Multi-scale fusion for underwater image enhancement using multi-layer perceptron

M. Sudhakara, M. Janaki Meena
Department of SCOPE, VIT University, Chennai, India

ABSTRACT

Underwater image enhancement (UIE) is an imperative computer vision activity with many applications and different strategies proposed in recent years. Underwater images are firmly low in quality by a mixture of noise, wavelength dependency, and light attenuation. This paper depicts an effective strategy to improve the quality of degraded underwater images. Existing methods for dehazing in the literature considering dark channel prior utilize two separate phases for evaluating the transmission map (i.e., transmission estimation and transmission refinement). Accurate restoration is not possible with these methods and takes more computational time. A proposed three-step method is an imaging approach that does not need particular hardware or underwater conditions. First, we utilize the multi-layer perceptron (MLP) to comprehensively evaluate transmission maps by base channel, followed by contrast enhancement. Furthermore, a gamma-adjusted version of the MLP recovered image is derived. Finally, the multi-scale fusion method was applied to two attained images. The standardized weight is computed for the two images with three different weights in the fusion process. The quantitative results show that significantly our approach gives the better result with the difference of 0.536, 2.185, and 1.272 for PCQI, UCIQE, and UIQM metrics, respectively, on a single underwater image benchmark dataset. The qualitative results also give better results compared with the state-of-the-art techniques.

Keywords:
Gamma correction
Image dehazing
Image fusion
Multi-layer perceptron
Multi-scale fusion

Corresponding Author:
M. Sudhakara
Department of SCOPE
VIT University, Chennai Campus, India
Email: malla.sudhakara2015@vit.ac.in

1. INTRODUCTION

Underwater image enhancement in the recent decade attained great popularity in processing images and underwater vision [1]. It is hard to enhance an underwater image because of the typical underwater conditions and lighting effects [2], [3]. The color distortion influences underwater situations, and contrast degrades because of absorption [4]. Many researchers have investigated underwater images to attain the best quality. The methods for the improvement of underwater images are classified into three categories. The first category applies the image degradation model [5] to enhance an underwater image by exploring the physical process of natural light underwater to build an underwater image degradation pattern. The second category improves image contrast by utilizing thresholding and histogram equalization to map the grey values into a uniform distribution for adjusting the bright or dark portions. There exist various enhanced methods based on histogram equalization [6]. Ancuti et al. [7], [8] focused fusion mechanism on underwater image and video, which comes under the third category.

Journal homepage: http://ijai.iaescore.com
Dehazing strategies, dark channel priors (DCP) [9], [10] are applied to improve quality. In [11], the authors presented a model to increase underwater image quality, and DCP helps estimate the transmission. A red channel model [12] recovers minimum wavelength colors during the restoration of skipped contrast. In [13], underwater DCP was proposed for image restoration. In [14], contrast limited adaptive histogram equalization is utilized to improve the picture quality at some rate, and the histogram of the preferred portion relatively matches the particular histogram. For multiple image development modes, a technique requires two polarization images with varying degrees [15], and the other approach uses many images to increase the underwater image clarity. A fusion-based mechanism [16] derived from adding multiple filters over input, and a fusion mechanism [17] applies the techniques such as white balancing and global contrast stretching for UIE. In [18], the authors proposed another fusion-based method to increase contrast and color, including the inclusion of two images from the color compensated and white-adjusted adoption of the original image. A two-phase method [19] proposes a colour correction mechanism of piece-wise transformation to restore the color. Afterwards, an ideal technique is utilized for contrast improvement. Drews et al. [20] developed upon DCP and considered it a significant source of visual data underwater in a blue color channel and green color channel. The underwater DCP of those two channels estimates better transmission of underwater images than the traditional DCP. Emberton et al. [21] developed a hierarchical rank-based technique by considering some features to detect image portions with the right quality image. The further fusion of these two HR images took place to get the final result. In [22], the authors proposed a Water-Net network, training on the UIEB data with convolution neural networks (CNNs). The fusion-based approaches give a better output than the existing dehazing techniques by utilizing a few weight maps to the low-quality images [23].

