Visibility Enhancement for Underwater Visual SLAM based on Underwater Light Scattering Model

Younggun Cho and Ayoung Kim[1]

Abstract— This paper presents a real-time visibility enhancement algorithm for effective underwater visual simultaneous localization and mapping (SLAM). Unlike an aerial environment, an underwater environment contains larger particles and is dominated by a different image degradation model. Our method starts with a thorough understanding of underwater particle physics (e.g., forward, back, multiple scattering, blur and noise). Targeting underwater image enhancement in a real-world application, we include an artificial light model in the derivation. The proposed method is effective for both color and gray images with substantial improvement in the process time compared to conventional methods. The proposed method is validated by using simulated synthetic images (color) and real-world underwater images (color and grayscale). Using two underwater image sets acquired from the same area but with different water turbidity, we evaluate the proposed visibility enhancement and camera registration improvement in SLAM.

I. INTRODUCTION

In recent years, there has been an increase in demand for underwater inspection and exploration using an autonomous underwater vehicle (AUV). Underwater is a typical global positioning system (GPS) denied environment, where visual information provides valuable navigation queues for robot navigation. Specifically, visual SLAM provides an effective solution to many underwater navigation applications including inspection [17], surveillance [20] and, place recognition [19]. However, underwater visual navigation has many issues due to the low visibility. The main drawback of using underwater images is their strong dependency on the water medium. The underwater images are often critically degraded by poor water medium conditions, such as tidal current and turbidity, and result in haze and blurring, contrast loss, and color distortion. Moreover, ocean environments change easily within just few hours (Fig. 1), which addresses the need for instant visibility enhancement to be used in robot navigation.

Many visibility enhancement methods have been recently proposed to solve image dehazing problems [9, 12, 25, 4, 1, 22, 3]. This line of research has mostly taken place in the computer vision fields, with the main focus upon natural color restoration of single image. Previous studies on haze removal can be divided into two classes: turbid atmosphere conditions and underwater environments.

Fattal [9] proposed a single image dehazing method using Independent Component Analysys (ICA) for turbid atmospheric conditions. ICA is the first breakthrough in haze

[1]Y. Cho and A. Kim are with the Department of Civil and Environmental Engineering, KAIST, Daejeon, S. Korea [yg.cho, ayoungk]@kaist.ac.kr
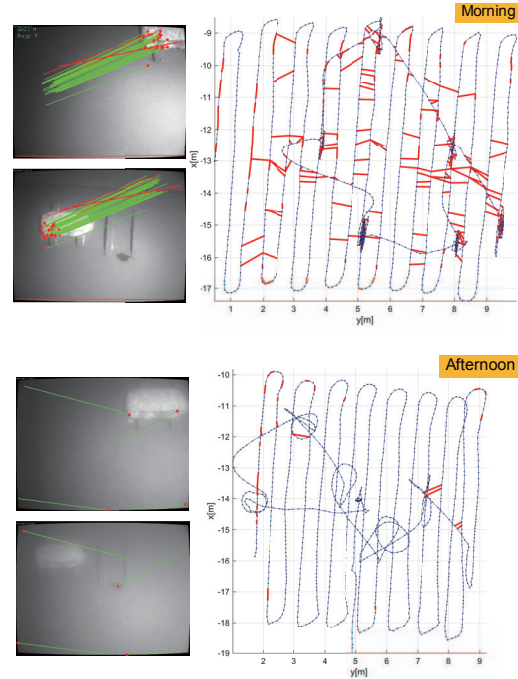
Fig. 1. Sample underwater images of the same area with different turbidity. On camera images (first column), red dots are the feature points detected and green lines show the correspondences established. The top images are from morning dive when the water was clear. On the contrary, the image set in the bottom is from afternoon dive when the water becomes turbid. Images are severely degraded providing few features. The correspondences could not be established. On the right, resulting SLAM trajectories are provided. Red line indicates successful image registration made by recognition. Note the critical difference in the number of registration.

removal methods using a single image, but performance is limited in densely haze scenes and requires a long computation time. Tarel and Hautiere [22] proposed a fast image dehazing method based on median filtering, but the method produces halo artifacts. He et al. [12] demonstrated single image dehazing using Dark Channel Prior (DCP) which is effective and simple. Zhu et al. [25] introduced Color Attenuation Prior (CAP) which represents the relation between medium transmission and image saturation for fast image dehazing. Cai et al. [3] proposed deep learning based approach for single image dehazing. These single image dehazing methods provide effective results in aerial images. However, applying single image dehazing to underwater images shows limited improvements, because the priors and assumptions may not hold on underwater environments

Dehazing methods aiming at underwater environments have also been developed to enhance the visibility of im-

ages. Carlevaris-Bianco et al. [4] developed a single-image dehazing method using wavelength-dependent attenuation of light in water [4]. They calculate the depth prior from the strong difference in attenuation between color channels, and then use the prior to recover the scene radiance by Markov Random Field (MRF). Ancuti et al. [1] proposed multi-scale fusion-based image dehazing. This method shows strong improvements in mild haze images, but boundary artifacts in strong haze images. Babaee and Negahdaripour [2] introduced a hybrid dehazing method by fusing of optical images and acoustic images. This method used MRF based depth map estimation to reconstruct underwater images with high turbidity. However, this method needs heavy computational burden to estimate depth maps from sparse data.

Due to the lack of environmental information, underwater image enhancement is an ill-posed problem and is therefore challenging. Conventional approaches fail on grayscale images and have a long computational time that is impractical for use in robotics. The proposed algorithm presents the following contribution to solving the underwater dehazing problem.

