

Hierarchical rank-based veiling light estimation for underwater dehazing

Simon Emberton
s.emberton@qmul.ac.uk

Lars Chittka
l.chittka@qmul.ac.uk

Andrea Cavallaro
a.cavallaro@qmul.ac.uk

Centre for Intelligent Sensing,
Queen Mary University of London,
Mile End Road, London, UK

Abstract

Current dehazing approaches are often hindered when scenes contain bright objects which can cause veiling light and transmission estimation methods to fail. This paper introduces a single image dehazing approach for underwater images with novel veiling light and transmission estimation steps which deal with issues arising from bright objects. We use features to hierarchically rank regions of an image and to select the most likely veiling light candidate. A region-based approach is used to find optimal transmission values for areas that suffer from oversaturation. We also locate background regions through superpixel segmentation and clustering, and adapt the transmission values in these regions so to avoid artefacts. We validate the performance of our approach in comparison to the state of the art in underwater dehazing through subjective evaluation and with commonly used quantitative measures.

1 Introduction

Veiling light is the atmospheric light that is scattered from particles in a hazy scene into the line of sight of an imaging device, contributing to image degradation and resulting in images with low contrast and colour cast [1]. Single image dehazing methods improve the visibility of scenes which have been degraded due to scattered light such as fog, haze and smoke by removing the haze in a scene so that objects can be seen more clearly [2]. Veiling light estimation is an important step in single image dehazing because it identifies the colour of the atmospheric light which the dehazing process aims to remove. The veiling light can be best estimated from the region of an image that is the most haze-opaque [3]. The *transmission* is the part of the light that is not scattered by the haze and reaches the camera [4].

Single image dehazing methods can be used to improve the visibility in underwater images [5, 6, 7, 8, 9]. A number of methods have also combined colour correction approaches with dehazing to further enhance the appearance of underwater scenes [6, 10]. One of the major challenges for dehazing approaches is to avoid selecting pixels from bright objects during the estimation of the veiling light. We propose a method which aims to avoid objects in the veiling light detection step. We find that regions with objects contain more edges and

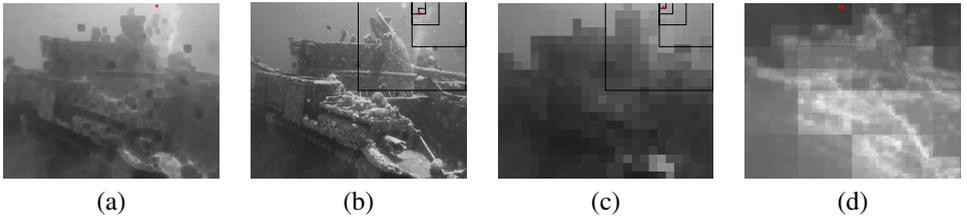


Figure 1: Comparison of veiling light estimation methods. Veiling light estimates are taken from locations highlighted in red. Current methods are biased toward areas of the Shipwreck image with bright objects (e.g. bubbles) whereas our approach is able to avoid these regions. (a) Underwater Dark Channel Prior [10]. (b) Quad-tree subdivision methods on the input image converted to greyscale [16] and (c) into blocks that are filled with their minimum value [20]. (d) Proposed method.

show more variation within each colour channel than pure veiling light regions. We also consider all regions of an image in the final decision of which region is the most heavily veiled. Fig. 1 compares state-of-the-art veiling light estimation methods together with our proposed approach.

Another challenge is to estimate a transmission map which avoids output dehazed images suffering from oversaturation and artefacts in the background regions. Under- and oversaturation occurs when the dehazed image has values that are unphysical (i.e. values outside of the range $[0,1]$). This often happens when bright objects are incorrectly estimated as being located in the distance and therefore having inappropriate transmission values [20]. Artefacts in the background regions are caused by low transmission values in these parts of the scene [2].

In this paper we introduce a single image dehazing approach for underwater scenes. We employ a hierarchical rank-based method to estimate the veiling light and an adapted transmission estimation step which prevents oversaturation and artefacts in the dehazed image. We make use of a set of features which allows us to locate veiling light regions and to avoid those regions which contain bright objects. We adapt a transmission map created with the green-blue dark channel to produce dehazed images which are free from oversaturation and noisy artefacts. Our method performs well in comparison to the state of the art in underwater dehazing on commonly used quantitative measures and a subjective evaluation.

2 Background

The most commonly used model for a hazy image is [14]

$$I(x) = S(x)t(x) + (1 - t(x))V, \quad (1)$$

where $I(x)$ is the observed intensity at pixel location x in an RGB image, $S(x)$ is the scene radiance (without haze), $t(x)$ is the transmission and V is the global atmospheric light (veiling light). The first term of the right-hand side describes how the scene radiance deteriorates in a medium. The second term models the global atmospheric light scattering in a hazy environment. The transmission can be expressed as $t(x) = e^{-\beta d(x)}$, where β is the scattering coefficient and $d(\cdot)$ is the estimated scene depth (distance between objects and the camera) [24]. The aim of dehazing methods is to recover $S(x)$, V and $t(x)$ from $I(x)$. Eq. 1 can be

rewritten as

$$S(x) = \frac{I(x) - V}{t(x)} + V, \quad (2)$$

where we aim to estimate the veiling light V and transmission t in a scene.

