Underwater Image Restoration with a Complementary Haar Wavelet Transform Restoration and an Ensemble of Triple Elite Correction Algorithms using Bootstrap Aggregation and Random Forests

Vahid Rowghanian (✉ v.rowghanian@gmail.com)

Research Article

Keywords: Underwater Image Restoration, Haar wavelet transform, Bootstrap Aggregation, Random Forests

Posted Date: January 24th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1286015/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Abstract

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 elite correction algorithm labels based on histogram quadrants' cumulative distribution feature instead of generating pixel intensities. The Haar wavelet transform is applied on the input image and its contrast stretched image to generate the degraded wavelet coefficients which are refined using Gaussian pyramid of the saliency weights to restore the original image. The ensemble of triple corrections exerts three elite color correction algorithms sequentially on the degraded image for restoration. The ensemble of algorithms covers 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 (DCP), fusion restoration, and our Haar wavelet transform restoration. The source is available at <https://github.com/vahidr213/Underwater-Image-Restoration-And-Enhancement-Collection>

1. Introduction

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.

In the past decade, machine vision has become universal in multiple disciplines into where the integration of automatic inspection and analysis of images is needed. Object detection [1], recognition [2] and tracking [3] are some of well-known fields of machine vision which are recently extending their dominance into the aquatic world.

The primary factor in underwater image degradation is the high attenuation of the radiated light from the surface of the objects [4]. 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) [5] is derived from the DCP [6] 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 zone [7].

The Sea-thru method [8] 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 prior [9] is a rearrangement of the original DCP equation as well as refining the estimated transmission map with guided filter instead of matting technique [10]. Recently, the transmission map of the DCP method for the red channel have been estimated [11] 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 [12], then the refined red channel is recombined with green and blue channels to form the restored image.

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 environments [4].

Other strategies such as multi-scale fusion has drawn much attention in recent years. The method proposed by Ancuti et al [13] 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 proposed [14] 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 [15]. An encoding-decoding convolutional neural network [16] 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 [17] network. They have accelerated the training process of the CNN as well as including low-level features by using skip connection [18], [19] 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 learning [21], generative adversarial network (GAN) [20], underwater convolutional NN [22], multi-scale dense block GAN [23] using batch normalization [24], underwater joint residual learning CNN [25], underwater image restoration network [26], WaterGAN [27], fusionGAN [28], MCycleGAN [29], underwater image enhancement with stacked conditional GAN [30].

The above mentioned networks still reflect a considerable rate of failure due to lack of handling the saturation and contrast problems. In [31], 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).

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 Elite Correction Algorithms (ETCA) and finally the experimental results and comparison with other state of the art algorithms.

2. Proposed Method

In this paper, our contribution to the problem of underwater image restoration includes a traditional complementary 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 refinement of input image's Haar wavelet coefficients. Our complementary approach should be applied after the multi-scale fusion restoration to further improve the quality of the restored image. The flowchart of this approach is shown in Figure 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 input feature instead of using pixel intensity. The input features are learned using Bootstrap Aggregation and Random Forests. We have trained three models separately. The responses of models are also not in the form 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.

2.1. 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 [32] 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:

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

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.

2.1.1. Haar wavelet inputs

Our proposed method uses Haar wavelet transform which is fed with the raw input image as well as a contrast stretched version of the raw input mentioned above. The input image for Haar transform is resized into 640x640 in order to have the Haar pyramid dimensions as equal as the Gaussian pyramid of the 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 1024x1024 to increase the accuracy.

2.1.2. Multi-scale fusion input

The first image for the pyramid of multi-scale fusion [13] is the contrast stretched version of the input image explained above and the second one is an equalized luminance image provided by adaptive histogram equalization [33].

2.2. Weights of Multi-Scale Fusion

The choice of correct weights reduces computations and artifacts. There are four weights in [13] that are explained below:

2.2.1. Laplacian contrast weight (W_L)

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

2.2.2. 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. 5x5, $[1, 4, 6, 4, 1]/16$ filter plus cutting off values above $\pi/2.75$ to $\pi/2.75$).

