An Online Platform for Underwater Image Quality Evaluation

Chau Yi Li, Riccardo Mazzon, Andrea Cavallaro

Centre for Intelligent Sensing, Queen Mary University of London
{chau.li, r.mazzon, a.cavallaro}@qmul.ac.uk

Abstract. With the miniaturisation of underwater cameras, the volume of available underwater images has been considerably increasing. However, underwater images are degraded by the absorption and scattering of light in water. Image processing methods exist that aim to compensate for these degradations, but there are no standard quality evaluation measures or testing datasets for a systematic empirical comparison. For this reason, we propose PUIQE, an online platform for underwater image quality evaluation, which is inspired by other computer vision areas whose progress has been accelerated by evaluation platforms. PUIQE supports the comparison of methods through standard datasets and objective evaluation measures: quality scores for images uploaded on the platform are automatically computed and published in a leaderboard, which enables the ranking of methods. We hope that PUIQE will stimulate and facilitate the development of underwater image processing algorithms to improve underwater images.

Keywords: Underwater image processing, evaluation platform, benchmark datasets, underwater image enhancement.

1 Introduction

Underwater image analysis is attracting an increasing level of attention [1] and supports applications such as underwater exploration, habitat monitoring and species identification [2]. However, the appearance of underwater scenes is degraded by scattering, which blurs the resulting image, and by wavelenght-dependent absorption, which reduces the energy of the light reaching the camera.

Image processing may be used to improve underwater image quality through restoration or enhancement methods. Restoration methods compensate for image distortions using prior information, such as the Dark Channel [3,4] or the Red Channel [5] prior. These methods may assume a uniform [3,4,5,6] or a more realistic non-uniform [7] background-light. Enhancement methods remove the colour cast [8,9,10] using for example global white balancing [8,9].

Figure 1 shows an underwater image processed by three different methods: an important issue is how to objectively assess the processed images as there are currently no benchmark datasets or standard image quality evaluation measures. In fact, while the evaluation of several computer vision tasks is supported by


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

Fig. 1: Image processing results on a sample underwater image. (a) Original image; (b) enhancement by global white balancing [9]; (c) enhancement by [10]; (d) restoration by [7]. Note the colour distortion in (b) and the overexposed areas (e.g. fishes) in (c).

online benchmarking platforms, such as the Middlebury platform for optical flow [11] and the Multiple Object Tracking (MOT) Challenge for multi-target tracking [12], no platform exists yet for underwater image processing evaluation.

In this paper, we present an online Platform for Underwater Image Quality Evaluation (PUIQE) that supports the development of underwater image processing algorithms by distributing commonly used test images and by calculating underwater image quality scores. The main contribution is having established a common protocol to compare processing results: researchers run their algorithms on the datasets downloaded from the platform and then submit the results, whose evaluation scores are then published on a leaderboard. PUIQE ensures a fair comparison by restricting submissions to contain the complete dataset with the correct image resolution and format. Moreover, unlike the Middlebury and the MOT platforms that enforce the automatic evaluation on complete datasets, PUIQE also allows users to use the platform for private development by testing single images in restricted sessions (see Fig. 2). PUIQE is available at `http://puiqe.eecs.qmul.ac.uk/`.

2 Evaluation measures

Most underwater image processing results are assessed subjectively by visual inspection [5,10] or by using no-reference image quality measures [6,8]. Artificial references can be created by taking stereo images with dual cameras to obtain the true scene-to-camera distance [9] or by using a colour chart as reference [9]. However, this additional information is not available outside controlled settings.

The current version of PUIQE includes two quality measures for underwater images, namely Underwater Color Image Quality Evaluation (*UCIQE*) [13] and Human-Visual-System-Inspired Underwater Image Quality Measures (*UIQM*) [14]. These measures quantify the colour degradation due to absorption of light in water and the blurring effect due to scattering. *UCIQE* and *UIQM* combine linearly measures of colour and sharpness, with coefficients obtained from subjective evaluation data. *UCIQE* evaluates the quality of an image only based on

