



OPEN

Underwater image restoration with Haar wavelet transform and ensemble of triple correction algorithms using Bootstrap aggregation and random forests

Vahid Rowghanian

This paper presents both a new strategy for traditional underwater image restoration using Haar wavelet transform as well as a new learned model that generates an ensemble of triple correction algorithm labels based on histogram quadrants' cumulative distribution feature instead of generating pixel intensities. The Haar wavelet transform is our tentative dynamic stretching plan that is applied on the input image and its contrast stretched image to generate the degraded wavelet coefficients which are blended using Gaussian pyramid of the saliency weights to restore the original image. The ensemble of triple corrections exerts three color correction algorithms sequentially on the degraded image for restoration. The ensemble of algorithms entails the superposition effect of the red channel mean shifting, global RGB adaptation, global luminance adaptation, global saturation adaptation, luminance stretching, saturation stretching, contrast stretching, adaptive Gamma correction for red spectrum, even to odd middle intensity transference using look-up table, green to red spectrum transference using histogram equalization, local brightening, Dark Channel Prior, fusion restoration, and our Haar wavelet transform restoration. The source is available at <https://github.com/vahidr213/Underwater-Image-Restoration-And-Enhancement-Collection>.

Various researchers have recently visualized several mathematical methods and thoughts to define an effective solution for color losses in underwater images and limitations by the extent to which several literatures in the machine vision science have approached the underwater image restoration problem from different perspectives. On the other hand, underwater image restoration is by no means trivial. It is actually as challenging as most in machine vision and image processing.

The primary factor in underwater image degradation is the high attenuation of the radiated light from the surface of the objects¹. An early and famous single image haze removal for outdoor images named Dark Channel Prior (DCP) introduced the assumption of darkness in at least one of the RGB channels in any non-sky patch of the image. Underwater Dark Channel Prior (UDCP)² is derived from the DCP³ for underwater images assuming the red channel as the dark channel, hence using green and blue channels to estimate the medium transmission matrix.

From the application point of view, 3D reconstruction of seafloor encounters objects in images that require automatic true color recovery since the effects can limit the ability to assess the changes happened to underwater organisms such as those living in benthic zone⁴.

The Sea-thru method⁵ examined the dark channel prior in detail via finding the optimized medium transmission matrix by estimating the range-dependent attenuation coefficient using local space average color. An alternative to dark channel prior named red channel prior⁶ is a rearrangement of the original DCP equation as well as refining the estimated transmission map with guided filter instead of matting technique. Recently, the transmission map of the DCP method for the red channel have been estimated⁷ by applying adaptive non-convex and non-smooth variation to a rough initial transmission map and then using this optimized transmission map to obtain red global background light via thermal exchange optimization, then the refined red channel is recombined with green and blue channels to form the restored image. As mentioned before, a complex image formation model is proposed¹ that has drawn attention to be utilized in⁸ for depth and background light estimation. In⁹, an

Independent Author, Ahvaz, Iran. email: v.rowghanian@gmail.com

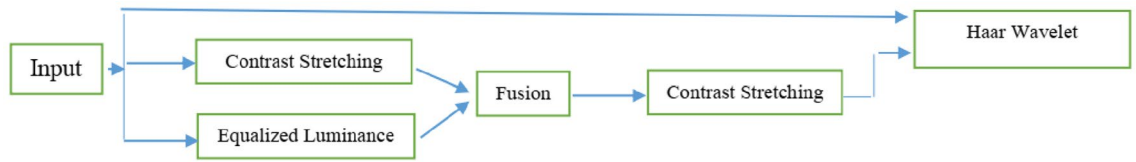


Figure 1. The flowchart of the multi-scale fusion¹⁰ and Haar wavelet transform for underwater image restoration.

attempt toward the estimation of background light and the transmission map is performed by using the feature priors of flatness, hue, and brightness.

Despite the prolific works as yet presented DCP and UDCP approaches, none have comprehensively sufficient positive effect on the underwater image restoration problem since the dark channel prior formation model is regulated for the atmospheric haze and due to being a weak function of wavelength, it's inadequate to be rendered for light propagation in harsh oceanic environments¹.

Other strategies such as multi-scale fusion has drawn much attention in recent years. The method proposed by Ancuti et al.¹⁰ has been built by fusing four weights (namely, exposedness, Laplacian of luminance, locally averaged luminance and saliency) extracted from both a contrast stretched image as well as an equalized luminance image to determine how each pixel should be modified.

Inspired by UDCP and multi-scale fusion, an algorithm is proposed¹¹ so that the medium transmission map produced by UDCP is decomposed into saliency and Laplacian weights. These weights are joined by multi-scale fusion to produce a refined medium transmission map which is used in DCP image formation model to restore the original image. However, fusion methods mentioned above produce equalized distributions in RGB spectrums that can cause over-enhancement or under-enhancement at some regions of image where the distribution is yet to remain untouched or shall have carried to the farthest ends of the intensity distribution. Moreover, saliency weights in both methods largely take up the principal structure of the weightings in contrast to other weights which have negligible effect.

Over the past decades, convolutional neural networks (CNN) have extended in many visual recognition fields. An encoding–decoding convolutional neural network¹² is used to restore underwater images where the encoding and decoding levels are visualized by convolution and deconvolution operations respectively. The inner structure of the convolutional part is similar to the popular Alexnet network. They have accelerated the training process of the CNN as well as including low-level features by using skip connection between the encoding and decoding stages. The skip connection has been used since few convolutions extract some important features, but it can smooth the image and destruct the edges if over applied especially in underwater images.

Neural networks are powerful and capable of constructing various architectures. Some of published works in the past few years, are domain adversarial learning¹³, generative adversarial network (GAN)¹⁴, underwater convolutional NN¹⁵, multi-scale dense block GAN¹⁶ using batch normalization¹⁷, underwater joint residual learning CNN¹⁸, underwater image restoration network¹⁹, WaterGAN²⁰, MCycleGAN²¹, underwater image enhancement with stacked conditional GAN²².

The above mentioned networks still reflect a considerable rate of failure due to lack of handling the saturation and contrast problems. In²³, an end-to-end CNN with three blocks named RGB (for basic operations), HSV (saturation and luminance adjustment) and attention (stage for quality enhancement) is proposed where the final restored image is produced by a weighted sum between the RGB block output and the attention block's RGB component as well as a weighted sum between HSV block output and the attention block's HSV component. The input of the attention block is a concatenation of the raw image and the images of other two blocks (RGB block and HSV block). As a good survey, some of above mentioned methods as well as other methods are recapitulated in²⁴.

The following literature explains the proposed method, the inputs, the structure of the multi-scale fusion restoration and 2-D Haar wavelet transform restoration, the structure of the Ensemble of Triple Correction Algorithms (ETCA) and finally the experimental results and comparison with other state of the art algorithms.

Proposed method

In this paper, our contribution to the problem of underwater image restoration includes a traditional method as well as a learned model using new features. The traditional methods solve portions of the whole problem quite topnotch. We used an alternative single image-based solution through blending the Haar wavelet coefficients of the raw image with its contrast stretched version (static tolerance) to convey a sense of dynamic contrast stretching. Our approach is applied after the multi-scale fusion restoration¹⁰ to further improve the quality of the restored image. The flowchart of this approach is shown in Fig. 1.

On the other hand, a comprehensive method for random patterns requires trained models. In this paper, we use the histogram quadrants' cumulative distribution as a new feature to characterize each image instead of using pixel intensity. The features are then learned using Bootstrap Aggregation and Random Forests. We have trained three models separately. The responses of models are in the form of a numeric label instead of pixel intensities. Each model generates a numeric label corresponding to a special color correction algorithm that should be applied sequentially to restore the original image.

Input images. The input images of fusion process can generally be different because the fusion is based on combining multiple sources to preserve the significant features of them. In this regard, the contrast stretching

is one of the most popular initial color correction tools as yet has overcome significant initial white balancing. Numerous white balancing methods have been suggested but none of them are always experimentally appropriate in underwater scenes. The Gray-World approach of Bachsbaum et al.²⁵ is a mediocre example of white balancing methods, yet creates color deviation by introducing reddish artifacts where the appearance is overall blue. The following method is used to adjust white balance initially which is simple and efficient and creates less red artifacts. At first, a three-element vector is defined as below:

$$ratio = K \left[\frac{I_{max}^R}{\bar{I}^R}, \frac{I_{max}^G}{\bar{I}^G}, \frac{I_{max}^B}{\bar{I}^B} \right] \quad (1)$$

where K is a constant (e.g. 0.005) to make a vector of probability ratios in the interval $[0, 1]$. The three cumulative probabilities presented in $ratio$ and their complements in $(1-ratio)$ determine the percentage of the data that should be saturated to the lowest and highest values in each of the three RGB channels respectively.

