IET Image Processing

Research Article

IET Journals
The Institution of Engineering and Technology

Depth estimation for underwater images from single view image

ISSN 1751-9659
Received on 2nd December 2019
Revised 12th November 2020
Accepted on 13th November 2020
doi: 10.1049/iet-ipr.2019.1533
www.ietdl.org

Jarina Raihan A[1] ✉, *Pg Emeroylariffion Abas*[1], *Liyanage C De Silva*[1]

[1]*Faculty of Integrated Technologies, Universiti Brunei Darussalam, Brunei Darussalam*

✉ *E-mail: jari3010@yahoo.in*

Abstract: Underwater images often undergo distortions from scattering, absorption, colour loss, diffraction, polarisation, and varying attenuation depending on the light frequency, due to the water medium. This study answers problems associated with the recovery of underwater images, by developing novel depth map estimation which may be used as an intermediary step for underwater image restoration. The depth map is an important factor for the recovery of the underwater image, as it has been shown that proper estimation of depth, results in better restoration of the underwater image. The proposed algorithm estimates a depth map from a single view image, using blurriness and lighting information, obtained using a simple background neutralisation method. Results from the algorithm on selected raw underwater image data set have been compared with other algorithms and it has been shown that the proposed method gives Pearson coefficients in the range of 0.75–0.95, which are considerably higher and better log root mean square value compared to other methods in the literature. This suggests a better estimation of depth by the proposed method.

1 Introduction

More than 71% of our planet is covered with water and about 95% of that proportion is made up of the oceans. However, only a very small fraction of this vast extent has been thoroughly explored; due to its unfriendly and dangerous environment, with extreme pressures and absence of light. Consequently, most of the underwater world is shown to humans through still pictures, videos and through words of mouth from professional underwater divers. Underwater images are also useful in many fields such as marine biological studies, underwater archaeology and underwater mineral exploration, to examine seafloor terrain and underwater pollution, as well as to examine underwater structures such as underground bridges, dams and pipeline. An increasing number of real-world applications such as fisheries, environmental and structural monitoring inspections, and oil and gas exploration, also requires knowledge of the underwater environments.

Restoration of underwater images is essential to obtain accurate and useful information from these images. In general, computer vision algorithms assume that the medium does not affect light propagation. However, this assumption does not hold in scattering media such as underwater scenes, with scattering and absorption affecting the propagation of light and degrading the quality of the captured images. Thus, research efforts in the field of image processing and computer vision are becoming increasingly important as more applications requiring underwater image processing, are developed.

Although both the visible light and infrared (IR) bands have been used in general photography, predominantly only the visible light spectrum has been used in underwater imaging [1]. This is because IR light is more strongly absorbed by water than visible light, and consequently, the amount of information captured using the IR band is comparatively less. Furthermore, cameras utilising the visible light are more common, and hence, relatively cheaper. Depending on the depth and light penetration, ocean water may be categorised into Euphotic, Disphotic, and Aphotic zones. The upper part of the ocean, which is exposed to a relatively strong amount of sunlight, is referred to as the Euphotic zone. In the two poles, this zone may be limited to a few metres depth, whereas in brightly lit areas along the equators, it may extend to a depth of 80 m. This layer is followed by the Disphotic zone; extending to a depth of 800 m in clear water, and below this zone is the Aphotic zone, which receives no sunlight and hence, in total darkness. In the uppermost Euphotic zone, flickers influence underwater images. This is due to the effect of direct sunlight on particles as well as absorption by floating micro-organisms. In the middle disphotic zone, underwater images are affected by Rayleigh and back scatterings as well as diffraction, whereas in the deepest zone, common issues include colour loss due to the total absence of sunlight and effect from benthic features. The different distortions affecting underwater images are depicted in Fig. 1 [2]. It can be seen that as the camera focuses on a target, object radiance and veiling light may be affected by particles in the water as well as by the various type of scatterings, before arriving at the camera. Similar distortions may be experienced due to artificial light, which is commonly used for underwater image capturing.

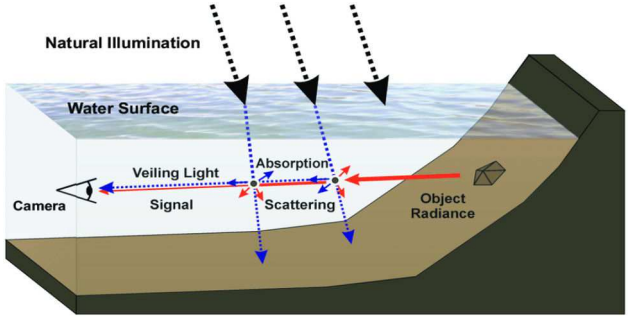

Fig. 1 *Types of distortion in underwater images, as of Garcia et al. [2]*

Depth map, defined as the distance of objects from the camera, of an underwater image needs to be determined in order to extract geometrical information of objects in the image, as well as to obtain the third dimension from a two-dimensional image. The depth cue of the scene acts as an important factor in determining the position of objects in an underwater image. Combined with other factors, the final restoration of the underwater image may be performed. Restoration of underwater images presents a challenge, in which many researchers are actively working on. Jarina Raihan *et al*. [3] provided a literature review of the different methods that have been proposed for underwater image restoration, classified according to the different approaches including hardware-, network-, and software-based approaches.

Examples of hardware approach are using polarisers, remotely operated vehicles and range-gated imaging. Some underwater images are also taken using a stereo camera, with the double lens to

capture two images of the objects, from the right and left perspective. In such case, true distance (TD) depth map may be generated based on pixel values of the left- and right-angled images. This information can then be used to restore the underwater images.

Some authors have proposed methods using the network-based approach. Anwar *et al.* [4] proposed convolutional neural network (CNN) by using end-to-end data-driven training mechanism, utilising generated synthetic underwater images from ten different marine image databases while taking into account the various optical properties and conditions of the underwater environment. Other authors have also proposed underwater restoration methods based on CNN. Hu *et al.* [5] proposed a CNN-based restoration method using two networks transmission estimation (T-Network) and global ambient light estimation network (A-Network), whereas Wang *et al.* [6] developed UIE-Net, a CNN-based method which considers colour correction and haze removal simultaneously. CNN using joint restoration filters have been proposed by Lu *et al.* [7]. Authors in literature works [8, 9] propose a generative adversarial network for image restoration. However, all their training data relies on purely synthetic data due to a problem in acquiring reliable data set. Consequently, this results in reduced quality of simulation models and hence, gives poor results considering the attenuation priors.

Software-based approach, generally, relies on the image formation model (IFM), by estimating transmission map and background lights. One of the most well-known methods is the dark channel prior (DCP); originally proposed by He *et al.* [10], by assuming scene points closer to the camera as dark images and vice versa, for estimating transmission map and background light. However, since, red light has a longer wavelength and has faster attenuation, DCP always ends up choosing the red channels as the darkest channel. Consequently, many variations of DCP have been proposed in the literature by using different priors, such as by considering green and blue channels only (DCP_{gb}) [11–13], and by considering inverse red channel with blue and green channels $DCP_{r'gb}$ [14], for prior information. Carlevaris-Bianco *et al.* [15] proposed maximum intensity prior (MIP), which considers the difference between the maximum intensity of the red channel and that of the green and blue channels to estimate the transmission map. However, performances of these methods vary according to the light absorption and different lighting conditions in different underwater images, which make many exceptions to those priors. An alternative to estimating the transmission map and background light directly is to first estimate depth map, as shown in [16]. In fact, it has been shown that proper estimation of the depth map, gives a superior restored image as compared to estimating transmission map and background light directly. However, results from Peng *et al.* [16] have shown that there is still room for improvement in the estimation of the depth map.

This paper proposes a novel depth estimation method. In Section 2, related works on image restoration algorithm are presented, highlighting the importance of depth estimation for accurate underwater image restoration. Subsequently, an algorithm for underwater image depth estimation is proposed in Section 3. Section 4 discusses results from qualitative and quantitative analysis of the proposed method, in comparison with other methods. Finally, Section 5 concludes the paper.

2 Related work

For an underwater image, the observed intensity $I^{c}(x)$ in the colour channel $c \in \{\mathrm{R, G, B}\}$ at pixel x of the acquired image is dependent on the radiance $J^{c}(x)$ of the object, the background light B^{c} and transmission map $t^{c}(x)$, as defined by the IFM:

$$I^{c}(x) = J^{c}(x)\,.\,t^{c}(x) + (1 - t^{c}(x))\,.\,B^{c},\ c \in \{\mathrm{R, G, B}\} \tag{1}$$

In the above equation, $J^{c}(x)\,.\,t^{c}(x)$ describes the radiance $J^{c}(x)$ of the object as it travels through the underwater medium, whereas $(1 - t^{c}(x))\,.\,B^{c}$ represents the scattering of background light B^{c} as it travels towards the camera. Transmission map $t^{c}(x)$ describes the part of the object radiance that reaches the camera, after considering absorption and scattering. As such, it is dependent on the object's distance from the camera $d(x)$ or its depth, and the water attenuation coefficient β_{c} in the colour channel $c \in \{\mathrm{R, G, B}\}$

$$t^{c}(x) = \mathrm{e}^{-\beta_{c}\, d(x)},\ c \in \{\mathrm{R, G, B}\} \tag{2}$$

Alternative and recently proposed IFM also exists, such as the revised model proposed by Akkaynak and Treibitz [17]. However, less emphasis is put on the variations of IFM as this paper primarily focuses on depth estimation method, which is not influenced by the choice of IFM.

