_d + E_f + E_b$$

The image formation process is depicted in Fig. 2. Besides, due to the less impact of the forward scattering quantitively, E_f normally is excluded. To

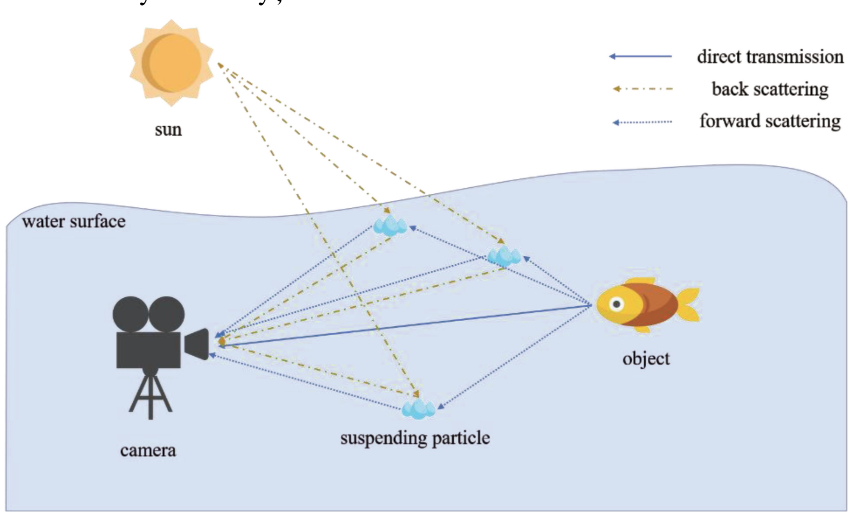

Fig. 2 General view of the underwater image formation process.

get a more explicit expression of Eq. (1), afterward, we would give more information about the formation process.

Denote $E(d,\lambda)$ as the ambient light on a small disk d_z , the scattered radiance on the disk be expressed as[15]

$$dL(z,\lambda) = b(\lambda)E(d,\lambda)dz$$

where $b(\lambda)$ denote the scattering coefficient.

Based on Eq. (2), the radiance received at distance $zdB(z,\lambda)$ from the small disk on the sensor is exponentially attenuated and can be formulated

$$dB(z,\lambda) = dL(z,\lambda)e^{-\beta(\lambda)z}$$

where $\beta(\lambda)$ is the attenuation coefficient.

Mathematically integrating from 0 to z for Eq. (3), the backscatter $B(z,\lambda)$ is obtained as:

$$B(z,\lambda) = \frac{b(\lambda)E(d,\lambda)}{\beta(\lambda)}(1 - e^{-\beta(\lambda)z})$$

When z is large enough, the backscatter can be expressed

$$B^{\infty}(\lambda) = \frac{b(\lambda)E(d,\lambda)}{\beta(\lambda)}$$

Therefore, the total radiance as the summation of the exponential attenuated signal and the backscatter can be expressed as

$$E_T = E(d,\lambda)e^{-\beta(\lambda)z} + B^{\infty}(\lambda)(1 - e^{-\beta(\lambda)z})$$

Denote the $\rho(\lambda)$ and $S_c(\lambda)$ as the object reflectance spectrum and spectral response, the total radiance from the object I can be expressed as

$$I = \frac{1}{\kappa}\int_{\lambda_1}^{\lambda_2} S_c(\lambda)\rho(\lambda)E(d,\lambda)e^{-\beta(\lambda)z}\mathrm{d}\lambda + \frac{1}{\kappa}\int_{\lambda_1}^{\lambda_2} S_c(\lambda)B^{\infty}(\lambda)(1 - e^{-\beta(\lambda)z})\mathrm{d}\lambda$$

where κ denotes the pixel geometry in the camera, and λ_1, λ_2 are the lower and upper spectrum bounds for integration, respectively.

