Contributions to the Automatic Restoration of Images from Scenes in Participating Media

Paulo L. J. Drews-Jr, Erickson R. Nascimento and Mario F. M. Campos
Programa de Pós-Graduação em Ciência da Computação (PPGCC)
Universidade Federal de Minas Gerais (UFMG)

***Abstract*—This work deals with the problem of image restoration of monocular images acquired in participating media, *i.e.* media that interfere with light propagation. Specifically, the proposed work focus on the automatic restoration of images acquired in underwater and foggy/hazy scenes. The proposed restoration process requires at least a pair of images as input and produces images in which the medium effects are attenuated and the visibility improved. Differently from previous works, our method does not need additional equipment or information – only a calibrated camera. Our method adopts a model-based approach by estimating of the depth map and the attenuation coefficient. We performed experimental evaluation in real and simulated environments with significant improvement in the quality of the images.**

***Keywords*-Image Restoration, Computer Vision, Participating Media**

I. Introduction

The computer vision and image processing fields have witnessed remarkable advances on a wide range of problems.Regardless of the advances, almost all of the techniques assumes that scene and camera are immersed in a non-participating medium, *i.e.* they assumed that the light rays travel through the medium without any alteration. However, there are some media that change the intensity and the direction of the light rays, called participating medium. Among them, the most important and discussed here are the water and haze/fog.

A myriad of real world problems need to deal with images acquired in participating media, *e.g* surveillance, mapping, autonomous vehicles [1] to name a few. The effects of absorption and light scattering in participating medium decrease the overall contrast on images and causes color shifting, which reduce visibility on underwater scenes, for instance.

The main contribution is a new automatic method capable of restoring monocular sequences of images acquired in participating medium without any additional information. The method is based on temporal relation, three-dimensional structure and medium properties. The main steps of our approach are depicted in the outline in Fig. 1. The method is initialized by using a new transmission prior that provides an initial estimation of the scene depth which allow us to compute the optical flow. Structure from motion techniques based on a novel optical flow model provide an estimation of the depth map, which is used for compute the attenuation coefficient and subsequently to restore the image sequence. The obtained results shows significant improvements in several quantitative metrics.

This PhD thesis was concluded on February 5th, 2016.

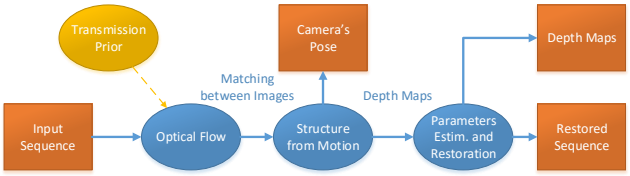


Figure 1. Outline of our image restoration methodology. We estimate a transmission map based on priors. This map allows us to compute the optical flow. Then, depth maps are predicted using structure from motion techniques. Finally, the medium parameters and the restored images are estimated. The orange boxes are the inputs/outputs data, the blue ellipses are the proposed steps, and the yellow ellipse is the prior that provides the optical flow initialization.

The rest of this work is organized as follows. After reviewing the state-of-the-art (Sec. II) and the light propagation model (Sec. III); Sec. IV presents the proposed methodology. Experimental results are presented in Sec. V. Finally, the main conclusions are drawn in Sec. VI.

II. Related Works

Several approaches have been proposed to tackle the problem of restoring images acquired in participating media, namely, specialized hardware [2], polarization filters [3] and stereo images [1]. Although the high quality of the reported results using specialized hardware, most of these methods are expensive and require complex setup. The use of polarizers is a cumbersome, even though images acquired with them present good results. The main drawback of this technique is the need to identify the maximum and minimum polarization states. In the case of stereo systems, the correspondence even harder an already difficult problem due to the effects imposed by the medium.

Several methods [4], [5], [6] based on single images have been proposed in the literature. While they show good results on foggy images, their performance degrades in underwater scenarios. The main issue of these methods is the estimation based on heuristics, which may hold true only for restricted conditions.

Very few studies have addressed the problem of image restoration of a sequence of image. In the work of [7], a method to estimate the medium transmission based on priors and optical flow is presented to enhance the visibility of hazy images. Despite the interesting results, it is based on

assumption of brightness constancy that does not hold true on common situations. The temporal coherence is take into account to estimate the transmission map in the work of [5], however this method also can also fail in typical situation.

