Underwater Video Dehazing using Hybrid Pyramid Decomposition with Multi Frame Fusion

B. Prabhakar
Department of ECE, JNTUH College of Engineering, Jagtial, Telangana, India.

Abstract - Underwater scenes captured by cameras are plagued with poor contrast and a spectral distortion, which are the result of the scattering and absorptive properties of water. In this paper we present a novel dehazing method that improves visibility in images and videos by detecting and segmenting video frame regions that contain only water. The colour of these regions, which we refer to as pure haze regions, is similar to the haze that is removed during the dehazing process. Moreover, we propose a semantic white balancing approach for illuminant estimation that uses the dominant colour of the water to address the spectral distortion present in underwater scenes. To validate the results of our method and compare them to those obtained with state-of-the-art approaches, we perform extensive subjective evaluation tests using images captured in a variety of water types and underwater videos captured onboard an underwater vehicle.

Keywords: Dehazing, video frame processing, Segmentation, Underwater, White balancing, Video processing

I. INTRODUCTION

Water is the real elixir of life. Ocean covers approximately 70 percentage of the earth's surface. Despite the importance of oceanic environment, humans are still unable to inquire the full depth of the ocean and discover its resources and wealth due to the dangerous, cold and dark, unfamiliar environment. The safety and surveillance of oceanic environment thus become very relevant for research. Underwater oil and gas resources account for around six percentage of global oil production. Once oil and gas are discovered in an underwater field, massive production platforms and specially designed systems and pipelines are required to extract and transport the oil and gas to shore [1] which makes it very mandatory to keep track of their physical condition. Recently, remotely operated vehicles (ROV) and autonomous underwater vehicles (AUV) are on the threshold of playing a key role in inspecting pipelines because they eliminate the need for humans to be at great depth or in dangerous conditions. It can collect and save required information without human intervention making it possible to introduce complete level of autonomy in inspecting pipelines. Vision based system is one of the cheapest methods for underwater pipeline inspection [2]. The photographer lowered a camera in housing in Weymouth Bay and the shutter was operated from an anchored boat. The exposure time used was 10 minutes. This experiment resulted in flooding of the camera, however the film was salvaged. Today, such underwater photography is performed using advanced cameras with scuba diving.

There are many difficulties for undersea optical imaging. The submerging of a camera underwater requires adequate housing. The maneuvering of camera with the help from remote place or in person at the site is likewise a complex task. However, the major challenge is imposed by underwater medium properties. The two foremost underwater phenomenon affecting the outcome and visual aspect is light attenuation and scattering [3-5]. As the distance between camera and object increases, the scattered light renders lower screen contrast in underwater images. As evident, scattered light component does not carry any scene information and thus underwater optical imaging becomes tedious. Research has been carried out to gauge the wideband attenuation coefficients per color channel in underwater images. However, these findings are relatively limited, as the parameters become sensitive to the original color and the distance between object and camera [6]. The parameters of scattering play a vital role in recovering the dehazed video frame. but, it has implications as these values tend to vary for the same type of water body at different places on account of turbidity, temperature, salinity and turbulence to name a few, which further demands precise calibration. For clean shallow water bodies, ambient light is sufficient to capture quality images. But, for deep sea underwater imaging, an artificial source of lighting is must to capture images. This source of light results in two problems. The first is a color cast of illumination source formed on the captured video frame, which requires a suitable white balancing approach to address the problem. Second, this artificial source of lighting tends to create non-uniform illumination, with a bright spot at the center which radially decreases from the center of the video frame. Underwater optical imaging suffers from light attenuation, which results on account of light absorption by water which increases exponentially with the depth and affects all the wavelengths to varying degrees [7]. The effect of wavelength dependency for gradual color attenuation is as shown in Figure 1.

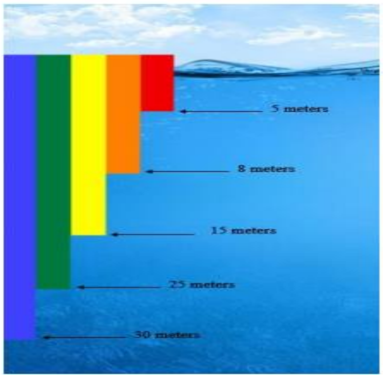

Figure 1: Illustration of the color attenuation for different wavelength at varying depths in underwater.

