Received: 19 April 2021 | Revised: 25 May 2023 | Accepted: 15 June 2023

CAAI Transactions on Intelligence Technology

DOI: 10.1049/cit2.12260

ORIGINAL RESEARCH

Underwater image clarifying based on human visual colour constancy using double-opponency

Bin Kong[1,2,3] | Jing Qian[1,2,4] | Pinhao Song[2,5] | Jing Yang[1,2,3] | Amir Hussain[6]

[1]Institute of Intelligent Machines, Chinese Academy of Sciences, Hefei, China

[2]Peng Cheng Laboratory, Shenzhen, China

[3]Anhui Key Laboratory of Biomimetic Sensing and Advanced Robot Technology, Hefei, China

[4]University of Science and Technology of China, Hefei, China

[5]Peking University Shenzhen Graduate School, Shenzhen, China

[6]Edinburgh Napier University, Edinburgh, Scotland

Correspondence

Bin Kong and Jing Qian.
Email: bkong@iim.ac.cn and qjjq@mail.ustc.edu.cn

Abstract

Underwater images are often with biased colours and reduced contrast because of the absorption and scattering effects when light propagates in water. Such images with degradation cannot meet the needs of underwater operations. The main problem in classic underwater image restoration or enhancement methods is that they consume long calculation time, and often, the colour or contrast of the result images is still unsatisfied. Instead of using the complicated physical model of underwater imaging degradation, we propose a new method to deal with underwater images by imitating the colour constancy mechanism of human vision using double-opponency. Firstly, the original image is converted to the LMS space. Then the signals are linearly combined, and Gaussian convolutions are performed to imitate the function of receptive fields (RFs). Next, two RFs with different sizes work together to constitute the double-opponency response. Finally, the underwater light is estimated to correct the colours in the image. Further contrast stretching on the luminance is optional. Experiments show that the proposed method can obtain clarified underwater images with higher quality than before, and it spends significantly less time cost compared to other previously published typical methods.

KEYWORDS

computers, computer vision, image processing, image reconstruction

1 | INTRODUCTION

Optical detection and sonar detection are two main means of environment sensing for underwater vehicles. Compared with sonar image, optical image has higher resolution and richer information with more direct expression [1, 2]. Therefore, the optical detection has advantages over the sonar detection in close range operations for underwater vehicles. However, when light propagates in water, the water medium and the impurity particles in it will cause the light absorption and scattering effects [3]. As a result, the underwater images will be degraded and often appear as blurred or hazed, contrast declined and colour biased. Consequently, these will cause great trouble in feature extraction, target detection and tracking. Therefore, improving the clarity of underwater images is very important to support the optical detection techniques in close range operations of underwater vehicles [4].

Most traditional underwater image clarifying methods can be divided into two categories: the type of image enhancement methods and the type of image restoration methods [5, 6]. The idea of image enhancement methods for underwater image clarifying is to deal with the underwater image through various image processing methods or image transformations, so as to eliminate the influence of light absorption and scattering on the image [7, 8] and get a higher sharpness image than the original one [9, 10]. The image restoration methods are based on establishing a mathematical model for the imaging degradation process. By estimating the model parameters and then inverting the degradation process, a clarified underwater image is expected to be obtained [11, 12]. Either image enhancement methods or image restoration methods require large amounts of calculation, and usually they would take a much longer time than desired in real-world applications to get the clarified result image.

With the remarkable success of deep learning methods in solving high-level and low-level vision problems, learning-based methods for underwater image clarifying are emerging [13, 14]. The learning-based methods rely on sufficient and effective training data with ground truth, and the truth value should be unique. However, for a real underwater image, it is difficult to obtain the corresponding clarified ground truth directly. When using manual annotation, different persons may give different outcomes, and it is hard to say what kind of image appearance should be taken as a unique ground truth of clear image for an underwater image. Therefore, we think that the learning-based methods are not suitable to be used for underwater image clarifying currently. Consequently, we mainly focus on the traditional non-learning methods in this paper.

In practical applications, there is a need to clarify the colour biased and hazed underwater images quickly to help automatic or manual underwater operations. The purpose of our work is not only to design a fast underwater image clarifying method to be used in real underwater scenes but also to make the method rely less on training data. In this paper, we propose a new idea for underwater image clarifying, which corrects the biased colour by imitating the colour constancy mechanism of the human eye.

The rest of our paper is arranged as follows: In Section 2, we give a brief review of the related works. In Section 3, we begin with an overview of the proposed method followed by an introduction of the double-opponency model for colour perception by the human eye, and then we describe in detail the process of our method. In Section 4, we show the experiment results with analysis and discussion, on ablation study and comparison with other methods. Finally, the conclusions and future work are presented in Section 5.

2 | RELATED WORK

With the development of underwater robots and facilities, underwater image clarifying has been gradually becoming one of the hot research topic since the end of the last century [15]. Various image enhancement methods or image restoration methods, and sometimes their different combinations, have been applied to deal with the problems emerging in underwater images, such as low contrast, colour distortion, blurred texture, uneven illumination, and so on. These methods can be classified into several types, including: colour correction and contrast stretching methods, Retinex based methods, domain transformation methods, dehazing based methods, optical model based methods, and hybrid methods. In recent years, since the success of learning-based methods in many areas, this type of method has also been brought into the field of underwater image clarifying. Here we give a brief review of these methods through some examples in the following.

Colour correction and contrast stretching methods: In 2004, Chambah et al. [16] proposed an unsupervised colour equalisation algorithm, which was inspired by human visual adaptation mechanism about lightness and colour, to help fish segmentation and feature extraction. In 2005, Torres-Méndez and Dudek [17] used a Markov Random Field to build the relation between the distorted and undistorted colour, then recovered the colour of underwater images based on training data and energy minimisation. In 2007, Iqbal et al. [18] gave a sliding stretch algorithm that based on the similar distribution hypothesis on the histograms of the red (R), green (G), and blue (B) components of an ideal image. It first stretches the value range of R and G to the same range of B in RGB colour space, then stretches the range of saturation (S) and intensity (I) to [0,255] in HSI colour space. In 2015, Singh et al. [19] presented two recursive algorithms using histogram equalisation to enhance the images acquired in low light condition such as underwater or night. Both the algorithms deal with the sub-image histograms, based on their exposure values and predefined thresholds in different ways. This type of method usually can improve the contrast of an underwater image effectively and correct the colours to some extent, but meanwhile, it is easy to blow up the noise and add artefacts to the image.

Retinex based methods: Retinex is a computational theory proposed by Edwin. H. Land to model colour and brightness perception and constancies in human vision. In 2010, Shi et al. [20] published the "Underwater image enhancement algorithm based on Contourlet transform and multi-scale Retinex". This algorithm first uses a non-subsampled Contourlet transform to decompose the underwater image, and then applies the multi-scale Retinex in the low frequency subband to enhance the global contrast. But it is unsuitable to handle the images with insufficient or uneven illumination. In 2017, Zhang et al. [21] extended the multi-scale Retinex to CIELAB colour space, and used the combination of bilateral filter and trilateral filter on different colour channels according to different constraints. It can suppress the generation of halo artefacts, but its computational complexity is high due to many parameters involved.

Domain transformation methods: This type of methods transform the image into other domain by some mapping and then take advantage of the special properties of the domain to process the image. The most typical mapping used is the wavelet transform. In 2010, Lan et al. [22] proposed a wavelet based method to reduce backscatter noise in underwater images. After wavelet decomposition, different filters were applied to filter the low and high frequency coefficients respectively to trade-off between eliminating noise and preserving details. In 2015, Singh et al. [23] used wavelet transform to correct the colour of underwater images. In 2017, Vasamsetti et al. [24] presented a wavelet based framework for underwater image enhancement. By applying wavelet decomposition and a set of energy functionals, they first modified the approximation coefficients of RGB components to adjust the average intensity of the image, and then modified these coefficients at finer scales to correct the colour and improve the contrast. This type of method usually can well remove the noise in underwater images, but it is not good for other image degradation problems.

Dehazing based methods: Many underwater images look like foggy image, so dehazing algorithms were borrowed to process underwater images by many researchers. In particular,

the dehazing algorithm proposed by He et al. [25] basing on dark channel prior (DCP) is widely adopted. In 2012, Chiang and Chen [26] used wavelength compensation and dehazing to reduce the influence of non-uniform auxiliary illumination and improve the colour fidelity. In 2015, Galdran et al. [27] applied DCP mainly to the red channel of underwater images to recover the lost contrast and colour. In 2017, Ni et al. [28] used a guided filter and two additional parameters to optimise adaptively the estimation of transmittance map. They also integrated colour temperature adjustment in the algorithm to rectify the colour of image. In 2018, Wang et al. [29] presented an algorithm that utilised sparse representation on the obtained dark channel image for deblurring and denoising, and consequently improved the entropy and average gradient of underwater images. A shortage in this type of method is that they all need a long time to fulfil the computation.

Optical model based methods: This type of methods builds the computer model of light propagation or image degradation in underwater, and clarify underwater images through estimating the model parameters and then inversing the imaging process. In 2004, Schechner and Karpel [30] proposed a method based on the characteristics of backscattered light to improve the visibility of underwater images. In 2006, Trucco and Olmos-Antillon [31] gave a self-tuning underwater image restoration algorithm based on a simplified Jaffe-McGlamery model of underwater imaging. The algorithm worked under two ideal assumptions that there was uniform illumination and only forward scattering, which rarely occur in real world. In 2018, Xie et al. [32] restored the image by first estimating the underwater optical parameters related to natural underwater light, and then calculating the transmittance of R, G and B channel respectively according to the relationship between scattering coefficient and wavelength of light. This type of method usually made some assumptions and hypotheses that were not always true in real underwater environment.

Hybrid methods: Due to the complexity of real underwater environment, a single method often cannot achieve the desired results. Many researchers applied several different techniques in combination to deal with different problems in underwater images, hoping to obtain better outputs. In 2006, Bazeille et al. [33] presented an automatic pre-processing algorithm that used in turn a homomorphic filter to correct uneven lighting and sharpen the edges, a wavelet filter to suppress the noise, an anisotropic filter to smooth textures and reduce artefacts, and then performed a histogram based intensity and colour adjusting. In 2012, Ancuti et al. [34] proposed a fusion strategy to enhance underwater images. They first produced a colour corrected version and a contrast enhanced version of the original image by white balancing, temporal bilateral filtering and local adaptive histogram equalising. Then they constructed four weight maps that aim to global and local contrast, saliency and exposedness, respectively, for further multiscale fusion. In 2015, Li and Guo [35] gave a hybrid approach that used a simple dehazing algorithm followed by colour compensation, histogram equalisation, saturation/intensity stretching and bilateral filtering to obtain better visibility of underwater images. In 2022, Gong and Hua [5] proposed a colour compensation method for underwater images. They first improved the low-brightness areas by fusing the polarisation image with the intensity image. Then they adopted the DCP principle to enhance and deblur the image. This type of method often gives better results than one means alone, but correspondingly takes longer time to process.

