A Dataset to Evaluate Underwater Image Restoration Methods

Amanda Duarte, Felipe Codevilla, Joel De O. Gaya, Silvia S. C. Botelho
Federal University of Rio Grande
Center of Computational Sciences (C3)
Rio Grande, Brazil
Email: a.duarte@furg.br

***Abstract*—Image restoration methods have been made to repair images that have some kind of degradation. Most of these methods are designed to deal with the degradation caused by the over-land effects. However, when the images was captured in underwater environments, there are different properties that can degrade the image in unusual ways. This work aims to evaluate how the popular image restoration methods behaves when applied in underwater images with the presence of turbidity in the water. For this, we propose a dataset where we are able to control the amount of image degradation due to underwater properties on a scenario with 3D objects that represents the seabed characteristics. After that, we evaluate the restoration of these methods and their behavior through the image degradation due to turbidity.**

I. Introduction

Optical images captured in underwater environment scenes, normally, lack of visual quality. Those environments have generally large numbers of suspended particles in the medium that causes "haziness" on the captured image, here called *turbidity*. When the light rays propagates on underwater environment, it interacts with the suspended particles being both *scattered* and *absorbed*. These phenomenas reduce the amount of image information culminating into a degraded version of the scene signal.

Underwater images are important on many applications such as: 3D reconstruction of scenes [1], coral image classification [2] [3] or robot navigation [4] [5]. However, frequently the raw data is not sufficient to sustain those applications. Thus, image processing algorithms are often used to increase the general quality of underwater images [6].

To recover general image visibility on underwater images, general enhancing methods can be used, *e.g*, contrast stretching, white balance, etc. However, besides producing some visually satisfying results, the enhancement methods do not invest into recovering the non-degraded signal properties. An alternative to this is the restoration methods. These methods are designed to recover the degraded image by removing the degradation relying on a physical model of image formation.

Independently of the method used to process underwater images, image quality evaluation is a hard matter. This matter makes the development of better restoration algorithms a hard task, since there is unknown way to accurately compare restoration algorithms. Usually one can evaluate the quality of images depending on the amount of the original, non-degraded, signal information [7]. This can be divided into three categories *i.* evaluation based on a noise free version of the image, *ii.* evaluation based on some statistical information of the noise free image, *iii.* evaluation based on just the degraded image. Underwater restoration algorithms can only fall on the third category of image evaluation. This happens since the degraded underwater image do not have the reference image, *i.e.* the same underwater image without degradation, as a way to compare.

In this paper we propose a way to access a reference image by producing an controlled underwater environment. By using this reference image, we are able to find the actual error obtained by restoration methods and, thus, accurately conclude about their efficiency since the algorithms are evaluated in function of the image degradation. This creates the possibility to evaluate underwater image processing algorithms under the category *i*, increasing the precision of the comparison.

The reference image is achieve by proposing a new version of the TURBID Dataset [8] called here as 3D TURBID, that contains different levels of image degradation on a planned seabed scenario with 3D objects, containing all different aspects found in a real sea floor.

After that, we evaluate the behaviour of the most popular image restoration and enhancement methods in the proposed Dataset. With this , we were able to observe how each algorithm behaves through these images, in order to determine which one presents better robustness through the increasing degradation.

The paper is presented as follows. The section II presents some image processing algorithms existent on the literature. The section III presents the description of the 3D TURBID dataset, and also some examples of the captured images. Section IV shows the chosen algorithms to be evaluated and how we get the restored image quality. Section V presents the results of the evaluation procedure. Finally the section VI presents the paper conclusions.

II. Underwater Image Restoration/Enhancement Algorithms

To recover the image visibility in degraded underwater images, general enhancing methods such as a Contrast Enhancement [9], Bilateral Filter [10] and Color Constancy [11] are used. Also a fusion of those proposed by Bazeille [12] and *Ancuti et al* [6].