- Dehazing considering underwater particle physics
- Artificial light model for underwater robot applications
- Independence to channel numbers
- Online processing performance

In this paper, we thoroughly investigate underwater particle physics to design an enhancement model per degradation factor. Since underwater robots operate with their own light source, artificial light modeling is needed for the veiling light estimation. Grayscale images have only one channel; therefore, it is challenging to estimate environmental information such as scene depth, attenuation coefficient, and veiling light. The proposed method works equally well for both color and grayscale images. The ultimate objective of the method is to implement it in the visual SLAM pipeline with real-time performance. The substantial improvement in processing time allows for direct implementation of the proposed method in the existing SLAM methods.

The remainder of this paper is organized as follows. Section II describes the underwater image model. Section III gives an overview of the proposed method and explains the detail parts of the method. In Section IV, experimental results for synthetic and real datasets are compared to previous methods. Finally, Section V offers conclusions of the proposed approach.

II. UNDERWATER HAZE MODEL

The most widely used haze image model is Single Scattering Atmospheric Model (SSAM), which has been used in many previous dehazing literatures [12, 22, 2, 25, 3]. In this model, the haze image $I(\mathbf{x})$ is modeled as:

$$I(\mathbf{x}) = J(\mathbf{x})t(\mathbf{x}) + A(1 - t(\mathbf{x})), \qquad (1)$$

where $\mathbf{x}$ is a pixel, J is the haze-free image (scene radiance), A is the global airlight, and $t(\mathbf{x}) = e^{-\beta d(\mathbf{x})}$ is the transmission which represents color or light attenuation due to the scattering medium. The level of attenuation is determined by
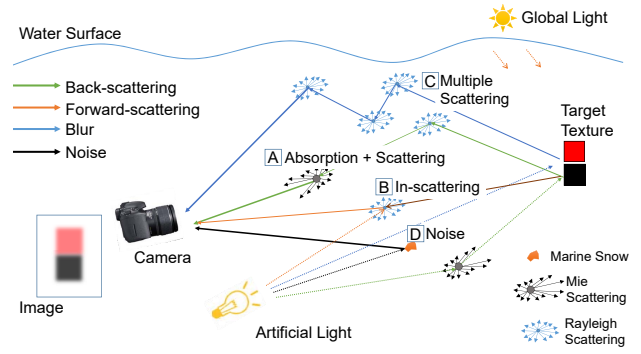

Fig. 2. Underwater scattering and noise model. Underwater image undergoes complex scattering model. Forward scattering, back scattering, multiple scattering and marine snow (noise) are described.

the scene depth $d(\mathbf{x})$ and the attenuation coefficient β which indicates thickness of haze.

A. Underwater Image Degradation

This SSAM performs well in aerial images with the homogeneous medium where the particles are small and identical. However, for underwater images, we found that simple atmospheric model (1) is not enough to deal with all degraded phenomena of underwater images. This addresses needs for thorough model using light transfer characteristics in underwater from particle physics [15]. As in Fig. 2, light photon characteristics change and join by irregularities in the heterogeneous medium with different sizes of the particles. Because mixtures of single photon characteristics lead to various degraded effects in underwater imaging, we classify degraded phenomena into four categories, absorption with scattering (green arrows, 'A'), in-scattering (orange arrows, 'B'), multiple scattering (blue arrows, 'C') and noise (black arrows, 'D') as described in Fig. 2. Absorption with scattering is related to back-scattering when light attenuation is based on both scattering and absorption. In turbid medium under artificial light source, in-scattering (or forward scattering) is mainly caused by luminance from the direct light. Multiple scattering occurs as a blurring effect when lights from the original target pixel (black square) is corrupted by the neighboring pixel (red square). The last phenomenon is noise. With large particle size, light reflects directly to the imaging sensor, causing a random noise, which is often called marine snow.

B. Underwater Image Model

The proposed image model in this paper is based on the underwater model in [5] and we adopt the light attenuation part as an exponentially decaying term as in [15]. The key differences of this model are three-fold, (*i*) non-uniform artificial light estimation, (*ii*) blur modeling with Point Spread Function (PSF) and (*iii*) random noises from underwater particles. Encapsulating the aforementioned properties in underwater imaging, we construct the unified model for underwater images as

$$I(\mathbf{x}) = \underbrace{\overbrace{[(E_A + E_L(\mathbf{x})) \cdot \rho(\mathbf{x})]}^{\text{Reflected Radiance}} * \overbrace{h_{\text{psf}}(\mathbf{x})}^{\text{blur}} \cdot e^{-\beta d(\mathbf{x})}}_{\text{Back Scattering}}$$

$$+ \underbrace{(1 - e^{-\beta d(\mathbf{x})}) \cdot E_L(\mathbf{x})}_{\text{Forward Scattering}} + \underbrace{N}_{\text{Noise}}$$

$$= J(\mathbf{x}) * h_{\text{psf}}(\mathbf{x}) \cdot e^{-\beta d(\mathbf{x})} + (1 - e^{-\beta d(\mathbf{x})}) \cdot E_L(\mathbf{x}) + N, \quad (2)$$

where E_A is the global veiling light bias from the water surface and is assumed to be constant. E_L is the light from artificial light source. The Linear sum of these two light sources are then scaled by reflectance of the surface $\rho(\mathbf{x})$. For simple derivation, we express reflected radiance as $J(\mathbf{x})$. Then, $J(\mathbf{x})$ is corrupted by a blur effect of multiple scattering ('C'). As blur is modeled as a convolution to a kernel function in the image space, we model PSF kernel as h_{psf} to model this multiple scattering. Overall, forward scattering component (E_L) and back scattering component ($J(\mathbf{x}) * h_{\text{psf}}$) are summed using the transmission $e^{-\beta d(\mathbf{x})}$ which represents light attenuation with the coefficient β. The additive term is randomly generated noise ('D') which is represented as N.