From (1) and (2), it can be seen that radiance $J^{c}(x)$ of the object is attenuated exponentially with depth and water type. Recovering the original object radiance $J^{c}(x)$ from the acquired image $I^{c}(x)$ at the camera requires knowledge of the background light B^{c} as well as the transmission map $t^{c}(x)$. This information is commonly estimated.

DCP was initially introduced by He *et al.* [10], for the recovery of hazy images. However, it has since been used for estimating both B^{c} and $t^{c}(x)$ for underwater image applications [18–20]. The method relies on the assumption that dark pixels in an image are those that are close to the camera; as they get less brightening effect, whereas bright pixels are those that are far-away from the camera. The DCP for an underwater image is given by

$$I^{c}_{\mathrm{dark}}(x) = \min_{y \in \Omega(x)} \left\{ \min_{c} I^{c}(y) \right\},\ c \in \{\mathrm{R, G, B}\} \tag{3}$$

where $\Omega(x)$ is the square patch centred at x.

Background light and transmission map of the acquired image $I^{c}(x)$ may then be estimated using (4) and (5); which are then used to determine radiance $J^{c}(x)$ of the object, using (6).

$$\widetilde{B^{c}} = I^{c}\left(\mathrm{argmax}_{x \in P_{0.1\%}} \sum_{c} I^{c}(x)\right),\ c \in \{\mathrm{R, G, B}\} \tag{4}$$

$$\widetilde{t^{c}}(x) = 1 - \min_{y \in \Omega(x)} \left\{ \min_{c} \frac{I^{c}(y)}{B^{c}} \right\},\ c \in \{\mathrm{R, G, B}\} \tag{5}$$

$$\widetilde{J^{c}}(x) = \frac{I^{c}(x) - \widetilde{B^{c}}}{\max\left(\widetilde{t^{c}}(x), t_{0}\right)} + \widetilde{B^{c}},\ c \in \{\mathrm{R, G, B}\} \tag{6}$$

where the lower bound t_{0} for $\widetilde{t^{c}}(x)$, is set to 0.1. Estimated transmission map $\tilde{t}(x)$ may be refined further using filtering [21], soft matting [22] etc. Depth map of the acquired image may be estimated using (2), by assuming certain water attenuation coefficient β_{c} value, which largely depends on water type.

Chiang and Chen [20] use DCP for estimating transmission map and background light, for underwater image restoration; using fixed attenuation coefficient measured for open ocean water. However, since DCP is based on RGB channels, restoration using the method on an underwater scene often ends up considering only the red channel for transmission map and background light estimations. This is because the red channel has the lowest transmission underwater. Consequently, the DCP method often gives erroneous depth estimation and poor restoration results. Furthermore, DCP method cannot handle the introduction of artificial lighting onto underwater images.

As a result, numerous variations of the DCP method have been proposed in the literature. In order to eliminate the problems caused by artificial lighting, Lu *et al.* [23] considered only dual-channel; red and blue channels, for estimation of the transmission map while using the same fixed water attenuation as in [20] to recover the image. Galdran *et al.* [14] worked with the effect of artificial light using red-saturation prior, by retrieving the information from the inverse of the red channel. This prior is then used to estimate background light and transmission map.

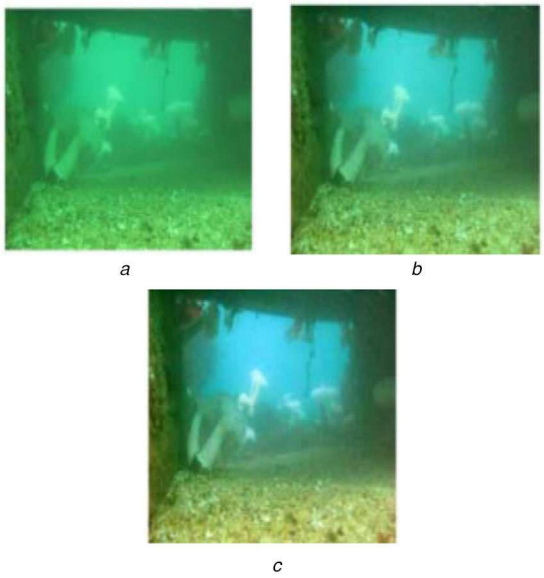

Fig. 2 *Examples of transmission map estimations, with and without depth estimation*
(a) Original image, ***(b)*** Its restored images obtained using DCP where transmission map estimation does not consider depth map, ***(c)*** Its restored image using depth map estimation, as of Peng *et al*. [16]

The red channel is neglected for transmission estimation by Drews *et al.* [12] and Wen *et al.* [11], by considering as prior $p(x)$, the green and blue channels only. Another single image de-hazing method for underwater images has been suggested by Emberton *et al.* [13], by ranking the regions of the image hierarchically to find the most likely veiling light regions in an image. Background regions are subsequently located using a superpixel segmentation and clustering process, with transmission values adapted to avoid oversaturation and artefacts. To avoid problems with the spectral dependency of water attenuation, water type is classified into blue, turquoise, or green-dominated. Ancuti *et al.* [24] used colour transfer strategy, to manipulate colour values of an input image so as to make an output image to share the appearance of a reference image. This approach preserves the salient regions of the input image as well as the subtracted details of the edge preserve smoothed filter version of the same image. By using the McGlamery underwater IFM with traditional DCP, colour transferred input is restored. However, the algorithms lack accuracy in choosing the proper reference image. Berman *et al.* [25] proposed a method for estimating transmission map, by considering different water type profiles. Multiple restored images are generated, with a final restored image chosen based on colour distribution.

Instead of single image transmission estimation, dual or multiple image transmission estimation method has also been used in literatures [26–28]. Images of the same scene are taken from different viewpoints to estimate attenuation coefficient, which are then considered for final restoration. These methods have the capability to produce good restoration results, however, collecting images of the same scene from different viewpoints is very difficult, especially underwater. Although this can be performed, albeit with much efforts, for static scene, it is almost impossible for mobile objects.

The image restoration methods mentioned above restore an underwater image by estimating background light and transmission map. Alternatively, the depth map of the underwater image may be estimated and used to determine the corresponding transmission map. As natural illumination undergoes a strong colour-dependent attenuation, which violates the assumption of wavelength-independent attenuation, Peng *et al.* [16] proposed an algorithm using blurriness information for depth map estimation. This depth map estimation is then used to find a transmission map of the underwater image using (2), by selecting attenuation coefficients β_c as defined by Zhao *et al.* [29]. The authors have subsequently shown that using depth estimation to determine the transmission

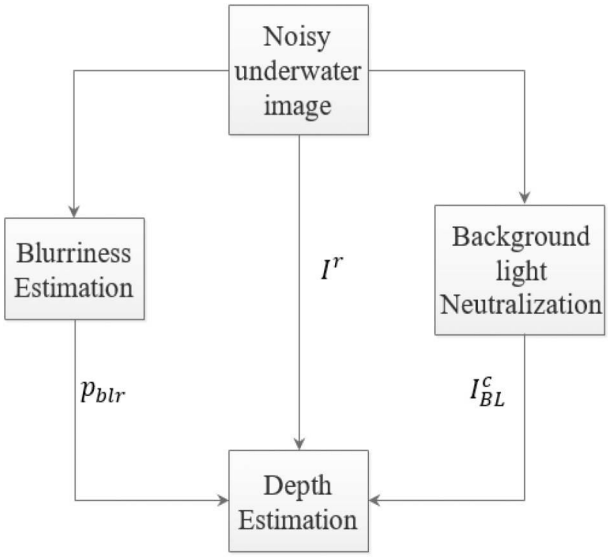

Fig. 3 *Flow chart of the proposed algorithm*

map produces better output than estimating transmission map directly without depth estimation, as shown in Fig. 2. This works well even for artificial lighting conditions.

On the other hand, other researchers have resorted to neural networks to solve the problem. Cao *et al.* [30] proposed a pixel-wise depth estimation method using convolutional networks to estimate efficient depth map from single images. A joint transmission and dehazing process have been proposed by Zhang *et al.* [31], using deep networks to eliminate the constant atmospheric light model and using a two-step prior-based method. Both literatures [30, 31] are, however, generalised methods and are not dedicated to underwater images which may make their performance lower for specific applications of underwater images.