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

2

2.2.3. 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 [34] is used which is defined in Lab color space by:

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

3

2.2.4. 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:

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

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:

$$W_1 = \frac{W_{L_1} + W_{LC_1} + W_{S_1} + W_{E_1}}{W_{L_1} + W_{LC_1} + W_{S_1} + W_{E_1} + W_{L_2} + W_{LC_2} + W_{S_2} + W_{E_2}}$$

5

$$W_2 = \frac{W_{L_2} + W_{LC_2} + W_{S_2} + W_{E_2}}{W_{L_1} + W_{LC_1} + W_{S_1} + W_{E_1} + W_{L_2} + W_{LC_2} + W_{S_2} + W_{E_2}}$$

6

2.3. Haar Wavelet Coefficient Refinement

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. Two images are derived from the initial degraded image (I_1, I_2). Therefore, each of which can generate slightly different type of wavelet coefficients which still carry degradation information. In this

paper, we propose a new method to refine or update the degraded Haar wavelet coefficients $\{A, H, V, D\}$ 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 updating weights.

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

$$W_1 = \frac{W_{S_1}}{W_{S_1} + W_{S_2}}$$

7

$$W_2 = \frac{W_{S_2}}{W_{S_1} + W_{S_2}}$$

8

The refinement or updating of the degraded Haar wavelet coefficients are performed according to the following equations:

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

9

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

10

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

11

$$A_{new} = \sum_{k=1}^K A^k$$

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 ($A_{\text{new}}, H_{\text{new}}, V_{\text{new}}, D_{\text{new}}$) to restore the original image.

The Haar wavelet restoration is independent and capable of restoring underwater images solely. 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 fusion. Figure 1 shows the flowchart of the Haar wavelet and multi-scale fusion 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.

2.4. Multi Scale Fusion

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

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

13

Where $I^k (M \times N)$ is the input image ($k=1,2,...,K$) with M rows and N columns, $W^{-k} (M \times N)$ is the normalized weights ($k=1,2,...,K$) ($\sum W^{-k} = 1$) 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:

$$R^l(x, y) = \sum_{k=1}^K \left\{ W^{-k}(x, y) \right\}^l \left\{ I^k(x, y) \right\}^l$$

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:

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

15

Where $L^l \left\{ I^k(x, y) \right\}$ is the Laplacian of the k th image at level l , and $G^l \left\{ W^k(x, y) \right\}$ is the Gaussian of the

k th weight at level l . The Gaussian pyramid is built by Gaussian filtering (or 5x5, [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.

3. Ensemble Of Triple Elite Correction

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 regard, we have trained three models. The training method is Bootstrap Aggregation and Random Forests. Each model is responsible for generating a numeric label that corresponds to a specific color correction method among the fifteen available methods. In this paper, we have also proposed new features to be learned which are explained in detail in the feature extraction section. The ensemble of color correction methods are 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, DCP [6], fusion restoration [13], and our Haar wavelet transform restoration.

It should be noted that the triple elite corrections are applied sequentially as the block diagram of the proposed method ETCA is showing in Figure 3.

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 collection of images which had a negligible difference with their reference image (negligible mean square error). Other color correction algorithms are labeled with a special number identifying the corresponding operation. The color correction methods are described in the feature extraction section in detail.

3.1. 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.

$$F_I^c(\Omega_1) = f^c(0 \leq I \leq \Omega_1) \quad (16)$$

$$F_I^c(\Omega_2) = f^c(\Omega_1 < I \leq \Omega_2) \quad (17)$$

$$F_I^c(\Omega_3) = f^c(\Omega_2 < I \leq \Omega_3) \quad (18)$$

$$F_I^c(\Omega_4) = f^c(\Omega_3 < I \leq \Omega_4) \quad (19)$$

Where F_I is the cumulative distribution function for the intensity I , 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 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 number of fifteen color correction methods used in evaluations. One of them is no action. The no action is labeled with zero showing that the input image requires no color correction. This was necessary due to existence of some fully enhanced images in the UIEB dataset. Hereafter, other color correction methods are described in detail.