The proposed method begins with understanding of underwater image degradation. To resolve the hazed and blurred image, we present a step-by-step procedure to recover the original image. The detailed process is described in the following section.

III. ONLINE IMAGE ENHANCEMENT

The ultimate objective of image enhancement is to improve the recognition ability for successful camera measurements and loop-closures. As described in Fig. 3, we propose two approaches to handle both temporal and non-temporal camera measurements used in visual SLAM. The first approach is to use the full image model described in (2) for non-temporal measurements with small image overlap. To extract a camera measurement form a nom-temporal pair, images need strong enhancement. In this case, we propose the full model based approach which skims each degradation property step by step. For the temporal measurements, image overlap is sufficiently large with fairly accurate navigation prior. Since all images belonging to sequential motion must be processed in real-time, faster image processing is required. For these temporal links, we propose the simplified model which is based on unsupervised image enhancement.
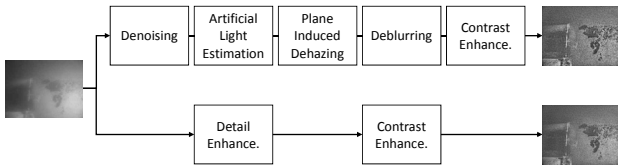
Fig. 3. Flow diagram of the full enhancement model for non-temporal links (top) and simplified model for temporal links (bottom).
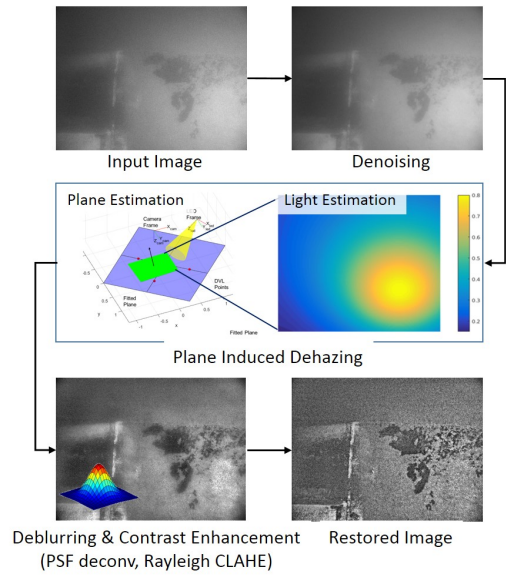
Fig. 4. Visibility enhancement using full degradation model. A step-by-step procedure skims noise and scattering effect to enhance image detail.

A. Full Enhancement Model

We introduce visibility enhancement based on (2) for non-temporal loop-closures (Fig. 4). This procedure is composed of four steps. With simple expression of $\mathbf{J}(\mathbf{x})$, we can rewrite (2) as

$$I(\mathbf{x}) = \underbrace{J(\mathbf{x}) * h_{\text{psf}}(\mathbf{x}) \cdot e^{-\beta d(\mathbf{x})}}_{\text{Back Scattering}} + \underbrace{(1 - e^{-\beta d(\mathbf{x})}) \cdot E_L(\mathbf{x})}_{\text{Forward Scattering}} + \underbrace{N}_{\text{Noise}}. \quad (3)$$

The full model based enhancement starts with removing noise N via initial denoising. Then, artificial light term $E_L(\mathbf{x})$ is estimated, assuming E_A to be constant. The transmission map is estimated using a planar surface assumption. Note that previous method such as [6] is applicable for the effective transmission map, but we exploit the surface condition for fast transmission estimation. Then, PSF is modeled to remove the blur effect. The estimated parameters are then used to solve for the enhanced image followed by contrast refinement with Contrast-limited Adaptive Histogram Equalization (CLAHE) with Rayleigh distribution.

1) Denoising: Uneven noise from the particles is corrupted in both scene radiance and veiling light term. Therefore we first apply guided filtering [13] as an initial denoising step. Guided filtering is an edge-preserving smoothing filter and can be expressed as below

$$\begin{aligned} I_{\text{de}}(\mathbf{x}) &= \text{GIF}_{k,\sigma}(I(\mathbf{x})) \\ &= \underbrace{J(\mathbf{x}) * h_{\text{psf}}(\mathbf{x}) \cdot e^{-\beta d(\mathbf{x})}}_{\text{Back Scattering}} + \underbrace{(1 - e^{-\beta d(\mathbf{x})}) \cdot E_L(\mathbf{x})}_{\text{Forward Scattering}}, \quad (4) \end{aligned}$$

where $\text{GIF}_{k,\sigma}(\cdot)$ refers guided filtering with kernel size $k = 4$ and smoothing level $\sigma = 0.1^2$.
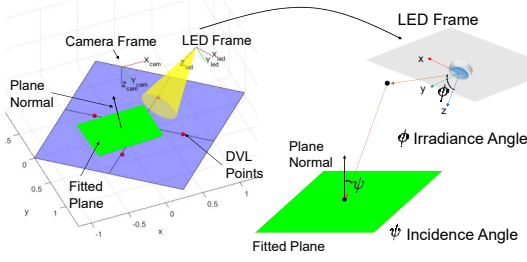
Fig. 5. Artificial veiling light estimation via luminous intensity estimation based on Lambertian radiation model. Estimated plane surface (blue plane) from DVL points (red dots) and fitted image plane (green plane).

2) Artificial Light Estimation: Underwater robots often carry their own light source. Because ambient light from the water surface decays rapidly, estimating this artificial light effect is critical. However, artificial light causes vignetting and biased veiling light on underwater image. Due to the effects of scattering and absorption of rays from the light source, pixels on image boundaries far from the light show lower intensities than vignetting center.

