Underwater Image Restoration using Deep Networks to Estimate Background Light and Scene Depth

Keming Cao, Yan-Tsung Peng and Pamela C. Cosman
Department of Electrical and Computer Engineering, UC San Diego, La Jolla, CA 92093-0407
Email: {k5cao, yapeng, pcosman}@ucsd.edu

***Abstract*—Images taken underwater often suffer color distortion and low contrast because of light scattering and absorption. An underwater image can be modeled as a blend of a clear image and a background light, with the relative amounts of each determined by the depth from the camera. In this paper, we propose two neural network structures to estimate background light and scene depth, to restore underwater images. Experimental results on synthetic and real underwater images demonstrate the effectiveness of the proposed method.**

***Keywords*-Underwater images, image restoration, depth estimation, convolutional neural networks**

I. Introduction

Images in fog, haze, sandstorms, or water can suffer from visibility degradation as light is scattered and absorbed with distance from the camera. In the image formation model (IFM) [1], the observed intensity $I^c(x)$ at pixel x consists of the scene radiance $J^c(x)$ blended with the background light (BL) B^c according to the transmission maps $t^c(x)$, where c is one of the red, green, and blue channels:

$$I^c(x) = J^c(x)t^c(x) + B^c\big(1 - t^c(x)\big), c \in \{r, g, b\} \quad (1)$$

where I^c, J^c, t^c and $B^c \in [0, 1]$. The transmission map (TM) describes the portion of the scene radiance which is not scattered or absorbed and which reaches the camera. Large values in the TM mean that the corresponding scene point is closer to the camera.

Using the IFM, He et al. [2] proposed the dark channel prior (DCP) to remove fog/haze in natural terrestrial images via estimation of the ambient light and TM. This motivated many underwater image restoration approaches [3]–[5]. However, estimating BL and TM for underwater images based on the DCP frequently fails since red light is more attenuated than other wavelengths underwater. In [6], image blurriness was adopted for estimation of BL and TM, and [7] considers both blurriness and light absorption. Some researchers exploited learning algorithms to generate a mapping function from a hazy input image to an output depth map [8]–[11]. Training using synthetic hazy images with bright ambient light may not suit underwater images. In addition, these methods adopted DCP-based BL estimation, which often fails for underwater images.

In this paper, deep networks are adopted to estimate BL and scene depth for underwater images, where a 5-layer ConvNet and a multi-scale deep network architecture similar to [12] are used to predict BL and scene depth. Using these, we are able to restore underwater images by reversing the image formation process based on the IFM. Experimental results on synthesized and real underwater images demonstrate the proposed method outperforms state-of-the-art underwater image restoration methods.

II. Proposed Method

Typical restoration based on the IFM involves estimating t^c and B^c, and then determining J^c based on approximately inverting the IFM:

$$J^c(x) = \frac{I^c(x) - B^c}{\max\big(t^c(x), t_0\big)} + B^c, \quad (2)$$

where t_0 is set to 0.1 to increase the exposure of J^c for display. We use a 5 layer ConvNet for BL estimation, and a multi-scale deep network architecture [12] to estimate underwater scene depth. Then, the TM can be computed using underwater attenuation coefficients and scene depth. With BL and the TM, we can restore images using Eq. (2).

BL Estimation: To estimate BL, we implement a 5-layer convolutional neural network, shown in Fig. 1. The first three layers are convolutional with filter sizes of 5×5, 5×5, and 3×3, each of which has a 2×2 pooling layer and a normalization layer. The last two layers are fully connected. Before the final output B^c, we threshold to the range $[0, 1]$.

TM Estimation: To predict scene depth, and hence the TM, we adopt a multi-scale architecture based on [12], which stacks two deep networks, a coarse global network and a refined network. The method [12] originally was proposed to predict depth from a single indoor image, and was used to estimate depth for hazy images in [10]. Here, we add an upsampling layer. The network architecture is shown in Fig. 2. The coarse global network first has five convolutional layers. The first two have a 2×2 pooling layer and a normalization layer, and the last has only a 2×2 pooling layer. The last two layers of the coarse network are fully connected layers, and the output coarse depth is concatenated to the output of the first layer of the refined network. The refined network has 3 convolutional layers and a $4\times$ upsampling layer. We place an upsampling layer before the final convolutional layer in expectation of refinement on