Bazeille propose an application of a pre-processing filter for underwater images. It is an automatic algorithm that aims to improve the quality of the segmented image and reduces the underwater perturbations. This methods is composed by successive independent processing steps that try to correct the non uniform illumination, reduce noise, enhance contrast and adjust the color.

Some problems can happen when you apply several filters on a image. First of all, the contributions of each filter are not used properly. Furthermore, when you enhance the contrast, this can increase the noise. To avoid these problems, *Ancuti et al* propose a fusion strategy. With that, images with different filters are considered. Based on that, a new enhance image is derived from a weight measure considering only the degraded versions of the image.

However, the enhancement methods do not invest into recover the image properties as a hazy-free image. With this, most of the restoration methods are designed to recover the degraded image by removing the degradation. For that, they are relying on a physical model of image formation defined as

$$T_i(x) = J(x)e^{-c_i d(x)} + A(1 - e^{-c_i d(x)}), \quad (1)$$

Where $J(x)$ is the signal with no degradation attenuated by $e^{-c_i d(x)}$, where is called *transmission*. The transmission can be understood as the turbidity portion that determines the amount of degradation in each part of the image in function of the distance from the object to the camera. The component A is the constant ambient light, that represents the color and radiance of the media. This constant can be altered by the depth and the characteristics of the environment.

To obtain a color recovering and a haze removal the authors starts solving this equation, which is a hard task since the transmission and the ambient light need to be estimated correctly. Being that one of the hardest task in the restoration problem since each patch of the image has lot of information ambiguity. For that, some properties of a image without the ambient interference should have need to be assumed. These properties are usually image priors or assumptions that are used to indicate the amount of turbidity each patch of the image has.

One of the main restoration method using this model was propose by [13]. This method was based on the Dark Channel Prior, where the author presents that minimum value of the image channels in a patch indicates the transmission. Using that, the author estimates the thickness of the haze and recover a haze-free image. This method have been adapted several times for underwater environments e.g. [14] [15] [16] and [17]. However all those adaptations do not consider the large range of colors that exist in underwater environments by assuming some specific condition such as the Red Channel Absorption [16]. This method aims to recover the short wavelengths of the colors leading to a recovery of the lost contrast.

Also, a general participative media method was proposed by *Codevilla et al* [18]. The authors proposed a joining of two different priors local contrast and color as a effective approach for image restoration.

A. Algorithms Evaluation

Besides the estimation problem, after recover the degraded image, one of the main problem is associated with how to evaluate the quality of the restoration obtained. Considering that the most of the degraded underwater image do not have the reference image, *i.e.* the same underwater image without degradation, to compare with the restored image.

Now a days, the quality evaluation of a restoration is made by subjective analysis, which can be tendentious. One example of that is two different people may have different opinion about the same image restoration.

Simulated images by computer are also used to evaluate the restoration. For that, a non degraded image is simulated to be a image taken in a underwater environment by using computer rendering techniques. However, this techniques are not able to simulate the complexity of the phenomena.

With this evaluation we believe that besides the finding of the actual error presents in the images restoration, the future methods can be develop to improve these errors and the final restoration.

III. The 3D TURBID Dataset

The *light attenuation* in the underwater environments is the gradual loss of the light rays intensity through the water-column. It is controlled by the amount and kind of particles that are dissolved and suspended in the water. This phenomena happens by two process, *absorption* and *scattering* [19]. *Absorption* fully removes the light rays while the *scattering* changes the direction of the light propagation. With that, when you imaged a scene in those environments some specific degradation in the image formation can happen.

Also other phenomenas can happen such as *Forward scattering* and *backscattering*. Forward scattering happens when the light rays coming from the scene are scattered in small angles creating a blurry effect on the image. This effect, however, has a small contribution to the total image degradation and it is frequently discarded [20]. The backscattering happens when the information of the sources from outside the captured scene scatters over the image plane creating a characteristic veil on the image which reduces the contrast.

