**Marine Snow Removal Benchmarking Dataset**

Reina Kaneko*, Yuya Sato†, Takumi Ueda†, Hiroshi Higashi*, and Yuichi Tanaka*

* Osaka University, Japan
† Tokyo University of Agriculture and Technology, Japan
E-mail: r.kaneko@sip.comm.eng.osaka-u.ac.jp, {higashi, ytanaka}@comm.eng.osaka-u.ac.jp

---

**Abstract**—This paper introduces a new benchmarking dataset for marine snow removal of underwater images. Marine snow is one of the main degradation sources of underwater images that are caused by small particles, e.g., organic matter and sand, between the underwater scene and photosensors. We mathematically model two typical types of marine snow from the observations of real underwater images. The modeled artifacts are synthesized with underwater images to construct large-scale pairs of ground truth and degraded images to calculate objective qualities for marine snow removal and to train a deep neural network. We propose two marine snow removal tasks using the dataset and show the first benchmarking results of marine snow removal. The Marine Snow Removal Benchmarking Dataset is publicly available online.

---

**I. INTRODUCTION**

Image restoration has been one of the main topics of computer vision for decades [10], [11], [28], [35]–[38]. In the current era of deep learning, the quality of image restoration for various tasks has been significantly improved under a sufficient number of pairs of ground truth and degraded images. Accordingly, the current focus on image restoration will be for images taken under extreme situations such as underwater, satellite, and medical images [1], [3], [6], [12], [14], [20], [23], [24], [26], [27], [31]–[34], [39], [41]. This problem is called **extreme image restoration** herein. Extreme images often have an untypical degradation beyond that of a Gaussian model, and the numbers of available images for training and modifying restoration algorithms are also limited. As a result, extreme image restoration is a challenging problem.

In this paper, we focus on underwater image enhancement. Its main challenge is mostly due to the fact that we generally have no ground truth images. This makes it difficult to evaluate the objective qualities of the restored images and measure the restoration performance. Moreover, the lack of pairs of ground truth and degraded image causes a difficulty in training a neural network (if we consider using a deep neural network for restoration) because the loss functions of the image restoration typically evaluate some objective image quality metrics such as mean squared error (MSE) and structured similarity index (SSIM) [40]. Hence, unsupervised restoration methods, including linear and nonlinear filtering, are still popular in underwater image enhancement [2], [6], [24].

Several studies for underwater image enhancement with deep neural networks have been proposed [14], [18], [20], [22], [23], [32], [33], [42]. Most focus on removing the color shift from underwater images. That is, the blueshift underwater images are restored by enhancing the red and green channels. Conventional underwater image enhancement methods using deep neural networks tackle the problem on the lack of large-scale underwater image datasets in various ways. For example, the studies in [3], [14], [32], [33] synthetically generate degraded images from clean images taken on the ground by simulating the degradation processes of absorption and scattering. Moreover, [15], [20], [25] use generative adversarial networks. And a number of challenges to improving the restoration performance still exist, some promising results have been presented. However, such studies ignore another main source of degradation for underwater images: marine snow.

In this paper, we consider a marine snow removal (MSR) problem for underwater images by constructing a large-scale dataset containing underwater image pairs. Marine snow is a typical underwater image degradation, examples of which are shown in Fig. 1. Marine snow artifacts are visible in digital images when we take an underwater image using a flashlight. Small particles in the scene, e.g., organic matter and sand, cause marine snow. The particles reflect the light from the flashlight, and the reflected light will be captured by the photosensors. Marine snow is often annoying in underwater photography because it affects the overall image quality. However, the particles are unevenly distributed in the scene, and the degradation process is difficult to model completely. Therefore, standard methods for marine snow removal are still limited to simple median-filter-based approaches [5], [13], [18], [21]. Moreover, as with standard underwater image enhancement, we do not have a pair of ground truth and degraded images to calculate the objective image qualities and tune the parameters of the restoration algorithms.

We propose marine snow synthesis methods for solving the above-mentioned problems of MSR. In other words, we intentionally append the marine snow artifacts to real underwater images without marine snow and generate pairs of ground truth images and degraded images (i.e., clean underwater images and their corresponding images with marine snow artifacts). The key for synthesizing artifacts is to mathematically model marine snow in digital images as accurately as possible. From actual marine snow images, we discover that many marine snow artifacts can be modeled through two typical pixel value
distributions. By utilizing this finding, we can synthetically generate many pairs of underwater images with/without marine snow artifacts.

