Marine Snow Detection and Removal: Underwater Image Restoration using Background Modeling

Fahimeh Farhadifard
University of Rostock, Germany
fahimeh.farhadifard@igd-r.fraunhofer.de

Martin Radolko
University of Rostock, Germany
martin.radolko@igd-r.fraunhofer.de

Uwe Freiherr von Lukas
Fraunhofer IGD Rostock, University of Rostock, Germany
uwe.freiherr.von.lukas@igd-r.fraunhofer.de

ABSTRACT

It is a common problem that images captured underwater (UW) are corrupted by noise. This is due to the light absorption and scattering by the marine environment; therefore, the visibility distance is limited up to few meters. Despite blur, haze, low contrast, non-uniform lightening and color cast which occasionally are termed noise, additive noises, such as sensor noise, are the center of attention of denoising algorithms. However, visibility of UW scenes is distorted by another source termed marine snow. This signal not only distorts the scene visibility by its presence but also disturbs the performance of advanced image processing algorithms such as segmentation, classification or detection. In this article, we propose a new method that removes marine snow from successive frames of videos recorded UW. This method utilizes the characteristics of such a phenomenon and detects it in each frame. In the meanwhile, using a background modeling algorithm, a reference image is obtained. Employing this image as a training data, we learn some prior information of the scene and finally, using these priors together with an inpainting algorithm, marine snow is eliminated by restoring the scene behind the particles.

Keywords

Underwater Image Processing, Marine Snow, Background Model, Inpainting

1 INTRODUCTION

The growing interest in UW image processing lies in the poor performance of devices used to capture UW scenes. The major barrier is that light, unlike sound, is poorly propagated in water. This is explained by the propagation properties of light in water ([McG80, Wel69]). Light is exponentially attenuated while traveling in water. This is caused by two factors: light absorption and scattering, which leads to poor contrast, haze, blur and color cast.

- Light absorption reduces the light energy; therefore, colors drop one by one based on their wavelength (color cast). One can augment the visibility range by using artificial lightening; however,
- water reflects a significant fraction of light back to the camera before it even reaches the object in the scene. This so-called backscattering yields degraded contrast scene and a foggy appearance. Furthermore,
- a fraction of light reflects from the object to the camera with a small angle (forward scattering) which generally leads to a blurry image. Finally,
- organic and inorganic floating particles in water distort the scene visibility as an unwanted signal and are considered as noise, although, they belong to the scene.

As a result, visibility UW is limited at a distance of about twenty meters in clear water and five meters or less in turbid water [ABMK05b]. Naming distortions for UW imaging, the one which is not well-researched and mostly neglected from image processing algorithms, is the presence of floating particles. Although, in orders of magnitude these particles together with Backscatter have the greatest degradation factor [ABMK05b].

Marine snow is the term which is used for the macroscopic aggregates of detritus, dead material and dissolved organic matter floating in water. According to the properties of light propagation in water, smaller particles scatter the light more, thus, marine snow is one of the main sources of scattering (more specifically backscatter). Light reflection on marine snow creates white bright spots that lead to an inhomogeneous

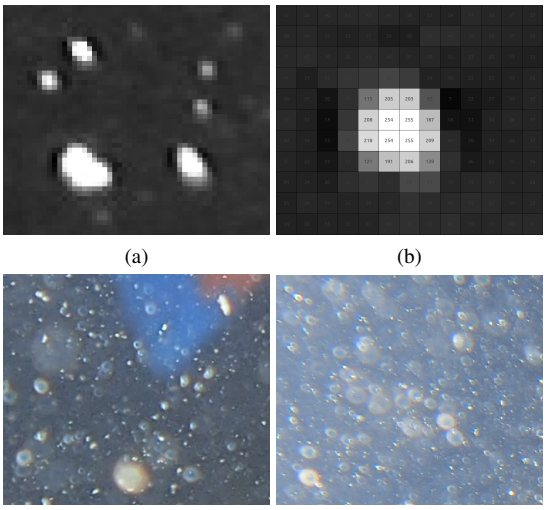

Figure 1: Illustration of physical characteristics of marine snow in acquired data. (a) particles with different sizes (3×3 to 20×20, (b) geometry of particles, (c) particles are present in different camera-scene depths (contrary to additive noise) and (d) strong light reflection of particles due to using artificial illumination.

medium [BG12]. Not only scattering and absorption are increased due to this phenomenon, but also it may appear dominant enough to reduce the scene perception (some examples are shown in Figure 1 (c) and (d)). Due to all the difficulties caused by marine snow, in this work, it is considered and treated as noise.