In [7], the author simulated Light-emitting diode (LED) light transfer model using designed LED lighting module. The basic characteristic of LED follows Lambertian cosine law proportional to the inverse of squared depth as depicted in Fig. 5. The luminance of the light source U_0 reaches the imaging surface depending on the irradiance angle (ϕ) and incidence angle (ψ). The luminous intensity at point $\mathbf{x}$ is represented as

$$L(\mathbf{x}) = \frac{U(\phi(\mathbf{x}), \mathbf{x}) \cdot e^{-\beta l(\mathbf{x})}}{l(\mathbf{x})^2 \cdot \cos(\psi(\mathbf{x}))} = \frac{U_0 \cos^m \phi(\mathbf{x}) \cdot e^{-\beta l(\mathbf{x})}}{l(\mathbf{x})^2 \cdot \cos(\psi(\mathbf{x}))},$$
(5)

where $l(\mathbf{x})$ is the range from an image surface to the light source, and $m = -\ln 2 / \ln(\cos 0.5\phi)$ indicates the order of Lambertian emission. The unit of $L(\mathbf{x})$ is the energy flux per a solid angle and $L(\mathbf{x})$ represents brightness with degradation by the scattering effect. The luminous intensity $L(\mathbf{x})$ is estimated by the predefined parameter U_0 which determined by the sensor specification and geometric parameters l, ϕ, ψ which estimated by the relative pose of the vehicle and the fitted plane. Then, $L(\mathbf{x})$ is fitted in the pixel value range $[\min(I), \max(I)]$ because the unit of the luminous intensity $L(\mathbf{x})$ does not match with image pixel value. By doing so, the restored image can be prevented from being over-saturated.

3) Plane Induced Dehazing: Most of conventional dehazing methods rely on the color channel value or intensity to predict scene depth. These methods restore clean images under the assumption that color or intensity level decays along a ray travels, therefore, underwater scene with various depth points are proper for these methods. However, this assumption is not suitable for grayscale images or for planar scenes with less variation of scene depth. Note that we can apply online depth estimation, as in [6], for scene depth estimation. Instead, we introduce plane induced dehazing aiming planar surface inspection scenario.

In the tested dataset, four Doppler velocity log (DVL)

beam points are available for plane fitting and scene depth measurements. For other types of inspection or seafloor mapping, similar techniques can be used. Using computed plane normal and known camera calibration, the scale parameter determines the depth map $d(\mathbf{x})$. As described in Section II, underwater attenuation coefficient includes scattering and absorption. In this paper, we use $\beta = 0.7$ for transmission.

Now we can estimate the first part of the image model in (4). For simple expression, we rewrite (4) to compute $J(\mathbf{x}) * h_{\text{psf}}$ as

$$J_{\text{plane}} = J(\mathbf{x}) * h_{\text{psf}} = \frac{I_{\text{de}} - E_L(\mathbf{x})}{e^{-\beta d(\mathbf{x})}} + E_L(\mathbf{x}),$$
(6)

where I_{de} is the denoised image from (4), and $E_L(\mathbf{x})$ is the estimated artificial light.

4) Deblurring: PSF $h_{\text{psf}}(\mathbf{x})$ represents blurring effect of ray by particles in turbid medium. In this paper, we use PSF based on generalized Gaussian distribution [24] and apply deconvolution on J_{plane} for deblurring. The mathematical model of h_{psf} is

$$h_{\text{psf}}(u, v, \eta, p, \sigma) = \frac{e^{-(\frac{u^2 + v^2}{A(p,\sigma)})^{0.5p}}}{2\Gamma(l + 1/p)A(p,\sigma)}, \ (u, v) \in \mathbb{R}^2 \quad (7)$$

where $p(T) = \sqrt{T}$, $A(p, \sigma) = \sigma(\Gamma(\frac{1}{p})/\Gamma(\frac{3}{p}))^{0.5}$, $\sigma = 0.5/\sqrt{q}$ and $\Gamma(\cdot)$ is the Gamma function. In [24], T refers optical thickness which means the level of turbidity. Also, the term q represents forward scattering parameter. These two parameters are selected according to thickness of turbidity. For hull inspection in dense underwater haze environment, we set the parameter as $T = 4$ and $q = 0.9$. Using the estimated PSF, the deblurred image can be obtain by deconvolution as $J_{\text{deblur}} = \text{DECONV}(J_{\text{plane}}, h_{\text{psf}})$.

5) Global Contrast Enhancement with CLAHE: To restore the final enhanced image, we apply CLAHE with Rayleigh distribution as global contrast enhancement method. In spite of estimating artificial light source estimation, salient regions such as edges of the image suffer from degraded contrast. Since haze images have been sufficiently improved in previous stages, we use CLAHE as the enhancement method instead of direct computation of added light term $(E_A + E_L(\mathbf{x}))$. This approach results in the final balanced scene radiance as $J_{\text{full}} = \text{CLAHE}(J_{\text{deblur}})$.

B. Simplified Enhancement Model

For a temporal image pair with large overlap and small view point changes, we propose a simplified enhancement model. We found that simplification of the aforementioned model can be expressed as two parts, background light with noise and contrast degradation.

