Enhancing Underwater Images and Videos by Fusion

Cosmin Ancuti, Codruta Orniana Ancuti, Tom Haber and Philippe Bekaert
Hasselt University - tUL -IBBT, EDM, Belgium

Abstract

This paper describes a novel strategy to enhance underwater videos and images. Built on the fusion principles, our strategy derives the inputs and the weight measures only from the degraded version of the image. In order to overcome the limitations of the underwater medium we define two inputs that represent color corrected and contrast enhanced versions of the original underwater image/frame, but also four weight maps that aim to increase the visibility of the distant objects degraded due to the medium scattering and absorption. Our strategy is a single image approach that does not require specialized hardware or knowledge about the underwater conditions or scene structure. Our fusion framework also supports temporal coherence between adjacent frames by performing an effective edge preserving noise reduction strategy. The enhanced images and videos are characterized by reduced noise level, better exposedness of the dark regions, improved global contrast while the finest details and edges are enhanced significantly. In addition, the utility of our enhancing technique is proved for several challenging applications.

1. Introduction

Underwater imaging is challenging due to the physical properties existing in such environments. Different from common images, underwater images suffer from poor visibility due to the attenuation of the propagated light. The light is attenuated exponentially with the distance and depth mainly due to absorption and scattering effects. The absorption substantially reduces the light energy while the scattering causes changes in the light direction. The random attenuation of the light is the main cause of the foggy appearance while the the fraction of the light scattered back from the medium along the sight considerably degrades the scene contrast. These properties of the underwater medium yields scenes characterized by poor contrast where distant objects appear misty. Practically, in common sea water, the objects at a distance of more than 10 meters are almost indistinguishable while the colors are faded since their characteristic wavelengths are cut according to the water depth.

There have been several attempts to restore and enhance the visibility of such degraded images. Mainly, the problem can be tackled by using multiple images [21], specialized hardware [15] and by exploiting polarization filters [25]. Despite their effectiveness to restore underwater images, these strategies have demonstrated several important issues that reduce their practical applicability. First, the hardware solutions (e.g. laser range-gated technology and synchronous scanning) are relatively expensive and complex. The multiple-image solutions require several images of the same scene taken in different environment conditions. Similarly, polarization methods process several images that have different degrees of polarization. While this is relatively feasible for outdoor hazy and foggy images, for the underwater case, the setup of the camera might be troublesome. In addition, these methods (except the hardware solutions) are not able to deal with dynamic scenes, thus being impractical for videos.

In this paper, we introduce a novel approach that is able to enhance underwater images based on a single image, as well as videos of dynamic scenes. Our approach is built on the fusion principle that has shown utility in several applications such as image compositing [14], multispectral video enhancement [6], defogging [2] and HDR imaging [20]. In contrast to these methods, our fusion-based approach does not require multiple images, deriving the inputs and the weights only from the original degraded image. We aim for a straightforward and computationally inexpensive that is able to perform relatively fast on common hardware. Since the degradation process of underwater scenes is both multiplicative and additive [26] traditional enhancing techniques like white balance, color correction, histogram equalization shown strong limitations for such a task. Instead of directly filtering the input image, we developed a fusion-based scheme driven by the intrinsic properties of the original image (these properties are represented by the weight maps). The success of the fusion techniques is highly dependent on the choice of the inputs and the weights and therefore we investigate a set of operators in order to overcome limitations specific to underwater environments. As a result, in our framework the degraded image is firstly white balanced in order to remove the color

978-1-4673-1228-8/12/$31.00 ©2012 IEEE

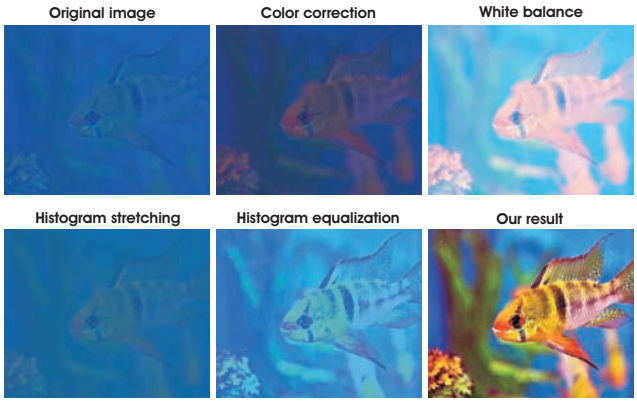

Figure 1. Traditional enhancing techniques that are found in commercial tools presents limitations when dealing with underwater images.

casts while producing a natural appearance of the sub-sea images. This partially restored version is then further enhanced by suppressing some of the undesired noise. The second input is derived from this filtered version in order to render the details in the entire intensity range. Our fusion-based enhancement process is driven by several weight maps. The weight maps of our algorithm assess several image qualities that specify the spatial pixel relationships. These weights assign higher values to pixels to properly depict the desired image qualities. Finally, our process is designed in a multi-resolution fashion that is robust to artifacts. Different than most of the existing techniques, our framework can deal with dynamic scenes. To preserve temporal coherence in videos we apply temporal bilateral filtering [6] between adjacent frames.