In this paper, we consider a novel approach to removing marine snow from frames of a video where the camera is assumed to be static. The information provided by the video sequence is used to eliminate marine snow from each individual frame. Our algorithms has three main steps, first, we employ our previously proposed background modeling algorithm [RG15] (which is based on the well-known Gaussian background modeling approach) and obtain an accurate model of the static components of the video. This model gives us the information about the background which is covered with the marine snow. Second, we detect the corrupted pixels based on our detection algorithm [FRvL17] and extract a mask which indicates the location of marine snow. Next, using the background model as a training data, some prior information about the scene are learned [RB05]. Finally, we employed the trained priors and the inpainting algorithm proposed by Roth and Black [RB05] together with the extracted mask, and eliminate marine snow by restoring the scene behind it with the most related prior information. Experiments show promising results where marine snow is almost completely removed and even small details are preserved.

The rest of this paper is structured as follows: in Section 2, we present a summary of the related works. Section 3 introduces marine snow and provides a short summary of its characteristics. Section 4.1 contains the explanation of background modeling method used to provide the training data [RG15]. In Section 4.2, we explain how to extract an accurate mask containing marine snow locations from a single frame. And at last, the inpainting algorithm in [RB05] is detailed in Section 4.3. Evaluation of the algorithm is provided in Section 5.

2 RELATED WORK

The popular approaches towards denoising consist of filtering [ABMK05a, LNHL15], wavelet decomposition and high-pass filtering [SZW11, PK10], a combination of curvelet and filtering [SSS13]. These methods assume that every kind of present noise including marine snow can be defined as one of the additive noises. Thus, salt & pepper, Gaussian and speckle noise are considered and with this assumption, authors provide a solution.

However, considering marine snow as an unwanted signal in UW images, these algorithms can not address eliminating of this phenomenon. This is due to their main assumptions (additive and single pixel noise) which do not match marine snow's characteristics. Marine snow is an object in the scene and has a structure of several pixels and covers the scene. Usually, these particles do not carry interesting information of the scene and therefore, are disturbing for image processing algorithms.

Banerjee et al. [BSG+14] proposed a probabilistic approach using median filtering to eliminate this phenomenon from single images. This approach checks the probability of the existence of marine snow in each patch. This is done by looking for high luminance pixels in a patch using a predefined threshold and calculating its probability as follows:

$$P(MS) = 1 - \frac{N_{HL}}{N} \tag{1}$$

where N_{HL} and N stand for the number of high luminance pixels and the total number of pixels in the current patch respectively. They consider a cross-checking to avoid misclassification of the true objects as marine snow. To this end, keeping the same center pixel, they increased the patch size by 2 (in both directions) and calculate the probability one more time. If the probability of having marine snow in the resized patch is still high (low number of high luminance pixels) then the center pixel is replaced by the median value of the local patch. The logic behind this is that they assume marine snow to have a structure of two or three pixels;

Figure 2: Comparison of different update schemes for the background modeling. In the top row are the first and 2000th frames of the Town Center Video. Second and the third rows correpond to the background models for the 2000th frame created with different updating mechanisms: the partial updating, a complete update, and GSM respectively.

therefore, if the probability of high luminance pixels increased it means that it is a bigger object which can not be marine snow.

However, this assumption does not hold always since usually marine snow, depending on the image resolution, have bigger structures (in our case it reaches to 20×20 pixels). Thus, considering it to have sizes bigger than three pixels, this criterion cannot differentiate between marine snow and other objects in the image. Increasing the patch size to take into account bigger sizes of marine snow may lead to a significantly blurred image. Moreover, this method does not use all the information provided in an image since it only considered gray scale image which could result in false detection of similar structures with different colors.

3 MARINE SNOW AND ITS CHARACTERISTICS

Decaying dead material and dissolved organic matter in water is referred to as marine snow since it is white and looks like snowflakes falling. These particles grow as they fall, some reaching several centimeters in diameter, this is due to aggregation of smaller particles. Thus, in an image, marine snow appears as white bright spots of different sizes and geometries randomly distributed in the image (Figure 1).

In view of Figure 1 (provided as an example), we could observe some physical characteristics of marine snow in captured images:

- it appears in different sizes depending on the image resolution. Usually between 3×3 to 20×20 pixels (Figure 1(a)).
- it can be roughly estimated as a Gaussian distribution in all directions, a high peak in the middle and lower intensities elsewhere proportional to the distance to the peak's location (Figure 1(b)).
- in contrary to additive noise, marine snow is present in all layers of a scene (considering the depth map of a scene consists of several layers) and can have a highly overlapped and non-uniform distribution over the image (Figure 1(c)).
- the most challenging fact about this phenomenon is that in the case of using an artificial light at the time of photography, it scatters the light to the camera and appears as circle shaped reflection (Figure 1(d)).

