

# Color Balance and Fusion for Underwater Image Enhancement

Codruta O. Ancuti<sup>1</sup>, Cosmin Ancuti, Christophe De Vleeschouwer<sup>2</sup>, and Philippe Bekaert

**Abstract**—We introduce an effective technique to enhance the images captured underwater and degraded due to the medium scattering and absorption. Our method is a single image approach that does not require specialized hardware or knowledge about the underwater conditions or scene structure. It builds on the blending of two images that are directly derived from a color-compensated and white-balanced version of the original degraded image. The two images to fusion, as well as their associated weight maps, are defined to promote the transfer of edges and color contrast to the output image. To avoid that the sharp weight map transitions create artifacts in the low frequency components of the reconstructed image, we also adapt a multiscale fusion strategy. Our extensive qualitative and quantitative evaluation reveals that our enhanced images and videos are characterized by better exposedness of the dark regions, improved global contrast, and edges sharpness. Our validation also proves that our algorithm is reasonably independent of the camera settings, and improves the accuracy of several image processing applications, such as image segmentation and keypoint matching.

**Index Terms**—Underwater, image fusion, white-balancing.

## I. INTRODUCTION AND OVERVIEW

UNDERWATER environment offers many rare attractions such as marine animals and fishes, amazing landscape, and mysterious shipwrecks. Besides underwater photography, underwater imaging has also been an important source of interest in different branches of technology and scientific research [1], such as inspection of underwater infrastructures [2] and cables [3], detection of man made objects [4], control of underwater vehicles [5], marine biology research [6], and archeology [7].

Different from common images, underwater images suffer from poor visibility resulting from the attenuation of the propagated light, mainly due to absorption and scattering effects. The absorption substantially reduces the light energy, while the scattering causes changes in the light propagation direction. They result in foggy appearance and contrast degradation,

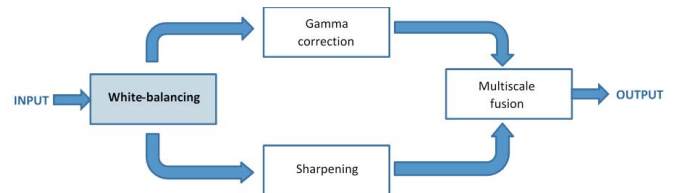


Fig. 1. Method overview: two images are derived from a white-balanced version of the single input, and are merged based on a (standard) multiscale fusion algorithm. The novelty of our approach lies in the proposed pipeline, but also in the definition of a white-balancing algorithm that is suited to our underwater enhancement problem.

making distant objects misty. Practically, in common sea water images, the objects at a distance of more than 10 meters are almost unperceivable, and the colors are faded because their composing wavelengths are cut according to the water depth.

There have been several attempts to restore and enhance the visibility of such degraded images. Since the deterioration of underwater scenes results from the combination of multiplicative and additive processes [8] traditional enhancing techniques such as gamma correction, histogram equalization appear to be strongly limited for such a task. In the previous works that are surveyed in Section II.B, the problem has been tackled by tailored acquisition strategies using multiple images [9], specialized hardware [10] or polarization filters [11]. Despite of their valuable achievements, these strategies suffer from a number of issues that reduce their practical applicability.

In contrast, this paper introduces a novel approach to remove the haze in underwater images based on a single image captured with a conventional camera. As illustrated in Fig. 1, our approach builds on the fusion of multiple inputs, but derives the two inputs to combine by correcting the contrast and by sharpening a white-balanced version of a single native input image. The white balancing stage aims at removing the color cast induced by underwater light scattering, so as to produce a natural appearance of the sub-sea images. The multi-scale implementation of the fusion process results in an artifact-free blending.

The rest of the paper is structured as follows. The next section briefly surveys the optical specificities of the underwater environment, before summarizing the work related to underwater dehazing. In Section III, we present our novel white-balancing approach, especially designed for underwater images. Section IV then describes the main components of our fusion-based enhancing technique, including inputs and associated weight maps definition. Before concluding, we present

Manuscript received August 16, 2016; revised December 20, 2016, March 13, 2017, May 9, 2017, August 1, 2017, and September 22, 2017; accepted September 25, 2017. Date of publication October 5, 2017; date of current version November 3, 2017. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. David Clausi. (Corresponding author: Cosmin Ancuti.)

C. O. Ancuti and C. Ancuti are with MEO, University Politehnica Timișoara, 300023 Timișoara, Romania (e-mail: cosmin.ancuti@upt.ro).

C. De Vleeschouwer is with ICTEAM, Université Catholique de Louvain, 1348 Louvain-la-Neuve, Belgium.

P. Bekaert is with the Expertise Centre for Digital Media, Hasselt University, 3590 Hasselt, Belgium.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2017.2759252

comparative qualitative and quantitative assessments of our white-balancing and fusion-based underwater dehazing techniques, as well as some results about their relevance to address common computer vision problems, namely image segmentation and feature matching.

## II. BACKGROUND KNOWLEDGE AND PREVIOUS ART

This section surveys the basic principles underlying light propagation in water, and reviews the main approaches that have been considered to restore or enhance the images captured under water.

### A. Light Propagation in Underwater

For an ideal transmission medium the received light is influenced mainly by the properties of the target objects and the camera lens characteristics. This is not the case underwater. First, the amount of light available under water, depends on several factors. The interaction between the sun light and the sea surface is affected by the time of the day (which influences the light incidence angle), and by the shape of the interface between air and water (rough vs. calm sea). The diving location also directly impacts the available light, due to a location-specific color cast: deeper seas and oceans induce green and blue casts, tropical waters appear cyan, while protected reefs are characterized by high visibility. In addition to the variable amount of light available under water, the density of particles that the light has to go through is several hundreds of times denser in seawater than in normal atmosphere. As a consequence, sub-sea water absorbs gradually different wavelengths of light. Red, which corresponds to the longest wavelength, is the first to be absorbed (10-15 ft), followed by orange (20-25 ft), and yellow (35-45 ft). Pictures taken at 5 ft depth will have a noticeable loss of red. Furthermore, the refractive index of water makes judging distances difficult. As a result, underwater objects can appear 25% larger than they really are.

The comprehensive studies of McGlamery [12] and Jaffe [13] have shown that the total irradiance incident on a generic point of the image plane has three main components in underwater mediums: direct component, forward scattering and back scattering. The direct component is the component of light reflected directly by the target object onto the image plane. At each image coordinate  $x$  the direct component is expressed as:

$$E_D(x) = J(x)e^{-\eta d(x)} = J(x)t(x) \quad (1)$$

where  $J(x)$  is the radiance of the object,  $d(x)$  is the distance between the observer and the object, and  $\eta$  is the attenuation coefficient. The exponential term  $e^{-\eta d(x)}$  is also known as the transmission  $t(x)$  through the underwater medium.

Besides the absorption, the floating particles existing in the underwater mediums also cause the deviation (scattering) of the incident rays of light. Forward-scattering results from a random deviation of a light ray on its way to the camera lens. It has been determined experimentally that its impact can be approximated by the convolution between the direct attenuated component, with a point spread function that

depends on the distance between the image plane and the object. Back-scattering is due to the artificial light (e.g. flash) that hits the water particles, and is reflected back to the camera. Back-scattering acts like a glaring veil superimposed on the object. The influence of this component may be reduced significantly by simply changing the position and angle of the artificial light source so that most of the reflected particle light do not reach the camera. However, in many practical cases, back-scattering remains the principal source of contrast loss and color shifting in underwater images. Mathematically, it is often expressed as:

$$E_{BS}(x) = B_\infty(x)(1 - e^{-\eta d(x)}) \quad (2)$$

where  $B_\infty(x)$  is a color vector known as the *back-scattered light*.

Ignoring the forward scattering component, the simplified underwater optical model thus becomes:

$$\mathcal{I}(x) = J(x)e^{-\eta d(x)} + B_\infty(x)(1 - e^{-\eta d(x)}) \quad (3)$$

This simplified underwater camera model (3) has a similar form than the model of Koschmieder [14], used to characterize the propagation of light in the atmosphere. It however does not reflect the fact that the attenuation coefficient strongly depends on the light wavelength, and thus the color, in underwater environments. Therefore, as discussed in the next section and illustrated in Fig. 11, a straightforward extension of outdoor dehazing approaches performs poorly at great depth, in presence of non-uniform artificial illumination and selective absorption of colors. This is also why our approach does not resort to an explicit inversion of the light propagation model.

### B. Related Work

The existing underwater dehazing techniques can be grouped in several classes. An important class corresponds to the methods using specialized hardware [1], [10], [15]. For instance, the divergent-beam underwater Lidar imaging (UWLI) system [10] uses an optical/laser-sensing technique to capture turbid underwater images. Unfortunately, these complex acquisition systems are very expensive, and power consuming.

A second class consists in polarization-based methods. These approaches use several images of the same scene captured with different degrees of polarization, as obtained by rotating a polarizing filter fixed to the camera. For instance, Schechner and Averbuch [11] exploit the polarization associated to back-scattered light to estimate the transmission map. While being effective in recovering distant regions, the polarization techniques are not applicable to video acquisition, and are therefore of limited help when dealing with dynamic scenes.

A third class of approaches employs multiple images [9], [16] or a rough approximation of the scene model [17]. Narasimhan and Nayar [9] exploited changes in intensities of scene points under different weather conditions in order to detect depth discontinuities in the scene. *Deep Photo* system [17] is able to restore images by employing the existing georeferenced digital terrain and

urban 3D models. Since this additional information (images and depth approximation) is generally not available, these methods are impractical for common users.

A fourth class of methods exploits the similarities between light propagation in fog and under water. Recently, several single image dehazing techniques have been introduced to restore images of outdoor foggy scenes [18]–[23]. These dehazing techniques reconstruct the intrinsic brightness of objects by inverting the Koschmieder’s visibility model [14]. Despite this model was initially formulated under strict assumptions (e.g. homogeneous atmosphere illumination, unique extinction coefficient whatever the light wavelength, and space-uniform scattering process) [24], several works have relaxed those strict constraints, and have shown that it can be used in heterogeneous lightning conditions and with heterogeneous extinction coefficient as long as the model parameters are estimated locally [20], [21], [25]. However, the underwater imaging is even more challenging, due to the fact that the extinction resulting from scattering depends on the light wavelength, i.e. on the color component.