2.1 Veiling light estimation

Early dehazing methods took the brightest pixel in a scene [10, 21]. However, this can lead to erroneous estimates as a pixel containing a white or bright object is often selected. The Dark Channel Prior (DCP), a method based upon the observation that haze-free images have at least one colour channel with very low intensity values and therefore the values in the dark channel of hazy images can be attributed to the haze, addressed this by choosing the 0.1% brightest pixels in the dark channel and then selecting the highest intensity pixel from this group [14]. However, this method fails in scenes with large bright objects. Underwater the red channel has little information at depth and therefore cannot perform correctly as a dark channel. Underwater specific methods deal with this issue by constructing the dark channel in various ways. One method introduces a novel prior, The Red Channel Prior (RCP) [15], which is based on the observation that intensities in the red channel deteriorate more quickly as distance increases underwater (see also Eq. 6). The veiling light is estimated as the furthest pixel in the scene which is taken to be the pixel with the maximum intensity in the red channel [15]. The Underwater Dark Channel Prior (UDCP) [2] estimates the veiling light from the pixel with the highest intensity in the green-blue dark channel. Carlevaris-Bianco *et al.* [9] estimate the veiling light from the pixel that contains the minimum value in the transmission map which indicates the furthest point in the scene, where the transmission is estimated from the difference between the maximum in the red channel and the maximum in the green and blue channels (see also Eq. 5).

Methods for veiling light estimation in terrestrial scenes include calibration [12], taking the maximum intensity in the sky region [24] and introducing a bright channel and then taking the difference between the dark channel and the bright channel [23]. A recent hierarchical veiling light estimation approach demonstrated it could improve current methods in the case of terrestrial scenes that contain bright lights [5]. Bright regions in an image are found by constructing a histogram in the luma channel of YCbCr colour space and outliers are discarded. Further refinement is achieved by searching for pixels with a high colour similarity. Veiling light is taken as the average value of the remaining candidates [5]. A number of approaches make use of a hierarchical search using quad-tree subdivision [16, 20]. A greyscale version of $I(x)$ can be divided into 4 equal size blocks, the block with the brightest mean pixel value is selected, then the process is repeated until the size of the block is under a pre-defined threshold. Finally, the veiling light is chosen from the input image as the brightest pixel in the selected block [16]. The quad-tree methods work on the assumption that the average brightness of the blocks in the veiling light regions is larger than the average brightness in the ground-based regions, which, although containing bright objects, are predominantly dark. However, when bright objects are present in the veiling light regions of an image, the increase in the average intensity values of the block containing the object makes a bias towards that block being chosen. In order to reduce the effect of bias from bright objects, a quad-tree search can be applied on a greyscale version of $I(x)$ which is divided into blocks of 30×30 pixels and the pixels in each block are given the minimum value from that block [20]. This method is still limited by the fact that bright objects may cause the estimate to drift.

2.2 Transmission estimation

In order to estimate the amount of transmission for each pixel in a scene the majority of approaches make use of the Dark Channel Prior (DCP) [14] or adaptations of it [4, 15, 16]. For an image I , its dark channel C^d can be described as [14]

$$C^d(x) = \min_{y \in \Omega(x)} \left(\min_{c \in \phi} I^c(y) \right), \quad \phi = \{R, G, B\}, \quad (3)$$

where $I^c(\cdot)$ is an intensity value of a specific pixel in colour channel c of I and $\Omega(x)$ is a local patch centred at x . Except in the sky regions, the C^d in a haze free image is close to zero [14]. Therefore in hazy images C^d is representative of the haze and, if the veiling light V is known, the DCP method estimates the transmission as [14]

$$\tilde{t}(x) = 1 - \min_{y \in \Omega(x)} \left(\min_{c \in \phi} \frac{I^c(y)}{V^c} \right), \quad \phi = \{R, G, B\}, \quad (4)$$

assuming the transmission in a local patch $\Omega(x)$ is constant, it is denoted as $\tilde{t}(x)$.

Long wavelength (red) light is attenuated in water more than short wavelength (blue) light [14], so the DCP method suffers from the limitation that the darkest channel is usually the red channel. Because at depth there is no red light and therefore no variation in information in the darkest channel, the transmission estimate will be corrupted [15]. In order to deal with this limitation the Underwater Dark Channel Prior (UDCP) [16] applies DCP only to the blue and green channels $C^{d^{GB}}(x)$.

