Underwater image dehaze using scene depth estimation with adaptive color correction

Xueyan Ding, Yafei Wang, Jun Zhang, Xianping Fu*
Information Science and Technology College
Dalian Maritime University
Dalian 116026, China
{dingxueyan_meow, wangyafei, william, fxp}@dlmu.edu.cn

***Abstract*—Underwater images suffer from poor visibility due to color casts and light scattering that caused by physical properties existing in underwater environments. Degraded underwater images lead to a low accuracy rate of underwater object detection and recognition. To solve this problem, a novel underwater image enhancement method which combines adaptive color correction and image dehazing based on atmospheric scattering model is proposed in this paper. As the most important component of dehazing model, a transmission map is derived from the color corrected image. Considering the exponential relationship between the transmission map and scene depth map, transmission map estimation will be naturally formulated into a scene depth map estimation problem. To predict scene depth map, a Convolutional Neural Network (CNN) is employed on image patches extracted from the color corrected image. The experimental results show that the proposed strategy improves the quality of underwater images efficiently and arrives at good results in underwater objects detection and recognition.**

***Keywords*—Underwater image dehazing; Scene depth; Convolutional Neural Network; Color correction; Atmospheric scattering model**

I. Introduction

Poor visibility of underwater images captured by cameras results in a difficult task for exploring ocean environments and recognizing underwater objects [1] [2]. To raise the accuracy rate of underwater objects detection and recognition, clear and high quality underwater images is necessary. However, capturing clear and high quality underwater images is challenging because of physical particles in underwater environments. As light travels in the water, it is absorbed and scattered. Light absorption reduces energy of light substantially and leads to color casts of underwater images. The degree of absorption varies with different wavelength of light. Besides, light traveling underwater suffers from two kinds of scattering: forward scattering and backward scattering. The former which deviates light from the object to the camera causes a blur scene, and the latter is a fraction of light reflected by water or floating particles towards the camera before it reaches the object and limits the contrast of the images. Affected by these properties of underwater imaging, it is not easy to acquire clear and high quality underwater images. Hence, an effective underwater image enhancement strategy is necessary for oceanic application of computer vision.

In this paper, an underwater image enhancement framework that involves adaptive color correction and model-based dehazing using scene depth estimation is proposed. The adaptive color correction algorithm with gain factor is firstly applied to degraded underwater image in order to compensate color casts. Then underwater image dehazing based on atmospheric scattering model is applied to color corrected image to remove the blurriness. In atmospheric scattering model, the most important component is the transmission map. Different from traditional transmission map estimation methods, Convolutional Neural Network is carried out to predict scene depth map which can be directly converted into transmission map.

II. Related Work

In order to improve visibility of underwater images, several methods have been presented. Traditional image enhancement methods, such as Histogram Equalization (HE) and Contrast Limited Adaptive Histogram Equalization (CLAHE), are widely utilized in improving the appearance of degraded underwater images. A mixture Contrast Limited Adaptive Histogram Equalization was proposed by Hitam et al. [3]. They operated CLAHE on RGB and HSV color models, and combined together the two results with Euclidean norm. Ahmad et al. [4] presented a method called dual-image Rayleigh-stretched contrast-limited adaptive histogram specification, which enhanced underwater images by integrating global and local contrast correction.

Underwater image enhancement methods based on atmospheric scattering model are also widely used. As a successful application of atmospheric scattering model, dark channel prior [5] is popular in underwater image enhancement because there are similarity between hazy images and underwater images. Li et al. [6] achieved underwater image enhancement using dark channel prior and luminance adjustment. Li et al. [7] restored underwater image using atmospheric scattering model by minimum information loss and histogram distribution prior.

Fusion-based underwater image enhancement is firstly presented by Ancuti and Ancuti [8] and can achieve better results that traditional single procedures in solving main problems such as noises, low contrast, blurring and color casts. They operated a multi-scale fusion strategy on underwater images

*Xianping Fu is the corresponding author (Email: fxp@dlmu.edu.cn).
This work was supported in part by the National Natural Science Foundation of China Grant 61370142 and Grant 61272368, by the Fundamental Research Funds for the Central Universities Grant 3132016352, by the Fundamental Research of Ministry of Transport of P. R. China Grant 2015329225300.