Learning based methods: In 2017, Wang et al. [13] presented a CNN based network called UIE-Net for colour correction and haze removal to enhance underwater images; In the same year, Sun et al. [36] proposed another underwater image enhancement model based on an Encoding-Decoding deep CNN network. This type of method needs image dataset to train and evaluate the network. Earlier, they used a synthetic dataset produced in a certain way. For example, literature [36] simulated water with varying degrees of turbidity by adding different amount of milk into fixed volume of water, literature [37] synthesised the underwater images from clean RGB-D images captured on the ground. Later, some datasets containing real underwater images emerged. In 2020, Li et al. [38] constructed an Underwater Image Enhancement Benchmark and proposed a gated fusion network called Water-Net as a baseline. Underwater Image Enhancement Benchmark includes 950 real-world underwater images, among which 890 have corresponding reference images that were selected by volunteers from the outputs of 12 existing image enhancement methods. In 2021, Li et al. [39] presented an underwater image enhancement network Ucolor, which was composed of a multi-colour space encoder with channel-attention modules and a medium transmission-guided decoder. Although in recent years, the learning-based method has been taken as a preferred solution for most applications by many researchers, but its performance is heavily influenced by training data. When the actual underwater scenario is quite different from that in training data, the outputs of such a method are often far from satisfactory.

In addition to these types of methods mentioned above, other techniques were also used for underwater image clarifying. For example, Abunaser et al. [40] utilised particle swarm optimisation to reduce the influence of light absorption and scattering on underwater images. Besides, researchers have also been working on developing new optical imaging techniques, such as underwater laser scanning imaging and range-gating imaging, for acquiring clearer underwater images of more distant targets. Although the development of such devices in recent years has improved the range and clarity of underwater imaging, image processing technology still keeps its position with its advantages of low cost, low load and easy implementation.

3 | METHOD

3.1 | Overview of proposed method

In an underwater application we involved, it expected to make quickly the colour biased and hazed underwater images clear. When we reproduced the existing methods given by other scholars, we found that either the image restoration methods or the image enhancement ones would consume much more

time than desired. We thought that human visual mechanism might help the quick image clarifying.

Human beings have prior knowledge of the normal colour expressions of objects in the world. The human visual system has the abilities of eliminating the influence of most colour-changing factors and giving the correct colour perception to objects. Visual physiologists have found that double-opponency colour sensing cells (double-opponent (DO) cells) in human primary visual cortex (V1) are the physiological basis of the stability of colour perception. Gao et al. showed in [41] that the response distribution of DO cells to the colour biased image is very consistent with the vector representing the colour of the light source, and established a computational colour constancy model that can efficiently estimate the colour of the scene light source. In the model, most of the calculations are linear. This inspired us an idea to quickly correct the biased colours in an image. Different from the traditional image restoration methods that are based on the physical model, we tried to use the double-opponency based colour constancy (DOCC) model of human's colour perception to deal with underwater images and avoid modelling the complex light propagation and imaging process in underwater environment.

In the following, we propose a new method to clarify underwater images by imitating human visual colour constancy using double-opponency. This method consists of two main stages: double-opponency calculation, which implements the response from the retina to the V1 area of the brain, and colour constancy transformation, which corrects the perception of colours. A further contrast enhancement stage is optional because the output of the former two stages has an accompanying effect of stretching the illuminance distribution. Figure 1 shows the flow diagram of the proposed method. We describe the details in the following subsections.

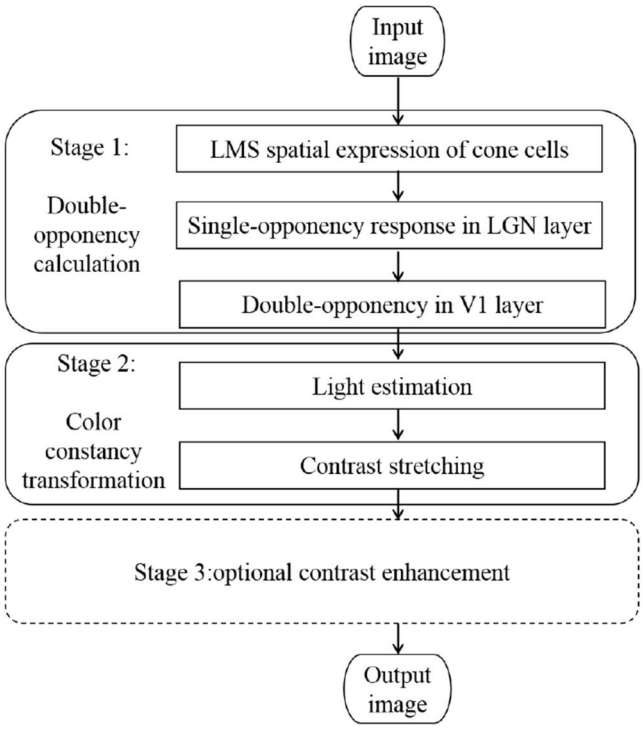


FIGURE 1 Flow diagram of our method.

3.2 | Double-opponency based colour constancy model

The colour perception in human vision system is a hierarchical process, which starts from the retina and, through the lateral geniculate nucleus (LGN) of the thalamus, reaches the V1 area of the primary and other advanced visual cortex to progress visual information analysis.

Based on the neural mechanism for colour processing in the early visual stages, Gao et al. [41] presented a computational DOCC model.

There are two types of photoreceptors on the retina, that is, rod cells and cone cells. Among them, cone cells are responsible for colour vision. They transform light signals entering the eye into bioelectrical signals that can be transmitted in the neuronal system. According to the spectral response characteristics of the three kinds of cone cells, each kind is more sensitive to different wavelengths of light, that is, the long-wavelength (L) cone cells, medium-wavelength (M) cone cells and short-wavelength (S) cone cells respond preferably to the red (R), green (G) and blue (B) colour respectively. In the visual computing model of DOCC, this stage is represented as converting the visual signal from RGB space to LMS space, where L, M, and S denote the light response to long, medium and short wavelengths respectively.

Then, the colour information is encoded by colour-opponency via the single-opponent (SO) cells and then the DO cells. The SO cells exist in the retinal ganglion layer and LGN. They code the colour information within their receptive fields (RFs) in the way of red-green, blue-yellow, and black-white opponency. The DO cells exist widely in V1. They are capable of detecting local colour contrast between the centre and surround of the RF through spatial transformation, thus helpful to colour constancy. Subsequently, the colour stream originally coded in colour opponency space (red-green, blue-yellow, and black-white) in early visual cortex is transformed to trichromatic space (red, green, and blue) in the higher visual cortexes, where the visual information processing and analysis is done.

3.3 | Stage 1, double-opponency calculation

In our method, we use the main part of the DOCC model to calculate the output of DO cells' responses in LMS space and call it double-opponency calculation. The schematic diagram of the double-opponency calculation is shown in Figure 2.

3.3.1 | LMS spatial expression of cone cells

Corresponding to the first stage of visual perception that L, M and S cones of the retina encode the information entering the eye, we need to convert the image from RGB space to LMS

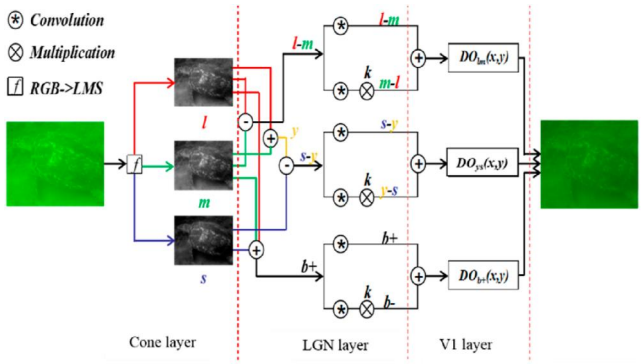


FIGURE 2 Schematic diagram of double-opponency calculation (modified based on [41]).

spatial expression of cone cells, with the help of XYZ colour space.

According to the standard issued by International Telecommunication Union, the relation between RGB colour space and XYZ colour space is given in formula (1) [42]:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4306 & 0.3415 & 0.1784 \\ 0.2220 & 0.7067 & 0.0713 \\ 0.0202 & 0.1295 & 0.9394 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{1}$$

Then, the conversion from XYZ colour space to LMS space is given in formula (2):

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.3897 & 0.6890 & -0.0787 \\ -0.2298 & 1.1834 & 0.0464 \\ 0 & 0 & 1.0000 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{2}$$

Combining formula (1) and (2), we can get formula (3) that gives the spatial transformation between RGB and LMS:

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.3192 & 0.6089 & 0.0447 \\ 0.1647 & 0.7638 & 0.0870 \\ 0.0202 & 0.1295 & 0.9391 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{3}$$

3.3.2 | Single-opponency response in LGN layer

According to the following formula (4), the spatial information of LMS is transformed into single-opponency (denoted as O) result:

$$\begin{bmatrix} O_{lm} \\ O_{ys} \\ O_{b+} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix}, \quad \begin{bmatrix} O_{ml} \\ O_{sy} \\ O_{b-} \end{bmatrix} = -\begin{bmatrix} O_{lm} \\ O_{ys} \\ O_{b+} \end{bmatrix} \tag{4}$$

where the subscript characters l, m and s represent the three components of the input image in LMS space respectively, y is the synthetic yellow component given by $y = m + s$, and b is the luminance component given by $b = l + m + s$.

