Underwater color restoration and dehazing based on deep neural network

Nan Wang
Huazhong Institute of Electro-Optics, Wuhan National Laboratory for Optoelectronics, Wuhan, China
nanwangmail@163.com

Abstract. In real-world underwater environment, exploration of seabed resources, underwater archaeology, and underwater fishing rely on a variety of sensors, vision sensor is the most important one due to its high information content, non-intrusive, and passive nature. However, wavelength-dependent light attenuation and back-scattering result in color distortion and haze effect, which degrade the visibility of images. To address this problem, firstly, we proposed an underwater unsupervised generative adversarial network call UWGAN for generating realistic underwater fake images from paired images in air and their corresponding depth maps based on an improved underwater imaging model. Secondly, U-Net, which is trained efficiently using synthetic underwater dataset, is adopted for color restoration and dehazing. Our model directly reconstructs underwater clear images using end-to-end autoencoder networks, while maintaining the structural similarity of underwater scenes. The results obtained by our method were compared with existing methods qualitatively and quantitatively. Experimental results obtained by the proposed model demonstrate well performance on open real-world underwater datasets, and the processing speed can reach up to 125FPS running on one NVIDIA 1060 GPU.

1. Introduction
In recent years, underwater vision plays an important role in a lot of different applications. Therefore, underwater image processing has received extensive attention and research due to the poor underwater imaging environment and image quality. The main reason is the scattering and attenuation of light, the scattering results in haze effect, and the attenuation of light leads to color cast.

So far many image enhancement algorithms have been proposed, such as white balance algorithm[1], gray world algorithm[2], histogram equalization[3] and fusion algorithm[4], however, these methods are not based on the underwater physical imaging model, so it is challenging and ineffective to apply these algorithms to different underwater scenes directly.

Many underwater image enhancement algorithms based on imaging models have been proposed. For instance, He et al [5] proposed a dark channel prior (Dark channel prior, DCP) dehazing algorithm based on many experiments. Chiang et al[6] apply DCP model on underwater image dehazing problem. These traditional methods are not intelligent, it is very time-consuming to calculate the characteristics of the image.

In these years, deep learning networks are increasingly used in the field of computer vision, especially CNNs, which are widely used in image classification[7], object detection[8], and motion recognition [9], the performance is much better than traditional methods. However, the recent research on underwater image enhancement (UIE) using CNNs are limited due to lack of underwater datasets.
It is difficult to obtain images without water in real-world underwater scenes. Therefore, using synthetic underwater datasets is an important approach[10][11][12]. Some model based on generative adversarial networks[13] are used to generating realistic underwater fake images. For instance, CycleGAN[14] generates images through style transfer. WaterGAN[15] takes in-air images, depth images and noise vectors as input, followed by a camera model, then output synthetic images. Based on our experimental results, the image generated by WaterGAN suffers color noise and they differ a lot from real world underwater images. Therefore, to generate realistic underwater fake images with both color cast and haze effect, we improved the underwater imaging model, and proposed an underwater unsupervised GAN based on this model to generate realistic underwater fake images from clear in-air images. Then, U-Net[16] with joint loss function is trained to enhance underwater images through synthetic datasets. Finally, the performance of the proposed algorithm is validated on real underwater images as well as underwater target detection datasets for both low-level and high-level computer vision tasks. The experimental results show that our proposed method can enhance underwater images as well as maintaining structural similarities. Experimental results obtained by the proposed model demonstrate well performance on open real-world underwater datasets, and the processing speed can reach high real-time on one NVIDIA 1060 GPU.

2. Our proposed method
To synthesis the realistic underwater fake images (color casts, low contrast and haze effect), we improved underwater imaging model, and proposed an underwater generative adversarial network (UWGAN), which takes indoor RGB-D images and several sample sets of real underwater images of a specific survey site as input to train UWGAN. Then, generated underwater fake images, which are used to train a restoration network based on U-Net that can enhance underwater images in real-time.

2.1. Improved underwater imaging model
As is well known, a simplified underwater imaging model is shown in Eq.(1).

\[
I(x) = J(x)T(x) + A(1 − T(x))
\]

\[
T(x) = e^{−\beta(\lambda)x}(1)
\]

where, I(x) is the light intensity of each pixel x. J(x) is the initial irradiance that not propagating through the water. T(x) is the transmission map of the scene. A is the atmospheric ambient light of the scene. \(\beta\) is attenuation coefficient of light of different wavelengths \(\lambda\), and \(d(x)\) is the range between the scene and the camera.

**Figure 1.** Synthetic underwater-style images through Eq.(2). (a) are in-air sample images, (b)-(d) are synthetic underwater-style sample images of different water types.