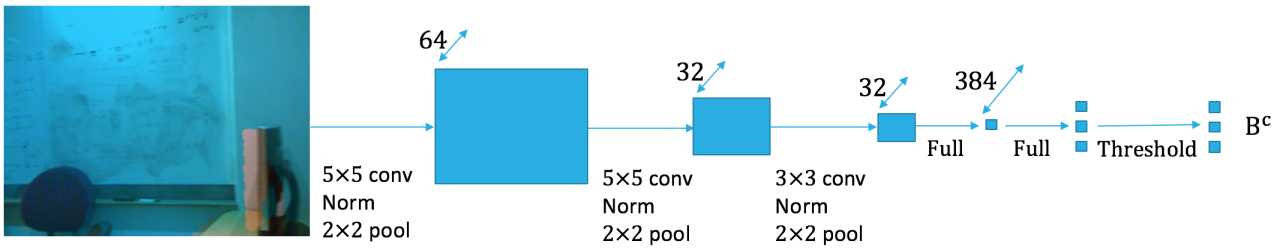

Figure 1: Network architecture for BL estimation

the upsampled feature map. The final output is a predicted depth map $d(x)$. After predicting scene depth, the TM for red channel of the input underwater image are computed using $t^r(x) = e^{-\beta^r d(x)}$ where β^r is the underwater attenuation coefficient for red channel which is set $1/5$ as default. Lastly, the TM for red channel is resized and further refined with guided filter. For green and blue channel, with their attenuation coefficient β^g and β^b computed using Eq. (10) from [7], their TMs are calculated by:

$$t^k(x) = t^r(x)^{\frac{\beta^k}{\beta^r}}, k \in \{g, b\} \tag{3}$$

Finally, the restored underwater color image is calculated using Eq. (2).

Training Data: To create sufficient data for training, we synthesize underwater images using the NYU depth dataset v2 [13] which contains 1449 densely labeled pairs of aligned RGB and depth images of indoor scenes. We choose 29 typical underwater ambient lights, shown in Fig. 3(a). With the paired depth map d, the TM can easily be computed. To increase the training set, we apply a scale factor s and offset f to generate a new depth map $d_{new} = s \times d + f$. In our experiment, s is randomly chosen from $[0.875, 1.125]$ and f is randomly chosen from $[0, 1.5]$. With Eq. (2), a synthesized underwater image is generated. Fig. 3(d) demonstrates example synthetic underwater images based on the scene radiances shown in Fig. 3(b) and their corresponding depth maps shown in Fig. 3(c).

Training Parameters: The minimizer used for both networks is AdamOptimizer with learning rate set to 10^{-4}. The dropout rate is 0.8. We use Tenserflow with GPU acceleration and take $12,000$ synthetic underwater images for training the BL estimation and depth prediction networks. For BL, we randomly choose initial weights from a normal distribution with standard deviation 0.01. The depth prediction network adopts weights from VGG net and [12] as initial values. The loss function for the BL network is:

$$Loss_{BL} = \sum_{c \in \{r,g,b\}} (B^c_{est} - B^c_{gt})^2 \tag{4}$$

where B^c_{est} is the estimated BL of channel c and B^c_{gt} is from ground truth. For depth estimation, we adopt the scale-invariant mean squared loss function from [12]:

$$Loss_{depth} = \frac{1}{n}\sum_x \delta(x)^2 - \frac{1}{n^2}(\sum_x \delta(x))^2 \tag{5}$$

where n is the total number of pixels and $\delta(x)$ is the difference of the logarithm between prediction and ground truth at pixel x: $\delta(x) = \log d_{est}(x) - \log d_{gt}(x)$.

III. Experimental Results

We first compare the proposed method against five underwater image restoration methods [3]–[7] using 3000 synthetic underwater images. The first row of Table I shows the average mean square error (MSE) per channel for BL estimation for all the compared methods. The next two rows show the average PSNR and SSIM results for restoring the images. The proposed method outperforms the others.

	Original	[3]	[4]	[5]	[6]	[7]	Proposed
MSE / ch		.045	.069	.070	.056	.042	**.006**
PSNR		16.83	17.08	16.46	14.67	16.72	**20.26**
SSIM		.72	.74	.72	.67	.74	**.85**
UCIQE	0.51	0.59	0.57	0.55	0.59	0.58	**0.60**
UIQM	3.79	4.17	4.10	4.02	4.09	4.24	**4.26**

Table I: Average results: (First three rows, synthesized images) MSE per channel for BL estimation, PSNR and SSIM for restoration, (Last two rows, real images) UCIQE and UIQM for restoration.

Next, we apply the proposed method to real underwater images. Fig. 4 shows examples. Although ground truth depth and BL are not available, they appear to be reasonable (all the TMs shown have undergone a linear stretching for display) and the restored images present better contrast and more vivid colors compared to the input images. Fig. 5 shows restored images with the methods in [3]–[7] and our method. The proposed method reveals more details. Lacking reference images for real underwater images, we apply two no-reference objective underwater image quality metrics, UCIQE [14] and UIQM [15] on 50 real underwater images (test images from Google Image). In Table I, the proposed method outperforms the other methods with these metrics.