More specially, the first term is also termed as the scene radiance J_c

$$J_c = \frac{1}{\kappa}\int_{\lambda_1}^{\lambda_2} S_c(\lambda)\rho(\lambda)E(d,\lambda)\mathrm{d}\lambda$$

The backscatter sensed by the camera is

$$B_c^{\infty} = \frac{1}{\kappa}\int_{\lambda_1}^{\lambda_2} S_c(\lambda)B_c^{\infty}\mathrm{d}\lambda$$

Finally, the underwater image formation model can be summarized as

$$I = J_c e^{-\beta_c(\lambda)z} + B_c^{\infty}(1 - e^{-\beta_c(\lambda)z}), c \in \{r,g,b\}$$

Based on Eq. (10), we could find that the key parameters are $\beta(\lambda)$, B_c^{∞} and z, which also play an important role in underwater image synthesis.

2.2 Existing underwater image database

Unlike images in the air, it is quite difficult to get the ground-truth or reference images for underwater scenes. Ideally, the ground-truth underwater images should be captured when the water is removed, which is nearly impossible for the ocean or the sea scenarios.

Instead, the intuitive idea has been implemented by some researchers in the controlled environment, for example, in the man-made water tank. In [19], the authors built a 1000L water tank with stones and objects on the floor. Milk was added into the tank to increase the water turbidity. In [20], the authors proposed the OUC-VISION dataset with 4400 images for salient object detection. A cube (1.5 m * 0.5 m * 1.5 m) was set up for image capturing with various illumination and turbidity. Similarly, authors in [21] captured 6240 images with the variation of lighting condition, turbidity, and depth in a man-made tank whose dimension was 1.5 m * 0.5 m * 1.5 m.

Although the methods mentioned above provide several underwater-style images, however, the diversity is limited by the size of the water tank. As shown in Eq. (10), both the attenuation and backscatter are dependent on the range or depth. Nevertheless, with the small-size man-made water tank, the image could only be captured in a very short range. Therefore, the short-range images may lose the characteristics of the physical image degradation.

Alternatively, researchers have endeavored to synthesize underwater images established on the physical image formation model, given B_c^{∞} , $\beta_c(\lambda)$ and the depth z. Nevertheless, there exist some limitations in their parameter settings. For example, the authors only vary the depth ranged from 0.5 m to 3 m in [22]. In [10], the image depth was fixed at

5 m. The limited depth range would result in a loss of diversity. What's worse, the depth value was randomly set for each pixel in the image in [10] and [22], which would be inconsistent with the image structure as well. In [23], the ambient light was arbitrarily selected but the values for three color channels were used as the same. However, unlike the ambient light in hazy scenarios, light is wavelength-dependent attenuated underwater, resulting in different light intensity in three color channels.

3 Our Method

In this section, we will present our systematic scheme to synthesis the underwater image benchmark dataset based on the image formation model. As discussed earlier, z, B_c^{∞}, $\beta_c(\lambda)$ are the unknown key parameters in the image formation model.

3.1 Depth z

To our knowledge, there lack theunderwater images with ground-truth depth information. Therefore, the indoor RGB-D dataset would be a good choice, providing the actual depth information for the RGB images. In the present paper, the NYU-V2 indoor dataset [25] is utilized for underwater image synthesis. This dataset provides 1449 indoor images labeled with corresponding depth values, which are in the range from 0 to 10 m.

3.2 Attenuation coefficients $\beta_c(\lambda)$

The attenuation coefficients are the summation of the wave length-dependent attenuation and scattering coefficients in water, which are related to various factors for instance the type of water, suspended particles, and so on. Therefore, $\beta_c(\lambda)$ should vary with the wavelength as well as the type of water [24].

Jerlov et al. have classified water into the coastal type and oceanic type. Further, the oceanic water can be classified into 5 types (Type-I, Type-IA, Type-IB, Type-II, Type-III) based on the clarity of the water. The coastal water can be further classified into five levels of turbidity (Type-1C, Type-3C, Type-5C, Type-7C, Type-9C). Besides, they provide the attenuation coefficients for different water types with different wavelengths. The coefficients for RGB color channels for all water types are listed in Table 1.

Table 1 Attenuation coefficients of the 10 water types [24].