Another method developed for dehazing underwater images focusses on the difference, $D(\cdot)$, between the maximum in the red channel and the maximum in the other channels (green and blue) [16]

$$D(x) = \max_{y \in \Omega(x)} I^R(y) - \max_{y \in \Omega(x)} \left(\max_{c \in \phi'} I^c(y) \right), \quad \phi' = \{G, B\}, \quad (5)$$

$$t'(x) = D(x) + (1 - \max_x D(x)),$$

where the estimated transmission at each pixel location $t'(x)$ is calculated by adjusting the values of $D(x)$ until the maximum difference (which indicates the nearest point in the scene) is one [16].

The Red Channel Prior method (RCP) modifies the DCP based on the observation that in underwater images intensities in the red channel deteriorate more quickly as distance increases. For a non-degraded image RCP states the following [16]

$$\tilde{t}(x) = 1 - \min \left(\frac{\min_{y \in \Omega(x)} (1 - I^R(y))}{1 - V^R}, \frac{\min_{y \in \Omega(x)} I^G(y)}{V^G}, \frac{\min_{y \in \Omega(x)} I^B(y)}{V^B} \right), \quad (6)$$

where V^R, V^G, V^B are the individual colour channels of V . For a degraded image the prior is also valid as the red channel retains some intensity. The prior rapidly becomes invalid as distance increases. This information is used to estimate a transmission map for each colour channel [16]. This method also colour corrects the dehazed image by weighting the contributions of each colour channel using the reciprocal of the veiling light coefficients [16].

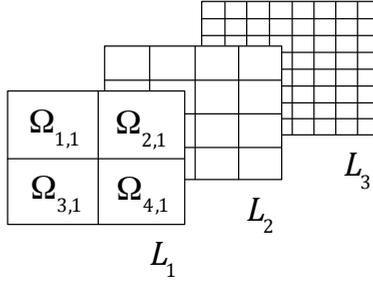


Figure 2: Hierarchical decomposition of an image. An image is represented as a set of layers L_1, L_2, \dots, L_J where each layer is divided into a different number of regions Ω .

Dark channel approaches fail when bright objects are present in a scene. When the dark channel is used to estimate t , nearby bright objects are classified as being far away. This then causes issues in the final dehazed image as these low values in the transmission map can lead to oversaturation in the output image. Other methods make use of block-based approaches where the optimum transmission for each local block is found by using an objective function [16, 20]. In [16] the transmission was selected to maximise the contrast via the standard deviation and the uniformness of the histogram on the luminance channel. Alternatively, [20] aims to increase the contrast whilst reducing the loss of information due to under- and oversaturation. The objective function aims to maximise the contrast measured with entropy and to minimise the amount of pixels that are truncated in the dehazed image. Another issue that has been highlighted in the literature is that the sky regions can suffer from noise and artefacts resulting from the dehazing process [22]. A number of ways have been suggested for how to deal with this problem, such as making use of semantic priors [22] or applying a binary adaptive threshold to the green-blue dark channel for underwater scenes [15].

3 Veiling light estimation

Given image I with intensity values between 0 and 1 we define a set of layers $L = \{L_j\}_{j=1}^J$, which in turn are composed of a set of regions $L_j = \{\Omega_{i,j}\}_{i=1}^{M_j}$, where i denotes the region, j the layer and $\Omega_{i,j}$ is the i^{th} region of the j^{th} layer which contains a set of pixels. We make an assumption that veiling light regions dominate certain areas of an image and so we use layers to incorporate global and local information. For each layer the whole image is divided into a different number of regions which ensures that information from different sized regions in an image is taken into consideration (Fig. 2). Each layer is composed of $M_j = 4^j$ regions which are of a uniform size and rectangular shape.

Veiling light regions that contain bright objects often have a lower average intensity in the green-blue dark channel in comparison to regions dominated by pure veiling light. However, because bright objects can bias veiling light estimates, we negatively weight regions that contain objects. Objects in comparison to areas of veiling light generally display more variation within each colour channel of the input image and contain more edges. We aim to use features to find the regions of an image which are the most haze-opaque, and which do not contain objects. These regions can be characterised by high intensities in the green-blue

dark channel, little variation within individual colour channels and an absence of edges. We extract 3 features f_1, f_2 and f_3 which make use of the green-blue dark channel, the standard deviation of each colour channel in I and the gradient magnitude of a greyscale version of I . We define $\mu_{i,j}^{dGB}$, the mean of the i^{th} region of the j^{th} layer of the green-blue dark channel, as

$$\mu_{i,j}^{dGB} = \frac{1}{|\Omega_{i,j}|} \sum_{x \in \Omega_{i,j}} C_{i,j}^{dGB}(x), \quad (7)$$

where $C^{dGB}(x)$ is the dark channel created with only the green and blue colour channels (from Eq. 3) and $|\Omega_{i,j}|$ is the number of pixels in $\Omega_{i,j}$. The first feature f_1 is defined as

$$f_{1,i,j} = 1 - \mu_{i,j}^{dGB} \in [0, 1]. \quad (8)$$

We subtract the mean from one so that smaller numbers indicate a higher likelihood of veiling light. We define $\sigma_{i,j}^c$, the standard deviation of a colour channel as

