Selective Colour Restoration of Underwater Surfaces

Chau Yi Li
chauyi.li@qmul.ac.uk
Andrea Cavallaro
a.cavallaro@qmul.ac.uk
Centre for Intelligent Sensing
Queen Mary University of London

Abstract

The appearance of surfaces in underwater images is degraded by the selective attenuation of light in water. The light intensity decays exponentially along the vertical depth and along the range between the surface and the camera. Images capturing a large vertical depth exhibit non-uniform water colour. Restoration methods that compensate for the resulting colour cast and reduced contrast generally ignore the degradation caused by depth, whereas methods that target the removal of the colour cast often distort the colour of the water mass. Furthermore, most methods assume a uniform water colour and under-compensate for the colour when the water colour is non-uniform. In this paper, we present a selective chromatic adaptation (SeCA) method that restores the colour appearance of underwater surfaces to that under an unattenuated light, as if the surfaces were captured in air. Using the Schechner-Karpel model, we restore the colour degraded along the range by estimating the extent of colour degradation for each colour channel. Moreover, we handle the case of uniform and non-uniform water colour with one single approach. We also derive a scene-adaptive map that restores the colour degraded along the vertical depth by selectively removing the cast on surfaces while maintaining the water colour. SeCA needs no knowledge of the range nor the vertical depth at which the surfaces are captured. SeCA outperforms state-of-the-art neural networks in terms of colour accuracy. Furthermore, we validate the stability by deploying SeCA on underwater videos without any temporal regularisation.

1 Introduction

Underwater images are degraded by the selective attenuation of light in water. Surfaces captured in the scene, such as fish, coral and divers, are under colour cast and have reduced contrast. The extent of attenuation is an exponential decay that depends on the distance light travels, its wavelength, and the water composition [20]. The attenuation along the vertical depth reduces the ambient light reaching the scene [4]. As the observed water colour changes ambient light, an image capturing a large vertical depth hence exhibits a non-uniform water colour. Moreover, the attenuated ambient light illuminating the surfaces causes the colour cast. The light reflected by a surface is further attenuated along the range before reaching the camera. Moreover, a portion of the ambient light, not reflected by any surfaces, is (back)scattered towards the camera and veils the surfaces with the water colour [32]. In Jerlov oceanic water types, red is most attenuated and surfaces are under blue cast [20].

Physics-based methods restore the colour appearance degraded along the range using *priors* that are learnt using neural networks [19, 37, 39], or derived from observations of the colour [5, 8, 10, 12, 13, 15] or texture appearance [34]. When the image exhibits a non-uniform water colour, restoring with a uniform background light under-compensates the colour at the bottom of the image. Existing methods that address this non-uniform water colour requires knowledge of the change in vertical depth in the scene, which is generally unavailable [10]. A few methods also remove the colour cast caused by attenuation along the depth by chromatic adaptation [3, 5] but often distort the water colour and result in unnatural appearance [29]. Using learning-based methods comes with the difficulty of determining a *reference image* with no degradation. Some methods instead synthesise the degraded image from indoor images (used as reference) that also provided the range of the scene, without modelling the attenuated illuminant [26, 38]. The networks trained on these images only compensate for the range but not the depth. Furthermore, the limited colour palette of indoor images results in a similarly limited colour palette in the processed images. An alternative source of reference images is computer-graphics images that allow full control of the colour palette [31]. Reference images can also be selected via subjective tests [14, 25, 27], for example, the images that are the most chosen among images processed with 12 processing methods [25, 27]. However, the learnt restoration is limited by the fact that these images do not represent the scenes without degradations.

In this paper, we propose a selective chromatic adaptation (SeCA) method for images taken in Jerlov oceanic water types with a visible water mass. We selectively restore the colour appearance of surfaces degraded by the attenuation along the vertical depth and the range between the surface and the camera, while preserving the water colour. To restore the colour appearance along the range, we establish an empirical colour constraint for candidate background light pixels using a physics-based model [35]. We show how, with a single approach, we can handle images with uniform as well as non-uniform background light, without prior knowledge on the vertical depth captured in the image. To estimate the transmission map of the red channel, which is the most attenuated, we use a colour appearance prior for oceanic water. To address the wavelength-dependency of attenuation, based on empirical oceanic data, we derive a transmission map for each colour channel. To compensate for the residual colour cast caused by the attenuation along the depth, we derive a scene adaptive map that selectively removes the residual cast on surfaces. The code is available at `https://github.com/janicelicy/SeCA/`.

2 Selective colour restoration

Let $\mathbf{I}$ be the (degraded) underwater image, that captures surfaces and the water mass, in the linear RGB colour space and $\mathbf{I}_k \in [0,1]$, $k \in \{R,G,B\}$, denote a colour channel. We will drop most subscripts for simplicity. Let (x,y) be a pixel in $\mathbf{I}$ representing a point of the scene at a vertical depth D under the water surface and a range z from the camera. The image $\mathbf{I}$ can be modelled as a convex combination of the light reflected by surfaces, $\mathbf{J}$, and the backscattered ambient or *background* light, $\mathbf{A}$ (Schechner-Karpel model [35]). The combination is modulated by a *transmission map*, $\mathbf{T}$, that describes the attenuation extent of each colour channel as $\mathbf{T}_k = e^{-\beta_k \mathbf{z}}$, $k \in \{R,G,B\}$, where β_k is the *wideband attenuation coefficient* that quantifies the attenuation of the light contributing to the colour channel and $\mathbf{z}$

is the range map of the scene. Therefore

$$\mathbf{I} = \mathbf{T} \cdot \mathbf{J} + (\mathbf{1} - \mathbf{T}) \cdot \mathbf{A}, \tag{1}$$

where $\cdot$ denotes element-wise multiplication. The range-compensated image $\mathbf{J}$, which represents the scene illuminated by an attenuated illuminant, can be obtained by solving Eq. 1.