In this context, a common term related with this phenomenas are *turbidity*. We consider it as the scattering and attenuation of the light which cause the loss of water transparency and clarity that causes the "haziness" on the captured image. We define *turbid image* as images where the visibility of the imaged scene is degraded by the turbidity.

To simulate these phenomenas becomes a challenge since it happens due to specific particles and property present in the oceans, rivers, lakes, etc. An study by [21] pointed out that the whole milk has a higher size of particles that induces a lot of wide angle scattering, increasing the *backscattering* effect.

Another challenge is related with to reproduce a untouched seabed in a controlled space with the specific underwater properties. It is important since we cannot take a small part of a real sea floor of a underwater environment to evaluate methods that will be used there.

As far as we know, only two experiments reproduce the underwater image degradation aspects in a controlled way [22][8]. Even so, the first one, used just a small set of structures to represent the seabed environments that not provided sufficient characteristics of a sea floor. The second one called TURBID, provide more information about the structures and characteristics with real seabed images, but it was printed in a pad resulting some noise unwanted from printing issues. Besides that, this experiment is not able to consider different distances to the camera, therefore it was difficult to validate the algorithms which depend on varying the distance.

The TURBID dataset was a initial dataset proposal for the algorithms evaluation procedure that will be present in the next section. We propose it using three different high quality printed real scenes previously photographed at the Bahamas. These images was called here as Photo1, Photo2 and Photo3. These scenes contains structures of the underwater floor and some human made objects. The pictures were re-photographed inside a 1000 litres tank made of plastic, illuminated by two 30 watts fluorescent light strips. As the image capture device we used a static Go Pro Hero3 Black Edition with 12 mega pixels (3000x4000) of resolution.

For that, we first photograph 30 images in a clean water. After that, the turbidity and consequently the amount of degradation are increased in a controlled way by successively adding whole milk into the water tank. This addition of milk began with 5ml to 190ml. We tested the amount of milk previously to obtain the required amount of turbidity. It was repeated 19 times with different amount of milk producing the different levels of turbidity. For each milk concentration, we photographed 30 images with 10 seconds delay between the shots, to avoid the illumination variance. To produce the set of the images with different levels of turbidity, we first calculated the average of the 30 first images taken in the clear water, it is our reference image I_0 (with no degradation). After that, we also calculate the average of the image taken in the same amount of turbidity producing the images I_1 ... I_19.

For the 3D TURBID[1] dataset just some adjustments was made. The capture device was upgraded to a Go Pro Hero3+ Black Edition, the tank was change by one made with glass, and the illumination was change by two Light-Emitting Diode (LED) lamps placed inside a softbox made with reflector and diffuser materials to obtain a continuous and uniform light. The main structure of the set is present in the image 1.

The photographed scene was also modify to a planned seabed scenario with 3D objects and some different aspects found in a sea floor. This scene contains stones to characterize the seafloor, decorations that imitate corals, seashells, marine rocks and other objects made by man to characterize a real seabed.

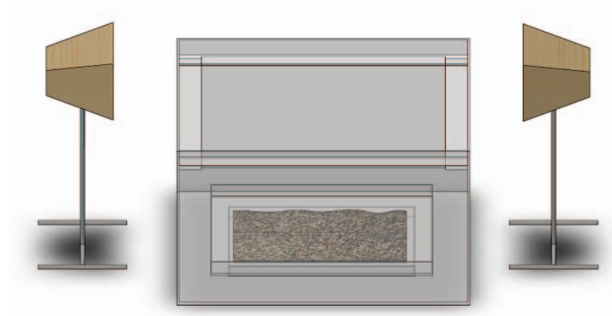


Fig. 1. The main structure of the experimental proposed. It is composed by a 1000 tank, two LEDs lamps and a planned scenario with 3D objects

Some examples of images obtained by this experiment can be seen at figure 2. The first row shows the Photo1, the second shows the Photo3, the third shows the Photo3 and the fourth shows the 3D TURBID version. These images also shown four different levels of degradation with their amount of milk added, the first column shows the reference images (clean image) I_0, the second shows the fourth level of degradation I_4, the third shows the tenth level of degradation I_{10} and the last column shows the sixteenth level of degradation I_{16}.