Differently from the aforementioned works, our approach is based on temporal relation, three-dimensional structure and medium properties. These information are fused to provide a more robust image restoration.

III. Light Propagation Model of Participating Media

As stated, our methodology makes use of a light propagation model to explain the image formation process in a participating media. Images captured from scene immersed in participating media can be modeled as a complex interaction between the light, the medium and the scene structure. This interaction is modelled as the sum of three main components [8]: Direct illumination (I_d), forward-scattering and backscattering (I_{bs}). The direct component is the fraction of light that reaches the camera. Part of the light that radiates from object is lost due to scattering and absorption. These effects are modeled by the direct component that can be written as:

$$I_d = Je^{-\eta d} = Jt_r, \tag{1}$$

where J and d are the scene radiance and the depth, respectively; η is the attenuation coefficient and t_r is the medium transmission, considered as the exponential term. According to Schechner and Karpel [3], backscattering is the prime reason for image contrast degradation, and for this reason the forward scattering can be usually neglected in underwater images. The backscattering component, I_{bs} results from the interaction between the ambient illumination sources and particles dispersed in the medium. It can be defined by Eq. 2:

$$I_{bs} = A(1 - e^{-\eta d}) = A(1 - t_r), \tag{2}$$

where A is the global light on the scene. This term is a scalar that depends on the wavelength and can be estimated by finding the brightest pixel. The final model describing the image formation in participating medium can be stated as:

$$I(\boldsymbol{x}) = J(\boldsymbol{x})t_r(\boldsymbol{x}) + A(1 - t_r(\boldsymbol{x})), \tag{3}$$

where $\boldsymbol{x}$ are the pixel coordinates (x, y) and $I(\boldsymbol{x})$ is the image obtained in the participating media.

IV. Methodology

In this work we present a new automatic methodology to restore images acquired in participating media. Assuming the model previously described, the problem of image restoration may be reduced to the problem of estimating the medium parameters and the depth map.

Our methodology is composed of three main steps: the dense correspondence between images by estimating the optical flow, the 3D structure estimation, and parameter estimation and image restoration. Fig. 1 depicted the proposed methodology. We estimate a transmission map based on a new prior. This map allows us to compute the optical flow. Depth maps are predicted using structure from motion techniques. Finally, the medium parameters and the restored images are estimated. In addition to the restored images, our method also produces an estimation of the depth maps, the camera's pose and the attenuation coefficient of the medium, which can be used in specific applications such as 3D reconstruction or tracking.

A. Optical Flow in Participating Media

The problem of correspondence between images is a key aspect for estimating the depth maps of monocular sequences. The majority of optical flow methods assumes constancy in the brightness patterns in the image. However, this assumption does not hold true for participating media.

Virtually all proposed methodologies are not catered for participating media, since the assumption of brightness constancy is not valid in these environments. Therefore, we proposed the Generalized Optical Flow Model (GOFM). GOFM assumes that the brightness in the image is not constant because of the effects of the medium. Nevertheless, it assumes that the scene radiance is approximately constant. The final model of the GOFM is:

$$(I_x^c u + I_y^c v + I_t^c) + (I^c - A^c)(D_x^c u + D_y^c v + D_t^c) = 0, \tag{4}$$

where the image derivatives I_x, I_y, and I_t can be estimated as proposed by [9]. The function $D(x, y, t) = -\log t_r(x, y, t)$ represents the depth map up to scale defined by the attenuation coefficient, where the derivatives in relation to x, y and t are defined as D_x, D_y and D_t, respectively. The index c represents each channel in the RGB domain. One important aspect of the GOFM model is related to the function $D(x, y, t)$ that needs to be previously known. It can be estimated using priors that allows to obtain the transmission map $t_r(x, y, t)$ based on a single image, detailed in the next section.

The accuracy of optical flow estimation algorithms has been improving steadily as evidenced by results on optical flow benchmarks. The key aspects of the modern techniques depends on mainly four factors [9]: objective function, implementation details, parameter tuning and optimization techniques. We performed a robust implementation based on this guidelines.

The optical flow model cannot be solved pointwise, since the number of parameters to be estimated is larger than the number of linearly independent equations. This indeterminacy is called the aperture problem. It remains true for GOFM formulation. Color information may provide valuable information, mainly in underwater medium that presents different absorption rates for each wavelength. Since the contribution of the color information in hazy scenes is limited, other constraint needs to be applied. We adopted the global smoothness and an incremental multi-resolution technique [9].