![Synthetic underwater-style images](image)

**Figure 2.** UnderwaterGAN architecture. UWGAN takes color image and its depth map as input, then it synthesizes underwater realistic images based on underwater optical imaging model by learning parameters through generative adversarial training.

We can generate underwater fake images using paired images in air and their corresponding depth maps by Eq.(1), which can well simulate color cast caused by light attenuation in water. However, it is
difficult to synthetic haze in images caused by light scattering in water. As shown in Figure 4, obvious haze effect can be observed on real underwater images. Inspired by related dehazing methods, we improved the second term in Eq.1. The improved imaging model is shown in Eq.(2).

\[ I(x) = J(x)T(x) + AT(x)(1 - T'(x)) \]

\[ T'(x) = e^{-\alpha d(x)} \]  

(2)

where, \(AT(x)\) is ambient light caused by the light attenuation of different wavelength. \(\alpha\) is the scene scattering coefficient, which corresponds to the scattering coefficient in the atmospheric imaging model, and \(\alpha\) is set by default to 1, corresponding to a moderate and homogeneous haze effect. Three types of realistic underwater images were synthesized with color cast and haze effect are shown in Figure 1. Underwater-style images are generated based on Eq.(2), whose parameters are estimated through adversarial learning using GAN, as shown in Figure 2.

2.2. Underwater image restoration based on U-Net

Figure 3. Proposed U-net Architecture for underwater image restoration and enhancement. U-Net is used for color restoration and haze removal of underwater images. A detailed description of U-Net architecture proposed in the paper is shown in Figure 3. Firstly, a degraded underwater RGB image is resized to 256x256 and then fed into the encoder part of U-net. In the encoder part, the image is finally downsampled into a 32x32x256-dimensional latent vector through a series of convolution and max-pooling operations. In each downsampling stage, 3x3 convolution with a stride of 1 followed by a rectified linear unit (ReLU) activation function are conducted twice, then a 2x2 max pooling with a stride of 2 is used. The number of feature maps are doubled after each convolution operation. In the decoder part, upsampling is done from the latent high dimensional vector back to the original input image size sequentially. Output tensor is concatenated to the corresponding symmetric layer in the encoder part, then followed by two consecutive convolution layers and a rectified linear activation layer. The number of feature maps is gradually reduced to three channels.

2.3. Dataset

The in-air datasets we used are images of indoor scenes that has been labeled in the NYU Depth dataset V1[17] and V2[18], which contain a total of 3733 image and depth map pairs. The underwater dataset contains real-world underwater images collected from marine organisms’ farms (including scallops, sea cucumbers, sea urchins, etc.), which can be roughly divided into two categories, one contains green hued images of near-field scenes (RealA), and the other contains blue-green hued images of far-field scenes (RealB). We also use underwater open datasets[19] (RealC) as testing sets, where RealA contains 2069 underwater images, RealB contains 2173 underwater images, and RealC contains 890 underwater images. Several typical images of the datasets are shown in Figure 4.
2.4. Experimental setup

The training settings of our proposed method are presented in details in this section. Our models are trained in the computer with the following configurations: Intel i7 HQ 8700 processor, 16GB RAM, NVIDIA 1080Ti 12GB graphics card.

Firstly, UWGAN is trained to synthesize underwater-style images using the NYU-Depth Dataset, RealA and RealB datasets. Our model was trained for 30 epochs, using Adam optimizer with a learning rate of 0.0001, and the momentum term was set to 0.5. The batch size was set to 64 with output images set to 256x256. Secondly, U-net is trained as an image enhancement network using synthetic pairs. The batch size was set to 32 and the output image size is 256x256. The learning rate is set to 0.0001 according to Adam optimizer, our model is trained for 200 epochs.

3. Result and discussion

In this section, we quantitatively and qualitatively compare our proposed method with several representative underwater image enhancement algorithms, including Unsupervised Color Correction Method (UCM)[20], Histogram equalization (HE)[21], Multi-Scale Retinex with Color Restoration (MSRCR)[22], Fusion[4], Underwater Dark Channel Prior (UDCP)[23], Image Blurriness and Light Absorption (IBLA)[24], Underwater Color Correction using GAN (UGAN)[25], WaterGAN-color-correction (WaterGAN)[15].

We employ a non-reference metric, UIQM[26], for the quantitative assessment of underwater image quality on RealA, RealB, and RealC datasets as no ground truth scenes are available as the reference for real-world underwater images. Besides, we employ three full-reference metrics, namely MSE, PSNR[27], SSIM, for assessment image quality on synthetic datasets. To reasonably assess the time spent on various algorithms, we resize all images to 256x256, which provides a stable output for enhancements in later experiments.