Then the response of RFs (denoted as SO) is calculated by Gaussian convolution. Formula (5) gives an example of the red-excitation/green-inhibition opponency:

$$SO_{l+m-}(x, y; \sigma) = O_{lm}(x, y) \otimes \mathrm{RF}(x, y; \sigma) \tag{5}$$

where "+" denotes excitation and "-" denotes inhibition, $\otimes$ denotes the convolution, σ gives the size of RF, and RF describes the structure of RF as a Gaussian function:

$$RF(x, y; \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \tag{6}$$

3.3.3 | Double-opponency in V1 layer

In V1 cortex, two single-opponency with different sizes of RFs work together to form a double-opponency response (denoted by DO), given by formula (7):

$$\begin{cases} DO_{lm}(x, y) = SO_{l+m-}(x, y; \sigma) + k \cdot SO_{m+l-}(x, y; \lambda\sigma) \\ DO_{ys}(x, y) = SO_{y+s-}(x, y; \sigma) + k \cdot SO_{s+y-}(x, y; \lambda\sigma) \\ DO_{b+}(x, y) = SO_{b+}(x, y; \sigma) + k \cdot SO_{b-}(x, y; \lambda\sigma) \end{cases} \tag{7}$$

where σ defines the size of centre RF, and $\lambda\sigma$ defines the size of surrounding RF. Generally, the size of the surrounding RF is set three times that of the central RF, that is, $\lambda = 3$. Parameter $k \in$ is the weight of surrounding RF. When $k = 0$, the effect of surrounding RF is ignored.

Then in the high-level visual cortex, the signals are transferred back to LMS space to obtain colour information according to the formula (8):

$$\begin{bmatrix} DT_l \\ DT_m \\ DT_s \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}^{-1} * \begin{bmatrix} DO_{lm} \\ DO_{ys} \\ DO_{b+} \end{bmatrix} \tag{8}$$

3.4 | Stage 2, colour constancy transformation

Colour bias in underwater images mainly results from light absorption by water. From the view of perception, it can be equivalent to a uniformly distributed single light source, which we call underwater light in this paper, shining on the scene. In order to eliminate the influence of such underwater light and obtain a constant perception of colour, we need to

estimate the light and wipe off its components in underwater images.

3.4.1 | Light estimation

The components of underwater light (E_l, E_m, E_s) is estimated as follow

$$E_i = \text{pooling}(DT_i) / \sum_{\{l,m,s\}} \text{pooling}(DT_i)\ i \in \{l, m, s\} \quad (9)$$

where pooling (·) denotes the canonical neural computation pooling the DT responses in separate channels over the whole image [41]. In Section 4.3, we will test the following four different pooling schemes and decide using which one to estimate the underwater light:

1) *Max*: Use the maximum value of each channel;
2) *Avg*: Use the average value of each channel;
3) *MaxMin*: Use the brightest value of all local minimum pixels of each channel;
4) *Top*: Take the top 1% brightest pixels in each channel, and among this set of pixels, choose the one that is brightest in the original image.

For the subsequently use of E_i, $i \in \{l,m,s\}$, it is necessary to transfer the result from LMS space expression to RGB space one. The transformation is as follows:

$$E_{RGB} = \begin{bmatrix} E_R \\ E_G \\ E_B \end{bmatrix} = \text{Tran.} * \begin{bmatrix} E_l \\ E_m \\ E_s \end{bmatrix} \quad (10)$$

where

$$\text{Tran.} = \begin{bmatrix} 5.3341 & -4.2829 & 0.1428 \\ -1.1556 & 2.2581 & -0.1542 \\ 0.0448 & -0.2195 & 1.0831 \end{bmatrix}$$

is the transfer matrix from LMS space to RGB space.

3.4.2 | Colour correction

The model preposed by Von Kries [43] is widely used in colour constancy after determining the light source. This model supposes that neurons can achieve colour constancy through a simple signal correction mechanism. According to this model, we use following transformation to correct the colour.

$$\begin{bmatrix} J_R \\ J_G \\ J_B \end{bmatrix} = \begin{bmatrix} E_r/E_R & 0 & 0 \\ 0 & E_g/E_G & 0 \\ 0 & 0 & E_b/E_B \end{bmatrix} * \text{Tran.} * \begin{bmatrix} DT_l \\ DT_m \\ DT_s \end{bmatrix} \quad (11)$$

where (J_R, J_G, J_B) denotes the corrected colour, and ($E_r = 1/3$, $E_g = 1/3$, $E_b = 1/3$) stands for a canonical white illuminant.

3.5 | Stage 3, optional contrast enhancement

Due to the scattering effect of suspended particles in the water, many underwater images are dim and appear as low contrast. In previous stages, we dealt with colour bias through the DOCC mechanism without considering the contrast. In this stage, we handle the contrast and should not change the colours in the image. So we only adjust the luminance contrast of the image and then fuse the result with colour components of the image.

3.5.1 | Adjust luminance

A simple way to enhance the image contrast is to stretch the distribution range of pixel values through proper mapping, for example, a linear mapping for uniform stretching. Many algorithms for grayscale contrast enhancement are available, and anyone could be used to meet our purpose. Here we use the following simple algorithm:

1) Calculate the histogram of grayscale values of an input image.
2) Find the minimum value $V_{\min}$ in the input image.
3) Find the maximum value $V_{\max}$ in the input image. The number of pixels with a value greater than $V_{\max}$ is about 2% of the total number of pixels.
4) Linearly map the pixel values from [$V_{\min}$, $V_{\max}$] to [0, 255]. Pixels with a value greater than $V_{\max}$ are set to 255.

3.5.2 | Fuse luminance with colour components

In order to adjust the contrast of an image without changing the colours in the image, we need to utilise a colour space in which brightness and chroma are expressed independently. We chose the HSV colour space because it intuitively represents the colour and is easy to convert from/to the RGB format.

In the conversion formula from RGB to HSV, the V component, which expresses brightness, takes the maximum value of the R, G and B components. In order to better maintain the relative brightness distribution of pixels in the image, we use the conventional luminance value (Y = 0.299 × R + 0.587 × G + 0.114 × B) instead of the V component for contrast enhancement operation and then transform the new HSV back to RGB.

The process of contrast enhancement is illustrated in Figure 3, where we denote the output of stage 2 (the colour-corrected image) as I_C, the enhanced grayscale image as I_E and the output of this stage (the fused contrast-enhanced image) as I_F. It is executed according to the following algorithm:

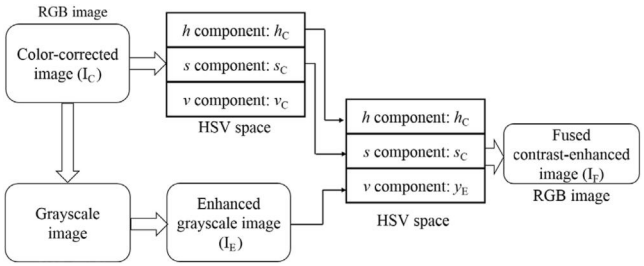


FIGURE 3 Flow chart of stage 3: Enhance the contrast of a colour image without changing its chromatic information.

1) Transform the colour-corrected image I_C from RGB to HSV and grayscale image respectively.
2) Enhance the grayscale image to get I_E using the algorithm described in subsection 3.5.1.
3) Take the hue component h_C and saturation component s_C of I_C as the hue component h_F and saturation component s_F of the fused image I_F respectively. Take the grayscale value y_E of I_E as the value component v_F of I_F. That is:

$$h_F = h_C, s_F = s_C, v_F = y_E \tag{12}$$

4) Transform the fused contrast-enhanced image I_F back to RGB from HSV space.

3.5.3 | Optionality of stage 3

In our initially designing of the underwater image clarifying method, the process of DOCC is followed by a contrast enhancing process. The DOCC process aims at correcting the biased colour caused by the light absorption of water. The contrast enhancing process deals with dim and low illumination caused by the light scattering and other factors.

In fact, the DOCC mechanism also has the side effects of changing the pixel brightness and enlarging the value distribution to some extent. Our experiments show that when the original image is dim and its grayvalues occupy a narrow span, the grayvalue distribution range can be significantly enlarged by the DOCC process. Stage 3 may stretch the range further. However, when large texture-free areas with approximately uniform or graded brightness appear in the original image, our method might introduce stepped texture or speckle noises in such areas. And stage 3 could intensify the steps and noises.

As the effect of stage 3 has both positive and negative sides, we set it to be an optional stage. For the sake of distinction, in this paper, we will refer to the proposed image clarifying method that includes only DOCC processing simply as DO method, while the method that includes all stages 1 through 3 as DO+. Because the computation of stage 3 consumes minimal computing resources, the DO and DO + take almost the same time to obtain the resulting image. Whether to use DO or DO+ in a practical application can be decided on a case-by-case basis. We recommend using the DO method if

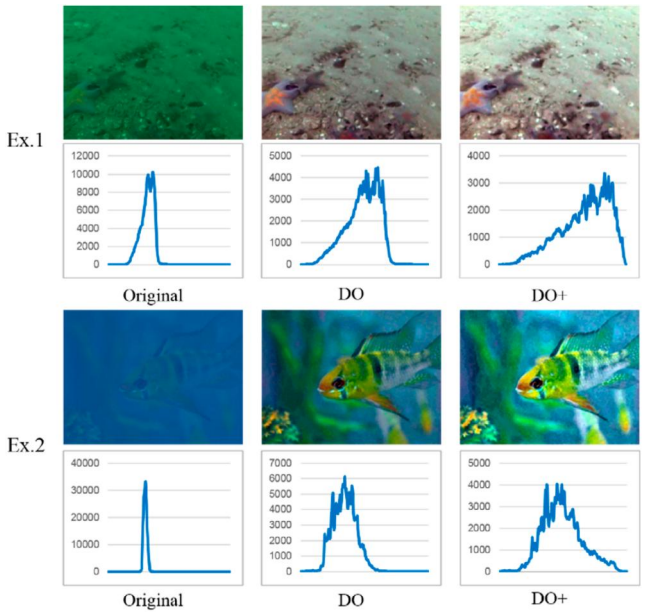


FIGURE 4 Examples of the effect of our proposed image clarifying method double-opponent (DO) and DO+.

the acquired underwater images often include areas of approximately uniform or graded colour (or brightness); Otherwise, using DO+ would get more distribution of pixel values. In the experiments Section 4, we will use DO+ unless otherwise specified.

Figure 4 uses two examples to demonstrate the effect of the DO and DO+. The images and their corresponding grayvalue histograms are presented. The left column in the figure is the original underwater images, the middle the output of DO and the right is DO+.