To fully restore the surfaces' colour appearance, the residual colour cast should be removed, for example, by chromatic adaptation. Chromatic adaptation transforms the image under the attenuated illuminant, $\mathbf{J}$, to the image under a target unattenuated illuminant, $\hat{\mathbf{J}}$, e.g. with a linear transformation [16]

$$\hat{\mathbf{J}} = \mathcal{M}^{-1}\mathcal{D}\mathcal{M}\mathbf{J}, \tag{2}$$

where the 3-by-3 matrix $\mathcal{M}$ maps a colour to the LMS colour space, that models the responses of the three types of cones in the human eye, and the 3-by-3 diagonal matrix $\mathcal{D}$ scales the intensity of each channel independently to achieve the cone responses under the unattenuated illuminant [16]. In particular, when all diagonal entries are 1, the colour cast remains (i.e. no chromatic adaptation is performed). The surfaces in $\hat{\mathbf{J}}$ would appear as if they were captured in air. Fig. 1 summarises the proposed method.

2.1 Restoration along the range

Transmission map. We use the observation that the red intensity of surfaces decreases with increasing range and that the blue and green intensities become more similar to the water colour [15]. Assuming that the degradation extent is constant in a local window $\Omega(x,y)$ centred at (x,y), we estimate a coarse transmission map of the red colour channel using the Red Channel Prior [15] as

$$\tilde{\mathbf{T}}_R(x,y) = 1 - \min\left(\min_{(s,t)\in\Omega(x,y)} \frac{1-\mathbf{I}_R(s,t)}{1-A^\star_R}, \min_{(s,t)\in\Omega(x,y)} \frac{\mathbf{I}_G(s,t)}{A^\star_G}, \min_{(s,t)\in\Omega(x,y)} \frac{\mathbf{I}_B(s,t)}{A^\star_B} \right), \tag{3}$$

where $\mathbf{A}^\star$ is the background light that represents the water colour. Pixels with a low red intensity have a smaller $\tilde{\mathbf{T}}_R$, whereas pixels with colour intensity different from that of the background light have $\tilde{\mathbf{T}}_R$ close to 1. Since the locally constant transmission assumption causes mis-estimation at surface boundaries, we obtain the refined map, $\mathbf{T}_R$, by applying[1] Laplacian matting [24] to $\tilde{\mathbf{T}}_R$. We then derive the transmission maps of blue and green channels using the *ratio* between the wideband attenuation coefficients of the colour channels, e.g. $\mathbf{T}_B = (\mathbf{T}_R)^{\beta_B/\beta_R}$. This ratio can be expressed in terms of the scattering coefficients b, the background light $A^\star$ and the peak wavelengths λ of the colour channel as [17, 41]:

$$\frac{\beta_k}{\beta_R} = \frac{b_k A^\star_R}{b_R A^\star_k}, \quad \text{where} \quad \frac{b_k}{b_R} = \frac{1.62517 - 0.00113\lambda_k}{1.62517 - 0.00113\lambda_R}, \ k \in \{G, B\}, \tag{4}$$

where λ for red, green and blue channels are 620nm, 540nm and 450nm, respectively [41].
Background light. We obtain the background light from water pixels that represent the visible water mass [5]. We establish a constraint for feasible water colour in oceanic water types from empirical data and present an approach that handles both uniform and non-uniform background light $\mathbf{A}$. We first identify candidate water pixels as flat regions [5] with similar

[1] For numerical stability, we also put in a lower bound of 0.1 to the transmission value to obtain $\mathbf{T}_R$.

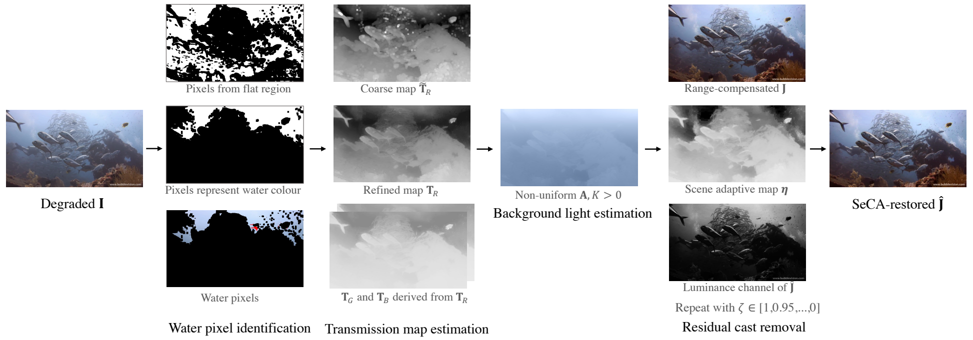


Figure 1: Proposed selective colour restoration from a single image. We first identify the water pixels and calculate the representative water colour (from pixels marked as red). We then estimate the transmission maps for the three colour channels. We then estimate the background light, which could be non-uniform, with linear regression on the change per pixel distance from the water pixels. Finally, we selectively remove the residual colour cast, caused by attenuation along the depth, from surfaces with the proposed scene adaptive map and estimated illuminant.

