Analysis of Object Distance Accuracy on Underwater Image Color Restoration

Kuan, H.W.[1*] and Jaw, J.J.[2]

[1]Master Student, Department of Civil Engineering, National Taiwan University, Taiwan

[2]Associate Professor, Department of Civil Engineering, National Taiwan University, Taiwan

*garykuan45@gmail.com

ABSTRACT

Underwater images often exhibit color distortion compared to those captured in air, primarily due to backscatter and light attenuation. Such distortion creates inconveniences in many fields, including engineering and archaeology. Therefore, algorithms based on physical models are often used for underwater image color restoration, which require accurate parameter estimation. The object distance between the scene and the camera is also crucial. However, current research lacks an analysis of the accuracy requirements for object distance. Therefore, this study aims to analyze the effect of object distance accuracy on underwater color restoration methods based on physical models, specifically the Sea-thru method. Using photos taken in the air as ground truth images and simulating underwater scenes, we introduce different errors in object distances to examine their effect on color restoration performance using the full-reference metric SSIM. The results of this research provide a reference for the minimum accuracy requirements of object distance to achieve SSIM of at least 0.8, considered a visual satisfaction level, in future research on underwater image color restoration.

Keywords: underwater image color restoration, Sea-thru method, object distance

Introduction

With the advancement of human technology, underwater environments are receiving increasing attention, as the energy generated from the sea and rivers is clean, sustainable, and a viable alternative to fossil fuels. Additionally, the underwater ecosystem plays a crucial role in maintaining environmental stability, with its significant potential for carbon storage and other benefits. Activities such as offshore wind power installation, submarine cable repairs, and coral reef conservation often require capturing images of specific areas, such as parts of facilities or corals, for further analysis. However, underwater images typically suffer from direct attenuation and backscatter, leading to color distortion and causing the images to

appear with a bluish or greenish tint. There are several approaches to addressing this problem, including deep learning-based methods, non-physical model-based methods, and physical model-based methods(Wang et al., 2024). Deep learning-based methods have gained popularity recently due to their convenience and efficiency; simply input an underwater image, and the model generates a restored image almost instantly. However, these methods are constrained by the availability of underwater images and corresponding reference images for training the model. Additionally, their "black-box" nature is often criticized for lacking transparency. Non-physical model-based methods, in contrast, focus on adjusting the image's color and contrast. While these methods allow users to enhance images without requiring additional on-site information, they rely heavily on experience to set parameters, which reduces the consistency of their performance. Physical model-based methods offer a clearer path for restoring underwater images by tracing and reversing the imaging processes to recover the original scene. However, the accuracy of these methods depends heavily on obtaining precise parameters; without them, the restoration process is likely to fail. Given that light attenuation is significantly influenced by the distance between the camera and the scene, object distance is one of the most crucial parameters in image restoration. However, no research has yet explored how precise the object distance must be to achieve satisfactory results.

This study aimed to determine the minimum required accuracy of object distance for the Sea-thru method (Akkaynak & Treibitz, 2019) , one of the state-of-the-art physical model-based methods. The process involved utilizing Agisoft Metashape Pro software to construct dense point cloud of the scene in air, then generated a depth map for each image. Using the depth map and corresponding images, simulated underwater images were created under different conditions. Finally, the Sea-thru method was applied to restore the images, and the performance is evaluated.

Literature Review

1. The Sea-thru method

The underwater image formation model can be expressed as Equation 1:

$$I_c = D_c + B_c \tag{1}$$

As a physical model-based method, the Sea-thru method (Akkaynak & Treibitz, 2019) divides the energy collected by the camera, I_c, into two components based on their sources:

D_c and B_c. As shown in Figure 1, D_c represents the direct signal, which is the attenuated light reflected by the scene, while B_c represents the backscatter, an additive signal caused by light reflected from particles suspended in the water.

Equation 1 can be further extended into Equation 2:

$$I_c = J_c e^{-\beta_c^D * z} + B_c^{\infty}\left(1 - e^{-\beta_c^B * z}\right) \quad (2)$$

Where z is object distance between camera and the scene, J_c is the unattenuated signal from the scene, B_c^{∞} is the veiling light, β_c^D is the wideband direct attenuation coefficient, and β_c^B is the wideband backscatter attenuation coefficient. As shown in Figure 2, it is clear that object distance z can significantly impact image formation, as it is a parameter in the exponential term.

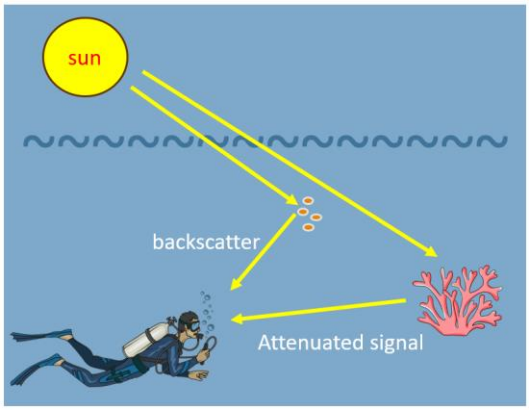
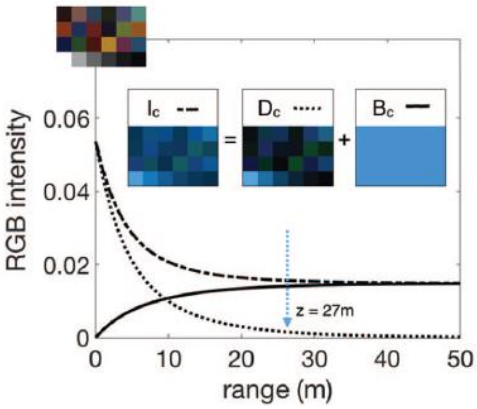


Figure 1: underwater image formation diagram



Source: (Akkaynak & Treibitz, 2018)

Figure 2: underwater image signal along with object distance

2. Inherent Optical Properties

When light passes through a medium such as seawater, it can be absorbed or scattered. Absorption occurs when the energy is converted into another form, such as heat or energy stored in chemical bonds. Scattering happens when the direction of the light is altered. Both absorption and scattering are properties of the medium itself and do not depend on the light field, which is why they are referred to as Inherent Optical Properties (IOPs).

Oceanographer Nils G. Jerlov established a global classification system for near-surface sea water (Jerlov, 1977). He divided sea water into five open-ocean classes (I-III) and five coastal-ocean classes (1C-9C), ranging from clear to turbid, based on the diffuse downwelling attenuation coefficients he measured.

3. Evaluation metrics

There are many metrics used to evaluate the performance of underwater image enhancement or restoration, which can be broadly divided into three types. The first type involves subjective scoring using the Mean Opinion Score (MOS), where participants quantify performance based on their subjective evaluations (Guo et al., 2021) . Although MOS effectively reflects human visual perception, it is time-consuming and can be unstable. The second type is Full Reference Image Quality Assessment, which compares the enhanced or restored image to a reference image of the scene. Common metrics include Mean Square Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM) (Wang et al., 2004) . While these metrics are reliable due to the use of a reference image, obtaining a reference image in underwater scenarios can be challenging, and these metrics are often not specifically designed for underwater images. The third type is Non-Reference Image Quality Assessment, which evaluates performance based on attributes such as mean value and saturation of the enhanced or restored images, using metrics like UCIQE (Yang & Sowmya, 2015) and UIQM (Panetta et al., 2015). This method is praised for its convenience, as it does not require a reference image, but it is criticized for relying on manually adjusted weights to extract feature attributes, which limits its versatility.

Methodology

In this study, dense point cloud of the scene in air was generated by taking photos with a few scattered control points in the scene. Afterwards, total stations were used to measure the control points to obtain their object space coordinates, which were subsequently used to correct the dense point cloud. Next, depth maps were generated from the dense point cloud and served as the true values for underwater image simulations, although they were also

simplified as data with errors. Finally, using the simulated images and depth maps, image restorations were performed, and the performance was assessed using the SSIM evaluation metric. Please note that although the reference image was used to generate the underwater image, the Sea-thru method cannot produce an identical result.