Contributions. This paper introduces the following main contributions:
1. A straightforward fusion-based framework that effectively blends different well known filters in order to enhance underwater images based on a single input.
2. Our strategy is able to enhance underwater videos of dynamic scenes. Until now, this was demonstrated only using hardware-based solutions.
3. A robust white balancing techniques specialized for underwater scenes that was validated based on an extensive study.
4. We demonstrate that the simple Laplacian pyramid yields effective results comparable with the recent edge preserving filters such as WLS [10].
5. To the best of our knowledge we are the first that demonstrate utility of an underwater enhancement/restoration technique for several complex applications such as segmentation, image matching by local feature points and image dehazing.

2. Our Enhancing Approach

In this work we propose an alternative single image-based solution built on the multi-scale fusion principles. We aim for a simple and fast approach that is able to increase the visibility of a wide variation of underwater videos and images. Even though we do not explicitly follow specialized optical models (e.g. McGlamery [19]), our framework blends specific inputs and weights carefully chosen in order to overcome the limitation of such environments. For the most of the processed images shown in this paper and in the supplementary material the back-scattering component (yielded in general due to the artificial light that hits the water particles and then is reflected back to the camera) has a reduced influence. This is generally valid for underwater scenes decently illuminated by natural light. However, even when artificial illumination is needed, the influence of this component can be easily diminished by modifying the angle of the source light [16].

Our enhancing strategy consists of three main steps: inputs assignment (derivation of the inputs from the original underwater image), defining weight measures and multi-scale fusion of the inputs and weight measures.

2.1. Inputs of the Fusion Process

When applying a fusion algorithm the key to obtain good visibility of the final result is represented by the well tailored inputs and weights. Different than most of the existing fusion methods (however, none of them designed to deal with underwater scenes), our fusion technique processes only a single degraded image. The general idea of image fusion is that the processed result, combines several input images by preserving only the most significant features of them. Thus, results obtained by a fusion-based approach fulfills the depiction expectation when each part of the result presents an appropriate appearance in at least one of the input images. In our single-based image approach two inputs of the fusion process are derived from the original degraded image. Our enhancing solution does not search to derive the inputs based on the physical model of the scene, since the existing models are quite complex to be tackled. Instead, we aim for a fast and simple technique that works generally. The first derived input is represented by the color corrected version of the image while the second is computed as a contrast enhanced version of the underwater image after a noise reduction operation is performed (see figure 2).

2.1.1 White Balancing of the Inputs

White balancing is an important processing step that aims to enhance the image appearance by discarding unwanted color casts, due to various illuminants. In water deeper than 30 ft, white balancing suffers from noticeable effects since the absorbed colors are difficult to be restored. Additionally,

underwater scenes present significant lack of contrast due to the poor light propagation in this type of medium.

Considering the large availability of white balancing methods [9] we have searched for a proper solution to our problem. In the following are briefly revised several important approaches that we have analyzed (more in-depth details are found on [9]).

The Finlayson's approach Shades-of-Grey [12] computes the illumination of the scene for each channel by using the Minkowski p -norm. For $p = 1$, this expression is a particular case of the Gray-World [9] while for $p = \infty$ it approximates the White-Patch hypothesis [9]. Despite of its simplicity, the low-level approach of Finlayson and Trezzi [12] has shown to yield comparative results to those of more complex white balance algorithms such as the recent method of [13] that relies on on natural image statistics. The Grey-Edge hypothesis of Weijer and Gevers [29], similarly with Shades-of-Grey [12] can also be formulated by extending the p-th Minkowski form.

In our experiments, we have noticed that solutions derived from the White-Patch algorithm [9] generally fail since the underwater images contain only reduced regions with specular reflection. Additionally, the solution of Gray-Edge algorithm [29] performs poorly in such cases, mainly due to the fact that underwater images are characterized by low contrast and less visible edges than natural images. However, we found that the most appropriate strategy is the Gray-World approach of Buchsbaum et al. [9]. One common problem noticed during our tests for most of the white-balancing techniques (observed either for the entire image or only for small regions of the image) is the color-deviation [9] that appears when the illumination is poorly estimated. For instance in the underwater images, where the appearance is overall blue, the parts that are miss-balanced will show reddish appearance(that corresponds to the opponent color of the illumination).

Our approach minimizes this effect of color shifting for the entire scene as can be noticed into the figure 7 that presents some comparative results. Related to these previous approaches, our solution is similar to the Shades-of-Grey [12] but much computationally effective. We found to be more robust to increase the average value estimated with a percentage λ instead to variate the norm value of p.

