2014 IEEE International Conference on Robotics & Automation (ICRA)
Hong Kong Convention and Exhibition Center
May 31 - June 7, 2014. Hong Kong, China

Simultaneous Underwater Visibility Assessment, Enhancement and Improved Stereo

Martin Roser[1], Matthew Dunbabin[2] and Andreas Geiger[1,3]

***Abstract*— Vision-based underwater navigation and obstacle avoidance demands robust computer vision algorithms, particularly for operation in turbid water with reduced visibility. This paper describes a novel method for the simultaneous underwater image quality assessment, visibility enhancement and disparity computation to increase stereo range resolution under dynamic, natural lighting and turbid conditions. The technique estimates the visibility properties from a sparse 3D map of the original degraded image using a physical underwater light attenuation model. Firstly, an iterated distance-adaptive image contrast enhancement enables a dense disparity computation and visibility estimation. Secondly, using a light attenuation model for ocean water, a color corrected stereo underwater image is obtained along with a visibility distance estimate. Experimental results in shallow, naturally lit, high-turbidity coastal environments show the proposed technique improves range estimation over the original images as well as image quality and color for habitat classification. Furthermore, the recursiveness and robustness of the technique allows implementation onboard an Autonomous Underwater Vehicle for improving navigation and obstacle avoidance performance.**

I. Introduction

Autonomous Underwater Vehicles (AUVs) are increasingly being utilized as tools to obtain visual observations of benthic habitats in highly unstructured and dynamic environments. These include shallow coral reefs, near-shore mangroves and marinas with degraded visibility making traditional diver surveys impractical and dangerous. The surveys, as considered here, typically consist of low-altitude (0.5 to 2 m above the seafloor) transects, similar to diver video transects, to allow habitat classification and mapping.

However, the ability to discriminate objects and perform classification from underwater images, whether collected by divers, AUVs, or Remote Operated Vehicles (ROV) depends on the visibility. Visibility in this paper means the distance which we can reliably discriminate an object in the scene. In shallow coastal environments the visibility can be less than 1 m in tidal embayments and up to 10-20 m with oceanic water. Whilst most literature referenced throughout this paper has considered image enhancement in relatively clear oceanic water, the focus of this paper is the significantly more challenging naturally lit turbid coastal environments.

[1]Dept. of Measurement and Control, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany roser@kit.edu

[2]Institute for Future Environments, Queensland University of Technology, Brisbane, Queensland, Australia m.dunbabin@qut.edu.au

[3]Max Planck Institute for Intelligent Systems, 72076 Tübingen, Germany andreas.geiger@tue.mpg.de

The authors would like to thank the CSIRO Sensors and Sensor Network Transformational Capability Platform, the Karlsruhe School of Optics and Photonics, and the Karlsruhe House of Young Scientists for their financial support to conduct this project.

A. Related Work

There has been a long history of research to enhance underwater images focusing on color restoration and dehazing. A comprehensive overview of methods for image enhancement and restoration along with subjective and quantitative assessment metrics is given in [1]. A particular problem with underwater images is that of color consistency and appearance changes due to preferential absorption of different color wavelengths as they travel through the water [2], [3]. Color techniques ranging from integrated color models [4] to hyper-spectral correction [5] have been proposed. These are typically multi-step approaches utilizing color histogram stretching and color correction to account for varying depth and natural or artificial illumination [4]. Distance-dependent contrast and color degradation [2] – which is the case in turbid media – is typically only addressed in relatively low turbidity images, although several works have achieved scene depth recovery and color enhancement [6]–[8].

Kocak *et al.* [9] provides an overview of improving long-range visibility in degraded images highlighting additional techniques such as structured and multi-source lighting. These methods can produce terrain maps in real-time but their depth of field can be limited.

Single image visibility assessment and enhancement is still a challenging and ill-posed problem. The observed brightness of each pixel depends on the observed scene point radiance, the scattering and attenuation properties of the circumfluent medium, the scene point distance as well as the ambient illumination. Related work for image enhancement (often termed dehazing) can be classified into four main categories depending on the additional information used for disambiguation.