The dataset and methodology that we propose can be useful not only to evaluate image restoration methods, but also to test many vision algorithms that need to be tested in multiple levels of turbidity. The main advantage of this dataset is the presence of the clean image that can be used as a ground truth for underwater restoration applications.

IV. Evaluation Procedure

A. Chosen Algorithms

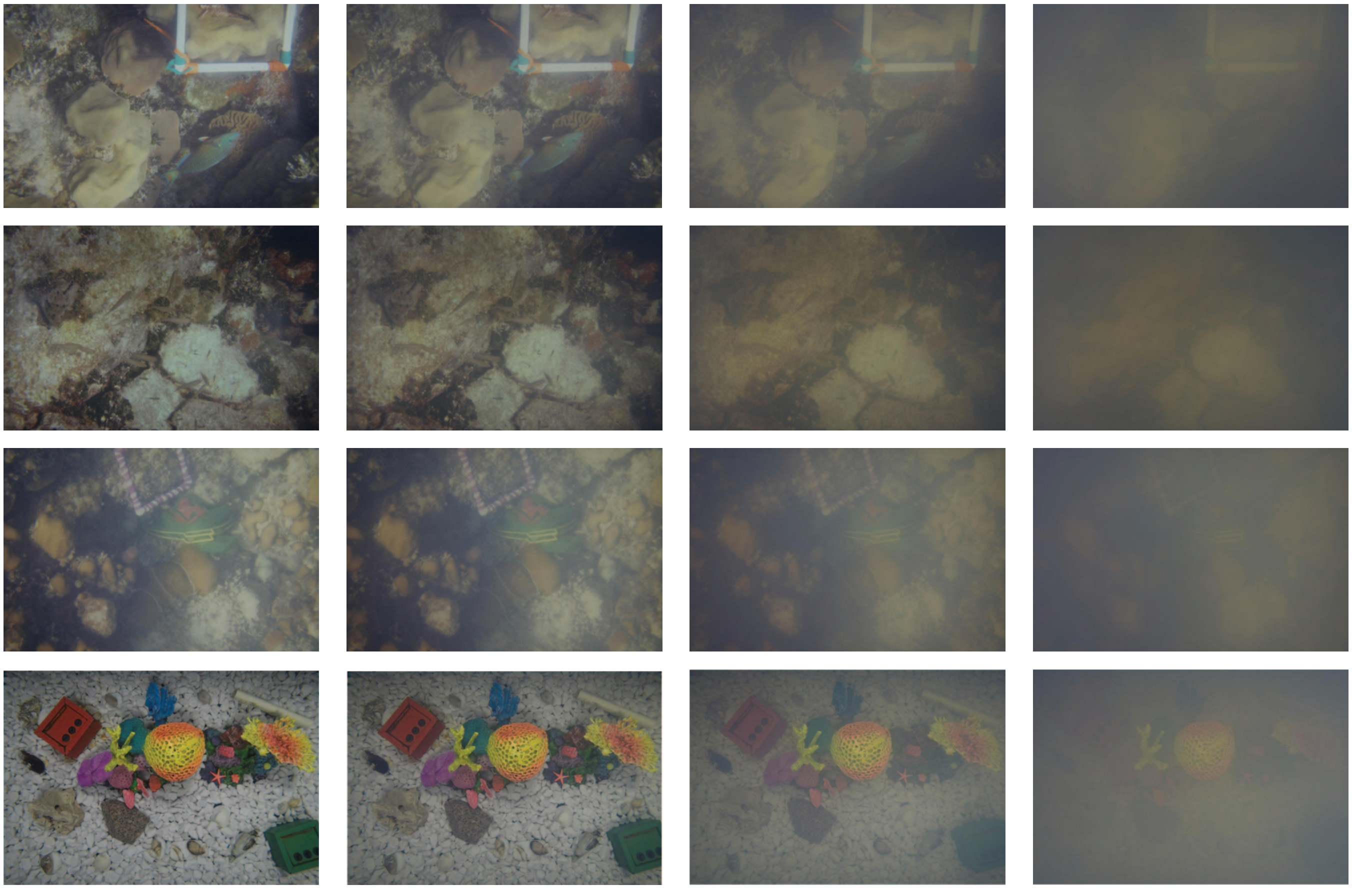
We decided to cover the most popular methods present in the state-of-the-art and new ones containing different paradigms. We choose to evaluate the the Dark Channel Prior (DCP)[13], the Red Channel Prior (RCP) [16], the method proposed by [6] (*Ancuti et al*), a General Participative Media Restoration Method proposed by [18] (*Codevilla et al*), a general enhancing method CLAHE [9], and also the white balance Shades-of-Gray [23].