- The contrast stretching in RGB color space, luminance stretching in Lab color space, and saturation stretching in HSV color space are performed simply by using the following equation which saturates the bottom 1% and the top 1% of all pixel values (with proper L and H).

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

20

Where L and H are the intensities corresponding to bottom 1% and top 1% of the image I .

- 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 law [35] and the retinex theory [36]. The global adaptation of an image I is expressed in by:

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

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:

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

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 are brightened more than brighter scenes through global adaptation.

- Adaptive Gamma correction is performed using the method proposed in [37] 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 .

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

23

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

24

$$I_{out} = c I_{in}^\gamma$$

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.

- 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:

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

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.

- 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 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 [8-250]. Other intensities are left intact. This method partially affects the image and has no extensive positive or negative effect on the image.
- 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:

$$\rho_{red} = \rho_{green}$$

27

Where ρ_{green} is the green spectrum which is used in histogram equalization.

- The local brightening [38] 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 [39]-[40] 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.

4. Experimental Results

A comprehensive and fair experiment of several traditional underwater image enhancement methods is performed in UIEB dataset paper [41] from which our raw and reference images are drawn. Likewise, a comprehensive evaluation of several recent learned models (deep learning) is performed in [49] from which we quote them in Table 1. The UIEB dataset [41] 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 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).

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

The speed performance depends on the image size and the methods applied on the image. 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 fusion [13] and our proposed Haar wavelet restoration go along with each other inside the ensemble of triple algorithms.

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 Table1. 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 you can see in Table 1, the best performance among all the 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.

Table 1. quotes the performance of our methods as well as the performance of other methods according to [41], [49].

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 images are well restored and only in very few cases some reddish artifacts are produced. An example of such case is shown in Figure 4. This situation is due to the type of the cost function. Since the criterion of our cost function relies only on the mean square error (MSE), the small possibility of finding a weak minimum is not avoidable.

An example of no operation is shown in Figure 5. In Figure 6 we compare our results with the reference images presented in UIEB.

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).

Table 1
FULL-REFERENCE IMAGE QUALITY EVALUATION

Method	MSE ↓	PSNR(dB) ↑	SSIM ↑
Original	1768.9	17.36	0.6168
Fusion-based [13]	867.9	18.7461	0.8162
Two-step-based [42]	1114.6	17.6596	0.7199
Retinex-based [43]	1353.1	16.8757	0.6233
UDCP [5]	5130.0	11.0296	0.4999
Regression-based [46]	1136.5	17.5751	0.6543
GDCP [47]	3634.5	12.5264	0.5503
Red Channel [9]	2107.3	14.8935	0.5973
Histogram prior [44]	1628.2	16.0137	0.5888
Blurriness-based [45]	1582.6	16.1371	0.6582
MCycleGAN [29]	1132.2	18.33	0.6138
URCNN [25]	2195.8	15.94	0.5972
UWGAN [48]	1853.7	16.06	0.2945
DUIENet [41]	1012.2	19.29	0.8093
DenseGAN [23]	1363.6	17.56	0.4239
UWCNN type-I [22]	2345.0	15.00	0.5306
UWCNN type-III [22]	2920.2	14.24	0.4945
Dive+	535.8	20.8408	0.8705
Our([13]+Haar)	745.3	22.1079	0.8105
Our(ETCA)	210.4	34.9780	0.8956