4 PROPOSED APPROACH

Having a video of reasonable length, we divide it into two parts. The first ~ 500 frames are used for training and the rest for testing. Although, the number of frames used for training can vary e.g. in the case of video in Figure 7 which is short (only 150 frames), we duplicate the training set by mirroring the order of training frames and conduct a bigger training set. The training frames are then used to learn a background model using Gaussian background modeling [RG15] (Section 4.1). Next, for each test frame, a mask containing marine snow locations is derived. The details of mask extraction are provided in Section 4.2. Once the background model of the scene and the mask corresponding to each test frame is available, the inpainting algorithm [RB05] is trained over the background model and recovers the scene behind marine snow using the corresponding mask (Section 4.3).

4.1 Gaussian Switch Model

It is a common practice to use a Gaussian distribution to model the color information of frame pixels in a video

sequence and extract one image which only contains the background. For this purpose, one can use the Mixture of Gaussian models [SG99, WBSP14], however, they are not ideal due to difficulty at unifying the different Gaussian distributions again. On the other hand, single Gaussian [WADP97] approaches lack accuracy. Thus, to keep the balance between accuracy and complexity, we use our Gaussian Switch Model (GSM) proposed in [RG15].

The idea behind this algorithm comes from the shortcoming of a single Gaussian approach which includes the information from foreground objects into the background model and corrupts it. Especially when there is a constant presence of many foreground objects. This can be solved by applying a *partial update*, this means that instead of updating the whole model, only the pixels that are classified as background are updated. Ideally, now only background information is included into the model which should lead to a more robust and precise model. For this, the segmentation of the current frame is computed by background subtraction before the model is updated with the information from this new frame. Then the segmentation can be used to update the background pixels and exclude foreground objects. In general, this improves the segmentation and stabilizes the model, but since the model is used to improve its updating process itself, a kind of self-fulfilling prophecy can occur.

An example of this is the presence of a foreground object in the initialization. This foreground object is a part of the model in the beginning and should slowly be overwritten with the background information during the updating process. However, when partial updating is applied, this usually does not happen because the actual background in that area will be marked as foreground; therefore, it does not get included into the model.

To still get the benefits from the partial updating without facing these problems, the GSM uses two Gaussians to model the background. The first Gaussian is partially updated and is taken as the background model and the second Gaussian is fully updated with every frame. The errors of the partially updated model can be discovered by a comparison between these two Gaussians since they always show the same characteristics:

- the means of the two Gaussians slowly diverge from each other as the Gaussian with the full update adapts to the new background and the other stays constant.
- for many successive frames a foreground object is detected at the same position.

If these characteristics are true for a specific pixel, the partial updated Gaussian for that pixel is overwritten with the values of the full updated Gaussian as it does not reflect the true background anymore. The represent version of the background model can then be simply extracted by taking the mean of the partially updated model for each pixel and color channel.

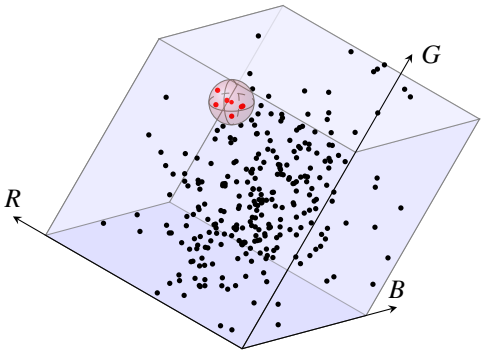

Figure 3: The pixels of current patch are visualized as points in RGB color space. The pink sphere demonstrates the search environment for the density calculation.

An example of the background modeling with the GSM compared to the partial and full update approaches can be seen in Figure 2 (the mean values of the Gaussians are displayed) where it is compared to the full and partial updating schemes on a video with many foreground objects. The parameters of the modeling are the same for all three methods and it can be seen that the complete update created a model which is very corrupted with the current foreground objects of the scene. The partial update eliminates this problem but many objects from the first frame can still be seen in the model there as they never get eliminated. The GSM can combine the advantages of both methods and can create an almost uncorrupted background model.

4.2 Mask Extraction

Knowing how marine snow appears in an image, we applied a detection algorithm to extract a mask indicating corrupted pixels in the image. This is done by looking for the pixels with the same characteristics as marine snow within the patches of an image. First, a rough detection of corrupted pixels is obtained and then, a voting algorithm conducts the final detection.

Marine snow is more visible and disturbing when the background is darker (lower intensity), although the particles are presented everywhere but they decrease the visibility of the scene especially when there is a higher contrast to the background. Thus, generally, the light reflection on the small particles are represented as bright spots in an image.