3 Proposed methodology

In this section, a depth estimation method that involves the use of blurriness information and background light neutralisation, for estimating depth $\tilde{d}(x)$, which is the distance of the object from the camera, has been proposed. In particular, a simple background light neutralisation method has been proposed, which requires only segmentation of the input image into four quadrants. Fig. 3 depicts the flowchart of the proposed method, involving (i) blurriness estimation, (ii) background light neutralisation, and (iii) depth estimation; using blurriness estimation, background light neutralisation and the red channel of the original image. Output from the depth estimation method shall be analysed, quantitatively and qualitatively, in the following section.

The red-light intensity $I^r(x)$, the blurriness estimated image p_{blr} and the background light neutralised image B^C are calculated from the input image $I^c(x)$. These are then used to estimate depth $\tilde{d}(x)$ of the underwater image. This is in contrast to the work in [16], which uses background light selection process, based on blurriness and variance selection. Furthermore, MIP has also been utilised by Peng *et al.* [16] as part of their final depth map selection process, which is absent in the proposed method, making the proposed depth estimation method simpler to implement.

3.1 Blurriness estimation

Blurriness estimation [16] involves the conversion of the original input image into grey-level image, which is then filtered using Gaussian filter to get the initial map. Max filter is then used to convert the initial map to rough map. The initial map $p_{\text{init}}(x)$ and rough map $p_r(x)$ are calculated as follows:

$$p_{\text{init}}(x) = I^g(x) - G(x) \tag{7}$$

$$p_{\mathrm{r}}(x) = \max_{y \in \Omega(x)} p_{\mathrm{init}}(y) \tag{8}$$

where $I^{\mathrm{g}}(x)$ is the grey-level image of the original input image $I^{\mathrm{c}}(x)$ and $G(x)$ is the 2×2 spatial Gaussian filter with variance 2. $\Omega(x)$ is a square local patch centred at x.

The rough map is then refined by hole filling morphological operator [32] and guided filter [21], to smoothen the rough map and give the blurriness map, p_{blr}. This blurriness map p_{blr} is further used for depth estimation of the image

$$p_{\mathrm{blr}} = F_{\mathrm{g}}\{C_{\mathrm{r}}(p_{\mathrm{r}}(x))\} \tag{9}$$

where C_{r} is a hole filling morphological operator and F_{g} is a guided filter.

3.2 Background light neutralisation

An underwater image with scene radiance $\widetilde{J^{\mathrm{c}}}(x)$, as shown in (3) is bounded between (0,1). Restoring an underwater image with dim background light would result in bright scene radiance while bright background light gives the opposite result. To determine the background light, the input image $I^{\mathrm{c}}(x)$ is segmented into four quadrants, which represents the least number of quadrant sub-divisions that can be made, and hence, simpler, more efficient, with reduced computational processing time. Output of the resulting algorithm shall then be compared with methods utilising quad-tree decomposition [16] and iterative decomposition process [10, 14], commonly used in traditional underwater image restoration methods. The average value of the pixels is calculated for each quadrant, as in (10), in order to determine average light intensity in the four quadrants

$$I^{\mathrm{c}}_{q_i} = \mathrm{avg}_{x \in q_i}\left(\sum_{\mathrm{c}} I^{\mathrm{c}}(x)\right) \tag{10}$$

where q_i is the four quadrants $i = 1,\ 2,,3,\ 4$, with $I^{\mathrm{c}}_{q_i}$ representing the average light intensity in the respective quadrant q_i. The brightest $q_{\max}$ and darkest $q_{\min}$ quadrants are neglected, as two extremes of the spectrum. Average light intensity in the remaining two quadrants is then calculated, and taken as average of the underwater image

$$I^{\mathrm{c}}_{q_{\mathrm{avg}}} = \mathrm{avg}_{q_i \in q_{\mathrm{mid}}}(I^{\mathrm{c}}_{q_i}),\ q_{\mathrm{mid}} \in \left\{q_{i=\{1,2,3,4\}} - q_{\max} - q_{\min}\right\} \tag{11}$$

$I^{\mathrm{c}}_{q_{\mathrm{avg}}}$ is used to modify all the pixels of the input image $I^{\mathrm{c}}(x)$ to retrieve the contrast neutralised image $I^{\mathrm{c}}_{\mathrm{cn}}(x)$, as follows:

$$I^{\mathrm{c}}_{\mathrm{cn}}(x) = I^{\mathrm{c}}(x) + I^{\mathrm{c}}_{q_{\mathrm{avg}}} \tag{12}$$

In order to denoise the newly formed contrast neutralised image, discrete wavelet transform (dwt) is performed on the newly formed contrast neutralised image $I^{\mathrm{c}}_{\mathrm{cn}}(x)$ and the grey version of the input image $I^{\mathrm{g}}(x)$. Three levels of filter banks are used for the dwt

$$\mathrm{dwt}(I^{\mathrm{c}}_{\mathrm{cn}}(x)) = a^{\mathrm{c}}(x) + \sum_{i=1}^{3} d^{\mathrm{c}}_i(x) \tag{13}$$

$$\mathrm{dwt}(I^{\mathrm{g}}(x)) = a^{\mathrm{g}}(x) + \sum_{i=1}^{3} d^{\mathrm{g}}_i(x) \tag{14}$$

where $a^{\mathrm{c}}(x)$ and $a^{\mathrm{g}}(x)$ are the approximation coefficients of the dwt of contrast neutralised image and grey image, respectively. $d^{\mathrm{c}}_i(x)$ and $d^{\mathrm{g}}_i(x)$ are the ith detailed coefficients of the dwt of contrast neutralised image and grey image, respectively.

The background light neutralised image may then be obtained by using inverse dwt, by setting its approximation coefficient based on the average of the approximation coefficients of the grey version of the input image and the contrast neutralised image, and its detailed coefficients based on the max rule between detailed coefficients of the grey version of the input image and the contrast neutralised image

$$I^{\mathrm{c}}_{\mathrm{BL}}(x) = \mathrm{idwt}(A^{c(x)} + \sum_{i=1}^{3} \max\left(d^{\mathrm{c}}_i(x), d^{\mathrm{g}}_i(x)\right) \tag{15}$$

where $A^{c(x)} = (a^{\mathrm{c}}(x) + a^{\mathrm{g}}(x))/2$, and idwt(.) is the inverse dwt.

3.3 Depth estimation

The estimated depth [16] map may be found, by combining three depth estimation methods using the red channel, blurriness map of the image, and background light neutralised image. These depth estimations are obtained using

$$d_{f(x)}(x) = 1 - F_{\mathrm{s}}(f(x)),\ f(x) \in \left\{r(x), p_{\mathrm{blr}}(x), I^{\mathrm{c}}_{\mathrm{BL}}(x)\right\} \tag{16}$$

where $f(x) \in \left\{r(x), p_{\mathrm{blr}}(x), I^{\mathrm{c}}_{\mathrm{BL}}(x)\right\}$ can either be the red channel map $r(x)$, blurriness map $p_{\mathrm{blr}}(x)$ or background neutralised image $I^{\mathrm{c}}_{\mathrm{BL}}(x)$, to give $d_{\mathrm{r}}(x)$, $d_{p_{\mathrm{blr}}}(x)$ and $d_{I^{\mathrm{c}}_{\mathrm{BL}}}(x)$, respectively. All $d_{r(x)}(x)$, $d_{p_{\mathrm{blr}}}(x)$ and $d_{I^{\mathrm{c}}_{\mathrm{BL}}}(x)$ need to be derived using (16) to obtain the estimated depth map. $F_{\mathrm{s}}(\boldsymbol{v})$ is a stretching function, which accepts vector $\boldsymbol{v}$ as its input, and defined in (17); with $\min(\boldsymbol{v})$ and $\max(\boldsymbol{v})$ giving the minimum and maximum value of the input vector $\boldsymbol{v}$, respectively

$$F_{\mathrm{s}}(\boldsymbol{v}) = \frac{\boldsymbol{v} - \min(\boldsymbol{v})}{\max(\boldsymbol{v}) - \min(\boldsymbol{v})} \tag{17}$$

The red channel map $r(x)$ is defined as the maximum intensity of the red channel of the image and is represented by

$$r(x) = \max_{y \in \varphi(x)} I^{\mathrm{r}}(y) \tag{18}$$

where I^{r} is the intensity of the red channel and $\varphi(x)$ is a square local patch centred at x.