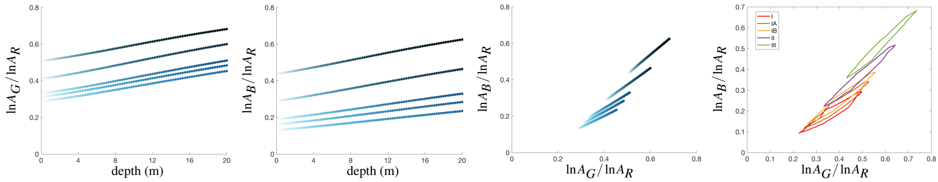
intensities, i.e. where the standard deviation in a local window $\Omega(x,y)$ of the greyscale image of $\mathbf{I}$ is smaller than 1% of the intensity range. As flat regions do not necessarily represent the water mass (see Fig. 1), we also aim to identify pixels representing the water colour. To this end, we synthesise the water colour as captured by different cameras at various depth for Jerlov water types. While the use of synthetic data has been adopted in underwater studies [2] and computational photography [21] to address the lack of real-world data for physics-based problems, we are the first to use these data in underwater image restoration.

The intensities of water pixels are that of the ambient light at a depth D sampled by the camera colour sensors (red, green, blue) over the visible spectrum of light, ω, as

$$A_k = \int_{\omega} S_k(\lambda) E(0,\lambda) e^{-K_d(\lambda) D} d\lambda, \quad k = \{R, G, B\}, \tag{5}$$

where $E(0,\lambda)$ is the light intensity at the water surface, $S_k(\lambda)$ is the spectral response function of the colour sensor, and $K_d(\lambda)$ is the diffuse attenuation coefficient. We synthesise the water colour (Eq. 5) using the known response function of a collection of 28 cameras [21] and the diffuse attenuation coefficients for the five oceanic water types, namely type I, IA, IB, II and III [20]. We synthesise for the depth D between 0.05m and 20m with a 0.05m interval, as most red light has diminished after 20m in oceanic water.

We investigate the *ratio* between colour channels changes to establish the feasible colour constraint. To minimise the effect of depth on the intensity, we take the natural logarithm of the intensity as $\ln A_G / \ln A_R$ and $\ln A_B / \ln A_R$. Fig. 2(a) and (b) visualise the ratios along the depth for a Canon D90 camera. Both ratios are shown to be increasing with the depth as the non-linear natural logarithm function, cannot completely remove the effect of depth from the integrated sampled intensities (Eq. 5). However, we can observe that the loci of the ratios for the Canon D90 are within the numerical value 0.7 (Fig. 2(c)). Moreover, the loci of $\ln A_G / \ln A_R$ and $\ln A_B / \ln A_R$ for the 28 cameras lie within the numerical values 0.8 and 0.7, respectively (Fig. 2(d)). We hence propose the following empirical ratio constraints to



(a) (b) (c) (d)

Figure 2: We investigated the feasible water colour by synthesising water colours along the depth 0.05m to 20m as captured by different cameras (Eq. 5). (a)(b) Relationship between $\ln A_B/\ln A_R$ and $\ln A_G/\ln A_R$ of a Nikon D90 camera for the 5 oceanic water types (from bottom to top: type I, IA, IB, II and III). Marker colour is the synthesised water colour. (c) Relationship between $\ln A_B/\ln A_R$ and $\ln A_G/\ln A_R$ of Nikon D90. (d) Contours of the loci of 28 cameras for the 5 oceanic water types.

identify water pixels:

$$\frac{\ln A_B}{\ln A_R} \leq 0.7 \quad \text{and} \quad \frac{\ln A_G}{\ln A_R} \leq 0.8. \tag{6}$$

From the identified water pixels, we derive the representative background light, $A^\star$, as average colour with the top 1% difference between the dominating water colour channel, i.e. the less attenuated between green and blue, and the most attenuated red colour channel.