We also propose two marine snow removal tasks and develop datasets corresponding to these tasks. This Marine Snow Removal Benchmarking Dataset (MSRB Dataset) is the first attempt in this field and is publicly available online. We also present the first MSR benchmarking results using the MSRB Dataset: We compare the objective performance of marine snow removal among median filters and a deep neural network. Real MSR results as well as the limitations are also described.

The remainder of this paper is organized as follows: Section II reviews several related studies for underwater image enhancement. Two representative marine snow models that we discovered are presented in Section III along with their synthesizing method. The MSRB Dataset specifications are introduced in Section IV and include the description of the two tasks. The first MSR benchmarking results with objective image quality metrics are shown in Section V. Real MSR results and limitations are described in Section VI. Finally, we provide some concluding remarks in Section VII.

II. RELATED STUDIES

There have been few studies on modeling marine snow in digital images. Its seminal study is described in [7], [8]. A marine snow artifact is modeled using a Gaussian function in which the artifact is less transparent in its center than in its surrounding area. This method is better than the traditional salt and pepper noise; however, it does not simulate the actual shapes of marine snow artifacts. As a result, the center part of marine snow becomes excessively thick and the generated images are far from the real ones.

MSR is also an underrepresented problem in underwater image enhancement. Here, we briefly review some existing approaches.

A widely used method for MSR is a median filter (MF) [9], [19]. If marine snow artifacts can be assumed sufficiently small (typically 1–3 pixels in diameter), MF is expected to work because small marine snow artifacts can be considered to have a salt-and-pepper effect. However, it significantly blurs the entire image, particularly when we use a large filter kernel size.

Specific to MSR, a modified version of a MF is proposed in [5], [17]. This method applies the MF selectively if the target pixel has a higher intensity than the surrounding pixels. However, it is still difficult to remove large marine snow artifacts.

A few MSR methods for video sequences have also been proposed. They utilize the fact that marine snow artifacts continuously move in consecutive video frames. In [16], background modeling is used to remove marine snow artifacts from a static scene. Marine snow artifacts are tracked and a customized MF is applied to the detected artifacts in [13]. However, they are not applicable to a single underwater image. Importantly, all methods above are based on a model-based approach. This is mainly due to a lack of high-quality datasets, as mentioned in Section I.

A deep learning-based MSR method is proposed in [18]. Its training dataset consists of synthesized marine snow images, corresponding ground truth images, marine snow masks. This dataset also contains real underwater images with marine snow (without ground truth). Marine snow particles are manually made with Adobe Photoshop: That results in a lack of randomness and scalability.

In the following, we address the contributions of this study; marine snow models and the dataset specifications.

---

1https://github.com/yachtanaka/marine-snow
There are various sources of marine snow and it is impractical to estimate the sources of all particles from a single underwater image. Instead of the estimation of the sources, we model the pixel value distributions of marine snow artifacts from observations of underwater images.

A. Real Marine Snow Examples

First, we start by showing marine snow examples in real underwater images. Figs. 2(a) and (c) show enlarged portions of representative marine snow artifacts cropped from real underwater images. Although they look similar, their pixel value distributions are slightly different.

Taking a closer look, the 3D plots of Figs. 2(a) and (c) are shown in Figs. 2(b) and (d), respectively. As clearly observed, they do not have a shape like a Gaussian function in contrast to the conventional assumption in [7], [8]. Rather than a Gaussian function, these 3D plots are similar to elliptic conical frusta, i.e., sliced elliptic cones. Furthermore, the top surfaces of the frusta have different characteristics between Figs. 2(b) and (d).

In our preliminary observation, most marine snow artifacts can be classified into these two representative shapes. In the following, we present the mathematical models of marine snow artifacts that reflect the above-mentioned observations.

B. Marine Snow Model 1: Highland Type

The first marine snow model is the Highland Type (type H). The type H marine snow corresponds to Fig. 2(a) and can be modeled as an elliptic conical frustum with a rough surface.

For modeling type H marine snow, we first suppose to have two ellipses $f_1$ and $f_2$, both centered at the coordinates $(k, l)$. As shown in Fig. 3, we assume the focal points of $f_1$ and $f_2$ are located on the horizontal axis, i.e., the ellipses are wider, for simplicity. Let $a_i$ and $b_i$ be the semi-major and semi-minor axes of $f_i$ where $i \in \{1, 2\}$, respectively. Suppose that $f_1$ is larger than $f_2$, i.e., $a_1 > a_2$ and $b_1 > b_2$ and the eccentricities of $f_1$ and $f_2$ are identical, i.e., $\frac{a_1}{b_1} = \frac{a_2}{b_2}$.