978-1-5090-5278-3/17/$31.00 ©2017 IEEE

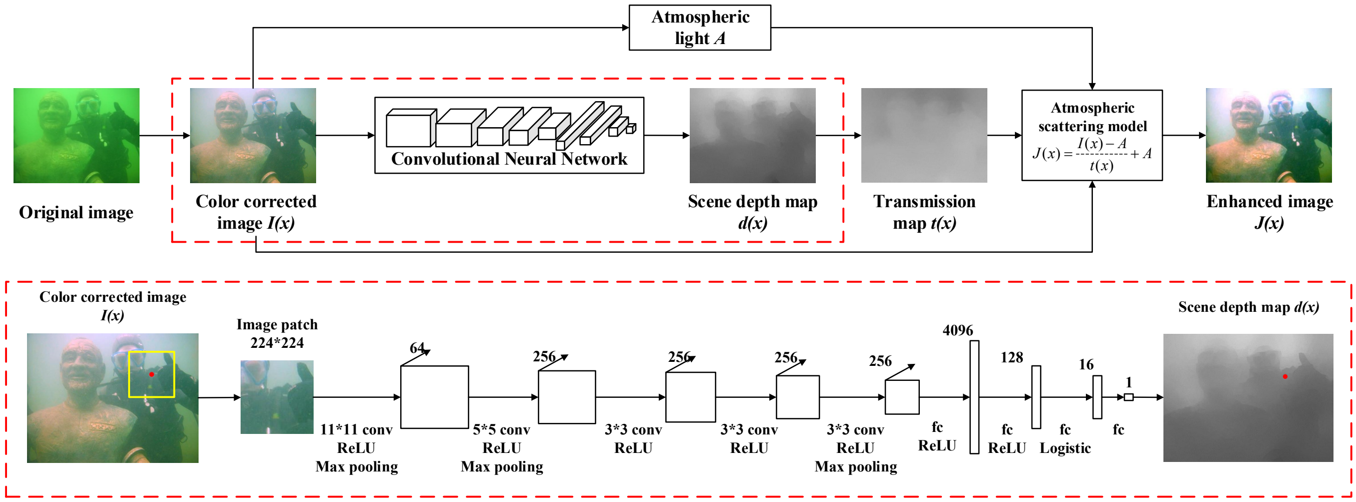


Fig. 1. The overall framework of proposed method.

using Gaussian pyramid. Wang et al. [9] proposed a fusion-based underwater image enhancement method using wavelet decomposition.

In addition, several depth prediction approaches are directly related to our work. Eigen et al. [10] predicted depth from a single image used two deep network stacks: global coarse-scale network and local fine-scale network. Liu et al. [11] employed deep convolutional neural fields to estimate depth from a single image.

III. The Proposed Enhancing Approach

Aiming to remove color casts and blurriness of underwater image, the proposed underwater enhancing approach involves two main steps: color correction (Section III-A) and model-based underwater image dehazing with scene depth (Section III-B).

In the proposed framework (shown as Fig.1), color correction algorithm is firstly applied to original underwater image to remove color casts and obtain color corrected image I. Then, color corrected image I is used to estimate atmospheric light A and transmission map t, which can be employed in atmospheric scattering model. Aiming to obtain the transmission map t, a convolutional neural network is carried out firstly to predict a scene depth map d from color corrected image I. Considering the exponential relationship between transmission and scene depth, scene depth map d can be naturally converted into transmission map t. Finally, combining color corrected image I, atmospheric light A and transmission map t, the final enhanced image J can be acquired by atmospheric scattering model.

A. Color Correction

In the proposed framework, color correction is firstly applied to original underwater image in order to discard color casts and produce a natural appearance. The proposed white balance algorithm is based on the [9] [12] with gain factor and can be described as:

$$I_{out} = \frac{I}{\lambda_1 \frac{\mu}{\mu_{ref}} + \lambda_2} \tag{1}$$

where I_{out} and I denote the color corrected image and the original underwater image. $\mu = \{\mu_R, \mu_G, \mu_B\}$ represents the sum of each channel of underwater image I, and $\mu_{ref} = ((\mu_R)^2 + (\mu_G)^2 + (\mu_B)^2)^{\frac{1}{2}}$. $\lambda_1 = \{\lambda_R, \lambda_G, \lambda_B\}$ is estimated by the maximum value of each channel of underwater image I. The value of λ_2 varies in the range of $[0, 0.5]$. The color corrected image will be brighter when λ_2 comes closer to 0. In other words, the higher λ_2 is, the lower the brightness is.