1) ***Polarization****:* Schechner *et al.* [10], [11] consider significantly varying scene distance image restoration using a polarizer attached to a camera. By taking two or more images with different degrees of polarization (DOP), the authors are able to nearly double the visible range and through image processing derive a distance map of the scene. However polarization filters attenuate the radiance transmission to the camera which is undesirable, particularly with time-varying adverse visibility and illumination conditions.

2) ***Turbidity****:* Narasimhan and Nayar [12], [13] as well as Cozman and Krotkov [14] analyse a static scene by taking multiple images under different visibility conditions. Even though they report impressive results it requires a static camera and a significant change in media turbidity for constant illumination conditions.

*3) **Scene Depth**:* Performing image enhancement utilizing the estimated scene depth (*e.g.* with a stereo camera system) is a chicken-and-egg dilemma. Contrast degradation leads to a low signal-noise-ratio (SNR) which makes disparity computation difficult. Conversely accurate and dense scene depth is a prerequisite for solving the ambiguities in the visibility enhancement problem. Whilst no literature on underwater stereo visibility enhancement has been found, some studies have considered it for seeing through fog [15]–[17]. These works make assumptions about the scene requiring either manual input or structure such as road or uniform geo-referenced terrain which is difficult to achieve underwater.

*4) **Priors**:* Single image underwater visibility enhancement often considers priors from image statistics. Fattal [18] and Tan [19] derive assumptions from different phenomena in order to develop a cost function in the framework of Markov random fields (MRFs). He *et al.* [20] derive a *Dark Channel Prior* from natural image statistics. This postulates that there will be a low response in at least one color channel for haze-free images. Turbid media scatter ambient light towards the observer and lead to an additive signal which is a function of the scattering coefficient and the scene distance. The subsequent image enhancement becomes a regional contrast stretching which can create halos. Since the ambiguity between visibility and scene distance cannot be solved, their approach still yields one unknown (depth scale). In addition the dark channel assumption is based on natural outdoor image statistics which might not hold in arbitrary underwater scenes particularly with man-made infrastructure. Notable works utilising the Dark Channel Prior include MRFs for estimating scene radiance variation [8], [21] and median filters [22] for generating depth maps to dehaze and color correct underwater images with relatively small amounts of haze.

B. Paper Outline

The remainder of the paper is as follows. Section II describes the proposed underwater visibility enhancement and assessment approach. Section III presents a simulation evaluation of the proposed approach, with Section IV evaluating the methodology using experimental field data collected by an AUV. Finally, Section V summarizes the paper and describes future research directions.

II. Underwater Visibility Estimation & Image Enhancement

The goal of this research is to allow real-time enhancement of underwater images which are naturally lit and degraded due to relatively high turbidity and other visibility reducing phenomena. Enhancement of underwater images requires modelling and estimation of the water absorption and scattering characteristics to remove haze. However it also requires a scene depth map. Many papers use a single image and the dark channel prior in the estimation of a depth map [8], [21], [22]. In our approach, we use stereo images in a two-stage enhancement process to improve overall image quality allowing visibility and range estimation. The following subsections describe the modelling and estimation steps in detail, with the overall image enhancement approach detailed in Section II-E.

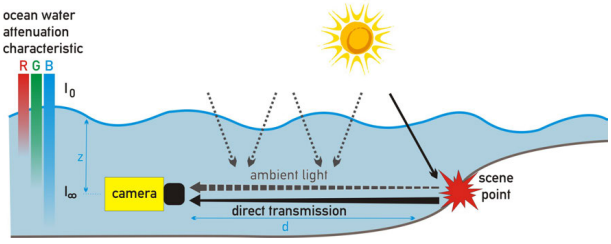


Fig. 1. Underwater Light Model. Light directly transmitted from the scene point will be wavelength dependent exponentially attenuated over distance d and superimposed by the ambient illumination I_∞ at depth z that will be refracted towards the observer.