Haar wavelet inputs. The Haar wavelet transform is fed with the raw input image as well as a %0.5 contrast stretched version of the raw image through the method described above. The input image for Haar transform is resized into 640×640 in order to have the Haar pyramid dimensions as equal as the Gaussian pyramid of the blending weights (W_1, W_2) as well as considering a power of 2 dimension to facilitate the wavelet transform. The input image can also be resized to 1024×1024 to increase the accuracy.

Multi-scale fusion input. The first image for the pyramid of multi-scale fusion¹⁰ is the contrast stretched version of the input image explained above and the second one is an equalized luminance image provided by adaptive histogram equalization²⁶.

Weights of multi-scale fusion. The choice of correct weights reduces computations and artifacts. There are four weights in¹⁰ that are explained below:

Laplacian contrast weight (W_L). Laplacian contrast weight is the absolute value of the Laplacian filter applied on the luminance component of the image.

Local contrast weight (W_{LC}). Local contrast weight is an additional contrast measure to recover the contrast in the regions where the Laplacian contrast weight (W_L) is not sufficient (e.g. ramp and flat regions). It improves the transitions between dark and bright regions. The (W_{LC}) is computed as the square of the difference between the luminance and its locally averaged luminance (e.g. 5×5 , $[1,4,6,4,1]/16$ filter plus cutting off values above $\pi/2.75$ to $\pi/2.75$).

$$W_{LC}(x, y) = (L - \bar{L})^2 \quad (2)$$

Saliency weights (W_S). Saliency weights emphasize on the pixels standing farthest relative to the mean of the reddish, bluish and luminance components collectively. In other words, the salient objects in a scene are biologically recognized and similarly mathematically defined as yet taken up a region of center-surround contrast. The saliency algorithm of Achanta et al.²⁷ is used which is defined in Lab color space by:

$$W_S(x, y) = (L(x, y) - \bar{L})^2 + (a(x, y) - \bar{a})^2 + (b(x, y) - \bar{b})^2 \quad (3)$$

Exposedness weights (W_E). Exposedness weights are a rough estimation of how much a pixel is exposed to light. This weight is defined by a Gaussian profile resembling (^) sign with mean value 0.5 and standard deviation 0.25. Therefore, pixels close to the average value are exaggerated to produce better results when combined with saliency weights. Exposedness weight (W_E) is defined as:

$$W_E(x, y) = \exp\left(-\frac{(I(x, y) - 0.5)^2}{2\sigma^2}\right) \quad (4)$$

The final weights W_1 and W_2 combine the above mentioned four weights and have the same size as input image (like all four weights). The final weights determine the amount of modifications that each pixel will receive after the fusion applied. The final weights W_1 and W_2 are computed as below:

$$W_1 = \frac{W_{L1} + W_{LC1} + W_{S1} + W_{E1}}{W_{L1} + W_{LC1} + W_{S1} + W_{E1} + W_{L2} + W_{LC2} + W_{S2} + W_{E2}} \quad (5)$$

$$W_2 = \frac{W_{L2} + W_{LC2} + W_{S2} + W_{E2}}{W_{L1} + W_{LC1} + W_{S1} + W_{E1} + W_{L2} + W_{LC2} + W_{S2} + W_{E2}} \quad (6)$$

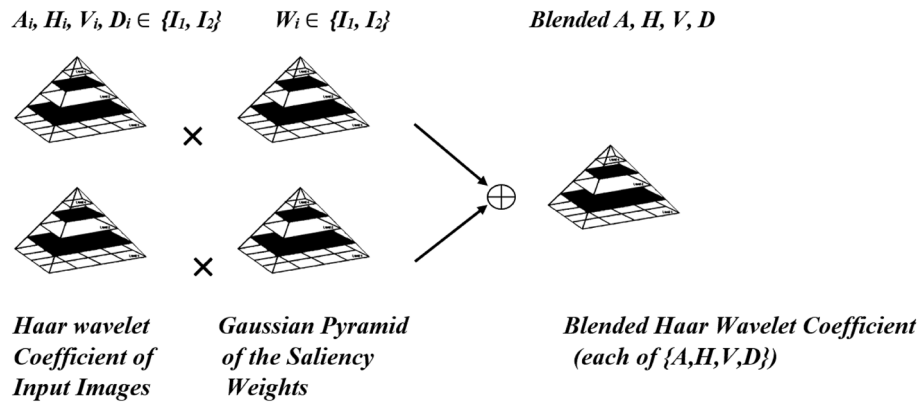


Figure 2. The block diagram for illustrating the Haar wavelet coefficient refinement.

Haar wavelet coefficient blending. Wavelet and inverse wavelet transformations have been designed as an analyzing and synthesizing tool for signals and images. The Haar wavelet transform has a good potential to decompose images into Approximation, Horizontal, Vertical and Diagonal components. Thus, it is suitable for multi-scale blending of coefficients to convey a sense of tentative dynamic stretching. A %0.5 contrast stretched image is derived from the initial degraded image (I_2). Therefore, it can generate slightly different type of wavelet coefficients which still carry degradation information. In this paper, we propose to blend the degraded Haar wavelet coefficients $\{A, H, V, D\}$ of the two images (I_1, I_2) using the Gaussian pyramid of the saliency weights (W_1, W_2) to achieve new coefficients ($A_{new}, H_{new}, V_{new}, D_{new}$). Figure 2 demonstrates the structure of the coefficient blending.

In the blending process, the saliency weights are used solely since the saliency weights take up the principal structure among the four weights described above. The refinement or blending weights (W_1, W_2) are computed according to the following equations:

$$W_1 = \frac{W_{S1}}{W_{S1} + W_{S2}} \tag{7}$$

$$W_2 = \frac{W_{S2}}{W_{S1} + W_{S2}} \tag{8}$$

The blending of the degraded Haar wavelet coefficients are performed according to the following equations:

$$H_{C_{new}}^l(x, y) = \sum_{k=1}^K G^l \{ W^k(x, y) \} H^l \{ I_c^k(x, y) \} \tag{9}$$

$$V_{C_{new}}^l(x, y) = \sum_{k=1}^K G^l \{ W^k(x, y) \} V^l \{ I_c^k(x, y) \} \tag{10}$$

$$D_{C_{new}}^l(x, y) = \sum_{k=1}^K G^l \{ W^k(x, y) \} D^l \{ I_c^k(x, y) \} \tag{11}$$

$$A_{new} = \sum_{k=1}^K G^l \{ W^k(x, y) \} A^k \tag{12}$$

where $H^l \{ I_c^k(x, y) \}, V^l \{ I_c^k(x, y) \}, D^l \{ I_c^k(x, y) \}$ are the horizontal, vertical and diagonal coefficients of k th image for the channel $c \in \{R, G, B\}$ at level l . A_k is the approximation coefficient matrix for image k .

The inverse Haar transform is then applied on the refined coefficients ($A_{new}, H_{new}, V_{new}, D_{new}$) to restore the original image.

The Haar wavelet restoration is able to restore underwater images independently. Nevertheless, our experiments shown the overall improvement of efficiency through using the Haar wavelet restoration method as a complementary algorithm. Thus, we applied the Haar wavelet restoration method on the restored image produced by multi-scale fusion¹⁰. Figure 1 shows the flowchart of our Haar wavelet restoration and multi-scale fusion¹⁰ structure. The overall algorithm shown a record breaking result representing an average mean square error (MSE) better than almost all trained models and traditional methods except for Dive + algorithm.

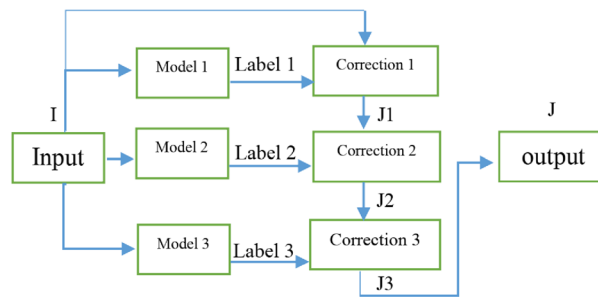


Figure 3. The block diagram illustrating the flow chart of the Ensemble of Triple Correction Algorithms. The three models generate three numeric labels for three stages (Correction 1, Correction 2, and Correction 3) to apply the corresponding color correction operation on the raw image.