As a result, in our framework, the illumination is estimated by the value μ_I that is computed from the average of the scene μ_{ref} and adjusted by the parameter λ:

$$\mu_I = 0.5 + \lambda \mu_{ref} \tag{1}$$

The average color μ_{ref} is used to estimate the illuminant color (a common solution derived from Gray-World [9]) and can be obtained based on Minkowski norm when $p = 1$. Furthermore, to assign parameter λ we analyze the density and the distribution on the color histogram. Consequently,

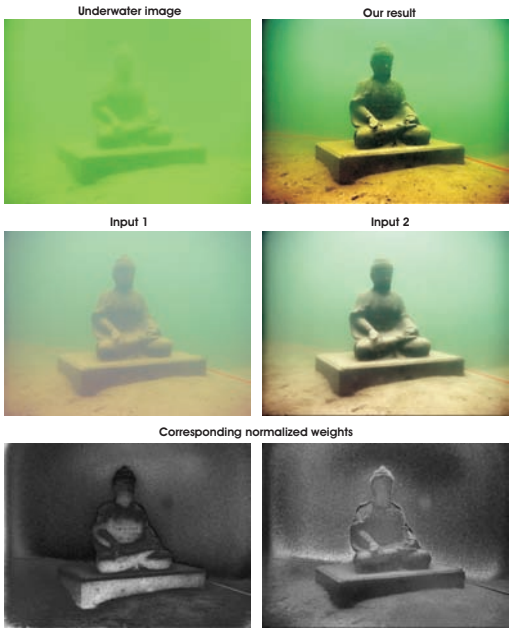

Figure 2. Top line: degraded image and our result; Middle line: the two inputs derived from the original image required by our fusion approach; Bottom line: the corresponding normalized weight maps $\bar{W}$.

we set a higher value for λ when the detected set of colors is small. The value that variates in the range $[0, 0.5]$ of λ decreases inverse-proportionally with the number of colors. In general, we have observed that a default value of 0.2 yields visually pleasing results (since most of the processed underwater images present a relative uniform color distribution). Despite of its simplicity, our white balance strategy is able to remove effectively the color cast but also to recover the white and gray shades of the image, while producing a natural appearance of the output. Our method overcomes the standard Gray-World as well as the other considered techniques (please refer to the supplementary material for an extensive study of different white balance techniques applied for underwater images).

Practically, the ***first input*** of the fusion process is computed based on this straightforward white balancing operation. Nevertheless, white balancing solely is not able to solve the problem of visibility, and therefore we derive an additional input (described in the next subsection) in order to enhance the contrast of the degraded image.

2.1.2 Temporal Coherent Noise Reduction

Due to the impurities and the special illumination conditions, underwater images are noisy. Removing noise while preserving edges of an input image enhances the sharpness and may be accomplished by different strategies such as median filtering, anisotropic diffusion and bilateral filtering. However, for videos this task is more challenging since

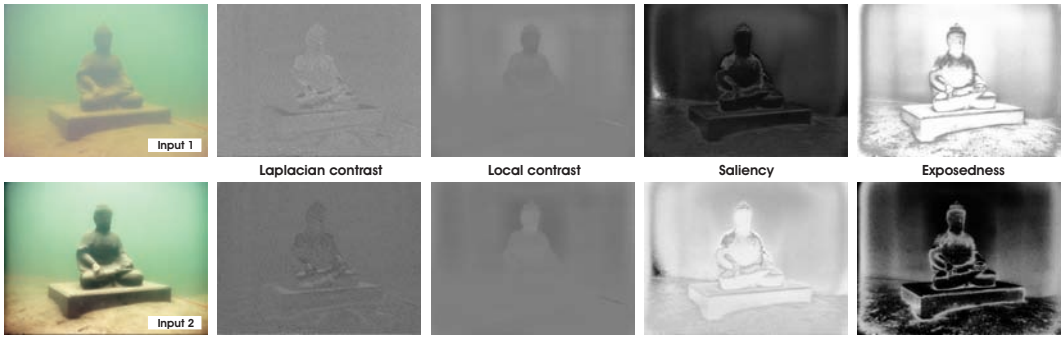

Figure 3. The two inputs derived from the original image presented in previous figure and the corresponding normalized weight maps.

both spatial and temporal coherence need to be taken into account. The bilateral filter [28, 23] is one of the common solutions being an non-iterative edge-preserving smoothing filter that has proven usefull for several problems such as tone mapping, mesh smoothing and dual photography enhancement. By considering the domain Ω of the spatial filter kernel f (Gaussian with standard deviation σ_f), the bilateral filter blends the center pixel s of the kernel with the neighboring pixels p that are similar to s:

$$J_s = \frac{1}{k(s)} \sum_{p \in \Omega} f(p - s, \sigma_f) g(D(p, s), \sigma_g) I_p \qquad (2)$$

where $D(p, s) = I_p - I_s$ is the difference in intensities, the normalization $k(s) = \sum_{p \in \Omega} f(p - s, \sigma_f) g(D(p, s), \sigma_g)$, g is the range kernel that is a Gaussian with standard deviation σ_g that penalizes pixels across edges that have large intensity differences.