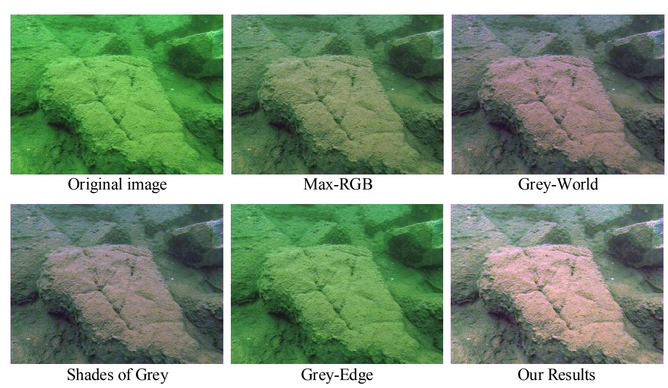


Fig. 2. Results of white balance methods.

This white balance method removes the color casts efficiently and produce a natural color corrected image (see Fig.2). Nevertheless, low contrast and blurriness still exist in the color corrected image. To obtain a better enhanced image, model-based underwater image dehazing with scene depth is enforced.

B. Model-based Underwater Image Dehazing with Scene Depth

After color correction, image dehazing is implemented to remove the blurriness because of the similarity between underwater image and hazy image. In Section III-B1, atmospheric scattering model described the formation of a hazy image is employed in underwater environments. In this model, there are two main factors: global atmospheric light A and transmission map t. Since exponential relationship between transmission and scene depth exists, transmission map estimation can be naturally formulated into a scene depth map estimation problem. In Section III-B2, a convolutional neural network is used to predict scene depth map.

1) Atmospheric Scattering Model: After color correction, atmospheric scattering model is used to remove the blurriness of color corrected image. The atmospheric scattering model implemented in underwater environments is as follows [5]:

$$J(x) = \frac{I(x) - A}{t(x)} + A \tag{2}$$

where J and I represent enhanced image and color corrected image. A is the global atmospheric light, and t is the transmission map. In this model, blurriness removal can be achieved by the color corrected image I, a transmission map t and a vector of global atmospheric light A. The global atmospheric light A can be evaluated by the mean values of each channel of the color corrected image. In addition, transmission map estimation is transformed into a scene depth map estimation problem. The expression that described the exponential relationship between the transmission map and the scene depth map is shown as:

$$t(x) = e^{-\beta d(x)} \tag{3}$$

where β is the scattering coefficient of the atmosphere and d represents the scene depth map. Since the value of β is difficult to calculate, maximum visibility d_{max} is introduced to obtain the transmission map. The transmission map will reach its minimum t_{min} with the maximum visibility by $t_{min} = e^{-\beta d_{max}}$. Then transmission map can be computed by:

$$t(x) = (t_{min})^{\frac{d(x)}{d_{max}}} \tag{4}$$

In this case, the transmission map t can be directly estimated by the scene depth map d, different from previous methods that use dark channel prior [6] [13]. Obviously, in underwater image dehazing, the main task is to predict the scene depth map.

2) Scene Depth Estimation with a Convolutional Neural Network: To obtain the scene depth map, a Convolutional Neural Network based on [11] [14] is employed. An outdoor scene depth dataset (Make3D dataset [15]) is used to train the proposed network. Make3D dataset involves 534 images in which the maximum depth is 81m with faraway objects are all mapped to the one distance of 81 meters. Since the color corrected image is similar to an outdoor hazy scene, it can be processed by the proposed network trained by Make3D dataset.

The convolutional neural network is shown in Fig.1. First of all, Input image is segmented into patches. These image patches are constrained in size $224 \times 224 \times 3$ to feed the proposed framework. The proposed framework contains 5 feature extraction layers and 4 fully-connected layers. The first feature extraction layer that involves convolution, ReLU (Rectified Linear Units) and max pooling is represented as:

$$F_1(x) = \max_{y \in \Omega(x)} f_1(y), f_1 = \max(0, W_1 * I + B_1) \tag{5}$$

where W_1 and B_1 represent the filters and biases of the first layer respectively, and $*$ denotes the convolutional operation. I is the input image patch of size $224 \times 224 \times 3$, and $\Omega(x)$ is a 2×2 neighborhood centered around pixel x. The second feature extraction layer is similar to the first one and can be written as:

$$F_2(x) = \max_{y \in \Omega(x)} f_2(y), f_2 = \max(0, W_2 * F_1 + B_2) \tag{6}$$

where W_2 and B_2 represent the filters and biases of the second layer. The next two layers both contain convolution and ReLU activation function:

$$\begin{aligned} F_3 &= \max(0, W_3 * F_2 + B_3) \\ F_4 &= \max(0, W_4 * F_3 + B_4) \end{aligned} \tag{7}$$

where W_3 , W_4, B_3 and B_4 represent the filters and biases of the third layer and fourth layer. The fifth layers is shown as:

$$F_5(x) = \max_{y \in \Omega(x)} f_5(y), f_5 = \max(0, W_5 * F_4 + B_5) \tag{8}$$

where W_5 and B_5 denote the filters and biases of the fifth layer. After these feature extraction layers, four fully-connected layers are carried out with output numbers 4096, 128, 16 and 1. In the first two fully-connected layers, ReLU ($\max(0, x)$)is used as activation function. In third fully-connected layer, logistic function($f(x) = (1 + e^{-x})^{-1}$) is used as activation function. There is no activation function in the last fully-connected layer. After these layers, the proposed network outputs an one-dimensional depth of an input image patch. Finally an estimated scene depth map can be obtained combining the predicted depths of image patches from a color corrected image.

IV. Results and Discussion

To verify the effectiveness of the proposed method, several experiments involving qualitative comparison, quantitative comparison and application test are conducted. Fig.3 shows the results of the proposed strategy and the comparison with other enhancement methods. In Fig.3, the first four columns contain original images, scene depth maps, transmission maps and our enhanced images. The last two columns show the enhanced images obtained by CLAHE [3] and Dark Channel Prior [5]. It can be observed that the proposed strategy is able to acquire more pleasing enhanced images with corrected color, enhanced contrast and sharpened details.

Moreover, several evaluation metrics are used to assess the performance of the proposed method. Following other researchers, MSE (Mean Square Error), PSNR (Peak Signal Noise Ratio), SSIM (Structural Similarity Index) and Patch-based Contrast Quality Index(PCQI) [16] are carried out. Generally, MSE and PSNR are utilized in assessing image noise. The lower MSE and higher PSNR values represent less noise. SSIM measures the visual impact of three characteristics of an image: luminance, contrast and structure. A higher SSIM