In order to obtain the restored images using the TURDID dataset images, we use for [13] [16] and [18] C++ implementations using OpenCV, and to obtain the [6], [9] and [23] results we used Matlab implementations. We reproduce these methods since it was not available from the original authors. To get the most approximate results presented by the authors in each paper, we get the images available for each method in the original papers, and try to reproduce the equal results. To promote a fair comparison, we avoided the estimation of the airligth constant for the restoration, setting a fixed value.

B. Image Quality Evaluation

One problem faced when we work with underwater turbid images is the lack of a good technique that is able to evaluate

[1]The dataset is available at: https://mega.nz/!ZkgkSDJZ!ES9LgsZz0oUMnsfQyzB-KgxamNx9KHrxAGyNvFjCwt8



(m) Clear(I_0)) (n) I_4 - 20 ml (o) I_{10} - 58 ml (p) I_{16} - 110 ml

Fig. 2. Some examples of the four different images obtained by the TURBID dataset. First row shows Photo1, second shows Photo2, third shows Photo3, and the fourth shows the 3D TURBID version. The first column shows the reference image I_0, the second shows I_4, the third shows I_{10} and the last one shows the sixteenth I_{16} and their amount of milk addition.

the quality of the image. The quality of a turbid image is understood as the imaged scene visibility, called as *visual clarity* by [22]. To quantify this visual clarity of the turbid images obtained by the TURBID dataset the best way that we find was using a index proposed by [22] defined from the Structure SIMilarity Index (SSIM) [24], called *Structural Degradation Index* (SDI). It was proposed as a more intuitive and easier to interpret way to expose the SSIM index. It was defined as

$$SDI = 100(1 - SSIM) \tag{2}$$

With this, we calculated the SDI between the reference image and each image that represent the different levels of degradation in the original set (without any kind of restoration). With the SDI index we can see a increasing integer scale for the image degradation that ranges from 0 in the reference image and 10 in the most degraded image (with almost no visibility). Using that, we can say that when the image degradation increase it leads to decrease the image similarity with its image reference.

After calculating the SDI between the reference image and the degraded images, the algorithms were evaluated by computing the Mean Square Error (MSE) between the reference image of the original set and the 20 restored image for each chosen method. Each set of restored image by different methods was plot as a different line. In this plot we can observe how each method behaves when the degradation of the image increase and compare the behavior of those methods.

V. Results

The Figures 3, 4, 5 and 6 shows the MSE in function of the SDI index, respectively for Photo 1 , Photo 2, Photo 3 and Photo 3D. The MSE was measured between the reference image (I_0) and each image restored by different algorithms ($I_1 ... I_{20}$). We also show the MSE with the degraded images as a comparison. Each set, corresponding to different restoration/enhancement algorithms, was plot as a different line. It is important to note that when the error is below to the degraded images line (represented here by the blue line) it means that the method performed an effective restoration since the image became closer to the image with no degradation (the reference image I_0).