We performed several experiments to evaluate the robustness of the GOFM using simulated and real image sequences using absolute endpoint error and the average angular error. The method is compared with the state-of-the-art Classic-NL and the modern implementation of Horn-Schunk method, both proposed by [9]. Results for both methods are obtained using degraded images, and the restored images using a single image method. Both error obtained using GOFM are significantly

smaller than the other methods. The results show GOFM is more robust to the increasing of the turbidity level, presenting an approximately linear behavior in endpoint error.

1) Initialization - Transmission Prior: The GOFM formulation assumes a previous knowledge about the function D. Recently, several priors for single images have been proposed. They enable us to estimate the medium transmission, and, consequently, the depth map up to scale. Among them, the most successful method is the Dark Channel Prior (DCP) [4]. It is adopted here for hazy/foggy images. DCP is a statistical prior based on the observation that local patch on clear day images contain some pixels whose intensity is very low in at least one color channel. These low intensity in the dark channel is mainly due to three factors, as described in [4]: a) shadows; b) colorful objects or surfaces where at least one color has low intensity and c) dark objects or surfaces.

The observation of a low Dark Channel in images acquired in non-participating medium is not easy to be tested underwater because of the difficulty to obtain real underwater scenes in out of water condition. Nevertheless, the assumptions made by [4] are still plausible.

Although the dark channel assumption sounds acceptable, the wavelength independence is clearly false in most of the cases. Therefore, we proposed a new prior called Underwater DCP (UDCP). The method only uses the green and blue channels due to the difficult to modeling the behavior of the red channel. This phenomenon is mainly related to the high absorption effect in the red channel [10] which imposes it to be near zero in many situation.

Similarly to DCP, we isolate the transmission in a local patch $\tilde{t_r}$ as:

$$\tilde{t_r}(\boldsymbol{x}) = 1 - \min_{\boldsymbol{y}\in\Omega(\boldsymbol{x})} (\min_{c\in G,B} \frac{I^c(\boldsymbol{y})}{A^c}), \tag{5}$$

where the global light A is estimated by finding the brightest pixel in the underwater dark channel, and the transmission is refined using the guided filter method [11]. This method also enables us to restore using a single image, and is adopted in the Sec. V.

We performed an experimental verification to evaluate the assumption of the UDCP based on two statements: a) the main assumption of the DCP for outdoor scenes remains valid if only applied to green and blue channels, and b) the behavior of the UDCP histogram in underwater scenes is plausible. We create a dataset composed of 1,022 outdoor landscape images. The results show the statistics for the UDCP assumption is a more general supposition than the DCP assumption. Another important characteristic concerns the blue channel, which in natural scenes tends to be darker than the other channels. The underwater medium is typically blue, thus increasing the intensities of this color channel. This fact corroborates the underwater dark channel assumption.

B. Structure from Motion

We estimated depth maps based on the correspondence between pixels provided by the optical flow. We assumed that the camera calibration is previously known. The adopted methodology is based on the continuous essential matrix, triangulation and bundle adjustment.

One key aspect of the structure from motion is the camera calibration. For haze/fog scenes, the medium does not significantly change the parameters of the camera. However, cameras immersed in water are usually confined in underwater housing filled with air, viewing the scene through a piece of glass. Thus, the perspective model might fail because the rays do not intersecting in one common center of projection. Nonetheless, the perspective model is the most adopted in the literature for underwater images because of its simplicity and robustness, approximating the refractive effect by calibrated parameters [12].

The epipolar constraint is a well studied tool to obtain the camera's pose by means of the discrete essential matrix. However, the continuous epipolar constraint is a more adequate tool for the case of correspondences obtained by optical flow [13]. We performed an experimental evaluation to guarantee the robustness of the constraint. We proposed a MSAC (M-estimator SAmple and Consensus) approach [14] to recover the continuous essential matrix (CEM). MSAC is an improvement in the classical RANSAC method that change the cost function to be minimized. This new function improves the accuracy with no additional computational burden [14]. The proposed method uses MSAC and the eight-point algorithm [13] for estimating camera's pose.