1. Depth map generation

In the Sea-thru method, the depth map is one of the key parameters, providing object distance information for color restoration. In this study, Agisoft Metashape Pro software was used to generate the depth maps. Since it is difficult to obtain accurate object distances in real underwater scenes, the study first used a scene in air and then simulated its underwater images. The scene was a flat outdoor ground with 4 cartons to simulate the objects in the sea. To evaluate the restoration performance for different colors, color boards with a wide range of colors were placed in the scene. Furthermore, to ensure the accuracy of the model, 32 markers were sticked on the cartons for coordinate measurement. as shown in Figure 3. First, a Canon EOS 760D camera was used to take 140 photos of the scene to reconstruct it in the software. The camera specifications are listed in Table 1. The distribution of the camera positions was arranged around the center of the scene, as shown in Figure 4, where each blue square represents one camera position.

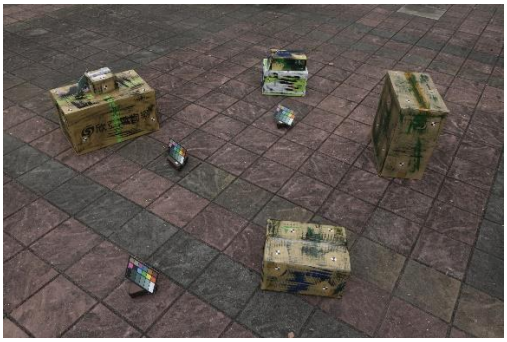


Figure 3: photo of the scene

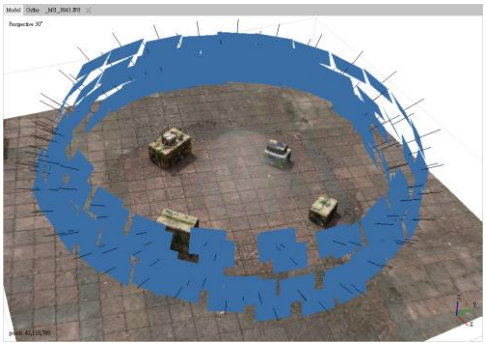


Figure 4: camera positions distribution

Table 1: Camera Specification

Category	Value
Model	Canon EOS 760D
CCD or CMOS	CMOS
CMOS Size	22.3 x 14.9 mm
Resolution	6000 x 4000 pixel
Principal distance	18 mm
Aperture diameter	F14.0

After taking photos of the scene, a total station was set up to measure the coordinates of the markers placed on the cartons. The total station was repositioned 4 times to ensure that each marker was measured at least twice. This redundancy helped to improve the accuracy of the measurements. The photos were then imported into Agisoft Metashape Pro for alignment. Using the coordinates of the markers, calculated through the least squares method, a high-accuracy dense point cloud of the scene was obtained, from which the depth map was generated.

2. Underwater image simulation

To simulate underwater images for the photos taken of the scene in air, this study applied the Sea-thru method's framework for simulation as well. This required data such as the sensor's spectral response, the scene's reflectance, and the diffuse downwelling attenuation coefficient(Solonenko & Mobley, 2015). Based on Jerlov's water type classification, eight water types were selected for simulation, ranging from clear to turbid: "Type I, Type IA, Type IB, Type II, Type III, Type 1C, Type 3C, and Type 5C." Additionally, two water depths were chosen for simulation: 1 meter and 5 meters, representing the depth from the water surface to the scene. As a result, each original image generated 16 simulated images. Although the same physical model was used for both underwater image simulation and color restoration, the algorithms employed were not identical. In underwater image simulation, the value of each pixel was calculated individually based on object distance and the environmental parameters mentioned above. In contrast, color restoration grouped pixels into different intervals according to object distance and directly accounted for backscatter and illumination, fitting the parameters to reverse the effects. This process did not involve solving for any environmental parameters.

First, backscatter required only object distance information and environmental parameter data, it can be derived from Equation 2 as Equation 3:

$$B_c = B_c^{\infty} * \left(1 - e^{-\beta_c^B * z}\right) \tag{3}$$

Where

$$B_c^{\infty}(\lambda) = \frac{b(\lambda)E(\lambda, d)}{\beta(\lambda)} \tag{4}$$

λ is the wavelength of visible light spectrum, $b(\lambda)$ is the beam scattering coefficient, $E(\lambda, d)$ is the ambient light at depth d, $\beta(\lambda)$ is the beam attenuation coefficient which equals to the sum of $b(\lambda)$ and beam absorption coefficient.

Then the wideband backscatter attenuation coefficient β_c^B can be expanded as Equation 5:

$$\beta_c^B = -\ln\left(1 - \frac{\int_{\lambda 1}^{\lambda 2} S_c(\lambda) B^{\infty}(\lambda)\left(1 - e^{-\beta(\lambda)z}\right) d\lambda}{\int_{\lambda 1}^{\lambda 2} B^{\infty}(\lambda)\, S_c(\lambda) d\lambda}\right) \Big/ z \tag{5}$$

$\lambda 1$ and $\lambda 2$ define the bounds of the integration over the spectrum.

Second, direct attenuation signal required object distance information and environmental parameter data and the original image, it can be derived from Equation 2 as Equation 6:

$$D_c = J_c e^{-\beta_c^D * z} \tag{6}$$

Where

$$\beta_c^D = \ln\left[\frac{\int_{\lambda 1}^{\lambda 2} S_c(\lambda)\rho(\lambda)E(\lambda)e^{-\beta(\lambda)z} d\lambda}{\int_{\lambda 1}^{\lambda 2} S_c(\lambda)\rho(\lambda)E(\lambda)e^{-\beta(\lambda)(z+\Delta z)} d\lambda}\right] \Big/ \Delta z \tag{7}$$

Finally, combined 2 images to get the simulation image.

3. Sea-thru method color restoration

To implement the Sea-thru method, several steps were required, including backscatter removal, depth map simplification, illumination map estimation, attenuation coefficient estimation, and image restoration. Each step involved estimating different coefficients in Equation 2, such as β_c^D and β_c^B. These coefficients are highly dependent on object distance, making the quality of the depth map crucial for accurate restoration. Eventually, to examine the impact of object distance accuracy on restoration performance, random errors were added

to the depth map with a mean value of 0 and five standard deviation intervals: [0.50, 0.30, 0.10, 0.05, 0.00] meters**.**

3.1 Backscatter removal

As shown in Equation 2, the signal received by the camera is composed of two parts: backscatter and direct attenuation. The direct attenuation contains information from the scene. This implies that if a pixel contains little to no information from the scene—meaning the scene is too dark to reflect light—then the pixel consists only of backscatter, which can be estimated and removed.

Given that backscatter is dependent on object distance, this study first partitioned the depth map evenly into 10 clusters, ranging from the minimum to the maximum depth values. Within each cluster, pixels whose RGB values fell in the bottom 1 percentile were identified and denoted as Ω. The signal for these pixels can be expressed by Equation 8:

$$\hat{B}_c = B_c^{\infty}\left(1 - e^{-\beta_c^B * z}\right) + J'_c e^{-\beta_c^{D'} * z} \tag{8}$$

$\hat{B}_c$ is the signal of Ω, J'_c is the very small residual term of the direct attenuation from the scene. This residual is retained to avoid interfering with subsequent analyses. We used non-linear least square fitting to estimate the parameters "$B_c^{\infty}, \beta_c^B, J'_c, \beta_c^{D'}$". In this step, the object distance-dependency of $\beta_c^{D'}$ was ignored. After parameter estimation, the backscatter values for each cluster were calculated.