For notational simplicity, $(k, l)$ is set to $(0, 0)$ hereafter. The $(m, n)$th pixel value of the type H marine snow $h(m, n)$ is formulated as follows:

$$h(m, n) = \begin{cases} c + \epsilon & \text{if } \frac{m^2}{a_1^2} + \frac{n^2}{b_1^2} \leq 1 \land \frac{m^2}{a_2^2} + \frac{n^2}{b_2^2} \leq 1 \land \frac{m^2}{a_1^2} + \frac{n^2}{b_1^2} > \frac{m^2}{a_2^2} + \frac{n^2}{b_2^2} \\ 0 & \text{otherwise} \end{cases}$$

where $c$ is a constant that determines the transparency of the marine snow, $d(\cdot, \cdot)$ calculates the Euclidean distance between two points, and $\epsilon$ is a small perturbation that mimics a rough surface. Furthermore, $p := [m, n]^T$ and $q_i$ is the intersection of the ellipse $f_i$ and the straight line crossing $(m, n)$, as illustrated in Fig. 3.

An example of the synthesized type H marine snow artifacts is shown in Fig. 4(a) along with its 3D plot in Fig. 4(b). See the similarity in Figs. 2(a) and (b). For creating a dataset, the major axis of $h(m, n)$ is randomly rotated and the transparency $c$ is also randomly chosen. These specifications are shown later in Section IV.

C. Marine Snow Model 2: Volcanic Crater Type

The second marine snow model is the volcanic crater type (type V). An example of type V marine snow is shown in Figs. 2(c) and (d). This can be modeled by a modified version of the type H with an overshoot top edge of the frustum.

For the type V marine snow, we consider another ellipse $f_3$, where $a_3$ and $b_3$ are its semi-major and semi-minor axes, respectively. It is located inside $f_2$, as illustrated in Fig. 3. As for the type H, focal points of all ellipses are assumed to be located along the horizontal axis and their eccentricities are identical, i.e., $\frac{a_1}{b_1} = \frac{a_2}{b_2} = \frac{a_3}{b_3}$.

Based on the type H function $h(m, n)$, the $(m, n)$th pixel value of the type V marine snow $v(m, n)$ is formulated as...
TABLE I: Parameter settings for marine snow synthesis where $\mathcal{U}(v_{\min}, v_{\max})$ is a continuous uniform distribution between $v_{\min}$ and $v_{\max}$.

| Parameter | Setting |
|-----------|---------|
| $c$ in (1) | $\mathcal{U}(10, 80)^*$ |
| $c_r$ in (2) | $\mathcal{U}(1, 40)^*$ |
| Gaussian filter size in (3) | $2(r + 1) \times (2r + 1)$ where $r \sim \mathcal{U}(1, 6)$ |
| $a_2$ | $\mathcal{U}(1, 3)$ for small marine snow artifacts |
| $b_2$ | $\mathcal{U}(3, 16)$ for large marine snow artifacts |
| $a_1$ | $\mathcal{U}(0.3a_2, a_2)$ for small marine snow artifacts |
| $a_3$ | $\mathcal{U}(0.5a_2, a_2)^{\dagger}$ for large marine snow artifacts |

‡ $b_1$ is obtained from the eccentricity of $f_3(m, n)$
† $b_3$ is obtained from the eccentricity of $f_2(m, n)$

$v(m, n) = h(m, n) + g(m, n)$, where $g(m, n)$ is the rim term defined as

$$g(m, n) = \begin{cases} c_r(1 - \frac{d((m, n) + q_3))}{2d(q_2, q_3)}) & 1 < \frac{a_2^2}{c_3^2} + \frac{b_2^2}{c_3^2} \wedge \frac{a_1^2}{c_3^2} + \frac{b_1^2}{c_3^2} \leq 1 \\ 0 & \text{otherwise}, \end{cases}$$

(2)

where $c_r$ is the constant for the maximum rim height and $q_3$ indicates the coordinates of the intersection of the ellipse $f_3$ and the straight line crossing $(m, n)$.

The synthesized type V marine snow artifact is shown in Figs. 2(c) and (d), which are similar to the real versions in Figs. 4(c) and (d).

D. Synthesizing Images with Marine Snow

By using the marine snow models introduced in the previous subsections, we synthesize the marine snow with underwater images.