To derive the non-uniform background light, we quantify the intensity change by a diffuse attenuation coefficient *per pixel distance*, K, of water pixels. The intensity of two water pixels that are vertically d pixels apart can be related as $\mathbf{A}(x,y+d) = \mathbf{A}(x,y)e^{-Kd}$. This can also be expressed in the linear form that depends on the pixel distance d, by taking the natural logarithm on both sides

$$\ln \mathbf{A}(x,y+d) = \ln \mathbf{A}(x,y) - Kd, \tag{7}$$

where K is the slope of the line and can be estimated with linear regression, using $\ln \mathbf{A}$ and d from the water pixels. The changes in $\mathbf{A}$ for pixels representing surfaces are smaller than K, as the same pixel distance corresponds to a smaller vertical change in depth. Fig. 3 shows the pinhole camera model capturing points representing surfaces ($\mathcal{P}$ and $\mathcal{S}$) and water mass ($\mathcal{Q}$ and $\mathcal{R}$). As a pixel can represent any point in the scene along the line passing through the camera lens' optical centre, $\mathcal{O}$, the change in $\mathbf{A}$ between pixels representing the surfaces can be expressed as a proportion of that between points representing water, as $\frac{\mathcal{PS}}{\mathcal{QR}}K$. We describe the derivation of the proportion $\mathcal{PS}/\mathcal{QR}$ when $\mathcal{S}$ and $\mathcal{R}$ are on the principal axis of the camera lens[2]. From the similar triangles $\Delta\mathcal{POS}$ and $\Delta\mathcal{QOR}$, the proportion can be obtained as $\mathcal{PS}/\mathcal{QR} = \mathcal{PO}/\mathcal{QO} = z_\mathcal{P}/z_\mathcal{Q}$, i.e. ratio between the range of points $\mathcal{P}$ and $\mathcal{Q}$. As the range is encoded in the exponent of transmission map $\mathbf{T}$, we can then obtain the ratio between ranges by taking natural logarithm of $\mathbf{T}$, that will eliminate the wideband attenuation coefficient in the exponent, as $z_\mathcal{P}/z_\mathcal{Q} = \ln \mathbf{T}(x_\mathcal{P}, y_\mathcal{P})/\ln \mathbf{T}(x_\mathcal{Q}, y_\mathcal{Q})$, where $\mathbf{T}(x_\mathcal{P}, y_\mathcal{P})$ and $\mathbf{T}(x_\mathcal{Q}, y_\mathcal{Q})$ are the transmission map values of the pixels representing $\mathcal{P}$ and $\mathcal{Q}$, respectively. As $\mathcal{P}$ is an arbitrary point representing a surface in $\mathbf{I}$ and $\mathcal{Q}$ represents the water mass, the

[2]The general case can be derived by projecting the the points onto the principal axis and considering the corresponding change between the pixels at the image centre and pixels representing the point.

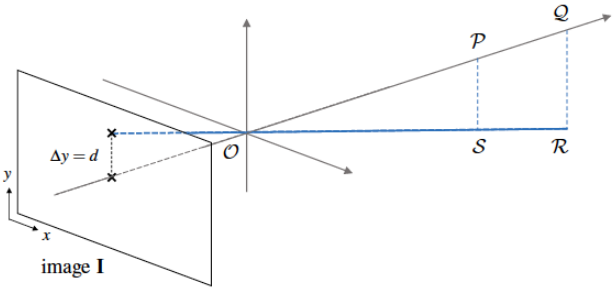


Figure 3: Pinhole camera model shows that the same pixel distance in the image **I** could refer to points at different ranges from the camera. For example, the points $\mathcal{P}$ and $\mathcal{S}$ represent surfaces in the scene, whereas the points $\mathcal{Q}$ and $\mathcal{R}$ represent the water mass. The points $\mathcal{PS}$ and $\mathcal{QR}$, with the same pixel distances in **I**, have different change in water colour in the scene. The change between the $\mathcal{PS}$ that is closer to the camera, $\mathcal{O}$, is smaller than that of $\mathcal{QR}$, which are farther away from the camera.

change in **A** for pixels representing *any* pixel can be expressed in the transmission map as

$$\frac{\mathbf{z}}{z^\star}K = \frac{\ln \mathbf{T}}{\ln \mathbf{T}^\star}K, \tag{8}$$

where $\mathbf{T}^\star$ is the average transmission map values of the pixels contributing to $\mathbf{A}^\star$. Finally, we obtain **A** for *all* pixels, indexed by (x,y), as:

$$\mathbf{A} = A^\star \exp\left(-\frac{\ln \mathbf{T}}{\ln \mathbf{T}^\star}K\,(y-y^\star)\right) \tag{9}$$

where $y-y^\star$ is the pixel distance. When **A** is in fact uniform, an estimated $K=0$ reduces all entries in **A** to $A^\star$. A numerical example can be found in the supplementary material. With the estimated **T** and **A**, we can obtain **J** that is still under a residual colour cast (Eq. 1).

2.2 Residual colour-cast removal

We perform a *selective* chromatic adaptation surfaces without any explicit surface and water-mass segmentation. To this end, we control the extent of cast removal by modifying the diagonal matrix $\mathcal{D}$ in Eq. 2.