Thereby, the intensity of pixels within a patch is checked and a sudden high-intensity occurrence is marked as potential marine snow. A candidate pixel p has to satisfy the following inequality:

$$\|p-\mu(\Omega)\|_2^2 > W_1 \cdot \sigma(\Omega) \quad p \in \Omega, \tag{2}$$

here W_1 is an empirical weight, $\sigma(\Omega)$ is the standard deviation, and $\mu(\Omega)$ denotes the mean value of the local patch Ω. W_1 is defined heuristically and is affected by the image resolution, in our case where the data has a resolution about 850×478, $W_1 = 1.7$ is the best choice.

Next, we look for general outliers within the candidate pixels from the last step. This is done to distinguish between a high-intensity outlier and a high-intensity object's edge. For this, the idea of Gutzeit et al. [GOK+10] is employed. The RGB color space is considered as Euclidean space. The pixels within the current patch are then represented in this space and the density surrounding each candidate pixel is calculated. Figure 3 demonstrates this process. A sphere covering an area surrounding each high-intensity pixel defined in the last step is explored.

The number of pixels within this sphere

$$\#\{v \in \Omega \mid \quad \exists p \in \Omega : \quad \|p - v\|_2^2 < \sigma(\Omega)\}, \qquad (3)$$

together with the volume of the sphere and the overall number of pixels in the patch gives us the density. Here v and p are pixels in the local patch where $p \neq v$, and $\sigma(\Omega)$ is the standard deviation. The radius of the sphere is defined dynamically based on the weighted standard deviation of Ω to make the approach adaptive.

Another observation can be derived from marine snow characteristics: it mostly appears having high intensity and low saturation. Thus, by applying the following inequality the pixels with high saturation are discarded:

$$|p_c - p_l| < T \quad \forall c, l \in \{R, G, B\} \ \wedge \ c \neq l. \qquad (4)$$

Thereby, the candidate pixels are limited to have colors close to white by using a predefined threshold T (e.g. $T = 2$). All the pixel values that satisfy the aforementioned conditions are then discarded and the median value of the remaining pixel values within the local patch Ω is calculated. For now, all the eliminated values in this patch are replaced by this median value. The filtering is done in a copy version of the original image. This procedure, initial filtering, is repeated for the whole image.

The patches are extracted highly overlapped; this means each pixel can be in $n \times n$ possible patches except for the pixels at the border of the image with fewer possibilities ($n \times n$ is the patch size). Therefore, each pixel of the image could have been filtered in different patches accordingly, which results in having several filtered versions for a single pixel. The final decision about each pixel is then made by using a voting algorithm. If the majority of the filtered versions correspond to each pixel in the original frame indicate that it contains marine snow then, the location of that pixel is marked as noisy in a mask image. A mask image is a matrix, the same size as the original frame, whose pixel values are binary (one indicating marine snow and zero elsewhere).

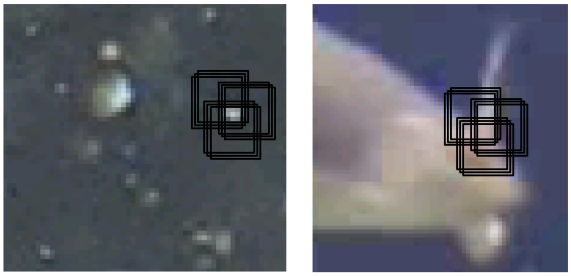

Figure 4: marine snow detection shows overlapping patches for marine snow (left) versus an object edge (right).

Figure 4 illustrates the condition with an example. This is applied on Laplacian pyramid of the image to detect marine snow with different sizes.

Once the final mask indicating marine snow locations is acquired, it is used to recover the denoised image using inpainting. At the end of this stage, we already can remove marine snow using filtering detected pixels. However, the results may suffer from smoothing the edges. In addition, in locations where the intensity of the image pixels varies (not only at the edges of the objects but also for example where color shades of a fish changes), the filtering can result in a wrong pixel value. An example of this situation is illustrated in the zoom-in presentation in Figure 6.

Thus, to improve the results, instead of filtering the image directly, we apply an inpainting algorithm which learns the most relevant priors to the test image by training over the background model of the same scene and restore the image accordingly.

4.3 Inpainting

Image inpainting is a useful application in several scenarios of image processing. It is used to fill in pixels which are missing in an image. Examples of inpainting in image manipulation include the removal of scratches on a photograph, unwanted occluding objects, superposed text, road-signs or publicity logos [ESQD05].

Generally, inpainting uses the information provided by the neighbors to fill in the missing pixels. However, whenever there is noise or any uncertainty, prior models of images such as depth maps, flow fields, *etc.* come into play. In our case where an object is considered as noise, prior information about the scene is advantageous. Therefore, we employ the inpainting algorithm proposed by Roth and Black [RB05]. This algorithm uses *Field of Experts* (FoE) to learn image priors from external data. We employ this algorithm rather than traditional inpainting algorithms so if the detection could not extract the particles precisely, the restoration will not be highly affected due to using direct neighborhoods.