For all cases, CLAHE performs better when it was applied in low levels of turbidity. CLAHE is surpassed by *Codevilla* *et al* in higher levels. DCP, RCP and *Ancuti et al* presents a good behaviour when the images have a high degradation by the turbidity. Also, both DCP and RCP tends to further degrade the image when the turbidity is low. The same happens with *Codevilla et al* and *Ancuti et al* when the degradation is even lower.

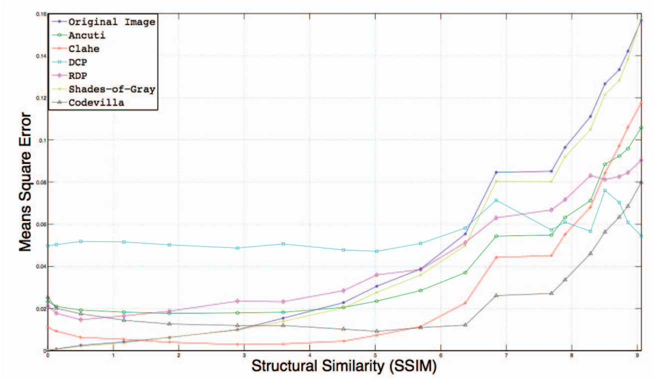


Fig. 3. The MSE between the Clean Image (I_0) and each restored Image in function of the SSIM in the dataset TURBID Photo1.

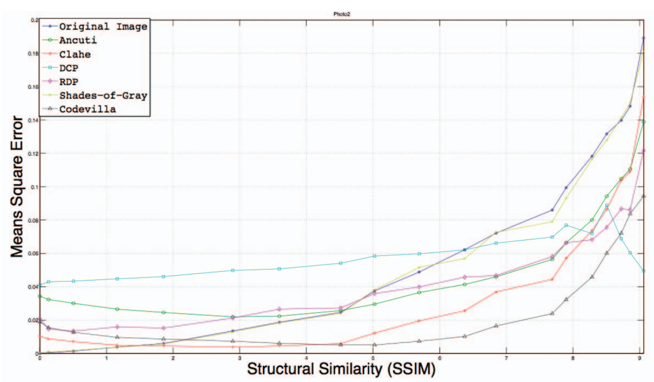


Fig. 4. The MSE between the Clean Image (I_0) and each restored Image in function of the SSIM in the dataset TURBID Photo2.

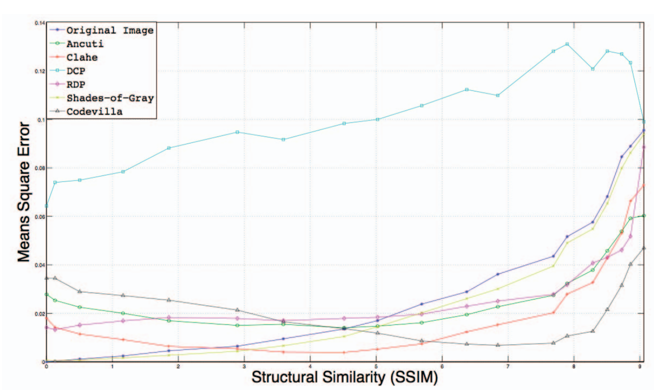


Fig. 5. The MSE between the Clean Image (I_0) and each restored Image in function of the SSIM in the dataset TURBID Photo3.