Type	I	IA	IB	II	III	IC	3C	5C	7C	9C
Red	0.341	0.342	0.349	0.375	0.426	0.439	0.498	2.43	0.635	0.755
Green	0.049	0.0503	0.0572	0.129	0.121	0.187	0.315	0.73	0.494	0.777
Blue	0.021	0.0253	0.0325	0.110	0.139	0.240	0.400	0.65	0.693	1.240

The light absorption is simulated based on the coefficients for all 10 types of water in Fig.3, where the depth is ranged from 0 to 20 m. From Fig.3, we could have the following observations. Around 5 types of water will become dark when the depth increases to 10 m. As the depth range of the NYU-V2 dataset is from 0 to 10 m, the depth range is relatively enough for diversity. Besides, Fig. 3 also verifies that the limited depth range would fail to simulate diverse images.

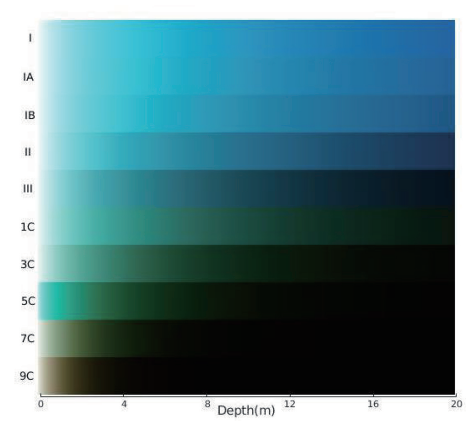

Fig. 3 Simulation results for light absorption underwater.

3.3 Ambient light B_c^{∞}

Although the underwater image formation model is analogous to that in the hazy scenario, one of the major differences lies in that the wavelength-dependent light absorption. Under this circumstance, the ambient light for different color channels is different.

In the present paper, we would first calculate the ambient light for a large number of underwater images, and the ambient light is set based on the statistical mean value. The calculation method is based on[8] for ambient light, and 3000 underwater images downloaded from the Internet were adopted. The average for ambient light for three channels is: $B_r^\infty = 0.6703$, $B_g^\infty = 0.7807$, $B_b^\infty = 0.7577$.

To add diversity, an additional 5% variation is added to the mean value for image synthesis.

3.4 Synthesizing underwater images

In total, we have synthesized 14490 images and 1449 underwater-style images for eachwater type. One sample is shown in Fig. 4. From Fig. 4, we could find the synthesized images follow the physical degradation process with the color cast and limited visibility. More specifically, the degradation would increase as the depth increase. For example, the synthesized type-9C image appears hazier on the right side compared to the left due to increased depth on the right, which is consistent with our previous analysis.

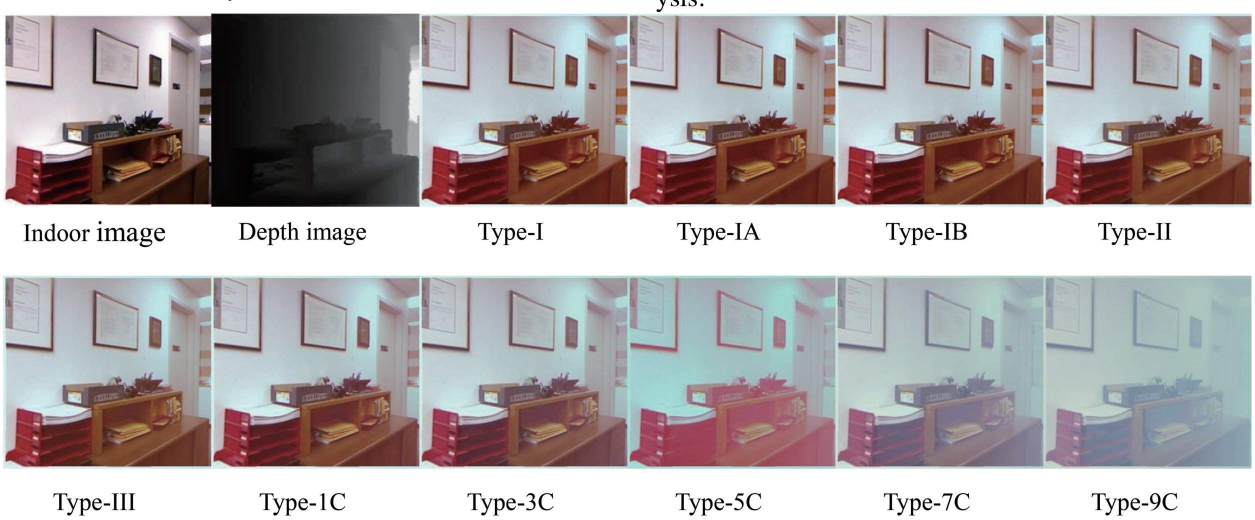

Fig. 4 One sample of the synthesized underwater images.