$$\sigma_{i,j}^c = \sqrt{\frac{1}{|\Omega_{i,j}| - 1} \sum_{x \in \Omega_{i,j}} (\mu_{i,j}^c - I_{i,j}^c(x))^2}, \quad (9)$$

where $\mu_{i,j}^c$ is the mean and $I_{i,j}^c(x)$ are pixels of the i^{th} region and j^{th} layer of each colour channel of the input image. For individual colour channels, we can find $\sigma_{i,j}^R, \sigma_{i,j}^G, \sigma_{i,j}^B$ using Eq. 9. Our second feature f_2 is defined as

$$f_{2,i,j} = \frac{\sigma_{i,j}^R + \sigma_{i,j}^G + \sigma_{i,j}^B}{3} \in [0, 1]. \quad (10)$$

Finally, we define our third feature f_3 , the mean of the gradient magnitude, as

$$f_{3,i,j} = \frac{1}{|\Omega_{i,j}|} \sum_{x \in \Omega_{i,j}} (G'_{i,j}(x)) \in [0, 1], \quad (11)$$

where $G'_{i,j}(\cdot)$ is the i^{th} region and j^{th} layer of the gradient magnitude of a greyscale version of the input image. The intermediate difference operator calculates the gradient magnitude between data points which are then interpolated [9].

For each feature $f_{1,i,j}, f_{2,i,j}, f_{3,i,j}$ we calculate the average value across all of the layers. For example, if we take a region $\Omega_{i,J}$ in the finest-grained layer J of feature 1. Let $\Omega_{k,j}$ be the region in layer j that contains the pixels of region $\Omega_{i,J}$, then we define the average feature as $\bar{f}_{1,i,J} = \frac{1}{J} \sum_{j=1}^J f_{1,k,j}$, where $\{\bar{f}_{1,i,J}\}_{i=1}^{M_J}$ is the size of 4^J . The same process is applied to all the features resulting in $\bar{f}_{1,i,J}, \bar{f}_{2,i,J}, \bar{f}_{3,i,J}$. To estimate veiling light regions more robustly than considering individual features alone we plan to fuse together the outputs from each feature. Initially we normalise them so that they are weighted with equal importance within $[0, 1]$. In order to standardise the data we fit a Gaussian distribution with mean μ and standard deviation σ . We define the limits of each distribution as $\mu \pm 3\sigma$ to discard outliers in the normalisation process. Data within these limits is stretched between 0 and 1 which results in the normalised outputs $\bar{f}_{1,i,J}^\alpha, \bar{f}_{2,i,J}^\beta, \bar{f}_{3,i,J}^\gamma$. An example of the normalised veiling light features (rescaled to image size) can be seen in Fig. 3. Finally, we add together the normalised outputs to produce $F_{i,J} = \frac{\bar{f}_{1,i,J}^\alpha + \bar{f}_{2,i,J}^\beta + \bar{f}_{3,i,J}^\gamma}{3}$, the fusion decision map. Where $\{F_{i,J}\}_{i=1}^{M_J}$ is the

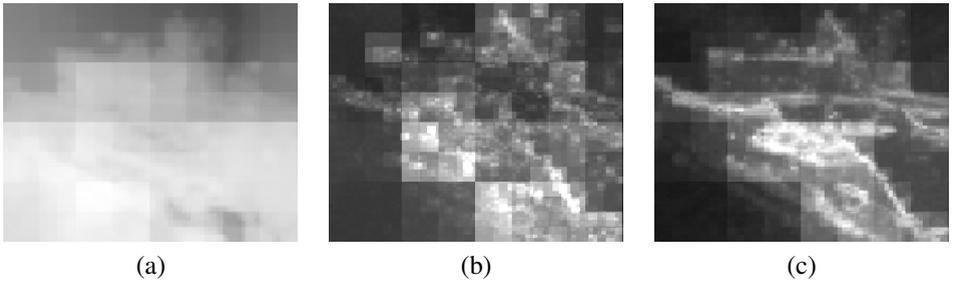


Figure 3: Veiling light estimation features (normalised and rescaled to image size) for Shipwreck using (a) the green-blue dark channel, (b) the standard deviation of each colour channel and (c) the gradient magnitude.

size of 4^J and lower values indicate a higher likelihood of veiling light. Fig. 1 (d) shows the fusion decision map (rescaled to image size). By assuming that the veiling light V is constant throughout the scene and that there are no other sources of illumination (e.g. artificial light sources), we choose $v^* = \underset{i}{\operatorname{argmin}}(F_{i,J})$ as the region which is most likely to contain the veiling light. Within v^* the brightest pixel in the green-blue dark channel is selected and then the RGB values of that pixel from I is taken to represent the veiling light $V = (x_v, c)$. Where $x_v \in \mathbb{R}^2$ is the location of the veiling light and $c \in \mathbb{R}^3$ is the RGB value in x_v .