Recently, several algorithms that specifically restore underwater images based on Dark Channel Prior (DCP) [19], [26] have been introduced. The DCP has initially been proposed for outdoor scenes dehazing. It assumes that the radiance of an object in a natural scene is small in at least one of the color component, and consequently defines regions of small transmission as the ones with large minimal value of colors. In the underwater context, the approach of Chiang and Chen [27] segments the foreground and the background regions based on DCP, and uses this information to remove the haze and color variations based on color compensation. Drews, Jr., *et al.* [28] also build on DCP, and assume that the predominant source of visual information under the water lies in the blue and green color channels. Their Underwater Dark Channel Prior (UDCP) has been shown to estimate better the transmission of underwater scenes than the conventional DCP. Galdran *et al.* [29] observe that, under water, the red component reciprocal increases with the distance to the camera, and introduce the Red Channel prior to recover colors associated with short wavelengths in underwater. Emberton *et al.* [30] designed a hierarchical rank based method, using a set of features to find those image regions that are the most haze-opaque, thereby refining the back-scattered light estimation, which in turns improves the light transmission model inversion. Lu *et al.* [31]. employ color lines, as in [32], to estimate the ambient light, and implement a variant of the DCP to estimate the transmission. As additional worthwhile contributions, bilateral filter is considered to remove highlighted regions before ambient light estimation, and another locally adaptive filter is considered to refine the transmission. Very recently, [31] has been extended to increase the resolution of its descattered and color-corrected output. This extension is presented in [33] and builds on super-resolution from transformed self-exemplars [34] to derive two high-resolution (HR) images, respectively from the output derived in [31] and from a denoised/descattered version of this output. A fusion-based strategy is then considered to blend those two intermediate HR images. This fusion aims at preserving the edges and detailed structures of the noisy

HR image, while taking advantage of the reduced noise and scatter in the second HR image. It however does not impact the rendering of color obtained with [31]. In contrast, our approach fundamentally aims at improving the colors (white-balancing component in Fig. 1), and uses the fusion to reinforce the edges (sharpening block in Fig. 1) and the color contrast (Gamma correction in Fig. 1). Actually, our solution provides an alternative to [31], while the HR fusion introduced in [33] should be considered as an optional complement to our work, to be applied to the output of our method when high-resolution is desired. An initial version of our fusion-based approach had been presented in our IEEE CVPR conference paper [35]. Compared to this preliminary work, this journal paper proposes a novel white balancing strategy that is shown to outperform our initial solution in presence of severe light attenuation, while supporting accurate transmission estimation in various acquisition settings. Our journal paper also revises the practical implementation of the fusion approach by proposing an alternative and simplified definition of the inputs and associated weight maps. The revised solution appears to significantly improve [35] in case of severe degradation of the underwater images (see the comprehensive qualitative and quantitative comparison provided in Section V.B). Moreover, additional experimental results reveal that our proposed dehazing strategy also improves the accuracy of conventional segmentation and point-matching algorithms (Section V.C), making our dehazing relevant for automatic underwater image processing systems.

To conclude this survey, it is worth mentioning that a class of specialized underwater image enhancing techniques have been introduced [36]–[38], based on the extension of the traditional enhancing techniques that are found in commercial tools such as color correction, histogram equalization/stretching, and linear mapping. Within this class, Chambah *et al.* [39] designed an unsupervised color correction strategy, Arnold-Bos *et al.* [37] developed a framework to deal with specific underwater noise, while the technique of Petit *et al.* [38] restores the image contrast by inverting a light attenuation model after applying a color space contraction. These approaches however appear to be effective only for relatively well illuminated scenes, and generally introduce strong halos and color distortions in presence of relatively poor lightning conditions.

### III. UNDERWATER WHITE BALANCE

As depicted in Fig. 1, our image enhancement approach adopts a two step strategy, combining white balancing and image fusion, to improve underwater images without resorting to the explicit inversion of the optical model. In our approach, white balancing aims at compensating for the color cast caused by the selective absorption of colors with depth, while image fusion is considered to enhance the edges and details of the scene, to mitigate the loss of contrast resulting from back-scattering. The fusion step is detailed in Section IV. We now focus on the white-balancing stage.

White-balancing aims at improving the image aspect, primarily by removing the undesired color castings due to various illumination or medium attenuation properties. In underwater,



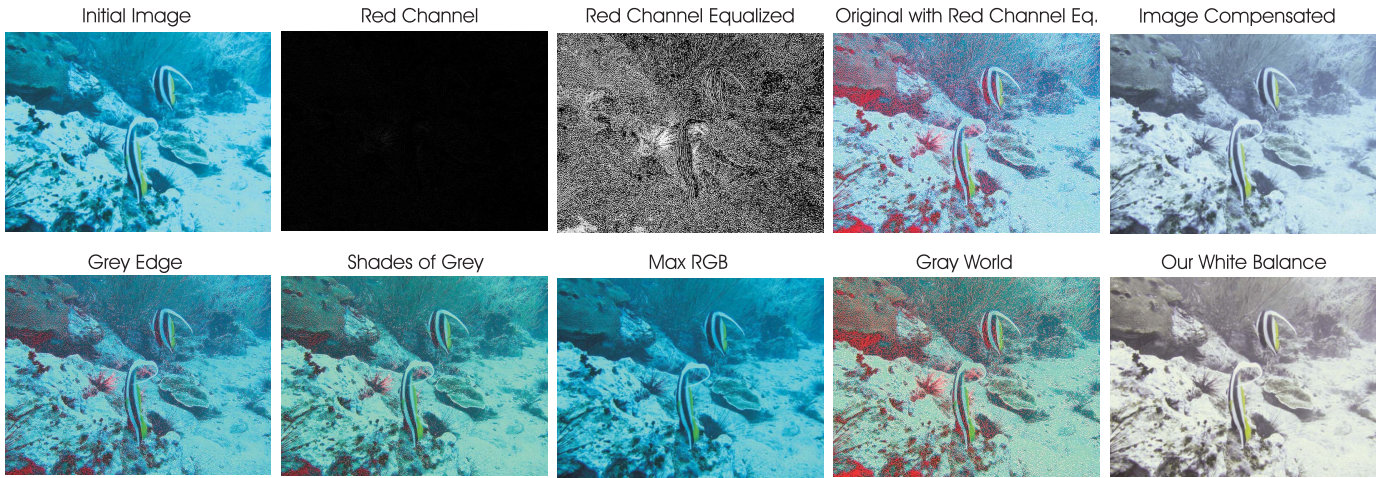


Fig. 2. Underwater white-balancing. Top row from left to right: initial underwater image with highly attenuated red channel, original red channel and red channel after histogram equalization, original image with the red channel equalized and the image after employing our compensation expressed by Eq.4. The image compensated is obtained by simply replacing the compensated red channel Eq.4 in the original underwater image. Bottom row from left to right: several well-known white balancing approaches (Gray Edge [40], Shades of Gray [41], Max RGB [42] and Gray World [43]). Please notice the artifacts (red colored locations) that appear due to the lack of information on the red channel.

the perception of color is highly correlated with the depth, and an important problem is the green-bluish appearance that needs to be rectified. As the light penetrates the water, the attenuation process affects selectively the wavelength spectrum, thus affecting the intensity and the appearance of a colored surface (see Section II). Since the scattering attenuates more the long wavelengths than the short ones, the color perception is affected as we go down in deeper water. In practice, the attenuation and the loss of color also depends on the total distance between the observer and the scene.

We have considered the large spectrum of existing white balancing methods [44], and have identified a number of solutions that are both effective and suited to our problem (see Fig. 2). In the following we briefly revise those approaches and explain how they inspired us in the derivation of our novel approach proposed for underwater scenes.

Most of those methods make a specific assumption to estimate the color of the light source, and then achieve color constancy by dividing each color channel by its corresponding normalized light source intensity. Among those methods, the Gray world algorithm [43] assumes that the average reflectance in the scene is achromatic. Hence, the illuminant color distribution is simply estimated by averaging each channel independently. The Max RGB [42] assumes that the maximum response in each channel is caused by a white patch [44], and consequently estimates the color of the light source by employing the maximum response of the different color channels. In their ‘Shades-of-Grey’ method [41], Finlayson *et al.* first observe that Max-RGB and Gray-World are two instantiations of the Minkowski  $p$ -norm applied to the native pixels, respectively with  $p = \infty$  and  $p = 1$ , and propose to extend the process to arbitrary  $p$  values. The best results are obtained for  $p = 6$ . The Grey-Edge hypothesis of van de Weijer *et al.* [40] further extends this Minkowski norm framework. It assumes the average edge difference in a scene to be achromatic, and computes the scene illumination color

by applying the Minkowski  $p$ -norm on the derivative structure of image channels, and not on the zero-order pixel structure, as done in Shades of Grey. Despite of its computational simplicity, this approach has been shown to obtain comparable results than state-of-the-art color constancy methods, such as the recent method of [45] that relies on natural image statistics.

When focusing on underwater scenes, we have found through the comprehensive study presented in Fig. 9 and Table I that the well-known Gray-World [43] algorithm achieves good visual performance for reasonably distorted underwater scenes. However, a deeper investigation dealing with extremely deteriorated underwater scenes (see Fig. 2) reveals that most traditional methods perform poorly. They fail to remove the color shift, and generally look bluish. The methods that best remove the bluish tone is the Gray World, but we observe that this method suffers from severe red artifacts. Those artifacts are due to a very small mean value for the red channel, leading to an overcompensation of this channel in locations where red is present (because Gray world divides each channel by its mean value). To circumvent this issue, following the conclusions of previous underwater works [28], [29], we therefore primarily aim to compensate for the loss of the red channel. In a second step, the Gray World algorithm will be adopted to compute the white balanced image.

To compensate for the loss of red channel, we build on the four following observations/principles:

1. The green channel is relatively well preserved under water, compared to the red and blue ones. Light with a long wavelength, i.e. the red light, is indeed lost first when traveling in clear water;
2. The green channel is the one that contains opponent color information compared to the red channel, and it is thus especially important to compensate for the stronger attenuation induced on red, compared to green. Therefore, we compensate the red attenuation by adding





Fig. 3. Compensating the red channel in Equation (4). Since for underwater scenes the blue channel contains most of the details, using this information will reduce the ability to recover certain colors such as yellow and orange and also it tends to transform the blue areas to violet shades. Compensating the red channel by masking only the green channel (Equation (4)) these limitations are reduced significantly. This is confirmed visually but also by computing the CIE2000 measure using as a reference the color checker (displayed with white ink on the resulted images).

a fraction of the green channel to red. We had initially tried to add both a fraction of green and blue to the red but, as can be observed in Fig. 3, using only the information of the green channel allows to better recover the entire color spectrum while maintaining a natural appearance of the background (water regions);

3. The compensation should be proportional to the difference between the mean green and the mean red values because, under the Gray world assumption (all channels have the same mean value before attenuation), this difference reflects the disparity/unbalance between red and green attenuation;
4. To avoid saturation of the red channel during the Gray World step that follows the red loss compensation, the enhancement of red should primarily affect the pixels with small red channel values, and should not change pixels that already include a significant red component. In other words, the green channel information should not be transferred in regions where the information of the red channel is still significant. Thereby, we want to avoid the reddish appearance introduced by the Gray-World algorithm in the over-exposed regions (see Fig. 3). Basically, the compensation of the red channel has to be performed only in those regions that are highly attenuated (see Fig. 2). This argument follows the statement in [29], telling that if a pixel has a significant value for the three channels, this is because it lies in a location near the observer, or in an artificially illuminated area, and does not need to be restored.