A. Underwater Light Propagation Modelling

Underwater light models generally follow a standard attenuation model [6] to accommodate wavelength attenuation coefficients [2]. In this study the *Koschmieder Model* [23] was adopted which has been established as a description of the atmospheric effects of weather on the observer [10], [12], [16], [24]. In outdoor clear weather conditions, the radiance from a scene point would reach the observer nearly unaltered. However when imaging underwater the irradiance observed by each pixel of the camera (E) is linear combination of directly transmitted scene object radiance that will be attenuated in the line of sight and scattered ambient light towards the observer as depicted in Figure 1. Therefore, the Koschmieder model has been adapted here for underwater lighting conditions (as in [2], [3]) and can be expressed as

$$E = I_\infty(z)\rho e^{-\beta(\lambda)d} + I_\infty(z)(1 - e^{-\beta(\lambda)d}), \tag{1}$$

where $I_\infty(z)$ is the ambient light at depth z, ρ is the normalized radiance of a scene point, d is the distance from the scene point to the camera and $\beta(\lambda)$ is the total beam attenuation coefficient which is nonlinear and dependent on the wavelength λ. The ambient illumination at depth z itself is subject to light attenuation in the following form:

$$I_\infty(z) = TI_0 e^{-\beta(\lambda)z}, \tag{2}$$

where I_0 is the atmospheric intensity at the surface, and T is the transmission coefficient at the water surface.

When dealing with atmospheric effects, like fog or haze, β is normally assumed to be constant for the whole visible range. However, in underwater applications different wavelengths of light (λ) are nonlinearly and preferentially absorbed relative to others. According to [3], the total beam attenuation coefficient in saltwater can be approximated as

$$\beta(\lambda) \geq K_w^{sw}(\lambda) + b_p(\lambda) + a_p(\lambda) + a_y(\lambda) \tag{3}$$

where K_w^{sw} is the diffuse attenuation coefficient for the clearest saltwater, b_p and a_p are the scattering and absorption coefficients of particles respectively, and a_y is the absorption coefficient of dissolved organic material. Figure 2 illustrates the wavelength dependency of β even in pure saltwater

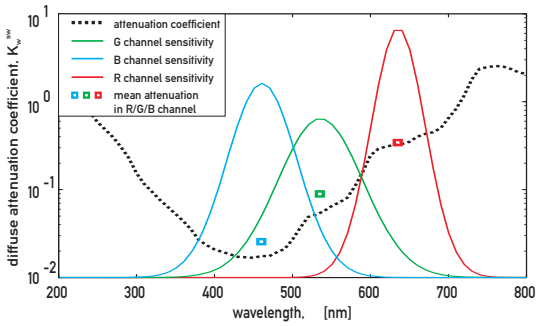


Fig. 2. Clear ocean water light attenuation characteristics deduced from [3]. Also illustrated is typical imaging sensor sensitivity and its overlap with attenuation coefficients which results in a mean channel chromaticity change over depth.

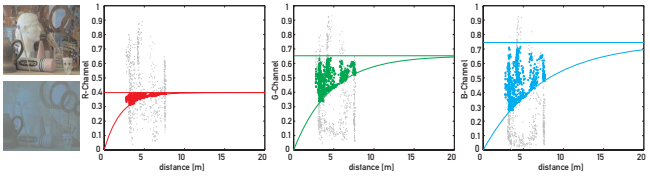


Fig. 3. The observed pixel brightness variation with distance for underwater images with and without haze. For illustrative purposes, the underwater image was obtained from the haze-free image as described in Section III. The range of scene point radiances in good visibility conditions (top left image) are independent from scene distance and possess a wide range of values [0..1] (grey points). The deviation of observed scene radiances in underwater images (bottom left image) will decrease with scene distance, depending on visibility and ambient light, which is subject to wavelength-dependent attenuation. This results in a narrow field of radiances (colored points), that converge to the current ambient light intensity. Conditional, non-linear estimation (trusted region) yields good estimates of the ambient light and attenuation coefficients (colored lines).

with the highest visibility ($b_p = 0, a_p = 0, a_y = 0$). This phenomena leads to a depth-dependent image chromaticity change.

However, in real ocean waters, the scattering and absorption coefficients are non-negligible, particularly in the highly turbid water scenarios considered in this paper. Therefore, these coefficients need to be approximated from the collected images as discussed in the following sections.

B. Estimating Visibility Coefficients

Estimating the visibility conditions ($I_\infty(z)$, ρ, $\beta(\lambda)$ and d) directly from single images is a highly ill-posed problem. For a color image with N pixels, we achieve N observable intensity values in each image color channel E^c where $c \in \{R, G, B\}$. However, the estimation problem in (1) yields $(c+1) \cdot N + 2c > c \cdot N$ unknowns for each image.