The primary objective of this work is to increase the quality of the underwater images. A wide variety of strategies are available for the mentioned requirement, each with its inherent advantages and disadvantages. Current state-of-the-art techniques utilize the dark channel prior based methods and many other methods using multiple images for processing. These techniques require a lot of computational time with less accurate restoration. In this regard, the authors have perceived a strong need to develop a suitable imaging methodology underwater. This paper introduces a novel procedure to eliminate the haze present in the underwater images captured by conventional cameras. The blueprint of the proposed work has appeared in Figure 1. The crucial steps of our proposed methodology are described below.

- A refined transmission map is estimated directly from the minimum channel to recover the image, and the multi-layer perceptron was used for training.
- Two variants of the recovered image with improving quality are produced by applying the contrast stretching and gamma correction algorithms.
- Multi-scale image fusion is applied to these two images alongside their weight maps to promote the transfer of the edges and colour contrast to the output image.
- The dominant colours are balanced by adapting a colour correction algorithm to the fused image.

![Figure 1. Proposed outline for UIE](image)

Terminology
The background concepts related to the underwater image, such as scattering and multi-layer perceptron, are discussed in this section.

- Scattering model
  An atmospheric scattering model is a mathematical model developed around the crude dynamical conditions that oversee underwater climatic movements-the atmospheric scattering model by [24] for image formation given in (1).

\[
I(p,q) = J(p,q) \ast t(p,q) + A[1-t(p,q)]
\]

(1)
Here, 'I' is the original image, 'J', 'A','t' are the real scene without impact, environmental light and transmission map separately. The transmission map is assessed as (2).

\[ t(p, q) = e^{-\beta d(p, q)} \]  

In (2), the 'β' is the scattering coefficient, and d (p, q) is scene depth. As shown in (1), modified as (3) to acquire a haze-free image.

\[ J(p, q) = I(p, q) - A_t(p, q) + A \]

It is recognizable that in (3), the original image contains the real scene, transmission map, and atmospheric light. The job here is recovering J (p, q) from the original image by eliminating the transmission and atmospheric light. The intensity of the dark channel proposed by DCP for a given image I (p, q) is (4).

\[ I_{dark}(p, q) = \min_{c \in \{R, G, B\}} \min_{z \in \Omega(p, q)} I_c(z) \]

Here Ω (p, q) is a part of m X m, centered over (p, q), IC p,q the colour of the channel of the image, and ‘z’ denotes pixel index in Ω(p, q) domain. \( \hat{t} (p, q) \) can be computed as (5).

\[ \hat{t}(p, q) = 1 - \omega \left( \frac{I_{dark}(p, q)}{A} \right) \]

In (5), 'ω' defines the estimated restoration levels, while 'I' is the maximum possible value. Here, ‘A’ is computed as (6).

\[ A = \max \sum_{c=1}^{3} I_c \left( \min \arg \max_{(p, q) \in (0.1 \times h \times w)} \left\{ I_{dark}(p, q) \right\} \right) \]

In the above equation, ‘w’ is the width, ‘h’ is the height of \( I_{dark}(p, q) \) and ‘C’ represents every RGB channel of an input image I (p, q). As shown in (3) helps in estimating an image with no affection \( \hat{J}(p, q) \). The utilization of windows of mXm in the computation of \( \hat{t} (p, q) \) creates artefacts over edges of the recovered image J(p, q).

Multi-layer perceptron

Artificial neural network (ANN) is a machine learning approach, a profoundly simplified scientific model from biological neural networks. The connection between the processing elements (neurons) is called a network. The inputs \( X_i \) multiplied by a weight \( W_{ij} \), and the bias \( b_j \) added to develop the neuron. This function illustrated as (7).

\[ Y_j = f \left( \sum_{i=1}^{3} W_{ij} X_i + b_j \right) \]  

To activate the perceptron, the activation function is utilized. In many networks, the tangent function is used for activation, and it is given as in (8).

\[ \tanh(x) = 2 \cdot \sigma(2x) - 1 \]  

where

\[ \sigma(x) = \frac{e^x}{1 + e^x} \]

As shown in (9) is the sigmoid activation function [25]. The preparation of ANN changes the weight and bias for getting the ideal outcome dependent on the different input combinations. Different perceptrons in one or many hidden layer topologies are connected; this system is known as MLP. Backpropagation is a well-known strategy for MLP, and it is a generalization of the least-squares component where the loads are altered dependent on mean square error (MSE). The idea is finding the smallest error in relationship with the associated weights and is given in (10).