Mathematically, to account for the above observations, we propose to express the compensated red channel  $I_{rc}$  at every pixel location ( $x$ ) as follows:

$$I_{rc}(x) = I_r(x) + \alpha \cdot (\bar{I}_g - \bar{I}_r) \cdot (1 - I_r(x)) \cdot I_g(x), \quad (4)$$

where  $I_r$ ,  $I_g$  represent the red and green color channels of image  $I$ , each channel being in the interval  $[0, 1]$ , after normalization by the upper limit of their dynamic range;

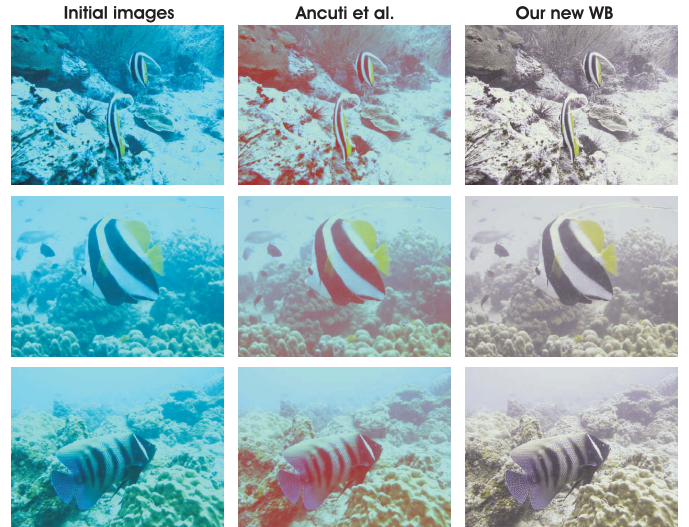


Fig. 4. Comparison to our previous white balancing approach [35].

while  $\bar{I}_r$  and  $\bar{I}_g$  denote the mean value of  $I_r$  and  $I_g$ . In Equation 4, each factor in the second term directly results from one of the above observations, and  $\alpha$  denotes a constant parameter. In practice, our tests have revealed that a value of  $\alpha = 1$  is appropriate for various illumination conditions and acquisition settings.

To complete our discussion about the severe and color-dependent attenuation of light under water, it is worth noting the works in [31] and [46]–[48], which reveal and exploit the fact that, in turbid waters or in places with high concentration of plankton, the blue channel may be significantly attenuated due to absorption by organic matter. To address those cases, when blue is strongly attenuated and the compensation of the red channel appears to be insufficient, we propose to also compensate for the blue channel attenuation, i.e. we compute the compensated blue channel  $I_{bc}$  as:

$$I_{bc}(x) = I_b(x) + \alpha \cdot (\bar{I}_g - \bar{I}_b) \cdot (1 - I_b(x)) \cdot I_g(x), \quad (5)$$

where  $I_b$ ,  $I_g$  represent the blue and green color channels of image  $I$ , and  $\alpha$  is also set to one. In the rest of the paper, the blue compensation is only considered in Figure 14. All other results are derived based on the sole red compensation.

After the red (and optionally the blue) channel attenuation has been compensated, we resort to the conventional Gray-World assumption to estimate and compensate the illuminant color cast.

As shown in Fig. 2, our white-balancing approach reduces the quantization artifacts introduced by domain stretching (the red regions in the different outputs). The reddish appearance of high intensity regions is also well corrected since the red channel is better balanced. As will be extensively discussed in Section V-A, our approach shows the highest robustness compared to the other well-known white-balancing techniques. In particular, whilst being conceptually simplest, we observe in Fig. 4 that, in cases for which the red channel of the underwater image is highly attenuated, it outperforms the white balancing strategy introduced in our conference version of our fusion-based underwater dehazing method [35].

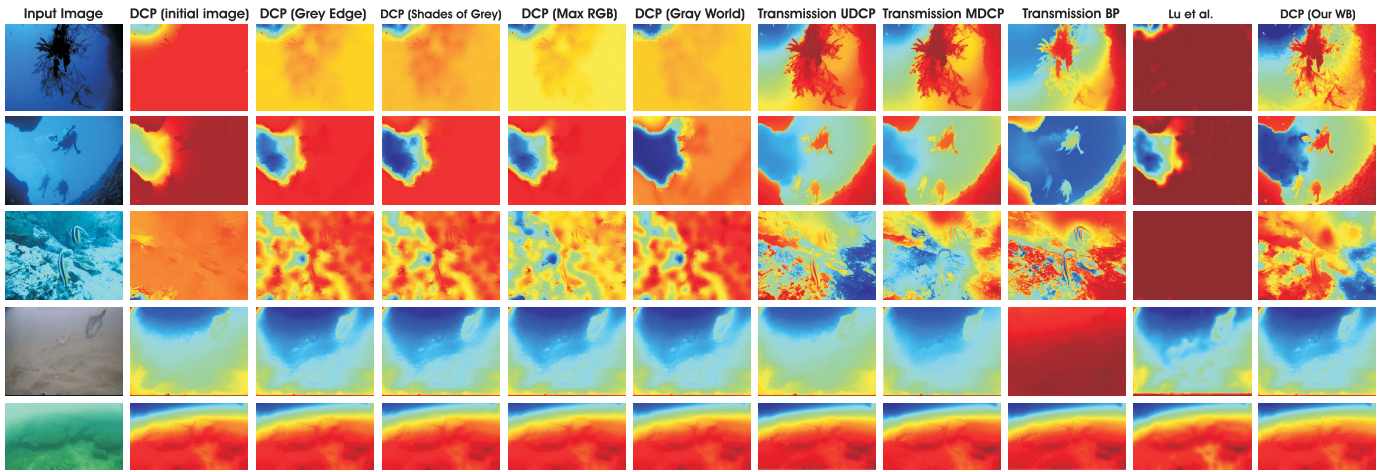


Fig. 5. Underwater transmission estimation. Columns 2 to 6, the transmission map is estimated based on DCP [26] applied to the initial underwater images but also from versions obtained with several well-known white balancing approaches (Gray Edge [40], Shades of Gray [41], Max RGB [42] and Gray World [43]) yield poor estimates. In contrast, applying DCP on our white balanced image version (last column), results in comparable and even better estimates compared with the specialized underwater techniques corresponding to UDCP [49], MDCP [50], BP [51] and Lu *et al.* [31].

Additionally, Fig. 5 shows that using our white balancing strategy yields significant improvement in estimating transmission based on the well-known DCP [26]. As can be seen in the first seven columns of Fig. 5, estimating the transmission maps based on DCP, using the initial underwater images but also the processed versions with several well-known white balancing approaches (Gray Edge [40], Shades of Gray [41], Max RGB [42] and Gray World [43]) yields poor estimates. In contrast, by simply applying DCP on our white balanced image version we obtain comparable and even better estimates than the specialized underwater techniques of UDCP [49], MDCP [50] and BP [51].

Despite white balancing is crucial to recover the color, using this correction step is not sufficient to solve the dehazing problem since the edges and details of the scene have been affected by the scattering. In the next section, we therefore propose an effective fusion based approach, relying on gamma correction and sharpening to deal with the hazy nature of the white balanced image.

#### IV. MULTI-SCALE FUSION

In this work we built on the multi-scale fusion principles to propose a single image underwater dehazing solution. Image fusion has shown utility in several applications such as image compositing [52], multispectral video enhancement [53], defogging [23], [54] and HDR imaging [55]. Here, we aim for a simple and fast approach that is able to increase the scene visibility in a wide range of underwater videos and images. Similar to [23] and [54], our framework builds on a set of inputs and weight maps derived from a single original image. In contrast to [23] and [54] however, those ones are specifically chosen in order to take the best out of the white-balancing method introduced in the previous section. In particular, as depicted in Fig.1, a pair of inputs is introduced to respectively enhance the color contrast and the edge sharpness of the white-balanced image, and the weight maps are defined to preserve the qualities and reject the

defaults of those inputs, i.e. to overcome the artifacts induced by the light propagation limitation in underwater medium.

This multi-scale fusion significantly differs from our previous fusion-based underwater dehazing approach published at IEEE CVPR [35]. To derive the inputs from the original image, our initial CVPR algorithm did assume that the back-scattering component (due to the artificial light that hits the water particles and is then reflected back to the camera) has a reduced influence. This assumption is generally valid for underwater scenes decently illuminated by natural light, but fails in more challenging illumination scenarios, as revealed by Fig. 11 in the results section. In contrast, this paper does not rely on the optical model and proposes an alternative definition of inputs and weights to deal with severely degraded scenes.

As depicted in Fig. 8 and detailed below, our underwater dehazing technique consists in three main steps: inputs derivation from the white balanced underwater image, weight maps definition, and multi-scale fusion of the inputs and weight maps.

##### A. Inputs of the Fusion Process

Since the color correction is critical in underwater, we first apply our white balancing technique to the original image. This step aims at enhancing the image appearance by discarding unwanted color casts caused by various illuminants. In water deeper than 30 ft, white balancing suffers from noticeable effects since the absorbed colors are difficult to be recovered. As a result, to obtain our *first input* we perform a gamma correction of the white balanced image version. Gamma correction aims at correcting the global contrast and is relevant since, in general, white balanced underwater images tend to appear too bright. This correction increases the difference between darker/lighter regions at the cost of a loss of details in the under-/over-exposed regions.

To compensate for this loss, we derive a *second input* that corresponds to a sharpened version of the white balanced image. Therefore, we follow the unsharp masking principle,



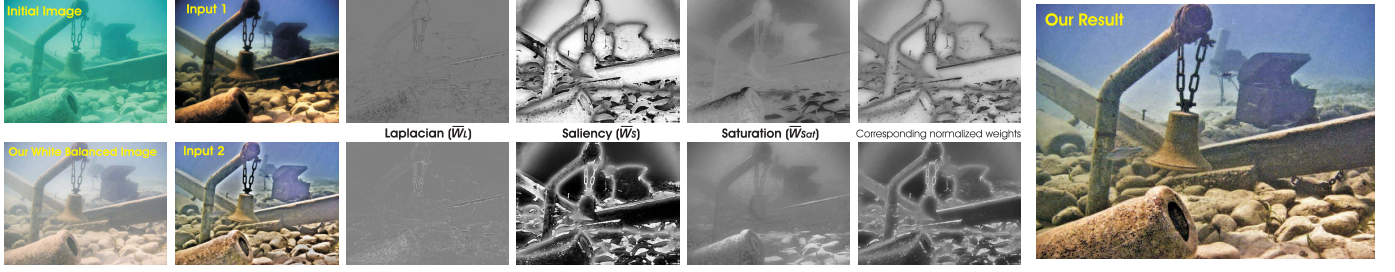


Fig. 6. The two inputs derived from our white balanced image version, the three corresponding normalized weight maps for each of them, the corresponding normalized weight maps and our final result.