4 Methods Evaluation Results

Based on the generated benchmark images, we can evaluate the state-of-the-art methods for underwater image enhancement. The following image enhancement methods which are published in the top conference or journals are compared: UVE (published in ICRA, 2017) [26], MBIE [27] (published in ICRA, 2018), Fusion [12] (published in CVPR, 2012), UHL [28] (published in BMVC, 2017), WCID [17] (published in TIP, 2012) and DCP [8] (published in CVPR, 2009).

Concerning the evaluation criteria for image quality, it is still an open problem to develop a good and unified metric for image quality evaluation. In the present paper, we adopt the full-reference image metric: peak-signal-to-noise-ratio (PSNR) and structural similarity index (SSIM), which are widely used for image quality evaluation. For both metrics, a higher value indicates a more desirable outcome.

The average value of PSNR and SSIM are reported in Table 2 and Table 3, respectively. Considering the page limitation, only 5 water types are presented in Fig. 5. From Fig. 5, we could find that the fusion-based enhancement method [12] could give the best visual pleasing results, as well as the highest

PSNR and SSIM against all the other methods.

Table 2 PSNR results for the state of the art methods.

Method\Type	I	IA	IB	II	III	IC	3C	5C	7C	9C
UVE	15.3886	15.3996	15.4013	15.3451	15.0771	14.9807	14.4290	12.6667	12.6398	11.6589
MBIE	16.0667	16.0794	16.0866	16.1386	15.2471	16.2771	16.2385	12.8334	13.8970	11.1727
Fusion	20.6516	20.6316	20.5499	20.1954	19.4084	19.1473	17.8650	13.5206	14.1309	11.9211
WCID	11.6825	11.6638	11.6250	11.4551	11.0293	10.9091	10.1878	8.8672	8.0597	7.2613
UHL	13.7663	13.9279	13.8241	13.8062	14.0902	14.1346	14.1230	12.7527	14.0028	13.2602
DCP	13.6527	13.6537	13.6392	13.5804	13.4599	13.3482	13.1819	11.5328	14.4655	10.9678