For this reason it is assumed a known 3D scene model to give d (the distance to each point in the image) which can be obtained to varying levels of granularity from the stereo images collected by the AUV (see Section II-E for detail).

In our approach, the visibility coefficients are estimated by first binning the intensity values in each color channel with scene distance for all pixels as illustrated in Figure 3. It can be seen that the range of scene point radiances in good visibility atmospheric conditions (top left image) are independent from scene distance and possess a wide range of values [0..1] (grey points). However, the deviation of observed scene radiances in underwater images (bottom left image) will decrease with scene distance, depending on visibility and ambient light, which is subject to wavelength-dependent attenuation through the water [2], [3]. This results in a narrow field of radiances (colored points), that converge to the current ambient light intensity.

Using dark and white patch assumptions, thresholds ($E^c_{X\%}$) are applied where $E^c_{X\%}$ is defined in a way that $X\%$ of intensity values in color channel c are smaller than $E^c_{X\%}$.

Black patch assumption: The lowest 1% of irradiances in each color channel and scene distance bin stem from black objects

$$\rho^c \approx 0 \qquad \forall E^c \leq E^c_{1\%} \tag{4}$$

Using this assumption directly would lead to halos and problems with uniformly colored surfaces like it is the case in regional contrast stretching techniques [20]. Image matting helps prevent halos [6], [8], however matting techniques are based on structure and intensity information which degrade the disparity map in the presence of large uniformly shaded structures or 3D objects that are covering a large scene depth-field. In contrast, we propose to use these scene points in a robust, non-linear estimation framework to achieve a global solution for all depth bins in one color channel.

White patch assumption: Unless retro-reflective surfaces are present, which is generally not the case in natural underwater environments, the observed scene point brightness is upper-bounded by the ambient illumination conditions

$$\rho^c \leq 1, \qquad \forall E^c \geq E^c_{99\%} \tag{5}$$

Inserting all assumptions in (1) yields a conditional, non-linear estimation problem

$$E^c = I^c_\infty(z)(1 - e^{-\beta^c d}), \qquad \forall E^c \leq E^c_{1\%} \tag{6}$$

with two unknowns per color channel ($I^c_\infty(z), \beta^c$) and an additional constraint on the ambient light such that

$$I^c_\infty(z) \geq E^c, \qquad \forall E^c \geq E^c_{99\%}. \tag{7}$$

We use a trust region approach for minimizing (6) subject to the bounds (7) [25]. The optimization is initialized using the brightest pixel in each color channel for the airlight (I^c_∞), and the theoretical clear ocean water values for (β^c). In practice, approximately 12 iterations are required for convergence. Figure 3 illustrates the estimation process as well as the validity of both proposed assumptions with the upper line showing the result of the ambient light estimate (7), and the lower curve the attenuation coefficient from (6).

C. Scene Radiance Recovery

Once β^c and $I_\infty(z)$ are estimated the scene radiance recovery of all image pixels is performed for each color channel according to

$$\rho^c = 1 + \left(\frac{E^c}{I^c_\infty(z)} - 1 \right) e^{\beta^c d} \tag{8}$$

For this step, a dense disparity map of the image scene is essential to determine the scene distances. The generation of dense disparity maps is described in Section II-E.

D. Visibility Assessment

To characterize the current underwater visibility condition, the mean attenuation coefficient (α) that is due to the scattering and absorption of particles is given by

$$\alpha = \frac{\sum_c (\beta^c - \kappa^c)}{\sum_c 1} \tag{9}$$

where the mean clear ocean water attenuation coefficients κ^c for each color channel is determined from the spectral response $s^c(\lambda)$ of the camera sensor as illustrated in Figure 2:

$$\kappa^c = \int_{200nm}^{800nm} s^c(\lambda) * K_w^{sw}(\lambda) d\lambda \tag{10}$$

This measure relates the current conditions to clear ocean water conditions ($\alpha = 0$) which possesses the best theoretical visibility. To estimate visibility distance we approximate human contrast discrimination. Human perception can distinguish two neighbouring scene points as soon as they exceed a contrast $\geq 5\%$ [26] such that

$$C_{\text{vis}} = \frac{E_{\text{max}} - E_{\text{min}}}{E_{\text{max}} + E_{\text{min}}} = 0.05 \tag{11}$$