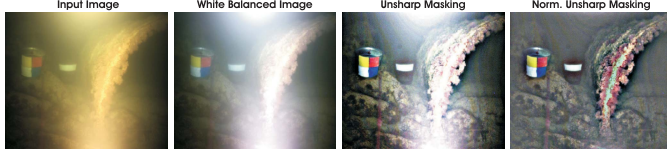


Fig. 7. Comparison between traditional unsharp masking and normalized unsharp masking applied on the white balanced image.

in the sense that we blend a blurred or unsharp (here Gaussian filtered) version of the image with the image to sharpen. The typical formula for unsharp masking defines the sharpened image  $S$  as  $S = I + \beta(I - G * I)$ , where  $I$  is the image to sharpen (in our case the white balanced image),  $G * I$  denotes the Gaussian filtered version of  $I$ , and  $\beta$  is a parameter. In practice, the selection of  $\beta$  is not trivial. A small  $\beta$  fails to sharpen  $I$ , but a too large  $\beta$  results in over-saturated regions, with brighter highlights and darker shadows. To circumvent this problem, we define the sharpened image  $S$  as follows:

$$S = (I + \mathcal{N}\{I - G * I\}) / 2, \quad (6)$$

with  $\mathcal{N}\{\cdot\}$  denoting the linear normalization operator, also named histogram stretching in the literature. This operator shifts and scales all the color pixel intensities of an image with a unique shifting and scaling factor defined so that the set of transformed pixel values cover the entire available dynamic range.

The sharpening method defined in (6) is referred to as *normalized unsharp masking* process in the following. It has the advantage to not require any parameter tuning, and appears to be effective in terms of sharpening (see examples in Fig. 7).

This second input primarily helps in reducing the degradation caused by scattering. Since the difference between white balanced image and its Gaussian filtered version is a highpass signal that approximates the opposite of Laplacian, this operation has the inconvenient to magnify the high-frequency noise, thereby generating undesired artifacts in the second input [56]. The multi-scale fusion strategy described in the next section will be in charge of minimizing the transfer of those artifacts to the final blended image.

### B. Weights of the Fusion Process

The weight maps are used during blending in such a way that pixels with a high weight value are more represented in the final image (see Fig. 6). They are thus defined based on a number of local image quality or saliency metrics.

**Laplacian contrast weight** ( $W_L$ ) estimates the global contrast by computing the absolute value of a Laplacian filter applied on each input luminance channel. This straightforward indicator was used in different applications such as tone mapping [55] and extending depth of field [57] since it assigns high values to edges and texture. For the underwater dehazing task, however, this weight is not sufficient to recover the contrast, mainly because it can not distinguish much between a ramp and flat regions. To handle this problem, we introduce an additional and complementary contrast assessment metric.

**Saliency weight** ( $W_S$ ) aims at emphasizing the salient objects that lose their prominence in the underwater scene. To measure the saliency level, we have employed the saliency estimator of Achanta *et al.* [58]. This computationally efficient algorithm has been inspired by the biological concept of center-surround contrast. However, the saliency map tends to favor highlighted areas (regions with high luminance values). To overcome this limitation, we introduce an additional weight map based on the observation that saturation decreases in the highlighted regions.

**Saturation weight** ( $W_{Sat}$ ) enables the fusion algorithm to adapt to chromatic information by advantaging highly saturated regions. This weight map is simply computed (for each input  $I_k$ ) as the deviation (for every pixel location) between the  $R_k, G_k$  and  $B_k$  color channels and the luminance  $L_k$  of the  $k^{th}$  input:

$$W_{Sat} = \sqrt{1/3 [(R_k - L_k)^2 + (G_k - L_k)^2 + (B_k - L_k)^2]} \quad (7)$$

In practice, for each input, the three weight maps are merged in a single weight map as follows. For each input  $k$ , an aggregated weight map  $W_k$  is first obtained by summing up the three  $W_L$ ,  $W_S$ , and  $W_{Sat}$  weight maps. The  $K$  aggregated maps are then normalized on a pixel-per-pixel basis, by dividing the weight of each pixel in each map by the sum of the weights of the same pixel over all maps. Formally, the normalized weight maps  $\bar{W}_k$  are computed for each input as  $\bar{W}_k = (W_k + \delta) / (\sum_{k=1}^K W_k + K \cdot \delta)$ , with  $\delta$  denoting a small regularization term that ensures that each input contributes to the output.  $\delta$  is set to 0.1 in the rest of the paper. The normalized weights of corresponding weights are shown at the bottom of Fig. 6.

Note that, in comparison with our previous work [35], we limit ourselves to these three weight maps only, and we do not compute the exposedness weight map anymore. In addition to reducing the overall complexity of the fusion process, we have observed that, when using the two inputs proposed in



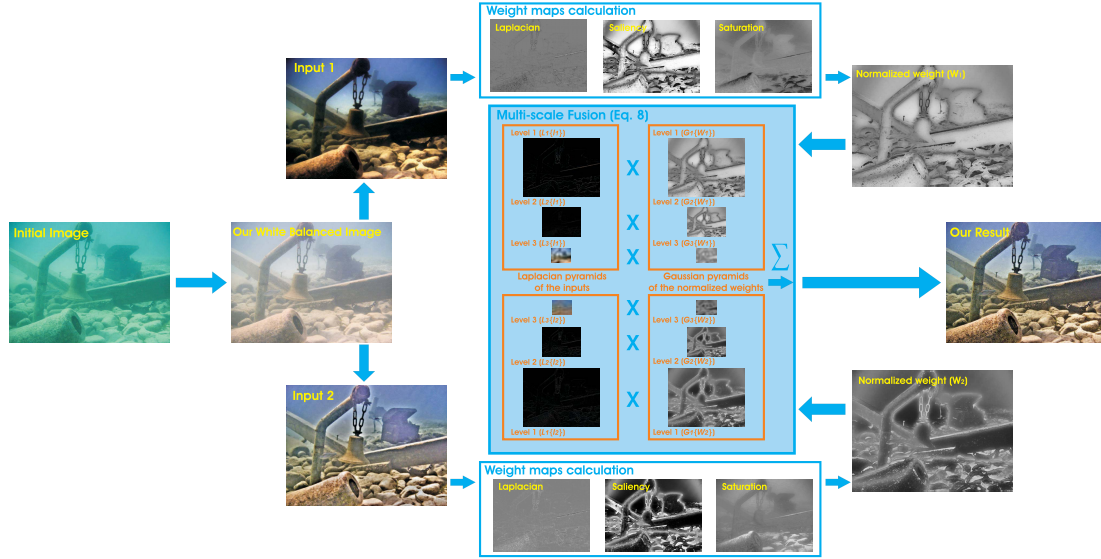


Fig. 8. Overview of our dehazing scheme. Two images, denoted Input 1 and Input 2, are derived from the (too pale) white balanced image, using Gamma correction and edge sharpening, respectively. Those two images are then used as inputs of the fusion process, which derives normalized weight maps and blends the inputs based on a multi-scale process. The multi-scale fusion approach is here exemplified with only three levels of the Laplacian and Gaussian pyramids.

this paper, the exposedness weight map tends to amplify some artifacts, such as ramp edges of our second input, and to reduce the benefit derived from the gamma corrected image in terms of image contrast. We explain this observation as follows. Originally, in an exposure fusion context [55], the exposedness weight map had been introduced to reduce the weight of pixels that are under- or over-exposed. Hence, this weight map assigns large (small) weight to input pixels that are close to (far from) the middle of the image dynamic range. In our case, since the gamma corrected input tends to exploit the whole dynamic range, the use of the exposedness weight map tends to penalize it in favor of the sharpened image, thereby inducing some sharpening artifacts and missing some contrast enhancements.

### C. Naive Fusion Process

Given the normalized weight maps, the reconstructed image  $\mathcal{R}(x)$  could typically be obtained by fusing the defined inputs with the weight measures at every pixel location ( $x$ ):

$$\mathcal{R}(x) = \sum_{k=1}^K \bar{W}_k(x) I_k(x) \quad (8)$$

where  $I_k$  denotes the input ( $k$  is the index of the inputs -  $K = 2$  in our case) that is weighted by the normalized weight maps  $\bar{W}_k$ . In practice, the naive approach introduces undesirable halos [35]. A common solution to overcome this limitation is to employ multi-scale linear [57], [59] or non-linear filters [60], [61].

### D. Multi-Scale Fusion Process

The multi-scale decomposition is based on Laplacian pyramid originally described in Burt and Adelson [57]. The pyramid representation decomposes an image into a sum of bandpass images. In practice, each level of the pyramid does

filter the input image using a low-pass Gaussian kernel  $G$ , and decimates the filtered image by a factor of 2 in both directions. It then subtracts from the input an up-sampled version of the low-pass image, thereby approximating the (inverse of the) Laplacian, and uses the decimated low-pass image as the input for the subsequent level of the pyramid. Formally, using  $G_l$  to denote a sequence of  $l$  low-pass filtering and decimation, followed by  $l$  up-sampling operations, we define the  $N$  levels  $L_l$  of the pyramid as follows:

$$\begin{aligned} I(x) &= I(x) - G_1 \{I(x)\} + G_1 \{I(x)\} \triangleq L_1 \{I(x)\} + G_1 \{I(x)\} \\ &= L_1 \{I(x)\} + G_1 \{I(x)\} - G_2 \{I(x)\} + G_2 \{I(x)\} \\ &= L_1 \{I(x)\} + L_2 \{I(x)\} + G_2 \{I(x)\} \\ &= \dots \\ &= \sum_{l=1}^N L_l \{I(x)\} \end{aligned} \quad (9)$$

In this equation,  $L_l$  and  $G_l$  represent the  $l^{th}$  level of the Laplacian and Gaussian pyramid, respectively. To write the equation, all those images have been up-sampled to the original image dimension. However, in an efficient implementation, each level  $l$  of the pyramid is manipulated at native subsampled resolution. Following the traditional multi-scale fusion strategy [55], each source input  $I_k$  is decomposed into a Laplacian pyramid [57] while the normalized weight maps  $\bar{W}_k$  are decomposed using a Gaussian pyramid. Both pyramids have the same number of levels, and the mixing of the Laplacian inputs with the Gaussian normalized weights is performed independently at each level  $l$ :

$$\mathcal{R}_l(x) = \sum_{k=1}^K G_l \{\bar{W}_k(x)\} L_l \{I_k(x)\} \quad (10)$$

where  $l$  denotes the pyramid levels and  $k$  refers to the number of input images. In practice, the number of levels  $N$  depends



Fig. 9. Comparative results of different white balancing techniques. From left to right: 1. Original underwater images taken with different cameras; 2. Gray Edge [40]; 3. Shades of Gray [41]; 4. Max RGB [42]; 5. Gray World [43]; 6. Our previous white balancing technique [35]; 7. Our new white balancing technique. The cameras used to take the pictures are Canon D10, FujiFilm Z33, Olympus Tough 6000, Olympus Tough 8000, Pentax W60, Pentax W80, Panasonic TS1. The quantitative evaluation associated to these images is provided in Table I.