In conclusion, we proposed an underwater image restoration method based on deep networks. BL and scene depth for an underwater image are estimated via a 5-layer ConvNet

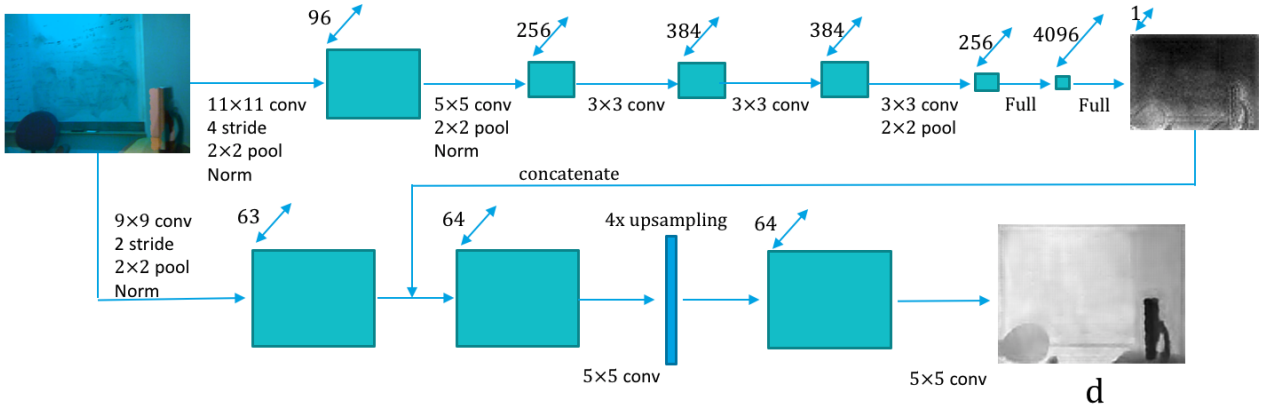

Figure 2: Network architecture for transmission estimation

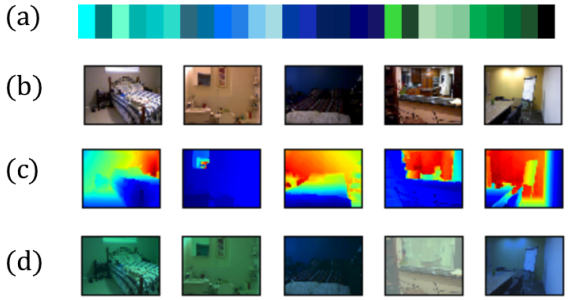

Figure 3: (a) 29 common underwater BLs. (b) Indoor images from [13] with (c) corresponding depth maps. (d) Example synthetic underwater images based on (a)-(c).

and a multi-scale deep network, allowing the input underwater image to be restored using the image formation model. Using synthetic underwater images with different BLs and real underwater images, we demonstrate that the proposed method produces satisfying restored results and outperforms other state-of-the-art IFM-based methods.

REFERENCES

[1] R. Fattal, "Single image dehazing," *ACM Trans. Graphics*, vol.27, no. 3, pp. 72 1-729, 2008.

[2] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2341-2353, Dec. 2011.

[3] H. Wen, Y. Tian, T. Huang, and W. Gao, "Single underwater image enhancement with a new optical model," *in Proc. IEEE Int. Symp. Circuits & Syst. (ISCAS)*, May 2013, pp. 753-756.

[4] P. Drews, E. do Nascimento, F. Moraes, S. Botelho, and M. Campos, "Transmission Estimation in Underwater Single Images," *in Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCVW)*, pp. 825-830, Dec. 2013.

[5] X. Zhao, J. Tao, and Q. Song. "Deriving inherent optical properties from background color and underwater image enhancement," *Ocean Eng.*, vol. 94, pp. 163-172, Jan. 2015.

[6] Y.-T. Peng, X. Zhao, and P. C. Cosman, "Single Underwater Image Enhancement using Depth Estimation based on Blurriness," *in Proc. ICIP*, pp. 4952-4956, Sep. 2015.

[7] Y.-T. Peng and P. C. Cosman."Underwater Image Restoration Based on Image Blurriness and Light Absorption." *IEEE Transactions on Image Processing* 26.4 (2017): 1579-1594.

[8] W. Ren, L. Si, H Zhang, J. Pan, X. Cao, and M.-H. Yang. "Single image dehazing via multi-scale convolutional neural networks," *ECCV*, pp. 154-169., 2016.