The estimation of the 3D points is still sensible to noise and misestimation of the optical flow. Thus, we performed a triangulation and a bundle adjustment optimization [15] to compute a reliable set of 3D points. Although the correspondences between points provided by the optical flow are dense, they are prone to outliers and missing values. Therefore, we applied two algorithms to surpass these problems. Firstly, the inpainting approach [16] is used to fill these gaps. After that, we applied the guided filter method [11] to smooth the depth maps and to improve the edges discontinuities.

C. Parameter Estimation and Restoration

This step presents two main objectives: estimating the attenuation coefficient and restoring the image. The restoration follows the model (Eq. 3). The direct component $I_d = J(\boldsymbol{x})t_r(\boldsymbol{x})$ is close to zero when $t_r(\boldsymbol{x})$ is small. Thus, the recovered scene radiance $J(\boldsymbol{x})$ is prone to noise. Thus, we restrict the transmission $t_r(\boldsymbol{x})$ by a lower bound t_{r0}. The final restored image $J^c(\boldsymbol{x})$ for each color channel c is given by:

$$J^c(\boldsymbol{x}) = \frac{I^c(\boldsymbol{x}) - A^c}{\max(t_{r0}, t_r(\boldsymbol{x}))} + A^c. \tag{6}$$

The estimation of the transmission is based on the knowledge about the depth map and the attenuation coefficient. An interpretation from the optical model allows us to estimate the attenuation coefficient, η, based on the depth maps in two consecutive frames. Assuming the same 3D point in the scene, the model is given by:

$$\eta = -\frac{1}{\Delta d} \ln \frac{I^{t+1} - A}{I^t - A}, \tag{7}$$

where I^t and I^{t+1} are the image acquired in a participating medium in the time t and $t+1$, respectively. d is the depth in time t, and Δd is the depth variation. The estimation of η can be reduced to line fitting because the attenuation coefficient is assumed as constant for the entire image.

This approach is highly dependent on the depth maps and, mainly, the optical flow. Thus, we imposed a new constraint inspired in the work of [1] to improve the robustness of the MSAC-based approach. We assume that the lowest 1% intensity pixels in each color channel for all depths are originated from black objects. Indeed, the object does not need to be black but dark in a specific color channel. Although this assumption is related with the dark channel assumption, it is slightly different because it is assumed as valid for a small portion of the pixels instead of all patches on image. The new error function to be minimized for each color channel using the proposed constraints is defined as:

$$E_{MSAC} = \sum \omega_1 (I^{t+1} - A(1 - e^{-\eta(d+\Delta d)})) + \omega_2 (I^t - A(1 - e^{-\eta d})) + \omega_3 \left(\frac{I^{t+1} - A}{I^t - A} - e^{-\eta(\Delta d)} \right), \quad (8)$$

where the weights ω_1, ω_2 and ω_3 define the importance of each term in the error function, and they are obtained using experimental estimation. It is worth to note that this function is valid only for black points, *i.e.* $J \approx 0$. This function is non-linear and complex to be optimized in terms of η. Furthermore, the number of outliers could be large due to the difficult to obtain depth and correspondence between pixels in participating medium. Thus, we also adopted an optimization based on MSAC approach [14].

V. Experimental Results

We obtained results using simulated and real datasets, all of them with qualitative and quantitative evaluation. We compared our approach with the single image methods DCP/UDCP and two enhancement techniques: histogram equalization and contrast-limited adaptive histogram equalization (CLAHE) [17].

Quantitative results are obtained using the metric proposed by [18]. They define three different indexes: e, $\overline{r}$ and s. The value of e evaluates the ability of a method to restore edges, which were not visible in the degraded image, but are visible in the restored image. The value of $\overline{r}$ measures the quality of contrast restoration. Finally, the value of s is obtained from the proportion of the number of pixels which are saturated after applying the restoration method but were not before. These three indexes allow us to estimate an empirical restoration score $\tau = e + \overline{r} + (1 - s)$ [18], where larger values mean better restoration.

We also perform quantitative results by matching SIFT [19] descriptors. It allow us to evaluate the ability of the descriptor to identify and match features from a raw and restored image. For each pair of images, we show the number of keypoints detected in both images and the number of correct matches.

For the simulated images whose ground truth is known, we adopted two state-of-art Image Quality Assessment (IQA) measures. IQA metrics compare two images using values in the interval $[0, 1]$, where one means the best quality while zero means the worst quality. The first metric, called FSIM [20], is based on the fact that the human visual system responds to an image mainly according to its low-level features, specifically the phase congruency and the gradient magnitude. Secondly, [20] also proposed an extension called FSIMc that uses the chrominance information in the YIQ color space.