4 Transmission estimation

We aim to find the transmission $t(x)$ where the dehazed image $S(x)$ does not suffer from oversaturation due to bright objects in the scene and where there are no artefacts and noise in the background regions. We make use of a transmission map created with the green-blue dark channel $C^{d^{GB}}(x)$ (from Eq. 3) and refined with the WLS filter [9] (see Fig. 4 (a)), an edge-preserving operator based on weighted least squares. We then modify this refined transmission map to account for bright objects and background regions in the scene.

We find that low t values can lead to bright objects in an image becoming oversaturated in the final dehazed image (see Fig. 4 (b)). We employ a region-based approach and make use of the regions from the layer L where $j = J$. Initially $C^{d^{GB}}(x)$ is used as the transmission map in Eq. 2 and we flag regions where pixels are truncated in any colour channel of S . For these flagged areas we then employ a range of values r between 0 and 2 for t to calculate S in Eq. 2. We aim to find the minimum possible transmission value without issues due to truncation. For each region i , we find the values r where S is between 0 and 1, then we take $\min(r)$ which ensures the minimum transmission value and results in the least amount of haze in S .

Artefacts often occur in the background regions of dehazed images (see Fig. 4 (b)). We assume that images contain background regions and that in these regions the haze is denser than in foreground regions. We attempt to adjust the transmission map within these regions so that noise and artefacts are not produced in the dehazed image. It is important to select only the background regions as altering the transmission in other regions would worsen the dehazing result. In order to find regions in an image with homogeneity and that are characterised by their similarity in colour to the veiling light estimate we employ SLIC super-

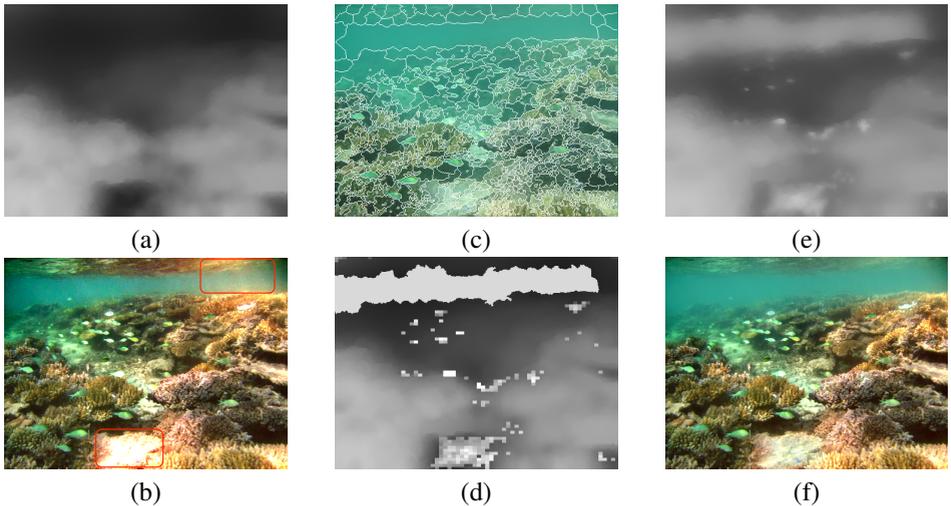


Figure 4: Transmission estimation. (a) Transmission map generated with green-blue dark channel (refined with weighted least squares (WLS) filter [10]) and (b) the resulting dehazed image. Red boxes highlight an area which is oversaturated due to a bright object being underestimated in the transmission map and a background region containing noise and artefacts. (c) Background region estimation using SLIC superpixel segmentation [11] and refined with DBSCAN clustering [12]. (d) Transmission map with adaptations for truncated and background regions. (e) Refined transmission map using WLS filter [10] and (f) the resulting proposed dehazed image.

pixel segmentation [11] and DBSCAN clustering [12], which is used to refine the segmentation (Fig. 4 (c)). The distance in CIELAB space is used to define which clusters are grouped together. We selected parameters between 1.6 and 4.5. We make use of the fusion decision map $F_{i,j}$ to find the superpixel segments \tilde{S} which are likely to predominantly contain veiling light. As the size of $F_{i,j}$ is smaller than the input image I we create an image P the same size as I . We define $P(x) = F_{i,j}$ if $x \in \Omega_{i,j}$, where $\Omega_{i,j}$ denotes the regions in the finest-grained layer. P is shown in Fig. 1 (d). For each \tilde{S} we find the corresponding mean value of P and if it is equal to or below a threshold it is designated as a background region, where the threshold is defined as the mean value of the \tilde{S} that contains the veiling light estimate. Background regions are given a high transmission value so that they are only slightly dehazed in the output image.

The new regions that form the adapted transmission map protrude from the image and require smoothing so that the transmission map is cohesive (Fig. 4 (d)). We again make use of the WLS filter [10] which provides a good balance between computation speed and performance. Smoothing and scaling parameters set to 1 and 1.2, respectively, removed sharp edges at the borders of regions whilst maintaining overall information (Fig. 4 (e)).