Inserting (6)-(7), we derive the visibility distance for AUV applications in the color channel with best visibility such that

$$d_{\text{vis}} = \frac{1}{\min_c(\beta^c)} \cdot \log\left(\frac{1 - C_{\text{vis}}}{2C_{\text{vis}}} + 1\right) \tag{12}$$

Additionally, we are able to obtain a depth estimate of the AUV (z_{est}) by comparing the estimated ambient lighting conditions with the light attenuation characteristic from clear ocean water as depicted in Figure 2, such that

$$z_{\text{est}} = \arg\min_z \left(\sum_c (\gamma^c - \gamma_{\text{th}}^c)^2\right) \tag{13}$$

where γ^c is the estimated ambient light chromaticity given by

$$\gamma^c = \frac{I_0^c}{\sum_c I_0^c} \tag{14}$$

and γ_{th} is the theoretical chromaticity-depth-dependency in clear ocean water.

Whilst z_{est} can be very uncertain as it assumes white sunlight and clear ocean water behaviour, in moderate to high visibility conditions it yields a good measure for plausibility checks against other sensors (such as an onboard depth sensor) in order to achieve higher accuracy and reliability.

E. Image Enhancement

In order to perform efficient enhancement of the visibility degraded underwater images, we propose an approach consisting of two visibility and enhancement steps utilizing stereo image pairs as depicted in Figure 4.

The first image enhancement step (step 1) recovers important scene content and image structures in areas with bad visibility. To start the process, a coarse 3D map of the scene needs to be created.

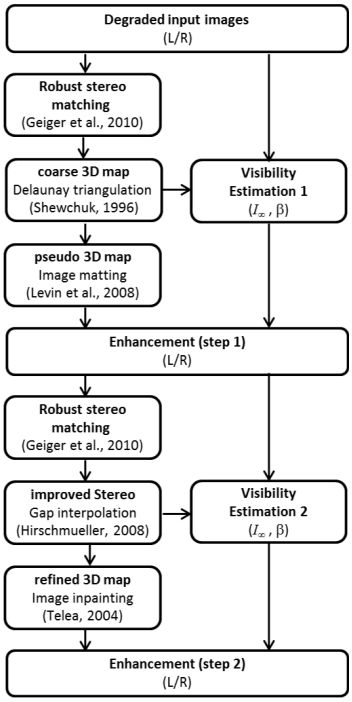


Fig. 4. Flowchart of the proposed image enhancement algorithm consisting of two visibility estimation and two enhancement steps.

We use the *LIBELAS*[1] implementation, proposed in [27], for all stereo retrieval. It is a Bayesian approach to stereo matching that creates a prior on disparities from robustly matched support points. The advantage of this implementation is that even in poor visibility where dense stereo estimation will be challenging, some scene points can be robustly matched. Hence, the prior already implicitly yields a coarse 3D map of the scene that can be used in the first visibility estimation step.

The coarse 3D map is obtained from the robustly matched scene points via Delaunay triangulation [28]. Using the coarse 3D map, the visibility coefficient as well as ambient illumination is estimated using the physical light attenuation model described in Sections II-A and II-B.

In order to recover the original scene radiances for image enhancement (Section II-C), a dense and edge preserving depth map is mandatory. However, due to the degraded input images, only a coarse disparity mesh is available with many disparity gaps present which need to be filled. Therefore, we utilize an image matting approach [29] that refines the coarse disparity map and extrapolates information to unmatchable areas, based on the observed intensity values to form a "pseudo 3D map". It is an edge preserving and halo preventing algorithm but with the disadvantage that the resulting dense map does not contain the observed disparity values anymore. However, it is a good starting point for relatively fast image enhancement and scene structure recovery. In

[1] **Lib**rary for **E**fficient **La**rge-scale **S**tereo Matching: http://www.cvlibs.net