Let $U \in \mathbb{R}^{H \times W \times 3}$ be an original underwater image. First, we randomly select a target pixel coordinate $(k, l)$ to append the marine snow. Second, we randomly select the major and minor axes of the ellipses $a_i$ and $b_i$. In addition, the direction of the major axis is randomly set. Third, for each channel, type H or V marine snow is independently generated, where the parameters $c$ and $c_r$ are independently set. We denote an image with one marine snow artifact as $M$, where pixel values in $M$ are all-zero except the marine snow artifact. Finally, $U$ and $M$ are combined as follows to yield the synthesized image $I$:

$$I_{k,l,t} = \begin{cases} M_{k,l,t} + G_{k,l,t} & M_{k,l,t} \geq 0 \\ U_{k,l,t} & \text{otherwise} \end{cases}$$

(3)

where $G \in \mathbb{R}^{H \times W \times 3}$ is the Gaussian filtered version of $U$ that reflects blur owing to marine snow.

We iterate the process until a predetermined number of marine snow artifacts are appended.

IV. DATASET SPECIFICATIONS

In this section, we present the specifications of synthesized marine snow artifacts in the MSRB Dataset. It is designed for two tasks:

Task 1: Removal of small-sized marine snow artifacts.

Task 2: Removal of various-sized marine snow artifacts.

Clearly, Task 2 is more difficult than Task 1.

First, detailed parameters and setups shared in both tasks are presented. Second, we introduce MSR Tasks with corresponding synthesized images.

A. General Setup

Each sub-dataset corresponding to an MSR Task contains 2,300 training image pairs and 400 test image pairs, all having a pixel resolution of $384 \times 384$. An image pair contains one original underwater image and one image containing synthesized marine snow artifacts. All original images are collected from flickr\(^2\) under a Creative Commons Attribution-NonCommercial-ShareAlike 2.0 Generic (CC BY-NC-SA 2.0) License.

Each synthesized image contains $N$ marine snow particles where $N$ is chosen from $\mathcal{U}(100, 600)$, in which $\mathcal{U}(v_{\min}, v_{\max})$ is a discrete uniform distribution between $v_{\min}$ and $v_{\max}$. In each synthesized image, type H and V marine snow particles are randomly generated with a probability of 0.7 for type H and 0.3 for type V, according to our preliminary observations. The other representative parameters introduced in the previous section are listed in Table I. Most of the parameters are determined according to our preliminary observations of real underwater images with marine snow artifacts.

B. MSR Task 1: Removing Small-Sized Marine Snow Artifacts

We introduce the first MSR Task: Removal of small-sized marine snow artifacts. In Task 1, the maximum width/height of the artifacts is restricted to 6 pixels, which corresponds to roughly 1.6% of the image width/height. Task 1 is designed for underwater scenes where particles are relatively far from the photosensors. Because the task is relatively simple, the conventional MF approach is expected to be successfully applied (at least partially).

Examples of test image pairs in MSR Task 1 are shown in Fig. 5. As clearly visualized, the synthesized images have various sized (but small) marine snow artifacts.

\(^{2}\)https://www.flickr.com/
C. MSR Task 2: Removing Various-Sized Marine Snow Artifacts

The second MSR task is designed for underwater scenes containing particles in various distances. As a result, the synthesized images for Task 2 have small- and large-sized marine snow artifacts. For large-sized artifacts, we set the largest width/height of marine snow to 32 pixels, which corresponds to 8.3% compared to the image width. Furthermore, the probabilities of small- and large-sized artifacts are set to 0.7 and 0.3, respectively.

Training image examples for Task 2 are also shown in Fig. 5. The images have various-sized artifacts and the degraded areas
In the following, we show the benchmarking results of MSR as well as the real MSR.

V. MSR BENCHMARKING RESULTS

In this section, we present the first benchmarking performance for MSR Tasks 1 and 2. The methods used for the benchmarking are 1) MF [19], 2) adaptive MF [5], and 3) U-Net [29]. The kernel size of MFs is set to $3 \times 3$ or $5 \times 5$ pixels. Thanks to the MSRB Dataset, we can train a deep neural network for removing (synthesized) marine snow. For a benchmarking purpose, we train U-Net [29] using the dataset. The parameter settings are slightly different from the original.

The MSR results for Tasks 1 and 2 are shown in Figs. 6 and 7, respectively. As clearly observed, MF oversmoothes the images, which results in blur even for regions without marine snow artifacts. The adaptive MF suppresses blur; however, many artifacts are not removed mainly owing to the threshold-based algorithm. Furthermore, both MFs do not operate successfully for large marine snow artifacts in Task 2. By contrast, U-Net successfully removes marine snow for both tasks. Even for Task 2, it is able to suppress the large marine snow artifacts because of its multiresolution structure and the large number of trainable parameters.

