UNDERWATER IMAGE ENHANCEMENT BASED ON STRUCTURE-TEXTURE DECOMPOSITION

Jingyu Yang†, *Xinyan Wang*†, *Huanjing Yue**†, *Xiaomei Fu*‡, *Chunping Hou*†

† School of Electrical and Information Engineering, Tianjin University
‡ School of Marine Science and Technology, Tianjin University

ABSTRACT

Underwater images generally suffer from low contrast, serious noise and color distortion. The main challenges of underwater image enhancement are to preserve details in dark regions while avoiding oversaturetion in bright regions. This paper proposes a novel underwater image enhancement method based on image decomposition. By decomposing the high-frequency texture and noise into the texture layer, the transmission map is estimated from the noise-free structure layer to avoid the noise amplification problem in underwater image enhancement. Both the structure layer and texture layer are descattered with the estimated transmission map. After denoising by gradient residual minimizition, the texture layer is enhanced and added back into the structure layer to recover the final enhanced image. Experimental results verify that the proposed approach can recover the high-quality images with fine details and edges while improving contrast and color naturalness, especially for images taken in the high turbidity environment.

Index Terms— Underwater images, color correction, structure-texture decomposition, contrast enhancement, denoising

1. INTRODUCTION

The quality of images heavily depends on the characteristics of the medium along the light path. In mediums with large suspended particles, such as haze or turbidity, the absorption and scatter severely degrade the image quality. Images taken in underwater environments (named underwater images hereafter) suffer from low contrast due to weak lighting or color cast due to absorption and scatter in underwater imaging. It is challenging to restore details and natural colors from degraded underwater images.

To some extent, the degradation mechanism in underwater images and foggy images are quite similar. Dehazing methods [1, 2] could be adapted to the enhancement of underwater images. However, directly applying dehazing methods to underwater images tends to lose details in the dark regions and causes oversaturation in the bright regions.

Recent years, numerous underwater image enhancement methods have been proposed with various considerations [3, 4]. Nicholas *et al.* [5] explore single image dehazing techniques for underwater images. Simon *et al.* [6] employ a hierarchical rank-based method to estimate the veiling light and prevent oversaturation and artifacts. Li *et al.* [7] introduce a global background light estimation algorithm based on the quad-tree subdivision and propose an estimation algorithm for media transmission mapping. Chiang *et al.* [8] propose a systematic approach to attenuate discrepancy along the light path by compensation. Ancuti *et al.* [9] enhance underwater image by fusing four weight maps that aim to increase the visibility of degraded details. These methods have improved the visual quality of underwater images with proper contrast, color and sharpness.

Despite these efforts, existing methods still have some problems of unnatural color, oversaturation, exposure and amplified noise, especially in the underwater environment with artificial illumination and high turbidity. This paper introduces a new underwater image enhancement method based on structure-texture decomposition. Firstly, the color of captured underwater image is corrected. Secondly, the corrected image is decomposed into two layers: the structure layer and texture layer. The transmission can be reliably estimated from the structure layer without the interference of texture and noise. Then denoising and contrast enhancement are conducted in the two layers respectively. Finally, we get the final restored image by recomposing the recovered layers and refined mask. Experimental results show that the proposed method can suppress the amplified noise and obtain excellent restored image with fine details. In addition, the reliability of the decomposition model is analysed in this paper.

The remainder of this paper is organized as follows. Section 2 gives the proposed image enhancement method. Validation experiments are presented in Section 3, and the paper is concluded in Section 4.

This work was supported by the National Natural Science Foundation of China under Grant 61372084, Grant 61672378, Grant 61571323, and Grant 61520106002. *Corresponding author: dayueer@tju.edu.cn.