TABLE I  
WHITE BALANCING QUANTITATIVE EVALUATION BASED ON CIEDE2000 AND  $Q_u$  [63] MEASURES

	Grey-Edge		Shades of Grey		max-RGB		Grey-World		Ancuti et al.		Our WB	
	CIE2000	$Q_u$	CIE2000	$Q_u$	CIE2000	$Q_u$	CIE2000	$Q_u$	CIE2000	$Q_u$	CIE2000	$Q_u$
Cannon D10	14.8368	0.5810	9.1694	0.6185	12.8245	0.5737	9.9290	0.6435	9.4534	0.6321	9.4055	0.6303
FujiFilm Z33	18.8104	0.5798	9.2788	0.7394	15.4398	0.6384	9.3965	0.7601	10.7183	0.6874	12.7031	0.6599
Olympus T6000	17.4335	0.6092	9.5245	0.6568	14.6394	0.6136	11.7903	0.6966	10.2549	0.6805	8.4598	0.7164
Olympus T8000	18.5834	0.5947	13.5059	0.6844	17.4925	0.6195	14.4812	0.6396	11.7469	0.7023	8.9903	0.7128
Panasonic TS1	12.8787	0.6501	12.9742	0.6672	15.3450	0.6756	13.3113	0.6419	12.9008	0.6814	8.9686	0.6858
Pentax W60	13.1949	0.5573	7.6364	0.6317	9.7164	0.5634	9.8853	0.7009	7.7004	0.6497	8.9194	0.6447
Pentax W80	12.9933	0.6816	8.3278	0.7510	10.7710	0.6816	8.7854	0.6993	8.3631	0.7263	7.2347	0.7652
<b>Average</b>	15.5330	0.6077	10.0596	0.6784	13.7469	0.6237	11.0827	0.6831	10.1625	0.6801	9.2402	0.6879

on the image size, and has a direct impact on the visual quality of the blended image. The dehazed output is obtained by summing the fused contribution of all levels, after appropriate upsampling.

By independently employing a fusion process at every scale level, the potential artifacts due to the sharp transitions of the weight maps are minimized. Multi-scale fusion is motivated by the human visual system, which is very sensitive to sharp transitions appearing in smooth image patterns, while being much less sensitive to variations/artifacts occurring on edges and textures (masking phenomenon). Interestingly, a recent work has shown that the multiscale process can be approximated by a computationally efficient and visually pleasant single-scale procedure [62]. This single-scale approximation should definitely be encouraged when

complexity is an issue, since it also turns the multiresolution process into a spatially localized procedure.

## V. RESULTS AND DISCUSSION

In this section, we first perform a comprehensive validation of our white-balancing approach introduced in Section IV. Then, we compare our dehazing technique with the existing specialized underwater restoration/enhancement techniques. Finally, we prove the utility of our approach for applications such as segmentation and keypoint matching.

### A. Underwater White Balancing Evaluation

Since in general the color is captured differently by various cameras, we first demonstrate that our white balancing





Fig. 10. Comparison to different outdoor (He *et al.* [19] and Ancuti and Ancuti [23]) and underwater dehazing approaches (Drews, Jr., *et al.* [28], Galdrian *et al.* [29], Emberton *et al.* [30] and our initial underwater approach [35]). The quantitative evaluation associated to these images is provided in Table II.

approach described in Section III is robust to the camera settings. Therefore, we have processed a set of underwater images that contain the standard *Macbeth Color Checker* taken by seven different professional cameras (see Fig. 9), namely Canon D10, FujiFilm Z33, Olympus Tough 6000, Olympus Tough 8000, Pentax W60, Pentax W80, Panasonic TS1. All the images have been taken approximately one meter away from the subject. The cameras have been set to their widest zoom setting, except for the Pentax W60, which was set to approximately 35mm. Due to the illumination conditions, flash was used on all cameras. On this set of images, we applied the following methods: the classical white-patch *max RGB* algorithm [42], the *Gray-World* [43], but also the more recent *Shades-of-Grey* [41] and *Gray-Edge* [40]. We compare those methods with our proposed white-balancing strategy in Fig. 9. To analyze the robustness of white balancing, we measure the dissimilarity in terms of color difference between the reference

ground truth Macbeth Color Checker and the corresponding color patch, manually located in each image.

Color differences are better represented in the perceptual  $CIEL^*a^*b^*$  color space, where  $L^*$  is the luminance,  $a^*$  is the color on a green-red scale and  $b^*$  is the color on a blue-yellow scale. Relative perceptual differences between any two colors in  $CIEL^*a^*b^*$  can be approximated by employing measures such as CIE76 and CIE94 that basically compute the Euclidean distance between them. A more complex, yet more accurate, color difference measure, which solves the perceptual uniformity issue of CIE76 and CIE94, is CIEDE2000 [68], [69]. CIEDE2000 yields values in the range [0,100], with smaller values indicating small color difference, and values less than 1 corresponding to visually imperceptible differences. Additionally, our assessment considers the index  $Q_u$  [63] that combines the benefits of SSIM index [70] and Euclidean color distance.



TABLE II  
UNDERWATER DEHAZING EVALUATION BASED ON PCQI [64], UCIQE [65] AND UIQM [66] METRICS. THE LARGER THE METRIC THE BETTER. THE CORRESPONDING IMAGES (SAME ORDER) ARE PRESENTED IN FIG.10

	He <i>et al.</i>			Ancuti&Ancuti			Drewns-Jr			Galdran <i>et al.</i>			Emberton <i>et al.</i>			Ancuti <i>et al.</i>			Our method		
	PCQI	UCIQE	UIQM	PCQI	UCIQE	UIQM	PCQI	UCIQE	UIQM	PCQI	UCIQE	UIQM	PCQI	UCIQE	UIQM	PCQI	UCIQE	UIQM	PCQI	UCIQE	UIQM
Shipwreck	1.012	0.565	0.565	0.998	0.629	0.578	0.649	0.550	0.492	0.920	0.646	0.605	0.945	0.632	0.588	1.131	0.634	0.629	1.172	0.632	0.668
Fish	1.023	0.602	0.509	1.047	0.650	0.532	0.863	0.623	0.571	0.835	0.527	0.528	1.156	0.705	0.759	1.089	0.669	0.598	1.117	0.667	0.624
Reef1	1.000	0.612	0.592	0.963	0.657	0.643	1.046	0.649	0.657	0.794	0.576	0.565	1.078	0.660	0.690	0.978	0.655	0.674	1.083	0.658	0.687
Reef2	0.774	0.702	0.749	0.899	0.683	0.724	0.483	0.659	0.653	0.769	0.633	0.671	0.607	0.718	0.757	0.983	0.718	0.733	1.075	0.711	0.781
Reef3	1.022	0.606	0.578	1.123	0.661	0.667	0.793	0.620	0.584	0.883	0.533	0.524	0.943	0.678	0.677	1.191	0.705	0.737	1.276	0.697	0.766
Galdran1	1.056	0.593	0.578	1.030	0.631	0.601	0.749	0.544	0.519	0.507	0.529	0.569	1.147	0.652	0.664	1.125	0.643	0.669	1.152	0.659	0.680
Galdran9	0.983	0.426	0.421	1.016	0.558	0.481	0.864	0.536	0.410	1.158	0.596	0.648	1.136	0.630	0.577	1.123	0.667	0.622	1.192	0.633	0.663
Ancuti1	0.860	0.485	0.353	1.032	0.561	0.412	0.909	0.499	0.383	0.962	0.641	0.458	1.036	0.499	0.407	1.074	0.588	0.547	1.022	0.594	0.507
Ancuti2	0.649	0.456	0.437	1.077	0.595	0.651	0.475	0.492	0.344	0.591	0.529	0.525	0.603	0.529	0.425	1.015	0.590	0.683	0.914	0.592	0.687
Ancuti3	1.071	0.577	0.596	1.071	0.643	0.616	0.973	0.535	0.492	1.021	0.614	0.646	1.129	0.555	0.563	1.171	0.652	0.693	1.207	0.664	0.651
Average	0.945	0.562	0.538	1.026	0.627	0.590	0.780	0.571	0.511	0.844	0.582	0.574	0.978	0.626	0.611	1.088	0.647	0.659	1.121	0.651	0.671

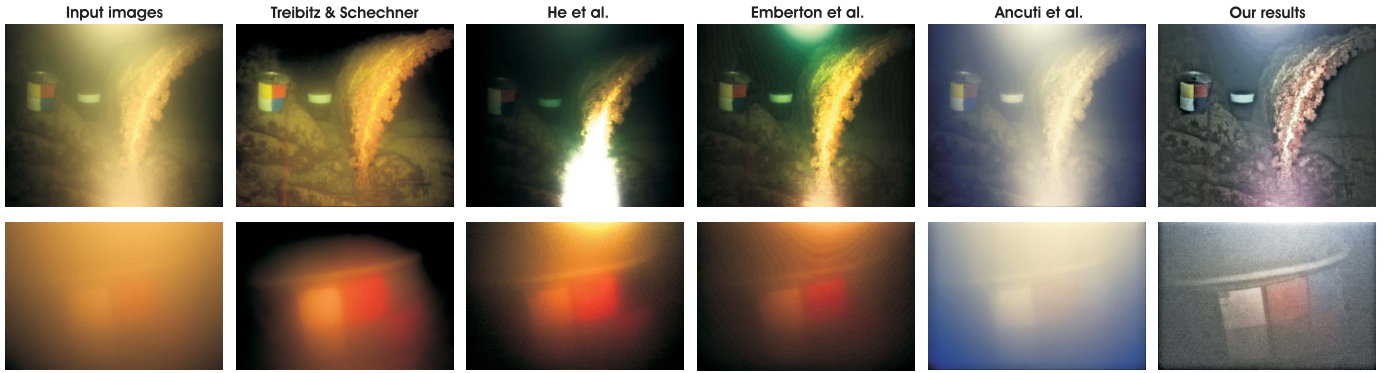


Fig. 11. Underwater dehazing of extreme scenes characterized by non-uniform illumination conditions. Our method performs better than earlier approaches of Treibitz and Schechner [67], He *et al.* [19], Emberton *et al.* [30] and Ancuti *et al.* [35].