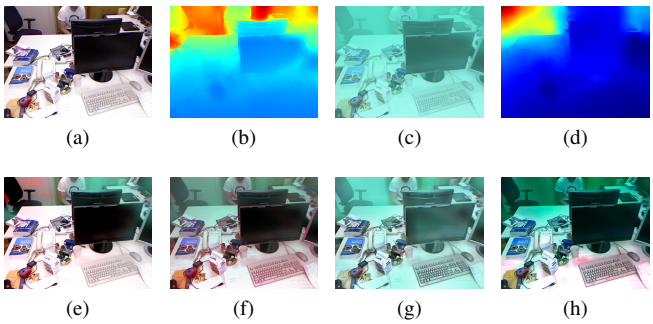


Figure 2. Simulation of underwater effects using images from RGB-D SLAM dataset [21]: (a) sample image from the dataset, (b) ground truth depth maps, (c) simulated underwater effects with strong chlorophyll concentration of $C = 2.0mg\,m^{-3}$ [10] at $5m$ of water depth, (d) estimated depth maps, and (e) restoration result obtained by **our method**. Finally, results obtained by histogram equalization techniques (f), CLAHE (g) and UDCP (h).

A. Simulated Results

One of the challenges with validating our methodology is the difficulty of acquiring images of participating medium with reliable ground truth. To tackle with this problem, we chose to validate our method by using RGB-D sequences [21], which includes the sequence of images and their respective depth maps. We performed artificial degradation of the images by simulating the effects of an underwater camera at a depth of $5m$. We simulated the attenuation coefficient η and the global illumination A as proposed in [10]. We assumed a strong turbidity with the medium being contaminated with a strong chlorophyll concentration $C = 2.0mg\,m^{-3}$ [10]. Since the depth maps provided by the dataset are not perfect, we also performed inpainting [16] and guided filtering [11].

A sample frame and the colorized ground truth depth maps are shown in figs. 2a and 2b. The result after simulation of the medium effects is shown in Fig. 2c. Notice that the space variant effect generated by the simulation, where the chair in the farthest region presents larger decreasing in visibility than the keyboard. Estimated depth map using our approach is shown in Fig. 2d which is prone to error, mainly in the top right area due to the movement of the camera. Our method were repeated one thousand times due to their stochastic nature and the difficult to adjust parameters that are typically unknown. The results presented a small standard deviation ($< 1\%$), and we are able to recover the coefficient with just a reduced mean squared error equal to 0.0109.

Qualitative results for the restoration of the simulated images are presented in figs. 2e-2h. We compared our approach with the single image method UDCP, histogram equalization

Table I
COMPARATIVE STUDY USING THE AVERAGE OF QUANTITATIVE METRICS IN SIMULATED IMAGES THAT INCLUDE IQA METRICS [20] AND SIFT MATCHING [19]. THE BEST AND SECOND BEST RESULTS ARE HIGHLIGHTED IN BLUE AND LIGHT BLUE LETTERS, RESPECTIVELY.

	FSIM	FSIMc	Correct Matches	Keypoints
UDCP	0.9217	0.85495	**49**	**2091**
Hist. Eq.	0.90545	**0.88405**	40	1844
CLAHE	**0.9354**	4.6273	40	**2067**
Our Method	**0.9796**	**0.96595**	**59**	1844

Table II
COMPARATIVE STUDY USING THE AVERAGE OF THE RESTORATION SCORE τ [18] AND SIFT MATCHING [19] FOR THE SEQUENCES IN FIG. 3. THE BEST AND SECOND BEST RESULTS ARE HIGHLIGHTED IN BLUE AND LIGHT BLUE LETTERS, RESPECTIVELY.