FoE employs both sparse coding and Markov Random Field (MRF) to learn rich, generic prior models of any

class of images. Sparse coding provides an elegant and powerful way of learning prior distributions on small image patches. However, the result does not generalize to give a prior model for the whole image. This is where MRF provides not very rich but general prior information of the whole image.

The key idea behind FoE [RB05] is to extend MRF by modeling the local field potentials with learned bases. These bases capture important structural properties of images, respond to various edge and texture features. For this, they used the idea of the Product of Experts (PoE) framework [Hin02] and trained a model on a data set and develop a diffusion-like scheme that exploits the prior for approximate Bayesian inference.

For more insight, consider the pixels in an image be presented by nodes V in a graph $G = (V,E)$, where E stands for the edges connecting nodes. A rectangle region of $m \times m$ neighborhood connecting all nodes is defined, where every such neighborhood is centered on a node (pixel). The probability density of such a graphical model is defined as a Gibbs distribution:

$$p(x) = \frac{1}{Z} \exp^{(-\sum_k V_k(x_k))} \tag{5}$$

here x stands for an image and $V_k(x_k)$ is the potential function for clique x_k. They further assumed that the MRF is homogeneous which leads to translation invariance of MRF model. The potential Function V is then learned using training images. And as a result, the probability density of a full image is obtained as follows:

$$p(x) = \frac{1}{Z(\Theta)} \prod_k \prod_{i=1}^{N} \Phi_i(J_i^T x_k; \alpha_i), \tag{6}$$

where $Z(\Theta)$ is a normalizing function, J_i is a linear filter, Φ_i the experts and N stands for the number of experts. This model works for different image sizes, enjoys translation invariant property which is desirable for generic image priors and has few parameters which need to be learned. The parameters α_i and linear filters J_i are learned from the training images by maximizing its likelihood.

Once the parameters are learned, the inpainting algorithm propagates information using only FoE prior and refills the pixels iteratively by introducing an iteration index t and an update rate η as follows:

$$x^{(t+1)} = x^{(t)} + \eta M \sum_{i=1}^{N} J_i^- * \Psi_i(J_i * x^{(t)}) \tag{7}$$

let $\Psi_i(y) = \frac{\mathrm{d}}{\mathrm{d}y} \log \Phi_i(y; \alpha_i)$, J_i^- denotes the filter obtained by mirroring J_i around its center pixel [ZM97], and $*$ stands for convolution.

5 EXPERIMENTAL RESULTS AND DISCUSSIONS

We have performed our experiments on two different scenes. One is taken at Ozeaneum Stralsund (Figure 6), and the other one is courtesy of GEOMAR which is taken in the Black sea (Figure 7). For the first video (Figure 6), the first 500 frames are used to train the background model and the rest (50 frames) are added to the testing set for evaluation of the algorithm. Second video (Figure 7) is shorter; therefore, only 150 frames are available for training (and 5 frames for testing). For this video, we have duplicated the training frames to expand the training set for a better result. The background models of both scenes are obtained using GSM background modeling algorithm [RFvL16] (an example: Figure 5). Once the background model is available, we applied [RB05] to learn the FoE priors. For each frame in the testing set, we obtain a mask containing marine snow locations by applying the method explained in section 4.2. Finally, using the inpainting algorithm and mask together with the priors, the scene behind marine snow is recovered.

To quantitatively evaluate our algorithm, we need ground truth data which is not available in our case. Therefore, we have provided simulated frames. To this end, one simple approach would be to generate a salt and pepper noise on the image. However, as it was discussed before, this model does not take into account various physical parameters such as the effect of water absorption and scattering on the signal backscattered by the particles, the size, and shape of the particles or the defocus effect. Boffety and Galland [BG12] consider most of these properties and proposed a method to model this phenomenon by assuming that these particles behave like white Lambertian scatters. However, it still does not cover different geometries of marine snow and gives an artificial look to the image.

Thus, we have employed a different strategy. First, marine snow is extracted from the a test frame of each scene. This is done by human experts where the particles are manually extracted with pixel accuracy. Then, we have restored the scene behind marine snow with information of the neighborhood pixels and frames and conduct a ground-truth image. Once this image is available, we have placed the extracted marine snow randomly in the frame. This simulated image together with the ground truth image is then used to evaluate the performance of our proposed method. This way, we obtain a simulated data with a very natural look where marine snow has the most accurate model and is highly correlated to the real frames.

Our result is compared to the method in [BSG$^+$14], the result by directly filtering marine snow using the extracted mask (explained at the end of Section 4.2), our proposed algorithm when inpainting is trained on an arbitrary training set of the same class (UW images), and finally our proposed algorithm when inpainting is trained using the background model. Comparison is done via PSNR and MSE calculation (Table 1).