We first compute the detail layer $\delta \mathbf{e}$ from foreground image (I_f) and background image (I_g). We apply dual guided image filtering method (GIF) by adopting detail enhancement [13] as

$$\Delta \mathbf{e} = I_f - I_g = \text{GIF}_{k_f, \sigma_f}(I) - \text{GIF}_{k_g, \sigma_g}(I),$$
(8)

where kernel sizes ($k_f = 2, k_g = 8$) and smoothing degrees ($\sigma_f = 0.1^2, \sigma_g = 0.4^2$) were chosen separately. With

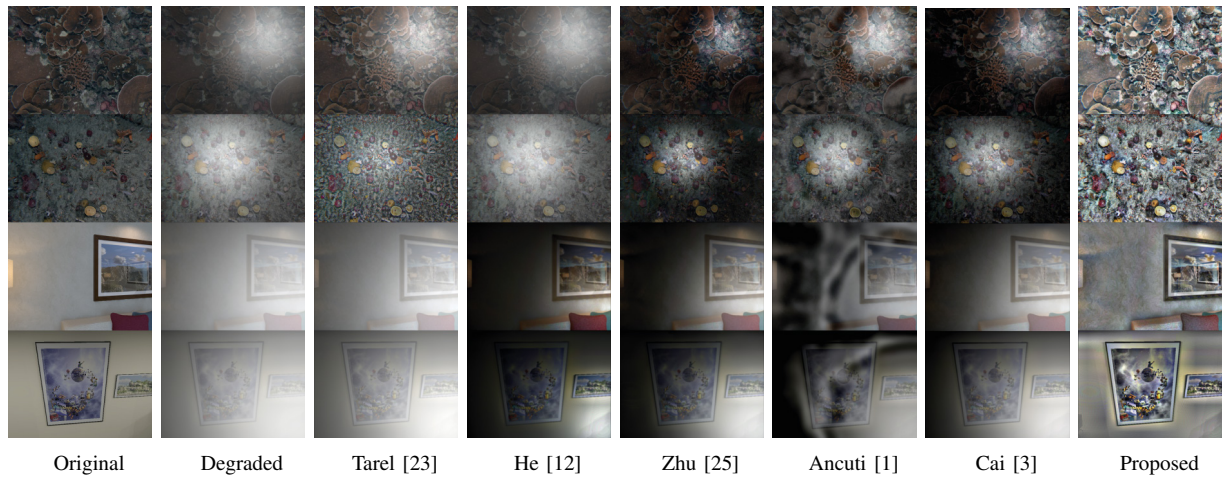| Original | Degraded | Tarel [23] | He [12] | Zhu [25] | Ancuti [1] | Cai [3] | Proposed |

Fig. 6. Image enhancement result under synthetic haze. From test datasets, selected `ACFR` (1st and 3rd row) and `ICL-NUIM` (4th and 5th row) are shown for the qualitative comparison.

denoised foreground and background images are obtained, final enhanced image then can be calculated as

$$I_{\text{simple}} = I_g + \alpha \cdot \Delta \mathbf{e}. \tag{9}$$

where α is the amplification factor which controls the level of details, and we use $\alpha = 5$ in the paper. Similar to the full model image enhancement, CLAHE is applied for contrast enhancement. Despite that this simplified model cannot fully enhance whole region in the image, the restored image is sufficiently improved for sequential image matching when a pair is with large overlap.

IV. EXPERIMENTAL RESULTS

We evaluate the proposed method by focusing on both (*i*) image enhancement and (*ii*) geometric vision application (feature matching and visual SLAM). First, we present both qualitative and quantitative comparison results to other conventional dehazing methods. In this comparison, we use MATLAB implementation for fair comparison. In the geometric vision application evaluation, we present C++ implementation that supports online performance, and evaluate the trajectory estimation improvement from using additional loop closures via enhanced vision. See also `https://youtu.be/GFgMdswWHAA` for these dehazing results.

A. Synthetic Degraded Color Images

First, we conduct an experiment using a synthetic dataset with ground truth. Because ground truth for underwater images in a turbid medium is extremely rare, we use publicly available datasets, `ICL-NUIM` RGBD benchmark dataset [11] and `ACFR` stereo datasets [10] to generate synthetic degraded images. The RGBD benchmark dataset offers a full depth map with and image resolution of 640×480 so that we can synthetically corrupt original images with various levels of a turbid medium. Images from `ACFR` datasets offer stereo disparity maps that indicate relative depths of images

resolution of 1360×1024. We apply a synthetic haze on haze-free images by (2) using a transmission map computed from a true depth map. We select test images that have a near planar surface with various colors as in Fig. 6. Ground truth provided from the dataset is used for evaluation in the comparison results to the other conventional approaches as in [22], [12], [1], [25], and [3].

The comparison result to other methods is given in Fig. 6. Zhu's and He's methods [25, 12] show intensity degradation on the left side of the images. This is because these methods strongly rely on dehazing priors, which is based on the assumption that depth of scene is proportional to color changes [12] or differences [25] among channels. Also, Zhu's [25] and He's [12] methods are based on the simple dehazing model (1); therefore the two methods cannot deal with the effect of an artificial light source. Despite the fact that Tarel's method [23] is also based on the simplified image model (1), the intensity distortion is not as large as with the previous two method's results. This is due to the fact that Tarel's method [23] has a regularization term for black and white patches with intensity correction by gamma correction. However, the result of Tarel's method [23] shows that the contrast is slightly enhanced, but back-scattering effects remain. Testing Ancuti's method [1] shows distorted results using five pyramids for fusion. The model shows weakness on highly dense haze regions with low intensity. Cai's method [3] also shows similar results as He's result due to the artificial light effect.

In comparison with conventional methods [25, 12, 23, 1, 3], our method presents better results, especially for contrast enhancement with edge preserving. The restored color seems a little distorted due to forcing the intensity balance by CLAHE in the full model. However, from geometric vision point of view, edge preserving with enhanced contrast will lead to accurate camera measurement extraction with haze-corrupted images.