Equation (19) gives the estimated depth map $\tilde{d}(x)$ of the original input image $I_{\mathrm{c}}(x)$

$$\tilde{d}(x) = \theta_b\left[\theta_a d_{I^{\mathrm{c}}_{\mathrm{BL}}}(x) + (1-\theta_a) d_{\mathrm{r}}(x)\right] + (1-\theta_b) d_{P_{\mathrm{blr}}} \tag{19}$$

where θ_b and θ_a are $\theta_a = S(\mathrm{avg}(I^{\mathrm{c}}_{\mathrm{BL}}), 0.5)$ and $\theta_b = S(\mathrm{avg}(I^{\mathrm{r}}), 0.1)$, respectively; with function avg(.) giving average value of the input and the sigma functions $S(a,v)$ given as:

$$S(a, v) = \left[1 + \mathrm{e}^{-s(a-v)}\right]^{-1} \tag{20}$$

The estimated depth map $\tilde{d}(x)$ in (19) is a combination of depth estimation using the red channel map $r(x)$, blurriness map $p_{\mathrm{blr}}(x)$ or background light neutralised image $I^{\mathrm{c}}_{\mathrm{BL}}(x)$. A few extreme cases may be highlighted as follows:

(i) when the background light of the original input image is dim with the image possessing a considerable level of red content, i.e. $\mathrm{avg}(I^{\mathrm{r}}) \gg 0.1$ and $\mathrm{avg}(I^{\mathrm{c}}_{\mathrm{BL}}) \ll 0.5$, then $d_{\mathrm{r}}(x)$ represents the estimated depth, i.e. $\tilde{d}(x) \simeq d_{\mathrm{r}}(x)$,

(ii) When the background light of the original input image is bright; maybe for the image taken in the first part of the ocean, with a considerable level of red content, i.e. $\mathrm{avg}(I^{\mathrm{r}}) \gg 0.1$ and $\mathrm{avg}(I^{\mathrm{c}}_{\mathrm{BL}}) \gg 0.5$, then $d_{I^{\mathrm{c}}_{\mathrm{BL}}}$ represents the estimated depth, i.e. $\tilde{d}(x) \simeq d_{I^{\mathrm{c}}_{\mathrm{BL}}}(x)$, and

(iii) Lastly, if the image is taken in the second or the third part of the ocean and there is very little red light in the scene with almost

dark background light, i.e. overall $\text{avg}(I_r) \ll 0.1$, then $d_{P_{\text{blr}}}$ represents the estimated depth, i.e. $\tilde{d}(x) \simeq d_{P_{\text{blr}}}(x)$.

4 Result and conclusion

The proposed method for depth estimation of underwater images, as shown in Fig. 4, is implemented, and compared to other methods. It consists of blurriness estimation, generation of background light neutralised image and red channel map of the original underwater image; which are then used to estimate depth map, as given in (19).

4.1 Blurriness estimation and background light neutralisation

Figs. 4 and 5 show the results of blurriness estimation and background light neutralisation process obtained using (9) and (15), respectively.

Initial map $p_{\text{init}}(x)$, in Fig. 4*b* is obtained from the grey-level image $I^{\text{g}}(x)$ of the original underwater image $I^{\text{c}}(x)$, in Fig. 4*a*, using (7). The rough map $p_{\text{r}}(x)$ is then obtained, before refining it further with morphological operators and passed through the guided filter, to give the blurriness map $p_{\text{blr}}(x)$, shown in Fig. 4*c*.

Fig. 5*c* shows the background light neutralised image $I^{\text{c}}_{\text{BL}}(x)$, which is generated from (15). The original underwater image and its corresponding grey-level images are depicted in Figs. 5*a* and *b*, respectively.

4.2 Depth estimation

Both the blurriness map $p_{\text{blr}}(x)$ and the background light neutralised image $I^{\text{c}}_{\text{BL}}(x)$, as well as the red channel map $r(x)$ of the original underwater image $I^{\text{c}}(x)$, are used for depth estimation using (19). Results of the proposed depth estimation are analysed qualitatively and quantitatively. The qualitative analysis relies on the visual comparison between the depth estimation obtained using the proposed method and the original underwater image. On the other hand, quantitative analysis mathematically evaluates the similarity between the depth estimation and the TD, via log value of root mean square error ($\text{RMSE}_{\log}$) and Pearson coefficient.

4.2.1 Qualitative analysis: Qualitative analysis is used as a performance comparison between estimated depth map and original underwater image, which lacks its TD information. As it relies on visual comparison, it is generally less reliable than quantitative analysis method. Various underwater images at different water conditions have been selected, in order to visually analyse the performance of the proposed depth estimation method. Considering the natural variabilities of the data chosen, the images which are used for comparison have been selected in such a way that the images have been taken at different seasons, depths, and water types (both in tropical waters and in murkier coastal waters). For instance, Figs. 6–10, represent underwater images with various water conditions. Fig. 11 shows the depth map of various algorithms, clearly showing that DCPs and MIP perform poorly, with the bright pixels are mistakenly judged to be faraway and the dark pixels are judged to be closer.

Fig. 12 shows the estimated depth maps of an underwater image, along with three estimates of $d_{r(x)}(x)$, $d_{p_{\text{blr}}}(x)$, and $d_{I^{\text{c}}_{\text{BL}}}(x)$ obtained using the proposed depth estimation method. Blue colour indicates nearby objects while fading colour which changes to yellow indicates farther objects. Estimation of the final estimated depth map $\tilde{d}(x)$ is determined using (19), with values $\text{avg}(I^{\text{c}}_{\text{BL}}) \simeq 0.376$ and $\text{avg}(I^{\text{r}}) \simeq 0.127$ for the input image shown.

Figs. 6 and 7 illustrate underwater images with various water conditions; deep water image with lighting and image with dim background, images with bluish background, with inclined foreground, with artificial lighting. Blue indicates that the object is close to the camera, while yellow indicates the object is farther from the camera. Fig. 6*a* illustrates a nice rendition of the proposed algorithm, with the hammerhead shark estimated as close to the camera; for deepwater image captured with lighting. In the final image in Fig. 6*b* with dim background, the camera was angled from the top so the upper part of the crab is indicated in blue compared to the background which is shown as yellow, and the crab has been estimated as closer compared to the rock. Generally, the proposed algorithm performs well in estimating the depth of the object, however, there is still some limitation in constructing the true depth representation, as can be visually observed from the figures.

In Fig. 7*a*, with the camera located on the lower right side, even though some sections of the upper part of the shark have been incorrectly merged with the bluish background, the front part of the shark (with blue colour) and the background water (with yellow colour) have been correctly estimated as nearer to, and far away from the camera, respectively. Similarly, the lower part of the inclined stingray has been correctly estimated as nearby, with the background as far from the camera; although the right-wing of the stingray has been incorrectly merged with the background, as shown in Fig. 7*b*. For the underwater image with artificial lights in Fig. 7*c*, objects in the picture have been inaccurately estimated to be the closest to the camera as compared to the rocks. These indicate that the estimated depth maps are influenced by colour of the objects; with lighter colour resulting in the object to be estimated as closer to the camera than it actually is, and darker object to be estimated as farther away from the camera than it actually is.

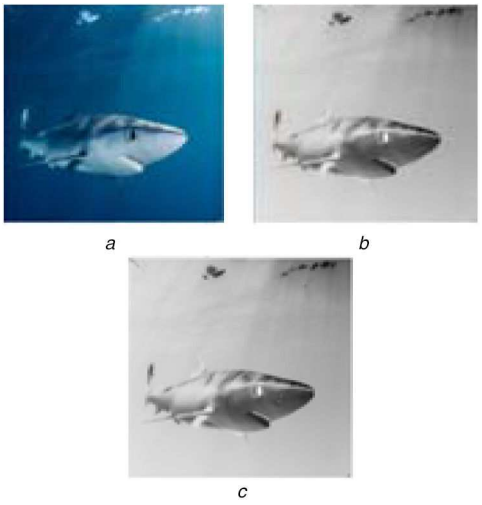

Fig. 4 *Blurriness estimation process, showing*
(a) Input image $I^{\text{c}}(x)$, ***(b)*** Initial map $p_{\text{init}}(x)$, ***(c)*** Refined blurriness map $p_{\text{blr}}(x)$

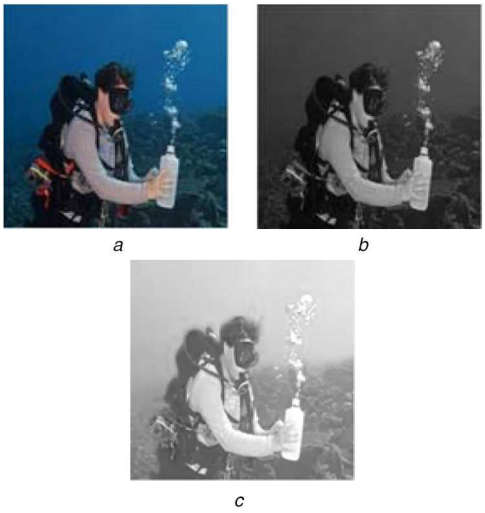

Fig. 5 *Background light neutralisation process, showing*
(a) Input underwater image $I^{\text{c}}(x)$, ***(b)*** Grey version $I^{\text{g}}(x)$ of the underwater image, ***(c)*** Background light neutralised image $I^{\text{c}}_{\text{BL}}(x)$