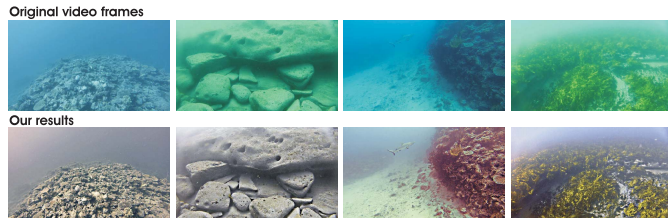


Fig. 12. Underwater video dehazing. Several video frames processed by our approach. The comparative videos can be visualized on: <https://www.youtube.com/watch?v=qspdHSTsCuQ&feature=youtu.be>

Table I shows the quantitative results obtained with the CIEDE2000 metric and index  $Q_u$ . As can be seen, these professional underwater cameras introduce various color casts, and *max RGB* and *Grey-Edge* methods are not able to remove entirely these casts. The *Gray-World* and the *Shades-of-Grey* strategy show better results, but our proposed white balance strategy shows the highest robustness in preserving the color appearance for different cameras.

### B. Underwater Dehazing Evaluation

The proposed strategy was tested for real underwater image and videos taken from different available amateur

and professional photographer collections, captured using various cameras and setups. Note that we process only 8-bit data format, making our validation relevant for common low-end cameras. For videos, the reader is referred to Fig. 12. Interestingly, our fusion-based algorithm has the advantage to employ only a reduced set of parameters that can be automatically set. Specifically, the white balancing process relies on the single parameter  $\alpha$ , which is set to 1 in all our experiments. For the multi-scale fusion, the number of decomposition levels depends on the image size, and is defined so that the size of the smallest resolution reaches a few tenth of pixels (e.g. 7 levels for an  $600 \times 800$  image size).

Fig. 10 presents the results obtained on ten underwater images, by several recent (underwater) dehazing approaches. Table II provides the associated quantitative evaluation, using three recent metrics: PCQI [64], UCIQE [65], and UIQM [66]. While PCQI is a general-purpose image contrast metric, the UCIQE and UIQM metrics are dedicated to underwater image assessment. UCIQE metric was designed specifically to quantify the nonuniform color cast, blurring, and low-contrast that characterize underwater images, while UIQM addresses three important underwater image quality criterions: colorfulness, sharpness and contrast.

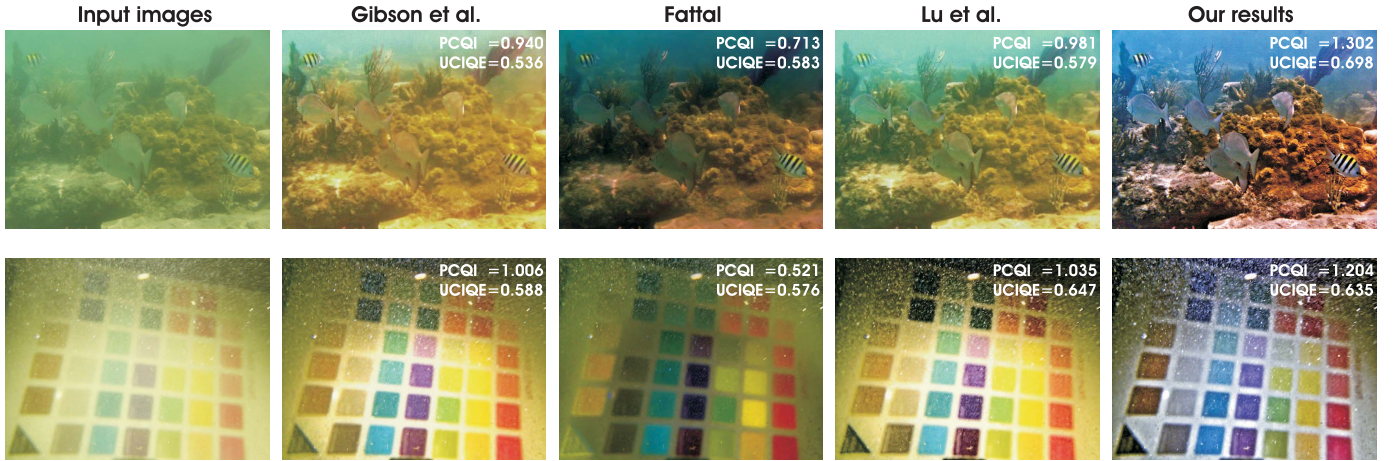


Fig. 13. Comparative results with the recent techniques of Lu *et al.* [31], Fattal [32] and Gibson *et al.* [50].

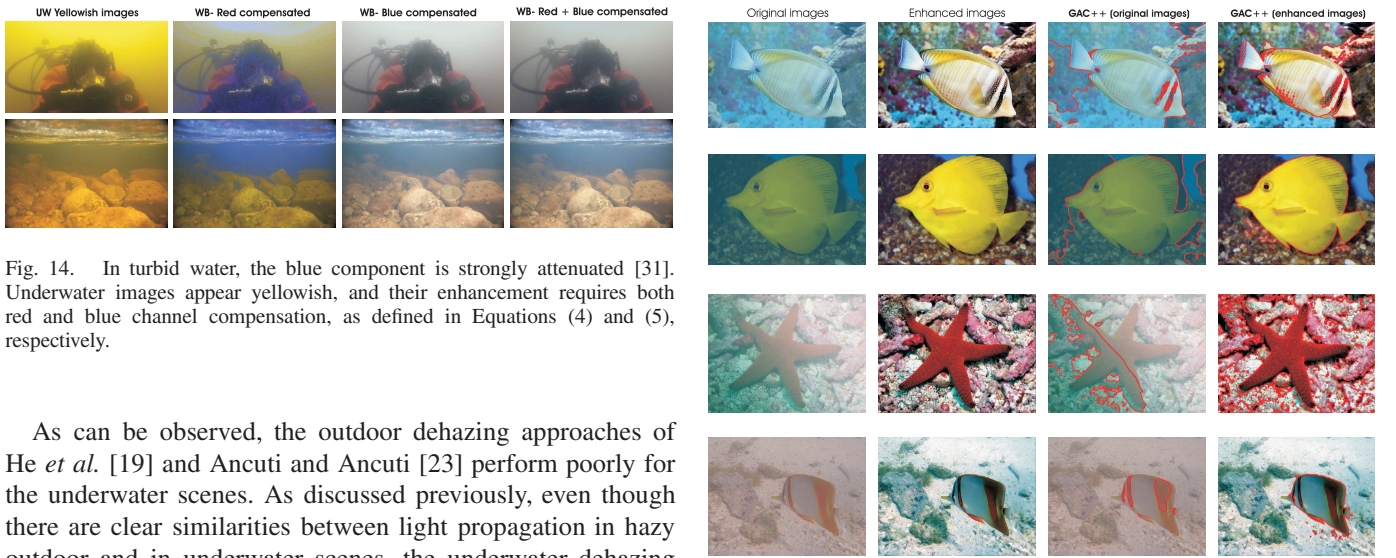


Fig. 14. In turbid water, the blue component is strongly attenuated [31]. Underwater images appear yellowish, and their enhancement requires both red and blue channel compensation, as defined in Equations (4) and (5), respectively.

As can be observed, the outdoor dehazing approaches of He *et al.* [19] and Ancuti and Ancuti [23] perform poorly for the underwater scenes. As discussed previously, even though there are clear similarities between light propagation in hazy outdoor and in underwater scenes, the underwater dehazing problem is much more challenging.

On the other hand, recent specialized underwater dehazing approaches of Emberton *et al.* [30] and Galdan *et al.* [29] show higher robustness than outdoor methods in recovering the visibility of the considered scenes. However, as attested by the last column in Fig. 10, our fusion-based approach outperforms [29], [30], having similar and in general higher values of the PCQI, UCIQE and UIQM metrics.

Compared with our initial multiscale approach presented in [35], the method introduced in this journal paper is characterized by higher robustness for extreme underwater cases, with turbid sea water and non-uniform artificial illumination. This is demonstrated by the results shown in Fig. 11 that depict two challenging underwater scenes. As can be seen, the proposed approach is able to perform better than our previous approach both in terms of contrast and color enhancement. Moreover, Fig. 11 presents the results yielded by the polarization technique of Treibitz and Schechner [67], which uses two frames taken with wide-field polarized illumination. Even though it processes only one image, our technique is able to produce an enhanced image with more details.

To complete our visual assessment, Fig. 13 compares our work with the three recent techniques of Lu *et al.* [31],

Fig. 15. **Image segmentation.** Processing underwater images with our method helps in segmenting properly. The segmentation result is more consistent while the filtered boundaries are perceptually more accurate. Here we employ the GAC++ [71] segmentation that represents a seminal geodesic active contours method.

Fattal [32], and Gibson *et al.* [50]. The PCQI and UCIQE metrics are provided for each picture, and are generally better for our approach.

Fig. 14 considers the extreme cases observed in turbid water, where the images appear yellowish due to a strong attenuation of the blue channel. For such images, compensating the red attenuation appears to be insufficient. However, interestingly, we observe that extending the color compensation to the blue component, as defined in Equation (5), significantly improves the enhanced images.