Multi scale fusion. The fusion is represented by a weighted sum of images at every location (x,y) :

$$R(x,y) = \sum_{k=1}^K \bar{W}^k(x,y) I^k(x,y) \quad (13)$$

where $I^k (M \times N)$ is the input image ($k = 1, 2, \dots, K$) with M rows and N columns, $\bar{W}^k (M \times N)$ is the normalized weights ($k = 1, 2, \dots, K$) ($\sum \bar{W}^k = 1$) and R is the restored image. I^k can generally be any color corrected form of the raw input image.

Due to halos and artifacts that are introduced in $R(x,y)$ by directly applying equation above, both weights and input images are decomposed into a multi-scale pyramid defined below:

$$R^l(x,y) = \sum_{k=1}^K \{ \bar{W}^k(x,y) \}^l \{ I^k(x,y) \}^l \quad (14)$$

where l represents the number of pyramid levels. Each scale is derived by down sampling the previous level. The initial level is the original image.

To preserve the important and desired information, an operation such as Gaussian filtering or Laplacian filtering can be applied before down sampling. In this regard, we decomposed weights into a Gaussian pyramid and decomposed images into a Laplacian pyramid. Therefore, the equation above can be rewritten by:

$$R^l(x,y) = \sum_{k=1}^K G^l \{ \bar{W}^k(x,y) \} L^l \{ I^k(x,y) \} \quad (15)$$

where $L^l \{ I^k(x,y) \}$ is the Laplacian of the k th image at level l , and $G^l \{ \bar{W}^k(x,y) \}$ is the Gaussian of the k th weight at level l . The Gaussian pyramid is built by Gaussian filtering (or 5×5 , $[1,4,6,4,1]/16$) and down sampling the image repetitively until we reach to the maximum level, considering the first layer to be the same size but only filtered version of the input image.

The Laplacian pyramid is created by building a pyramid of down sampled images at first, then calculating the difference of the lower level image from the resized (enlarged) upper level image (using bi-cubic interpolation). The first level for both the Gaussian pyramid and the Laplacian pyramid is as the same size as the input image to increase the accuracy.

Ensemble of triple corrections

Traditional methods present a global solution (or limited joint solutions) for a typical prior assumption such as homogeneous lighting along the line of sight, unbalanced color spectrum in RGB channels, saturated colors, low contrast and brightness, blurriness, and especially degradation based on underwater image formation. On the other hand, the comprehensiveness of a specific traditional method solving different patterns of underwater degradation is as yet undecided.

There are several conventional color correction techniques aimed at increasing the visual quality of an image. These techniques include stretching, global adaptation, sharpening, unsharp masking, histogram equalization and many others. In this paper, we propose to predict and choose the best three combinations from an ensemble of conventional color correction techniques as well as our current proposed methods to be applied on a single image according to the features extracted from the image. In this paper, we have also proposed new features to be learned which are explained in detail in the feature extraction section. Three models are trained with these extracted features. The training method is Bootstrap Aggregation and Random Forests. Each model is responsible for generating a numeric label for its corresponding stage that indicates a specific color correction method among the fifteen available methods. It should be noted that the triple corrections are applied sequentially as the block diagram of the proposed method ETCA shows in Fig. 3. The ensemble of color correction methods entails

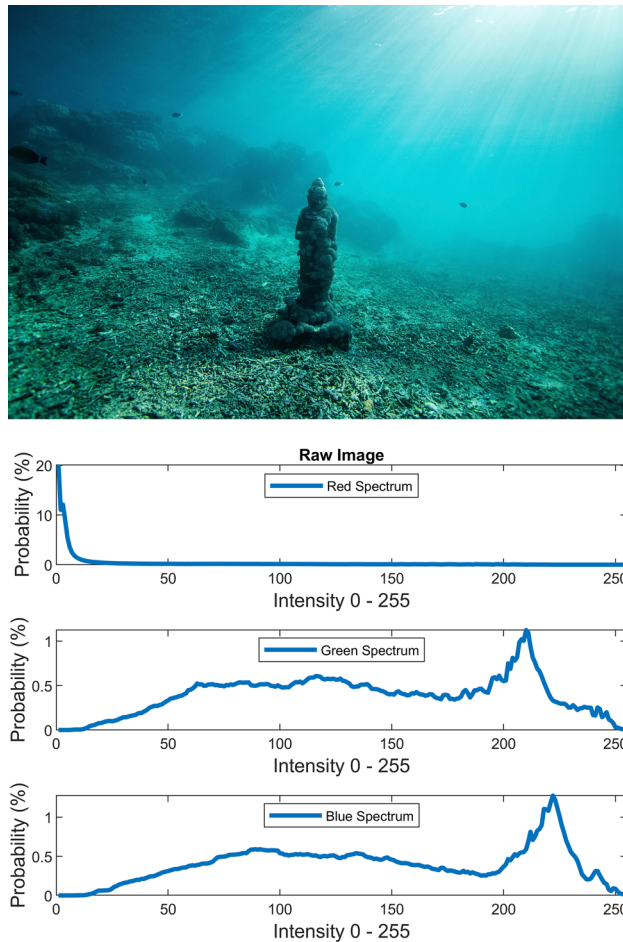


Figure 4. The raw image and the RGB histograms (from unsplash.com).

the superposition effect of the red channel mean shift and sharpening, global RGB adaptation, global luminance adaptation, global saturation adaptation, luminance stretching, saturation stretching, contrast stretching, adaptive Gamma correction for red spectrum, our even to odd middle intensity transference using look-up table, green to red spectrum transference using histogram equalization, local brightening, DCP³, fusion restoration¹⁰, and our Haar wavelet transform restoration. For comparison purposes, Fig. 4 is used for all the subjective visualizations. Figures 5, 6 illustrate the visual and spectral effects of DCP and local brightening methods applied on Fig. 4.

The total number of the available methods including no action is equal to fifteen. It should be noted that the null operation leaves the image intact. This was necessary due to existence of some fully enhanced images in the UIEB dataset which had a negligible difference with their reference image (negligible mean square error). Other color correction algorithms are assigned a special numeric label identifying the corresponding operation. The color correction methods are described in the feature extraction section in detail.

Features extraction. The feature extraction plays a key role in the quality of the deep learning-based underwater enhancement methods. As far as we know, most of the recent deep learning models have used the RGB values as their features. A majority of the state of the art underwater deep learning methods map pixel to pixel or pixel to difference. Pixel to pixel mapping mutually produces an output pixel in response to an input pixel. Pixel to difference mapping produces a positive or negative amount in the output to be added to or subtracted from the input pixel intensity. However, the possibility of generating multiple RGB combinations in the outputs of these networks still remain unsolved and will fall behind the generalization of the such conventional deep learning-based underwater restoration methods.

In this regard, we propose a probability based solution to be an alternative feature for pixel intensity. We have used cumulative distribution function (*CDF*) in 4 closed intervals or regions of the probability distribution function (*PDF*) of the input image. Therefore, each quadrant of the image's *PDF* is assigned with a *CDF* scalar. Since 8-bit RGB images have three channels, a total of twelve scalars are extracted for each input image using equations below.

$$S_1^c = f^c(I \leq \Omega_1) \quad (16)$$

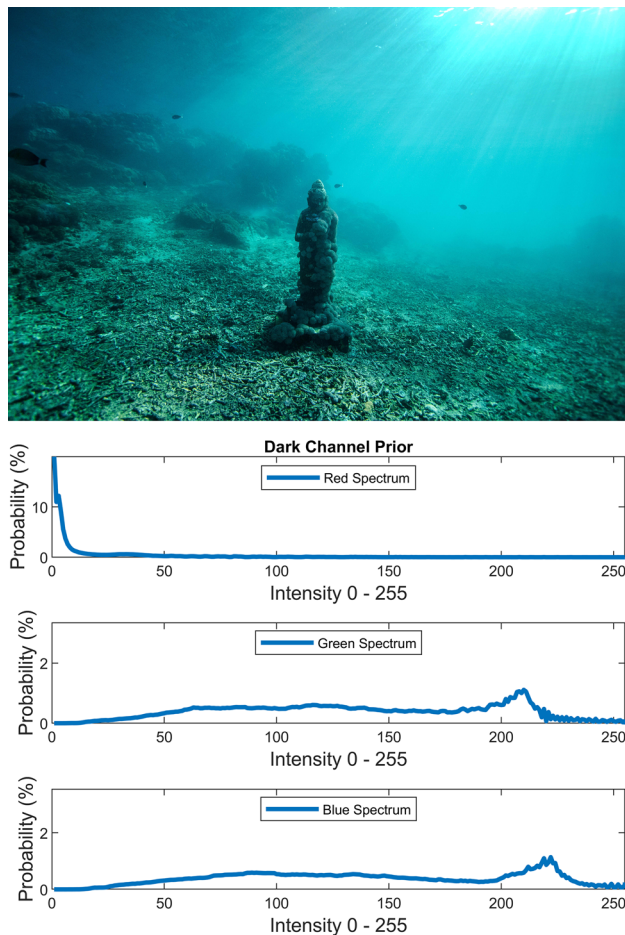


Figure 5. Dark Channel Prior.

$$S_2^c = f^c(I \leq \Omega_2) - f^c(I \leq \Omega_1) \quad (17)$$

$$S_3^c = f^c(I \leq \Omega_3) - f^c(I \leq \Omega_2) \quad (18)$$

$$S_4^c = f^c(I \leq \Omega_4) - f^c(I \leq \Omega_3) \quad (19)$$

where S_n^c is the sum of probabilities in the n th interval, f is the probability distribution function or histogram, $c \in \{R, G, B\}$, and $\Omega_1, \dots, \Omega_4$ are the maximum boundary of each quadrant which their numeric values are in the experimental results.