	Underwater Sequence					**Foggy Sequence**				
	e	$\overline{r}$	τ	Match.	Pts.	e	$\overline{r}$	τ	Match.	Pts.
UDCP/DCP	1.7089	1.2591	3.9678	1	18	1.4185	0.7877	3.2063	10	30
Hist. Eq.	**2.5430**	2.4544	**5.9885**	10	197	**1.5540**	**1.8267**	**4.3667**	**70**	**286**
CLAHE	1.8821	**2.8494**	5.7315	**11**	**391**	0.4996	1.5081	3.0077	65	265
Our Method	**2.7057**	**2.6493**	**6.3488**	**33**	**375**	**1.5804**	**1.9689**	**4.5430**	**78**	**360**

and CLAHE [17]. A sample restored image by our method is presented in Fig. 2e. The restoration is not perfect but we obtain a significant improvement, *e.g.* the chair in the top left. Histogram equalization (Fig. 2f) and CLAHE (Fig. 2g) produce limited restoration in terms of visibility and color fidelity. Results obtained by UDCP (Fig. 2h) show improvement in terms of visibility, but the method distorts the colors, *e.g.* white desk with some red halos.

Table I shows quantitative results using FSIM metrics [20] and SIFT matching [19]. One can readily see that our methodology produces the largest values in the IQA metrics. Results for matching show that our approach provides the largest number of correct matches, while the number of detected keypoints for UDCP and CLAHE approaches are the largest. The increasing in the contrast and, mainly, in the noise produce these results.

B. Real Results

We show real results in two scenarios. Firstly, we captured underwater images in the Brazil's Southeast Coast with depth ranging from $12m$ to $20m$. Furthermore, we captured a sequence from a residential area in a foggy day.

For the underwater sequence, the estimated attenuation coefficient is $\eta = [0.0335, 0.0331, 0.0289]$ for each RGB channel, respectively. The blue channel typically has a smaller attenuation value while the red channel has a larger value as shown in the results. It is worth noting that this coefficient is obtained up to a scale factor due to the depth map estimation, and they are similar to each other due to the water characteristics and depth of the image acquisition.

Fig. 3a is a sample underwater image with limited visibility and significant color distortion. Fig. 3b shows restored image by our method, where the quality is improved. Results obtained using CLAHE are shown in Fig. 3d, where the contrast and the noise are increased, and the colors are distorted. Restored images using histogram equalization and UDCP are shown in fig. 3c and 3e, respectively.

Table II shows qualitative evaluation for the underwater sequence. Our methods outperforms the others in term of the τ metric. Our method obtains improvement in terms of contrast and slightly smaller values in terms of new edges. CLAHE obtains the largest values of $\overline{r}$ because of the increase in the overall contrast, however with some color distortion and increasing the noise. Our approach also provides the largest number of correct matches using SIFT [19]. However, the number of detected keypoints for the CLAHE method is the largest. This results is expected since the restoration obtained by CLAHE presented a large $\overline{r}$. However, this restoration is not stable, thus it does not increase the number of correct matches. Our method is able to significantly improve the number of matches, as well as the number of detected keypoints. It is also corroborated by the improvement in terms of contrast, $\overline{r}$.

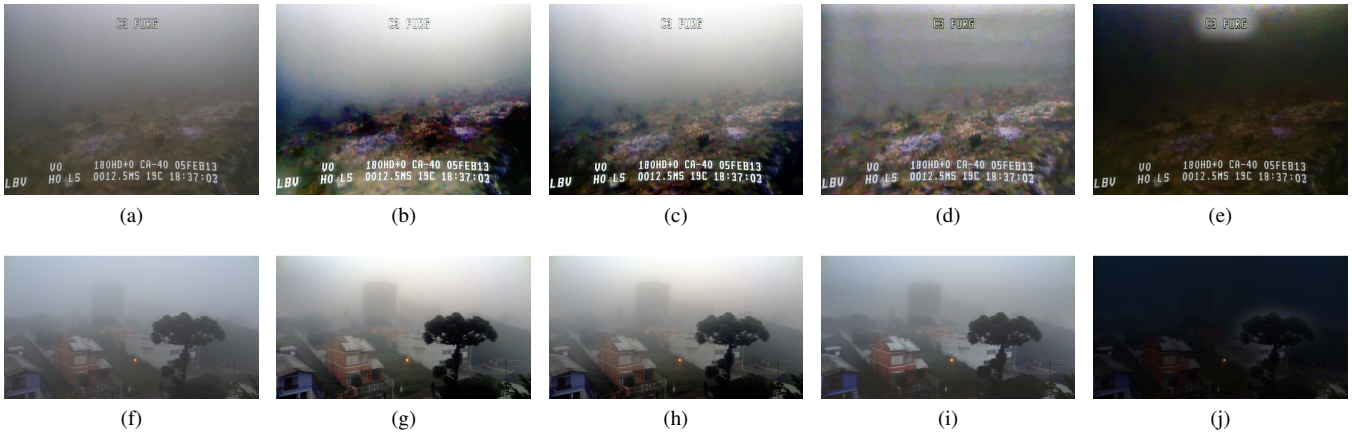
A foggy sequence is also shown in Fig. 3. Fig.3f shows a sample image that presents limited visibility and color distortion. Fig. 3g shows our restoration. The visibility and the color are improved, mainly in the houses (bottom left). The buildings in the center of image are under a strong "haze" layer, thus limiting the capability of restoration due to the loss of information. However, the contours of the buildings are recovered by our approach. The result for histogram equalization (Fig. 3h) is similar to our result. One important difference can be noted in the largest tree that the our method is able to improve. The result produced by CLAHE (Fig. 3i) is imperceptible. DCP fails to estimate the global light (Fig. 3j.), thus the image becomes darker with limited restoration.