Overall, we conclude that our approach generally results in good perceptual quality, with significant enhancement of the global contrast, the color, and the image structure details. The main limitations are related to the fact that: (i) color can not always be fully restored, and (ii) some haze is maintained, especially in the scene regions that are far from the camera. Those limitations are however limited, especially when compared to previous works. The good performance



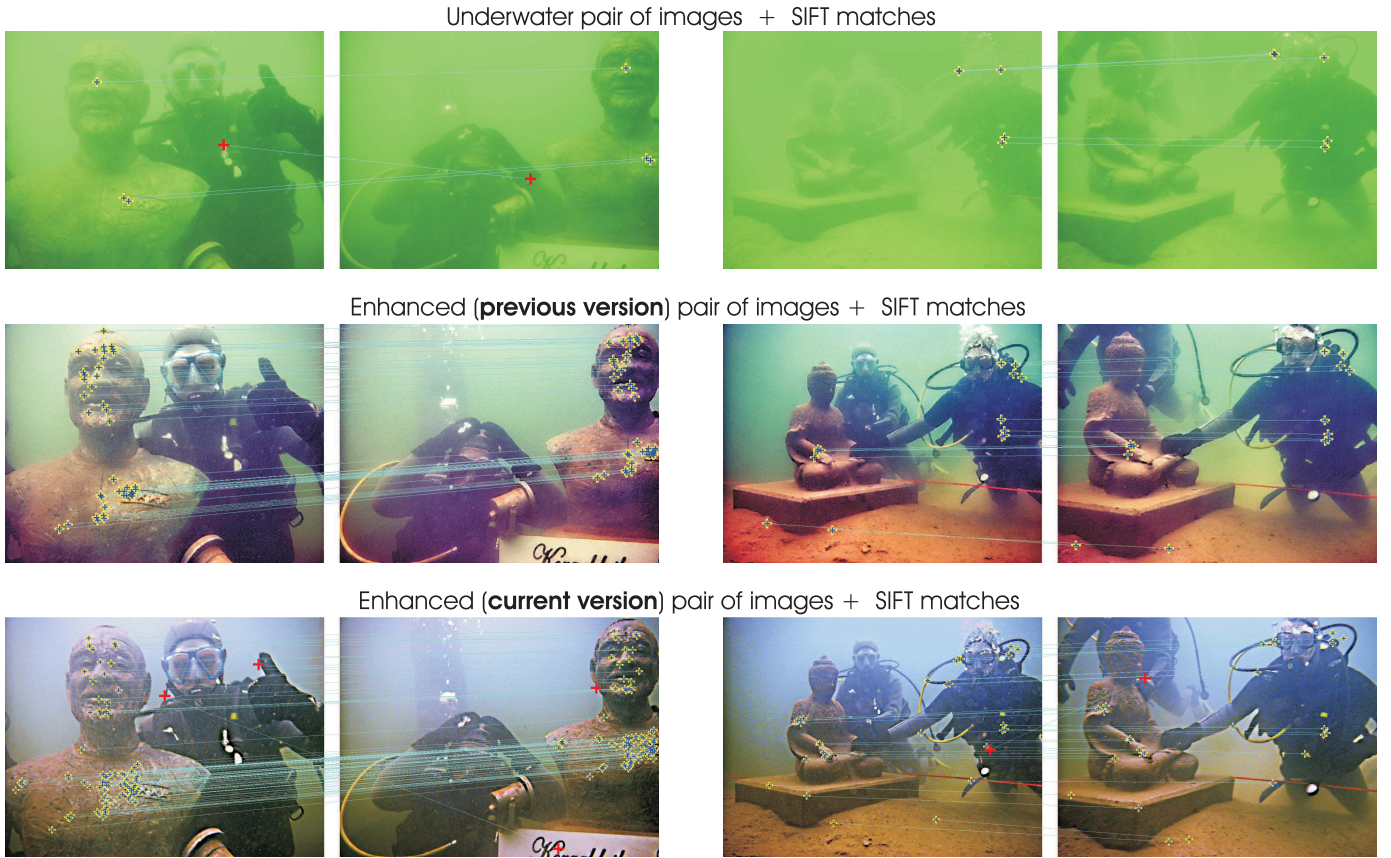


Fig. 16. **Local feature points matching.** To prove the usefulness of our approach for an image matching task we investigate how the well-known SIFT operator [72] behaves on a pair of original and dehazed underwater images. Applying standard SIFT on our enhanced versions (bottom) improves considerably the point matching process compared to processing the initial image (top), and even compared with our initial conference version of our proposed fusion-based dehazing method [35] (middle). For the left-side pair, the SIFT operator defines 3 good matches and one mismatch when applied on the original image. In contrast, it extracts 43 valid matches (without any mismatch) when applied to our previous conference version [35], and 90 valid matches (with 2 mismatches) when applied on the pair of images derived from this journal paper. For the right-side pair, SIFT finds 4 valid matches between the original images, to be compared with 14 valid matches for images enhanced with our initial approach [35], and 34 valid matches for the images derived from this paper.

of our approach is confirmed by the few more examples presented in Fig. 16 and 15. Through this visual assessment, we also observe that, despite the gray-world assumption might obviously not always be strictly valid, the enhanced images derived from our white-balanced image are constantly visually pleasant. It is our belief that the gamma correction involved in the multiscale fusion (see Fig. 1) helps in mitigating the color cast induced by an erroneous gray-world hypothesis. This is for example illustrated in Fig. 6, where input 1- obtained after gamma correction- appears more colorful than second input 2, which corresponds to a sharpened version of the white balanced image.

### C. Applications

We found our technique to be suitable for computer vision applications, as briefly described in the following section.

**Segmentation** aims at dividing images into disjoint and homogeneous regions with respect to some characteristics (e.g. texture, color). In this work, we employ the *GAC++* segmentation algorithm [71], which corresponds to a seminal geodesic active contours method (variational PDE). Fig. 15 shows that the segmentation result is more consistent when segmentation is applied to images that have been processed by our approach.

**Local feature points matching** is a fundamental task of many computer vision applications. We employ the SIFT [72] operator to compute keypoints, and compare the keypoint computation and matching processes for a pair of underwater images, with the one computed for the corresponding pair of images enhanced by our method (see Fig. 16). We use the original implementation of SIFT applied exactly in the same way in both cases. The promising achievements presented in Fig. 16 demonstrate that, by enhancing the global contrast and the local features in underwater images, our method significantly increases the number of matched pairs of keypoints.

## VI. CONCLUSIONS

We have presented an alternative approach to enhance underwater videos and images. Our strategy builds on the fusion principle and does not require additional information than the single original image. We have shown in our experiments that our approach is able to enhance a wide range of underwater images (e.g. different cameras, depths, light conditions) with high accuracy, being able to recover important faded features and edges. Moreover, for the first time, we demonstrate the utility and relevance of the proposed image enhancement technique for several challenging underwater computer vision applications.