We have used these twelve CDFs as input features. The cumulative distribution (CDF) of each histogram quadrant decreases the number of training features significantly in contrast to huge training data used in above mentioned networks. The histogram quadrants also decrease the training time significantly.

The training process also needs the corresponding responds for which the input features have been extracted. In this work, the respond of the network is also not a pixel intensity. The respond to each input feature is an ensemble of optimized triple labels minimizing the cost function (in our case, the Mean Square Error between the reference image and the restored image). Each numeric label corresponds to a single special color correction method and each ensemble of triple labels has to be applied on the input image sequentially to restore the original image. Therefore, an optimization has been performed to generate the optimum triple correction methods (labels) for each input image.

A total of fifteen color correction methods used in evaluations. One of them is no action. Hereafter, other color correction methods are described in detail. It should be mentioned that the superiority of contrast stretching and pyramidal methods over other proposed methods with shallow positive effects is obviously not a strange phenomenon but the essence of beneficiary complementary approaches is exclusively indispensable for other aspects of visual improvements.

- The static contrast stretching in RGB color space, static luminance stretching in Lab color space, and static saturation stretching in HSV color space are performed simply by using the following equation which satu-

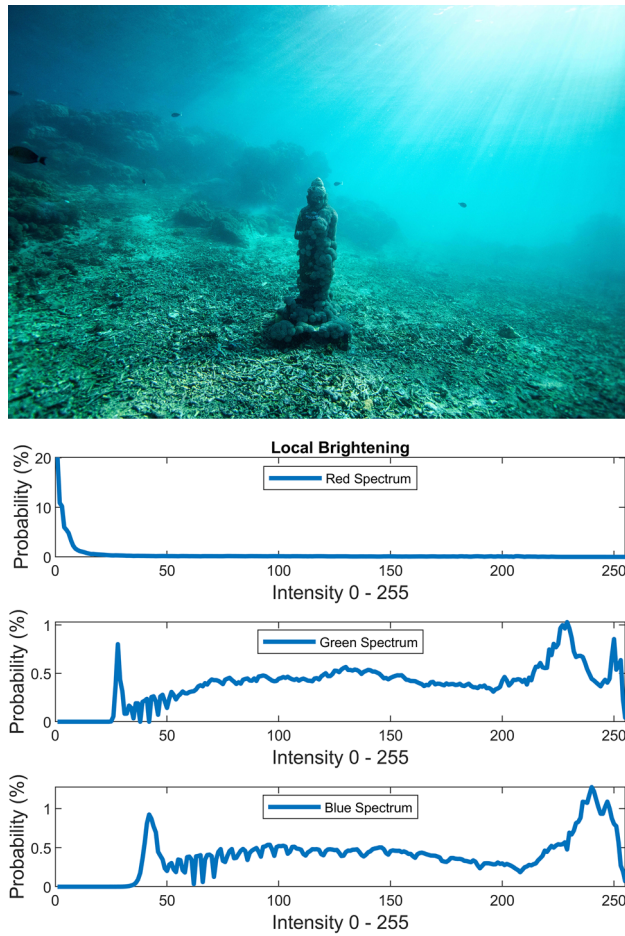


Figure 6. Local brightening.

rates the bottom 1% and the top 1% of contrast, luminance, and saturation values (with proper L and H) to the lowest and highest values.

$$J = \frac{I - L}{H - L} \tag{20}$$

where L and H are the intensities corresponding to bottom 1% and top 1% of the image I . Figures 7, 8, 9 illustrate the visual and spectral effects of these stretching methods applied on Fig. 4.

- The global RGB adaptation, global luminance adaptation and global saturation adaptation are performed using the logarithmic approximate function of human visual system (in early stages) according to the Weber-Fechner law²⁸ and the retinex theory²⁹. The global adaptation of an image I is expressed by:

$$I_g(x, y) = \frac{\log\left(1 + \frac{I(x, y)}{\bar{I}}\right)}{\log\left(1 + \frac{I_{max}}{\bar{I}}\right)} \tag{21}$$

where scalar I_{max} is image maximum pixel intensity and the scalar \bar{I} is the log-average of the image that enables the I_g to adapt to the scene and is computed by equation below:

$$\bar{I} = \exp\left(\frac{1}{MN} \sum \sum \log(I(x, y) + \delta)\right) \tag{22}$$

where $I(x, y)$ is an $M \times N$ gray scale image and δ is small value to avoid singularity (e.g. $\delta = 10^{-3}$). As the log-average \bar{I} increases to larger values, the global adaptation equation I_g behaviors more in linear shape than logarithm shape. Therefore, darker scenes (i.e. contrast, saturation and luminance spectrums) are brightened more than brighter scenes through global adaptation. Figures 10, 11, 12 illustrate the visual and spectral effects of these global adaptation methods applied on Fig. 4.

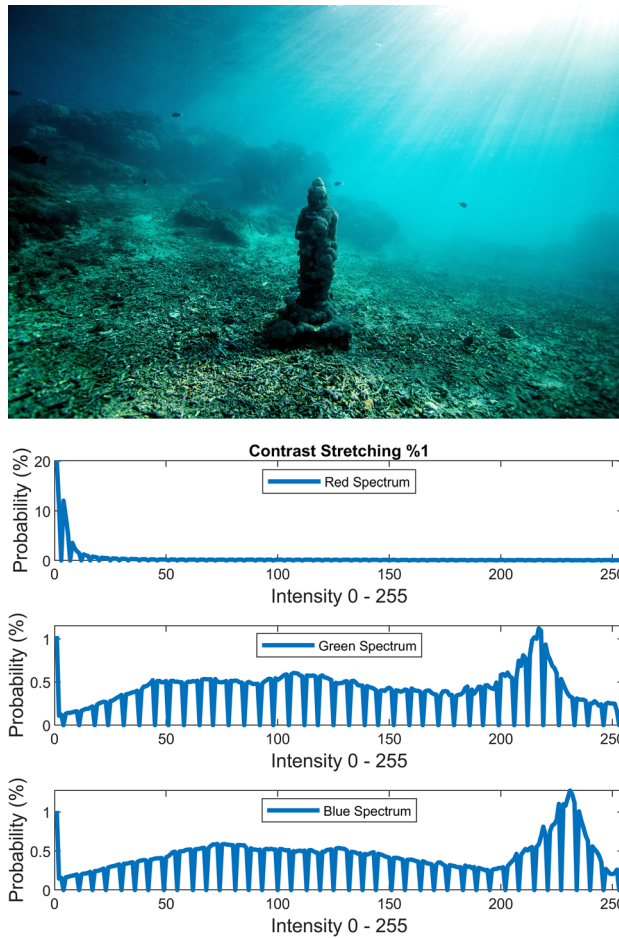


Figure 7. %1 Contrast stretching.

- Adaptive Gamma correction is performed using the method proposed in³⁰ which provides an automatic way to adaptively compute the Gamma for a given image instead of using a constant scalar. This method is computed through equations below for a gray scale image I :

$$\gamma = \exp\left(\frac{1 - (\mu + \sigma)}{2}\right) \tag{23}$$

$$c = \frac{1}{1 + (I^\gamma + (1 - I^\gamma)\mu^\gamma - 1) \times \text{heaviside}(0.5 - \mu + \delta)} \tag{24}$$

$$I_{out} = cI_{in}^\gamma \tag{25}$$

where scalar μ is the average value of the image, scalar σ is the standard deviation of the image, and δ is an infinitesimal value to avoid zero input to Heaviside function. Figure 13 illustrates the visual and spectral effect of this method applied on Fig. 4.