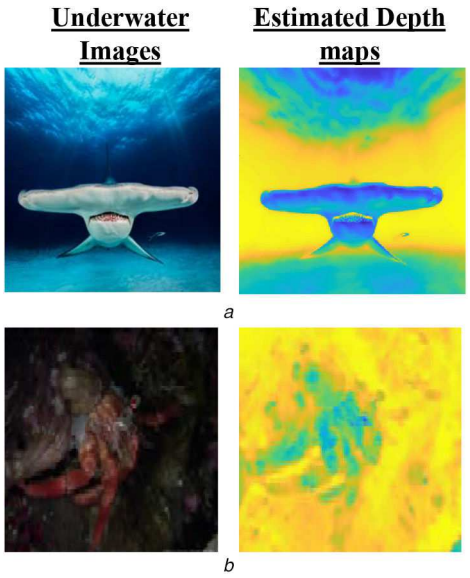

Fig. 6 *Depth map at various water conditions*
(a) Deep water with lighting, ***(b)*** Dim background image

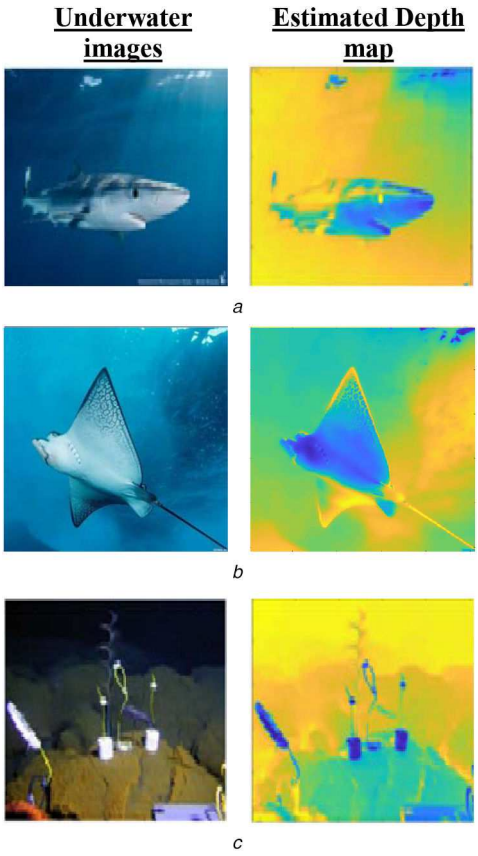

Fig. 7 *Depth map at various water conditions*
(a) Bluish background, ***(b)*** Inclined foreground, ***(c)*** Artificial lighting

Fig. 8*a* shows an underwater image of a wrecked ship covered with coral reefs. The background colour of the image has a greenish tone, with the object being horizontal to the camera. Hence, in the depth map, the bulkier part of the reefs appears bluish indicating that these are nearer to the camera, compared to the part of the ship which is yellowish in colour. Fig. 8*b* shows a cracked underwater image of a seabed with artificial lighting. The estimated depth map clearly shows three clear shades of colours: yellow, blue, and dark blue, representing different depths based on their distances.

4.2.2 Quantitative analysis: For the quantitative analysis, the resolution of images used in Figs. 9, 10, 13 and 14 are 72 dpi

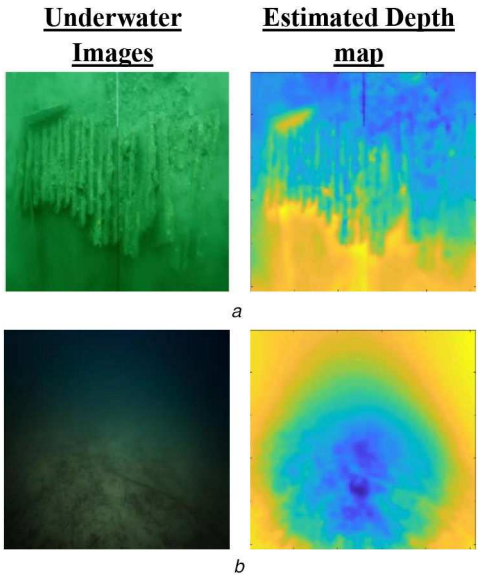

Fig. 8 *Depth map at various water conditions*
(a) Greenish background, ***(b)*** Minimal lighting

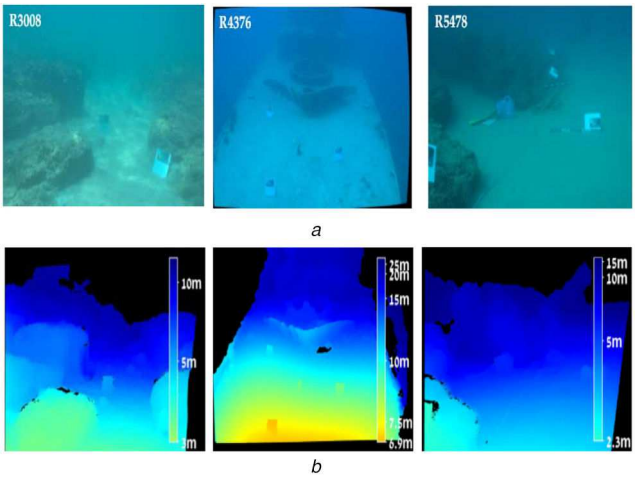

Fig. 9 *Input underwater images: R3008, R4376 and R5478, with their respective TD maps [25]*

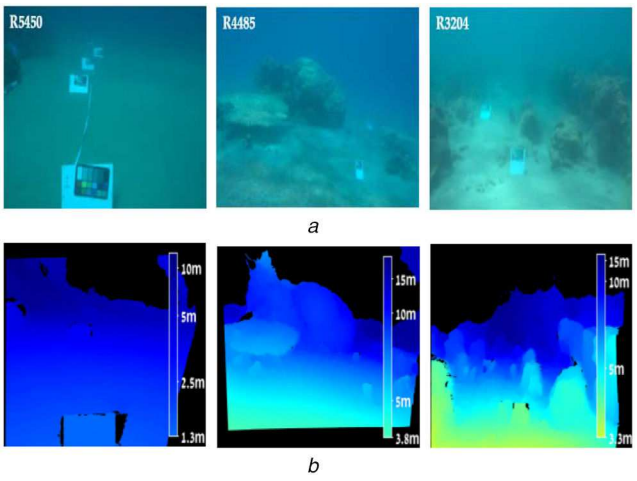

Fig. 10 *Input underwater images: R5450, R4485 and R3204, with their respective TD maps [25]*

(horizontal and vertical) and are run with the same resolution using different methods; to find the efficiency of the proposed method as compared to other state-of-the-art methods. Both the log value of $RMSE_{log}$ and Pearson coefficient have been used for quantitative

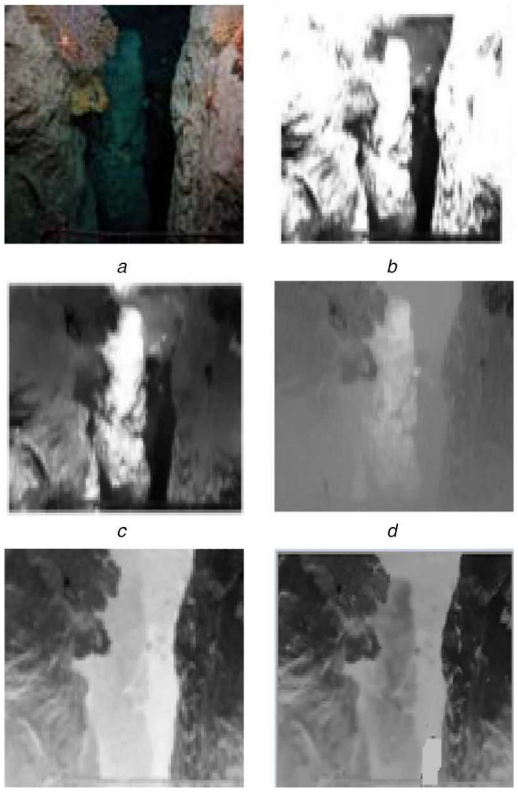

Fig. 11 *Qualitative comparison using data set from Peng et al. [16]*
(a) Input image, ***(b)*** Drews *et al.* [12], ***(c)*** Galdran *et al.* [14], ***(d)*** Carlevaris-Bianco *et al.* [15], ***(e)*** Peng *et al.*, [16], ***(f)*** Proposed method