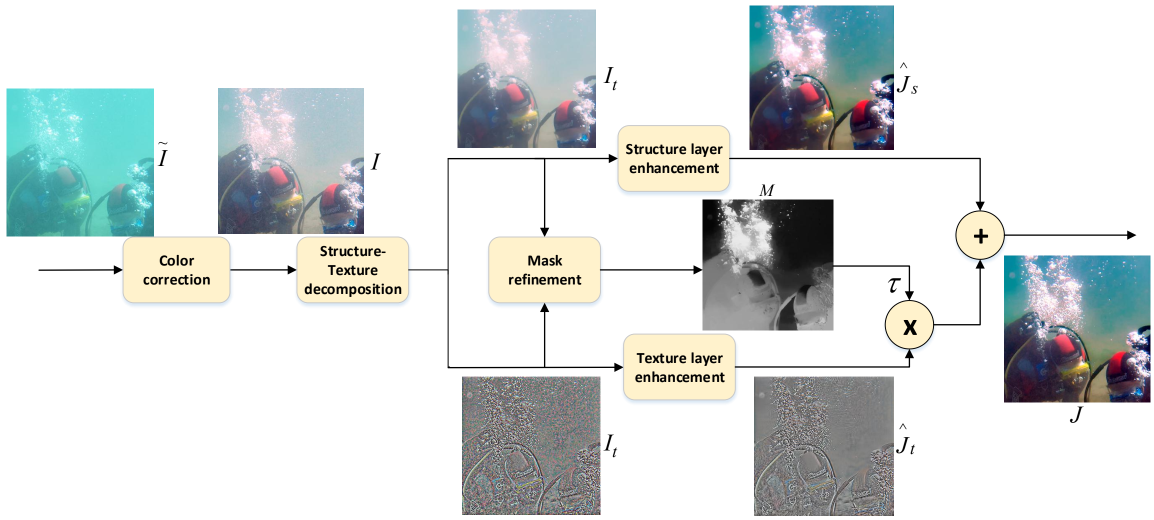

Fig. 1. Framework of the proposed enhancement method based on structure-texture decomposition.

2. PROPOSED METHOD

The main challenge of underwater image enhancement is to restore visually-natural details without amplifying artifacts. To this end, we propose an underwater image enhancement method based on decomposition. The framework of our approach is shown in Fig. 1. Firstly, the color is corrected by histogram equalization. Secondly, we use the total-variation and L1 norm minimization (TV-L1) to decompose the image into texture and structure layers. Then the structure and texture layers are enhanced and denoised separately, and combined to refine texture mask. Finally, after proper scaling, the cleaned texture layer is added back to the enhanced structure layer to generate the artifacts free output.

2.1. Color Correction

Since the light wave suffers from strong absorption and attenuation under water, and the light energy decreases as the water depth, the captured images tend to appear a bluish tone.

We utilize the simple histogram equalization [10] to solve the color shift problem. The rationale is quite intuitive: shift the mean color intensity of each channel to the middle of the desired range, and linearly normalize the histogram with truncation. For each color channel $\tilde{I}^c$, we calculate the lower bound $\tilde{I}^c_{min}$ and upper bound $\tilde{I}^c_{max}$ as follows:

$$\tilde{I}^c_{min} = \mu^c - \lambda\sigma^c, \tag{1}$$

$$\tilde{I}^c_{max} = \mu^c + \lambda\sigma^c, \tag{2}$$

where $c \in \{R, G, B\}$, μ^c and σ^c represent respectively the mean value and the standard deviation of each color channel. λ is a tune parameter to adjust the valid intensity range, beyond which pixel values would be truncated. Finally, the corrected image $I(x)$ is obtained by normalization.

$$I^c(x) = \text{clip}(255(\tilde{I}^c(x) - \tilde{I}^c_{min})/(\tilde{I}^c_{max} - \tilde{I}^c_{min})), \tag{3}$$

where clip$(\cdot)$ is a clip function that truncates the operand to the range of [0, 255].

2.2. Structure-Texture Decomposition

2.2.1. *Decomposition Model*

Since the degradation of the light path along turbid water is the same as that along hazy air, we use the image hazing model [1] for the observation model of underwater images, namely

$$I^c(x) = J^c(x)t(x) + A^c(1 - t(x)) + E(x), \tag{4}$$

where x is the pixel index, $I^c(x)$ is the corrected underwater image in channel c, $J^c(x)$ is the restored image, A^c is the global background light, and $t(x)$ is the medium transmission map. We denote $E(x)$ the artifacts of input image. The purpose of image enhancement is to recover $J^c(x)$ from $I^c(x)$, which usually involves estimating transmission map $t(x)$ and background light A^c.