For the datasets with degraded images, we conduct a quantitative analysis as in Table. I. We consider metrics such
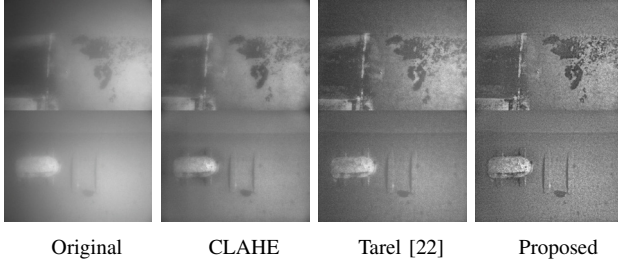
| Original | CLAHE | Tarel [22] | Proposed |

Fig. 7. Comparison result on real underwater gray images. Original images are corrupted by haze, blur and noise. Compared to CLAHE and Tarel's method [22], our method shows contrast enhanced and edge restored results.

TABLE I

ENHANCEMENT COMPARISON (AVERAGE) RESULT WITH EVALUATION METRICS (HIGHER VALUE INDICATES BETTER PERFORMANCE). SEE ALSO FIG. 6 FOR QUALITATIVE COMPARISON.

	Tarel	He	Zhu	Ancuti	Cai	Proposed
SSIM	0.6589	0.3611	0.42	0.45	0.4465	0.5846
PSNR	21.729	11.08	12.10	16.58	13.56	17.85
CNR	284.95	121.64	152.00	221.57	190.09	294.88
e	0.409	0.485	0.431	0.1345	0.491	0.564
$\bar{r}$	2.32	1.24	1.28	1.60	1.32	4.29
m	0.058	0.0095	0.0169	0.0349	0.0229	0.0795
time (sec) 640×480	8.64	34.45	1.66	1.34	2.75	0.85
time (sec) 1360×1024	158.11	94.34	6.39	2.75	7.38	2.87

as Structural Similarity Index (SSIM) [25], Peak Signal-to-Noise Ratio (PSNR) [8], Contrast-to-Noise Ratio (CNR) [8], visible edge indicators (e, $\bar{r}$, m) [23, 21]. For image quality evaluation, we chose SSIM, PSNR, and CNR, whereas SSIM measures color (intensity) similarity between a ground truth and a result, and CNR analyzes the contrast level that affects the visibility of objects. Additionally, we add three more metrics e, $\bar{r}$, and m for the edge and gradient level. For all metrics, higher values indicates better performance. As can be seen in Table. I, our method shows improved scores on contrast and edge related metrics (i.e., CNR, e, $\bar{r}$, and m). For SSIM and PSNR, Tarel's method [23] shows better scores because our method have the contrast enhancement step. The result of our method may not be as natural as ground truth images. However, as in Table. I, our method shows outperforming results on visible edge (e, $\bar{r}$, m) and computation time which strongly related to visual SLAM.

B. Underwater Gray Images and Geometric Vision Test

Next, we validate the proposed method with real underwater gray images in the application of visual SLAM. The dataset used in the validation is from the University of Michigan during its in-water hull inspection projects [14]. The dataset consists of two dives with grayscale underwater images and navigational sensor data. Two dives are performed following the same lawn-mower pattern over the same area but at different times of the day. Due to the tidal changes, images in dive2 (afternoon) are highly corrupted, whereas images are clear in dive1 (morning). The sample images are as in Fig. 1. We include the proposed
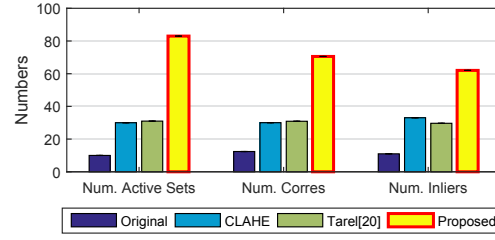

Fig. 8. Comparison result on feature matching. Num. Active sets mean meaningful image pairs which have more 10 correspondences. Num. Corres and Num. Inliers mean average of correspondence points and homography inliers in the active sets.

TABLE II

TEMPORAL AND NON-TEMPORAL CAMERA MEASUREMENTS

	Original	Enhanced
Temporal links	30	68 (simplified)
Non-temporal links	7	63 (full)

enhancement algorithm in the existing SLAM pipeline [17], and test the performance using the real underwater robotic application.

Similar to the synthetic dataset, we compare the image enhancement results in comparison to conventional methods. Since the images have only one channel, we exclude color-oriented algorithms and compare them only with CLAHE and Tarel's method[1] [22]. Our method presents improvements compared to the results of CLAHE and Tarel [22] (Fig. 7) for full resolution images (1360 × 1024). Both CLAHE and Tarel's method [22] show limited improvements. CLAHE shows only balancing intensity; the details of textures are still degraded.

We verify the performance of image enhancement via feature matching. 1103 matching pairs are used from the Michigan dataset, and the performance is compared using three comparisons as shown in Fig. 8. The number of active sets refers to image pairs with more than 10 correspondence points, and the number of correspondences and inliers represents the average correspondences and homography inliers within the active sets. As described in Fig. 8, when the proposed method (yellow) is applied, it can be confirmed that the performance is more than twice as high as that of CLAHE (blue) or Tarel (green) which is frequently used in the conventional monochrome image.

We further evaluate the proposed method within the SLAM framework. We build a SLAM graph using only odometry and camera measurements while applying the visibility enhancement in the camera registration engine. The pose graph is optimized using incremental smoothing and mapping (iSAM) library [16] with the conventional 5 degree of freedom (DOF) camera measurements as used in [17]. Tarel's method [22] is not evaluated by the SLAM framework because the computation speed is slower than the camera framerate (2 Hz). The processing time of the proposed method (implementation on C++) supports about

[1]For the grayscale image set, [22] showed better performance than [23] and was chosen in the analysis.