- Red channel mean shifting and sharpening is composed of two operations. It should be noted that this approach is used in only few cases since this method requires a user defined scalar for blending. The initial operation shifts the mean of the red channel toward the gray image's mean and then blends the green spectrum into red spectrum according to the following equation:

$$I^R = \frac{\mu_{Gray}}{\mu_R} I^R + 0.1(\mu_G - \mu_R) I^G \left(1 - \frac{\mu_{Gray}}{\mu_R} I^R\right) \tag{26}$$

The second operation sharpens the red channel I^R using unsharp masking. Unsharp masking is performed by subtracting a blurred (unsharp) version of the image from the initial image. Figure 14 illustrates the visual and spectral effect of this method applied on Fig. 4.

- In this work, we propose the idea of even to odd middle intensity transference using look-up table to augment the probability of middle intensities (roughly doubling). After transferring the even middle intensities

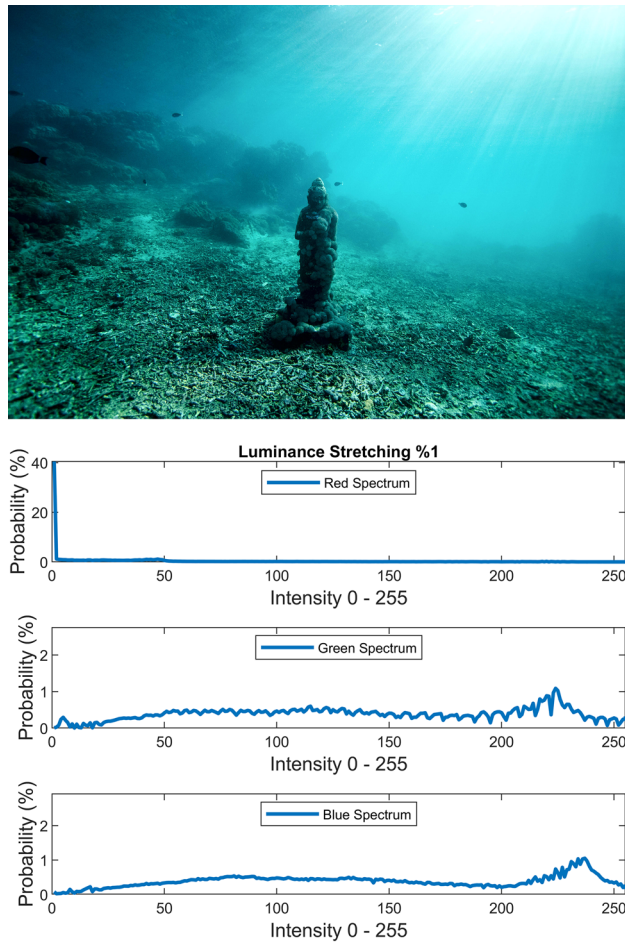


Figure 8. %1 Luminance stretching.

to their adjacent odd intensities, every odd middle intensity will roughly have a doubled probability. This phenomenon is due to the continuity of the cumulative distribution (cdf) of the discrete and continuous data. In other words, two adjacent intensities have quite slight difference in their probabilities. The selected middle intensities are between [8 and 250]. Other intensities are left intact. This method partially affects the image and has no extensive positive or negative effect on the image. Figure 15 illustrates the visual and spectral effect of this method applied on Fig. 4.

- Green to red spectrum transference using histogram equalization is performed since we have observed many positive effects using green spectrum in contrast to blue spectrum. The transference is performed with the following equation:

$$\rho_{red} = \rho_{green} \quad (27)$$

where ρ_{green} is the green spectrum which is used in histogram equalization. Figure 16 illustrates the visual and spectral effect of this method applied on Fig. 4.

- The local brightening³¹ is used as one of our color correction methods but it should be noted that local brightening method quite degrades the overall underwater images. Due to this effect, we have set the blending option of the brightening method to (0.1).

We have used the Bootstrap Aggregation (Bagging) and Random Forests^{32,33} to create three trained models with which the three labels are generated. The first, second and third labels are generated separately through their designated trained model and then these labels are sequentially exerted on the input according to their order.

Experimental results

A comprehensive and fair experiment of several traditional underwater image enhancement methods is performed in UIEB dataset paper³⁴ from which our raw and reference images are drawn. Likewise, a comprehensive evaluation of several recent learned models (deep learning) is performed in³⁵ from which we quote them in Table 1. The UIEB dataset³⁴ has 890 real-world paired images (raw and reference) plus 60 challenging raw images. The images are 8-bit and taken in different situations and viewing angles and have different dimensions. The paired images facilitate error measuring in different forms such as Mean Square Error (MSE), Peak Signal

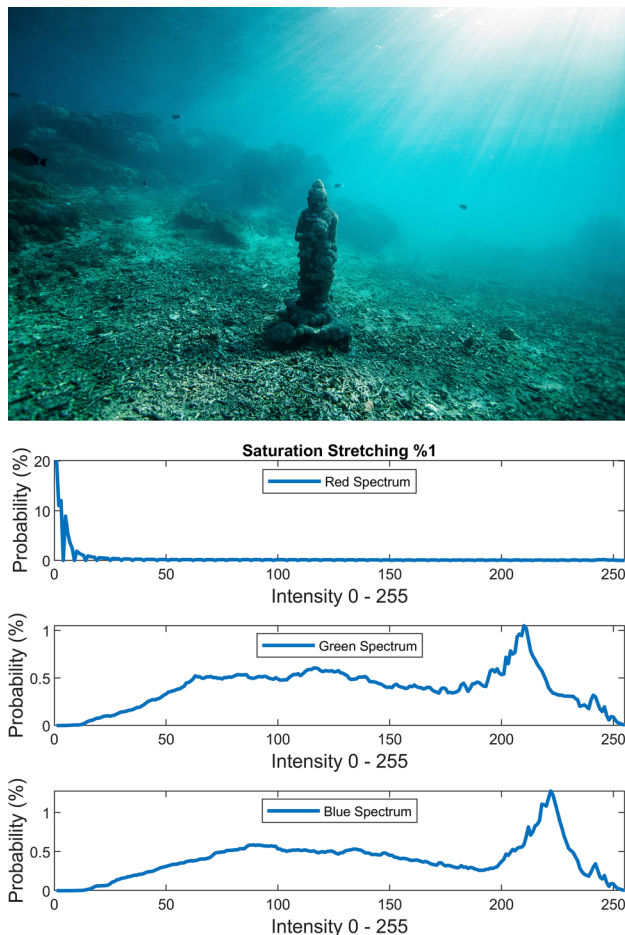


Figure 9. %1 Saturation stretching.

to Noise Ratio (PSNR) and Structural Similarity Index for Measuring image quality (SSIM). In UIEB dataset, the degraded images are paired with the references by asking 50 volunteers to select one image out of 12 enhanced results. The enhanced results were produced by 12 traditional methods while one of the was Dive + app (with predefined settings).

Table 1 quotes the performance of our methods as well as the performance of other methods according to^{34,35}.

Our simulations are executed via Matlab software and the source is freely available for evaluation.

The speed performance depends on the image size, methods applied on the image and the processor. The larger the image, the slower the result is produced since all of our color correction algorithms are applied on the whole image. Therefore, the number of pixels has direct influence on the speed performance. The maximum latency happens when both multi-scale fusion¹⁰ and our proposed Haar wavelet restoration go along with each other inside the ensemble of triple algorithms. The frame rate (FPS) for red channel mean shift, contrast stretching, luminance stretching, saturation stretching, global RGB adaptation, global luminance adaptation, global saturation adaptation, DCP, even to odd transference, local brightening, green to red transference, multi-scale fusion¹⁰, and Haar wavelet is 0.60, 35, 0.70, 4.95, 3.08, 0.67, 3.87, 0.28, 312.0, 2.70, 5.5, 0.24, and 0.93 respectively which is the average measured with an Intel Core-i3 2100–3.1 GHz processor on a 1120 × 1380 image after 300 repetitions.

The number of levels used in decomposing the input images and weights in both multi-scale fusion and Haar wavelet transform is equal to $l=5$. The maximum boundary of each quadrant depends on the number of histogram bins. We have used 256-bin histogram. Therefore, the maximum boundaries $\Omega_1, \Omega_2, \Omega_3, \Omega_4$ will be {63,127,191,255}.