The longer wavelength corresponding to reddish tone is attenuated first, while the shortest one corresponding to blue color travels the longest [8]. This loss in color results in a greenish or bluish appearance as the distance between the camera and the object in the underwater scene increases. Organic particles such as phytoplankton found in coastal waters absorb light predominantly in the shortest wavelength such as blue color, allowing only the greenish tint to persist. Apart from this, according to Rayleigh scattering theory, the intensity of scattering is inversely proportional to fourth power of wavelength, and as such the colors with shorter wavelength scatter more prominently than the longer wavelengths. So from the above discussion, we can conclude that water absorbs colors of longer wavelength and scatters the shorter wavelength for visible light spectrum.

II. LITERATURE SURVEY

The autonomous underwater vehicle incorporates vision system for various applications. A review of several applications is carried out. All these applications require pre-processing of an acquired video or video frame. Various techniques for enhancing underwater images are also reviewed. Authors [9] suggested that to have a better advancement in underwater imaging the following areas have to be advanced. Affordable, high quality cameras needed, compact, efficient and easy to program digital signal processors should be available and better modelling and simulation software is required. The authors have set up camera system to monitor coral reef communities in marine parks. Not all cameras deployed in the sea are stationary or mounted to underwater vehicle. Since mid-1980's, animals have become imaging platforms. In [10] authors described the need of pre-processing for underwater images because of the poor quality of the captured video frame. Various filters such as homomorphic, wavelet denoising etc. are compared. For Gaussian noise, wavelet method was better, and for salt and pepper noise homomorphic filter has shown a better result. Speckle noise can easily be removed by using wavelet denoising.

In [11] authors proposed YUV color space based turbid underwater video frame enhancement using filtering approach in the frequency domain. The YUV color model is defined in terms of one luma (Y) and two chrominance (UV) components. The technique is comprised of numerous independent algorithms. In the first step repetitive wave patterns are removed using filtering based on spectral analysis. The problem of non-uniform lighting is corrected using homomorphic filtering followed with wavelet noise reduction. The resultant video frame requires smoothing yet preserving the edge information. This is performed by applying anisotropic filtering. Further, the video frame intensity is enhanced by contrast stretching. The video frame is subsequently transformed back to RGB color space followed with color normalization. However, the resulting video frame exhibits distorted output in terms of color fidelity.

In [12] authors presented an ICM model wherein the underwater video frame is dynamically stretched for entire range in RGB color space followed with contrast stretching of the resultant video frame for I and S component in hue intensity saturation (HIS) color space. This method is simple and effective and is ideally suited for underwater images with minimal haze component. Color cast issue was addressed by modifying red and green channels in RGB color space using von Kries hypoproject followed by contrast correction in Unsupervised Color Model (UCM) algorithm [13]. The major shortcoming of this technique is a generation of high video frame noise affecting the pixels of the resultant video frame and uneven enhancement producing areas with dark regions. Such areas exhibit less information content of the underlying images. Another drawback encountered in this technique, is that the processed images retain blue-green illumination. This problem was pointed and addressed by authors [14-15] by applying the stretching limits only to blue and red color channels of the hazy underwater video frame.

III. PROPOSED METHOD

In contrast, this paper introduces a novel approach to remove the haze in underwater images based on a single video frame captured with a conventional camera. As illustrated in Fig. 2, this work 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 video frame. 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.

Since the color correction is critical in underwater, we first apply this work white balancing technique to the original video frame. This step aims at enhancing the video frame 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 This work first input we perform a gamma correction of the white balanced video frame 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 video frame. Therefore, we follow the unsharp masking principle in the sense that we blend a blurred or unsharp (here Gaussian filtered) version of the video frame with the video frame to sharpen. The typical formula for unsharp masking defines the sharpened video frame S as $S = I + \beta(I - G * I)$, where I is the video frame to sharpen (in This work case the white balanced video frame), $G * I$ denotes the Gaussian filtered version of I, and β is a parameter. In practice, the selection of β is not trivial. A small β fails to sharpen I, but a too large β results in over-saturated regions, with brighter highlights and darker shadows.