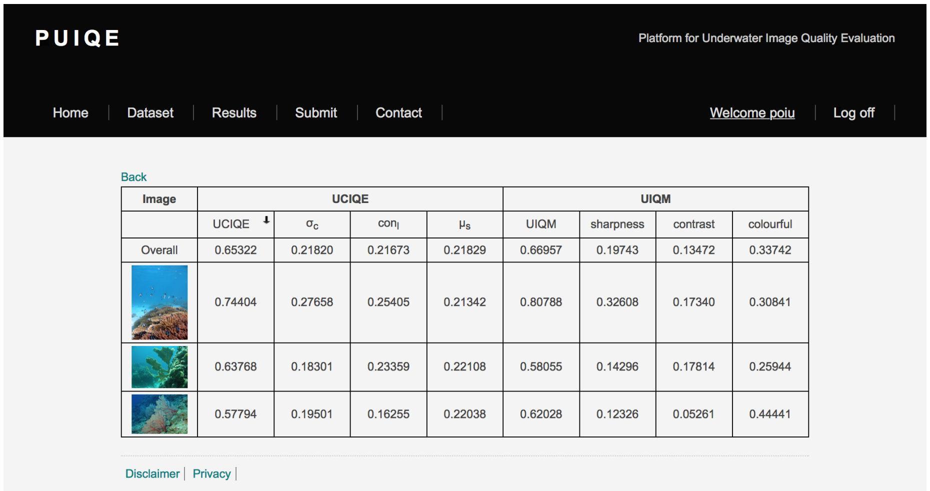


Image	UCIQE				UIQM			
	UCIQE	σ_c	con_l	μ_s	UIQM	sharpness	contrast	colourful
Overall	0.65322	0.21820	0.21673	0.21829	0.66957	0.19743	0.13472	0.33742
	0.74404	0.27658	0.25405	0.21342	0.80788	0.32608	0.17340	0.30841
	0.63768	0.18301	0.23359	0.22108	0.58055	0.14296	0.17814	0.25944
	0.57794	0.19501	0.16255	0.22038	0.62028	0.12326	0.05261	0.44441

Fig. 2: Example of private session in PUIQE. Individual images are evaluated with *UCIQE* and *UIQM* (Sec. 2).

the colour distortion caused by light attenuation, whereas *UIQM* is modelled on the human vision system and also considers the loss of contrast.

Let $I_p = [L_p, a_p, b_p]$ be the value of pixel p in the CIELab space, and L_p, a_p and b_p be the intensity values in the L, a and b channels, respectively. *UCIQE* is a linear combination with weights obtained from a subjective evaluation of 12 subjects on 44 images [13]:

$$UCIQE = 0.4680 \times \sigma_c + 0.2745 \times con_l + 0.2576 \times \mu_s, \tag{1}$$

where con_l is the contrast of luminance, i.e. the difference between the top 1% and the bottom 1% of the values in $\{L_p | p = 1...N\}$, where N is the number of pixels in the image; σ_c is the standard deviation of chroma:

$$\sigma_c = \frac{1}{N} \sum_{p=1}^{N} \left(C_p^2 - \mu_c^2 \right), \tag{2}$$

with chroma, C_p, defined as [15]:

$$C_p = \sqrt{a_p^2 + b_p^2}; \tag{3}$$

and μ_s is the average of saturation:

$$\mu_s = \frac{1}{N} \sum_{p=1}^{N} S_p, \tag{4}$$

with saturation, S_p, defined as [15]:

$$S_p = \frac{C_p}{L_p}. \tag{5}$$

(a) (b) (c)

Fig. 3: Sample images for illustrating the behaviour of the evaluation measures currently implemented in PUIQE: (a) original; (b) blurred with Gaussian filter ($\sigma = 2.0$); (c) gamma corrected ($\gamma = 2.2$).

Table 1: Effect of blurring and reduced colour intensity on the *UCIQE* (Eq. 1) and *UIQM* (Eq. 6) measures, and their components. Numerical values and trends. Key (comparison with respect to the original image): ↓: value decreased; ↑: value increased; -: measure unchanged (up to fourth decimal digit).

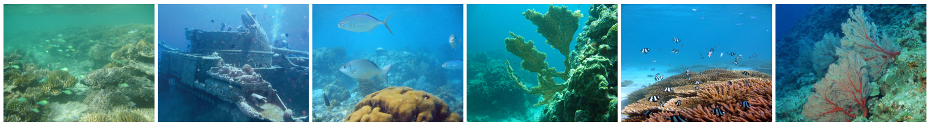
Image	*UCIQE*	σ_c	con_l	μ_s	*UIQM*	*UISM*	*UIConM*	*UICM*
Fig. 3(a)	.6392	.2049	.2218	.2129	.6855	.2073	.1069	.3712
Fig. 3(b)	.6327 ↓	.2048 ↓	.2153 ↓	.2126 ↓	.6579 ↓	.2069 ↓	.0399 ↓	.4111 ↑
Fig. 3(c)	.6326 ↓	.2072 ↑	.1830 ↓	.2424 ↑	.6617 ↓	.1831 ↓	.1452 ↑	.3334 ↓

UIQM combines linearly colourfulness, *UICM*, sharpness, *UISM*, and contrast, *UIConM*:

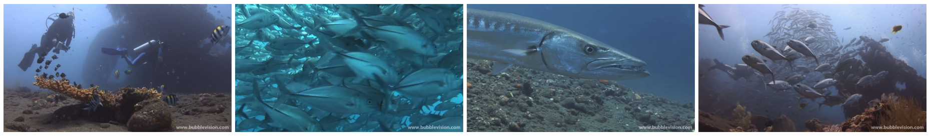
$$UIQM = 0.0282 \times UICM + 0.2953 \times UISM + 3.3753 \times UIConM, \qquad (6)$$

with weights obtained from a subjective evaluation of 10 subjects on 14 images. *UICM*, which quantifies the degradation caused by light absorption, is defined by the statistics of the differences between red-green and yellow-blue planes. *UISM* and *UIConM* account for the degradation due to scattering: *UISM* depends on the strength of Sobel edges computed on each colour channel independently; whereas *UIConM* is obtained using the logAMEE operation [16], which is considered consistent with human visual perception in low light conditions.

In Fig. 3, we mimic the effect of scattering and absorption by artificially distorting an underwater image with Gaussian blur and gamma correction, respectively. The corresponding *UCIQE* and *UIQM* values are shown in Table 1. For the Gaussian-blurred image, *UISM* calculated by *UIQM* decreases, as expected; for the gamma-corrected image, both the values of colourfulness σ_c of *UCIQE* and *UICM* of *UIQM* decrease, as expected.



(a)



(b)

Fig. 4: Datasets currently available in PUIQE: (a) ReefnFish; (b) Bali.

3 Datasets

As no established image dataset for underwater image processing is currently available, researchers use different images for testing, and, even when images of the same scene are used, they might have different resolutions. For example, there is no overlap between the test images used by Berman et al. [9] and Galdran et al. [5], whereas Chiang and Chen [10] used images with lower resolution than the original video [17]. Therefore valid comparisons are in general not possible from published results.

To allow an easy access to frequently used images as first step towards the establishment of standard testing datasets, we gathered two sets of images used in various publications, namely *ReefnFish* and *Bali* (Fig. 4). The ReefnFish dataset consists of 6 images that include man-made and natural objects, such as shipwreck, fish and coral, and scenes under low-lighting conditions. These images were used in [5,6,8]. The Bali dataset consists of 4 images used in [6,10] and extracted from a video with scenes with changing background light, scuba divers and varying scene depth. Note that while these images were employed in multiple publications [5,8,10], there is no guarantee that they had the same image resolution and format when used in the experiments. For instance, using the same image but saved with different JPEG compression distortions [18] does not allow for a fair comparison of algorithms.

By sharing datasets online, researchers can easily access them along with a standard procedure for comparison. The consistency of testing will be ensured by constraining the entries to the leaderboard only to those algorithms tested on the full datasets, and on images of the same resolution and format as the originals.

4 Leaderboard

Evaluation tables facilitate the comparison of methods, and the identification of their strengths and limitations. PUIQE lists on a leaderboard the details of the methods being tested and their performance scores, as discussed in Sec. 2. To be included in the leaderboard, methods have to process all the images in the ReefnFish or Bali datasets. Submission results can be either associated with the authors and the related publication or remain anonymous at the authors' wish. The online system checks whether the image format is the same as the original datasets and rejects the submission otherwise.

The processed images of the datasets, and the *UCIQE* and *UIQM* measures along with their contributing terms (see Eq. 1 and Eq. 6), are presented in the form of a table summarising the results. Users can select to view a compact table where the details of the methods are hidden, and sort the methods by measure or by their contributing components.

Evaluation results can be displayed either as the average of all processed images in a dataset or as individual scores for each image. Moreover, original and processed images are shown to enable visual comparison.

As example, we uploaded the results of two published methods [9,19] generated from their implementations and one of our previous works [7]. An anonymous submission is also included. Figure 5(a) reports a snapshot of this leaderboard and Fig. 5(b) the evaluation of an individual image, both from the ReefnFish dataset.

5 Conclusion

We presented PUIQE, the first online platform for underwater image quality evaluation. PUIQE supports the development of underwater image processing algorithms by facilitating their comparisons, with results presented in the form of a leaderboard. PUIQE allows an easy access to images frequently used in publications and provides a simple-to-use evaluation with existing performance measures. We expect that the proposed platform will boost the development of new algorithms and the convergence towards standardised procedures for underwater image quality evaluation.