Note that natural images usually contain considerable high-frequency textures and non-ignorable amount of noise, particularly in weak light conditions. Such interfering factors would make the estimation of the transmission map less reliable, which affects the enhancement performance in turn. To overcome this problem, we firstly decompose the image into structure and texture layers, and then estimate the transmission map from the noise-free structure layer. In the experiments, we show the validness of such decomposition for underwater images. Let J_s and J_t denote the structure layer and texture layer respectively, and $J = J_s + J_t$. Then, the image model in Eq. (4) can be expressed as

$$\begin{aligned} I^c(x) &= ({J_s}^c(x) + {J_t}^c(x))t(x) + A^c(1 - t(x)) + E(x) \\ &= {J_s}^c(x)t(x) + A^c(1 - t(x)) + {J_t}^c(x)t(x) + E(x). \end{aligned} \tag{5}$$

Meanwhile, the observed underwater image could be also decomposed into two layers, $I = I_s + I_t$. Considering that the transmission map $t(x)$ is piecewise smooth signal having nearly the same structure as $I_s(x)$. Therefore, we suppose $t(x)$ can be estimated from $I_s(x)$. Namely, the decomposition of I_s and I_t satisfies the following relationship:

$$I_s{}^c(x) = J_s{}^c(x)t(x) + A^c(1 - t(x)), \tag{6}$$

$$I_t^c(x) = J_t^c(x)t(x) + E(x), \tag{7}$$

where $I_s{}^c$ is the structure layer containing most features of the captured image. The texture layer image $I_t{}^c$ contains a majority of textures and noise. In this way, $t(x)$ is estimated from $I_s(x)$ without interference of high-frequency details and noise.

2.2.2. *Structure-Texture Decomposition*

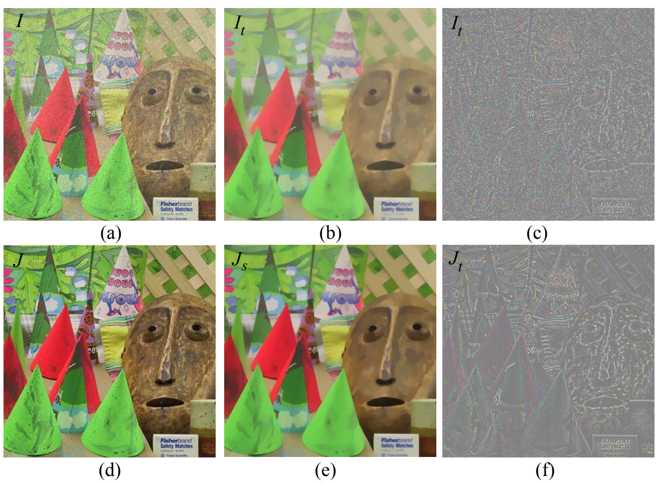

Fig. 2. One example of structure-texture decomposition. (a) Synthetic degraded image I, (b) Structure layer I_s of I, (c) Texture layer I_t of I, (d) Original image J, (e) Structure layer J_s of J, and (f) Texture layer J_t of J.

The structure layer has larger gradient magnitudes at the contours and boundaries of objects, while the texture layer captures fine details that exhibit smaller gradient magnitudes. To separate the structure and texture, we use the popular TV-L1 model:

$$\min_{I_s} \sum_x (I_s(x) - I(x))^2 + \xi |\nabla I_s(x)|_1, \tag{8}$$

where the gradient of structure component is regularized by the sparsity-promoted L1 norm, and ξ is a weighting parameter that controls the balance between accuracy and smoothness. The TV-L1 minimization is solved by the alternating direction method [11]. Once the structure layer I_s is obtained, the texture layer with noise is calculated by $I_t = I - I_s$.

A natural question is whether the TV-L1 model could separate the two layers as expected in Eq. (6) and Eq. (7). Fig. 2 shows one example for the decomposition. The input image I is synthesized from J by Eq. (4). Both I and J are decomposed into two layers using the TV-L1 minimization. Fig. 2 shows that the structure layer (I_s) of the degraded image (I) does not contain high-frequency details and noise, and is visually consistent with the structure layer (J_s) of the original image (J). In the paper, we amplified the texture by a factor of 10 and shift it by +0.5 for better visualization.