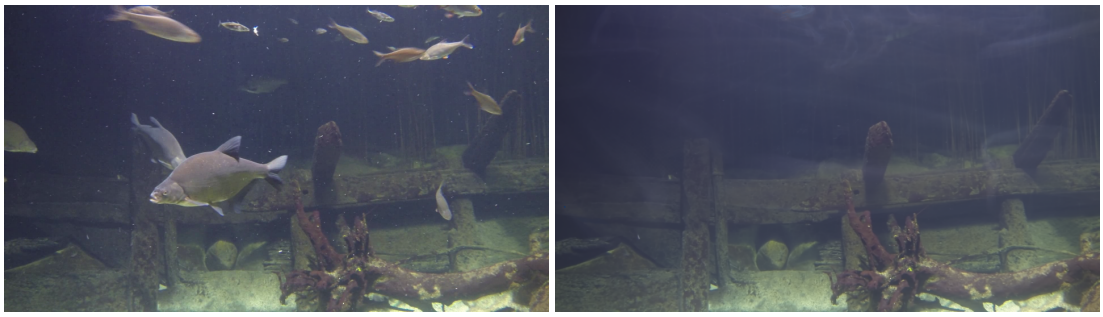

Figure 5: Results of the GSM background modeling on an UW video. On the left is the original frame of the video and on the right the background model is depicted.

In view of Table 1, one notices that filtering marine snow using the extracted mask has already succeeded by about 4 and 6.3 dB improvement on the results of [BSG$^+$14] for simulated images of scene 1 and 2 respectively. The advantage of using inpainting algorithm together with the extracted mask has been proved by achieving further improvement of 1.2 and 3.4 dB respectively. It can be seen that training data for the inpainting algorithm plays an effective role, where the result of the proposed algorithm differs when the training data changes. When the training data has a high correlation with the input image, the algorithm can achieve about 0.2 dB and 0.5 dB improvement, for the first and the second scene respectively, compared to the situation where a set of arbitrary images of the same class (UW images) are used.

Figures 6 and 7 illustrate the qualitative results corresponding to the table 1. The improvement over [BSG$^+$14] is clear, where the edges are smoothed and marine snow is not removed completely. Filtering the image using the extracted mask has already succeeded to remove marine snow effectively without smoothing the edges too much. However, a closer look demonstrates the shortcoming of this approach. It fails at removing marine snow correctly where it lies on the edges of objects or where there is no edge but the intensity values change (e.g. color shades of a fish). It happens due to the fact that median value of different candidate patches may not exactly match the true value. An example can be seen in the zoom-in presentation of Figure 6. In these situations, inpainting provides smoother results because it does not use the neighborhood pixels directly but the learned priors of the whole scene.

PSNR Values		
Approach	Scene 1	Scene 2
[BSG$^+$14]	43.6504	37.4235
Mask + Filtering Only	47.4818	43.7806
P. A. without BG model	48.6403	47.1723
P. A. with BG model	48.8275	47.6833
MSE Values		
[BSG$^+$14]	1.6750	3.4306
Mask + Filtering Only	1.0776	1.6501
P. A. without BG model	0.9430	1.1167
P. A. with BG model	0.9229	1.0529

Table 1: Evaluation using PSNR and MSE values.

6 CONCLUSION

In this paper, we have proposed an algorithm to eliminate marine snow from UW videos where the camera is static. The approach has three main steps; first, our background modeling algorithm [RG15] has been employed to extract an accurate model of the static components of a video. Second, we have detected marine particles in each individual test frame and a mask containing marine snow locations is derived. Finally, the background model is used as training data for an inpainting algorithm to extract the generic distribution of the scene which is then used together with the mask to restore the information behind the marine snow.

The results have illustrated the success of the proposed algorithm at eliminating marine snow. Simple filtering using the extracted mask has shown superior to the results of [BSG$^+$14] both quantitatively and qualitatively. In addition, employing inpainting has enhanced the results by restoring the image more accurately.

However, there is still space for improvement especially in extracting the mask. If the mask is not accurate enough to cover marine snow completely, inpaining may rebuild it back. Furthermore, when marine snow lies at the edge of two regions with low and high luminance, the algorithm may not be able to detect it. This is due to our first assumption that a sudden high luminance occurrence in a patch is a candidate to be marine snow.

7 ACKNOWLEDGMENTS

This research has been supported by the German Federal State of Mecklenburg-Western Pomerania and the European Social Fund under grant ESF/IV-BM-B35-0006/12.