References

1. Horgan, J. and Daniel T.: Computer Vision Applications in the Navigation of Unmanned Underwater Vehicles. Underwater Vehicles. (2012).
2. Strachan, N.: Recognition of fish species by colour and shape. Image and Vision Computing, vol. 11, pp. 2-10 (1993).
3. Emberton, S., Chittka, L., Cavallaro, A.: Hierarchical rank-based veiling light estimation for underwater dehazing. British Machine Vision Conference (2015).
4. Emberton, S., Chittka, L., Cavallaro, A.: Underwater image and video dehazing with pure haze region segmentation. Computer Vision and Image Understanding, Vol. 168, pp. 145-156 (2018)
5. Galdran, A., Pardo, D., Picn, A., Alvarez-Gila, A.: Automatic red-channel underwater image restoration. Journal of Visual Communication and Image Representation. vol. 26, pp. 132-145 (2015).

Dataset: ReefnFish (click to see results for individual images)

Team	UCIQE				UIQM			
	UCIQE	σ_c	con_l	μ_s	UIQM	sharpness	contrast	colourful
UWHL	0.673	0.217	0.238	0.218	0.691	0.241	0.152	0.298
Dana Berman, Tali Treibitz, Shai Avidan (result generated by code available in Author's GitHub) Diving into Haze-Lines: Color Restoration of Underwater Images								
DCP	0.669	0.222	0.230	0.217	0.658	0.200	0.151	0.308
He Kaiming (result generated with implementation by MATLAB) Single Image Haze Removal Using Dark Channel Prior								
bglight_icip	0.613	0.200	0.201	0.212	0.625	0.156	0.098	0.371
Chau Yi Li, Andrea Cavallaro Background Light Estimation for Depth-dependent Underwater Image Restoration								
Anonymous	0.606	0.195	0.196	0.215	0.627	0.140	0.098	0.389
Anonymous entry								

[Top]

(a)

Image: fish.png

	UCIQE				UIQM			
Team	UCIQE	σ_c	con_l	μ_s	UIQM	sharpness	contrast	colourful
UWHL	0.681	0.247	0.235	0.199	0.810	0.338	0.104	0.368
DCP	0.662	0.241	0.216	0.205	0.615	0.162	0.132	0.321
bglight_icip	0.639	0.240	0.194	0.205	0.669	0.190	0.063	0.416
Anonymous	0.629	0.230	0.192	0.208	0.684	0.152	0.064	0.468

[Top]

(b)

Fig. 5: (a) Snapshot of the leaderboard for the dataset ReefnFish. Three methods are associated to their authors and publications [7,9,19]. (b) Snapshot of the leaderboard for an image in the ReefnFish dataset.

6. Peng, Y., Cosman, P.: Underwater image restoration based on image blurriness and light absorption. IEEE Transactions on Image Processing. Vol. 26, pp. 1579-1594 (2017).
7. Li, C. Y., Cavallaro, A.: Background light estimation for depth-dependent underwater image restoration. IEEE International Conference on Image Processing (2018).
8. Ancuti, C., Ancuti, C., De Vleeschouwer, C., Bekaert, P.: Color balance and fusion for underwater image enhancement. IEEE Transactions on Image Processing. vol. 27, pp. 379-393 (2018).
9. Berman, D., Treibitz, T., Avidan, S.: Diving into haze-lines: Color restoration of underwater images. British Machine Vision Conference. (2017)
10. Chiang, J., Chen, Y.: Underwater image enhancement by wavelength compensation and dehazing. IEEE Transactions on Image Processing. vol. 21, pp. 1756-1769 (2012).

11. The Middlebury Computer Vision Pages, http://vision.middlebury.edu/flow/. Last accessed June 2018.
12. Multiple Object Tracking Benchmark, https://motchallenge.net/. Last accessed June 2018.
13. Yang, M., Sowmya, A.: An underwater color image quality evaluation metric. IEEE Transactions on Image Processing. vol. 24, pp. 6062-6071 (2015).
14. Panetta, K., Gao, C., Agaian, S.: Human-visual-system-inspired underwater image quality measures. IEEE Journal of Oceanic Engineering, vol. 41, pp. 541-551 (2016).
15. Hasler, D., Suesstrunk, S.: Measuring colorfulness in natural images. SPIE. pp. 87-95 (2003).
16. Panetta, K., Agaian, S., Zhou, Y., Wharton, E.: Parameterized Logarithmic Framework for Image Enhancement. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 41 (2011).
17. BubbleVision Youtube Channel, https://www.youtube.com/user/bubblevision. Last accessed June 2018.
18. Bins, J., Draper, B., Bohm, W., Najjar, W.: Precision vs. error in JPEG compression. SPIE. pp. 3817 (1999)
19. He, K., Sun, J., Tang, X.: Single image haze removal using dark channel prior. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, pp. 2341-2353 (2011).