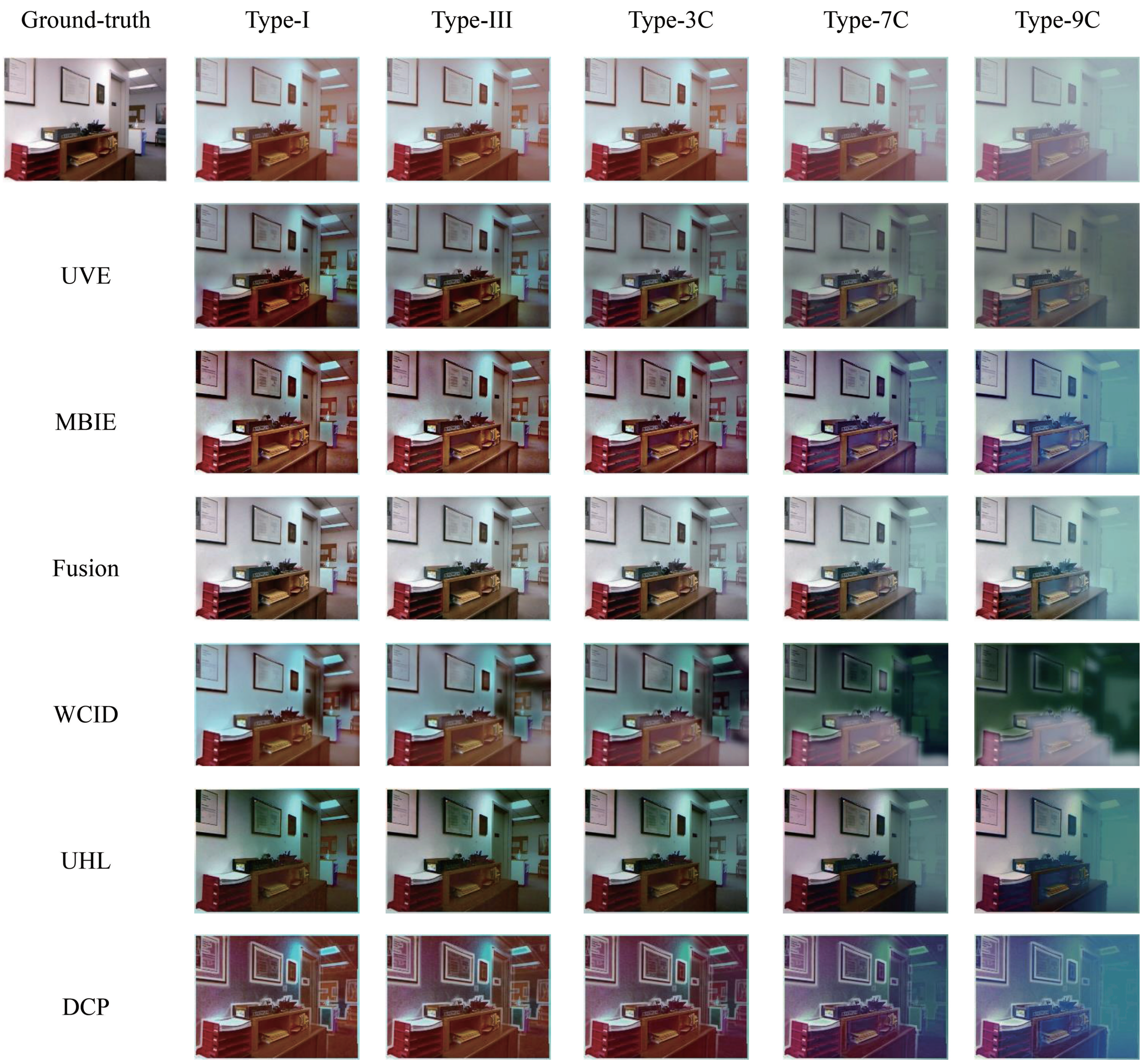

Fig. 5 Methods evaluation of the proposed benchmark dataset.

Table 3 SSIM results for the state of the art methods.

Method\Type	I	IA	IB	II	III	IC	3C	5C	7C	9C
UVE	0.6028	0.6035	0.6013	0.5901	0.5623	0.5510	0.5000	0.2892	0.2539	0.1079
MBIE	0.5653	0.5670	0.5677	0.5744	0.5875	0.5943	0.6112	0.4235	0.3349	0.0222
Fusion	0.7489	0.7498	0.7493	0.7461	0.7351	0.7304	0.6960	0.4328	0.4114	0.2014
WCID	0.3606	0.3610	0.3606	0.3578	0.3476	0.3423	0.3072	0.1962	0.0353	0.0192
UHL	0.4079	0.4124	0.4094	0.4147	0.4106	0.4205	0.4257	0.2708	0.2756	0.1236
DCP	0.5274	0.5290	0.5294	0.5317	0.5363	0.5371	0.5461	0.3755	0.2550	0.0684

Besides, almost all methods could only handle the intermediate turbidity. We could observe the performance decrease from water type-5C to type-9C with the increase of the turbidity. Some researchers have also reported such difficulty for the restoration of the highly degraded images due to the low signal to noise ratio. There still needs further efforts and attention to tackle this challenge.

5 Conclusion

Confronted with the lack of a benchmark dataset for underwater images, in the present paper, we propose a systematic scheme to produce the large undersea images established on the physical image formation model. NYU-V2 indoor dataset is adopted. The ambient light is estimated by the statistical value of the underwater images. Based on 10 types of underwater attenuation coefficients, 1449 images are synthesized for each type and overall 14490 images in total.