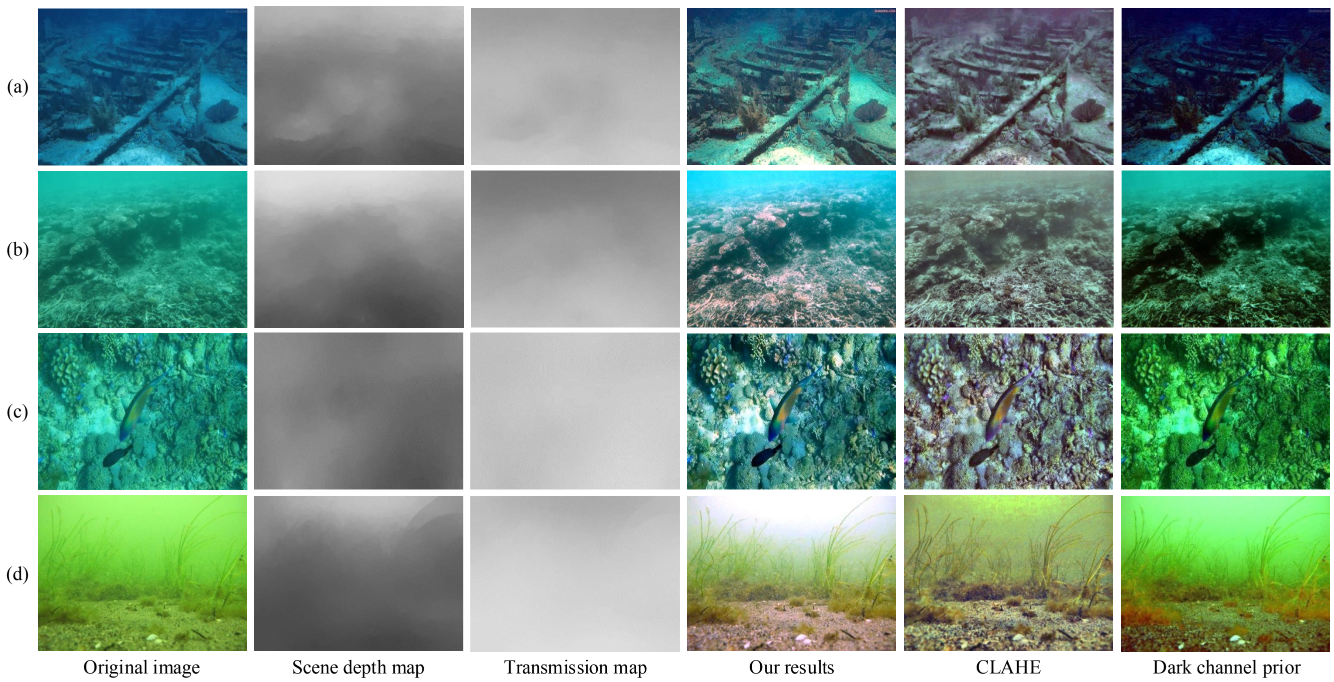


Fig. 3. The experimental results and comparison of underwater image enhancement methods.

TABLE I. THE MSE, PSNR, SSIM AND PCQI OF ENHANCED IMAGES

Image	Method	MSE	PSNR	SSIM	PCQI [16]
Fig.3(a)	CLAHE [3]	2462.90	14.22	0.31	**1.15**
	Dark Channel Prior [5]	1970.55	15.18	0.56	1.00
	Our method	**1650.81**	**15.95**	**0.69**	1.02
Fig.3(b)	CLAHE [3]	2587.11	14.00	0.33	1.07
	Dark Channel Prior [5]	**2062.72**	**14.99**	0.61	0.93
	Our method	2105.38	14.90	**0.72**	**1.22**
Fig.3(c)	CLAHE [3]	3027.17	13.32	0.37	**1.27**
	Dark Channel Prior [5]	1843.79	15.47	0.69	1.11
	Our method	**1160.95**	**17.48**	**0.81**	1.09
Fig.3(d)	CLAHE [3]	1420.57	16.61	0.66	**1.02**
	Dark Channel Prior [5]	**681.04**	**19.80**	**0.88**	0.74
	Our method	4355.68	11.74	0.46	**1.02**

indicates a better result of enhancement. Wang et al. [16] proposed PCQI, and the higher PCQI means that the image has better contrast. TABLE I shows the MSE, PSNR, SSIM and PCQI values of enhanced images in Fig.3.

SIFT [17] feature detection and matching is highly configurable and applicable to object detection and recognition. In Fig.4, SIFT operator is applied to two pairs of original degraded underwater images and as well as the corresponding enhanced images. The comparison in Fig.4 shows that the detected and matched feature points of the enhanced images are greatly increased. It can be observed that the original degraded underwater images show great limitations in underwater objects detection and recognition. By contrast with original underwater images, enhanced images help to reveal more feature points. Therefore, the proposed underwater image enhancement method is helpful for underwater objects detection and recognition.

V. CONCLUSION

In this paper, an efficient underwater image enhancement strategy combining adaptive color correction and model-based dehazing is presented. In model-based dehazing, a convolutional neural network is used to predict scene depth map which can be transformed into transmission map. The experimental results show that the proposed method can enhance underwater image effectively and performs well in underwater objects detection and recognition. In future work, we will consider to improve the convolutional neural network to achieve better enhancement, robust adaptation and real-time performance.

REFERENCES

[1] R. Wang, Y. Wang, J. Zhang and X. Fu, Review on underwater image restoration and enhancement algorithms. In *Proc. 7th International Conference on Internet Multimedia Computing and Service (ICIMCS)*, ACM, pages 56-62, 2015.

[2] R. Schettini, S. Corchs, Underwater image processing: state of the art of restoration and image enhancement methods. *EURASIP Journal on Advances in Signal Processing*, 1: pages 1-14, 2010.

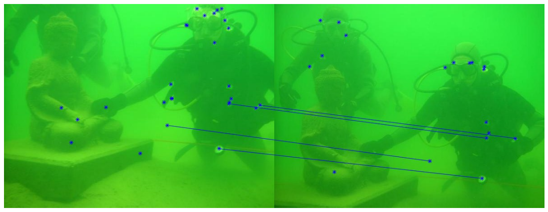
[3] M. S. Hitam, E. A. Awalludin, W. N. J. H. W. Yussof and Z. Bachok, Mixture contrast limited adaptive histogram equalization for underwater image enhancement. In *Proc. International Conference on Computer Applications Technology (ICCAT)*, pages 1-5, 2013.