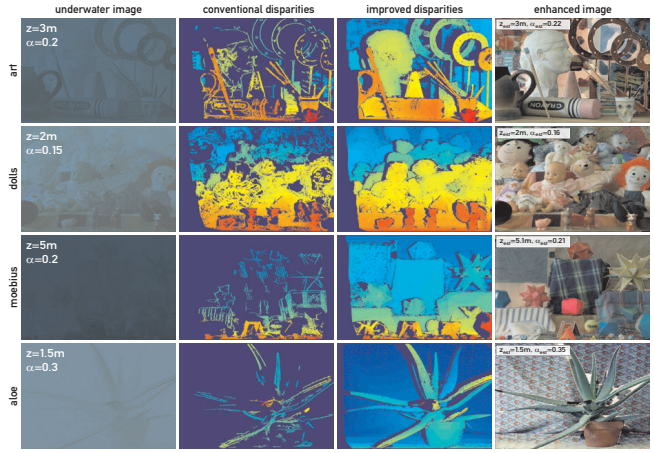


Fig. 5. Results on Middlebury Dataset. First column shows degraded underwater images, rendered with $z = 3m$, $\alpha = 0.2$ (*art*), $z = 2m$, $\alpha = 0.15$ (*dolls*), $z = 5m$, $\alpha = 0.2$ (*moebius*), $z = 1.5m$, $\alpha = 0.3$ (*aloe*). Second column shows their disparity maps. Third column shows the improved disparities (after step 1). Last column shows enhanced image (step 2).

very challenging images, as in some experimental results of Section IV, where the coarse 3D map covers only a minor part of the image, we additionally fill disparity gaps by robustly estimating the predominant 3D plane using *Random Sample Consensus (RANSAC)* [30].

In the second enhancement (Fig. 4, step 2), the procedure is repeated using the enhanced left and right images from (step 1) to create an "improved stereo" map. This map is further refined utilizing a discontinuity preserving gap interpolation [31], followed by image inpainting [32] for large areas with missing 3D information. Using this improved stereo map, detailed visibility parameters are estimated. These are then used with the gap interpolated map to recover scene radiances in the final enhancement (step 2).

III. Simulation Results

Images of complex underwater scenes are difficult to precisely ground-truth. Therefore, the accuracy of our method was investigated on the Middlebury data set[2], using the image pairs and ground-truth disparities from *aloe*, *art*, *dolls* and *moebius* at full resolution ($\approx 1300 \times 1100$ pixels). Artificial degradation of the images was performed to approximate the underwater survey requirements of an AUV in a depth range of $0 - 5m$ and a mean attenuation coefficient $\alpha = 0 - 0.3$ utilizing the model from Section II-A (without added noise).

Figure 5 shows results after applying the methodology described in Section II-E for different visibility conditions and depths. Although visibility degradation leads to a significant information loss in the acquired images they still provide sufficient information to estimate the visual range and to some extent considering SNR restore lost image parts. The original image chromaticity could also be restored. The main limitations of image enhancement stem from occluded image regions where no disparities can be matched (see Figure 5, *art*).

[2]http://vision.middlebury.edu/stereo/

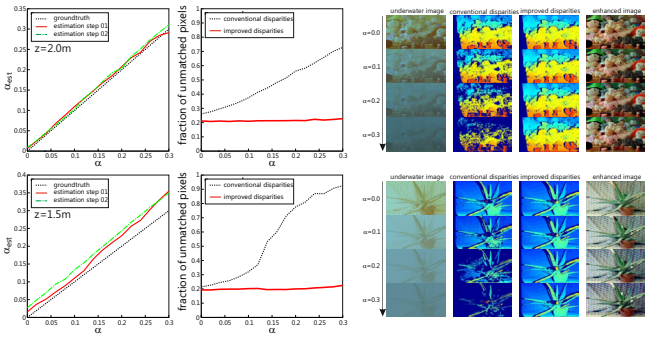


Fig. 6. Improved stereo results for decreasing visibility. Left column shows the estimated mean attenuation coefficient for decreasing visibility $\alpha = 0 - 0.3$ at depth $z = 2.0m$ (*dolls*) and $z = 1.5m$ (*aloe*), respectively. Center column depicts the fraction of unmatched pixels which is rapidly increasing in a conventional stereo matching framework. Upstream visibility estimation enables stereo matching with almost constant quality.