Underwater image enhancement methods are fairly evaluated based on the proposed benchmark dataset where the fusion-based method [12] achieved the best both visually and quantitatively. Besides, we also demonstrated the remaining challenge for the restoration of the severely degraded images. Future work could investigate the deep-learning-based methods based on the proposed benchmark dataset, as well as the image restoration for the challenging scenes.

References

[1] Jacobi, M., &Karimanzira, D. (2013, June). Underwater pipeline and cable inspection using autonomous underwater vehicles. In 2013 *MTS/IEEE OCEANS-Bergen* (pp. 1-6). IEEE.

[2] Johnson-Roberson, M., Bryson, M., Friedman, A., Pizarro, O.,Troni, G., Ozog, P., & Henderson, J. C. (2017). High-resolution underwater robotic vision-based mapping and three-dimensional reconstruction for archaeology. *Journal of Field Robotics*, 34 (4), 625-643.

[3] Hover, F. S., Eustice, R. M., Kim, A.,Englot, B., Johannsson, H., Kaess, M., & Leonard, J. J. (2012). Advanced perception, navigation, and planning for autonomous in-water ship hull inspection. *The International Journal of Robotics Research*, 31 (12), 1445-1464.

[4] Bryson, M., Johnson-Roberson, M., Pizarro, O., & Williams, S. (2013, November). Automated registration for multi-year robotic surveys of marinebenthic habitats. In 2013 *IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 3344-3349). IEEE.

[5] Yemelyanov, K. M., Lin, S. S., Pugh Jr, E. N., & Engheta, N. (2006). Adaptive algorithms for two-channel polarization sensing under various polarization statistics with nonuniform distributions. *Applied optics*, 45(22), 5504-5520.

[6] Tan, C. S.,Sluzek, A., GL, G. S., & Jiang, T. Y. (2007, May). Range gated imaging system for underwater robotic vehicle. In *OCEANS* 2006-*Asia Pacific* (pp. 1-6). IEEE..

[7] Pizer, S. M., Amburn, E. P., Austin, J. D., Cromartie, R., Geselowitz, A., Greer, T., ... & Zuiderveld, K. (1987). Adaptive histogram equalization and its variations. *Computer vision, graphics, and image processing*, 39(3), 355-368.

[8] He, K., Sun, J., & Tang, X. (2011). Single image haze removal using dark channel prior.*IEEE transac-*

tions on pattern analysis and machine intelligence, 33(12), 2341-2353.

[9] Cheng, C. Y., Sung, C. C., & Chang, H. H. (2015, October). Underwater image restoration by red-dark channel prior and point spread function deconvolution. In2015 *IEEE International Conference on Signal and Image Processing Applications* (*ICSIPA*) (pp. 110-115). IEEE.

[10] Drews, P., Nascimento, E., Moraes, F., Botelho, S., & Campos, M. (2013). Transmission estimation in underwater single images. In *Proceedings of the IEEE international conference on computer vision workshops* (pp. 825-830).

[11] Łuczyński, T., & Birk, A. (2018). Underwater Image Haze Removal and Color Correction with an Underwater-ready Dark Channel Prior. *arXiv preprint arXiv*:1807.04169.

[12] Ancuti, C., Ancuti, C. O., Haber, T., & Bekaert, P. (2012, June). Enhancing underwater images and videos by fusion. In 2012 *IEEE Conference on Computer Vision and Pattern Recognition* (pp. 81-88). IEEE.

[13] Li, J., Skinner, K. A., Eustice, R. M., & Johnson-Roberson, M. (2018). WaterGAN: unsupervised generative network to enable real-time color correction of monocular underwater images. *IEEE Robotics and Automation Letters*, 3(1), 387-394.

[14] McGlamery, B. L. (1980, March). A computer model for underwater camera systems. In*Ocean Optics VI* (Vol. 208, pp. 221-232). International Society for Optics and Photonics.

[15] Jaffe, J. S. (1990). Computer modeling and the design of optimal underwater imaging systems. *IEEE Journal of Oceanic Engineering*, 15(2), 101-111..