However, the bilateral filter does not guarantee the preservation of the temporal coherence for videos. Even though, in the absence of motion, a simple average of all pixels at each coordinate through time would represent a decent solution, for real dynamic scenes this naive strategy yields undesirable ghosting artifacts. Inspired by Bennet et al. [6] where the solution was used in context of multi-spectral video fusion, we employ a *temporal bilateral filter* strategy on the white balanced version of the frames that aims to reduce noise and smoothing frames while preserving temporal coherence. Choosing an appropriate value of σ_g to simultaneously deal with noise while still preserving edges is difficult. High values of σ_g would yield halos while small values of σ_g are not able to reduce sufficiently undesired noise. Instead of just comparing the intensities as $D(p, s) = I_p - I_s$, we compute the sum of squared differences (SSD) between small spatial neighborhood Ψ around s and p weighted by a Gaussian $\Gamma(x, y)$:

$$D(p, s) = \sum_{x}^{\Psi} \sum_{y}^{\Psi} \Gamma(x, y)(I_p - I_s)^2 \qquad (3)$$

Typically, the size of neighborhood Ψ is 3×3 or 5×5. This simple approach significantly reduces the ambiguity between noise and edges since the larger neighborhood Ψ reduces the impact of single-pixel temporal noise.

In our fusion framework, the ***second input*** is computed from the noise-free and color corrected version of the original image. This input is designed in order to reduce the degradation due to volume scattering. To achieve an optimal contrast level of the image, the second input is obtained by applying the classical contrast local adaptive histogram equalization [30]. To generate the second derived image common global operators can be applied as well. Since these are defined as some parametric curve, they need to be either specified by the user or to be estimated from the input image. Commonly, the improvements obtained by these operators in different regions are done at the expense of the remaining regions. We opted for the local adaptive histogram since it works in a fully automated manner while the level of distortion is minor. This technique expands the contrast of the feature of interest in order to simultaneously occupy a larger portion of the intensity range than the initial image. The enhancement is obtained since the contrast between adjacent structures is maximally portrayed. To compute this input several more complex methods, such as the gradient domains or gamma correction multi-scale Retinex (MSR) [8], may be used as well.

2.2. Weights of the Fusion Process

The design of the weight measures needs to consider the desired appearance of the restored output. We argue that image restoration is tightly correlated with the color appearance, and as a result the measurable values such as salient features, local and global contrast or exposedness are difficult to integrate by naive per pixel blending, without risking to introduce artifacts. Higher values of the weight determines that a pixel is advantaged to appear in the final image (see figure 3).

Laplacian contrast weight (W_L) deals with global contrast by applying a Laplacian filter on each input luminance channel and computing the absolute value of the filter result. This straightforward indicator was used in different applications such as tone mapping [20] and extending depth

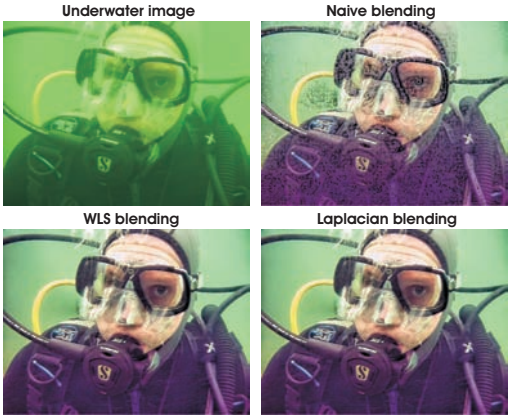

Figure 4. ***Blending strategies***. Considering the underwater image (top-left) by directly performing the naive blending (equation 6) yields unpleasing artifacts (top-right). On the other hand, by employing the multi-scale approaches based on WLS filter [Farbman et al 2008] (bottom-left) and Laplacian pyramid yield significant improvements. As may be observed, the difference between WLS and Laplacian pyramid is negligible.

of field [7] since it assigns high values to edges and texture. For the underwater restoration task, however, this weight is not sufficient to recover the contrast, mainly because it can not distinguish between a ramp and flat regions. To handle this problem, we searched for an additional contrast measurement that independently assess the local distribution.