The underwater image restoration using Haar wavelet coefficients refinement has shown a superior improvement in terms of average MSE, PSNR and SSIM on the whole dataset over traditional and deep learning methods as you can see in Table 1. The only superior method over our proposed traditional method is the commercial app called Dive +. As mentioned before, the images are resized to 640 × 640 in image restoration using Haar wavelet transform. Due to this, the accuracy decreases while the processing speed increases. It is possible to increase the input image to higher dimensions to increase the precision since the Gaussian pyramid of the saliency weights will carry more information to be used in the modification of the final pixel.

Our method ETCA is flexible and it is possible to use other color correction algorithms that are more robust instead of some of our current fifteen methods. As it can be seen in Table 1, the best performance among all the

Method	MSE ↓	PSNR(dB) ↑	SSIM ↑
Original	1768.9	17.36	0.6168
Fusion-based ¹⁰	867.9	18.7461	0.8162
Two-step-based ³⁶	1114.6	17.6596	0.7199
Retinex-based ³⁷	1353.1	16.8757	0.6233
UDCP ²	5130.0	11.0296	0.4999
Regression-based ³⁸	1136.5	17.5751	0.6543
GDCP ³⁹	3634.5	12.5264	0.5503
Red Channel ⁶	2107.3	14.8935	0.5973
Histogram prior ⁴⁰	1628.2	16.0137	0.5888
Blurriness-based ⁴¹	1582.6	16.1371	0.6582
MCycleGAN ²¹	1132.2	18.33	0.6138
URCNN ¹⁸	2195.8	15.94	0.5972
UWGAN ⁴²	1853.7	16.06	0.2945
DUIENet ³⁴	1012.2	19.29	0.8093
DenseGAN ¹⁶	1363.6	17.56	0.4239
UWCNN type-I ¹⁵	2345.0	15.00	0.5306
UWCNN type-III ¹⁵	2920.2	14.24	0.4945
Dive+	535.8	20.8408	0.8705
Our ^(10 + Haar)	745.3	22.1079	0.8105
Our(ETCA)	210.4	34.9780	0.8956

Table 1. Full-reference image quality evaluation.

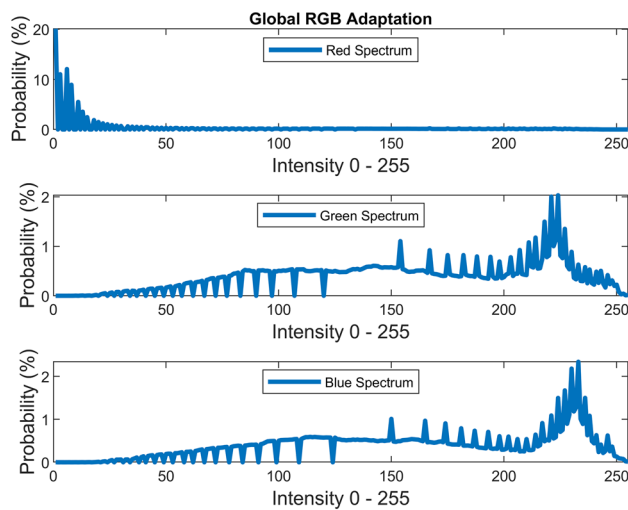


Figure 10. Global RGB adaptation.

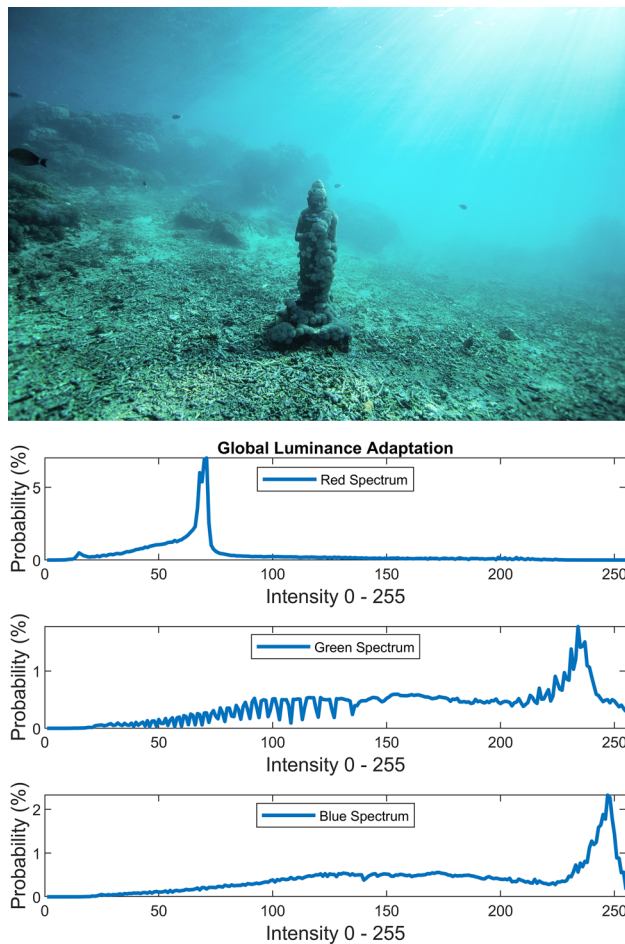


Figure 11. Global luminance adaptation.

presented methods belongs to our proposed ETCA. Our evaluations and calculations shown that the ETCA has an average percent of MSE improvement of nearly %83.4 which is a high percentage. This means that most of the MSEs are low and only few cases are least improved.

During optimization, more than 1460 triple permutations are evaluated to reach the best ensemble of labels. The number of images on which the ETCA has applied null operation is 11.

Our technique shown limitations when the optimization process has produced weak ensemble of correction algorithms in few cases. By evaluating the restored images visually, we found that most of the UIEB dataset images are well detected and restored by histogram quadrants and only in very few cases some reddish artifacts are produced. In Fig. 17 evaluation of multi-scale fusion¹⁰ and our proposed methods on some test images that are not from UIEB dataset is shown (from unsplash.com).

Out of fifteen color correction algorithms, only two of them rely quite a lot on experimental user defined parameters. These two algorithm can create extreme changes when applied with inappropriate parameters. These methods include red channel mean shifting and local brightening. Both of them require a parameter for blending amount to which we have set a scalar value of (0.1).

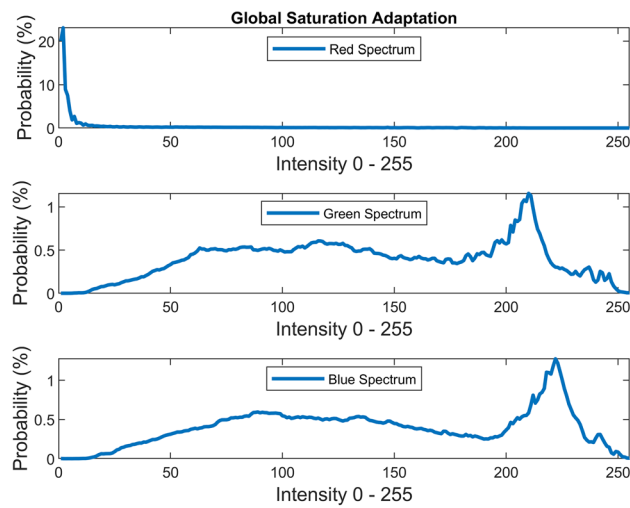


Figure 12. Global saturation adaptation.

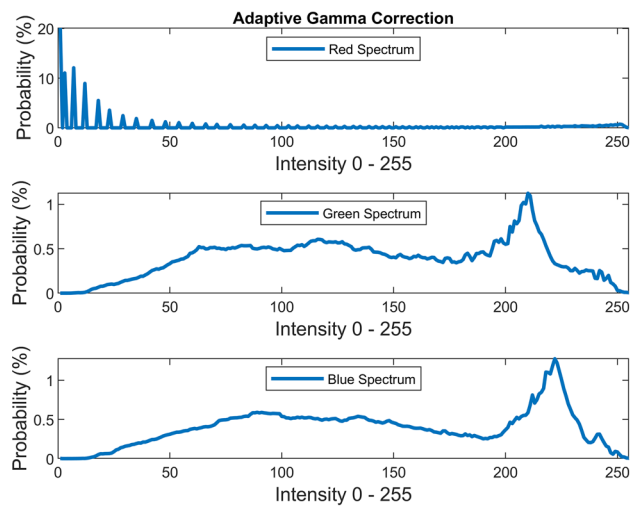


Figure 13. Adaptive Gamma correction.