3.2 Depth map simplification

Simplification was applied after removing backscatter by dividing the object distance by an interval value of 0.05 meters, a minimum value to visually distinguish differences in object distance and taking the quotient. Following this, the "Gray World Assumption" (Buchsbaum, 1980) was used. This assumption suggests that scenes with a diverse set of colors will have a mean color close to gray. If the mean color deviates from gray, it indicates an influence from the illuminant. In this study, attenuation was treated as an illuminant that altered the scene's mean color. The Local Space Average Color (LSAC) method (Ebner, 2009) was then used to calculate the mean color in each object distance interval, as attenuation is dependent on object distance.

3.3 Illumination map estimation

As mentioned in the previous section, this study used the LSAC method to compute the mean colors for each interval, forming the illumination map of the scene. First, the local space average color a_c of each interval was calculated iteratively using the following equations:

$$a'_c(x,y) = \frac{1}{N_e}\sum_{N_e} a_c(x',y') \tag{9}$$

$$a_c(x,y) = D_c(x,y)p + a'_c(x,y)(1-p) \tag{10}$$

Here, N_e represents the number of pixels in the same interval, D_c is the direct attenuation signal received by the camera, and p controls the iteration speed, with larger p resulting in slower convergence and smaller p leading to faster convergence. The initial value of a_c was set to 0.

After calculating the local space average color a_c for each interval, the illumination map $\hat{E}_c$ was found using Equation 11:

$$\hat{E}_c = f a_c \tag{11}$$

Where f is a factor based on geometry condition, scaling all color channels equally. Here we set $f = 2$, assuming a perpendicular direction between scene and the camera.

3.4 Attenuation coefficient estimation

Illumination map was regarded as the relationship between direct attenuation and the original signal:

$$\hat{E}_c = e^{-\hat{\beta}_c^D * z} \tag{12}$$

Estimate attenuation coefficient $\hat{\beta}_c^D$ can be derived from Equation 12:

$$\hat{\beta}_c^D = -\log(\hat{E}_c)/z \tag{13}$$

β_c^D can be expressed as 2-term exponential form:

$$\beta_c^D = a * e^{(b*z)} + c * e^{(d*z)} \tag{14}$$

Parameters ”a,b,c,d” can be estimated from $\hat{\beta}_c^D$ with non-linear least square method, then used to refine the attenuation coefficient.

3.5 Image restoration

Finally, after obtaining backscatter signal value, attenuation coefficient, the signal camera received I_c was able to restore the color through Equation 15 derived from Equation 2:

$$J_c = (I_c - B_c) * e^{\beta_c^D * z} \tag{15}$$

A white balance adjustment was then applied, based on the top 10% average values of the blue and green channels.

4. Evaluation

This study applied SSIM, a Full Reference Image Quality Assessment metric, to evaluate performance because it considers neighboring pixels, unlike MSE and PSNR, which focus on individual pixels. The value of SSIM ranges from 0 to 1, with higher values indicating greater similarity to the reference image. Based on our previous research (Kuan & Jaw, 2024), we developed an adjusted SSIM framework, called underwater SSIM, that better suits the characteristics of underwater images. SSIM consists of three components: luminance, contrast, and structure. However, analysis shows that the structure component is not sensitive to underwater image distortion, so it was removed. Additionally, since the primary distortion in underwater images is a shift in color, we introduced the luminous efficiency function (Wyszecki & Stiles, 2000) into the framework. This function reflects the human eye's sensitivity to different colors by converting pixel values from the sRGB color space to the CIE1931 XYZ color space (Daniel, 2011). The Y value, representing luminance, was used as a weight for each pixel in the reference image to calculate a weighted average, resulting in the underwater SSIM score:

$$underwater\ SSIM = \frac{\sum w_i\left([l_i(x,y)]^{\alpha}[c_i(x,y)]^{\beta}\right)}{\sum w_i} \tag{16}$$

$$l(x,y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2\mu_y^2 + C_1} \tag{17}$$

$$c(x,y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2\sigma_y^2 + C_2} \tag{18}$$

$$C_1 = (K_1 L)^2 \tag{19}$$

$$C_2 = (K_2 L)^2 \tag{20}$$

where l = luminance factor

c = contrast factor

w_i = luminous efficiency of each pixel color

x, y = two images input

μ = mean value of image

σ = standard deviation of image

$K_1 = constant\ close\ to\ 0$

$L = dynamic\ range\ of\ pixel\ values$

$K_2 = constant\ close\ to\ 0$

α and β are set to 1. K_1 and K_2 are set to 0.01 and 0.03, respectively. In 8-bit images, L is 255.

The weight w_i is calculated by first applying gamma correction to convert the pixel values from sRGB to linear RGB. Then, using Equation 21, the Y value of pixel in the CIE1931 XYZ color space is obtained, which corresponds to w_i.

$$w_i = [0.2126729 \quad 0.7151522 \quad 0.0721759] * [linear\ RGB] \tag{21}$$

Considering that existing algorithms do not provide a standard SSIM value to determine what is considered good enough, we reviewed five studies that used SSIM to establish a benchmark. The first study, when excluding highly turbid data, showed SSIM values close to or above 0.8 (Anwar et al., 2018) . The second study compared 12 algorithms using 3 datasets, with the top 2 algorithms in each dataset scoring over 0.76 (Lyu et al., 2022) . The third study introduced a learning-based method that outperformed 7 other algorithms and achieved an SSIM score of 0.86, with the second-highest scoring 0.80 (Liu et al., 2022) . The fourth study compared 10 algorithms, with the top 2 scoring 0.88 and 0.81, showing little visual difference (Islam et al., 2020) . The final study proposed an enhancement method for marine creature images, with SSIM scores near or above 0.8 (Jamadandi & Mudenagudi, 2019) . Based on this review, we adopted an SSIM score of 0.8 as a reasonable standard for underwater image restoration.

Results and Discussion

1. Depth map generation

The scene measurement was conducted at sunny noon, providing sufficient light to ensure adequate feature points in the photos. The area of the scene was approximately 8 * 8 m², and the expected accuracy of the dense point cloud was at the millimeter level. For each total

station setup, we measured 19 to 21 markers and collected observations including slope distances from the total station to the markers, azimuth angles, and horizontal angles between the markers and a reference point, resulting in a total of 237 observations. The unknown parameters included the coordinates of 32 markers and the transition parameters between two total station setups, such as rotation angles κ and translations ΔX, ΔY, and ΔZ, making a total of 108 unknown parameters. This gave a redundancy of 237 - 108 = 129. The coordinates of the markers were solved using the least squares method, with the iteration termination threshold set to one-tenth of the expected accuracy. When this threshold was reached, it indicated that the corrections to the observations had become minimal, and the iteration had converged.

The iteration converged at second loop, the threshold was 0.00007m (0.07mm), as shown in Figure 5, and the a posteriori standard deviation of unknown parameters was 0.00005m (0.05mm), which met the expected accuracy. The coordinates of markers were then inputted into the Agisoft Metashape Pro as the true value of the control points and the check points to adjust the dense point cloud. Figure 6 shows the dense point cloud of the scene in Agisoft Metashape Pro after adjustment. The Root Mean Square Difference (RMSD) between control points and check points, comparing the dense point cloud and total station-measured coordinates, are shown in Table 2.