Local contrast weight (W_{LC}) comprises the relation between each pixel and its neighborhoods average. The impact of this measure is to strengthen the local contrast appearance since it advantages the transitions mainly in the highlighted and shadowed parts of the second input. The (W_{LC}) is computed as the standard deviation between pixel luminance level and the local average of its surrounding region:

$$W_{LC}(x,y) = \left\| I^k - I^k_{\omega_{hc}} \right\| \tag{4}$$

where I^k represents the luminance channel of the input and the $I^k_{\omega_{hc}}$ represents the low-passed version of it. The filtered version $I^k_{\omega_{hc}}$ is obtained by employing a small 5×5 ($\frac{1}{16}[1,4,6,4,1]$) separable binomial kernel with the high frequency cut-off value $\omega_{hc} = \pi/2.75$. For small kernels the binomial kernel is a good approximation of its Gaussian counterpart, and it can be computed more effectively.

Saliency weight (W_S) aims to emphasize the discriminating objects that lose their prominence in the underwater scene. To measure this quality, we have employed the saliency algorithm of Achanta et al. [1]. This computationally efficient saliency algorithm is straightforward to be implemented being inspired by the biological concept of center-surround contrast. However, the saliency map tends to favor highlighted areas. To increase the accuracy of results, we introduce the exposedness map to protect the mid tones that might be altered in some specific cases.

Exposedness weight (W_E) evaluates how well a pixel is exposed. This assessed quality provides an estimator to preserve a constant appearance of the local contrast, that ideally is neither exaggerated nor understated. Commonly, the pixels tend to have a higher exposed appearance when their normalized values are close to the average value of 0.5. This weight map is expressed as a Gaussian-modeled distance to the average normalized range value (0.5):

$$W_E(x,y) = exp\left(-\frac{(I^k(x,y) - 0.5)^2}{2\sigma^2}\right) \tag{5}$$

where $I^k(x,y)$ represents the value of the pixel location (x,y) of the input image I^k, while the standard deviation is set to $\sigma = 0.25$. This map will assign higher values to those tones with a distance close to zero, while pixels that are characterized by larger distances, are associated with the over- and under- exposed regions. In consequence, this weight tempers the result of the saliency map and produces a well preserved appearance of the fused image.

To yield consistent results, we employ the normalized weight values $\bar{W}$ (for an input k the normalized weight is computed as $\bar{W}^k = W^k / \sum_{k=1}^{K} W^k$), by constraining that the sum at each pixel location of the weight maps W equals one (the normalized weights of corresponding weights are shown at the bottom of figure 2).

2.3. Multiscale Fusion Process

The enhanced image version $\mathcal{R}(x,y)$ is obtained by fusing the defined inputs with the weight measures at every pixel location (x,y):

$$\mathcal{R}(x,y) = \sum_{k=1}^{K} \bar{W}^k(x,y) I^k(x,y) \tag{6}$$

where I^k symbolizes the input (k is the index of the inputs - $K = 2$ in our case) that is weighted by the normalized weight maps $\bar{W}^k$. The normalized weights $\bar{W}$ are obtained by normalizing over all k weight maps W in order that the value of each pixel (x,y) to be constrained by unity value ($\sum \bar{W}^k = 1$).

As can be seen in figure 4 the naive approach to directly fuse (to apply directly equation 6) the inputs and the weights introduces undesirable halos. A common solution to overcome this limitation is to employ multi-scale linear [7, 24] or non-linear filters [23, 10]. The class of non-linear filters are more complex and has shown to add only insignificant improvement for our task (applying WLS [10] yields minor improvements compared with Laplacian pyramid as depicted in figure 4). Since it is straightforward to implement and computationally efficient, in our experiments the classical multi-scale Laplacian pyramid decomposition [7] has been embraced. In this linear decomposition, every input

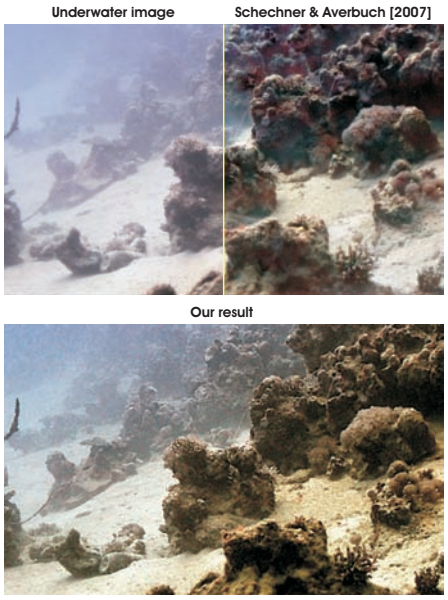

Figure 5. ***Comparison with polarization methods of [Schechner and Averbuch 2007]***. We applied our technique on the white balanced version of one of the employed inputs provided by the authors.

image is represented as a sum of patterns computed at different scales based on the Laplacian operator. The inputs are convolved by a Gaussian kernel, yielding a low pass filtered versions of the original. In order to control the cut-off frequency, the standard deviation is increased monotonically. To obtain the different levels of the pyramid, initially we need to compute the difference between the original image and the low pass filtered image. From there on, the process is iterated by computing the difference between two adjacent levels of the Gaussian pyramid. The resulting representation, the Laplacian pyramid, is a set of quasi-bandpass versions of the image.