Let the attenuated and the target illuminant in the *LMS space* be $(\rho_s, \gamma_s, \beta_s)$ and $(\rho_t, \gamma_t, \beta_t)$, respectively. Chromatic adaptation typically remove the colour cast to the same extent for all pixels, with the diagonal matrix is typically defined as $\mathcal{D} = \mathrm{diag}(\rho_t/\rho_s, \gamma_t/\gamma_s, \beta_t/\beta_s)$. We instead use an exponent $\eta \in [0,1]$ to control the removal extent of each pixel, as

$$\mathcal{D} = \mathrm{diag}\left((\rho_t/\rho_s)^\eta, (\gamma_t/\gamma_s)^\eta, (\beta_t/\beta_s)^\eta\right), \tag{10}$$

where $\eta=0$ corresponding to no cast removal and $\eta=1$ corresponding to full cast removal. The distinction between surfaces and water-mass is implicitly encoded in the range map **z**, that can be derived from the transmission map, as surfaces are at shorter ranges from the

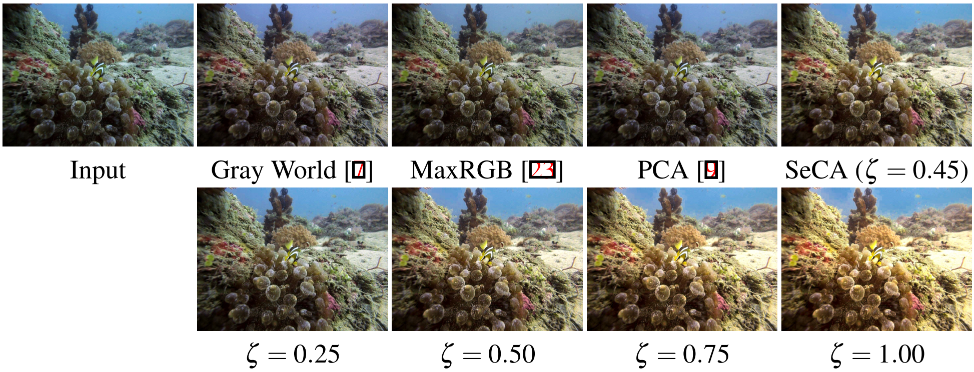


Figure 4: Comparison with illuminant estimation algorithms and the proposed results with different ζ (Eq. 12). The cast removal effect increases with ζ, as the attenuated illuminant is estimated to have a lower intensity. The input is the range-compensated image.

camera than the water-mass. We hence derive a *scene adaptation map*, $\boldsymbol{\eta}$, from the transmission map of $\mathbf{J}$, denoted by $\hat{\mathbf{T}}$, and maximise the removal effect by normalising to $[0,1]$ as

$$\boldsymbol{\eta} = \frac{\ln \hat{\mathbf{T}} - \min_{(x,y)\in\hat{\mathbf{T}}} \ln \hat{\mathbf{T}}}{\max_{(x,y)\in\hat{\mathbf{T}}} \ln \hat{\mathbf{T}} - \min_{(x,y)\in\hat{\mathbf{T}}} \ln \hat{\mathbf{T}}}. \tag{11}$$

Pixels representing the water mass would have $\eta = 0$ and hence no cast removal, whereas pixels representing surfaces would have η close to 1 and hence most cast removal.

The next step is to estimate the attenuated illuminant $\mathbf{E}$ on surfaces. Instead of estimating the illuminant based on reflectance hypothesis, we note that the background light $\mathbf{A}$ is the attenuated illuminant backscattered along the line of sight [41], i.e. $0 \leq \mathbf{A} \leq \mathbf{E} \leq 1$. We can hence express $\mathbf{E}$ in $\mathbf{A}$ as

$$\mathbf{E} = \mathbf{A}^{\zeta}, \quad \zeta \in [0,1]. \tag{12}$$

The larger ζ is, the smaller the intensity of the estimated $\mathbf{E}$ and the larger the cast removal. To adjust for different ambient light environment, we adaptively choose ζ to maximise the cast removal while ensuring the restored image $\hat{\mathbf{J}}$ would not lose details due to overexposure. We define overexposure as too many pixels with high luminance. In our implementation, we repeatedly search for the largest ζ such that no more than 1% of the restored image's pixels exceeds the range of luminance, starting from $\zeta = 1$ and decrease the value by 0.05. Fig. 4 compares the cast removal result of using $\mathbf{E}$ estimated with different algorithms and proposed cast removal with different ζ. All of the existing illuminant estimation approaches over-estimate $\mathbf{E}$ and could only remove the colour cast partially.

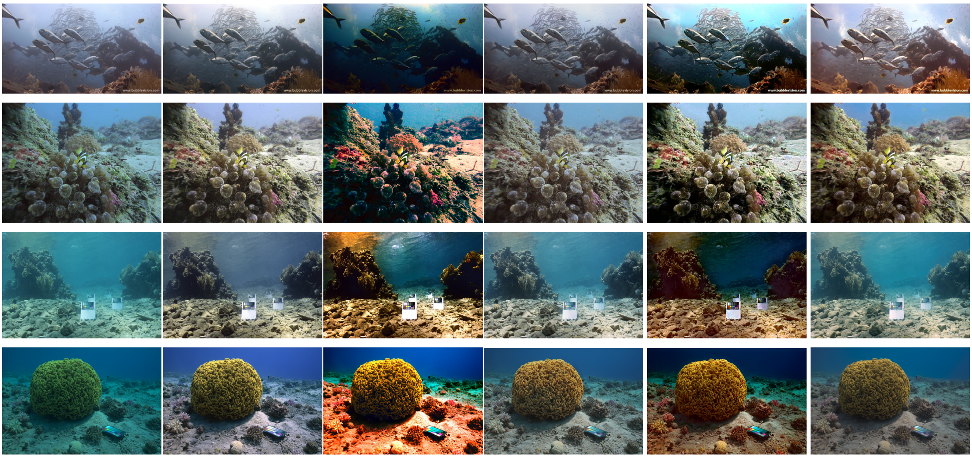
Finally, to deploy Eq. 2 in the LMS space, we first transform $\mathbf{J}$ and $\mathbf{E}$ to the CIEXYZ space then use the Bradford chromatic adaptation transform [22] as $\mathcal{M}$. We use the CIE illuminant D65, that approximates midday light in open-air [11], as the target illuminant.