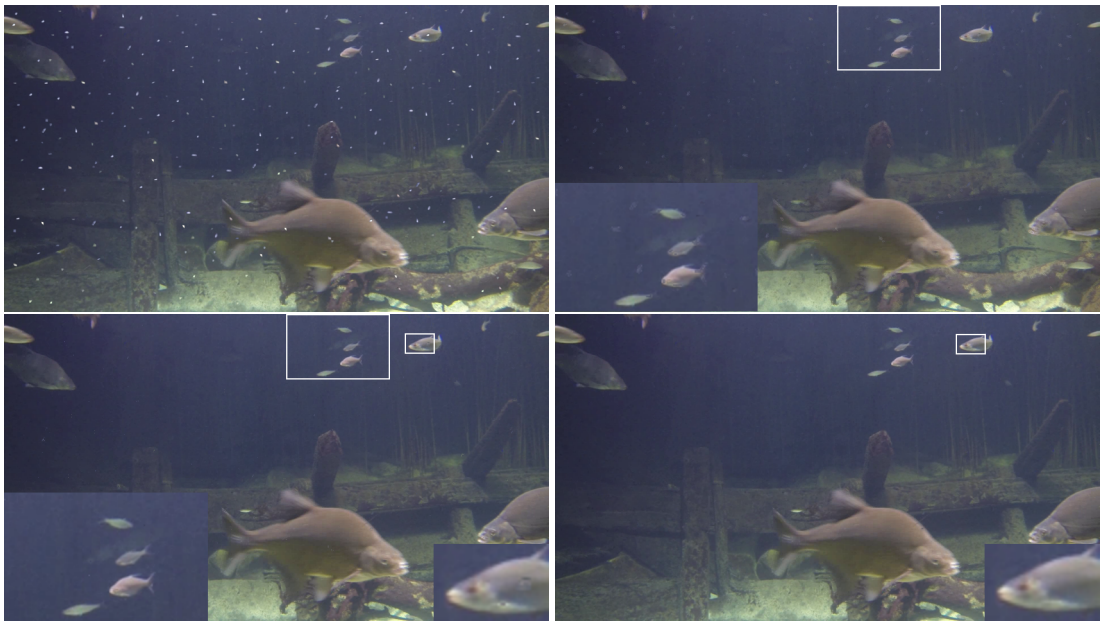

Figure 6: Results of marine snow removal for a frame of the first video. From top to bottom and left to right: input corrupted image, result of [BSG$^+$14], result of filtering using the proposed mask and final result of the proposed algorithm. Marked areas illustrate the improvement of the proposed method on previous work.

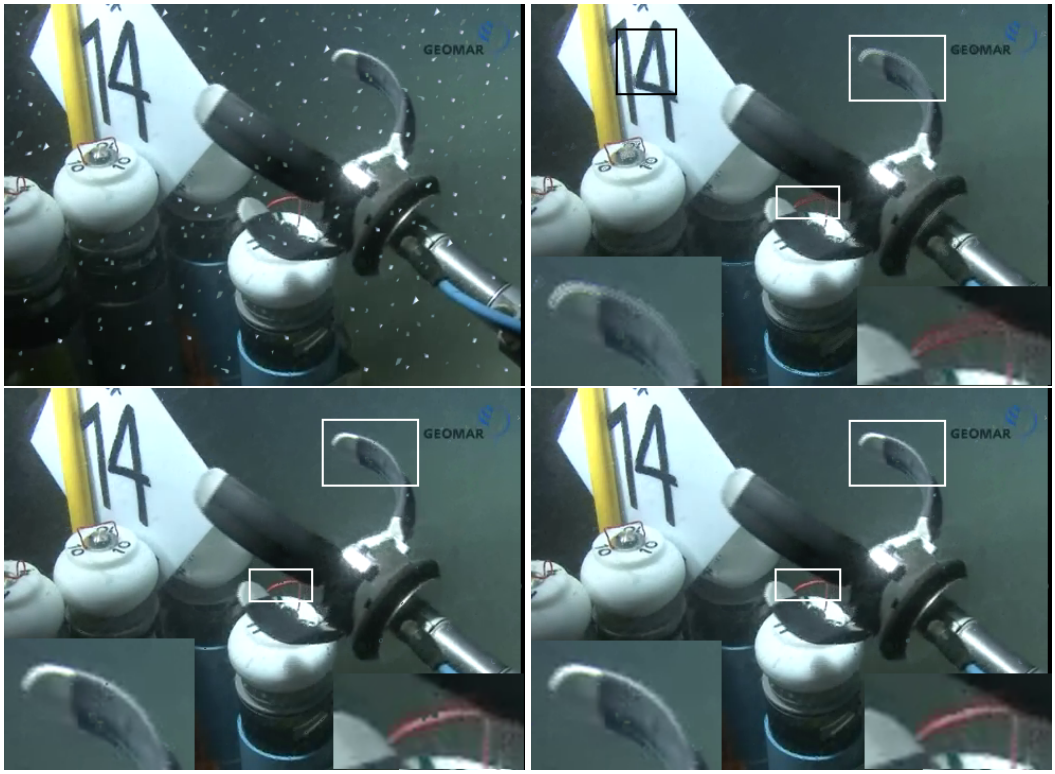

Figure 7: Results of marine snow removal for a simulated frame of the second video. From left to right and top to bottom: input corrupted image, result of [BSG$^+$14], result of filtering using the proposed mask and final result of the proposed algorithm. The marked areas show some examples of [BSG$^+$14] failure which are improved using the proposed method. Full video of this scene is courtesy of JAGO-Team, GEOMAR Kiel, Germany.

