DehazeGAN: Underwater Haze Image Restoration using Unpaired Image-to-image Translation

Younggun Cho∗ Ramavtar Malav∗∗ Gaurav Pandey∗∗∗ Ayounh Kim∗

∗ Department of Civil and Environmental Engineering, KAIST, Daejeon, Korea, (e-mail: {yg.cho,ayoungk}@kaist.ac.kr).
∗∗ Department of Electrical Engineering, IIT, Kanpur, India, (e-mail: {rmalav}@iitk.ac.in).
∗∗∗ Ford Motor Company, Palo Alto, CA, USA, (e-mail: {gpandey.iitk}@gmail.com).

Abstract:
In this paper, we propose a Generative Adversarial Networks (GAN)-based image restoration method. Our method adopts an unpaired image-to-image translation network to learn the characteristics of underwater haze images. To enhance restoration, we propose multiple cyclic consistency losses that capture the detail of images and suppress distortion image translation. To prepare unpaired images of clean and degraded scenes, we collected images from Flickr and filter out false images using image characteristics. The proposed network is tested on public underwater images and shows promising results under severe image distortion.

1. INTRODUCTION
Optic cameras and sonar are widely applied sensors for underwater perception. Sonar is a specialized sensor for underwater environments, showing a strong advantage against optical imaging sensors with their robustness to the water turbidity and visible range. With the acoustic-based sensing, target object detection in the submerged condition is operated in the best condition. Nevertheless, optic cameras have been exploited in many underwater applications, since optical images are intuitive to apply and use for robot and human vision. Especially for target detection, inspection, and human-robot interaction, detecting feature and changes from the optic images are still important. Furthermore, the higher resolution image can be exploited in the post-processing phase. For example, feature extraction and detection, depth estimation, and visual simultaneous localization and mapping (SLAM) can be performed using optical images.

Securing reliable visual perception has been challenging for many robotics applications, including visual SLAM, visual odometry, and visual servoing. Since Fattal (2008) first proposed the single image dehazing using Independent Component Analysis (ICA), many approaches have been introduced to enhance the visibility of an image. A fast dehazing introduced by Tarel and Hautiere (2009) works regardless of the channel using median filtering. The well-known Dark Channel Prior (DCP) by He (He et al., 2011) is also widely used. Later Zhu et al. (2015) developed Color Attenuation Prior (CAP) that provides the prior using the saturation-transmission relation. Recently, Fattal (2014) exploits color lines when predicting the ambient light.

This line of researches in dehazing is either model based or non-model based. The model-based approach is initiated since the work by Jaffe (1990); McGlamery (1980). Many underwater dehazing methods Carlevaris-Bianco et al. (2010); Chiang and Chen (2012); Galdran et al. (2015) used a simplified degradation model which focus on single scattering attenuation in a water medium.

The non-model-based approaches assume no prior information. Histogram Equalization (HE) and Contrast-limited Adaptive Histogram Equalization (CLAHE) are mostly widely used method for overall image enhancement including dehazing Zuiderveld (1994). For underwater applications, Eustice et al. (2002) presented using CLAHE for underwater visual SLAM. More recently, Ancuti and Ancuti (2013) proposed fusion-based dehazing using Laplacian and Gaussian spaces.

Among non-model-based approaches, interests in deep learning-based dehazing have been increased. DehazeNet by Cai et al. (2016) shows estimating image transmission in the application to the dehazing. Later Shin et al. (2016) proposed a Convolutional Neural Network (CNN)-based approach to estimate both colored ambient light and transmission map. Motivated by the work of Li et al. (2017a) who proposed a method to create synthetic underwater images using GAN, we try to solve both problems (generating datasets and training networks) using the unpaired image-to-image translation method.

2. PROPOSED METHOD

We now describe an underwater image restoration using a generative model. The basic notation of the paper is described as follows. We adopt cyclic architecture Zhu et al. (2017a) for unpaired image translation across two domains A and B. Generators (G_{AB}, G_{BA}) transform images from the source domain to target domain (G_{AB}(a) \rightarrow b) where a
To obtain a detail-enhanced image, we apply Laplacian loss (Li et al., 2017b) to the networks. In this paper, we set domain \( A \) as degraded (hazy) images and domain \( B \) as clean (aerial) images.

\[ L_{\text{rec}} = E_{a,a'}[||a - a'||_1]. \] (2)

\( L_1 \) norm was used for pixel space distance.

To obtain a detail-enhanced image, we apply Laplacian loss (Li et al., 2017b) to the networks. In this paper, we set domain \( A \) as degraded (hazy) images and domain \( B \) as clean (aerial) images.

\[ L_{\text{rec}} = E_{a,a'}[||a - a'||_1]. \] (2)

\( L_1 \) norm was used for pixel space distance.