5 Results

We validate our approach through comparison with underwater dehazing methods [4, 6, 10]. The 10 images used for the evaluation have all previously been used to assess underwater

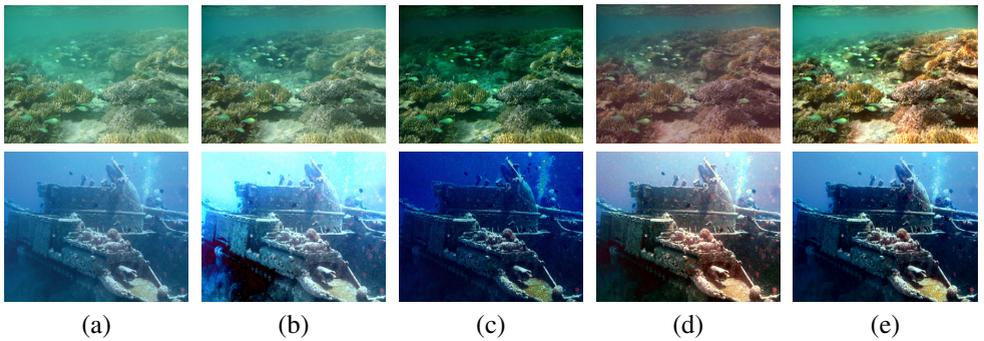


Figure 5: Dehazing results. Top row Galdran1. Bottom row Shipwreck. (a) Input hazy images. Dehazed images with: (b) Carlevaris-Bianco'10 [10], (c) Drews-Jr'13 [13], (d) Galdran'15 [15] and (e) Proposed method.

image processing methods [10, 13, 15, 16]. The dehazed results for two of the images can be seen in Figure 5.¹ For our experiments we define the number of layers J as 6. For the generation of all the dark channels we make use of a patch size of 15×15 pixels. In the transmission estimation step, the segments \tilde{S} that are assigned to the background region are set with a transmission value of 0.85.

A method that is commonly used to quantitatively evaluate dehazing approaches is blind contrast enhancement assessment [16]. This measure compares the gradient ratios of edges between greyscale versions of the input and output images and results in three scores e , Σ and \bar{r} . The score e is defined as the newly visible edges after dehazing, Σ is the percentage of pixels that become completely black or white after dehazing and \bar{r} is the mean ratio of the gradient norms at the visible edges [16]. We do not include the results from Σ as the values were all very close to zero. One limitation with this assessment method is that undesirable artefacts produced by dehazing methods are counted as edges. Due to the unreliability of this measure we use it together with other contrast measures, the Global Contrast Factor (GCF) [16] and a visibility metric (VM) based on contrast-to-noise ratio [17]. We also undertake a subjective evaluation of the methods. Ten participants (none of which had experience developing software for underwater image processing) were shown two dehazed images at a time and asked to decide which they preferred. One point was awarded for the best method, zero for the worst and if both methods were deemed the same, half a point for each. This was repeated so that all of the methods were compared for all of the images. For all of the measures the higher the value the better the performance. The results can be seen in Table 5.

The method of Carlevaris-Bianco *et al.* [10] achieves a high performance on images taken in clear waters (e.g. Fish, Reef1, Reef2 and Reef3) with high scores on both the \bar{r} and VM measures. However, this approach performs badly with turbid scenes (e.g. Shipwreck, Galdran1, Galdran9, Ancuti1, Ancuti2 and Ancuti3). Another limitation is that the background regions in some of the images are oversaturated (e.g. see Shipwreck in Fig. 5 (b)) and this could explain why the high \bar{r} and VM scores coincide with poor performance on the subjective evaluation. In contrast the method of Galdran *et al.* [15] performs well in the more turbid waters. This approach makes use of a colour correction step which permits a more desirable output for these water types. In less turbid waters, the dehazed images display a

¹The dehazed results for all of the images can be found in <http://www.eecs.qmul.ac.uk/~andrea/dehazing.html>