4 | EXPERIMENTS AND DISCUSSION

In this section, we show the experimental results on underwater images from two different sources. One is a dataset we constructed ourselves, and the other is two open datasets. We want to test our method on underwater images with various qualities in different scenarios. In the early work, considering that most open underwater image datasets have few scenes and quality changes, we constructed a dataset containing images with different qualities and scene styles. As the images in this dataset were collected from the webpages of various contents and purposes, we briefly call the dataset UISI (Underwater Images by Searching the Internet). Later, to verify the effectiveness of our method, the second source we selected is two open datasets as additional experimental objects. That is, Real-world Underwater Image Enhancement (RUIE) [44] and URPC2019 (the third Underwater Robot Picking Contest) [45].

In the following, we first describe the UISI dataset in Section 4.1, then give some evaluation indicators we used to assess the performance of our method in Section 4.2. In Section 4.3, we compare the pooling schemes for estimating

underwater light. In Sections 4.4 and 4.5, we present the comparison of our method to several other typical methods.

The computer we used to do the experiments is MSI micro star GP62 2QE-215XCN, with Intel Core i7 5700HQ CPU, 2.7 GHz frequency, 8G memory, and dual graphics card.

4.1 | UISI dataset

UISI dataset contains 310 underwater images with different qualities. All the images come from webpages of various contents and purposes. Some are snapshots of underwater videos. Their original sizes and aspect ratios were different. We resized all the images to 640 × 480 pixels. In order to facilitate analysing the effect of clarifying results, we divided these 310 images into six groups and assigned each image an index number. See the following description and Table 1 for details:

182 images are from illustrations of webpages. Based on our observation, we divided these images into four groups according to the degree of clarity from high to low and named them Group-1 to Group-4, for simplicity.
114 images are snapshots of three different videos on webpages. We call them Group-5.
14 images are randomly selected from the URPC2019 dataset. These images belong to Group-6.

In Figure 5, six example images are shown for an overview of UISI, each from one group described above.

4.2 | Assessment indicators

Researchers have proposed many indicators for evaluating image quality. Different indicators are suitable for different purposes and applications. Peak Signal to Noise Ratio (PSNR) is a commonly used indicator for measuring image quality, which came from the engineering of assessing image compression methods. However, we do not think PSNR is suitable for evaluating underwater image clarifying methods. Different from evaluating an image compression method, where the original image is regarded as a known pure signal, we cannot know what the noise-free signal should be in an underwater image, especially in that captured in natural environments. Using a colour biased or blurry original underwater image as the baseline for evaluating and comparing underwater image clarifying methods is obviously unreasonable. Therefore, we did not use PSNR as a quality indicator for performance evaluation in this paper.

Considering the characteristics of underwater images and the purpose of image clarifying, we adopted five indicators to assess the performance of our algorithm and the comparison algorithms. Two are general image evaluation indexes, colour bias and contrast. The other two are quality evaluation indicators designed for underwater images, Underwater Colour Image Quality Evaluation (UCIQE) [46] and UIQM (Underwater Image Quality Measure) [47]. The fifth is the time cost of an algorithm.

Colour bias: Sometimes, colour affects the outcome that we recognise and judge objects and their properties. Similarly, different colours let objects reveal different characteristics in images and may result in different recognition outputs. Colour bias often occurs in underwater images mainly because of the absorption of light energy by water, and it is consistent across the entire image in this case. Images taken at a depth below 20 m are mostly green- or blue-bias. Therefore, the elimination of colour bias is one of the essential goals of underwater image clarifying algorithms, and the degree of colour bias should be

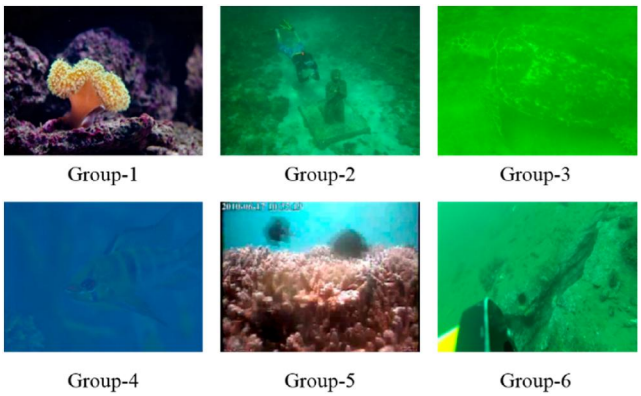


FIGURE 5 Example images in UISI dataset.

TABLE 1 Description of the composition of UISI dataset.

Group name	Number of images	Index assigned	Image quality	Original source
Group-1	32	1–32	Clear	Illustrations on webpages of various contents and purposes
Group-2	83	33–115	Moderate	
Group-3	54	116–169	Dim	
Group-4	13	170–182	Very dim	
Group-5	53	183–235	-	Snapshots of three different videos on webpages
	50	236–285	-	
	11	286–296	-	
Group-6	14	297–310	-	Randomly selected from the URPC2019

taken as one of the assessment indicators to evaluate the output of the clarifying algorithms.

Based on the formula given by Xu et al. [48], we calculate the measure of colour bias *Bias* in CIE Lab colour space as the following:

$$Bias = D_{ab} / R_d * 100 \quad (13)$$

where D_{ab} is the distance from the average chroma point to the origin in a-b plane, R_d is the radius of the equivalent circle of the chroma distribution. The constant 100 is used to enlarge the range of *Bias* values for easy comparison of the results, and it does not change the essence of the original formula. The lower the value of *Bias*, the better.

Contrast: Another widely used indicator of image quality is contrast. A high-contrast image reveals more brightness levels and object details, making objects detected easier than a low-contrast image. A commonly used formula to calculate the contrast value C of an image is as the following [49]:

$$C = \sum_{\delta} \delta(i,j)^2 P_\delta(i,j) \quad (14)$$

where $\delta(i,j) = |i-j|$ is the grayscale difference between adjacent pixels, $P_\delta(i,j)$ is the distribution probability of the difference. In this paper, we use 4-neighbour to define adjacent pixels.

As the formula (14) only considers the intensity of an image, we use the classic equation Y = 0.299 × R + 0.587 × G + 0.114 × B to transform the underwater image to a grey one before calculating its contrast value. Generally, the higher the value of C, the better.

UCIQE: A widely used quantitative indicator for assessing the quality of no-reference underwater images is UCIQE, given by Yang and Sowmya in 2015 [46]. It is calculated in the CIELab colour space, which is more in line with human visual perception, using the following formula:

$$UCIQE = \omega_1 * \sigma + \omega_2 * \delta + \omega_3 * \mu \quad (15)$$

where σ is the standard variance of chroma of the underwater image, δ is the contrast of brightness, μ is the average of saturation, and ω_1, ω_2 and ω_3 are the weights of σ, *con* and μ respectively. A commonly used setting is $\omega_1 = 0.4680$, $\omega_2 = 0.2745$ and $\omega_3 = 0.2576$. The higher the value of *UCIQE*, the better.

UIQM: Another quantitative quality indicator of no-reference underwater images was proposed by Panetta et al. in 2016 [47] and called UIQM. Inspired by the human visual system and considering the characteristics of different aspects of underwater images, three metrics compose the visual quality UIQM according to the following formula:

$$UIQM = \beta_1 * UICM + \beta_2 * UISM + \beta_3 * UIConM \quad (16)$$

where *UICM* is the colourfulness measure of the underwater image, *UISM* is the sharpness measure, *UIConM* is the contrast measure, and β_1, β_2 and β_3 are the weights of *UICM*, *UISM*, and *UIConM* respectively. A typical setting is $\beta_1 = 0.0282$, $\beta_2 = 0.2953$ and $\beta_3 = 3.5753$. The higher the value of *UIQM*, the better.

Time cost: We want to spend as little time as possible clarifying the underwater images in real-world applications whenever there are remote-controlled or automated operations. As generally, the less the time cost, the better.

4.3 | Pooling schemes

In subsection 3.4.1, we gave four possible pooling schemes—*Max*, *Avg*, *MaxMin* and *Top*—to estimate the underwater light. Now we use the assessment indicators described in Section 4.2 to help select which scheme will be used in our later experiments.

We present in Figure 6, the colour bias values of each original image in the UISI dataset and their corresponding clarified images derived from using the four pooling schemes respectively. The abscissa value is the image's index number in the dataset. The ordinate value is the *Bias* value. We also give the group averages of *Bias* values in Table 2.

The data in Figure 6 and Table 2 show that our image clarifying method is able to correct the biased colour and reduce the *Bias* values in most cases. We inspected the original and resulting images and found that the increased *Bias* values mainly occurred on those images for appreciation, which usually have the same high picture quality, and the increase was not statistically significant. Most of those images belong to Group-5. The data also show that, among the four pooling schemes, the *Max* scheme seems the best and the *Top* the worst if Group-5 is not considered.

Figure 7 shows the contrast values of each original image in the UISI dataset and their corresponding clarified images derived from using the four pooling schemes respectively. The abscissa value is the image's index number in the dataset. The ordinate value is the C value. We also give the group averages of C values in Table 3.

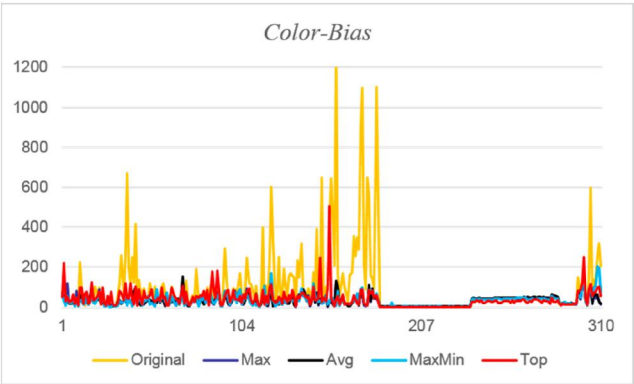


FIGURE 6 Using the *Bias* value to compare different pooling schemes on each image in the UISI dataset.

The data in Figure 7 and Table 3 show that our image clarifying method can enhance the contrast of dim images in which the colour values distribute in a limited range. For those with both the brightest and darkest pixel values, that is, close or equal to 255 and 0, the *C* values of the resulting images may drop. Checking with the images, we found that higher original *C* values also occurred on those for appreciation, where there often were many contrasting spots or edges, and our method would smooth them. The data also show that the *Max* scheme is the best on average and the *Top* is the worst among the four pooling schemes.