Declarations

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

References

1. Zou, Z., et al., *Object detection in 20 years: A survey*. arXiv preprint arXiv:1905.05055, 2019.
2. Szegedy, C., et al. *Going deeper with convolutions*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.
3. Rowghanian, V. and K. Ansari-Asl, *Object tracking by mean shift and radial basis function neural networks*. *Journal of Real-Time Image Processing*, 2018. **15**(4): p. 799–816.
4. Akkaynak, D. and T. Treibitz. *A revised underwater image formation model*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.
5. Drews, P.L., et al., *Underwater depth estimation and image restoration based on single images*. *IEEE computer graphics and applications*, 2016. **36**(2): p. 24–35.
6. He, K., J. Sun, and X. Tang, *Single image haze removal using dark channel prior*. *IEEE transactions on pattern analysis and machine intelligence*, 2010. **33**(12): p. 2341–2353.
7. Bryson, M., et al., *True color correction of autonomous underwater vehicle imagery*. *Journal of Field Robotics*, 2016. **33**(6): p. 853–874.
8. 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.
9. Galdran, A., et al., *Automatic red-channel underwater image restoration*. *Journal of Visual Communication and Image Representation*, 2015. **26**: p. 132–145.
10. Levin, A., D. Lischinski, and Y. Weiss, *A closed-form solution to natural image matting*. *IEEE transactions on pattern analysis and machine intelligence*, 2007. **30**(2): p. 228–242.
11. Jiao, Q., et al., *Underwater Image Restoration via Non-Convex Non-Smooth Variation and Thermal Exchange Optimization*. *Journal of Marine Science and Engineering*, 2021. **9**(6): p. 570.
12. Kaveh, A. and A. Dadras, *A novel meta-heuristic optimization algorithm: thermal exchange optimization*. *Advances in Engineering Software*, 2017. **110**: p. 69–84.
13. Ancuti, C., et al. *Enhancing underwater images and videos by fusion*. in *2012 IEEE conference on computer vision and pattern recognition*. 2012. IEEE.
14. Zhang, M. and J. Peng, *Underwater image restoration based on a new underwater image formation model*. *IEEE Access*, 2018. **6**: p. 58634–58644.

15. Nebauer, C., *Evaluation of convolutional neural networks for visual recognition*. IEEE transactions on neural networks, 1998. **9**(4): p. 685–696.
16. Sun, X., et al., *Deep pixel-to-pixel network for underwater image enhancement and restoration*. IET Image Processing, 2019. **13**(3): p. 469–474.
17. Krizhevsky, A., I. Sutskever, and G.E. Hinton, *Imagenet classification with deep convolutional neural networks*. Advances in neural information processing systems, 2012. **25**: p. 1097–1105.
18. Long, J., E. Shelhamer, and T. Darrell. *Fully convolutional networks for semantic segmentation*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.
19. Hariharan, B., et al. *Hypercolumns for object segmentation and fine-grained localization*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.
20. 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.
21. Uplavikar, P.M., Z. Wu, and Z. Wang. *All-in-One Underwater Image Enhancement Using Domain-Adversarial Learning*. in *CVPR Workshops*. 2019.
22. Anwar, S., C. Li, and F. Porikli, *Deep underwater image enhancement*. arXiv preprint arXiv:1807.03528, 2018.
23. Guo, Y., H. Li, and P. Zhuang, *Underwater image enhancement using a multiscale dense generative adversarial network*. IEEE Journal of Oceanic Engineering, 2019. **45**(3): p. 862–870.
24. Ioffe, S. and C. Szegedy. *Batch normalization: Accelerating deep network training by reducing internal covariate shift*. in *International conference on machine learning*. 2015. PMLR.
25. Hou, M., et al. *Joint residual learning for underwater image enhancement*. in *2018 25th IEEE International Conference on Image Processing (ICIP)*. 2018. IEEE.
26. 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.
27. Li, J., et al., *WaterGAN: Unsupervised generative network to enable real-time color correction of monocular underwater images*. IEEE Robotics and Automation letters, 2017. **3**(1): p. 387–394.
28. Li, C. and M. Wand. *Precomputed real-time texture synthesis with markovian generative adversarial networks*. in *European conference on computer vision*. 2016. Springer.
29. Lu, J., et al., *Multi-scale adversarial network for underwater image restoration*. Optics & Laser Technology, 2019. **110**: p. 105–113.
30. Ye, X., et al. *Underwater image enhancement using stacked generative adversarial networks*. in *Pacific Rim Conference on Multimedia*. 2018. Springer.
31. Wang, Y., et al., *UIEC²-Net: CNN-based underwater image enhancement using two color space*. Signal Processing: Image Communication, 2021. **96**: p. 116250.
32. Ebner, M., *Color constancy*. Vol. 7. 2007: John Wiley & Sons.