The proposed network was implemented using PyTorch and trained with an NVIDIA GTX 1080ti. We employed the Adam solver [Kingma and Ba (2014)] with \( \beta_1 = 0.9 \), \( \beta_2 = 0.99 \), and \( \omega_{\text{decay}} = 10^{-5} \). We started the training with an initial learning rate of \( 10^{-4} \) and controlled it by a step scheduler with a step-size of 20 and \( \gamma = 0.5 \).

For dataset preparation, we gathered unpaired training images from Flickr. Training images were collected using
Fig. 2. Comparison result of the proposed method. Test images were collected from publicly open underwater images of Youtube. We evaluated different underwater images with various water conditions. As in the figure, our method shows meaningful performance on severe distortion images.

several keywords, such as underwater haze, distortion and clear scene. To get reliable datasets, we filtered outlier images using previous haze relevant priors and the level of color distortion.

3.2 Results

For validation, we compared the proposed network with other dehazing methods as Fig. 2. We tested each method on real underwater images available online. As expected, we get meaningful results that the dehazing network presents enhanced images with corrected color. Recent dehazing methods show relatively similar results on hazy scenes without color distortion (first row), however when color is severely distorted, other methods show limitation on haze parameter estimation (transmission and ambient light). The result in third row shows the promising performance of the proposed method. Unlike other methods, our method learn the characteristics of underwater haze images from various underwater scenes. Since the dehazing network (generator $G_{AB}$) directly translates the haze image into the clean aerial image, and the proposed method does not suffer from significant distortion and degradation.

4. CONCLUSION

In this paper, we propose an underwater image restoration method using deep generative models. We utilized cyclic architecture and constructed multiple consistency losses for detail enhanced images. The proposed method shows promising restoration performance on underwater images with significant distortion. For further works, we evaluated our methods on various underwater images and utilized with other robot vision application.

REFERENCES

Ancuti, C.O. and Ancuti, C. (2013). Single image dehazing by multi-scale fusion. IEEE Transactions on Image Processing, 22(8), 3271–3282.
Cai, B., Xu, X., Jia, K.and Qing, C., and Tao, D. (2016). Dehazenet: An end-to-end system for single image haze removal. 25(11), 5187–5198.
Carlevaris-Bianco, N., Mohan, A., and Eustice, R.M. (2010). Initial results in underwater single image dehazing. 1–8. IEEE, Seattle.
Chiang, J.Y. and Chen, Y. (2012). Underwater image enhancement by wavelength compensation and dehazing. 21(4), 1756–1769.
Eustice, R., Pizarro, O., Singh, H., and Howland, J. (2002). UWIT: Underwater image toolbox for optical image processing and mosaiccing in matlab. 141–145. IEEE.
Fattal, R. (2008). Single image dehazing. ACM Transactions on Graphics, 27(3), 72–80.
Fattal, R. (2014). Dehazing using color-lines. ACM Transaction on Graphics, 34(1), 13:1–13:14.
Galdran, A., Pardo, D., Picón, A., and Alvarez-Gila, A. (2015). Automatic red-channel underwater image restoration. Journal of Visual Communication and Image Representation, 26, 132–145.
He, K., Sun, J., and Tang, X. (2011). Single image haze removal using dark channel prior. 33(12), 2341–2353.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, 770–778.

Isola, P., Zhu, J.Y., Zhou, T., and Efros, A.A. (2017). Image-to-image translation with conditional adversarial networks. arXiv preprint.

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

Johnson, J., Alahi, A., and Fei-Fei, L. (2016). Perceptual losses for real-time style transfer and super-resolution. In European Conference on Computer Vision, 694–711. Springer.

Kingma, D.P. and Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Li, J., Skinner, K.A., Eustice, R.M., and Johnson-Roberson, M. (2017a). WaterGAN: Unsupervised generative network to enable real-time color correction of monocular underwater images. arXiv preprint arXiv:1702.07392.

Li, S., Xu, X., Nie, L., and Chua, T.S. (2017b). Laplacian-steered neural style transfer. arXiv preprint arXiv:1707.01253.

McGlamery, B. (1980). A computer model for underwater camera systems. In Ocean Optics VI, 221–231. International Society for Optics and Photonics.

Shin, Y.S., Cho, Y., Pandey, G., and Kim, A. (2016). Estimation of ambient light and transmission map with common convolutional architecture. 1–7. Monterey, CA.

Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

Tarel, J. and Hautiere, N. (2009). Fast visibility restoration from a single color or gray level image. 2201–2208. IEEE.

Zhu, J.Y., Park, T., Isola, P., and Efros, A.A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. arXiv preprint arXiv:1703.10593.

Zhu, Q., Mai, J., and Shao, L. (2015). A fast single image haze removal algorithm using color attenuation prior. 24(11), 3522–3533.

Zuiderveld, K. (1994). Contrast limited adaptive histogram equalization. In Graphics gems IV, 474–485. Academic Press Professional, Inc.