Firstly, we compare the capabilities of different methods to improve real underwater images visibility on the RealA, RealB, and RealC datasets. The qualitative comparison is shown in Figure 5. Most methods can improve the visual quality of images of a slight haze effect. UCM, HE, and Fusion can improve underwater images brightness and contrast, but are less uniform for color restoration and seem to be over-enhanced in some areas of the image. The results of MSRCR appear to have a suitable hue but lack sufficient saturation and contrast. UDCP and IBLA do not recover well for green-toned images, they make the image darker but improve underwater images contrast. UGAN, WaterGAN can improve underwater images contrast, but they don’t recover color well and generate some artifacts, which destroy the structural information of the image. The proposed method recovers degraded underwater images color while keeping a proper brightness and contrast.

Table 1 quantitatively show the average quantized scores of UIQM evaluated on RealA, RealB, and RealC datasets. Our model achieves the best scores based on UIQM.
Table 1. Average quantitative UIQM values on underwater real image datasets RealA, RealB and RealC.

| Datasets | Assessments | Input | UCM | HE | MSRCR Fusion | UDCP | IBLA | UGAN | WaterGAN | Ours |
|----------|-------------|-------|-----|----|--------------|------|------|-------|----------|-----|
| RealA    | UIQM        | 4.22  | 4.574 | 4.508 | 3.967 | 4.622 | 4.721 | 4.533 | 4.832 | 4.280 | **4.936** |
| RealB    | UIQM        | 3.920 | 4.695 | 4.864 | 4.387 | 4.675 | 4.598 | 4.590 | 4.927 | 4.309 | **5.080** |
| RealC    |             | 4.561 | 4.803 | 4.816 | 3.932 | 4.891 | 5.251 | 4.942 | 4.996 | 4.979 | **5.636** |

UIQM is a non-reference assessment metric whose quantitative results depend largely on the value of scale factors. Structural information of images is not considered in these kinds of non-reference evaluation metrics. Although some enhanced images can get higher score, the visual quality is poor, the reason is that the metric is calculated from the pixels. Therefore, we also employ three full-reference assessment metrics MSE, PSNR, and SSIM to evaluate the performance of different methods on synthetic datasets without training. The comparison results in Table 2 demonstrate that our proposed method achieves the best results in terms of MSE, PSNR, and SSIM.

Table 2. Quantitative results evaluation on synthetic dataset by full-reference metrics: MSE, PSNR, SSIM values. The smaller the MSE values, the greater the PSNR and SSIM values, the better the enhanced results, with blue representing the best results.

| Datasets | Assessments | Input | UCM | HE | MSRCR Fusion | UDCP | IBLA | UGAN | WaterGAN | Ours |
|----------|-------------|-------|-----|----|--------------|------|------|-------|----------|-----|
| Synthesis| MSE         | 0.042 | 0.029 | 0.045 | 0.059 | 0.027 | 0.072 | 0.058 | 0.026 | 0.014 | **0.002** |
|          | PSNR        | 20.68 | 23.46 | 18.315 | 13.25 | 23.13 | 17.37 | 19.10 | 20.63 | 20.25 | **30.31** |
|          | SSIM        | 0.869 | 0.944 | 0.845 | 0.580 | 0.933 | 0.847 | 0.832 | 0.779 | 0.842 | **0.966** |

The average inference time of different algorithms are compared in one computer with following configuration: Intel i7-8750H CPU, 16GB RAM, and GTX1060 6G GPU. The results are shown in Table 3. Our model has the fastest processing speed compared to other methods. Moreover, the model we proposed has the fewest Params and FLOPs compared to other deep-learning-based methods. UGAN employs many convolution layers with 512 kernels, which causes that there are too many network parameters. WaterGAN employs multiple networks, resulting in slow processing speed.

Table 3. Testing time and parameters of generator of different enhancement methods.

| UCM | HE  | MSRCR Fusion | UDCP | IBLA | UGAN | WaterGAN | Ours |
|-----|-----|--------------|------|------|------|----------|-----|
| 1.284 | 0.009 | 0.076 | 0.118 | 2.051 | 4.561 | 0.022 | 10.347 | **0.080** |

4. Conclusion

Based on an improved underwater imaging model, an underwater unsupervised generative adversarial network (UWGAN) for synthesizing realistic underwater fake images is proposed in this paper. Then, U-net with joint loss functions is used for degraded underwater images enhancement. Our model is validated on open real-world underwater datasets, which demonstrate its effectiveness and time efficiency. In addition, our model has higher real-time performance than other deep learning models and has the potential to be deployed on edge devices.