3 Validation

We compare the proposed method with 2 physics-based methods, namely Fusion [3], and Underwater Haze Line (UWHL) [6], that restore the colour appearance along the range as well as the depth; and 2 neural networks, namely UColor [28] and Cast-GAN [31]. Fusion first uses the opponent colour theory [18] to compensate for the colour lost along the range. On this compensated image, 3 different image processing techniques, namely chromatic adaptation, contrast stretching and image unsharpening, were applied. Fusion then combines the three images using a saliency map. UWHL derives the transmission map from a prior that models the distribution of colour in outdoor images, and assumes a uniform background light. UWHL then applies chromatic adaptation to the entire image. UColor is trained on the UIEB dataset [25] of underwater image pairs and their (preferred) processed version selected by a panel of observers, whereas Cast-GAN is trained on computer-graphics images. We first conduct visual inspection on the restored images, followed by quantitative measures.

Fig. 5 shows restored images of oceanic water. For images with non-uniform background light (top row), methods that only address uniform background light (Fusion) fail to restore the colour at the bottom of the image. UWHL produces a darkened image with its global white balancing. Cast-GAN produces an unnatural water colour (top region) as it was only trained on images with uniform background light. The estimation of non-uniform background light (Eq. 9) ensures the restoration at the bottom of the image for those capturing a large depth. In general, Fusion restores the red colour but also tends to introduce a blue tint and modifies the water colour. UWHL overcompensates for the red colour and produces unnatural red regions. This is caused by the use of outdoor prior that mis-estimates the transmission map. Despite being trained on a large dataset, UColor produces images that have a grey tint and lack contrast (second row). Cast-GAN produces images with high contrast but significantly darkens most of the surfaces. The proposed SeCA shows consistent performance in bright and low ambient light that correspond to shallow and deep water, respectively. The adaptive selection of ζ removes the colour cast in both shallow water (third row) and deeper water (fourth row), where only compensating for the range is insufficient to restore the colour appearance.

As for quantitative evaluation, it has been widely reported that underwater-specific image quality measures, such as UIQM [33] and UCIQE [40], do not reflect subjective judgement [6, 27, 30]. Following the latest literature [30], we quantify the colour accuracy in the processed images. Among the only two underwater datasets available for assessing colour accuracy, we chose Sea-thru [1] over SQUID [6] for the scene diversity. We segmented and labelled all colour patches individually in 55 colour chart in images with visible water mass. Of the 55 charts, 41 are captured in bright ambient light (D3 subset) and 14 are in low ambient light (D5 subset). We measure the colour accuracy, with respect to the reference chart under D65 illuminant, as CIEDE2000, which measures the perceptual difference in the CIELab colour space, and angular error, which measures the angular distance between two colour vectors in the RGB space. The lower the values, the better the colour accuracy. The range of CIEDE2000 is between 0 and 100. However, a CIEDE2000 value bigger than 10 is meaningless as it is designed for measuring small colour distances. The angular error ranges between $0°$ to $180°$. Table 1 reports the colour accuracy on the colour charts (sample shown in third and fourth row of Fig. 5) where all methods have CIEDE2000 higher than 10. We hence do not discuss the accuracy based on CIEDE2000 and focus on the intensity independent angular error measure. For images in bright ambient light, UWHL achieves the best accuracy in angular error ($15.53°$) whereas UColor has the worst accuracy ($21.47°$),



Degraded Fusion [3] UWHL [6] UColor [28] Cast-GAN [31] SeCA

Figure 5: Sample images in oceanic water. First row: Image that captures a non-uniform water colour. Second & third row: Image that captures a uniform water colour in bright ambient light. Third & fourth rows: Sea-thru image with colour charts [1] used in colour accuracy validation, captured in bright and low ambient light, respectively.

Amb.	Acc.	Degraded	Fusion [3]	UWHL [6]	UColor [28]	Cast-GAN [31]	SeCA
Bright	ΔE	29.92 (13.25)	21.06 (8.87)	**20.39** (9.58)	25.31 (10.98)	26.08 (12.58)	23.60 (10.04)
	Φ	26.38 (13.79)	18.09 (11.66)	**15.53** (10.85)	21.47 (12.21)	18.04 (12.35)	18.87 (11.72)
Low	ΔE	27.23 (11.36)	21.44 (6.43)	23.59 (8.24)	18.99 (7.11)	**16.30** (10.62)	19.43 (7.20)
	Φ	32.51 (11.09)	17.96 (9.90)	21.10 (10.85)	18.25 (7.24)	17.81 (12.11)	**17.18** (8.08)