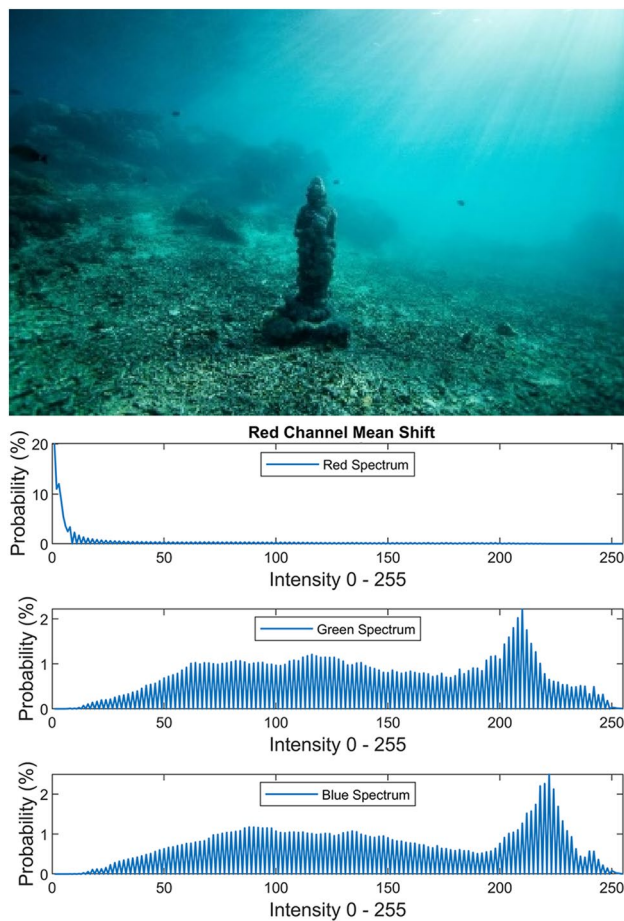


Figure 14. Red channel mean shifting.

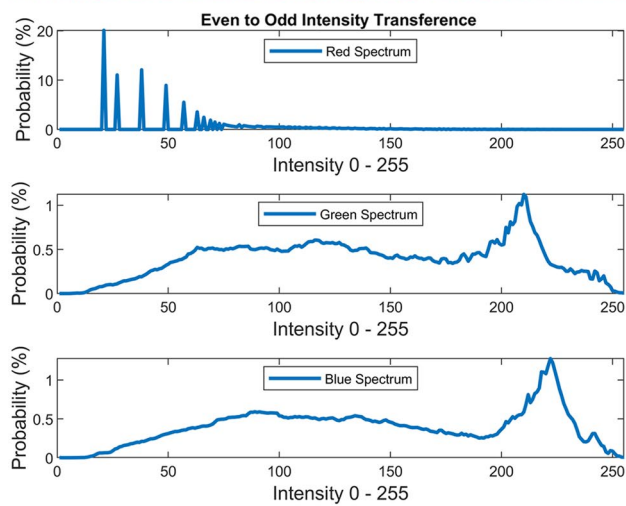
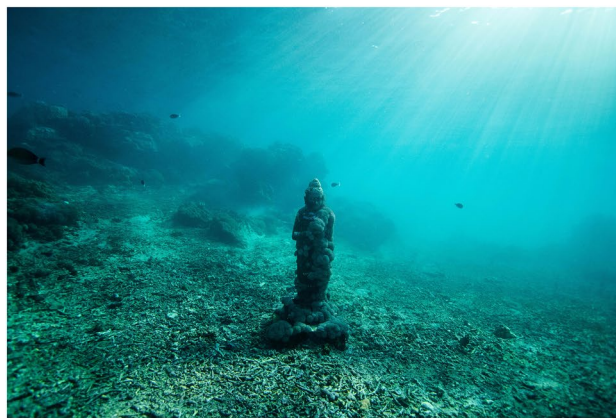


Figure 15. Even to odd transference.

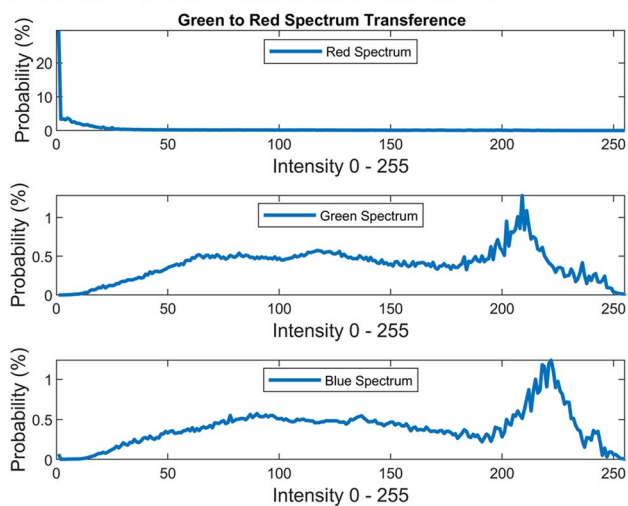
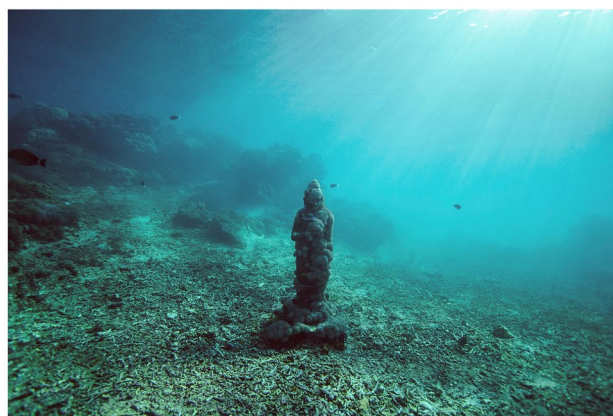


Figure 16. Green to red spectrum transference.



Figure 17. Evaluation of four methods on test images from unsplash.com: multi-scale Fusion¹⁰, our Haar wavelet, multi-scale Fusion¹⁰ plus our Haar wavelet, and proposed ETCA (from left to right).

Conclusion. The ensemble of triple correction algorithms proves the advantageous of superposition effect of different robust and weak approaches. Our evaluations on robustness of histogram's quadrants have proven the veracity of these features to be discriminative enough for raw images (before any correction). This indicates that with each stage of correction, the histogram disguises into a shape that causes the histogram quadrants become more indistinguishable. This situation entails a considerable risk if images are required to be characterized before the second stage. To confront this situation, we opt histogram octants instead of quadrants. Another feature of the ETCA is that it relies on few user-defined parameters.

The alternative approach to random forests and bootstrap aggregation is the pattern recognition neural networks. We have built the ETCA using pattern recognition neural networks with under hundred neurons in a network with a single hidden layer. While the quadrants are learned easily to characterize raw images, it is strikingly not feasible for characterizing a largely corrected image.

The Haar wavelet coefficient blending is our tentative plan to incorporate a dynamic contrast stretching method through blending the coefficients of a %0.5 contrast stretched image with the raw image (as a counterpart to static %1 stretching).

The idea of merging multi-scale fusion with wavelet coefficient refinement has shown a mediocre performance since the method generally produces stretched/equalized spectrums that is insufficient for all types of degradations. Nevertheless, among all the state of the art methods in Table 1, Dive + app merely outperforms ours.

The numerical and visual inspections via carrying out the global RGB adaptation, global luminance adaptation and global saturation adaptation on the raw images have shown to cause displeasing yields as an initial color correction operation for underwater images.

The global RGB adaptation exerted as the second or the third correction operation yields an image spectrum with confined upward compression based on the spectrum's average provided that the first color correction operation broadens the initial spectrum. Likewise, the global saturation adaptation and global luminance adaptation exerted as the second or the third operation yield upward compressed saturation and luminance spectrums respectively which will dynamically amplify the saturation and luminance spectrums according to the log-average value.

The contrast stretching, luminance stretching and saturation stretching are by orders of magnitude more engrossing than three global adaptations as initial color correction methods since most of the underwater images suffer initially from a dense narrow red spectrum. These three stretching methods also to some extent, overwhelm the three global adaptation methods as the second or the third correction methods.

The idea of static %1 contrast stretching is prone to produce reddish artifacts or sometimes supplant the green water pixels with blue water pixels. This confinement brings up the idea of dynamic contrast stretching such as with variable stretching tolerance.