References

[1] Liu Y C, Chan W H, Chen Y Q. Automatic white balance for digital still camera[J]. IEEE Transactions on Consumer Electronics, 1995, 41(3): 460-466.
[2] Rizzi A, Gatta C, Marini D. Color correction between gray world and white patch[C]//Human Vision and Electronic Imaging VII. International Society for Optics and Photonics, 2002, 4662: 367-375.
[3] Pizer S M, Amburn E P, Austin J D, et al. Adaptive histogram equalization and its variations[J]. Computer vision, graphics, and image processing, 1987, 39(3): 355-368.
[4] Ancuti C, Ancuti Č O, Haber T, et al. Enhancing underwater images and videos by fusion[C]//2012 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2012:
[5] He K, Sun J, Tang X. Single image haze removal using dark channel prior[J]. IEEE transactions on pattern analysis and machine intelligence, 2010, 33(12): 2341-2353.

[6] Chiang J Y, Chen Y C. Underwater image enhancement by wavelength compensation and dehazing[J]. IEEE Transactions on Image Processing, 2011, 21(4): 1756-1769.

[7] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[C]//Advances in neural information processing systems. 2012: 1097-1105.

[8] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 779-788.

[9] Kuehne H, Jhuang H, Garrote E, et al. HMDB: a large video database for human motion recognition[C]//2011 International Conference on Computer Vision. IEEE, 2011: 2556-2563.

[10] Anwar S, Li C, Porikli F. Deep underwater image enhancement[J]. arXiv preprint arXiv:1807.03528, 2018.

[11] Ancuti C, Ancuti C O, De Vleeschouwer C. D-hazy: A dataset to evaluate quantitatively dehazing algorithms[C]//2016 IEEE International Conference on Image Processing (ICIP). IEEE, 2016: 2226-2230.

[12] Uplavikar P, Wu Z, Wang Z. All-In-One Underwater Image Enhancement using Domain-Adversarial Learning[J]. arXiv preprint arXiv:1905.13342, 2019.

[13] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[C]//Advances in neural information processing systems. 2014: 2672-2680.

[14] Zhu J Y, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]//Proceedings of the IEEE international conference on computer vision. 2017: 2223-2232.

[15] Li J, Skinner K A, Eustice R M, et al. WaterGAN: Unsupervised generative network to enable real-time color correction of monocular underwater images[J]. IEEE Robotics and Automation letters, 2017, 3(1): 387-394.

[16] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015: 234-241.

[17] Silberman N, Fergus R. Indoor scene segmentation using a structured light sensor[C]//2011 IEEE international conference on computer vision workshops (ICCV workshops). IEEE, 2011: 601-608.

[18] Silberman N, Hoiem D, Kohli P, et al. Indoor segmentation and support inference from rgbd images[C]//European Conference on Computer Vision. Springer, Berlin, Heidelberg, 2012: 746-760.

[19] Li C, Guo C, Ren W, et al. An underwater image enhancement benchmark dataset and beyond[J]. arXiv preprint arXiv:1901.05495, 2019.

[20] Iqbal K, Odetayo M, James A, et al. Enhancing the low quality images using unsupervised colour correction method[C]//2010 IEEE International Conference on Systems, Man and Cybernetics. IEEE, 2010: 1703-1709.

[21] Hummel R. Image enhancement by histogram transformation[J]. Computer Graphics and Image Processing, 1977, 6(2):184-195.

[22] Rahman Z, Jobson D J, Woodell G A. Multi-scale etinex for color image enhancement[C]//Proceedings of 3rd IEEE International Conference on Image Processing. IEEE, 1996, 3: 1003-1006.

[23] Drews P, Nascimento E, Moraes F, et al. Transmission estimation in underwater single images[C]//Proceedings of the IEEE international conference on computer vision workshops. 2013: 825-830.

[24] Peng Y T, Cosman P C. Underwater image restoration based on image blurriness and light absorption[J]. IEEE transactions on image processing, 2017, 26(4): 1579-1594.
[25] Fabbri C, Islam M J, Sattar J. Enhancing underwater imagery using generative adversarial networks[C]/2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2018: 7159-7165.

[26] Panetta K, Gao C, Agaian S. Human-visual-system-inspired underwater image quality measures[J]. IEEE Journal of Oceanic Engineering, 2015, 41(3): 541-551.

[27] Hore A, Ziou D. Image quality metrics: PSNR vs. SSIM[C]/2010 20th International Conference on Pattern Recognition. IEEE, 2010: 2366-2369.