For an objective comparison, we compute the average PSNRs and SSIMs over the test datasets. We summarize the results in Table II. The objective measure indicates that U-Net is superior to the MFs for both MSR tasks. Note that all MFs present smaller PSNRs/SSIMs than those of the synthesized images because they filtered areas without marine snow. By contrast, U-Net achieves a better objective quality than the synthesized images, which implies the effect of a deep neural network for MSR.

VI. REAL MARINE SNOW REMOVAL EXAMPLES

We construct the MSRB Dataset to mimic real marine snow artifacts. In this section, we show the real MSR results and present limitations of the dataset.

A. Real Marine Snow Removal

Fig. 8 shows MSR results for some real images. We use MF, adaptive MF, and U-Net as in the previous section. In addition, Deep WaveNet [30] is applied to validate the regular underwater enhancement model works for marine snow removal. U-Net is trained using MSRB Task 2 Dataset. The trained model of Deep WaveNet is taken from the original authors’ GitHub repository.

As clearly observed, MFs cannot suppress marine snow artifacts because of their various sizes and different transparencies. Deep WaveNet also fails to remove marine snow artifacts. That means marine snow cannot be suppressed by underwater image enhancement techniques alone. U-Net presents a better

| Method       | Task 1 PSNR | Task 1 SSIM | Task 2 PSNR | Task 2 SSIM |
|--------------|-------------|-------------|-------------|-------------|
| MF ($3 \times 3$) | 28.55       | 0.846       | 22.81       | 0.770       |
| MF ($5 \times 5$) | 25.98       | 0.711       | 21.93       | 0.645       |
| Adaptive MF ($3 \times 3$) | 29.88       | 0.910       | 23.35       | 0.842       |
| Adaptive MF ($5 \times 5$) | 28.08       | 0.861       | 22.83       | 0.794       |
| U-Net        | 36.82       | 0.978       | 30.95       | 0.932       |
| Synthesized image | 32.20       | 0.945       | 23.83       | 0.876       |
restoration image quality than MFs: It (partly) suppresses artifacts having both small and large sizes. This implies the effectiveness of deep-learning-based methods trained using the MSRB Dataset for real MSR. Note that thick marine snow artifacts remain in the restored images. The reason for this could be two-fold: First, U-Net is not specifically designed for MSR although we slightly customized its structure. A neural network designed specifically for MSR would improve the performance. Second, our proposed model classifies marine snow artifacts into two representative types. Marine snow artifacts not fitted to these types are not well removed. This limitation is described in the next section.

B. Limitations

Because our MSR benchmarking dataset is based on synthesizing marine snow artificially into real underwater images, some limitations exist. First, the dataset is designed for removing marine snow artifacts and does not aim at achieving other underwater image enhancements such as color correction. The design of a simultaneous restoration of MSR and color correction is an interesting area of future study.

Although U-Net trained using the MSRB Dataset presents promising results, the dataset itself is imperfect. Specifically, real MSR performances are occasionally limited if MSR artifacts are not fitted to our marine snow models. Examples are shown in Fig. 9, which compares real underwater images and their restoration results using U-Net. Typically, U-Net fails to suppress artifacts when marine snow artifacts are larger and denser than our MSR benchmarking dataset. Moreover, if the lighting conditions differ from the dataset, the current version of U-Net will fail. Through a future study, this can be improved by carefully updating the MSRB Dataset to reflect various underwater scenes.

VII. CONCLUSIONS

In this paper, the Marine Snow Removal Benchmarking Dataset was proposed. We mathematically modeled two representative marine snow artifacts and synthesized them in real underwater images. Two MSR tasks, the removal of small-sized and various-sized marine snow artifacts were also proposed. The first benchmarking results were also shown, which revealed the effectiveness of a deep neural network for MSR compared to median filter-based methods. MSR for real underwater images indicates that our MSRB Dataset contributes to real MSR problems. Our future studies will include improving the marine snow models as well as designing deep neural networks specifically designed for MSR.

REFERENCES

[1] Derya Akkaynak and Tali Treibitz. A revised underwater image formation model. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 6723–6732, 2018.

[2] Derya Akkaynak and Tali Treibitz. Sea-thru: A method for removing water from underwater images. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 1682–1691, 2019.

[3] Saeed Anwar and Chongyi Li. Diving deeper into underwater image enhancement: A survey. Signal Processing: Image Communication, 89:115978, Nov. 2020.

[4] Saeed Anwar, Chongyi Li, and Fatih Porikli. Deep Underwater Image Enhancement. arXiv:1807.03528, July 2018.

[5] Soma Banerjee, Gautam Sanyal, Shatadal Ghosh, Ranjit Ray, and Nath Shome, Shome. Elimination of Marine Snow effect from underwater