Figure 8 shows the *UCIQE* values of each original image in the UISI dataset and their corresponding clarified images derived from using the four pooling schemes respectively. The abscissa value is the image's index number in the dataset. The ordinate value is the *UCIQE* value. We also give the group averages of *UCIQE* values in Table 4.

We can see from Figure 8 and Table 4 that, except for the *Top* scheme and if Group-5 is not considered, our image clarifying method can improve the image's *UCIQE* indicator by the other three schemes, with values close to each other. The *UCIQE* values of clarified images in Group-5 are not always better than that of the original images. The reason is that the images in Group-5 are snapshots of videos for people to enjoy, which usually have the same high picture quality across the whole video.

Figure 9 shows the *UIQM* values of each original image in the UISI dataset and their corresponding clarified images derived from using the four pooling schemes respectively. The abscissa value is the image's index number in the dataset. The ordinate value is the *UIQM* value. We also give the group averages of *UIQM* values in Table 5.

TABLE 2 The group averages of *Bias* values of the images in the UISI dataset with different pooling schemes.

	Original	Max	Avg	Maxmin	Top
Group-1	45.78	37.88	29.60	27.68	49.53
Group-2	90.91	32.84	31.24	33.04	52.55
Group-3	193.84	36.22	36.80	38.20	56.14
Group-4	470.10	38.49	44.28	54.49	46.72
Group-5	20.39	21.34	24.44	21.49	14.77
Group-6	170.23	55.48	63.43	92.00	89.33
All average	99.01	31.10	31.67	32.92	40.60

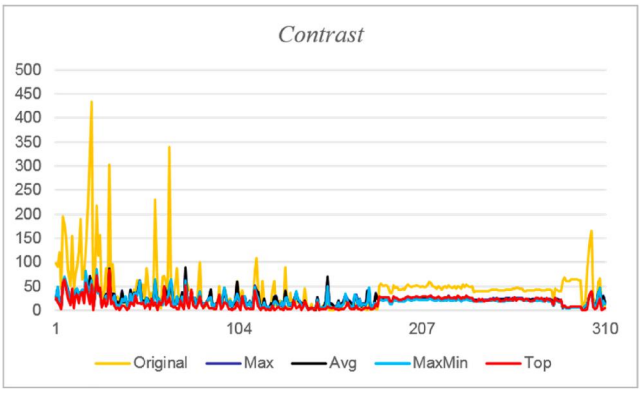


FIGURE 7 Using the constrast *C* value to compare different pooling schemes on each image in the UISI dataset.

TABLE 3 The group averages of *C* values of the images in the UISI dataset with different pooling schemes.

	Original	Max	Avg	Maxmin	Top
Group-1	124.30	35.88	36.24	37.34	30.38
Group-2	38.30	24.15	23.77	22.02	11.73
Group-3	14.42	14.94	14.21	12.36	4.55
Group-4	3.68	20.24	18.39	15.68	4.79
Group-5	46.86	22.91	21.54	18.87	22.39
Group-6	43.43	19.07	18.13	17.23	9.33
All average	44.96	22.91	22.09	20.28	15.89

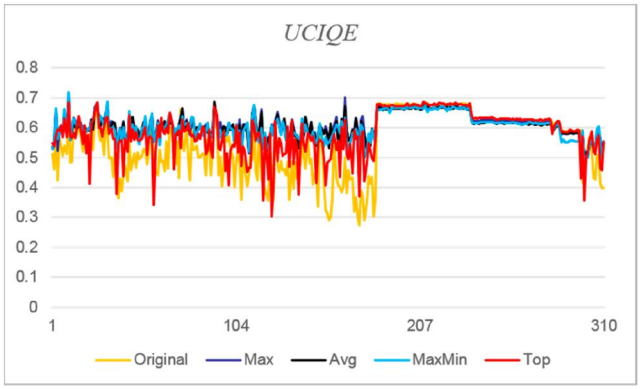


FIGURE 8 The underwater colour image quality evaluation value comparison on different pooling schemes, for images in the UISI dataset.

TABLE 4 The group averages of underwater colour image quality evaluation values of the images in the UISI dataset by using different pooling schemes.

	Original	Max	Avg	Maxmin	Top
Group-1	0.5591	0.6112	0.6092	0.6144	0.5912
Group-2	0.5021	0.5987	0.5974	0.5929	0.5554
Group-3	0.4454	0.5874	0.5840	0.5763	0.5206
Group-4	0.3556	0.5748	0.5656	0.5596	0.5210
Group-5	0.6444	0.6386	0.6346	0.6325	0.6466
Group-6	0.4776	0.5414	0.5389	0.5408	0.5113
All average	0.5432	0.6091	0.6060	0.6030	0.5831

Similar to the *UCIQE* indicator, the *UIQM* values of clarified images in Group-5 are also not always better than that of the original images, and the *Top* scheme again reveals unstable performance even though it increases the *UIQM* value in most cases. By the other three schemes, our image clarifying method can significantly lift the *UIQM* values, and the *Max* scheme has higher values than the other two on average.

Figure 10 shows the time spent for clarifying each image by different pooling schemes, and Table 6 gives the group averages of the time spent. We can see that the *Max* and *Avg* schemes reveal the advantage among the four.

According to the above experiments, it is easy to decide using the *Max* pooling scheme in our proposed method.

4.4 | Compare with other methods on UISI

We mentioned in Section 1 that it is difficult to directly obtain the corresponding clarified ground truth for an actual underwater image. We could use some technical means to produce a clear image for any fuzzy one; however, different methods could be used to yield several images of similar high quality. We could not say one was the ground truth and the others were not. In addition, the learning-based methods tend to be limited to the content of the training dataset. Therefore, we do not think it is reasonable to use so-called ground truth images in open datasets to determine whether our results are good or bad, and we do not consider the learning-based methods suitable for comparison with our method. Consequently, we mainly focus on the traditional non-learning methods in this paper. Experiments are conducted on our no-reference underwater image dataset UISI, and the methods used to compare are traditional non-learning methods.

We choose three typical methods for comparison, which represent different types of ideas and algorithm principles. The first is based on DCP [10], and we call it DCP later for simplicity. It belongs to the classic underwater image restoration method. The second is based on colour compensation [3], and we call it CC. It falls into the classic underwater image enhancement method. The third involves information fusion [50], and we call it IF. We reproduced these three methods. Unless otherwise specified, DO+ is used in the following experiments.

We present in Figure 11 several examples of UISI images and their corresponding clarified results by different methods. The column from left to right is the original images, the output images by DCP, CC, IF, and our DO+ respectively. Table 7 gives the corresponding evaluation data of the images in Figure 11.

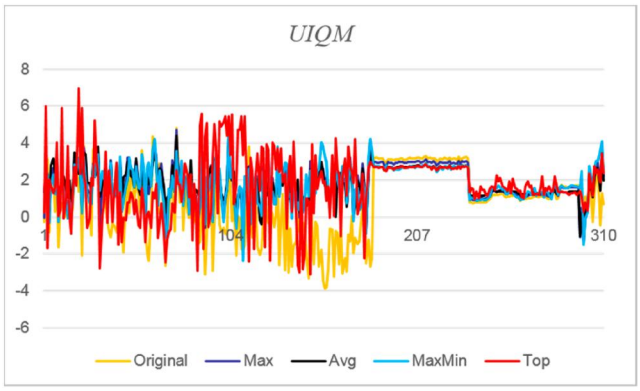


FIGURE 9 The *UIQM* value comparison on different pooling schemes, for images in the UISI dataset.

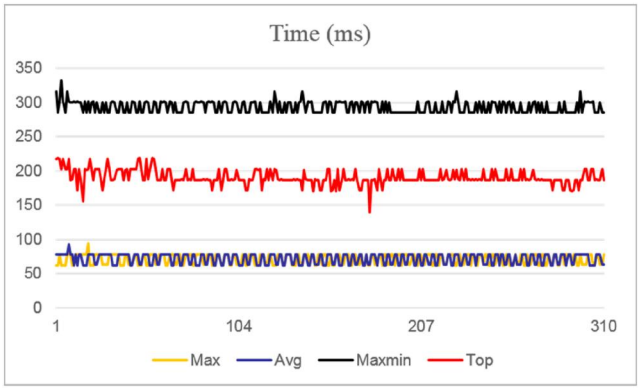


FIGURE 10 The time spent for clarifying each image in the UISI dataset by different pooling schemes.

TABLE 5 The group averages of *UIQM* values of the images in the UISI dataset by using different pooling schemes.

	Original	Max	Avg	Maxmin	Top
Group-1	1.2365	1.7678	2.1037	1.6381	1.9170
Group-2	0.4874	1.8060	1.8556	1.6288	1.5218
Group-3	−1.0726	1.6233	1.5401	1.7325	1.5804
Group-4	−1.1081	1.9947	1.7098	1.7445	1.7338
Group-5	2.0754	2.0487	1.9551	1.9499	2.0379
Group-6	1.0686	1.4221	1.3199	2.0010	1.7413
All average	0.8363	1.8500	1.8326	1.7876	1.7814

TABLE 6 The group averages of time spent for clarifying each image in the UISI dataset by different pooling schemes (unit: ms).

	Max	Avg	Maxmin	Top
Group-1	71.63	75.03	298.03	199.50
Group-2	71.04	71.40	294.43	192.86
Group-3	71.04	70.46	294.48	187.54
Group-4	68.46	69.69	297.77	181.23
Group-5	71.15	71.05	291.31	188.97
Group-6	70.36	71.29	295.86	192.79
All average	70.99	71.41	293.90	190.69

Figure 11 shows that our method DO + can correct the biased colours and stretch the contrast, making the result images clearer than before. Figure 11 also reveals that the effect of de-biasing by DO+ is similar to that by CC and better than that by DCP and IF, and the brightness distribution in the output image by DO + looks more moderate than that by the other three methods.

However, sometimes the quality expressed by the indicators *Bias* and *C* is inconsistent with the image's observation quality. For example, in Figure 11, the four output images by DCP still have a visible colour bias compared to that by DO +, but three of them have a lower *Bias* value except the fourth. Moreover, all four outputs by DCP have a higher *C* value than that by DO+. The third image shows better contrast and more details of seagrass in the DO + output than in the DCP output, yet its *C* value is much lower than that of the DCP output.