2.3. Enhancement

We estimate A by the global back-scattered light estimator method [14]:

$$A^c = \max_{y \in M_{DC}^{I_s}} \left(\min_{x \in \Omega(y)} I_s^c(x) \right), \tag{9}$$

where $M_{DC}^{I_s}$ is the location of the maximum value of the dark channel in the structure layer, namely $M_{DC}^{I_s} = \{y | I_{s_{DC}}(y) = I_{s_{DC}}^{max}\}$ and $\Omega(y)$ is a neighborhood around y.

Without the disturbance of noise, the pixels of similar colors in the haze-free image of I_s, i.e. J_s, form a line in I_s. Therefore we adopt the haze-line based dehazing method [2] to remove the haze in I_s. We first convert $\bar{I}_s(x)$ ($\bar{I}_s(x) = I_s(x) - A$) to spherical coordinates ($[r(x), \theta(x), \phi(x)]$) and cluster the pixels according to their longitude ($\theta(x)$) and latitude ($\phi(x)$) values. Then the transmission $\bar{t}(x)$ is calculated according to the radius ($r(x)$) values. Since $\bar{t}(x)$ is calculated pixel-wise, it may be not smooth in spatial space, which is not consistent with the smooth assumption of transmission maps. Therefore, we regularize $\bar{t}(x)$ by minimizing

$$\alpha \sum_x \frac{[\hat{t}(x) - \bar{t}(x)]^2}{\sigma^2(x)} + \beta \sum_x \sum_{y \in \mathcal{N}(x)} \frac{[\hat{t}(x) - \hat{t}(y)]^2}{\|I_s(x) - I_s(y)\|^2}, \tag{10}$$

where α and β are the weighting parameters to balance the data and smooth terms, $\sigma(x)$ is the standard deviation of the estimated transmission $t(x)$ calculated along each haze-line, and $\mathcal{N}(x)$ represents the four nearest neighbors of x. In this way, we obtain the refined transmission map $\hat{t}(x)$. For more information, pls. refer to [2].

Hereafter, the enhanced structure layer $\hat{J}_s(x)$ is obtained by

$$\hat{J}_s(x) = \{I_s(x) - [1 - \hat{t}(x)]A\}/\hat{t}(x). \tag{11}$$

Since the texture layer contains much noise, we utilize the gradient residual minimization method [13] to enhance the texture layer. The enhanced texture layer $\hat{J}_t(x)$ is obtained by minimizing

$$\begin{aligned} &\min_{\hat{J}_t, E} \sum \{ \|\hat{J}_t(x)\hat{t}(x) + E(x) - I_t(x)\|_2^2 + \delta \|E(x)\|_1 \\ &+ \eta \|\nabla \hat{J}_t(x) - \nabla I_t(x)\|_1 \}, \end{aligned} \tag{12}$$

where δ and η are weighting parameters. We solve it utilize a classic TV solver [15].

We observe that the enhanced texture layer $\hat{J}_t(x)$ may still contain undesirable artifacts, especially in smooth regions, as shown in Fig. 1. Therefore, we utilize a binary mask M to separate the original texture layer into smooth and detail regions, and remove the remained details in smooth regions of enhanced texture layer. Similar to [16], we utilize the discrete cosine transform (DCT) coefficients to check whether