The estimated attenuation coefficient for the foggy sequence is $\eta = [0.1752;$
$0.2026; 0.1882]$. Differently to the underwater sequence, the red channel is the smallest attenuation coefficient while the green channel presents the largest value. These three coefficients are relatively similar ($\approx 15\%$) as expected.

Table II also shows qualitative evaluation for the foggy sequence. Our method outperforms the others in term of the τ metric. Our method obtains similar results to histogram equalization method. However, our method presents a small advantage in all metrics. The results of CLAHE technique is limited, presenting a small improvement. Therefore, CLAHE obtains the smallest values in the number of edges and the tau metric. DCP distorts the colors, but this fact is not take into account by this metric. Our method obtains the largest number of correct matches using SIFT [19], as well as detected keypoints. The histogram equalization also presents good results.

VI. CONCLUSIONS

This work proposed a new methodology to restore sequence of images acquired in participating media. We explore the temporal relation between the images that allow us to estimate scene structure, camera's pose and depth maps. The relation is obtained by a new formulation of optical flow adapted for participating media that depends on the knowledge of the medium transmission, which is computed by UDCP/DCP. Finally, we developed a new robust methodology to estimate the most critical parameter of the medium: the attenuation



(a) (b) (c) (d) (e)

(f) (g) (h) (i) (j)

Figure 3. Qualitative comparison using a sample image acquired in naturally lit shallow oceanic water and in a foggy day: original image (a,f), restored using **our method** (b,g), histogram equalization (c,h) , CLAHE [17] (d,i), UDCP (e), and DCP [4] (j).

coefficient. Qualitative and quantitative results in simulated and real images show the quality of the restoration obtained by our approach. The estimated depth maps is still limited, but enough to the restoration task. The proposed method to estimate the attenuation coefficient is robust even in the presence of outliers.

Future work will focus on investigating the structure from motion method to improve the depth map estimation and the inclusion of artificial illumination in the scene. Furthermore, a new method to quantitatively evaluate the image restoration methods will be investigated.

The results of this work were partly published in international conferences and an international journal[1]. Furthermore, some papers are under submission/review[2]. This work is awarded with a sandwich doctorate scholarship from PDSE-CAPES in CSIRO-Australia.

REFERENCES

[1] M. Roser, M. Dunbabin, and A. Geiger, "Simultaneous underwater visibility assessment, enhancement and improved stereo," in *IEEE ICRA*, 2014, pp. 3840–3847.

[1] • Drews-Jr, P. L. J. ; Nascimento, E. R. ; Botelho, S. S. C.; Campos, M. F. M. . Underwater Depth Estimation and Image Restoration Based on Single Images. IEEE Computer Graphics and Applications, vol. 36(2), pp. 50-61, 2016 – **A2** in the Qualis-CC.

•Drews-Jr, P. L. J. ; Nascimento, E. R. ; Campos, M. F. M. ; Elfes, A. . Automatic Restoration of Underwater Monocular Sequences of Images. In: IEEE/RSJ IROS, 2015 – **A1** in the Qualis-CC.

•Concha, A. ; Drews-Jr, P. L. J. ; Campos, M. F. M. ; Civera, J. . Real-time localization and dense mapping in underwater environments from a monocular sequence. In: IEEE OCEANS, 2015 – Most Important Conf. in the Oceanic Engineering field.