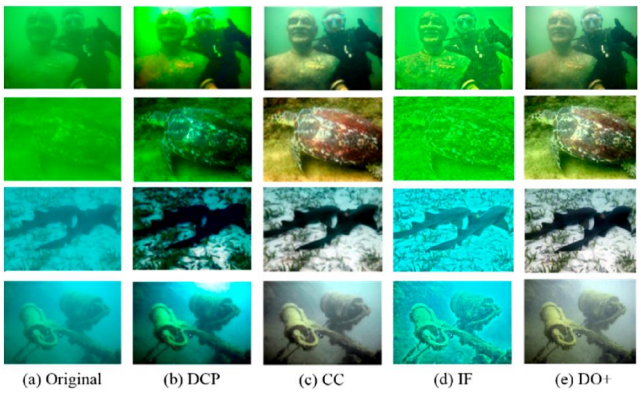


FIGURE 11 Examples of the clarifying results by different methods.

TABLE 7 The corresponding quality indicator values of the images in Figure 11.

	Original	DCP	CC	IF	DO+
Bias	397.57	11.769	40.765	156.63	51.579
C	4.9699	25.440	8.3487	62.802	5.6191
UCIQE	0.4237	0.6316	0.5932	0.4379	0.5765
UIQM	−1.0856	0.1893	1.2002	0.7262	1.0155
Bias	347.73	27.996	23.259	114.49	29.985
C	5.7696	60.392	30.114	72.394	32.171
UCIQE	0.3205	0.5720	0.6175	0.3615	0.5849
UIQM	−1.3089	0.8408	3.9121	0.0482	3.2222
Bias	674.28	8.8221	41.380	296.54	21.497
C	17.203	186.87	44.029	110.07	40.432
UCIQE	0.3641	0.6430	0.5809	0.3952	0.5641
UIQM	−1.0997	1.6444	2.8651	−0.3970	3.0894
Bias	256.86	39.732	43.506	124.28	31.468
C	7.9404	46.348	8.0367	90.879	5.3104
UCIQE	0.4568	0.6236	0.5831	0.4484	0.5513
UIQM	−0.8345	0.0558	1.6462	0.6662	1.4694

We analysed such a phenomenon by looking inside the example images, and found the following explanation:

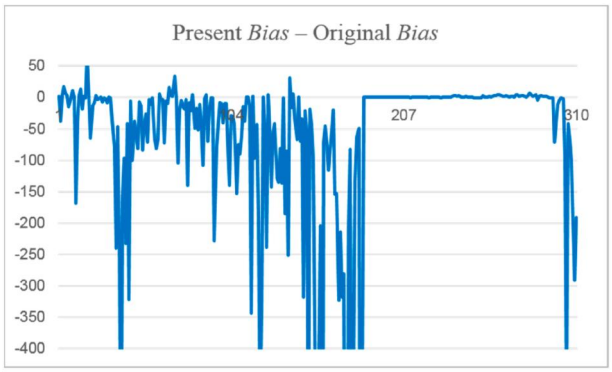
The DCP method brightens the brighter areas and darkens the darker areas with no special treatment for colour. As a result, it tends to increase the difference between the sides of an edge and enlarge the dark areas. Consequently, its output image might form a second dark peak near the origin in addition to the chromatic peak in the a-b plane. As in Figure 11, the black wetsuit in the first image with green bias and the dark fish and seagrass in the third image with blue bias drag the average chroma to the origin and extend the equivalent circle of the chroma distribution. This means a smaller D_{ab} and bigger R_d in formula (13) and comes out with a smaller *Bias* value. Meanwhile, more sharper edges result in higher *C* value. In the third image in Figure 11, the grayscale differences between seagrass and seabed in DCP output are much greater than that in DO+ output and the amount of seagrass edges is large, leading to a much higher *C* value of DCP output than that of DO+ output. However, the DO+ output appears better with more levels than the DCP output. The formula (14) does not consider the contrast between different colours. A colour image with a sharp chrominance contrast may have little change in brightness, resulting in a small contrast value calculated by the formula (14).

Though the indicators *Bias* and *C* cannot always accurately reflect the quality of an image, they can still provide reference to some extent. Therefore, we still present the evaluation results on these two indicators in this paper.

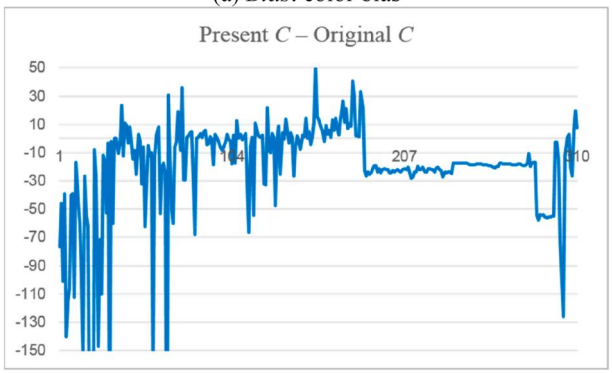
We plot in Figure 12 the improvement of the values of the four quality indicators after using our method (DO+). The abscissa value is the index number of an image in the UISI dataset. The ordinate value is the difference between the assessed values of a clarified image and its original one. A negative value means that the assessed value is reduced. A lower *Bias* value is usually better, so a lower negative value would be better. As to the other three indicators, a negative value means the image quality has fallen after the DO+ process under the specific indicator. Figure 12 shows that our method is effective in most cases. Fallen quality values are found mainly in the video snapshots, consistent with the fact that videos for enjoyment generally have high quality.

In Figure 13, we give the average quality values of each group in UISI dataset before and after different clarifying processes. Here we give both the values of DO and DO+. The figure shows that all the group average values of DO+ are greater than that of DO, which indicates that the additional contrast enhancement stage would get further improvement on the indicators *C*, *UCIQE* and *UIQM* but lose a small amount on the indicator *Bias*. When we look at the individual corresponding value pairs of each image in the UISI dataset, the values show that 306 out of 310 contrast indicator *C* values of DO+ are greater than that of DO, indicating the statistical advantage of DO+. On the other three indicators, there are both, more or less, positive and negative values of the pair difference, and some of the difference values are very small. It

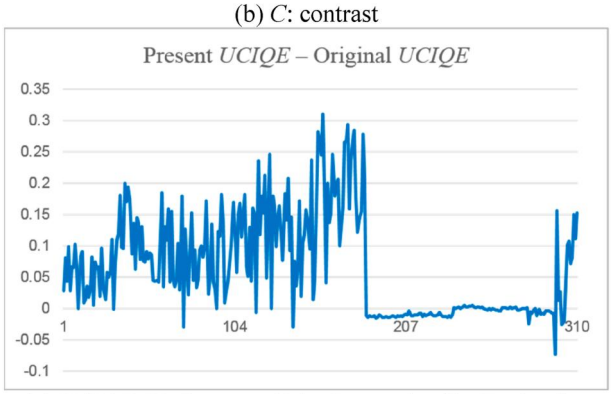
does not tell whether DO+ or DO has a statistical advantage. Combined with the visual perception of the 310 pairs of resulting images (two examples are given in Figure 4), we conclude that DO and DO+ have comparable performance. Whether to use DO or DO+ in a practical application can be decided on a case-by-case basis. We recommend using the DO



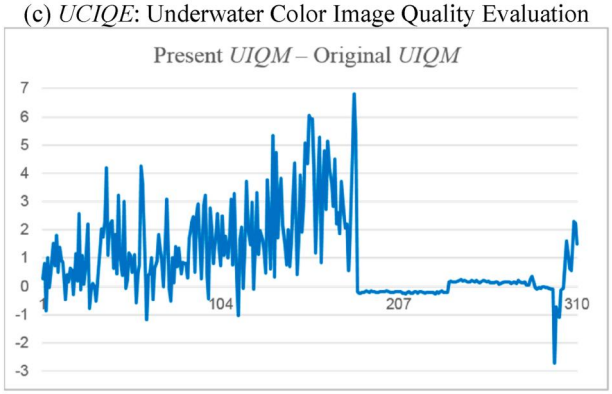
(a) *Bias*: color bias



(b) *C*: contrast

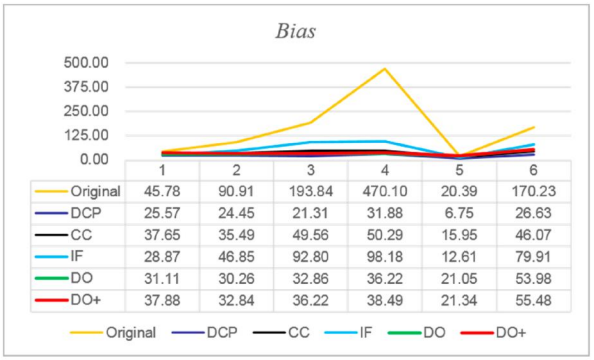


(c) *UCIQE*: Underwater Color Image Quality Evaluation

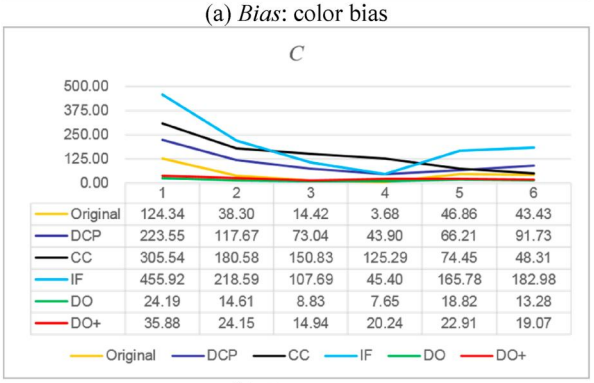


(d) *UIQM*: Underwater Image Quality Measure

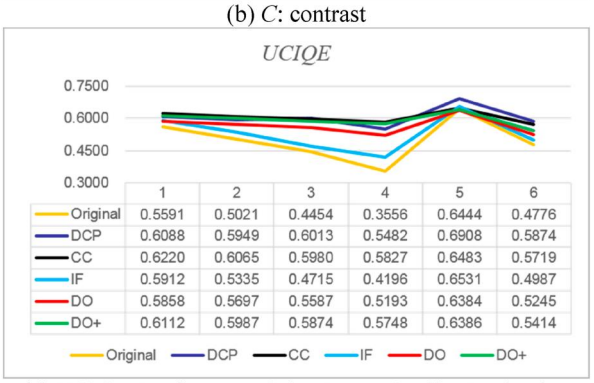
FIGURE 12 Improvement of the four quality indicators by our method DO+: (a) *Bias*, (b) *C*, (c) *UCIQE*, and (d) *UIQM*.