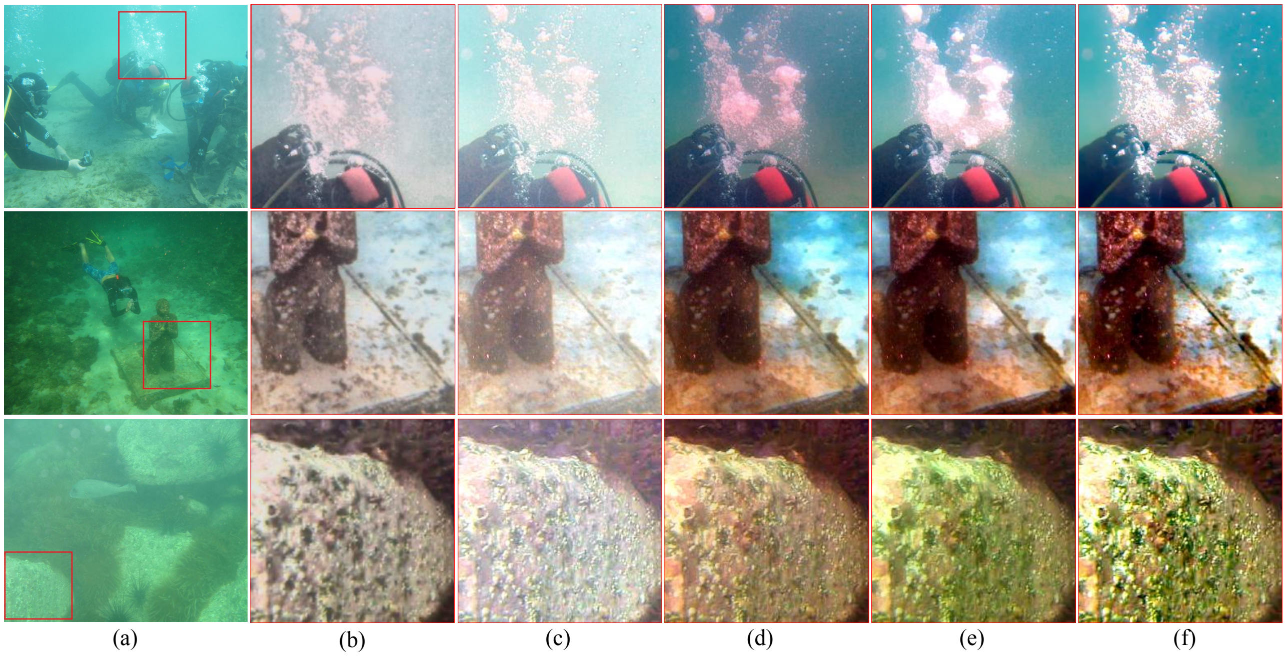

Fig. 3. Comparison of enhancement results. From left to right: (a) the input image, the results of [10] (b), [12] (c), [13] (d), proposed method without decomposition (e), and (f) proposed method with decomposition.

the block is smooth or not. For each 8×8 block in I_t, we denote its DCT coefficients as B. The likelihood of this block to be details can be expressed as

$$\rho = \sum_{x,y} B_{x,y}^2 - B_{1,1}^2 - B_{1,2}^2 - B_{2,1}^2, \tag{13}$$

where x, y denotes the coordinates. If ρ is larger than κ, we set M to be 1 for this block. Otherwise M is set to 0. κ is set to 0.1 in our experiments. Considering M is a coarse estimation, we utilize soft matting to refine the mask M with I_s generating the matting Laplacian matrix [17]. The refined mask is denoted as $\hat{M}$.

Finally, the enhanced underwater image $\hat{J}(x)$ is recovered by:

$$\hat{J}(x) = \hat{J}_s(x) + \tau \hat{J}_t(x) \hat{M}(x), \tag{14}$$

where τ is a scale factor. $\hat{J}_s(x)$ and $\hat{J}_t(x)$ are the enhanced structure and texture layers, respectively. Note that τ is introduced to further enhance the details, which is a common strategy in image super-resolution. In experiments, we set τ to $\frac{1}{\hat{t}(x)}$.

3. EXPERIMENTAL RESULTS

The parameters in the proposed method are set as: $\lambda = 2.3$, $\xi = 1$, $\alpha = 1$, $\beta = 0.1$, $\delta = 0.03$, $\eta = 0.1$. To demonstrate the effectiveness of the proposed method, we compare it with state-of-the-art underwater image enhancement methods [10, 12] and the haze-line based dehazing method [13]. Fig. 3 presents the comparison results. It can be observed that the results of [10] amplify the artifacts, and the tone of enhanced images is not natural. The results of [12] have low contrast and the artifacts are obvious. The dehazing method [13] could improve image contrast and reduce artifacts to some extent. However, since it is not designed for underwater image enhancement, it has color deviation and the details are not enhanced well.