8 REFERENCES

[ABMK05a] A. Arnold-Bos, J. P. Malkasse, and G. Kervern. A preprocessing framework for automatic underwater images denoising. In *European Conference on Propagation and Systems*, 2005.

[ABMK05b] A. Arnold-Bos, J. P. Malkasse, and G. Kervern. Towards a model-free denoising of underwater optical images. In *Europe Oceans 2005*, pages 527–532, June 2005.

[BG12] M. Boffety and F. Galland. Phenomenological marine snow model for optical underwater image simulation: Applications to color restoration. In *OCEANS, 2012-Yeosu*, pages 1–6. IEEE, 2012.

[BSG+14] S. Banerjee, G. Sanyal, S. Ghosh, R. Ray, and S. N. Shome. Elimination of marine snow effect from underwater image - an adaptive probabilistic approach. In *Electrical, Electronics and Computer Science (SCEECS), 2014 IEEE Students' Conference on*, pages 1–4, 2014.

[ESQD05] M. Elad, J. L. Starck, P. Querre, and D. L. Donoho. Simultaneous cartoon and texture image inpainting using morphological component analysis (mca). *Applied and Computational Harmonic Analysis*, pages 340–358, 2005.

[FRvL17] F. Farhadifard, M. Radolko, and U. F. von Lukas. Single image marine snow removal based on a supervised median filtering scheme. to appear in 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications *VISAPP 2017*, 2017.

[GOK+10] E. Gutzeit, S. Ohl, A. Kuijper, J. Voskamp, and B. Urban. Setting graph cut weights for automatic foreground extraction in wood log images. In *VISAPP 2010*, pages 60–67, 2010.

[Hin02] G. E. Hinton. Training products of experts by minimizing contrastive divergence. *Neural computation*, 14(8):1771–1800, 2002.

[LNHL15] X Liu, R. Nian, B. He, and A. Lendasse. A rapid weighted median filter based on saliency region for underwater image denoising. In *OCEANS 2015 - MTS/IEEE Washington*, pages 1–4, Oct 2015.

[McG80] B. L. McGlamery. A computer model for underwater camera systems. volume 0208, pages 221–231, 1980.

[PK10] C. J. Prabhakar and P. P. Kumar. Underwater image denoising using adaptive wavelet subband thresholding. In *Signal and Image Processing (ICSIP), 2010 International Conference on*, pages 322–327. IEEE, 2010.

[RB05] S. Roth and M. J. Black. Fields of experts: A framework for learning image priors. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, pages 860–867, 2005.

[RFvL16] M. Radolko, F. Farhadifard, and U. F. von Lukas. Dataset on underwater change detection. 2016.

[RG15] M. Radolko and E. Gutzeit. Video segmentation via a gaussian switch background-model and higher order markov random fields. In *Proceedings of the 10th International Conference on Computer Vision Theory and Applications* VISAPP 2015, pages 537–544, 2015.

[SG99] C. Stauffer and W.E.L. Grimson. Adaptive background mixture models for real-time tracking. In *Proceedings 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Vol. Two*, pages 246–252. IEEE Computer Society Press, June 1999.

[SSS13] M. Shanmugasundaram, S. Sukumaran, and N. Shanmugavadivu. Fusion based denoise-engine for underwater images using curvelet transform. In *Advances in Computing, Communications and Informatics (ICACCI), 2013 International Conference on*, pages 941–946, 2013.

[SZW11] F. Sun, X. Zhang, and G. Wang. An approach for underwater image denoising via wavelet decomposition and high-pass filter. In *Intelligent Computation Technology and Automation (ICICTA), 2011 International Conference on*, pages 417–420. IEEE, 2011.

[WADP97] C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland. Pfinder: Real-time tracking of the human body. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19:780–785, 1997.

[WBSP14] R. Wang, F. Bunyak, G. Seetharaman, and K. Palaniappan. Static and moving object detection using flux tensor with split gaussian models. In *2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 420–424, June 2014.

[Wel69] W. H. Wells. Loss of resolution in water as a result of multiple small-angle scattering. *JOSA*, 59(6):686–691, 1969.

[ZM97] S. C. Zhu and D. Mumford. Prior learning and gibbs reaction-diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(11):1236–1250, 1997.