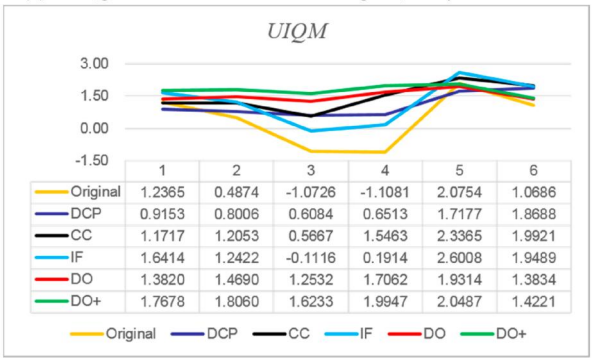
(a) *Bias*: color bias



(b) *C*: contrast



(c) *UCIQE*: Underwater Color Image Quality Evaluation



(d) *UIQM*: Underwater Image Quality Measure

FIGURE 13 The group averages of the four quality indicators in the UISI dataset before and after different clarifying processes: (a) *Bias*, (b) *C*, (c) *UCIQE*, and (d) *UIQM*.

method if the acquired underwater images often include areas of approximately uniform or graded colour (or brightness); Otherwise, using DO+ would get more distribution of pixel values.

In Figure 14, we give the time cost of clarifying each image in UISI dataset by different methods. We also counted the average processing times of each group and listed them in Table 8.

The figures show that the quality performance of our method is, on average, at a top or medium level without considering the video snapshots, and the time cost is far less than other methods. According to the formulae given in Section 3, the time complexity of the main steps of our method DO+ is $O\ (MN)$, where $M \times N$ is the image size. Our experiment on the UISI dataset showed that, among the methods used for comparison, CC is the fastest and IF is the slowest. We calculated the ratio of the time used by DO+ to that by CC for each image and got the maximum value of 13.97% and the minimum value of 7.09%. On average, the proposed method DO+ reduces the time cost to less than one-10th of the CC needed when the image size is 640 × 480.

Based on the above evaluations, our proposed underwater image clarifying method presents significant advantages.

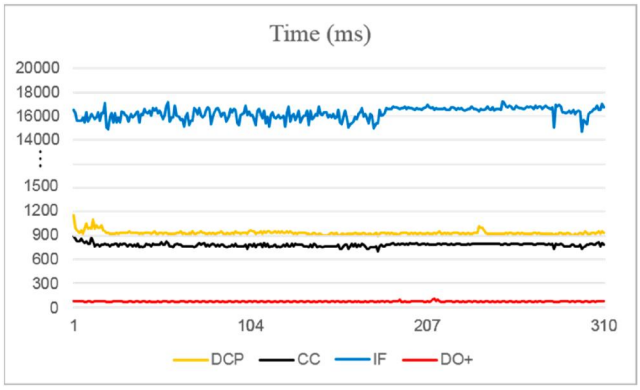


FIGURE 14 The processing times of different methods on each image in UISI.

TABLE 8 Average time of each group used by different methods (unit: ms).

	DCP	CC	IF	DO+	DO+/CC
Group-1	971.59	800.97	15,923.88	66.84	8.35%
Group-2	932.98	775.10	16,131.31	65.78	8.49%
Group-3	925.91	768.17	15,978.31	64.70	8.42%
Group-4	921.62	753.69	15,877.31	64.92	8.61%
Group-5	927.39	787.66	16,637.87	64.32	8.17%
Group-6	931.71	781.14	16,220.86	66.79	8.55%
All average	933.14	780.55	16,262.92	65.17	8.35%

4.5 | Experiment on real-world underwater image enhancement and URPC2019 data

To verify our method, we made testing and comparison on underwater images in RUIE and URPC2019 datasets respectively. Some results are shown in Figures 15 and 16 and Tables 9 and 10. The quality values in Tables 9 and 10 are calculated according to formula (13)–(16).

Figures 15 and 16 show that the quality of all images assessed by visual perception was improved and got consistent colour correction by our method DO+. Comparing the four methods, we found that DO+ and CC could correct the biased colour better than DCP and IF. DO+ and DCP could keep the colour of the same object in different scenes more similar and closer to its actual colour than CC and IF (see the object at the bottom-right corner of the images in Figure 16). We also noticed that the colours of the output images by DO+ might be wrong when the motion blur in an image was noticeable.

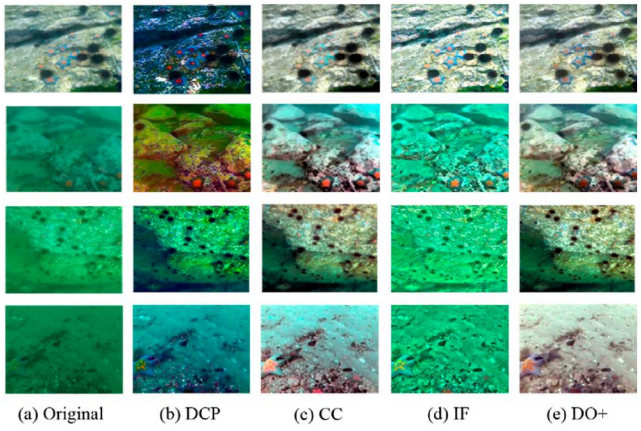


FIGURE 15 Some examples of the images in the real-world underwater image enhancement (RUIE) dataset: (a) original images, and the corresponding clarified images by (b) DCP, (c) CC, (d) IF, and (e) our method DO+.

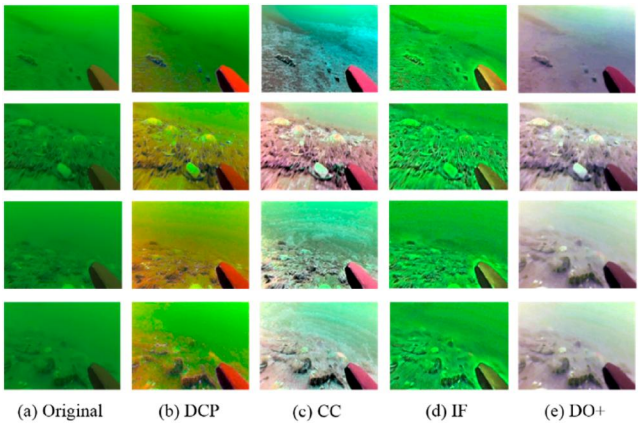


FIGURE 16 Some examples of the images in the URPC2019 dataset: (a) original images, and the corresponding clarified images by (b) DCP, (c) CC, (d) IF, and (e) our method DO+.

TABLE 9 The quality indicator values corresponding to the images in Figure 14.

	Original	DCP	CC	IF	DO+
Bias	40.798	7.7204	39.640	20.424	24.142
C	71.759	255.909	41.241	219.26	33.728
UCIQE	0.5112	0.6300	0.5671	0.5179	0.5403
UIQM	3.5711	3.2164	3.5441	4.2884	3.2335
Bias	271.54	10.446	15.716	85.182	21.461
C	6.6187	48.751	39.860	87.225	26.846
UCIQE	0.3887	0.5750	0.6081	0.4576	0.5727
UIQM	0.4875	2.2807	2.6194	1.8046	2.7384
Bias	137.56	21.816	30.530	84.491	43.583
C	45.229	206.22	38.317	148.10	42.230
UCIQE	0.4100	0.6164	0.5982	0.4400	0.5896
UIQM	0.8046	1.5058	3.3439	1.6692	3.1263
Bias	424.37	29.742	13.589	111.35	15.755
C	3.9693	48.618	46.931	71.150	19.813
UCIQE	0.3516	0.4945	0.5811	0.4483	0.4980
UIQM	0.0611	0.4493	2.4301	1.1776	2.6600

TABLE 10 The quality indicator values corresponding to the images in Figure 15.

	Original	DCP	CC	IF	DO+
Bias	80.303	14.575	9.9124	59.335	12.086
C	1.4494	18.551	16.599	31.125	2.5715
UCIQE	0.4555	0.5694	0.5951	0.4720	0.5773
UIQM	−1.1867	0.8356	0.9134	0.3227	1.3033
Bias	137.14	17.212	7.398	87.917	0.6758
C	3.8653	42.188	26.039	65.360	11.893
UCIQE	0.4054	0.5638	0.5763	0.4764	0.5369
UIQM	0.3140	2.9685	3.2171	1.5566	2.3627
Bias	129.77	21.803	21.498	102.505	25.022
C	2.9092	39.070	31.630	58.609	5.7139
UCIQE	0.4115	0.5427	0.5473	0.4709	0.5277
UIQM	−0.5766	2.7096	1.3365	1.1184	1.4759
Bias	136.16	14.820	15.609	111.91	6.0916
C	1.4367	51.483	14.707	32.208	4.6362
UCIQE	0.4118	0.5590	0.5587	0.4693	0.5268
UIQM	−0.6964	1.8287	1.4512	0.9407	1.4354

For areas of dense texture, the objects such as echini, starfish and holothurians in the clarified images by all methods can be seen more clearly. In contrast, for large texture-free areas with approximately uniform or graded colour and brightness, DO+ and CC may introduce stepped texture.

Tables 9 and 10 show that most of the values of the quality indicators are improved by our method DO+, and DO+ is in the middle level compared with other methods.

Generally speaking, the experiment on RUIE and URPC2019 data can support the validity of our method.

5 | SUMMARY

In this paper, we propose an underwater image clarifying method based on human visual colour constancy using double-opponency. This method consists of two main stages. In the first stage, we execute the double-opponency calculation. First, convert the original image from RGB space to LMS space. Next, carry out the single-opponency response in a single RF through a linear combination of the signals followed by a Gaussian convolution. Then, two RFs with different sizes work together to constitute the double-opponency response. In the second stage, we perform the colour constancy transformation. First, estimate the underwater light through an appropriate pooling scheme. Then, correct the colours according to the model of *Von Kries*. An optional contrast enhancement stage can be added. This stage only adjusts the luminance contrast of the image and then fuses the result with colour components of the colour-corrected image.

Experiments on the UISI dataset and other images captured from real underwater scenes show that our proposed method could reduce the influence of the light absorption of water, the light scattering and other image degrading factors, resulting in corrected colour perception and enlarged brightness distribution. Meanwhile, our method presents significant advantages, with comparable performances on the four quality indicators and superior performance on time cost to other typical traditional image restoration or enhancement methods.