## REFERENCES

- [1] M. D. Kocak, F. R. Dalgleish, M. F. Caimi, and Y. Y. Schechner, "A focus on recent developments and trends in underwater imaging," *Marine Technol. Soc. J.*, vol. 42, no. 1, pp. 52–67, 2008.
- [2] G. L. Foresti, "Visual inspection of sea bottom structures by an autonomous underwater vehicle," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 31, no. 5, pp. 691–705, Oct. 2001.
- [3] A. Ortiz, M. Simó, and G. Oliver, "A vision system for an underwater cable tracker," *Mach. Vis. Appl.*, vol. 13, pp. 129–140, Jul. 2002.
- [4] A. Olmos and E. Trucco, "Detecting man-made objects in unconstrained subsea videos," in *Proc. BMVC*, Sep. 2002, pp. 1–10.
- [5] B. A. Levedahl and L. Silverberg, "Control of underwater vehicles in full unsteady flow," *IEEE J. Ocean. Eng.*, vol. 34, no. 4, pp. 656–668, Oct. 2009.
- [6] C. H. Mazel, "In situ measurement of reflectance and fluorescence spectra to support hyperspectral remote sensing and marine biology research," in *Proc. IEEE OCEANS*, Sep. 2006, pp. 1–4.
- [7] Y. Kahanov and J. G. Royal, "Analysis of hull remains of the Dor D Vessel, Tantura Lagoon, Israel," *Int. J. Nautical Archeol.*, vol. 30, pp. 257–265, Oct. 2001.
- [8] R. Schettini and S. Corchs, "Underwater image processing: state of the art of restoration and image enhancement methods," *EURASIP J. Adv. Signal Process.*, vol. 2010, Dec. 2010, Art. no. 746052.
- [9] S. G. Narasimhan and S. K. Nayar, "Contrast restoration of weather degraded images," *IEEE Trans. Pattern Anal. Mach. Learn.*, vol. 25, no. 6, pp. 713–724, Jun. 2003.
- [10] D.-M. He and G. G. L. Seet, "Divergent-beam LiDAR imaging in turbid water," *Opt. Lasers Eng.*, vol. 41, pp. 217–231, Jan. 2004.
- [11] Y. Y. Schechner and Y. Averbuch, "Regularized image recovery in scattering media," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 9, pp. 1655–1660, Sep. 2007.
- [12] B. L. McGlamery, "A computer model for underwater camera systems," *Proc. SPIE*, vol. 208, pp. 221–231, Oct. 1979.
- [13] J. S. Jaffe, "Computer modeling and the design of optimal underwater imaging systems," *IEEE J. Ocean. Eng.*, vol. 15, no. 2, pp. 101–111, Apr. 1990.
- [14] H. Koschmieder, "Theorie der horizontalen sichtweite," *Beiträge Phys. Freien Atmos.*, vol. 12, pp. 171–181, 1924.
- [15] M. Levoy, B. Chen, V. Vaish, M. Horowitz, I. McDowall, and M. Bolas, "Synthetic aperture confocal imaging," in *Proc. ACM SIGGRAPH*, Aug. 2004, pp. 825–834.
- [16] S. K. Nayar and S. G. Narasimhan, "Vision in bad weather," in *Proc. IEEE ICCV*, Sep. 1999, pp. 820–827.
- [17] J. Kopf *et al.*, "Deep photo: Model-based photograph enhancement and viewing," *ACM Trans. Graph.*, vol. 27, Dec. 2008, Art. no. 116.
- [18] R. Fattal, "Single image dehazing," in *Proc. ACM SIGGRAPH*, Aug. 2008, Art. no. 72.
- [19] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," in *Proc. IEEE CVPR*, Jun. 2009, pp. 1956–1963.
- [20] J.-P. Tarel and N. Hautiere, "Fast visibility restoration from a single color or gray level image," in *Proc. IEEE ICCV*, Sep. 2009, pp. 2201–2208.
- [21] J.-P. Tarel, N. Hautiere, L. Caraffa, A. Cord, H. Halmaoui, and D. Gruyer, "Vision enhancement in homogeneous and heterogeneous fog," *IEEE Intell. Transp. Syst. Mag.*, vol. 4, no. 2, pp. 6–20, Aug. 2012.
- [22] L. Kratz and K. Nishino, "Factorizing scene albedo and depth from a single foggy image," in *Proc. IEEE ICCV*, Sep. 2009, pp. 1701–1708.
- [23] C. O. Ancuti and C. Ancuti, "Single image dehazing by multi-scale fusion," *IEEE Trans. Image Process.*, vol. 22, no. 8, pp. 3271–3282, Aug. 2013.
- [24] H. Horvath, "On the applicability of the koschmieder visibility formula," *Atmos. Environ.*, vol. 5, no. 3, pp. 177–184, Mar. 1971.
- [25] C. Ancuti, C. O. Ancuti, C. De Vleeschouwer, and A. Bovik, "Night-time dehazing by fusion," in *Proc. IEEE ICIP*, Sep. 2016, pp. 2256–2260.
- [26] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2341–2353, Dec. 2011.
- [27] J. Y. Chiang and Y.-C. Chen, "Underwater image enhancement by wavelength compensation and dehazing," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1756–1769, Apr. 2012.
- [28] P. Drews, Jr., E. Nascimento, F. Moraes, S. Botelho, M. Campos, and R. Grande-Brazil, "Transmission estimation in underwater single images," in *Proc. IEEE ICCV*, Dec. 2013, pp. 825–830.
- [29] A. Galdran, D. Pardo, A. Picón, and A. Alvarez-Gila, "Automatic red-channel underwater image restoration," *J. Vis. Commun. Image Represent.*, vol. 26, pp. 132–145, Jan. 2015.
- [30] S. Emberton, L. Chittka, and A. Cavallaro, "Hierarchical rank-based veiling light estimation for underwater dehazing," in *Proc. BMVC*, 2015, pp. 125.1–125.12.
- [31] H. Lu, Y. Li, L. Zhang, and S. Serikawa, "Contrast enhancement for images in turbid water," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 32, no. 5, pp. 886–893, May 2015.
- [32] R. Fattal, "Dehazing using color-lines," *ACM Trans. Graph.*, vol. 34, Nov. 2014, Art. no. 13.
- [33] H. Lu, Y. Li, S. Nakashima, H. Kim, and S. Serikawa, "Underwater image super-resolution by descattering and fusion," *IEEE Access*, vol. 5, pp. 670–679, 2017.
- [34] J.-B. Huang, A. Singh, and N. Ahuja, "Single image super-resolution from transformed self-exemplars," in *Proc. IEEE CVPR*, Jun. 2015, pp. 5197–5206.
- [35] C. Ancuti, C. O. Ancuti, T. Haber, and P. Bekaert, "Enhancing underwater images and videos by fusion," in *Proc. IEEE CVPR*, Jun. 2012, pp. 81–88.
- [36] S. Bazeille, I. Quidu, L. Jaulin, and J.-P. Malkasse, "Automatic underwater image pre-processing," in *Proc. Caracterisation du Milieu Marin (CMM)*, Brest, France, 2006.
- [37] A. Arnold-Bos, J.-P. Malkasset, and G. Kervern, "Towards a model-free denoising of underwater optical images," in *Proc. IEEE Eur. Oceans Conf.*, Jun. 2005, pp. 527–532.
- [38] F. Petit, A.-S. Capelle-Laizé, and P. Carre, "Underwater image enhancement by attenuation inversion with quaternions," in *Proc. IEEE ICASSP*, Apr. 2009, pp. 1177–1180.
- [39] M. Chambah, D. Semani, A. Renouf, P. Courtellemont, and A. Rizzi, "Underwater color constancy: Enhancement of automatic live fish recognition," *Proc. SPIE*, vol. 5293, pp. 157–169, Dec. 2003.
- [40] J. van de Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy," *IEEE Trans. Image Process.*, vol. 16, no. 9, pp. 2207–2214, Sep. 2007.
- [41] G. D. Finlayson and E. Trezzi, "Shades of gray and colour constancy," in *Proc. 12th Color Imag. Conf., Color Sci., Syst. Appl., Soc. Imag. Sci. Technol.*, 2004, pp. 37–41.
- [42] E. H. Land, "The Retinex theory of color vision," *Sci. Amer.*, vol. 237, no. 6, pp. 108–128, Dec. 1977.
- [43] G. Buchsbaum, "A spatial processor model for object colour perception," *J. Franklin Inst.*, vol. 310, no. 1, pp. 1–26, Jul. 1980.
- [44] M. Ebner, *Color Constancy*, 1st ed. Hoboken, NJ, USA: Wiley, 2007.
- [45] A. Gijsenij and T. Gevers, "Color constancy using natural image statistics and scene semantics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 4, pp. 687–698, Apr. 2011.
- [46] H. Lu *et al.*, "Underwater image enhancement method using weighted guided trigonometric filtering and artificial light correction," *J. Vis. Commun. Image Represent.*, vol. 38, pp. 504–516, Jul. 2016.
- [47] F. Bonin, A. Burguera, and G. Oliver, "Imaging systems for advanced underwater vehicles," *J. Maritime Res.*, vol. 8, no. 1, pp. 65–86, 2011.
- [48] S. Serikawa and H. Lu, "Underwater image dehazing using joint trilateral filter," *Comput. Elect. Eng.*, vol. 40, no. 1, pp. 41–50, 2014.
- [49] P. L. J. Drews, Jr., E. R. Nascimento, S. S. C. Botelho, and M. F. M. Campos, "Underwater depth estimation and image restoration based on single images," *IEEE Comput. Graph. Appl.*, vol. 36, no. 2, pp. 24–35, Mar./Apr. 2016.
- [50] K. B. Gibson, D. T. Vo, and T. Q. Nguyen, "An investigation of dehazing effects on image and video coding," *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 662–673, Feb. 2012.
- [51] N. Carlevaris-Bianco, A. Mohan, and R. M. Eustice, "Initial results in underwater single image dehazing," in *Proc. IEEE OCEANS*, Sep. 2010, pp. 1–8.
- [52] M. Grundland, R. Vohra, G. P. Williams, and N. A. Dodgson, "Cross dissolve without cross fade: Preserving contrast, color and salience in image compositing," *Comput. Graph. Forum*, vol. 25, no. 3, pp. 577–586, 2006.
- [53] E. P. Bennett, J. L. Mason, and L. McMillan, "Multispectral bilateral video fusion," *IEEE Trans. Image Process.*, vol. 16, no. 5, pp. 1185–1194, May 2007.
- [54] C. O. Ancuti, C. Ancuti, and P. Bekaert, "Effective single image dehazing by fusion," in *Proc. IEEE ICIP*, Sep. 2010, pp. 3541–3544.
- [55] T. Mertens, J. Kautz, and F. Van Reeth, "Exposure fusion: A simple and practical alternative to high dynamic range photography," *Comput. Graph. Forum*, vol. 28, no. 1, pp. 161–171, 2009.
- [56] G. C. Rafael and W. E. Richard, *Digital Image Processing*. Englewood Cliffs, NJ, USA: Prentice-Hall, 2008.

- [57] P. Burt and T. Adelson, "The laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. COM-31, no. 4, pp. 532–540, Apr. 1983.
- [58] R. Achanta, S. Hemamiz, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," in *Proc. IEEE CVPR*, Jun. 2009, pp. 1597–1604.
- [59] T. Pu and G. Ni, "Contrast-based image fusion using the discrete wavelet transform," *Opt. Eng.*, vol. 39, pp. 2075–2082, Aug. 2000.
- [60] S. Paris and F. Durand, "A fast approximation of the bilateral filter using a signal processing approach," *Int. J. Comput. Vis.*, vol. 81, no. 1, pp. 24–52, Jan. 2009.
- [61] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski, "Edge-preserving decompositions for multi-scale tone and detail manipulation," in *Proc. ACM SIGGRAPH*, 2008, Art. no. 67.
- [62] C. O. Ancuti, C. Ancuti, C. De Vleeschouwer, and A. C. Bovik, "Single-scale fusion: An effective approach to merging images," *IEEE Trans. Image Process.*, vol. 26, no. 1, pp. 65–78, Jan. 2017.
- [63] Y. Li, H. Lu, J. Li, X. Li, Y. Li, and S. Serikawa, "Underwater image de-scattering and classification by deep neural network," *Comput. Electr. Eng.*, vol. 54, pp. 68–77, Aug. 2016.
- [64] S. Wang, K. Ma, H. Yeganeh, Z. Wang, and W. Lin, "A patch-structure representation method for quality assessment of contrast changed images," *IEEE Signal Process. Lett.*, vol. 22, no. 12, pp. 2387–2390, Dec. 2015.
- [65] M. Yang and A. Sowmya, "An underwater color image quality evaluation metric," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 6062–6071, Dec. 2015.
- [66] K. Panetta, C. Gao, and S. Agaian, "Human-visual-system-inspired underwater image quality measures," *IEEE J. Ocean. Eng.*, vol. 41, no. 3, pp. 541–551, Jul. 2015.
- [67] T. Treibitz and Y. Y. Schechner, "Active polarization descattering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 3, pp. 385–399, Mar. 2009.
- [68] G. Sharma, W. Wu, and E. N. Dalal, "The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations," *Color Res. Appl.*, vol. 30, no. 1, pp. 21–30, Feb. 2005.
- [69] S. Westland, C. Ripamonti, and V. Cheung, *Computational Colour Science Using MATLAB*, 2nd ed. Hoboken, NJ, USA: Wiley, 2005.
- [70] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [71] G. Papandreou and P. Maragos, "Multigrid geometric active contour models," *IEEE Trans. Image Process.*, vol. 16, no. 1, pp. 229–240, Jan. 2007.
- [72] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

**Codruta O. Ancuti** received the Ph.D. degree from Hasselt University, Belgium, in 2011. From 2015 to 2017, she was a Research Fellow with the ViCOROB Group, University of Girona, Spain. She is a Senior Researcher/Lecturer with the Faculty of Electrical and Telecommunication Engineering, University Politehnica Timisoara. Her main interest of research includes image understanding and visual perception. She is the first that introduced several single images-based enhancing techniques built on the multi-scale fusion (e.g., color-to-grayscale, image dehazing, underwater image, and video restoration). Her work received the Best Paper Award at NTIRE 2017 (CVPR workshop).

**Cosmin Ancuti** received the Ph.D. degree from Hasselt University, Belgium, in 2009. He was a Post-Doctoral Fellow with IMINDS and Intel Exascience Laboratory (IMEC), Leuven, Belgium, from 2010 to 2012, and a Research Fellow with the University Catholique of Louvain, Belgium, from 2015 to 2017. He is currently a Senior Researcher/Lecturer with University Politehnica Timisoara. He is the author of over 50 papers published in international conference proceedings and journals. His area of interests includes image and video enhancement techniques, computational photography, and low-level computer vision.

**Christophe De Vleeschouwer** was a Senior Research Engineer with IMEC from 1999 to 2000, a Post-Doctoral Research Fellow with the University of California at Berkeley from 2001 to 2002 and EPFL in 2004, and a Visiting Faculty with Carnegie Mellon University from 2014 to 2015. He is a Research Director with the Belgian NSF, and an Associate Professor with the ISP Group, Universite Catholique de Louvain, Belgium. He has co-authored over 35 journal papers or book chapters, and holds two patents. His main interests lie in video and image processing for content management, transmission, and interpretation. He is enthusiastic about nonlinear and sparse signal expansion techniques, the ensemble of classifiers, multiview video processing, and the graph-based formalization of vision problems. He served as an Associate Editor of the IEEE TRANSACTIONS ON MULTIMEDIA, and has been a (Technical) Program Committee Member for most conferences dealing with video communication and image processing.

**Philippe Bekaert** is a Professor in visual computing with the Expertise Center for Digital Media, Hasselt University, Belgium. His research areas include omnidirectional and free viewpoint video, immersive environments, and general purpose GPU computing. His group is participating in several large scale EU and national research and development projects on the intersection of arts, cinema, TV broadcasting, and visual computing. He strongly advocates joint creative and technological research. He is cofounder and CTO of a spin-off company AZilPix, commercializing a broadcast video production system following a novel approach.