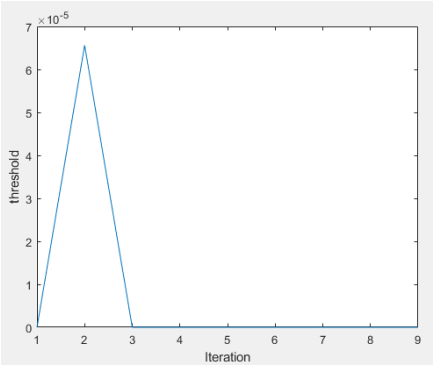


Figure 5: threshold diagram

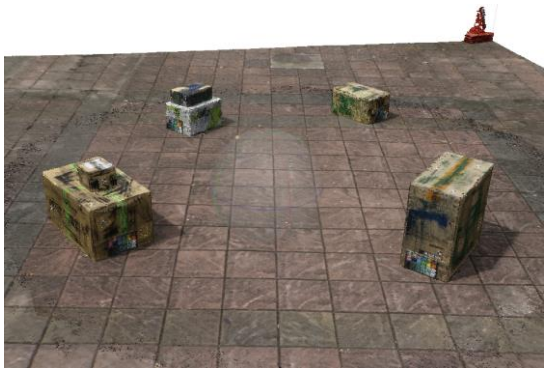


Figure 6: dense point cloud of the scene

Table 2: RMSD of control points and check points

Point type	X(m)	Y(m)	Z(m)	Total(m)
control point	0.004	0.004	0.003	0.006
check point	0.004	0.002	0.002	0.006

With dense point cloud of the scene, we generated the depth map within Agisoft Metashape Pro. Each pixel in depth map represents the object distance between camera and corresponding object point of the scene. Comparison of reference image and its depth map are as shown in Figure 7 (a), (b):

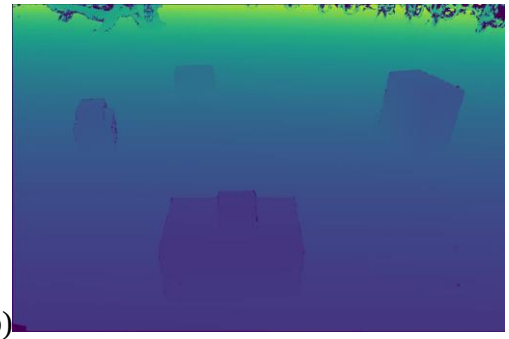
(a)  (b) 

Figure 7: Comparison of reference image and its depth map, (a) reference image; (b) corresponding depth map

Inevitably, some data were missing in the depth map, either because those areas were not the area of interest in the scene and thus captured by an insufficient number of photos, or due to the uniform texture of the object. These missing data points were not considered in the subsequent processes.

2. Underwater image simulation

We used MATLAB software to simulate underwater images by utilizing a reference image, depth map, and the inherent optical properties of various types of seawater. One of the environmental parameters required to calculate β_c^D, reflectance, is impractical to measure for every surface color in each simulation. Therefore, in our previous research (Kuan & Jaw, 2024), we calculated the distance between each pixel's color and a color database in the RGB color space, then directly assigned the reflectance value of the nearest match to each pixel. Partial simulation results are shown in Figure 8. In Jerlov type I, the details of the scene are still visible. However, in Jerlov type IB, the images become more turbid, and in Jerlov type 1C, the objects in the scene are barely distinguishable. Images at a water depth of 5 meters

experience more severe attenuation compared to those at 1 meter, due to the longer light travel path.

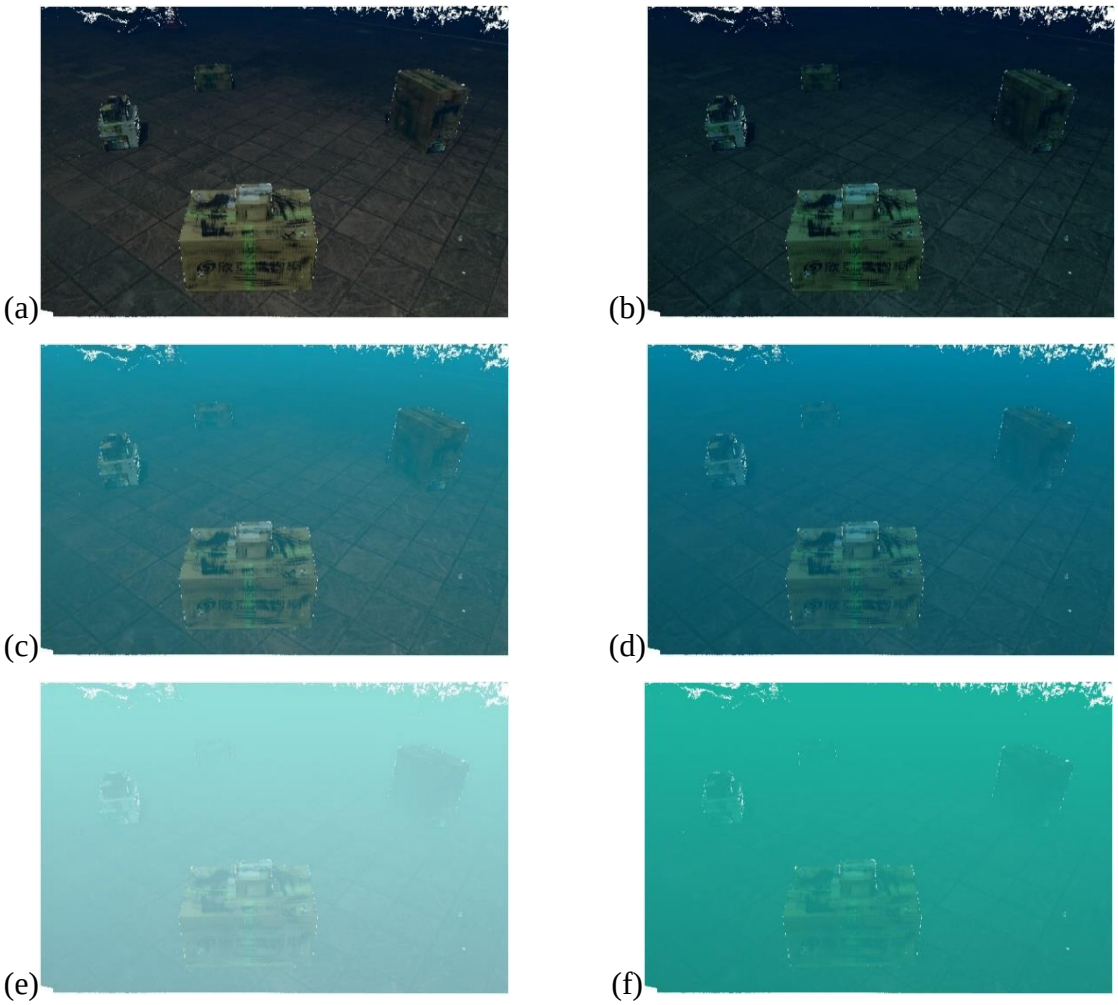


Figure 8: Partial simulation results, (a) Jerlov type I_water depth 1m; (b) Jerlov type I_water depth 5m; (c) Jerlov type IB_water depth 1m; (d) Jerlov type IB_water depth 5m; (e) Jerlov type 1C_water depth 1m; (f) Jerlov type 1C_water depth 5m

3. Color restoration & Evaluation

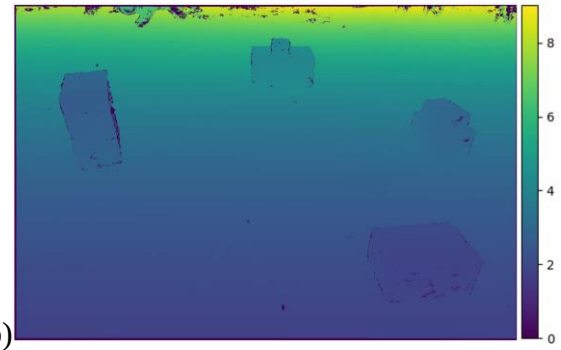
For each water type, we selected five standard deviation intervals for error added on object distance: [0.50, 0.30, 0.10, 0.05, 0.00] meters. The depth map values were simplified by dividing each pixel value by the interval of 0.5 meters and taking the quotient.