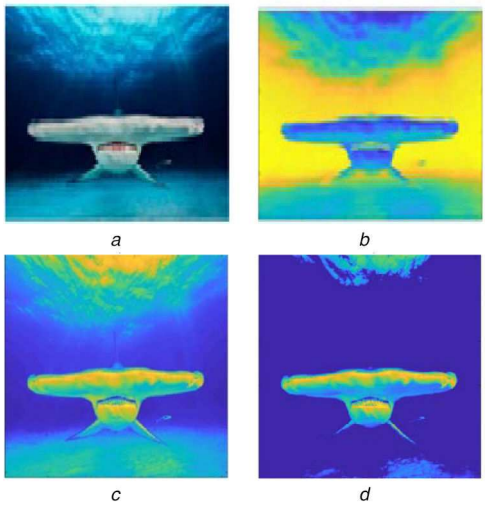

Fig. 12 *Illustration of*
(a) Input image, with estimated depth maps, ***(b)*** $d_{r(x)}(x)$, ***(c)*** $d_{p\text{blr}}(x)$, ***(d)*** $d_{\text{BL}}^{c}(x)$

analysis of the different methods. $\text{RMSE}_{\log}$ follows the below equation:

$$\text{RMSE}_{\log} = \frac{1}{T}\sum_{p}(\log d_{\text{e}} - \log d_{\text{o}} + \alpha(d_{\text{e}}, d_{\text{o}}))^2 \tag{21}$$

where d_{e} and d_{o} are estimated and original ground truth depths, respectively, at a pixel p, T denotes the total number of pixels and $\alpha(d_{\text{e}}, d_{\text{o}}) = 1/n\sum_{p}(\log d_{\text{o}} - \log d_{\text{e}})$ addresses the scale alignment factor. The lower the value of $\text{RMSE}_{\log}$, the better is the depth map estimations.

On the other hand, Pearson coefficient requires knowledge of the TD of the image for comparison with the estimated depth map of the image. TD of an underwater image may be obtained by

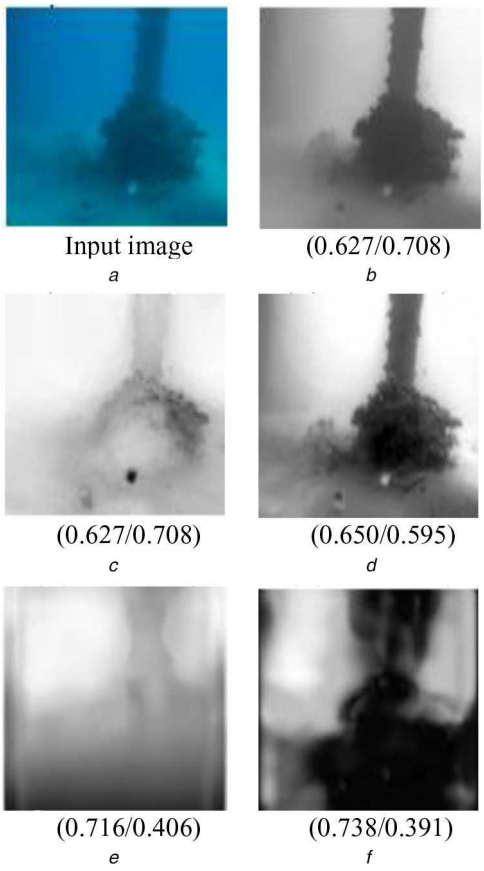

Fig. 13 *Quantitative analysis using Berman et al. [25] data set*
(a) Input image, with (Pearson coefficient/$\text{RMSE}_{\log}$) values calculated for, ***(b)*** Drews *et al.* [12], ***(c)*** Peng *et al.* [16], ***(d)*** Berman *et al.* [25], ***(e)*** Godrad *et al.* [33], ***(f)*** Godrad *et al.* [34]

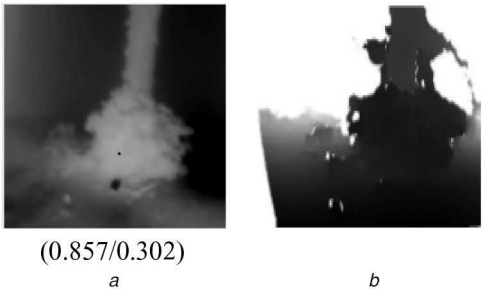

Fig. 14 *Quantitative analysis*
(a) Proposed method with (Pearson coefficient/$\text{RMSE}_{\log}$), ***(b)*** Ground truth

employing underwater stereo cameras and then using the stereo camera calibration parameters as well as the distance between stereo cameras for its calculation. All images used for the analysis are taken from Berman data set [25], which were taken using a pair of the digital single-lens reflex camera; Nikon D810 with an AF-S NIKKOR 35 mm f/1.8G ED lens, encased in Hugyfot housing with a dome port. The camera setup was set on a rigid and very stable rig, and hence, assumed to be free from any motion artefacts. Based on stereo imaging, the original distances of objects from the camera may be recovered. These are then used as a basis for quantitative evaluation of depth maps obtained from the proposed method, which utilises a single image only.

Pearson correlation coefficient ρ is calculated between TD and the estimated depth map $\tilde{d}(x)$, using:

$$\rho_{\text{TD},\tilde{d}(x)} = \frac{\text{cov}\left((\text{TD}), \tilde{d}(x)\right)}{\sigma_{\text{TD}} * \sigma_{\tilde{d}(x)}} \tag{22}$$

where cov(.) is covariance while σ_{TD} and $\sigma_{\tilde{d}(x)}$ are standard deviations of the TD and estimated depth map, respectively.

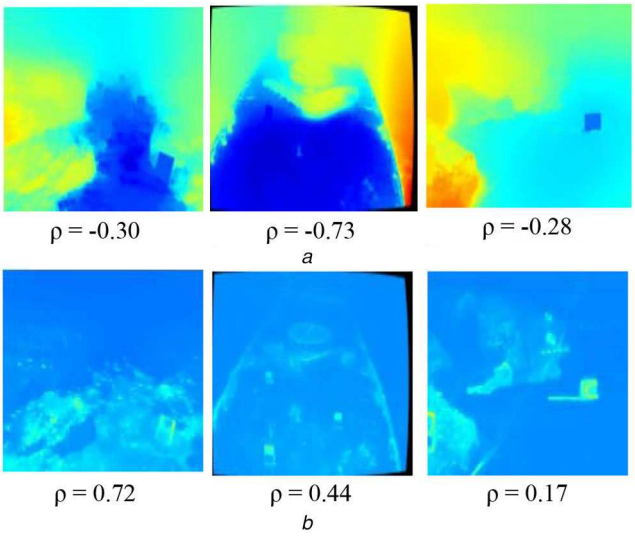

Fig. 15 *Calculated depth maps using*
(a) Drews *et al.* [12], ***(b)*** Peng *et al.* [16]

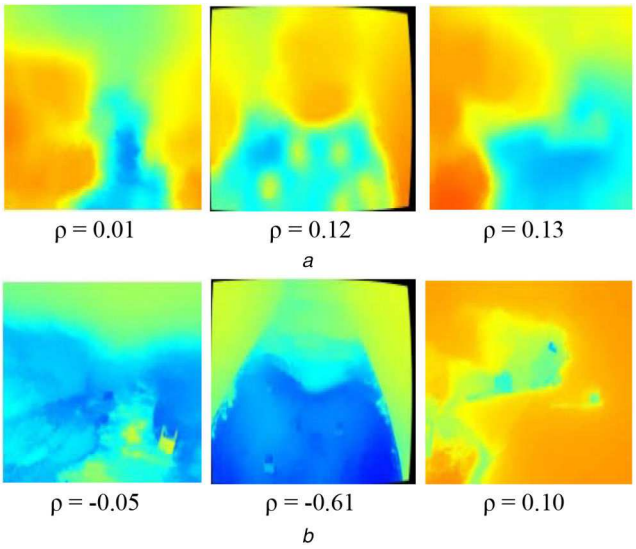

Fig. 16 *Calculated depth maps using*
(a) Ancuti *et al.* [24], ***(b)*** Emberton *et al.* [13]

Pearson coefficient takes a value between −1 and +1, with +1 indicating a perfect correlation between the estimated depth map and TD. Both low and negative Pearson coefficient values indicates incorrect estimation of the depth map.

Figs. 13 and 14 shows an input image, as well as, the depth maps estimated using state of the art methods; with their respective Pearson coefficients and $RMSE_{log}$ values given for the different methods. It can be clearly seen that depth map estimation using the proposed method, as shown in Fig. 14, gives better Pearson coefficient and $RMSE_{log}$ values as compared to other methods. Particularly, the proposed method performs better than methods proposed by Peng *et al.* [16] (which utilises quad-tree decomposition for its background light estimation stage) and Drews *et al.* [12] (which chooses brightest pixel in the red channel), He *et al.* [10], Galdran *et al.* [14] (which chooses bright pixels in dark channel for their background light estimation). This shows that using only four quadrants for the background light neutralisation stage gives better quality output, with reduced complexity.