In our case, each input is decomposed into a pyramid by applying the Laplacian operator to different scales. Similarly, for each normalized weight map $\bar{W}$ a Gaussian pyramid is computed. Considering that both the Gaussian and Laplacian pyramids have the same number of levels, the mixing between the Laplacian inputs and Gaussian normalized weights is performed at each level independently yielding the fused pyramid:

$$\mathcal{R}^l(x,y) = \sum_{k=1}^{K} G^l\left\{\bar{W}^k(x,y)\right\} L^l\left\{I^k(x,y)\right\} \qquad (7)$$

where l represents the number of the pyramid levels (typically the number of levels is 5), $L\{I\}$ is the Laplacian version of the input I, and $G\left\{\bar{W}\right\}$ represents the Gaussian version of the normalized weight map $\bar{W}$. This step is performed successively for each pyramid layer, in a bottom-up manner. The restored output is obtained by summing the fused contribution of all inputs.

The Laplacian multi-scale strategy performs relatively fast representing a good trade-off between speed and accuracy. 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 that is primarily sensitive to local contrast changes such as edges and corners.

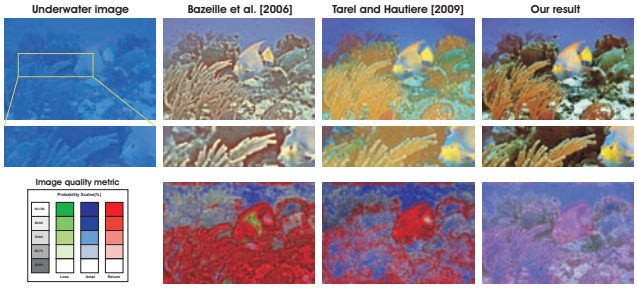

Figure 6. ***Comparative results.*** Compared with the outputs of other enhancing methods our result is less prone to halos and color distortions. This also results (bottom row) when applied IQM metric [4]. As may be observed, based on this metric our approach mainly amplifies the contrast (blue).

3. Results and Discussion

The proposed strategy was tested for real underwater videos and images taken from different available amateur photographer collections. As a result, images and videos have been captured using various cameras and setups. However, an important observation is that we process only 8-bit data format even though many professional cameras have the option to shoot in the RAW mode that usually stores the unprocessed data of the camera's sensor in 12-bit format.

Our technique is computationally effective taking approximately 2 seconds (Matlab code) for a 800×600 frame but we believe that an optimized implementation could run real-time on common hardware. The reader is referred to the supplementary material for additional results (images and videos). By a general visual inspection it can be observed that our technique is able to yield accurate results with enhanced global contrast, color and fine details while the temporal coherence of the videos is well preserved.

In figure 5 we compare our results with the polarization technique of Schechner and Averbuch [25] that uses two frames taken with wide-field polarized illumination. By employing our technique on the provided white balanced version of one of their inputs we are able to produce a more pleasing image version.

In Figure 6 we compare our technique with several specialized underwater enhancing techniques. We considered the specialized single underwater image enhancing techniques [5] but also the recent specialized dehazing technique [27]. By a closer inspection (please observe as well

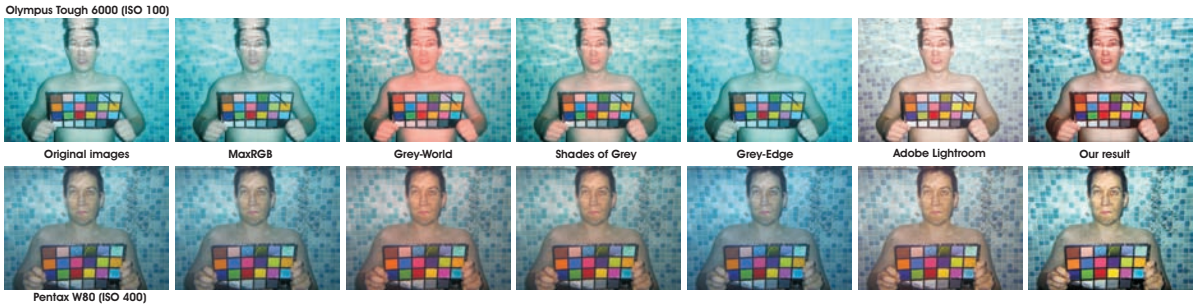

Figure 7. ***Robustness to different cameras.*** We have employed our algorithm to a set of underwater images that contain the standard Macbeth Color Checker taken by different specialized cameras (Olympus Tough 6000 and Pentax W80 are shown in this figure, for the complete set please refer to the supplementary material). In addition, several white balance algorithms are employed to the images.