Through the Sea-thru method, we picked images taken from 3 different angles “IMG_3804, IMG_3815, IMG_3819”. Reference images, depth maps and their restored images are as shown in Figure 9. Only the results for water types I, IB, and 1C at water depth 1 meter are presented, as the performance in turbid water types ware unsatisfactory. Additionally, we present the results for water type I at a depth of 5 meters to compare the performance of the

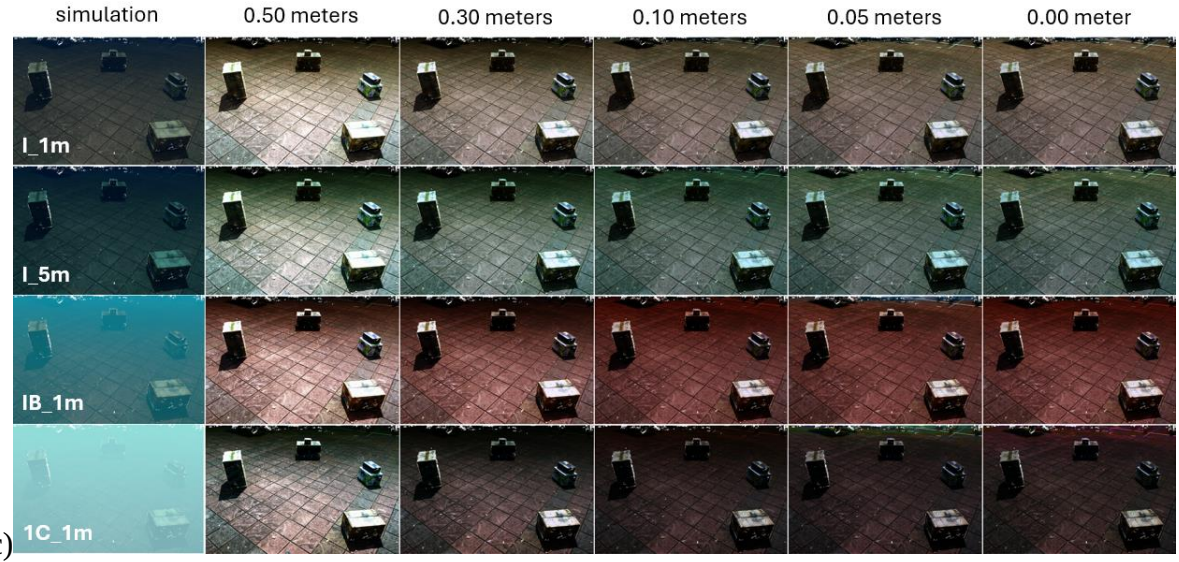
Sea-thru method under different water depths. The SSIM values are shown in Table 3 and its corresponding diagram is shown in Figure 10, where environments such as IMG_3804 with Jerlov type I at a depth of 1 meter are denoted as “04_I_1m”.

(a)

(b)

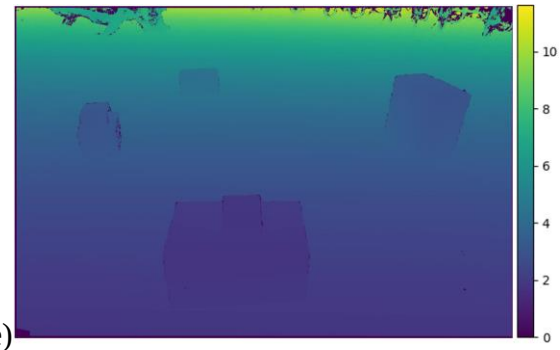


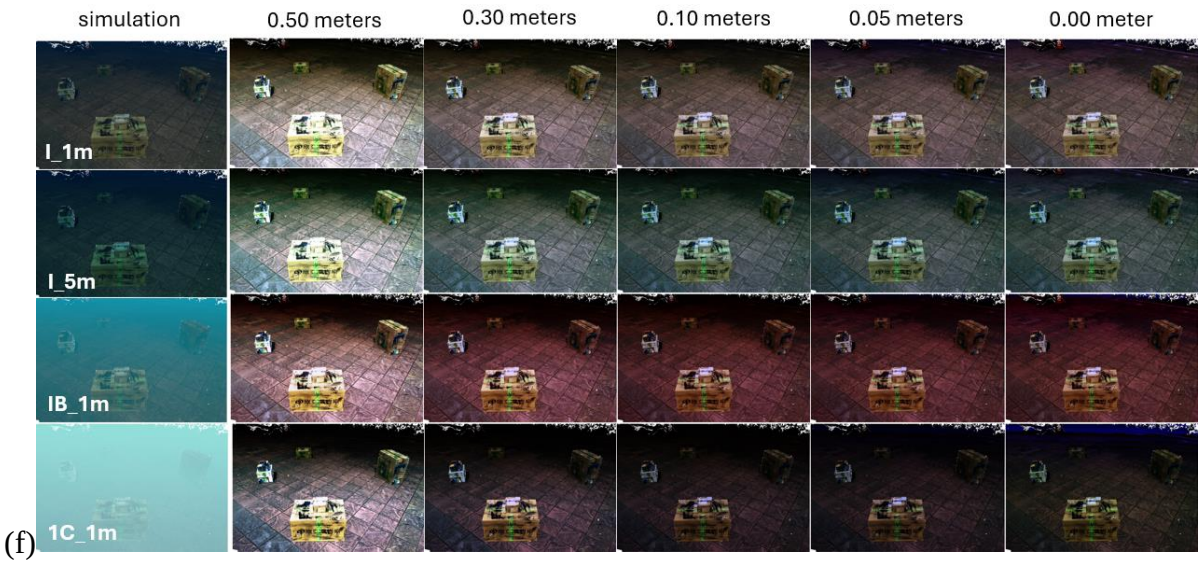
(c)



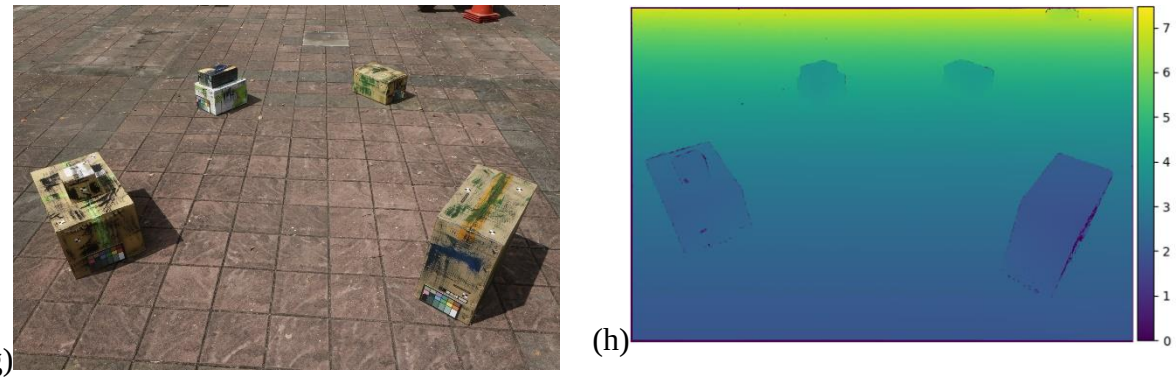
(d)

(e)