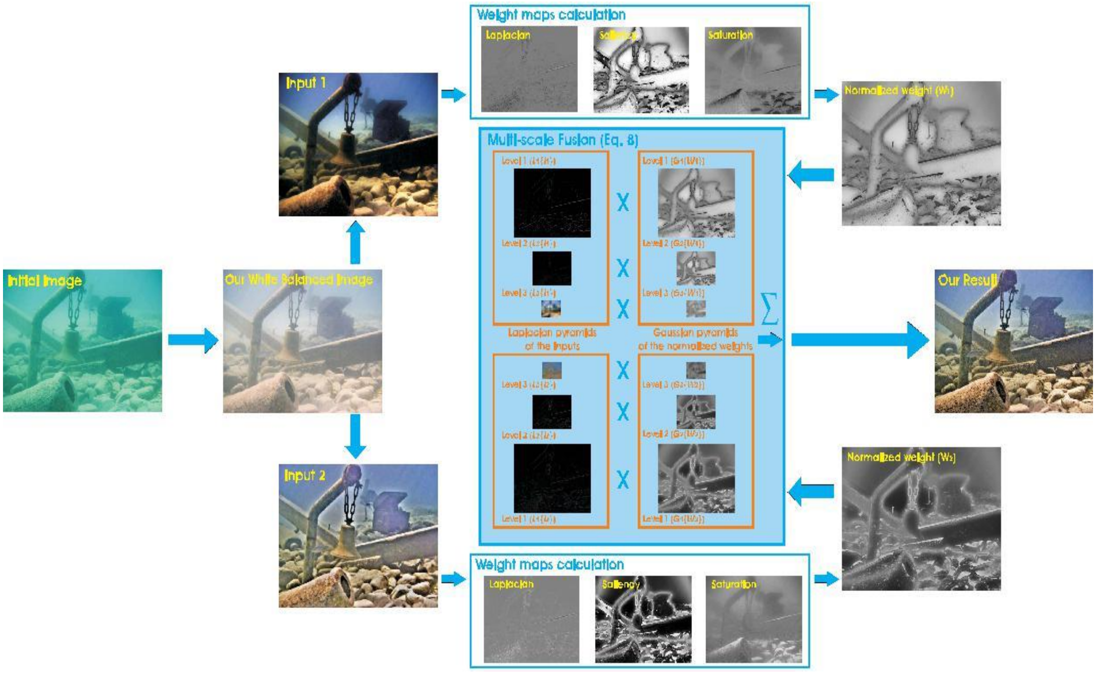

Figure 2: Flow chart of proposed multi-scale video frame enhancement.

White Balancing

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

Weight Maps

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

Laplacian contrast weight (WL) 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 and extending depth of field 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 (WS) 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. 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 (WSat) enables the fusion algorithm to adapt to chromatic information by advantaging highly saturated regions. This weight map is simply computed (for each input Ik) as the deviation (for every pixel location) between the Rk ,Gk and Bk color channels and the luminance Lk of the kth input:

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

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 Wk is first obtained by summing up the three WL , WS, and WSat 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

Multi scale Fusion

The multi-scale decomposition is based on Laplacian pyramid originally described. The pyramid representation decomposes an video frame into a sum of bandpass images. In practice, each level of the pyramid does filter the input video frame using a low-pass Gaussian kernel G, and decimates the filtered video frame by a factor of 2 in both directions. It then subtracts from the input an up-sampled version of the low-pass video frame, thereby approximating the (inverse of the) Laplacian, and uses the decimated low-pass video frame as the input for the subsequent level of the pyramid. Formally, using Gl to denote a sequence of l low-pass filtering and decimation, followed by l up-sampling operation. In practice, the number of levels N depends on the video frame size, and has a direct impact on the visual quality of the blended video frame. The dehazed output is obtained by summing the fused contribution of all levels, after appropriate upsampling.