\[ E = \frac{1}{2} \sum_{j \in M} (\hat{y}_m - y_m)^2 \]
2. PROPOSED METHODOLOGY

The proposed work is mainly a three-step process. First, the input image is acquired from the underwater. Second, an MLP is utilized to recover the marine image. For the MLP retrieved image, we used two contrast enhancement techniques to improve the contrast (i.e., contrast stretching, gamma correction). Before applying the contrast stretching, the RGB version of the MLP recovered image is altered to the L*a*b model, and an algorithm is used on the luminance channel of the L*a*b image. Later multi-scale fusion is applied to the derived images to acquire the result. Further, simple colour balancing is used to balance the colour before reaching the final output image.

The process of MLP on the given input image is represented in Figure 2. For a given input image, air-light is estimated. It is nothing but applying the morphological erosion operation on the input image and obtaining the dark channel map. The methods belong under the DCP mechanism yielded good results when compared with existing models. As per DCP, the transmission map is estimated under two phases. Initially, a coarse transmission map is estimated through square-patches with the user-defined size. The transmission map is enhanced with a different filter and soft matting, and median filter in the next step. According to [26], the transmission map is updated, as shown in (11), (12).

\[ t_{\min}(p,q) = 1 - \omega I_{\min}(p,q) \]  
\[ I_{\min}(p,q) = \min_{c \in (R,G,B)} I_c(p,q) \]  

The \( t_{\min} \) is equal to the dark channel only when \( \Omega = 1 \). \( t_{\min}(p,q) \) has a decent resolution. MLP for estimating a transmission map \( t'(p,q) \) is (13).

\[ t'(p,q) = MLP[t_{\min}(p,q)] \]  

![Figure 2. Dehazing process using MLP](image)

The primary thought of this strategy is estimating the refined transmission map. The recovered image is converted from RGB to L*a*b format, and the contrast of the luminance is increased. Once air-light is estimated, we estimated the transmission map using the input image, air-light. Further, we applied a multi-layer perceptron on the transmission estimated image to get the MLP recovered image. Then two images are derived from the MLP improved model. The first input image is a gamma-corrected version of the MLP recovered image. Gamma correction expects to improve the global contrast in the picture, and it is appropriate to our work since many marine photos look too bright. Hence, the correction will increase the contrast between the darker regions by reducing the more elegant areas. As mentioned earlier, contrast stretching is applied to the MLP recovered image on the Luminance channel to get the second image for the fusion process. Then we utilized the multi-scale fusion to get a solution for the single underwater image. The weight maps are calculated based on the saliency metrics in such a way that the highest weights are represented in the image (i.e., Laplacian weight (\( W_L \)), saliency weight (\( W_S \)), and saturation weight (\( W_{sat} \))). Subsequently, Laplacian Filter is applied to each luminance channel, and \( W_L \) estimates the global contrast by computing the total value. This weight allocates high costs to edges, and textures. This weight is not satisfactory because it cannot differentiate among the different regions (i.e., ramp, flat). \( W_L \) highlights the objects that miss their importance in the input image and is derived from [27]. \( W_{sat} \) adopts the chromatic information from the image, which is advantageous to the fusion process. Saturation weight is nothing but the deviation between each colour channel in the RGB image and is computed, as shown in (14).

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

For every input, these three weight maps are combined into a single weight map, and the weights are normalized, and finally, a normalized weight map is obtained. With these standardized weight maps, we utilized multi-scale fusion to fuse the derived images. According to this, the Laplacian and Gaussian pyramids are constructed and combining with Laplacian and Gaussian normalized weights achieved independently at every level $l$, as shown in (15).

$$R_i(x) = \sum_{k=1}^{K} G_i \left( \overline{W}_k (x) \right) L_i \left( I_k (x) \right)$$ \hspace{1cm} (15)

Here, 'T' signifies the pyramid levels, and 'k' symbolizes input images. The pyramid levels may be varied depending upon the size of the picture. By combining all the pyramid level images with the proper upsampling, we get the final fused image.