Table 1: Colour accuracy measured as CIEDE2000 (ΔE) [36] and angular error (Φ, in degree) on the colour charts images captured in different ambient light (Amb.) from Sea-thru dataset [1]. We report the average and standard deviation (in bracket). The lower the value, the better the colour accuracy. Best result is bold.

which is half way between that of UWHL and that in the degraded images (26.38°). Fusion, Cast-GAN and the proposed SeCA have errors between 18.09° to 18.87°. As for images in low ambient light, SeCA has the best performance, improving the accuracy from the original error of 32.51° to 17.18°. To summarise, the proposed SeCA has a stable performance in bright and low ambient light and outperforms state-of-the-art neural network in terms of colour accuracy. We also note that the colour accuracy measures only consider the colour chart but not other area in the image. Obvious artifacts, such as the overly red sand in UWHL results, are not reflected by the measure. This demonstrates the difficulty of extending conventional measures to underwater image evaluation and also urgent need to develop holistic measures appropriate for the task.

Finally, we demonstrate the stability of SeCA by applying it to video frames independently (Fig. 6). The temporally consistent results, without enforcing any temporal constraint, indicate SeCA's applicability to different ambient light conditions. Additional videos are available in the GitHub repository.

Figure 6: SeCA restores the videos capturing scenes with non-uniform background light without any temporal constraint. Left panel shows the degraded video, right panel shows the restored video. The stability of SeCA is a result of its accurate background light estimation. Click on the image to start video clip (Adobe Reader only). Additional videos are available in the GitHub repository.

4 Conclusion

We address two main problems in underwater image restoration, namely the non-uniform water colour and the removal of residual colour cast from surfaces without distorting the water colour. The proposed approach, SeCA, restores the colour degradations along the range and selectively removes the colour cast on surfaces. We propose a novel approach to handle both uniform and non-uniform background light to compensate for the colour appearance degradation along the range. To remove the colour cast along the depth, we estimate the illuminant on surfaces from the background light. We then derive a scene adaptive map from the range-compensated image to selectively remove the colour cast on surfaces. SeCA shows consistent restoration results in images taken in different ambient light and outperforms state-of-the-art neural networks in colour accuracy. Furthermore, we demonstrated the stability by extending SeCA directly to video frames without imposing any temporal consistency. SeCA demonstrated how combining oceanic domain knowledge and image processing techniques could address physics-based problems and outperform neural networks when accurate, ideal training images are scarce.

References

[1] D. Akkaynak and T. Treibitz. Sea-thru: A method for removing water from underwater images. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pages 1682–1691, June 2019.

[2] D. Akkaynak, T. Treibitz, T. Shlesinger, Y. Loya, R. Tamir, and D. Iluz. What is the space of attenuation coefficients in underwater computer vision? In *Proceedings IEEE conference on Computer Vision and Pattern Recognition*, pages 568–577, Honolulu, HI, USA, July, 2017.

[3] C. O. Ancuti, C. Ancuti, C. De Vleeschouwer, and P. Bekaert. Color balance and fusion for underwater image enhancement. *IEEE Transactions on Image Processing*, 27(1): 379–393, Jan. 2018. ISSN 1057-7149.

[4] A. Beer. Bestimmung der absorption des rothen lichts in farbigen flüssigkeiten. *Annalen der Physik*, 162:78–88, January 1852.

[5] D. Berman, T. Treibitz, and S. Avidan. Diving into haze-lines: Color restoration of underwater images. In *Proceedings of the British Machine Vision Conference*, pages 44.1–44.12, Sep. 2017.

[6] D. Berman, D. Levy, S. Avidan, and T. Treibitz. Underwater Single Image Color Restoration Using Haze-Lines and a New Quantitative Dataset. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 2822–2837, Aug. 2021.

[7] G. Buchsbaum. A spatial processor model for object colour perception. *Journal of The Franklin Institute*, 310(1):1–26, July 1980.

[8] N. Carlevaris-Bianco, A. Mohan, and R. M. Eustice. Initial results in underwater single image dehazing. In *OCEANS 2010 MTS/IEEE Seattle*, pages 1–8, Sep. 2010.

[9] D. Cheng, D. K. Prasad, and M. S. Brown. Illuminant estimation for color constancy: why spatial-domain methods work and the role of the color distribution. *Journal of the Optical Society of America A*, 31(5):1049–1058, May 2014.

[10] J. Chiang and Y. Chen. Underwater image enhancement: Using wavelength compensation and image dehazing. *IEEE Transactions on Image Processing*, 21(4):1756–1769, Apr. 2012.

[11] Commission Internationale de L'Eclairage. Cie standard llluminants for colorimetry, Jan. 2008.

[12] P. Drews Jr, E. do Nascimento, F. Moraes, S. Botelho, and M. Campos. Transmission estimation in underwater single images. In *Proceedings of the International Conference on Computer Vision Workshop*, pages 825–830, Dec. 2013.

[13] S. Emberton, L. Chittka, and A. Cavallaro. Underwater image and video dehazing with pure haze region segmentation. *Computer Vision and Image Understanding*, 168: 145–156, Mar. 2018.

[14] C. Fabbri, M. J. Islam, and J. Sattar. Enhancing underwater imagery using generative adversarial networks. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 7159–7165, May 2018.