With this evaluation we can see that there is a clear difference between the restoration and simpler enhancement

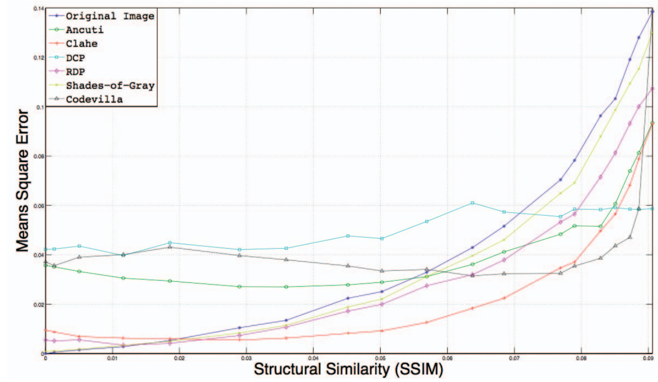


Fig. 6. The MSE between the Clean Image (I_0) and each restored Image in function of the structural similarity (SSIM) in the dataset 3D TURBID.

methods. All restoration methods DCP, RDP and *Codevilla et al* are based on Priors. We show that the estimation of these priors need certainly level of turbidity to be estimate correctly. Most of the problems of these methods are associate with this estimation. In low levels of turbidity, when these priors are not estimated correctly, they tend to include non-signal information. In the other hand, in higher levels of degradation, when the visibility is poor and the priors are estimated correctly, they present a good behaviour. *Ancuti et al* is not based on prior, but the results also shows that this method needs a certain level of turbidity to correctly measure the weights present in the method.

The white balance just considers the correction of the light in the scene. It is a good solution to over-land images with wrong light estimation, but in the underwater environments do not presents a sufficient performance on recovering original signal properties. The general contrast stretching method, CLAHE, presented as a good method to improving the visibility of a degraded underwater image since it is unlikely that it would add information to a scene. There is just a move into the histogram, creating a more robust method.

VI. Conclusions and Future Work

This paper presented a novel dataset of turbid underwater images where it is possible to access the reference image,*i.e.* the same underwater images with no degradation was acquired and put available. The proposed dataset created possibility for a novel evaluation on underwater restoration/enhancement algorithms. With this, we compared some of the most popular image restoration/enhancement methods on their capacity to approximate a turbid image with the clean image.

The evaluation shows that general and simple enhancing method such as CLAHE [**?**] can improve the image visibility as much as a specific restoration methods, having a more robust behaviour. With this evaluation we show that for restoration algorithms it is hard to estimate model parameters when bigger range of environments conditions is considered. The existent methods in the state-of-the-art just seems to not deal with different levels of degradation that a underwater image can have. Yet we showed that for recent works [18], based on joining different priors, there is a more robust parameter estimation for multiple turbidity conditions, culminating on better image restoration.

As a future work we believe that restoration methods should consider different turbidity conditions as a way to propose priors. For that, we think that learning approaches can be the most suited since is hard to design multiple priors by hand.

Finally, the dataset and methodology proposed in this work can be useful not only to evaluate image restoration methods, but also to test any vision algorithms that are sensitive to turbidity on underwater vision applications.

Acknowledgment

The authors would like to thank the Brazilian Petroleum Corporation - Petrobras, the Brazilian National Agency of Petroleum, Natural Gas and Biofuels (ANP), to the Funding Authority for Studies and Projects (FINEP) and to Ministry of Science and Technology (MCT) for their financial support through the Human Resources Program of ANP to the Petroleum and Gas Sector - PRH-ANP/MCT. This paper is also a contribution of the Brazilian National Institute of Science and Technology - INCT-Mar COI funded by CNPq Grant Number 610012/2011-8.

References

[1] V. Brandou, A.-G. Allais, M. Perrier, E. Malis, P. Rives, J. Sarrazin, and P.-M. Sarradin, "3d reconstruction of natural underwater scenes using the stereovision system iris," in *OCEANS 2007 - Europe*, June 2007, pp. 1–6.

[2] A. Shihavuddin, N. Gracias, R. Garcia, J. Escartin, and R. B. Pedersen, "Automated classification and thematic mapping of bacterial mats in the north sea," in *OCEANS-Bergen, 2013 MTS/IEEE*. IEEE, 2013, pp. 1–8.

[3] M. D. Stokes and G. B. Deane, "Automated processing of coral reef benthic images," *Limnol. Oceanogr.: Methods*, vol. 7, no. 157, pp. 157–168, 2009.