the middle row of Figure 6) our result presents less halos and color distortions. For this example to visualize how contrast is modified we employed the IQM metric [4] that was originally developed to evaluate tone mapping operators. This metric utilizes a model of the human visual system being sensitive to three types of structural changes: loss of visible contrast (green), amplification of invisible contrast (blue) and reversal of visible contrast (red). As a general remark, compared with the other considered approaches, the most predominant structural change characteristic to our method is the amplification of the contrast (blue) and only very few locations exhibit reverse (red) and loss (green) of the contrast.

Since in general the color is captured differently by various cameras we demonstrate that our algorithm is independent of certain camera settings. We have employed our algorithm to a set of underwater images that contain the standard *Macbeth Color Checker* taken by seven different professional cameras (see figure 7 while for complete set please refer to the supplementary material). At first glance, these professional cameras introduce various color casts. Our approach shows high robustness to preserve the uniformity in the color appearance for different cameras. The reader is referred to the supplementary material for the study that interprets statistically the disparity between the reference color patches and the registered results of different methods.

Our technique shown limitations when dealing with images of very deep scenes taken with poor strobe and artificial light.In such cases, even some enhancement could be obtained, the bluish appearance however still remains. Moreover, when the illumination is poor the very distant parts of the scene cannot be recovered reliably. The restoration of distant objects and regions represents also a general limitation of our approach compared with hardware and polarization-based techniques that in general perform better in such cases due to the additional available information.

Figure 8. ***Local feature points matching.*** Compared with the initial images (top) by applying standard SIFT on our enhanced versions (bottom) the matching result is improved considerable.

3.1. Applications

We found our technique suitable for several other applications that are briefly described in the following section. More results are included as well in the supplementary material.

Matching images by local feature points is a fundamental task of many computer vision applications. We employ the SIFT [18] operator for an initial pair of underwater images and as well for the restored versions of the images (see figure 8). We use the original implementation of SIFT applied exactly in the same way in both cases. For the initial case SIFT filters 3 good matches and one mismatch while the matching of the enhanced image versions yields 43 valid matches and no mismatches. These promising achievements demonstrate that our technique does not introduce artifacts but mainly restores both global contrast and local features of underwater images.

Segmentation aims to divide images into disjoint and homogeneous regions with respect to some characteristics (e.g. texture, color). In this work we employ the $GAC++$

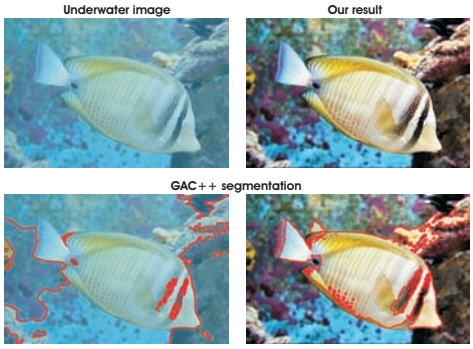

Figure 9. ***Image segmentation.*** Processing underwater images with our method, the segmentation result is more consistent while the filtered boundaries are perceptually more accurate.

[22] that represents a state-of-the-art geodesic active contours method (variational PDE). Figure 9 proves that by processing underwater images with our approach the segmentation result is more consistent while the filtered boundaries are perceptually more accurate. This task demonstrates that our technique does not introduce halos close to object boundaries.

Image dehazing [11, 3] is the process of removing the haze and fog effects from the spoilt images. Because of similarities between hazy and underwater environments due to the light scattering process, we found our strategy appropriate for this challenging task. However, as explained previously, since the underwater light propagation is more complex we believe that image dehazing could be seen as a subclass of the underwater image restoration problem. Comparative results with state-of-the-art single image dehazing techniques [11, 17] are shown in Figure 10.

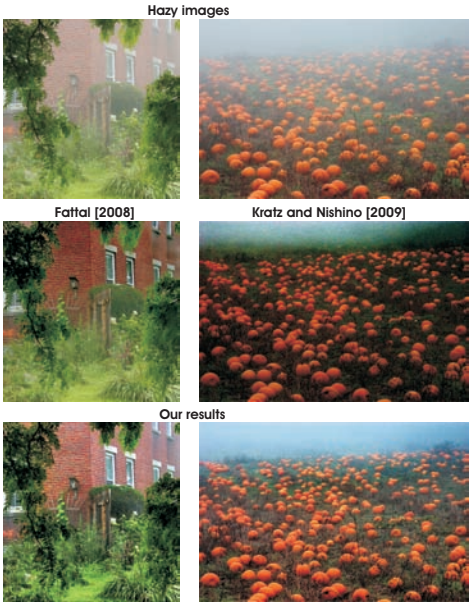

Figure 10. ***Single image dehazing.*** Our approach is able to dehaze images comparable to and even more accurate than the specialized single dehazing techniques.