(a) SLAM with original images
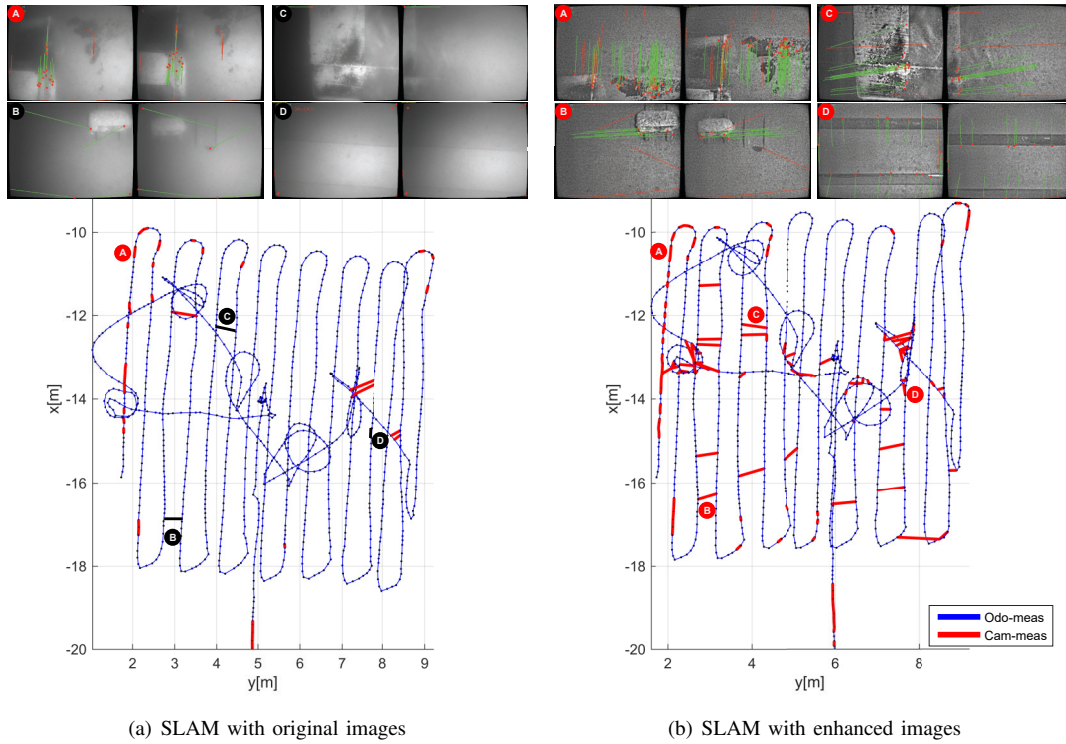

(b) SLAM with enhanced images

Fig. 9. SLAM results with original degraded images versus enhanced images. Each graph is composed of pose nodes (black dots), odometry measurements (blue lines) and camera measurements (red lines). Both graphs are displayed with sampled camera proposal links. Red circles are successfully extracted camera measurements while black circles are not. In sampled figures, red dots refer keypoints, green and red lines mean inliers and outliers. Unit of x, y axis is meter, and the trajectory is represented as top-view with z-axis down.

(a) Trajectory Comparison

(b) Time elevation

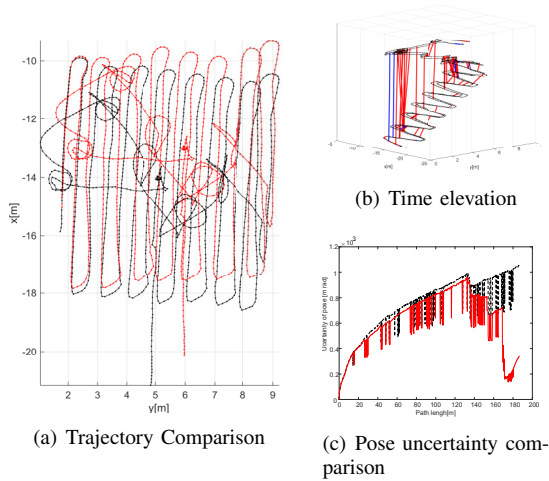(c) Pose uncertainty comparison

Fig. 10. SLAM results analysis. (a) Trajectory of proposed method (red lines) and original method (black lines) are plotted. (b) Time elevation graph for loop-closure visualization. The xy component represents trajectory in meter scale and z-axis represents mission time. Blue lines and red lines indicate camera measurements from original and enhanced images. (c) Uncertainty of pose with respect to the mission path length is expressed as covariance determinant ($\sqrt[6]{\Sigma}$ in [m · rad] [18]). Pose uncertainties using proposed method is better bounded (red lines) compared to SLAM without visibility enhancement (black dashed lines).

10 Hz for the simplified model and 3 Hz for the full model. Table. II summarizes the number of measurements made by each enhancement model.

The performance with the SLAM framework is shown in Fig. 9. We compare the result of our method in Fig. 9(b) to the SLAM result with the original images Fig. 9(a). With same number of nodes (1228) and odometry measurements (1227), our method extracts 131 camera measurements versus only 37 measurements for the original images. Comparison with sampled matching image pairs clearly illustrates improvement from using our method. Our method demonstrates improved matching for temporal measurements with both feature-rich and featureless image pairs (A and D in Fig. 9). Stronger improvement can be found when matching cross-track non-temporal pairs (B and C in Fig. 9).

We compare trajectories between the original SLAM to the visually enhanced one as in Fig. 10(a). Pose drift from odometry measurements is compensated by camera measurements (red lines). However, trajectory with original images is biased and drifted due to the lack of camera measurements. To better depict non-temporal camera measurements, we compare a time elevation graph by plotting the mission time on the z-axis as in Fig. 10(b). Vertical lines are from non-temporal measurements, and the enhanced images (red lines) produce more loop closures than original images (blue lines). Through the proposed image enhancement method, camera measurements are well extracted not only on lateral cross-links but also on loop closure links. Because we do not have a ground truth for this dataset, we evaluate the SLAM performance with the uncertainty plot as in Fig. 10(c). The uncertainty of the proposed method is well bounded by the newly added cross track measurements.