[16] Schechner, Y. Y., & Karpel, N. (2004, June). Clear underwater vision. In *Proceedings of the* 2004 *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2004. *CVPR* 2004. (Vol. 1, pp. I-I). IEEE.

[17] Chandrasekhar, S. (2013).*Radiative transfer*. Courier Corporation.

[18] Huynh, C. P., & Robles-Kelly, A. (2007, December). Comparative colorimetric simulation and evaluation of digital cameras using spectroscopy data. In9*th Biennial Conference of the Australian Pattern Recognition Society on Digital Image Computing Techniques and Applications* (*DICTA* 2007) (pp. 309-316). IEEE.

[19] Codevilla, F., Gaya, J. D. O., Duarte, N., & Botelho, S. (2004). Achieving turbidity robustness on underwater images local feature detection. *International journal of computer vision*, 60(2), 91-110.

[20] Jian, M., Qi, Q., Dong, J., Yin, Y., Zhang, W., & Lam, K. M. (2017, July). Theouc-vision large-scale underwater image database. In 2017 *IEEE International Conference on Multimedia and Expo* (*ICME*) (pp. 1297-1302). IEEE.

[21] Ma, Y., Feng, X., Chao, L., Huang, D., Xia, Z., & Jiang, X. (2018, November). A New Database for Evaluating Underwater Image Processing Methods. In2018 *Eighth International Conference on Image Processing Theory, Tools and Applications* (*IPTA*) (pp. 1-6). IEEE.

[22] Zhao, X., Jin, T., & Qu, S. (2015). Deriving inherent optical properties from background color and underwater image enhancement. *Ocean Engineering*, 94, 163-172.

[23] S. Anwar, C. Li and F.Porikli, (2018) 'Deep underwater image enhancement, arxiv preprint.

[24] Solonenko, M. G., & Mobley, C. D. (2015). Inherent optical properties of Jerlov water types. *Applied optics*, 54(17), 5392-5401.

[25] Silberman, N., Hoiem, D., Kohli, P., & Fergus, R. (2012, October). Indoor segmentation and support inference from rgbd images. In *European Conference on Computer Vision*(pp. 746-760). Springer, Berlin, Heidelberg.

[26] Cho, Y., & Kim, A. (2017, May). Visibility enhancement for underwater visual SLAM based on underwater light scattering model. In2017 *IEEE International Conference on Robotics and Automation* (*ICRA*) (pp. 710-717). IEEE.

[27] Cho, Y., Jeong, J., & Kim, A. (2018). Model-Assisted Multiband Fusion for Single Image Enhancement and Applications to Robot Vision. *IEEE Robotics and Automation Letters*, 3(4), 2822-2829.

[28] Berman, D., Treibitz, T., & Avidan, S. (2017). Diving into haze-lines: Color restoration of underwater images. In *Proc. British Machine Vision Conference* (*BMVC*) (Vol. 1, No. 2).

Authors' Biographies

Xiaodong LIU received his master's degree from Beihang University in 2016. Now he is a Ph.D. candidate at the National University of Singapore. His main research interests include image processing, image enhancement and robotics.
E-mail: xiaodongliu@u.nus.edu

Ben M. CHEN is currently a Professor in the Department of Mechanical and Automation Engineering at the Chinese University of Hong Kong. He was a Provost's Chair Professor in the Department of Electrical and Computer Engineering, the National University of Singapore (NUS), where he was also serving as the Director of Control, Intelligent Systems and Robotics Area, and Head of Control Science Group, NUS Temasek Laboratories. His current research interests are in unmanned systems, robust control and control applications.

Prof. Chen is an IEEE Fellow. He has published more than 400 journal and conferencearticles and authored/co-authored 10 research monographs in systems and control theory, industrial applications, unmanned systems and financial market modeling. He had served on the editorial boards of several international journals including IEEE Transactions on Automatic Control and Automatica. He currently serves as an Editor-in-Chief of Unmanned Systems. Prof. Chen has received a number of research awards nationally and internationally. His research team has actively participated in international unmanned aerial vehicles competitions and won many championships in the contests.
E-mail: bmchen@cuhk.edu.hk