[4] A. S. A. Ghani and N. A. M. Isa, Enhancement of low quality underwater image through integrated global and local contrast correction. *Applied Soft Computing*, 37: pages 332-344, 2015.

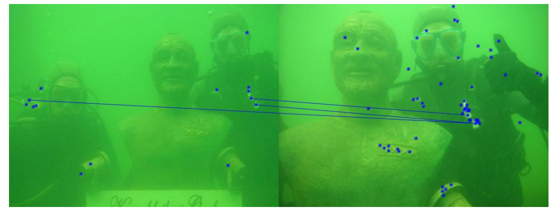
[5] K. He, J. Sun and X. Tang, Single image haze removal using dark channel prior. *IEEE transactions on pattern analysis and machine intelligence*, 33(12): pages 2341-2353, 2011.

[6] X. Li, Z. Yang, M. Shang. and J. Hao, Underwater image enhancement via dark channel prior and luminance adjustment. In *OCEANS 2016-Shanghai*, pages 1-5, 2016.

Original image

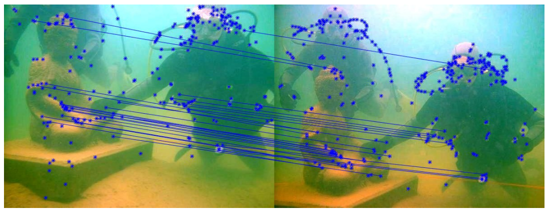


Detected feature points: 42 (left) and 25 (right)
Matched feature points: 6

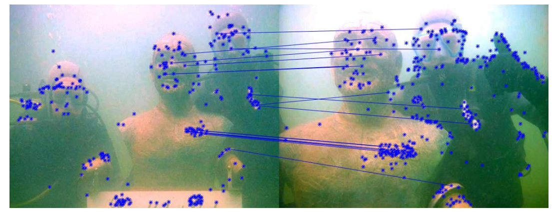


Detected feature points: 20 (left) and 72 (right)
Matched feature points: 4

Ehanced image



Detected feature points: 297 (left) and 396 (right)
Matched feature points: 31



Detected feature points: 354 (left) and 445 (right)
Matched feature points: 19

Fig. 4. Local feature points detection and matching. The first row is the original images, and the second row is the enhanced images processed by our method.

[7] C. Li, J. Guo, R. Cong, Y. Pang and B. Wang, Underwater Image Enhancement by Dehazing With Minimum Information Loss and Histogram Distribution Prior. *IEEE Transactions on Image Processing*, 25(12): pages 5664-5677, 2016.

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

[9] Y. Wang, X. Ding, R, Wang, J. Zhang and X. Fu, Fusion-based underwater image enhancement by wavelet decomposition. In *Proc. IEEE International Conference on Industrial Technology (ICIT)*, pages 1013-1018, 2017.

[10] D. Eigen and R. Fergus, Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture. In *Proc. IEEE International Conference on Computer Vision (ICCV)*, pages 2650-2658, 2015.

[11] F. Liu, C. Shen and G. Lin, Deep convolutional neural fields for depth estimation from a single image. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pages 5162-5170.

[12] G. D. Finlayson and E. Trezzi, Shades of gray and colour constancy. In *Color and Imaging Conference*, 2004(1): pages 37-41, 2004.

[13] H. Zheng, X. Sun, B. Zheng, R. Nian and Y. Wang, Underwater image segmentation via dark channel prior and multiscale hierarchical decomposition. In *OCEANS 2015-Genova*, pages 1-4, 2015.

[14] A. Krizhevsky, I. Sutskever and G. Hinton, Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems (NIPS)*, pages 1097-1105, 2012.

[15] A. Saxena, M. Sun, and A. Y. Ng, Make3d: Learning 3d scene structure from a single still image. *IEEE transactions on pattern analysis and machine intelligence*, 31(5): pages 824-840, 2009.

[16] S. Wang, K. Ma, H. Yeganeh, Z. Wang and W. Lin, A patch-structure representation method for quality assessment of contrast changed images. *IEEE Signal Processing Letters*, 22(12): pages 2387-2390, 2015.

[17] D. Lowe, Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2): pages 91-110, 2004.