References

[1] R. Achantay, S. Hemamiz, F. Estraday, and S. Susstrunky. Frequency-tuned salient region detection. *IEEE CVPR*, 2009.

[2] C. O. Ancuti, C. Ancuti, and P. Bekaert. Fusion-based restoration of the underwater images. In *In IEEE ICIP*, 2011.

[3] C. O. Ancuti, C. Ancuti, C. Hermans, and P. Bekaert. A fast semi-inverse approach to detect and remove the haze from a single image. *ACCV*, 2010.

[4] T. O. Aydin, R. Mantiuk, K. Myszkowski, and H.-S. Seidel. Dynamic range independent image quality assessment. In *SIGGRAPH*, 2008.

[5] S. Bazeille, L. J. I. Quidu, and J. P. Malkasse. Automatic underwater image pre-processing. *In Proc. of CMM*, 2006.

[6] E. P. Bennett, J. L. Mason, and L. McMillan. Multispectral bilateral video fusion. *IEEE Trans. on Image Processing*, 2007.

[7] P. Burt and T. Adelson. The laplacian pyramid as a compact image code. *IEEE Transactions on Communication*, 1983.

[8] D.J.Jobson, Z.Rahman, and G.A.Woodell. A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Transactions on Image Processing*, 1997.

[9] M. Ebner. *Color Constancy, Wiley 1st edition*. 2007.

[10] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski. Edge-preserving decomposition for multi-scale tone and detail manipulation. *SIGGRAPH 08*.

[11] R. Fattal. Single image dehazing. *SIGGRAPH*, 2008.

[12] G. D. Finlayson and E. Trezzi. Shades of gray and colour constancy. *IS&T/SID Twelfth Color Imaging Conference: Color Science,Systems and Applications, Society for Imaging Science and Technology*, pages 37–41, 2004.

[13] A. Gijsenij and T. Gevers. Color constancy using natural image statistics and scene semantics. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(4):687–698, 2011.

[14] M. Grundland, R. Vohra, G. P. Williams, and N. A. Dodgson. Cross dissolve without cross fade. *Computer Graphics Forum*, 2006.

[15] D.-M. He and G. G. L. Seet. Divergent-beam lidar imaging in turbid water. *Optics and Lasers in Engineering*, 2004.

[16] D. M. Kocak, F. R. Dalgleish, F. M. Caimi, and Y. Y. Schechner. A focus on recent developments and trends in underwater imaging. *Marine Technology Society Journal*, 2008.

[17] L. Kratz and K. Nishino. Factorizing scene albedo and depth from a single foggy image. *ICCV*, 2009.

[18] D. Lowe. Distinctive image features from scale-invariant keypoints. *Int. Journal of Comp. Vision*, 2004.

[19] B. L. McGlamery. A computer model for underwater camera systems. *Ocean Optics*, 1979.

[20] T. Mertens, J. Kautz, and F. V. Reeth. Exposure fusion. *Comp. Graph. Forum*, 2009.

[21] S. Narasimhan and S. Nayar. Contrast restoration of weather degraded images. *IEEE Trans. on Pattern Analysis and Machine Intell.*, 2003.

[22] G. Papandreou and P. Maragos. Multigrid geometric active contour models. *In IEEE Trans. on Image Processing*, 2007.

[23] S. Paris and F. Durand. A fast approximation of the bilateral filter using a signal processing approach. *Int. Journal of Computer Vision*, 2009.

[24] T. Pu and G. Ni. Contrast-based image fusion using the discrete wavelet transform. *Opt. Eng.*, 2000.

[25] Y. Schechner and Y. Averbuch. Regularized image recovery in scattering media. *IEEE Trans Pattern Anal Mach Intell.*, 2007.

[26] R. Schettini and S. Corchs. Underwater image processing: State of the art of restoration and image enhancement methods. *J. on Adv. in Sig. Proc.*, 2010.

[27] J.-P. Tarel and N. Hautiere. Fast visibility restoration from a single color or gray level image. *In IEEE International Conference on Computer Vision*, 2009.

[28] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. *In IEEE International Conference on Computer Vision*, 1998.

[29] J. van de Weijer, T. Gevers, and A. Gijsenij. Edge based color constancy. *IEEE Transactions on Image Processing*, 2007.

[30] J. Zimmerman. An evaluation of the effectiveness of adaptive histogram equalization for contrast enhancement. *IEEE Trans. on Med. imaging*, 1988.