33. Zimmerman, J.B., et al., *An evaluation of the effectiveness of adaptive histogram equalization for contrast enhancement*. IEEE Transactions on Medical Imaging, 1988. **7**(4): p. 304–312.
34. Achanta, R., et al. *Frequency-tuned salient region detection*. in *2009 IEEE conference on computer vision and pattern recognition*. 2009. IEEE.
35. Drago, F., et al. *Adaptive logarithmic mapping for displaying high contrast scenes*. in *Computer graphics forum*. 2003. Wiley Online Library.
36. Land, E.H. and J.J. McCann, *Lightness and retinex theory*. Josa, 1971. **61**(1): p. 1–11.
37. Rahman, S., et al., *An adaptive gamma correction for image enhancement*. EURASIP Journal on Image and Video Processing, 2016. **2016**(1): p. 1–13.
38. Dong, X., et al. *Fast efficient algorithm for enhancement of low lighting video*. in *2011 IEEE International Conference on Multimedia and Expo*. 2011. IEEE.
39. Breiman, L., *Bagging predictors*. Machine learning, 1996. **24**(2): p. 123–140.
40. Breiman, L., *Random forests*. Machine learning, 2001. **45**(1): p. 5–32.
41. Li, C., et al., *An underwater image enhancement benchmark dataset and beyond*. IEEE Transactions on Image Processing, 2019. **29**: p. 4376–4389.
42. 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.
43. Fu, X., et al. *A retinex-based enhancing approach for single underwater image*. in *2014 IEEE international conference on image processing (ICIP)*. 2014. IEEE.
44. Li, C.-Y., et al., *Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior*. IEEE Transactions on Image Processing, 2016. **25**(12): p. 5664–5677.
45. Peng, Y.-T. and P.C. Cosman, *Underwater image restoration based on image blurriness and light absorption*. IEEE transactions on image processing, 2017. **26**(4): p. 1579–1594.
46. Li, C., et al., *A hybrid method for underwater image correction*. Pattern Recognition Letters, 2017. **94**: p. 62–67.
47. Peng, Y.-T., K. Cao, and P.C. Cosman, *Generalization of the dark channel prior for single image restoration*. IEEE Transactions on Image Processing, 2018. **27**(6): p. 2856–2868.
48. Li, C., J. Guo, and C. Guo, *Emerging from water: Underwater image color correction based on weakly supervised color transfer*. IEEE Signal processing letters, 2018. **25**(3): p. 323–327.
49. Anwar, S. and C. Li, *Diving deeper into underwater image enhancement: A survey*. Signal Processing: Image Communication, 2020. **89**: p. 115978.

Figures

Figure 1

The flowchart of the complementary Haar wavelet transform for underwater image restoration

Figure 2

The block diagram for illustrating the Haar wavelet coefficient refinement

Figure 3

The block diagram illustrating the flow chart of the Ensemble of Triple Elite Correction. M_1 , M_2 , and M_3 are the trained models generating a numeric label for C_1 , C_2 , and C_3 sequential color correction blocks respectively.

Figure 4

An example of our method's failure in finding the best three color correction algorithms. The red artifacts are introduced in the image due to weak ensemble of algorithms.

Figure 5

An example of output where our proposed method has applied no color correction on the input image.

Figure 6

The comparison between the restored images of ETCA in the left vs the reference image