3. RESULTS AND DISCUSSION

The proposed method was evaluated on real-time underwater images used as a benchmark dataset for the single UIE methods. For a multi-scale fusion, the levels depend on the size of the input image. The six underwater images with different backgrounds are taken and applied our multi-layer perceptron based multi-fusion strategy. The input image, recovered image 1, recovered image 2, and the fusion of these dual images are shown in Figure 3. The experimental outcomes are shown using the human visual system (HVS) and automatic evaluation measures. As per the best of our knowledge, generating ground-truth images for underwater images is so difficult. Due to the lack of this ground-truth information, HVS by specialists helps evaluate the quality of the image by incorporating the perceptual metrics. We compared our method with four approaches used for dehazing underwater and outdoor images. We have tested our result with many images, but only four image results are shown in Figure 4. In this work, three quantitative measures are considered (PCQI, UCIQE, and UIQM). Among these metrics, PCQI is a universal metric, and the other two parameters UCIQE, UIQM [28], are the metrics for evaluating the underwater images. PCQI depends on the patch-based model depends on global statistics. It considers three quantities of the image which are not dependent (mean, signal strength, and structure). Mathematically PCQI can be computing using (16).

$$PCQI = Q_i(p, q) * Qc(p, q) * Qs(p, q)$$ \hspace{1cm} (16)

In (16), $Q_i(p, q), Qc(x, y), Qs(x, y)$ are the contrast mean intensity, structural distortion, and contrast change. This metric is computationally less costly compared with other evaluation metrics. UCIQE on three attributes such as Chroma, contrast, and saturation of CIELab and evaluated by using (17).

$$UCIQE = K1 \times \sigma \gamma + K2 \times contrast + K3 \times \mu$$ \hspace{1cm} (17)

‘$\sigma$’ is the standard deviation of Chroma, and ‘$\mu$’ is the mean of the saturation. The acronym of UIQM is an underwater image quality measure and is dissimilar from predefined computation measures. UIQM computed by considering three attributes colorfulness, contrast, and sharpness. It can be evaluated by (18).

$$UIQM = k1 \times UICM + k2 \times UISM + l3 \times UIConM$$ \hspace{1cm} (18)

In (18), $k1$, $k2$, and $k3$ are the application dependent parameters; for example, extra weight must be put on $k1$ for submerged colour correct and $k2$ to increase the perceptibility. The comparison metrics are shown in Table 1, and the parameters are based on (16), (17), and (18). The larger values correspond to the better result.

<table>
<thead>
<tr>
<th>Author/Image</th>
<th>Shipwreck</th>
<th>Fish</th>
<th>Reef1</th>
<th>Reef2</th>
<th>Reef3</th>
<th>Galdran1</th>
<th>Ancuti1</th>
<th>Ancuti2</th>
<th>Ancuti3</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>He, et al. [32] PCQI</td>
<td>1.012</td>
<td>1.023</td>
<td>1.000</td>
<td>0.774</td>
<td>1.022</td>
<td>1.056</td>
<td>0.860</td>
<td>0.649</td>
<td>1.071</td>
<td>0.940</td>
</tr>
<tr>
<td>UIQM</td>
<td>0.565</td>
<td>0.509</td>
<td>0.592</td>
<td>0.749</td>
<td>0.578</td>
<td>0.578</td>
<td>0.353</td>
<td>0.437</td>
<td>0.538</td>
<td>0.480</td>
</tr>
<tr>
<td>Drew-Jr [13] UCIQE</td>
<td>0.649</td>
<td>0.863</td>
<td>1.046</td>
<td>0.483</td>
<td>0.793</td>
<td>0.749</td>
<td>0.909</td>
<td>0.475</td>
<td>0.973</td>
<td>0.771</td>
</tr>
<tr>
<td>UIQM</td>
<td>0.550</td>
<td>0.623</td>
<td>0.649</td>
<td>0.659</td>
<td>0.620</td>
<td>0.544</td>
<td>0.499</td>
<td>0.492</td>
<td>0.535</td>
<td>0.574</td>
</tr>
<tr>
<td>Galdran, et al. [12] UCIQE</td>
<td>0.492</td>
<td>0.571</td>
<td>0.657</td>
<td>0.653</td>
<td>0.584</td>
<td>0.519</td>
<td>0.383</td>
<td>0.344</td>
<td>0.492</td>
<td>0.521</td>
</tr>
<tr>
<td>UIQM</td>
<td>0.646</td>
<td>0.527</td>
<td>0.576</td>
<td>0.633</td>
<td>0.533</td>
<td>0.529</td>
<td>0.641</td>
<td>0.529</td>
<td>0.614</td>
<td>0.580</td>
</tr>
</tbody>
</table>