[4] J. Aulinas, Y. R. Petillot, X. Lladó, J. Salvi, and R. Garcia, "Vision-based underwater slam for the sparus auv."

[5] R. M. Eustice, O. Pizarro, and H. Singh, "Visually augmented navigation for autonomous underwater vehicles," *Oceanic Engineering, IEEE Journal of*, vol. 33, no. 2, pp. 103–122, 2008.

[6] T. H. P. B. C. Ancuti, C. O. Ancuti, "Enhancing underwater images and videos by fusion," *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, 2012.

[7] W. Lin and C.-C. J. Kuo, "Perceptual visual quality metrics: A survey," *Journal of Visual Communication and Image Representation*, vol. 22, no. 4, pp. 297–312, 2011.

[8] N. D. S. B. F. Codevilla, J. Gaya, "Achieving turbidity robustness on underwater images local feature detection," *BMVC 2015*, vol. 26, p. 157, 2015.

[9] R. Hummel, "Image enhancement by histogram transformation," *Computer graphics and image processing*, 1977.

[10] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," *Sixth International Conference on Computer Vision, 1998.*, 1998.

[11] J. Van De Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy," *Image Processing, IEEE Transactions on*, vol. 16, no. 9, pp. 2207–2214, 2007.

[12] Q. I. J. L. Bazeille, S. and J. P. Malkasse, "Automatic underwater image pre-processing," *Proceedings of the Caracterisation du Milieu Marin*, 2006.

[13] X. T. K. He, J. Sun, "Single image haze removal using dark channel prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011.

[14] N. Carlevaris Bianco, A. Mohan, and R. M. Eustice, "Initial results in underwater single image dehazing," in *OCEANS 2010*. IEEE, 2010, pp. 1–8.

[15] J. Y. Chiang and Y.-C. Chen, "Underwater image enhancement by wavelength compensation and dehazing," *Image Processing, IEEE Transactions on*, vol. 21, no. 4, pp. 1756–1769, 2012.

[16] A. Galdran, D. Pardo, A. Picón, and A. Alvarez-Gila, "Automatic red-channel underwater image restoration," *Journal of Visual Communication and Image Representation*, vol. 26, pp. 132–145, 2015.

[17] H. Lu, Y. Li, L. Zhang, and S. Serikawa, "Contrast enhancement for images in turbid water," *JOSA A*, vol. 32, no. 5, pp. 886–893, 2015.

[18] A. C. D. S. B. Joel D.O. Gaya, Felipe Codevilla, "General participative media single image restoration," 2016, arXiv:submit/1493653 [cs.CV] 6 Mar 2016.

[19] J. T. O. Kirk, *Light and Photosynthesis in Aquatic Ecosystems, 3rd Edition*. cambridge, 2011.

[20] Y. Schechner and N. Karpel, "Recovery of underwater visibility and structure by polarization analysis," *Oceanic Engineering, IEEE Journal of*, vol. 30, no. 3, pp. 570–587, July 2005.

[21] C. D. R. R. S. K. N. Srinivasa G. Narasimhan, Mohit Gupta and H. W. Jensen, "Acquiring scattering properties of participating media by dilution," *ACM Transactions on Graphics (TOG) - Proceedings of ACM SIGGRAPH 2006*, vol. 25, pp. 1003–1012, 2006.

[22] R. Garcia and N. Gracias, "Detection of interest points in turbid underwater images." *OCEANS 2011 IEEE Spain*, pp. 1–9, 2011.

[23] G. D. Finlayson and E. Trezzi, "Shades of gray and colour constancy," *IST/SID Twelfth Color Imaging Conference: Color Science,Systems and Applications, Society for Imaging Science and Technology*, 2004.

[24] H. R. S. Z. Wang, A. C. Bovik and E. P. Simoncelli, "Image quality assessment: From error measurement to structural similarity," *IEEE Transactions on Image Processing*, 2003.