•Drews-Jr, P. L. J. ; Nascimento, E. R. ; Xavier, A. ; Campos, M. F. M. . Generalized Optical Flow Model for Scattering Media. In: IEEE/IAPR ICPR, 2014 – **A1** in the Qualis-CC.

•Drews-Jr, P. L. J. ; Nascimento, E. R. ; Moraes, F. C. ; Botelho, S. S. C.; Campos, M. F. M. . Transmission Estimation in Underwater Single Images. In: IEEE ICCV - Workshop on Underwater Vision, 2013 – Workshop of **A1** in the Qualis-CC

[2] •Drews-Jr, P. L. J. ; Hernàndez, E. ; Elfes, A. ; Nascimento, E. R. ; Campos, M. F. M.. Real-Time Monocular Underwater Obstacle Avoidance. In: IEEE/RSJ IROS, 2016, 6 pages – **A1** in the Qualis-CC (Under Review).

•Drews-Jr, P. L. J. ; Nascimento, E. R. ; Campos, M. F. M. Restoration of Sequences of Monocular Images Acquired in Participating Media. Computer Vision and Image Understanding, 22 pages – **A1** in the Qualis-CC (Under Submission).

•Drews-Jr, P. L. J. ; Nascimento, E. R. ; Campos, M. F. M. A Survey of Visual Image Restoration in Scattering Media. ACM Computing Surveys, 38 pages – **A1** in the Qualis-CC (Under Submission).

[2] Z. Murez, T. Treibitz, R. Ramamoorthi, and D. Kriegman, "Photometric stereo in a scattering medium," in *IEEE ICCV*, 2015, pp. 3415–3423.

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

[4] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior." in *IEEE CVPR*, 2009, pp. 1956–1963.

[5] J. Kim, W. Jang, Y. Park, D. Lee, J. Sim, and C. Kim, "Temporally coherent real-time video dehazing," in *IEEE ICIP*, 2012, pp. 969–972.

[6] D. Berman, T. Treibitz, and S. Avidan, "Non-local image dehazing," in *IEEE CVPR*, 2016, pp. 1–9.

[7] J. Zhang, L. Li, Y. Zhang, G. Yang, X. Cao, and J. Sun, "Video dehazing with spatial and temporal coherence," *The Visual Computer*, vol. 27, no. 6-8, pp. 749–757, 2011.

[8] J. Jaffe, "Computer modeling and the design of optimal underwater imaging systems," *IEEE JOE*, vol. 15, no. 2, pp. 101–111, 1990.

[9] D. Sun, S. Roth, and M. J. Black, "A quantitative analysis of current practices in optical flow estimation and the principles behind them," *IJCV*, vol. 106, no. 2, pp. 115–137, 2014.

[10] C. Mobley, *Light and Water: Radiative Transfer in Natural Waters*. Acad. Press, 1994.

[11] K. He, J. Sun, and X. Tang, "Guided image filtering," in *ECCV*, 2010, pp. 1–14.

[12] A. Sedlazeck and R. Koch, "Perspective and non-perspective camera models in underwater imaging - overview and error analysis," in *Outdoor and Large-Scale Real-World Scene Analysis*, ser. LNCS. Spring-Verlag, 2012, vol. 7474, pp. 212–242.

[13] Y. Ma, S. Soatto, J. Kosecka, and S. S. Sastry, *An Invitation to 3-D Vision: From Images to Geometric Models*. Spring-Verlag, 2006, vol. 26.

[14] P. Torr and A. Zisserman, "MLESAC: A new robust estimator with application to estimating image geometry," *CVIU*, vol. 78, no. 1, pp. 138–156, 2000.

[15] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, 2003.

[16] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *ACM SIGGRAPH*, 2000, pp. 417–424.

[17] K. Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics Gems IV*. Acad. Press, 1994, pp. 474–485.

[18] N. Hautière, J.-P. Tarel, D. Aubert, and E. Dumont, "Blind contrast enhancement assessment by gradient ratioing at visible edges," *ISS IAS*, vol. 27, no. 2, pp. 87–95, 2008.

[19] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *IJCV*, vol. 60, no. 2, pp. 91–110, 2004.

[20] L. Zhang, D. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE TIP*, vol. 20, no. 8, pp. 2378–2386, 2011.

[21] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, "A benchmark for the evaluation of RGB-D SLAM systems," in *IEEE IROS*, 2012, pp. 573–580.