Fig. 9 shows underwater images: R3008, R4376, and R5478 and Fig. 10 shows underwater images: R5450, R4485, and R3204, together with their respective TD maps [25]. Figs. 15–17, illustrate depth map estimations of the images in Fig. 9, using methods given in [12, 13, 16, 24, 25], as well as the proposed method. Similarly, Figs. 18–20, illustrate depth map estimations of the images in

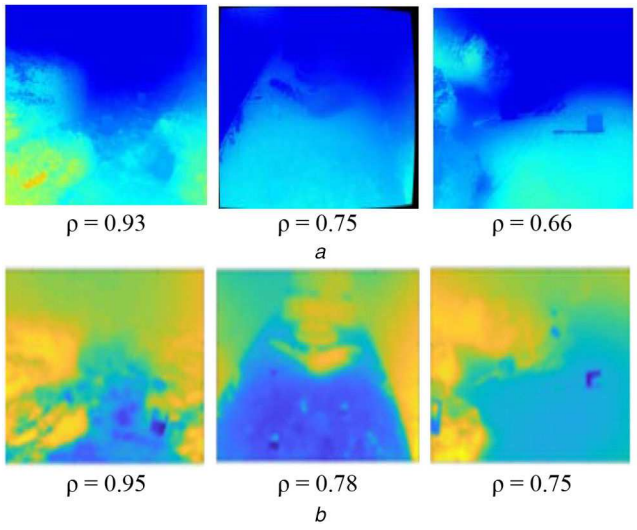

Fig. 17 *Calculated depth maps using*
(a) Berman *et al.* [25], ***(b)*** Proposed method

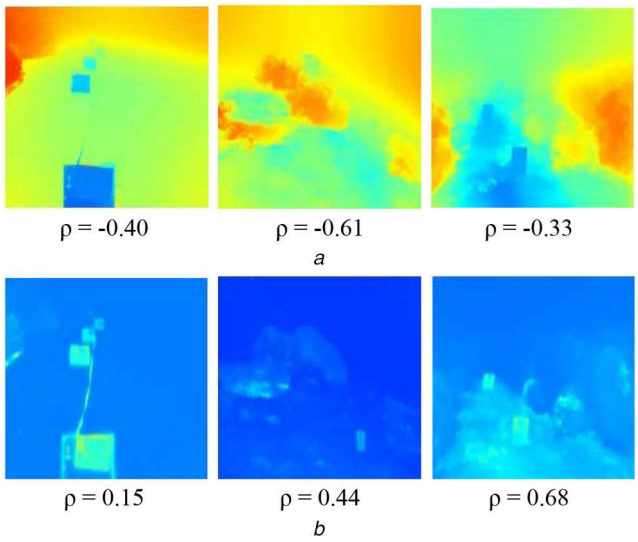

Fig. 18 *Estimated depth maps, using*
(a) Drews *et al.* [12], ***(b)*** Peng *et al.* [16]

Fig. 10 using methods given in [12, 13, 16, 24, 25], as well as the proposed method. It is noted that distance map is colour coded, changing from light yellow to dark blue, indicating increase in distance. Black indicates missing TD due to limitation of the stereo camera as well as its orientation. With the exception of Peng *et al.* [16] and our proposed method, most authors [12, 13, 24, 25] have directly estimated transmission map. In such cases, depth map estimations are accomplished using (2). Table 1 indicates Pearson coefficients for the various methods on the different underwater images. It can be observed that the proposed depth estimation method performs very well, relative to other methods; with positive Pearson coefficients between 0.75 and 0.95, indicating good estimation of depth purely from original underwater images.

For all underwater images, the proposed method gives the highest positive Pearson coefficient. Peng *et al.* [16] develop depth map estimation using blurriness information, by iteratively splitting input image into blocks based on lowest variance and largest blurriness, for background light estimation. The iterative process makes the method computationally more intensive than the proposed method. Highest Pearson coefficient using the method is 0.72; good correlation with the TD, however, for some underwater images, it gives lowly 0.15 Pearson coefficients, indicating that the estimated depth maps give incorrect representation of the TD.

Drews *et al.* [12] rely to some extent on the DCP, which has been empirically proven on outdoor scenes. However, for underwater image, the method is not able to estimate depth map of

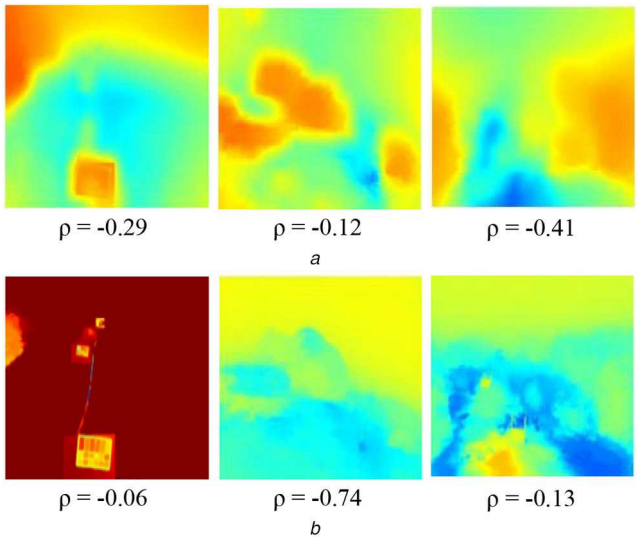

Fig. 19 *Estimated depth maps using*
(a) Ancuti *et al.* [24], ***(b)*** Emberton *et al.* [13]

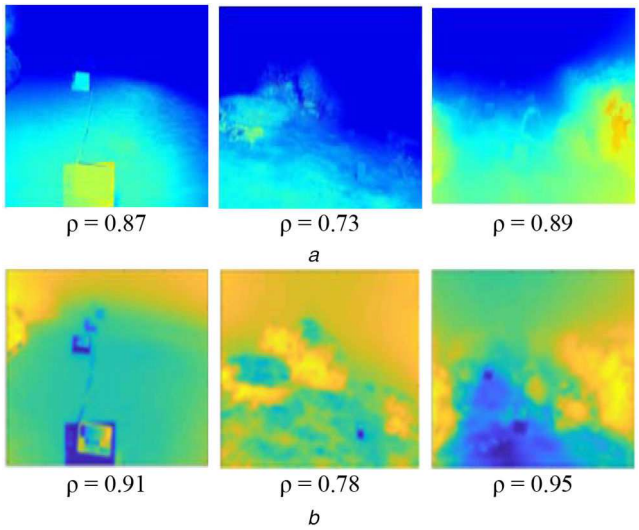

Fig. 20 *Estimated depth maps using*
(a) Berman *et al.* [25], ***(b)*** Proposed method

Table 1 Pearson coefficient value for different methods of estimating depth maps, for various underwater images

Images	Drews *et al*. [12]	Peng *et al*. [16]	Ancuti *et al*. [24]	Emberton *et al*. [13]	Berman *et al*. [25]	Proposed method
R3008	−0.30	0.72	0.01	−0.05	0.93	0.95
R4376	−0.73	0.44	0.12	−0.61	0.75	0.78
R5478	−0.28	0.17	0.13	0.10	0.66	0.75
R5450	−0.40	0.15	−0.29	−0.06	0.87	0.91
R4485	−0.61	0.44	−0.12	−0.74	0.73	0.78
R3204	−0.33	0.68	−0.41	−0.13	0.89	0.95

Table 2 Average pearson coefficients and $RMSE_{log}$ values of different methods, using images from Berman *et al*. [25] data set

	He *et al*. [10]	Drews *et al*. [12]	Berman *et al*. [25]	Peng *et al*. [16]	Godrad *et al*. [33]	Godrad *et al*. [34]	Proposed method
Pearson coefficient	0.086	0.556	0.669	0.722	0.803	0.732	0.857
$RMSE_{log}$	2.022	0.450	0.518	0.485	0.311	0.363	0.301

the image effectively, as depicted from the negative and low Pearson coefficients in the table. Both Ancuti *et al.* [24] and Emberton *et al.* [13], also perform poorly in depth estimation of underwater image with low Pearson coefficients. On the other hand, Berman *et al.* [25] estimate underwater depth well; albeit producing generally lower Pearson coefficients as compared to the proposed depth estimation methods.

Table 2 shows the average Pearson coefficient and $RMSE_{log}$ calculated on Berman *et al*. [25] data set, using different methods. It can be clearly seen that the proposed method gives the best performance, with average Pearson coefficient of 0.857 and

$RMSE_{log}$ of 0.301. DCP [10] shows poor performance, due to its constant selections of red as its darkest channel. Both methods proposed in [33, 34], give better performances than DCP, however, with the methods being not dedicated to underwater images, these methods cannot be expected to give superior performances for dedicated underwater data set. Furthermore, the methods are considered as somewhat generalised models and fine-tuning them to fit to a particular condition may result in loss of accuracy and performance. Other methods [12, 16, 25] can be seen to produce inferior performances.