IV. SIMULATION RESULTS

Figure 3: Input video frame

As shown in the Figure above is the given input video frame from this single input video frame two images are derived from the white balancing. White balancing aims at compensating for color cast caused by the selective absorption of colors with depth.

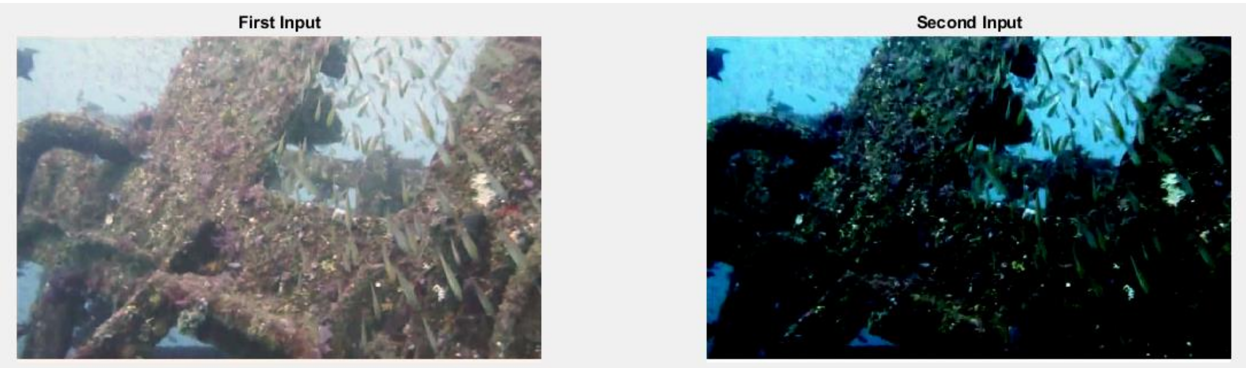

Figure 4: white balance output

Underwater images are highly correlated with the depth and typically exhibits color distortion and low contrast. Depending on wavelengths and attenuation rates red color is the one that attenuates the fastest. So, from the input video frame the red channel video frame is recovered.

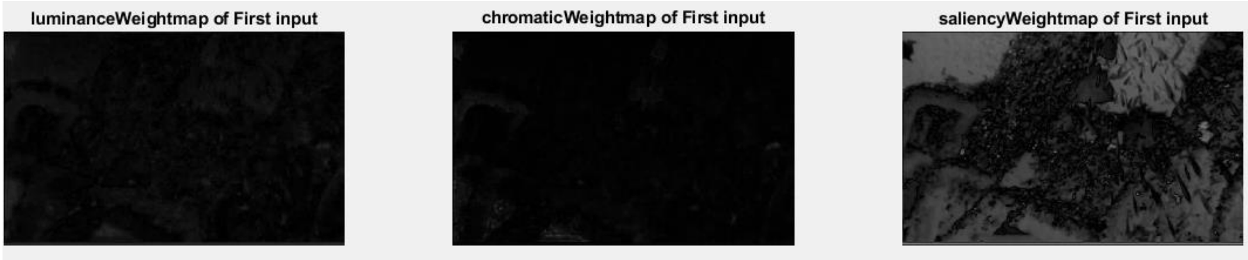

Figure 5: weight maps output

In fact that turbid waters blue channel may be significantly attenuated due to the absorption by organic matter then the red channel may appears to be insufficient if the blue is strongly attenuated.

Figure 6. Fused enhanced video frame

As shown in the Figure 6 fused video frame builds on the set of inputs and weight maps derived from single video frame. This fused video frame is defined based on number of local video frame quality. In this Figure we can see the difference between the input video frame and fused video frame as the fused video frame is clear with colour and quality of the video frame. Finally we calculate and get the values of MSE and PSNR to know the quality of the video frame.

Table 1: performance evaluation

Method	PSNR	SSIM	MSE
HE [11]	32.34	0.81	0.0938
CLAHE [13]	36.84	0.83	0.0927
Proposed	**49.234**	**0.912**	**0.0282**

From the table 1 it is observed, the proposed method gives the maximum quality evaluation compared to the state of art approaches.