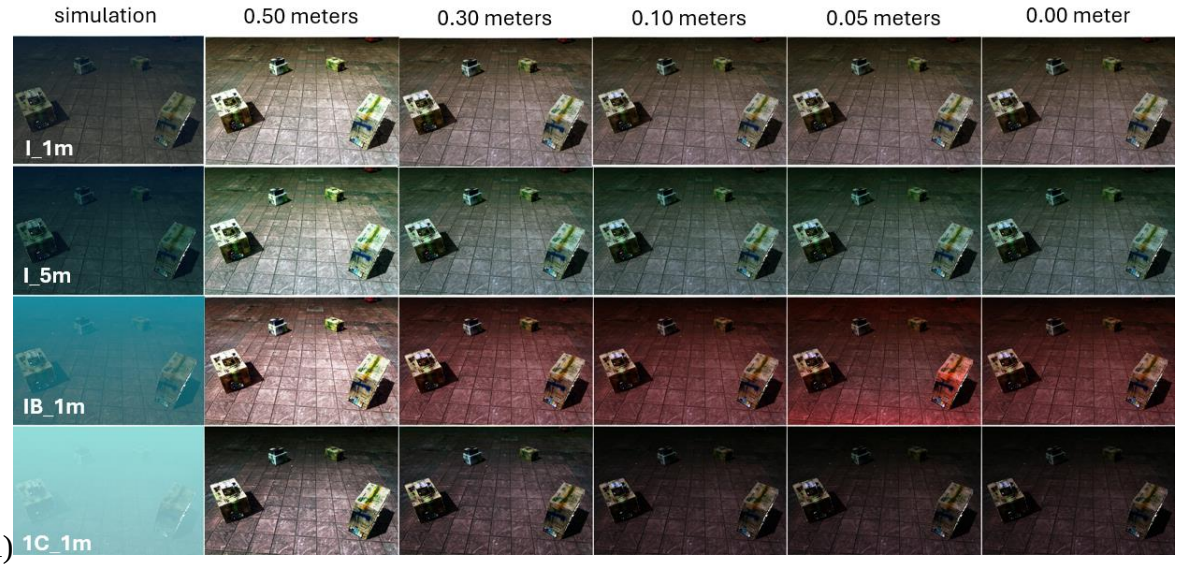




(f)



(g) (h)



(i)

Figure 9: Results comparison, (a), (d), (g) reference images from 3 different angles; (b), (e), (h) corresponding depth maps; (c), (f), (i) restored images under different sets of environments

Table 3: SSIM of restored images with different error standard deviations

Environment	0.50 m	0.30 m	0.10 m	0.05 m	0.00 m
“04_I_1m”	0.638	0.790	0.899	0.931	0.933
“04_I_5m”	0.619	0.757	0.852	0.893	0.897
“04_IB_1m”	0.567	0.691	0.706	0.780	0.711
“04_1C_1m”	0.493	0.595	0.497	0.551	0.548
“15_I_1m”	0.555	0.743	0.793	0.827	0.854
“15_I_5m”	0.525	0.712	0.742	0.766	0.798
“15_IB_1m”	0.532	0.652	0.634	0.641	0.658
“15_1C_1m”	0.464	0.542	0.438	0.431	0.477
“19_I_1m”	0.681	0.819	0.822	0.808	0.802
“19_I_5m”	0.631	0.788	0.772	0.750	0.738
“19_IB_1m”	0.537	0.702	0.649	0.635	0.610
“19_1C_1m”	0.547	0.619	0.451	0.413	0.385

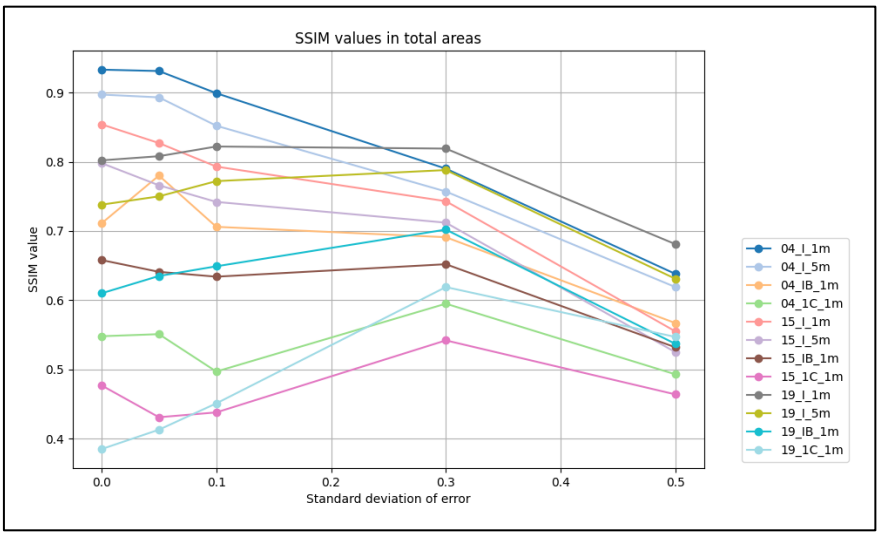


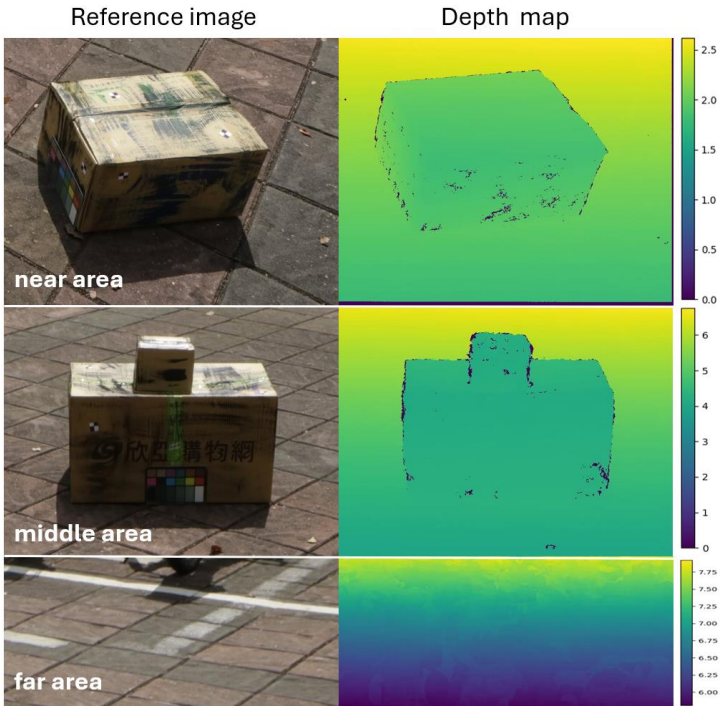
Figure 10: SSIM values in total areas

The results showed that only images taken in clear water types, such as type I, could be properly restored. Turbid water types, like type IB or type IC, usually failed to meet the 0.8 SSIM standard. For type I, as the error standard deviation decreased, SSIM increased at both 1-meter and 5-meter water depths. Overall, the pattern of SSIM changes with error standard deviation was consistent for IMG_3804 and IMG_3815, while IMG_3819 peaked at a

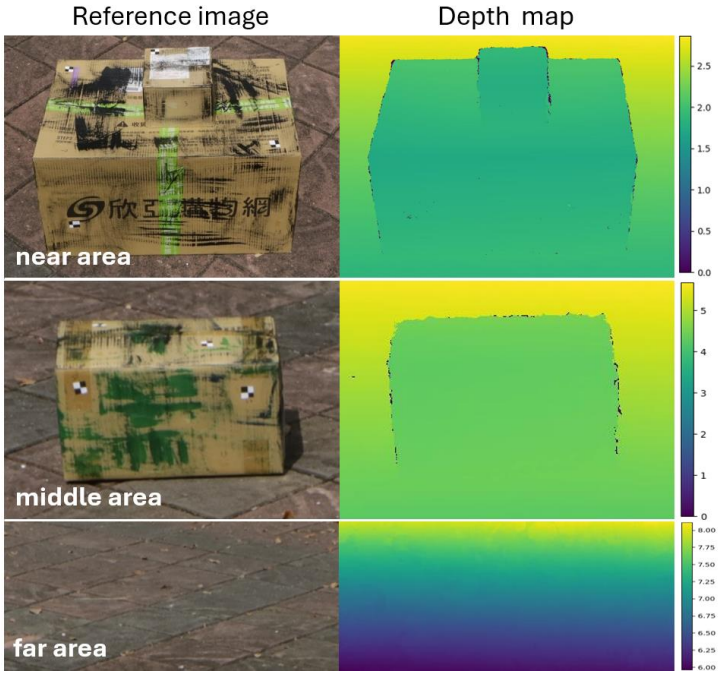
standard deviation of around 0.30 meters across four environments. This may be due to the higher camera position in IMG_3819, causing the object distances in the near scene to be farther, resulting in a darker image. Additionally, higher error standard deviations tended to increase the attenuation coefficient β_c^D, making the images appear brighter. Together, these factors offset the negative effects for the 0.30-meter sets.