[15] A. Galdran, D. Pardo, A. Picón, and A. Alvarez-Gila. Automatic Red-Channel underwater image restoration. *Journal of Visual Communication and Image Representation*, 26:132–145, 2015.

[16] A. Gijsenij, T. Gevers, and J. van de Weijer. Computational color constancy: Survey and experiments. *IEEE Transactions on Image Processing*, 20(9):2475–2489, 2011.

[17] R. W. Gould, R. A. Arnone, and P. M. Martinolich. Spectral dependence of the scattering coefficient in case 1 and case 2 waters. *Applied Optics*, 38(12):2377–2383, Apr. 1999.

[18] E. Hering. *Grundzűge der Lehre vom Lichtsinn*. Springer, Berlin, Heidelberg, Jan. 1920.

[19] Y. Hu, K. Wang, X. Zhao, H. Wang, and Y. Li. Underwater image restoration based on convolutional neural network. In *Proceedings Asian Conference on Machine Learning*, volume 95, pages 296–311, Nov. 2018.

[20] N. G. Jerlov. *Marine optics*. Elsevier, Jan. 1976.

[21] J. Jiang, D. Liu, J. Gu, and S. Süsstrunk. What is the space of spectral sensitivity functions for digital color cameras? In *Proceedings of the IEEE Workshop on Applications of Computer Vision*, pages 168–179, Jan. 2013.

[22] K.M. Lam. Metamerism and colour constancy, 1985. PhD thesis, University of Bradford.

[23] E. H. Land and J. J. McCann. Lightness and retinex theory. *Journal of the Optical Society of America*, 61(1):1–11, Jan. 1971.

[24] A. Levin, D. Lischinski, and Y. Weiss. A closed-form solution to natural image matting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2):228–242, Feb. 2008.

[25] C. Li, C. Guo, W. Ren, R. Cong, J. Hou, S. Kwong, and D. Tao. An underwater image enhancement benchmark dataset and beyond. *IEEE Transactions on Image Processing*, 29(11):4376–4389, Nov. 2019.

[26] C. Li, S. Anwar, and F. Porikli. Underwater scene prior inspired deep underwater image and video enhancement. *Pattern Recognition*, 98:1–11, Feb. 2020. ISSN 0031-3203.

[27] C. Li, S. Anwar, J. Hou, R. Cong, C. Guo, and W. Ren. Underwater image enhancement via medium transmission-guided multi-color space embedding. *IEEE Transactions on Image Processing*, 30:4985–5000, May 2021.

[28] C. Li, S. Anwar, J. Hou, R. Cong, C. Guo, and W. Ren. Underwater image enhancement via medium transmission-guided multi-color space embedding. *IEEE Transactions on Image Processing*, 30:4985–5000, May 2021.

[29] C. Y. Li and A. Cavallaro. Background light estimation for depth-dependent underwater image restoration. In *Proceedings of the IEEE International Conference on Image Processing*, pages 1528–1532, Oct. 2018.

[30] C. Y. Li and A. Cavallaro. On the limits of perceptual quality measures for enhanced underwater images. *arXiv:2207.05470*, July 2019.

[31] C. Y. Li and A. Cavallaro. Cast-GAN: Learning to remove colour cast in underwater images. In *Proceedings of the IEEE International Conference on Image Processing*, pages 1083–1087, Oct. 2020.

[32] B. L. McGlamery. A computer model for underwater camera systems. In *Ocean Optics VI, SPIE*, volume 0208, Mar. 1980.

[33] K. Panetta, C. Gao, and S. Agaian. Human-visual-system-inspired underwater image quality measures. *IEEE Journal of Oceanic Engineering*, 41(3):541–551, July 2016.

[34] Y. T. Peng and P. C. Cosman. Underwater image restoration based on image blurriness and light absorption. *IEEE Transactions on Image Processing*, 26(4):1579–1594, Apr. 2017.

[35] Y. Y. Schechner and N. Karpel. Clear underwater vision. In *Proceedings IEEE conference on Computer Vision and Pattern Recognition*, pages 536–543, June 2004.

[36] G. Sharma, W. Wu, and E. Dalal. The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Research & Application*, 30(1):21–30, Dec. 2005.

[37] Y. Shin, Y. Cho, G. Pandey, and A. Kim. Estimation of ambient light and transmission map with common convolutional architecture. In *OCEANS 2016 MTS/IEEE Monterey*, pages 1–7, Sep. 2016.

[38] P. M. Uplavikar, Z. Wu, and Z. Wang. All-in-one underwater image enhancement using domain-adversarial learning. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition Workshop*, June 2019.

[39] Y. Wang, J. Zhang, Y. Cao, and Z. Wang. A deep CNN method for underwater image enhancement. In *Proceedings of the IEEE International Conference on Image Processing*, pages 1382–1386, Sep. 2017.

[40] M. Yang and A. Sowmya. An underwater color image quality evaluation metric. In *IEEE Transactions on Image Processing*, volume 24, pages 6062–6071, Dec. 2015.

[41] X. Zhao, T. Jin, and S. Qu. Deriving inherent optical properties from background color and underwater image enhancement. In *Ocean Engineering*, volume 94, pages 163–172, Jan. 2015.