[9] Q. Zhu, J. Mai, and L. Shao, "A Fast Single Image Haze Removal Algorithm Using Color Attenuation Prior," *IEEE Trans. Image Process.*, vol. 24, pp. 3522-3533, Nov. 2015.

[10] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "DehazeNet: An End-to-End System for Single Image Haze Removal," *IEEE Trans. Image Process.*, vol. 25, pp. 5187-5198, Nov. 2016.

[11] X. Fan, Y. Wang, X. Tang, R. Gao, and Z. Luo, "Two-Layer Gaussian Process Regression with Example Selection for Image Dehazing," *IEEE Trans. Circuits Syst. Video Technol.*, 2016.

[12] D. Eigen, C. Puhrsch, and R. Fergus, "Depth map prediction from a single image using a multi-scale deep network," *Advances in neural information processing systems*, pp. 2366-2374, 2014.

[13] http://cs.nyu.edu/∼silberman/datasets/nyu_depth_v2.html

[14] M. Yang and A. Sowmya. "An underwater color image quality evaluation metric." *IEEE Transactions on Image Processing* 24.12 (2015): 6062-6071.

[15] K. Panetta, C. Gao, and S. Agaian. "Human-visual-system-inspired underwater image quality measures." *IEEE Journal of Oceanic Engineering* 41.3 (2016): 541-551.

Figure 4: Examples of restoration of real underwater images based on the proposed method. A: Original underwater image. B: Estimated BL. C: Estimated transmission map. D: Restored images.

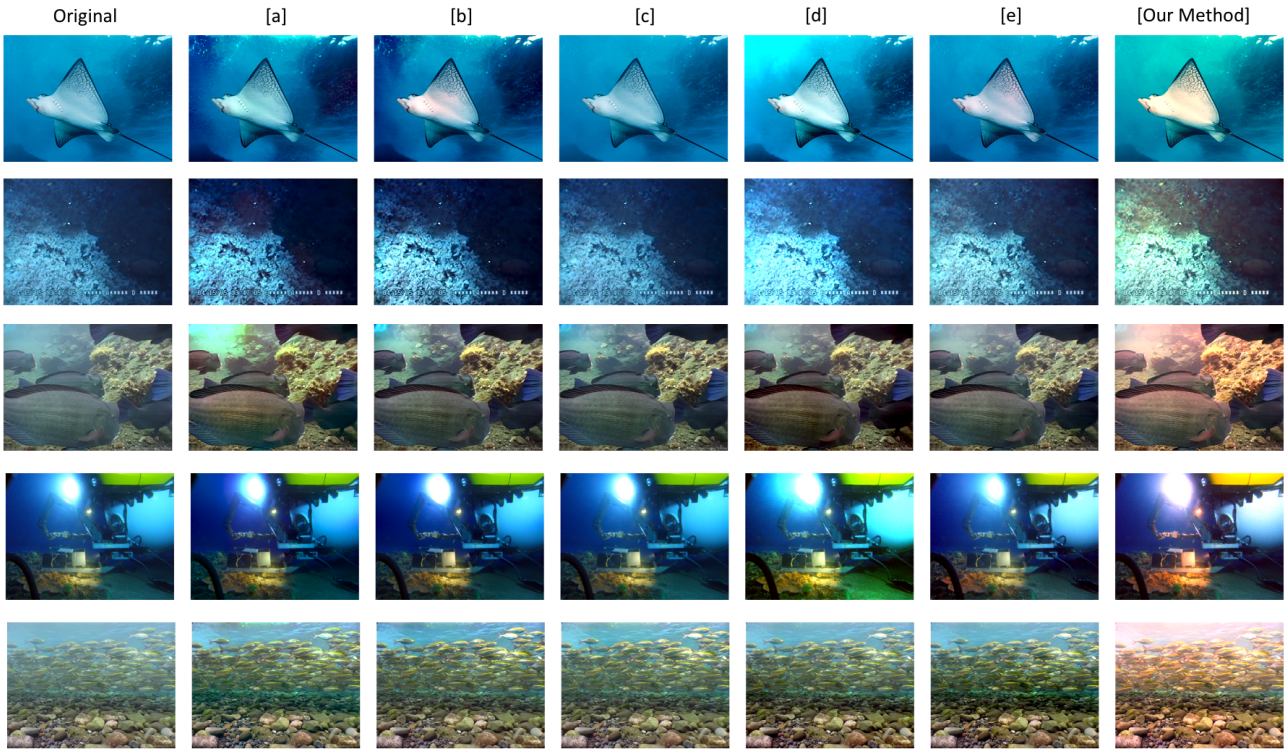

Figure 5: Restored images visual comparision. (a) [3]. (b) [4]. (c) [5]. (d) [6]. (e) [7]