| | Carlevaris-Bianco '10 [10] | | | | | Drews-Jr '13 [13] | | | | | Galdran '15 [15] | | | | | Proposed | | | | |
|-----------|----------------------------|-----------|--------|------|-------|-------------------|-----------|-------|------|-------|------------------|-----------|-------|------|-------|----------|-----------|-------|------|-------|
| | e | \bar{r} | VM | GCF | SE(%) | e | \bar{r} | VM | GCF | SE(%) | e | \bar{r} | VM | GCF | SE(%) | e | \bar{r} | VM | GCF | SE(%) |
| Shipwreck | 0.9 | 2.4 | 53.23 | 7.55 | 4.17 | 1.69 | 1.16 | 41.47 | 6.16 | 20.83 | 0.86 | 2.09 | 48.81 | 7.89 | 37.5 | 0.94 | 1.34 | 44.64 | 8.52 | 37.5 |
| Fish | 1.08 | 2.39 | 52.79 | 9.11 | 34.17 | 1.74 | 1.66 | 41.06 | 8.88 | 29.17 | 0.27 | 0.71 | 42.94 | 3.11 | 5 | 0.68 | 1.95 | 53.11 | 9.45 | 31.67 |
| Reef1 | 0.18 | 1.33 | 76.64 | 8.85 | 24.17 | 0.2 | 1.32 | 80.74 | 8.92 | 37.5 | 0.04 | 0.72 | 72.29 | 2.86 | 8.33 | 0.12 | 1.33 | 79.45 | 8.66 | 30 |
| Reef2 | 0.73 | 2.13 | 73.33 | 9.59 | 28.33 | 0.92 | 1.22 | 55.54 | 7.9 | 29.17 | 0.04 | 0.67 | 58.34 | 3.96 | 9.17 | 0.05 | 1.15 | 58.63 | 7.84 | 33.33 |
| Reef3 | 0.08 | 1.92 | 104.98 | 9.01 | 15.83 | 0.12 | 1.19 | 83.24 | 8.92 | 40 | 0.1 | 0.87 | 88.39 | 4.01 | 16.67 | 0.05 | 1.78 | 89.93 | 9.83 | 27.5 |
| Galdran1 | 0.22 | 1.51 | 70.08 | 6.15 | 16.67 | 0.51 | 1.35 | 67.34 | 8.78 | 21.67 | 0.24 | 1.16 | 59.11 | 4.9 | 25 | 0.29 | 1.99 | 93.35 | 8.02 | 36.67 |
| Galdran9 | 0.86 | 1.51 | 43.73 | 5.9 | 22.5 | 3.46 | 1.58 | 30.65 | 4.53 | 27.5 | 1.6 | 2.29 | 41.81 | 3.03 | 17.5 | 1.96 | 2.28 | 36.13 | 7.59 | 32.5 |
| Ancut1 | 2.03 | 1.35 | 26.86 | 5.9 | 6.67 | 1.88 | 1.28 | 19.84 | 6.8 | 19.17 | 5.04 | 2.01 | 27.99 | 8.21 | 44.17 | 3.67 | 1.58 | 23.79 | 7.05 | 30 |
| Ancut2 | 0.46 | 1.07 | 13.21 | 3 | 7.5 | 8.98 | 2.21 | 17.63 | 4 | 21.67 | 6.05 | 2.34 | 18.53 | 3.62 | 39.17 | 7.51 | 2.2 | 16.66 | 4.98 | 31.67 |
| Ancut3 | 0.34 | 1.14 | 44.26 | 2.83 | 1.67 | 1.85 | 1.8 | 40.97 | 4.93 | 24.17 | 2.18 | 2.4 | 48.56 | 6.83 | 42.5 | 1.8 | 1.97 | 43.85 | 4.68 | 31.67 |
| Average | 0.69 | 1.68 | 55.91 | 6.79 | 16.7 | 2.13 | 1.48 | 47.85 | 6.98 | 27.08 | 1.64 | 1.53 | 50.68 | 4.84 | 24.5 | 1.71 | 1.76 | 53.95 | 7.66 | 32.25 |

Table 1: Results for underwater dehazing methods. e : newly visible edges after dehazing [13], \bar{r} : mean ratio of the gradient norms at the visible edges [13], visibility metric (VM): difference based on contrast-to-noise ratio from the input image [25], Global Contrast Factor (GCF): difference in contrast from the input image [18] and SE(%): percentage results for the subjective evaluation. **Red**: best performance. **Green**: second best performance.

reddish veil.

The method of Drews-Jr *et al.* [13] performs consistently well across all of the images, especially on the e and GCF measures. However, the high e scores can be attributed to the production of artefacts in background regions which are counted as edges. Another limitation is that due to incorrectly estimating the veiling light many of the images appear overly dark green and blue (see Fig. 5 (c)). The proposed approach achieves good scores for all of the measures and, on average, outperforms the compared approaches on the \bar{r} and GCF metrics and the subjective evaluation. In comparison to [13] the low e values for our method can be attributed to fewer newly visible edges appearing in the background regions. Limitations of our method include the fact that false colours are introduced in some dehazed images (e.g. Fish and Galdran1) and that in the turbid images further colour correction is required. Another drawback with our approach is the step to locate background regions which makes use of computationally expensive segmentation and clustering, and the clustering method requires input parameters to be set. We plan to automate this step and improve the computational efficiency.

6 Conclusion

Our approach takes into account scene features, and thereby avoids estimating the veiling light from regions that contain objects. It introduces a transmission estimation step which deals with issues such as oversaturation and the production of artefacts. The dehazed images produced by our method show good performance in comparison to the state of the art across a range of underwater images when evaluated both subjectively and on a number of commonly used quantitative measures. In the future we aim to further develop the approach to improve the colour cast of the output images and to increase the computational efficiency to facilitate video dehazing.

Acknowledgement

This work is supported by Queen Mary University of London's Media and Arts Technology programme, EPSRC Doctoral Training Centre EP/G03723X/1.