We further analyzed the near, middle, and far areas of images from three angles. The object distances in the near areas were around 2 meters, in the middle areas around 4 meters, and in the far areas around 7 meters, as shown in Figure 11. The SSIM values for the near areas are presented in Table 4, those for the middle areas in Table 5, and those for the far areas in Table 6. Their corresponding diagrams are shown in Figure 12 to 14.

(a)



(b)



(c)

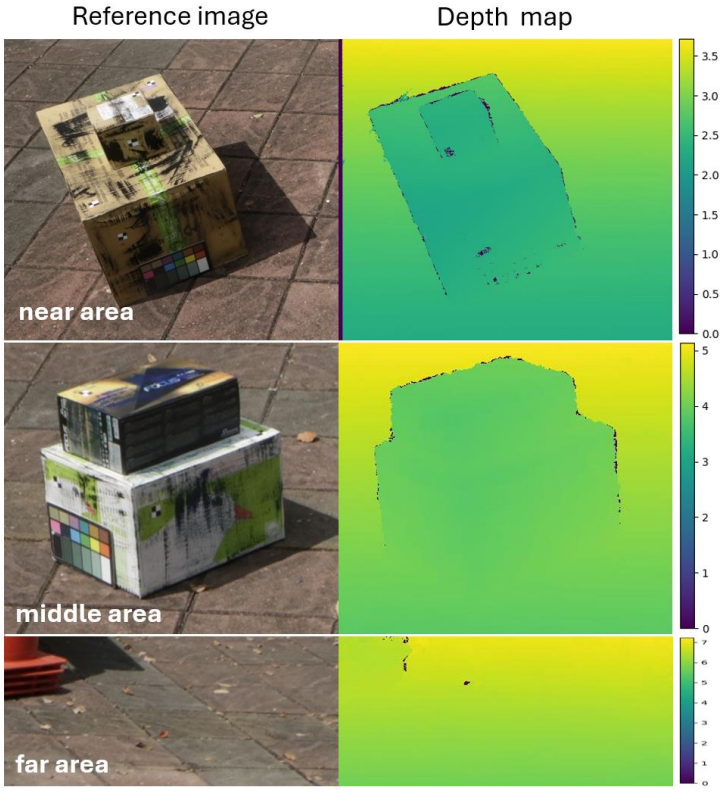


Figure 11: Areas of reference images selected and corresponding depth maps, (a) near, middle, far areas in IMG_3804; (b) near, middle, far areas in IMG_3815; (c) near, middle, far areas in IMG_3819

Table 4: SSIM of near areas in restored images with different error standard deviations

Environment	**0.50 m**	**0.30 m**	**0.10 m**	**0.05 m**	**0.00 m**

“04_I_1m”	0.607	0.822	0.923	0.939	0.941
“04_I_5m”	0.637	0.806	0.916	0.920	0.917
“04_IB_1m”	0.593	0.799	0.861	0.868	0.861
“04_1C_1m”	0.515	0.711	0.712	0.716	0.708
“15_I_1m”	0.559	0.780	0.929	0.946	0.952
“15_I_5m”	0.537	0.757	0.914	0.925	0.927
“15_IB_1m”	0.531	0.770	0.862	0.868	0.865
“15_1C_1m”	0.512	0.725	0.699	0.694	0.714
“19_I_1m”	0.702	0.875	0.922	0.919	0.914
“19_I_5m”	0.653	0.843	0.884	0.874	0.862
“19_IB_1m”	0.559	0.797	0.780	0.758	0.746
“19_1C_1m”	0.692	0.686	0.579	0.544	0.511

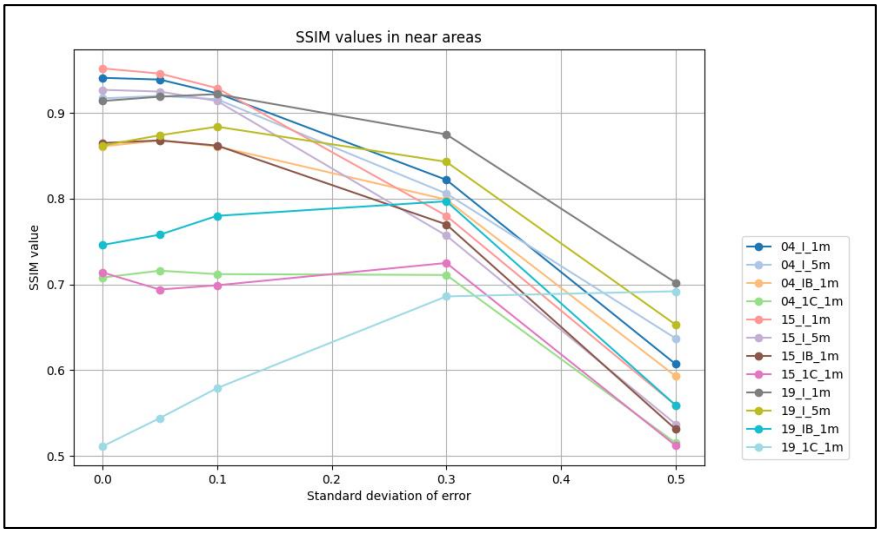


Figure 12: SSIM values in near areas

Table 5: SSIM of middle areas in restored images with different error standard deviations

Environment	**0.50 m**	**0.30 m**	**0.10 m**	**0.05 m**	**0.00 m**
“04_I_1m”	0.747	0.721	0.822	0.881	0.889
“04_I_5m”	0.640	0.626	0.742	0.838	0.856

“04_IB_1m”	0.649	0.508	0.506	0.668	0.517
“04_1C_1m”	0.591	0.386	0.314	0.491	0.515
“15_I_1m”	0.738	0.765	0.734	0.783	0.826
“15_I_5m”	0.698	0.712	0.626	0.674	0.748
“15_IB_1m”	0.718	0.487	0.428	0.464	0.523
“15_1C_1m”	0.521	0.253	0.205	0.214	0.301
“19_I_1m”	0.858	0.877	0.791	0.762	0.749
“19_I_5m”	0.821	0.842	0.720	0.678	0.661
“19_IB_1m”	0.718	0.703	0.556	0.534	0.504
“19_1C_1m”	0.676	0.648	0.361	0.298	0.280

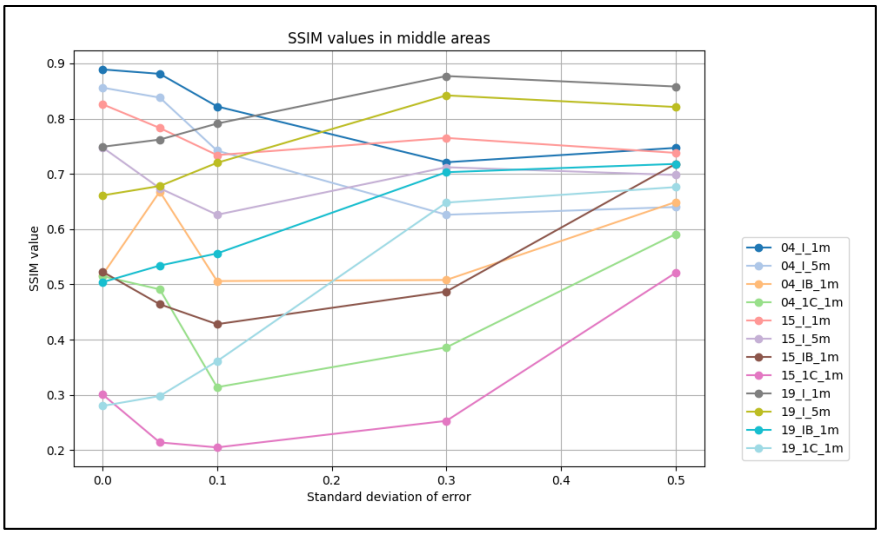


Figure 13: SSIM values in middle areas

Table 6: SSIM of far areas in restored images with different error standard deviations