Likewise, the ample green and blue pixels in a scene may occasionally turn the luminance stretched image into a more displeasing greenish or bluish scene if it is applied as the initial color correction method. On the other hand, luminance stretching mostly confronts the red spectrum with an indispensable zenith (sufficiently large bin but not the peak) at zero while one or two zeniths may spawn at each end of green and/or blue spectrums (introducing wide range of pixel exposures) provided that it is not applied as the first color correction operation.

Saturation stretching confronts the red spectrum's first quadrant (Q1) whether supplanted with an indispensable non-vain Q1 or turned down to the lower values due to imbalance between RGB spectrums, therefore, confining its usage as an initial color correction.

Saturation stretching is also prone to introduce darker colors or turn the blue underwater sunlight into bright white sunlight when it is applied as the second or third operation.

Adaptive Gamma correction supplants the initial repressed red spectrum (having large Q1 cdf and low mean) with a striking wider and turned up spectrum whose peak has been settled between Q2 and Q3. This method may occasionally create displeasing reddish artifacts.

The veracity of green and red spectrums' resemblance is frequently seen in different environments and scenes by visual inspections of image histogram. Therefore, an engrossing tentative solution for red channel compensation is to convey the green spectrum's distribution to red channel.

Data availability

The presented work is open access and is available at: <https://github.com/vahidr213/Underwater-Image-Restoration-And-Enhancement-Collection/18.00>. The UIEB dataset is downloadable from: https://li-chongyi.github.io/proj_benchmark.html.

Received: 22 January 2022; Accepted: 25 April 2022

Published online: 27 May 2022

References

1. Akkaynak, D. and T. Treibitz. *A revised underwater image formation model*. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. (2018).
2. Drews, P. L. et al. Underwater depth estimation and image restoration based on single images. *IEEE Comput. Graphics Appl.* **36**(2), 24–35 (2016).
3. He, K., Sun, J. & Tang, X. Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **33**(12), 2341–2353 (2010).
4. Bryson, M. et al. True color correction of autonomous underwater vehicle imagery. *J. Field Robot.* **33**(6), 853–874 (2016).
5. Akkaynak, D. and T. Treibitz. *Sea-thru: A method for removing water from underwater images*. in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. (2019).
6. Galdran, A. et al. Automatic red-channel underwater image restoration. *J. Vis. Commun. Image Represent.* **26**, 132–145 (2015).
7. Jiao, Q. et al. Underwater image restoration via non-convex non-smooth variation and thermal exchange optimization. *J. Marine Sci. Eng.* **9**(6), 570 (2021).
8. Zhou, J. et al. Underwater image restoration via depth map and illumination estimation based on a single image. *Opt. Express* **29**(19), 29864–29886 (2021).
9. Zhou, J. et al. Underwater image restoration via feature priors to estimate background light and optimized transmission map. *Opt. Express* **29**(18), 28228–28245 (2021).
10. Ancuti, C., et al. *Enhancing underwater images and videos by fusion*. in *2012 IEEE conference on computer vision and pattern recognition*. (2012). IEEE.
11. Zhang, M. & Peng, J. Underwater image restoration based on a new underwater image formation model. *IEEE Access* **6**, 58634–58644 (2018).
12. Sun, X. et al. Deep pixel-to-pixel network for underwater image enhancement and restoration. *IET Image Proc.* **13**(3), 469–474 (2019).
13. Uplavikar, P.M., Z. Wu, and Z. Wang. *All-in-One Underwater Image Enhancement Using Domain-Adversarial Learning*. in *CVPR Workshops*. (2019).
14. Fabbri, C., M.J. Islam, and J. Sattar. *Enhancing underwater imagery using generative adversarial networks*. in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. (2018). IEEE.
15. Anwar, S., C. Li, and F. Porikli, *Deep underwater image enhancement*. arXiv preprint <http://arxiv.org/abs/1807.03528>, (2018).
16. Guo, Y., Li, H. & Zhuang, P. Underwater image enhancement using a multiscale dense generative adversarial network. *IEEE J. Oceanic Eng.* **45**(3), 862–870 (2019).
17. Ioffe, S. and C. Szegedy. *Batch normalization: Accelerating deep network training by reducing internal covariate shift*. in *International conference on machine learning*. (2015). PMLR.
18. Hou, M., et al. *Joint residual learning for underwater image enhancement*. in *2018 25th IEEE International Conference on Image Processing (ICIP)*. (2018). IEEE.

19. Cao, K., Y.-T. Peng, and P.C. Cosman. *Underwater image restoration using deep networks to estimate background light and scene depth*. in *2018 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI)*. (2018). IEEE.
20. Li, J. *et al.* WaterGAN: Unsupervised generative network to enable real-time color correction of monocular underwater images. *IEEE Robot. Automat. Lett.* **3**(1), 387–394 (2017).
21. Lu, J. *et al.* Multi-scale adversarial network for underwater image restoration. *Opt. Laser Technol.* **110**, 105–113 (2019).
22. Ye, X., *et al.* *Underwater image enhancement using stacked generative adversarial networks*. in *Pacific Rim Conference on Multimedia*. (2018). Springer.
23. Wang, Y. *et al.* UIEC²-Net: CNN-based underwater image enhancement using two color space. *Signal Process. Image Commun.* **96**, 116250 (2021).
24. Zhou, J.-C., Zhang, D.-H. & Zhang, W.-S. Classical and state-of-the-art approaches for underwater image defogging: A comprehensive survey. *Front. Inf. Technol. Electron. Eng.* **21**(12), 1745–1769 (2020).
25. Ebner, M. *Color Constancy* (Wiley, 2007).
26. Zimmerman, J. B. *et al.* An evaluation of the effectiveness of adaptive histogram equalization for contrast enhancement. *IEEE Trans. Med. Imaging* **7**(4), 304–312 (1988).
27. Achanta, R., *et al.* *Frequency-tuned salient region detection*. in *2009 IEEE Conference On Computer Vision and Pattern Recognition*. (2009). IEEE.
28. Drago, F., Myszkowski, K., Annen, T. & Chiba, N. Adaptive logarithmic mapping for displaying high contrast scenes. *Comput. Graph. Forum* <https://doi.org/10.1111/1467-8659.00689> (2003).
29. Land, E. H. & McCann, J. J. Lightness and retinex theory. *Josa* **61**(1), 1–11 (1971).
30. Rahman, S. *et al.* An adaptive gamma correction for image enhancement. *EURASIP J. Image Video Process.* **2016**(1), 1–13 (2016).
31. Dong, X., *et al.* *Fast efficient algorithm for enhancement of low lighting video*. in *2011 IEEE International Conference on Multimedia and Expo*. (2011). IEEE.
32. Breiman, L. Bagging predictors. *Mach. Learn.* **24**(2), 123–140 (1996).
33. Breiman, L. Random forests. *Mach. Learn.* **45**(1), 5–32 (2001).
34. Li, C. *et al.* An underwater image enhancement benchmark dataset and beyond. *IEEE Trans. Image Process.* **29**, 4376–4389 (2019).
35. Anwar, S. & Li, C. Diving deeper into underwater image enhancement: A survey. *Signal Process. Image Commun.* **89**, 115978 (2020).
36. Fu, X., *et al.* *Two-step approach for single underwater image enhancement*. in *2017 international symposium on intelligent signal processing and communication systems (ISPACS)*. (2017). IEEE.
37. Fu, X., *et al.* *A retinex-based enhancing approach for single underwater image*. in *2014 IEEE international conference on image processing (ICIP)*. (2014). IEEE.
38. Li, C. *et al.* A hybrid method for underwater image correction. *Pattern Recogn. Lett.* **94**, 62–67 (2017).
39. Peng, Y.-T., Cao, K. & Cosman, P. C. Generalization of the dark channel prior for single image restoration. *IEEE Trans. Image Process.* **27**(6), 2856–2868 (2018).
40. Li, C.-Y. *et al.* Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior. *IEEE Trans. Image Process.* **25**(12), 5664–5677 (2016).
41. Peng, Y.-T. & Cosman, P. C. Underwater image restoration based on image blurriness and light absorption. *IEEE Trans. Image Process.* **26**(4), 1579–1594 (2017).
42. Li, C., Guo, J. & Guo, C. Emerging from water: Underwater image color correction based on weakly supervised color transfer. *IEEE Signal Process. Lett.* **25**(3), 323–327 (2018).

Author contributions

The corresponding author of the presented paper is responsible for the entire process including the source code, and manuscript. The figures and images are from our computations (spectrums) and also the unsplash.com (images).

Competing interests

The author declares no competing interests.

Additional information

Correspondence and requests for materials should be addressed to V.R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022