In addition, we compare with the results generated with proposed method without decomposition, namely directly enhancing the color corrected image. It can be observed that the details are lost to some extent, such as the bubbles in the first image. It further demonstrates the effectiveness of proposed texture-structure decomposition based enhancement method. In a word, our method achieves the best results with reduced artifacts, natural color, and fine details.

4. CONCLUSION

This paper proposes an underwater image enhancement method based on structure-texture decomposition. We separate noise into the texture layer and the structure layer could satisfy the haze-line model. We enhance the structure layer by a haze-line based dehazing method and enhance the texture layer by utilizing the gradient residual minimization method. After proper scaling, the enhanced texture layer is added back to the enhanced structure layer, generating the final enhanced image. Experimental results show that our method generates the most pleasing enhancement results with high contrast, natural color and fine details.

5. REFERENCES

[1] Kaiming He, Jian Sun, and Xiaoou Tang, “Single image haze removal using dark channel prior,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, pp. 2341–2353, 2010.

[2] Dana Berman, Tali Treibitz, and Shai Avidan, “Non-local image dehazing,” in *IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 1674–1682.

[3] Mark Sheinin and Yoav Y. Schechner, “The next best underwater view,” *Computer Science*, 2015.

[4] Karthik Seemakurthy, “Deskewing of underwater images.,” *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, vol. 24, no. 3, pp. 1046–59, 2015.

[5] Nicholas Carlevaris-Bianco, Anush Mohan, and Ryan M. Eustice, “Initial results in underwater single image dehazing,” in *Oceans*, 2010, pp. 1–8.

[6] Simon Emberton, Lars Chittka, and Andrea Cavallaro, “Hierarchical rank-based veiling light estimation for underwater dehazing,” in *British Machine Vision Conference*, 2015, pp. 125.1–125.12.

[7] Chongyi Li, Jichang Guo, Shanji Chen, and Yibin Tang, “Underwater image restoration based on minimum information loss principle and optical properties of underwater imaging,” in *IEEE International Conference on Image Processing*, 2016, pp. 1993–1997.

[8] J. Y. Chiang and Y. C. Chen, “Underwater image enhancement by wavelength compensation and dehazing.,” *Image Processing IEEE Transactions on*, vol. 21, no. 4, pp. 1756–69, 2012.

[9] P. Bekaert, T. Haber, C. O. Ancuti, and C. Ancuti, “Enhancing underwater images and videos by fusion,” in *IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 81–88.

[10] Xueyang Fu, Peixian Zhuang, Yue Huang, and Yinghao Liao, “A retinex-based enhancing approach for single underwater image,” in *IEEE International Conference on Image Processing*, 2015, pp. 4572–4576.

[11] Yilun Wang, Junfeng Yang, Wotao Yin, and Yin Zhang, “A new alternating minimization algorithm for total variation image reconstruction,” *Siam Journal on Imaging Sciences*, vol. 1, no. 3, pp. 248–272, 2008.

[12] Huimin Lu, Yujie Li, and S Serikawa, “Single underwater image descattering and color correction,” in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2015, pp. 1623–1627.

[13] Chen Chen, Minh N. Do, and Jue Wang, *Robust Image and Video Dehazing with Visual Artifact Suppression via Gradient Residual Minimization*, Springer International Publishing, 2016.

[14] Christophe De Vleeschouwer Cosmin Ancuti, Codruta O. Ancuti, “Multi-scale underwater descattering,” in *International Conference on Pattern Recognition (ICPR), oral, Cancun, Mexico*, 2016.

[15] A Beck and M Teboulle, “Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems.,” *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, vol. 18, no. 11, pp. 2419–34, 2009.

[16] Yu Li, Fangfang Guo, Robby T. Tan, and Michael S. Brown, *A Contrast Enhancement Framework with JPEG Artifacts Suppression*, Springer International Publishing, 2014.

[17] Anat Levin, Dani Lischinski, and Yair Weiss, “A closed-form solution to natural image matting,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 2, pp. 228–42, 2008.