Environment	**0.50 m**	**0.30 m**	**0.10 m**	**0.05 m**	**0.00 m**
“04_I_1m”	0.424	0.346	0.739	0.865	0.871
“04_I_5m”	0.259	0.236	0.628	0.810	0.835
“04_IB_1m”	0.338	0.154	0.311	0.662	0.330
“04_1C_1m”	0.282	0.079	0.144	0.502	0.441

“15_I_1m”	0.192	0.213	0.244	0.388	0.515
“15_I_5m”	0.145	0.161	0.166	0.279	0.429
“15_IB_1m”	0.107	0.065	0.069	0.112	0.189
“15_1C_1m”	0.053	0.022	0.028	0.068	0.141
“19_I_1m”	0.624	0.467	0.345	0.314	0.310
“19_I_5m”	0.596	0.411	0.269	0.234	0.228
“19_IB_1m”	0.633	0.254	0.151	0.211	0.133
“19_1C_1m”	0.156	0.240	0.067	0.050	0.049

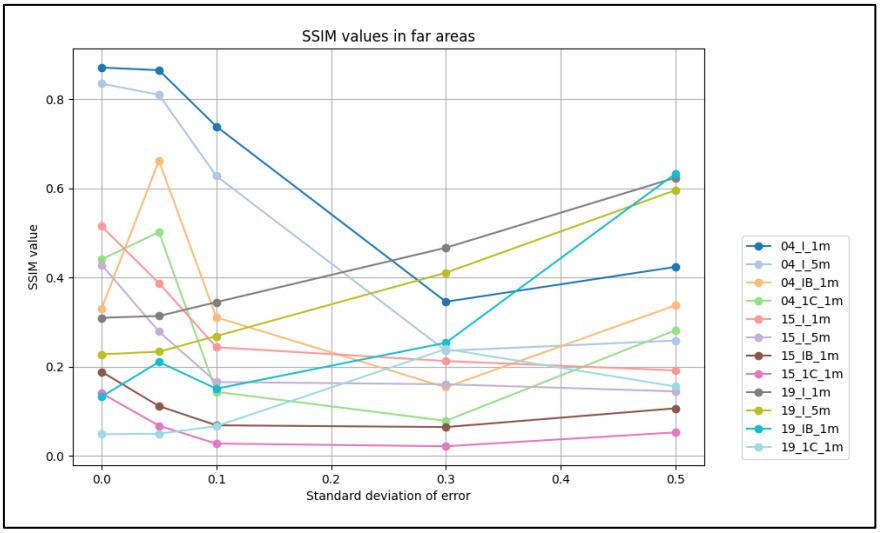


Figure 14: SSIM values in far areas

Results show that the near areas exhibited a similar pattern to the total area, likely due to their large percentage in the image. However, the middle and far areas displayed inconsistent patterns, indicating that the performance of the Sea-thru method in these regions is unstable and less correlated with object distance accuracy compared to the near areas. This suggests that the Sea-thru method may be more suitable for scenes at closer distances

Finally, for type I at a 1-meter depth, with an error standard deviation between 0.30 and 0.10, the SSIM just reached the 0.8 standard, indicating the limit of accuracy for this condition. However, at a 5-meter depth, type I only met the standard in one case, and another set nearly reached it when no error was present. This suggests that object distance sensitivity is higher in deep-water environments. For types IB and IC, the low SSIM values indicate that the Sea-

thru method may not be suitable for turbid water types, which could lead to unexpected results.

In comparison with the original paper (Akkaynak & Treibitz, 2019), they did not use SSIM for evaluation, as their dataset lacked reference images. Instead, they used the RGB angular error on a color chart placed in the scene to evaluate the restoration performance for specific colors, for which they had ground truth values in air. Some water types in our dataset, such as Jerlov type 1C, are more turbid than theirs, which may help clarify the applicable conditions for the Sea-thru method. Additionally, for scenes in deeper water, their results were brighter than ours, possibly because the variety of colors in our scenes was insufficient to apply the Gray World Assumption effectively.

Conclusion and Recommendation

This study analyzed the required object distance accuracy for the Sea-thru method under different environmental conditions. A controlled field setup was used to ensure depth map accuracy, and the generated depth map was then utilized to simulate underwater images for color restoration tasks. The Sea-thru method was applied to the simulated images with simplified depth maps, and random errors with a mean value of 0 and five intervals of standard deviation—[0.50, 0.30, 0.10, 0.05, 0.00] meters—were added to the depth map to determine the accuracy needed to meet the 0.8 SSIM standard.

The results showed that the minimum accuracy requirement for object distance is 0.30 meters, as lower accuracy may lead to unsatisfactory performance. However, this finding is applicable only to clear water environments, suggesting that the Sea-thru method may require further adjustments to perform effectively in different conditions. Analysis of areas at varying object distances indicates that the Sea-thru method performs more consistently in near areas compared to middle and far areas, with its correlation to object distance accuracy diminishing as the object distance increases.

References

Akkaynak, D., & Treibitz, T. (2018). A revised underwater image formation model. Proceedings of the IEEE conference on computer vision and pattern recognition,

Akkaynak, D., & Treibitz, T. (2019). Sea-thru: A method for removing water from underwater images. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition,

Anwar, S., Li, C., & Porikli, F. (2018). Deep underwater image enhancement. *arXiv preprint arXiv:1807.03528*.

Buchsbaum, G. (1980). A spatial processor model for object colour perception. *Journal of the Franklin institute, 310*(1), 1-26.

Daniel, M. (2011). Color Vision and Colorimetry: Theory and Applications, Second Edition: Front Matter. In. SPIE Press. https://doi.org/10.1117/3.881172.fm

Ebner, M. (2009). Color constancy based on local space average color. *Machine Vision and Applications*, *20*(5), 283-301.

Guo, P., He, L., Liu, S., Zeng, D., & Liu, H. (2021). Underwater image quality assessment: Subjective and objective methods. *IEEE Transactions on Multimedia*, *24*, 1980-1989.

Islam, M. J., Xia, Y., & Sattar, J. (2020). Fast underwater image enhancement for improved visual perception. *IEEE Robotics and Automation Letters*, *5*(2), 3227-3234.

Jamadandi, A., & Mudenagudi, U. (2019). Exemplar-based underwater image enhancement augmented by wavelet corrected transforms. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops,

Jerlov, N. G. (1977). Classification of sea water in terms of quanta irradiance. *ICES Journal of Marine Science*, *37*(3), 281-287.

Kuan, H., & Jaw, J. (2024). Establishing Full Reference Image Quality Assessment for Underwater Image Enhancement. *International Symposium of Remote Sensing 2024, 24-26 April 2024, Taichung, Taiwan*.

Liu, Y., Xu, H., Zhang, B., Sun, K., Yang, J., Li, B., Li, C., & Quan, X. (2022). Model-based underwater image simulation and learning-based underwater image enhancement method. *Information*, *13*(4), 187.

Lyu, Z., Peng, A., Wang, Q., & Ding, D. (2022). An efficient learning-based method for underwater image enhancement. *Displays*, *74*, 102174.

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

Solonenko, M. G., & Mobley, C. D. (2015). Inherent optical properties of Jerlov water types. *Applied optics*, *54*(17), 5392-5401.

Wang, H., Zhang, W., & Ren, P. (2024). Self-organized underwater image enhancement. *ISPRS Journal of Photogrammetry and Remote Sensing*, *215*, 1-14.

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, *13*(4), 600-612.

Wyszecki, G., & Stiles, W. S. (2000). *Color science : concepts and methods, quantitative data, and formulae / Günther Wyszecki, W.S. Stiles* (Wiley classics library ed.). John Wiley & Sons.

Yang, M., & Sowmya, A. (2015). An underwater color image quality evaluation metric. *IEEE Transactions on Image Processing*, *24*(12), 6062-6071.