References

- [1] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. on PAMI*, 34(11):2274–2282, Nov. 2012.
- [2] C. Ancuti, C. O. Ancuti, T. Haber, and P. Bekaert. Enhancing underwater images and videos by fusion. In *Proc. of IEEE CVPR*, pages 81–88, Jun. 2012.
- [3] M. Bosma, J. Smit, and J. T. van Scheltinga. Super resolution volume rendering hardware. In *Proc. of 10th Eurographics conference on Graphics Hardware*, pages 117–112, Aug. 1995.
- [4] N. Carlevaris-Bianco, A. Mohan, and R. M. Eustice. Initial results in underwater single image dehazing. In *Proc. of IEEE OCEANS*, pages 1–8, Sept. 2010.
- [5] F. C. Cheng, C. C. Cheng, P. H. Lin, and S. C. Huang. A hierarchical airlight estimation method for image fog removal. *Engineering Applications of Artificial Intelligence*, 43: 27–34, Aug. 2015.
- [6] J. Y. Chiang and Y. C. Chen. Underwater image enhancement by wavelength compensation and dehazing. *IEEE Trans. on Image Processing*, 21(4):1756–1769, Apr. 2012.
- [7] P. Drews-Jr, E. Nascimento, F. Moraes, S. Botelho, M. Campos, R. Grande-Brazil, and B. Horizonte-Brazil. Transmission estimation in underwater single images. In *Proc. of IEEE ICCV*, Dec. 2013.
- [8] M. Ester, H. P. Kriegel, J. Sander, and X. Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proc. of KDD*, pages 226–231, Aug. 1996.
- [9] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski. Edge-preserving decompositions for multi-scale tone and detail manipulation. *ACM Trans. on Graphics*, 27(3):67, Aug. 2008.
- [10] R. Fattal. Single image dehazing. *ACM Trans. on Graphics*, 27(3):72, Aug. 2008.
- [11] A. Galdran, D. Pardo, A. Picon, and A. Alvarez-Gila. Automatic red-channel underwater image restoration. *Journal of Visual Communication and Image Representation*, 26:132–145, Jan. 2015.
- [12] K. B. Gibson, D. T. Vo, and T. Q. Nguyen. An investigation of dehazing effects on image and video coding. *IEEE Trans. on Image Processing*, 21(2):662–673, Feb. 2012.
- [13] N. Hautiere, J. P. Tarel, D. Aubert, and E. Dumont. Blind contrast enhancement assessment by gradient ratioing at visible edges. *Image Analysis & Stereology Journal*, 27(2):87–95, May 2008.
- [14] K. He, J. Sun, and X. Tang. Single image haze removal using dark channel prior. *IEEE Trans. on PAMI*, 33(12):2341–2353, Dec. 2011.
- [15] B. Henke, M. Vahl, and Z. Zhou. Removing color cast of underwater images through non-constant color constancy hypothesis. In *Proc. of IEEE ISPA*, pages 20–24, Sept. 2013.

- [16] J. H. Kim, J. Y. Sim, and C. S. Kim. Single image dehazing based on contrast enhancement. In *Proc. of IEEE ICASSP*, May 2011.
- [17] H. Koschmieder. *Theorie der horizontalen Sichtweite: Kontrast und Sichtweite*. Keim & Nemnich, 1925.
- [18] K. Matkovic, L. Neumann, A. Neumann, T. Psik, and W. Purgathofer. Global contrast factor - a new approach to image contrast. In *Proc. of Computational Aesthetics in Graphics, Visualization and Imaging*, pages 159–167, May 2005.
- [19] W. N. McFarland. Light in the sea - correlations with behaviors of fishes and invertebrates. *American Zoologist*, 26(2):389–401, 1986.
- [20] D. Park, H. Park, D. K. Han, and H. Ko. Single image dehazing with image entropy and information fidelity. In *Proc. of IEEE ICIP*, Oct. 2014.
- [21] R. T. Tan. Visibility in bad weather from a single image. In *Proc. of IEEE CVPR*, pages 1–8, Jun. 2008.
- [22] K. Wang, E. Dunn, J. Tighe, and J. M. Frahm. Combining semantic scene priors and haze removal for single image depth estimation. In *Proc. of IEEE WACV*, pages 800–807, Mar. 2014.
- [23] C. H. Yeh, L. W. Kang, C. Y. Lin, and C. Y. Lin. Efficient image/video dehazing through haze density analysis based on pixel-based dark channel prior. In *Proc. of IEEE ISIC*, volume 238, pages 14–16, Aug. 2012.
- [24] J. Yu, C. Xiao, and D. Li. Physics-based fast single image fog removal. In *Proc. of IEEE ICSP*, pages 1048–1052, Oct. 2010.
- [25] D. Zhengguo. <http://uk.mathworks.com/matlabcentral/fileexchange/33529-a-new-visibility-metric-for-haze-images>. Last accessed 20th July 2015.