Table 1. Comparison results with the proposed method (continue)
Table 1. Comparison results with the proposed method

<table>
<thead>
<tr>
<th>Author/Image</th>
<th>Shipwreck</th>
<th>Fish</th>
<th>Reef1</th>
<th>Reef2</th>
<th>Reef3</th>
<th>Galdran1</th>
<th>Ancuti1</th>
<th>Ancuti2</th>
<th>Ancuti3</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emberten, et al. [21]</td>
<td>PCQI</td>
<td>0.945</td>
<td>1.156</td>
<td>1.078</td>
<td>0.607</td>
<td>0.943</td>
<td>1.147</td>
<td>1.036</td>
<td>0.603</td>
<td>1.129</td>
</tr>
<tr>
<td>Ancuti, et al. [33]</td>
<td>UCIQE</td>
<td>0.632</td>
<td>0.705</td>
<td>0.660</td>
<td>0.718</td>
<td>0.678</td>
<td>0.652</td>
<td>0.499</td>
<td>0.529</td>
<td>0.555</td>
</tr>
<tr>
<td>Ancuti, et al. [18]</td>
<td>UIQM</td>
<td>0.588</td>
<td>0.759</td>
<td>0.690</td>
<td>0.757</td>
<td>0.677</td>
<td>0.664</td>
<td>0.407</td>
<td>0.425</td>
<td>0.563</td>
</tr>
<tr>
<td>Our result</td>
<td>PCQI</td>
<td>1.131</td>
<td>1.089</td>
<td>0.978</td>
<td>0.893</td>
<td>1.191</td>
<td>1.125</td>
<td>1.074</td>
<td>1.015</td>
<td>1.171</td>
</tr>
<tr>
<td></td>
<td>UCIQE</td>
<td>0.634</td>
<td>0.669</td>
<td>0.655</td>
<td>0.718</td>
<td>0.705</td>
<td>0.643</td>
<td>0.588</td>
<td>0.590</td>
<td>0.652</td>
</tr>
<tr>
<td></td>
<td>UIQM</td>
<td>0.629</td>
<td>0.598</td>
<td>0.674</td>
<td>0.733</td>
<td>0.737</td>
<td>0.669</td>
<td>0.547</td>
<td>0.683</td>
<td>0.693</td>
</tr>
</tbody>
</table>

Input | Recovered Image 1 | Recovered Image 2 | Fused Image

Figure 3. The results of multi-layer perceptron based multi fusion strategy. Recovery image 1 is a gamma-corrected version of MLP improved image. Recovery image 2 is a contrast stretching of MLP improved image. The fused image is the balanced colour image of multi-scale fusion approach.
4. CONCLUSION

Image enhancement is the best approach to increase the quality of the submerged images distorted by limited light range, blurs, and reduced contrast. This paper introduces a multi-scale fusion strategy by utilizing the multi-layer perceptron to improve the underwater image quality. A single image dehazing technique works based on the fusion and needn't bother about extra data. Firstly, the multi-layer perceptron is used to dehaze an image from the originally captured image. Then fusion technique is applied to the two images derived from contrast stretching and gamma corrected version of the images. The two images are mixed with a multi-scale fusion procedure with the standardized weight, followed by a color balancing strategy to adjust the color. The upgraded submerged images progressively more visible, interpretable with decent clarity. Though this method uses synthetic images using MLP, it took less computation time, and qualitative observations of this approach show better outcomes compared with the current UIE techniques. The quantitative results show that significantly our approach gives the better result with the difference of 0.536, 2.185, and 1.272 for PCQI, UCIQE, and UIQM metrics, respectively, on the Ancuti dataset. However, the method is to be improved to develop a single UIE technique that can make applicable to all sorts of distorted images underwater.
REFERENCES


Multi-scale fusion for underwater image enhancement using multi-layer perceptron (M. Sudhakara)