To analyse the time complexity of the algorithm, running times are recorded for the different methods and the proposed method is found to be more efficient; with an average running time of 12.43 s when run on Berman *et al*. [25] data set, with Peng *et al.* [16] and Berman *et al.* [25] recording average running times of 14.94 and 18.56 s, respectively. The experiment was conducted using MATLAB R2018b on a PC with Intel(R) i5, 32 GB RAM.

5 Conclusion

It has been previously shown that estimating depth map of an underwater image is important for image restoration. Consequently, in this paper, blurriness map and background light neutralisation process have been proposed to estimate depth map for underwater images. Good visual results of the resulting depth map, have been shown, for various sea level conditions. Quantitative analysis has also been performed to measure correlation of depth map, estimated using the proposed method with the TD of the image, as well as to compare against other methods in the literature. It has been shown that the proposed depth estimation method shows high positive correlations to the TD; indicating the potential of the proposed method to estimate the depth map accurately. Pearson coefficients between 0.75 and 0.95, are obtained from the six (6) different underwater images using the proposed method. Furthermore, it has also been shown to perform much better than other methods that have been proposed in the literature. Estimated depth using the proposed method can be further used to estimate transmission map and image restoration, using (2) and (1), respectively. Naturally, the improvement in depth map obtained using the proposed method would consequently result in improvement in both transmission map and image restoration. However, further development of transmission map and image restoration methods shall be proposed as the future work, towards the whole process.

6 References

[1] Biegański, W., Kasiński, A.: 'Image acquisition in an underwater vision system with NIR and VIS illumination', *J. Comput. Sci. Inf. Technol.*, 2014, **4**, pp. 1–10

[2] Garcia, R., Nicosevici, T., Gracias, N.*, et al.*: 'Exploring the seafloor with underwater robots: land, sea & air', in Lopez, A.M., Imiya, A., Pajdla, T.*, et al.* (Eds): '*Computer vision in vehicle technology*' (John Wiley & Sons Ltd, Barcelona, 2017), pp. 75–99

[3] Jarina Raihan, A., Emeroylariffion, A., Liyanage, C.: 'Review of underwater image restoration algorithms', *J. IET Image Process.*, 2019, **13**, (10), pp. 1587–1596

[4] Anwar, S., Li, C., Porikli, F.: 'Deep underwater image enhancement', *J. Comput. Vis. Pattern Recogn.*, 2018, **10**, pp. 1–12

[5] Hu, Y., Wang, K., Zhao, X.: 'Underwater image restoration based on convolutional neural network'. Proc. Asian Conf. Machine Learning, Beijing, China, November 2018, pp. 296–311

[6] Wang, Y., Zhang, J., Cao, Y.*, et al.*: 'A deep CNN method for underwater image enhancement'. Proc. Int. Conf. Image Processing, Beijing, China, September 2017, pp. 1382–1386

[7] Lu, H., Xu, X., Kim, H.: 'Underwater light field depth map restoration using deep convolutional neural fields', in Lu, H., Xi, X., (Eds): '*Artificial intelligence and robotics*' (Springer Int. publishing press, Switzerland, 2018), pp. 305–312

[8] Fabbri, C., Islam, M.J., Sattar, J.: 'Enhancing underwater imagery using generative adversarial networks'. Proc. Int. Conf. Robotics and Automation, Brisbane, Australia, May 2018, pp. 7159–7165

[9] Li, J., Skinner, K.A., Eustice, R.M.: 'Watergan: unsupervised generative network to enable real-time color correction of monocular underwater images', *IEEE Robot. Autom. Lett.*, 2018, **3**, (1), pp. 387–394

[10] He, K., Sun, J., Tang, X.*, et al.*: 'Single image haze removal using dark channel prior', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2011, **33**, (12), pp. 2341–2353

[11] Wen, H., Tian, Y., Huang, T.*, et al.*: 'Single underwater image enhancement with a new optical model'. Proc. Int. Symp. Circuits and Systems, Beijing, China, May 2013, pp. 753–756

[12] Drews, P.Jr., Nascimento, E.D., Moraes, F.*, et al.*: 'Transmission estimation in underwater single images'. Proc. Int. Conf. Computer Vision Workshops, Sydney, Australia, December 2013, pp. 825–830

[13] Emberton, S., Chittka, L., Cavallaro, A.: 'Hierarchical rank-based veiling light estimation for underwater dehazing'. Proc. British Conf. Machine Vision, Swansea, UK, September 2015, pp. 125.1–125.12

[14] Galdran, A., Pardo, D., Picón, A.*, et al.*: 'Automatic red-channel underwater image restoration', *J. Vis. Commun. Image Represent.*, 2015, **26**, pp. 132–145

[15] Carlevaris-Bianco, N., Mohan, A., Eustice, R.M.: 'Initial results in underwater single image dehazing'. Proc. OCEANS MTS, Seattle, USA, September 2010, pp. 1–8

[16] Peng, Y., Zhao, X., Cosman, P.C.: 'Underwater image restoration based on image blurriness and light absorption', *IEEE Trans. Image Process.*, 2017, **26**, (4), pp. 1579–1594

[17] Akkaynak, D., Treibitz, T.: 'A revised underwater image formation model'. Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, Salt Lake City, USA, June, 2018, pp. 6723–6732

[18] Liu, C., Meng, W.: 'Removal of water scattering'. Proc. Int. Conf. Computer Engineering and Technology, Chengdu, China, April 2010, pp. 35–39

[19] Yang, H., Chen, P., Huang, C.*, et al.*: 'Low complexity underwater image enhancement based on dark channel prior'. Proc. Int. Conf. Innovations in Bio-inspired Computing and Applications, Shenzan, China, December 2011, pp. 17–20

[20] Chiang, J.Y., Chen, Y.: 'Underwater image enhancement by wavelength compensation and dehazing', *IEEE Trans. Image Process.*, 2012, **21**, (4), pp. 1756–1769

[21] He, K., Sun, J., Tang, X.: 'Guided image filtering', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2013, **35**, (6), pp. 1397–1409

[22] Levin, A., Lischinski, D., Weiss, Y.: 'A closed form solution to natural image matting'. Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, New York, USA, December 2006, pp. 1–10

[23] Lu, H., Li, Y., Nakashima, S.*, et al.*: 'Turbidity underwater image restoration using spectral properties and light compensation', *IEICE Trans. Inf. Syst.*, 2016, **E99.D**, (1), pp. 219–227

[24] Ancuti, C.O., Ancuti, C., De Vleeschouwer, C.*, et al.*: 'Color transfer for underwater dehazing and depth estimation'. Proc. IEEE Int. Conf. Image processing, Beijing, China, September 2017, pp. 695–699

[25] Berman, D., Levy, D., Avidan, S.: 'Underwater single image color restoration using haze-lines and a new quantitative dataset', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2018, **1**, pp. 1–13

[26] Bryson, M., Johnson, R., Matthew Pizarro, O.*, et al.*: ' true color correction of autonomous underwater vehicle imagery', *J. Field Robot.*, 2015, **33**, (6), pp. 853–874

[27] Bryson, M., Johnson, R., Matthew Pizarro, O.*, et al.*: 'Colour-consistent structure-from-motion models using underwater imagery', *J. Robot.: Sci. Syst.*, 2012, pp. 1–8

[28] Yamashita, A., Fujii, M., Kaneko, T.: 'Color registration of underwater images for underwater sensing with consideration of light attenuation'. Proc. IEEE Int. Conf. Robotics and Automation, Rome, Italy, April 2007, pp. 4570–4575

[29] Zhao, X., Jin, T., Qu, S.: 'Deriving inherent optical properties from background color and underwater image enhancement', *J. Ocean Eng.*, 2015, **94**, pp. 163–172

[30] Cao, Y., Wu, Z., Shen, C.: 'Estimating depth from monocular images as classification using deep fully convolutional residual networks', *IEEE Trans. Circuits Syst. Video Technol.*, 2018, **28**, (11), pp. 3174–3182

[31] Zhang, H., Sindagi, V., Patel, V.M.: 'Joint transmission map estimation and dehazing using deep networks', *IEEE Trans. Circuits Syst. Video Technol.*, 2020, **30**, (7), pp. 1975–1986

[32] Soille, P.: '*Morphological image analysis. Principles and applications*' (Springer-Verlag Berlin Heidelberg, Berlin, Germany, 1999, 2nd edn. 2004), pp. 173–174

[33] Godard, C., Aodha, O., Gabriel, J.: 'Unsupervised monocular depth estimation with left-right consistency'. Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Honolulu, USA, November 2017, pp. 1–10

[34] Godard, C., Aodha, O., Firman, M.*, et al.*: 'Digging into self-supervised monocular depth estimation'. Proc. IEEE Int. Conf. Computer Vision, Seoul, South Korea, October 2019, pp. 1–18