V. CONCLUSION

Fusion based underwater video frame restoration approach has been considered for underwater images. Detailed comparisons revealed that the fusion based technique used for under water images has greatly enhanced the visibility as compared to the dehazing technique which involves the amalgamation of two input images. Thorough evaluation also reveals that dehazing technique does not also solve the problem of restoration of underwater images; the problem of restoration of the color along with the enhancement of contrast has been sorted out better by using the fusion based technique then dehazing technique. Usage of different polarization filters also yielded better results as compared to White balancing method in above mentioned problem of restoration of color and the enhancement of contrast. By the use of diverse polarization filter systems better viewing of the under video frame water has been generated, as compared to the conventional techniques thus far being used while keeping in check that no special environment factors and hardware issues taken into consideration.

VI. REFERENCES

[1] G. L. Foresti, “Visual inspection of sea bottom structures by an autonomous underwater vehicle,” IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics), vol. 31, no. 5, pp. 691– 705, 2001.

[2] G. Kim, D. Kim, D. Kim, H. T. Choi, D. Lee, and H. Myung, “Path Planning and T racking of an Autonomous Underwater Vehicle using Virtual Way-points,” The Abstracts of the international conference on advanced mechatronics : toward evolutionary fusion of IT and mechatronics : ICAM, vol. 2010.5, no. 0, pp. 118–123, 2010.

[3] M. Lanzagorta and Marco, “Quantum imaging for underwater arctic navigation,” in Proceedings of the SPIE, Volume 10188, id. 101880G 27 pp. (2017)., vol. 188, p. 101880G, 2017..

[4] H. Zhu, F. H. Y. Chan, and F. K. Lam, “video frame Contrast Enhancement by Constrained Local Histogram Equalization,” Computer Vision and video frame Understanding, vol. 73, no. 2, pp. 281– 290, Feb. 1999.

[5] S. Aslani and H. Mahdavi-Nasab, “Optical flow based moving object detection and tracking for traffic surveillance,” World Academy of Science, Engineering, 2013.

[6] Q. Fu and M. Celenk, “Marked watershed and video frame morphology based motion detection and performance analysis,” in 2013 8th International Symposium on video frame and Signal Processing and Analysis (ISPA), pp. 159–164, 2013.

[7] N. Mohamed, I. Jawhar, J. Al-Jaroodi, and L. Zhang, “Sensor Network Architectures for Monitoring Underwater Pipelines,” Sensors, vol. 11, no. 12, pp. 10738–10764, Nov. 2011.

[8] J. Villasenor and B. Hutchings, “The flexibility of configurable computing,” IEEE Signal Processing Magazine, vol. 15, no. 5, pp. 67–84, 1998.

[9] Takashi Saegusa, Tsutomu Maruyama, and Yoshiki Yamaguchi, “How fast is an FPGA in video frame processing?,” in 2008 International Conference on Field Programmable Logic and Applications, pp. 77– 82, 2008.

[10] S. Hasan, “Performance-Aware Architectures for Parallel 4D Color fMRI Filtering Algorithm: A Complete Performance Indices Package,” IEEE Transactions on Parallel and Distributed Systems, vol. 27, no. 7, pp. 2116–2129, Jul. 2016.

[11] S. Boussakta, “A novel method for parallel video frame processing applications,” Journal of Systems Architecture, vol. 45, no. 10, pp. 825–839, Apr. 1999.

[12] 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, vol. 42, no. 1, pp. 52–67, 2008.

[13] M. Kumar and S. Kumar Thakur, “Comparison of Filters used for Underwater video frame Pre-Processing,” IJCSNS International Journal of Computer Science and Network Security, vol. 10, no. 1, 2010.

[14] C. J. Prabhakar and P. U. P. Kumar, “An video frame Based Technique for Enhancement of Underwater Images,” arXiv preprint arXiv:1212.0291, Dec. 2012.

[15] K. Plakas and E. Trucco, “Uncalibrated computer vision technique in the underwater environment,” in IEE Colloquium on Underwater Applications of video frame Processing, vol. 18, pp. 11–15, 1998.