V. Conclusion

This paper proposes an end-to-end pipeline for online image visibility enhancement specifically targeting underwater visual SLAM. We present a method with the model that considers particle physics and applies to both synthetic and real-world images. We validate that the algorithm performs equally well regardless of the channel numbers with substantially reduced process time.

Acknowledgment

References

[1] C. Ancuti, C. O. Ancuti, T. Haber, and P. Bekaert. Enhancing underwater images and videos by fusion. In *Proc. IEEE Conf. on Comput. Vision and Pattern Recog.*, pages 81–88. IEEE, 2012.

[2] M. Babaee and S. Negahdaripour. Improved range estimation and underwater image enhancement under turbidity by opti-acoustic stereo imaging. In *Proc. IEEE/MTS OCEANS Conf. and Exhibit.*, pages 1–7. IEEE, 2015.

[3] B. Cai, X. Xu, C Jia, K.and Qing, and D. Tao. Dehazenet: An end-to-end system for single image haze removal. *IEEE Trans. Image Processing*, 25(11):5187–5198, 2016.

[4] N. Carlevaris-Bianco, A. Mohan, and R. M. Eustice. Initial results in underwater single image dehazing. In *Proc. IEEE/MTS OCEANS Conf. and Exhibit.*, pages 1–8. IEEE, 2010.

[5] J. Y. Chiang and Y. Chen. Underwater image enhancement by wavelength compensation and dehazing. *IEEE Trans. Image Processing*, 21(4):1756–1769, 2012.

[6] Y. Cho, Y. Shin, and A. Kim. Online depth estimation and application to underwater image dehazing. In *Proc. IEEE/MTS OCEANS Conf. and Exhibit.*, pages 1–7. IEEE, 2016.

[7] A. Davis. Light emitting diode source modeling for optical design. *Reflexite Display Optics*, 2004.

[8] C. Deng, L. Ma, W. Lin, and K. Ngan. *Visual Signal Quality Assessment: Quality of Experience (QoE)*. Springer Publishing Company, Incorporated, 2014.

[9] R. Fattal. Single image dehazing. *ACM Trans. on Graphics (TOG)*, 27(3):72, 2008.

[10] Australian Centre for Field Robotics. Marine robotics datasets. marine.acfr.usyd.edu.au/datasets/, 2009.

[11] A. Handa, T. Whelan, J.B. McDonald, and A.J. Davison. A benchmark for RGB-D visual odometry, 3D reconstruction and SLAM. In *Proc. IEEE Intl. Conf. on Robot. and Automat.*, pages 1524–1531, Hong Kong, China, May 2014.

[12] K. He, J. Sun, and X. Tang. Single image haze removal using dark channel prior. *IEEE Trans. Pattern Analysis and Machine Intell.*, 33(12):2341–2353, 2011.

[13] K. He, J. Sun, and X. Tang. Guided image filtering. *IEEE Trans. Pattern Analysis and Machine Intell.*, 35 (6):1397–1409, 2013.

[14] F. Hover, R. Eustice, A. Kim, B.Englot, H. Johannsson, M. Kaess, and J. Leonard. Advanced perception, navigation and planning for autonomous in-water ship hull inspection. *Intl. J. of Robot. Research*, 31(12): 1445–1464, October 2012.

[15] M. Jonasz and G. Fournier. *Light scattering by particles in water: theoretical and experimental foundations*. Academic Press, 2011.

[16] M. Kaess, H. Johannsson, and J. Leonard. Open source implementation of iSAM. people.csail.mit.edu/kaess/isam, 2010.

[17] A. Kim and R. M. Eustice. Real-time visual SLAM for autonomous underwater hull inspection using visual saliency. *IEEE Trans. Robot.*, 29(3):719–733, 2013.

[18] A. Kim and R. M. Eustice. Active visual SLAM for robotic area coverage: Theory and experiment. *Intl. J. of Robot. Research*, 34(4-5):457–475, 2015.

[19] J. Li, R. M. Eustice, and M. Johnson-Roberson. Underwater robot visual place recognition in the presence of dramatic appearance change. In *Proc. IEEE/MTS OCEANS Conf. and Exhibit.*, Washington, D.C., USA, October 2015.

[20] D.M Parry, M.A Kendall, D.A Pilgrim, and M.B Jones. Identification of patch structure within marine benthic landscapes using a remotely operated vehicle. *J. of Exp. Mar. Biology and Ecology*, 285-286(0):497–511, 2003.

[21] I. Shim, J. Lee, and I. S. Kweon. Auto-adjusting camera exposure for outdoor robotics using gradient information. In *Proc. IEEE/RSJ Intl. Conf. on Intell. Robots and Sys.*, pages 1011–1017. IEEE, 2014.

[22] J. Tarel and N. Hautiere. Fast visibility restoration from a single color or gray level image. In *Proc. IEEE Intl. Conf. on Comput. Vision*, pages 2201–2208. IEEE, 2009.

[23] J. Tarel, N. Hautiere, A. Cord, D. Gruyer, and H. Halmaoui. Improved visibility of road scene images under heterogeneous fog. In *Proc. IEEE Intell. Vehicle Symposium*, pages 478–485. IEEE, 2010.

[24] R. Wang, R. Li, and H. Sun. Haze removal based on multiple scattering model with superpixel algorithm. *Signal Processing*, 127:24–36, 2016.

[25] Q. Zhu, J. Mai, and L. Shao. A fast single image haze removal algorithm using color attenuation prior. *IEEE Trans. Image Processing*, 24(11):3522–3533, 2015.