Further work is still needed. For example, when large texture-free areas with approximately uniform or graded colour and brightness appear in an original image, our method might introduce stepped texture or speckle noises in such areas. We should study on how to eliminate the pseudo texture phenomenon. In addition, the realisation of our method made a hypothesis that an underwater image had either a uniformly distributed single light source or a uniform background light. However, in the real world, the light is often uneven, and sometimes there are multiple light sources. As a result, the colour bias may not be uniform in the image. Therefore, another future work would be implementing a region-adaptive light estimation to achieve better correction of non-uniform colour bias.

ACKNOWLEDGEMENTS

The authors thank Prof. Xianping Fu from Information Science and Technology College, Dalian Maritime University, for providing real-world underwater images and videos.

CONFLICT OF INTEREST STATEMENT

None.

DATA AVAILABILITY STATEMENT

Some or all data, models, or code generated or used during the study are available from the corresponding author by request.

ORCID

Bin Kong https://orcid.org/0000-0002-7678-3123
Jing Qian https://orcid.org/0000-0002-6885-0090
Pinhao Song https://orcid.org/0000-0002-3950-0403

REFERENCES

1. Yahya, M.F., Arshad, M.R.: Image-based visual servoing for docking of an autonomous underwater vehicle. In: 2017 IEEE 7th International Conference on Underwater System Technology: Theory and Applications, pp. 18–20 (2017)
2. Zhang, S., et al.: Enhanced visual perception for underwater images based on multistage generative adversarial network. The Visual Computer, 86–95. (2022)
3. Dai, C.G., et al.: Color compensation based on bright channel and fusion for underwater image enhancement. Acta Opt. Sin., 86–95 (2018)
4. Mohamed, N.M.A., et al.: Underwater image quality: enhancement and evaluation. In: 2020 Cross Strait Radio Science & Wireless Technology Conference, pp. 13–16 (2020)
5. Gong, K., Hua, D.: Research on the method of color compensation and underwater image restoration based on polarization characteristics. In: 2022 3rd International Conference on Computer Vision, Image and Deep Learning & International Conference on Computer Engineering and Applications (CVIDL & ICCEA), pp. 746–751 (2022)
6. Cheng, Y.F., et al.: Underwater Video Image Clarity Method, pp. 668–677. Journal of Jilin University (Engineering Edition) (2019)
7. Hou, W.L., et al.: Automated underwater image restoration and retrieval of related optical properties. In: IEEE International Geoscience and Remote Sensing Symposium, pp. 1889–1892 (2008)
8. Liu, Y.B., et al.: Underwater image dehazing using the color space dimensionality reduction prior. In: 2020 IEEE International Conference on Image Processing, pp. 25–28 (2020)
9. Wu, M.W., et al.: Underwater image restoration using color correction and non-local prior. In: OCEANS 2017, pp. 19–22. Aberdeen (2017)
10. Cai, C.D., et al.: Underwater image restoration based on scene depth estimation and white balance. In: Progress in laser and optoelectronics, pp. 137–144 (2019)
11. Guo, J.C., et al.: Research progress of underwater image enhancement and restoration methods. Chinese Journal of Image Graphics, 273–287 (2017)
12. Li, L., Wang, H.G., Liu, X.: Underwater image enhancement based on improved dark channel prior and color correction. Acta Opt. Sin., 121–129 (2017)
13. Wang, Y., et al.: A deep CNN method for underwater image enhancement. In: 2017 IEEE International Conference on Image Processing, pp. 1382–1386 (2017)
14. Li, C.Y., et al.: An underwater image enhancement benchmark dataset and beyond. IEEE Trans. Image Process. 29, 4376–4389 (2020). https://doi.org/10.1109/tip.2019.2955241
15. Fairweather, A.J.R., Hodgetts, M.A., Greig, A.R.: Robust scene interpretation of underwater image sequences. In: Sixth International Conference on Image Processing and its Applications, pp. 660–664 (1997)
16. Chambah, M., et al.: Underwater color constancy: enhancement of automatic live fish recognition proceedings of the SPIE 5293. Color Imaging IX: Processing, Hardcopy, and Applications, 157–168 (2004)
17. Torres-Méndez, L.A., Dudek, G.: Color correction of underwater images for aquatic robot inspection. In: Proceedings of the 5th International Conference on Energy Minimization Methods in Computer Vision and Pattern Recognition, pp. 60–73 (2005)
18. Iqbal, K., et al.: Enhancing the low quality images using unsupervised Colour correction method. In: 2010 IEEE International Conference on Systems, Man and Cybernetics, pp. 1703–1709 (2010)
19. Singh, K., Kapoor, R., Sinha, S.K.: Enhancement of low exposure images via recursive histogram equalization algorithms. In: Optik-International Journal for Light and Electron Optics, pp. 2619–2625 (2015)
20. Shi, D., et al.: Underwater image enhancement algorithm based on Contourlet transform and multi-scale Retinex. In: Laser & Optoelectronics Progress, pp. 41–45 (2010)
21. Zhang, S., et al.: Underwater image enhancement via extended multi-scale Retinex. In: Neurocomputing, pp. 1–9 (2017)
22. Lan, G.N., Li, J., Ji, F.: Underwater image backscatter noise reduction based on wavelets. In: Ocean Technology, pp. 43–47 (2010)
23. Singh, G., et al.: Underwater image/video enhancement using wavelet based color correction (WBCC) method. In: Proceedings of 2015 IEEE Underwater Technology, pp. 1–5 (2015)
24. Vasamsetti, S., et al.: Wavelet based perspective on variational enhancement technique for underwater imagery. In: Ocean Engineering, pp. 88–100 (2017)
25. He, K.M., Sun, J., Tang, X.O.: Single image haze removal using dark channel prior. In: IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 2341–2353 (2011)
26. Chiang, J.Y., Chen, Y.C.: Underwater image enhancement by wavelength compensation and dehazing. In: IEEE Transactions on Image Processing, pp. 1756–1769 (2012)
27. Galdran, A., et al.: Automatic red-channel underwater image restoration. J. Vis. Commun. Image Represent. 26, 132–145 (2015). https://doi.org/10.1016/j.jvcir.2014.11.006
28. Ni, J.Y., et al.: Underwater image restoration based on transmittance optimization and color temperature adjustment. In: Progress in Laser and Optoelectronics, pp. 96–103 (2017)
29. Wang, X., et al.: Combination of dark-channel prior with sparse representation for underwater image restoration. J. Electron. Inf. Technol., 264–271 (2018)
30. Schechner, Y.Y., Karpel, N.: Clear underwater vision. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 89–97 (2004)
31. Trucco, E., Olmos-Antillon, T.: Self-tuning underwater image restoration. IEEE J. Ocean. Eng. 31(2), 511–519 (2006). https://doi.org/10.1109/joe.2004.836395
32. Xie, H.L., et al.: Underwater image restoration based on background light estimation and dark channel prior. Acta Opt. Sin., 1–10 (2018)
33. Bazeille, S., et al.: Automatic underwater image pre-processing. In: Proceedings of the Caracterisation du Milieu Marin, pp. 1–7 (2006)
34. Ancuti, C., et al.: Enhancing underwater images and videos by fusion. In: Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition, pp. 81–88 (2012)
35. Li, C., Guo, J.: Underwater image enhancement by dehazing and color correction. J. Electron. Imag. 24(3), 23–33 (2015). https://doi.org/10.1117/1.jei.24.3.033023
36. Sun, X., Liu, L., Dong, J.: Underwater image enhancement with encoding-decoding deep CNN networks. In: 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation, pp. 1–6 (2017)
37. Ueda, T., Yamada, K., Tanaka, Y.: Underwater image synthesis from RGB-D images and its application to deep underwater image restoration', 2019 IEEE International Conference on Image Processing, Taipei, 2019, pp. 2115–2119
38. Li, C.Y., et al.: An underwater image enhancement benchmark dataset and beyond. IEEE Trans. Image Process. 29, 4376–4389 (2020). https://doi.org/10.1109/tip.2019.2955241
39. Li, C.Y., et al.: Underwater image enhancement via medium transmission-guided multi-color space embedding. IEEE Trans. Image Process. 30, 1–16 (2021). https://doi.org/10.1109/tip.2021.3076367
40. Abunaser, A., et al.: Underwater image enhancement using particle swarm optimization. J. Intell. Syst. 24(1), 99–115 (2015). https://doi.org/10.1515/jisys-2014-0012
41. Gao, S.B., et al.: Color constancy using double-opponency. In: Pattern Analysis and Machine Intelligence, pp. 1973–1985 (2015)

42. Reinhard, E., et al.: Color transfer between images. IEEE Compute. Graph, 34–41 (2001)
43. von Kries, J.: Influence of adaptation on the effects produced by luminous stimuli. In: MacAdam, D. (ed.) Sources of Color Vision, pp. 109–119. MIT Press, Cambridge (1970)
44. Liu, R., et al.: Real-world underwater enhancement: challenges, benchmarks, and solutions. In: IEEE Transactions on Circuits and Systems for Video Technology, pp. 4861–4875 (2020)
45. URPC2019 Dataset, https://github.com/mousecpn/DG-YOLO, accessed 11 November 2020
46. Yang, M., Sowmya, A.: An underwater color image quality evaluation metric. In: IEEE Transactions on Image Processing, pp. 6062–6071 (2015)
47. Panetta, K., Gao, C., Agaian, S.: Human-visual- system-inspired underwater image quality measures. IEEE J. Ocean. Eng. 41(3), 541–551 (2016). https://doi.org/10.1109/joe.2015.2469915
48. Xu, X.Z., et al.: Color Cast Detection and Color Correction Methods Based on Image Analysi, pp. 10–13. Measurement&Control Technology (2008)
49. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features for image classification. IEEE Trans. Syst, Man, Cybern. SMC- 3(6), 610–621 (1973). https://doi.org/10.1109/tsmc.1973.4309314
50. Marques, T.P., Albu, A.B.: L^2UWE: a framework for the efficient enhancement of low-light underwater images using local contrast and multi-scale fusion. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–10 (2020)

How to cite this article: Kong, B., et al.: Underwater image clarifying based on human visual colour constancy using double-opponency. CAAI Trans. Intell. Technol. 1–17 (2023). https://doi.org/10.1049/cit2.12260