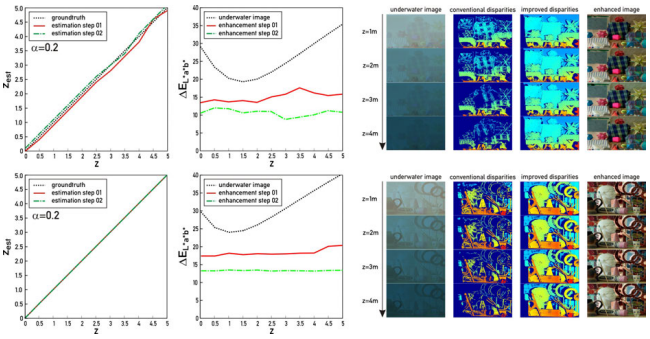


Fig. 7. Image restoration for increasing depth. Left column shows the estimated mean attenuation coefficient for increasing depth $z = 0 - 5m$ for $\alpha = 0.2$. Centre column depicts the color difference to the haze-free groundtruth image. Visual results are shown in the right column.

The improvement for stereo computation was evaluated for *aloe* at $z = 1.5m$ and *dolls* at $z = 2.0m$ with increasing mean attenuation coefficient $\alpha = 0 - 0.3$ (Figure 6). As shown the visibility conditions could be estimated reliably. Due to missing black/dark objects in *aloe*, the black patch assumption was violated which leads the proposed algorithm to slightly overestimate the current attenuation coefficient. When performing conventional stereo matching, the fraction of unmatched pixels is rapidly increasing from 0.21 to 0.92 (*aloe*) and from 0.26 to 0.72 (*dolls*) with adverse visibility. In contrast, our method proves to be robust to visibility changes, that is the fraction of unmatched pixels stays below 0.23 for both images over the whole range of visibility conditions.

The attenuation characteristic of ocean water leads to significant light chromaticity changes for increasing depths. A popular measure for perceived color differences with respect to the original image (above the water surface) is the *CIE 1976 L*,a*,b* color difference formula* [33] that computes the Euclidean point distance between two color points in the color-opponent and perceptually uniform CIE L*a*b* (ΔE_{L*a*b*}). A color difference of $\Delta E_{L*a*b*} \approx 2.3$ is the *Just noticeable difference (JND)* and is the minimum color deviation which can be distinguished by a human observer. As depicted in Figure 7, ΔE_{L*a*b*} between original image

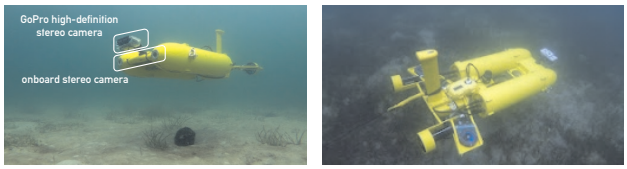


Fig. 8. The Starbug AUV was equipped with an onboard forward looking stereo camera pair. In addition, a GoPro high-definition stereo camera head with 32 mm baseline was mounted to the front of the vehicle and recorded during the experiments.

and underwater image rises with increasing depth $z = 1-5m$ from 19.28 to 35.3 (*moebius*) and from 23.87 to 40.36 (*art*), respectively. A first image enhancement leads to an almost constant color difference $\Delta E_{L*a*b*} \approx 15$ (*moebius*) and $\Delta E_{L*a*b*} \approx 18.1$ (*art*). A second restoration step, using the improved stereo information further reduces the color deviation to 10.7 and 13.1, respectively.

IV. Experimental Results

The underwater image enhancement algorithm presented above was evaluated on an AUV in uncontrolled ocean conditions. The purpose of these experiments was to evaluate the robustness and performance of the technique in terms of effective stereo range improvement and ability to correct color under different naturally lit operating scenarios.

A. AUV Platform

The Starbug MkIII AUV [34] was used for these experiments (Figure 8). The AUV's onboard vision system consists of two stereo camera pairs, one looking forward and the other looking downwards. For this study an external high-resolution stereo system was used for image collection. These calibrated GoPro™cameras (baseline 32mm) collected images at 30 fps with resolution 1280×960 and were time-synchronised to the AUV's onboard computer.

B. Experiments

A set underwater transects were performed at two sites in Queensland, Australia. The first was a shallow water, rather turbid, seagrass survey site off Peel Island (153.377^{o}E, 27.494^{o}S) in Moreton Bay. The second site was in the relatively clearer waters off Heron Island (151.8975^{o}E, 23.4504^{o}S) at the Southern end of the Great Barrier Reef.

1) Peel Island Transects: This area is adjacent to a marine protected zone and contains a range of seafloor types such as hard coral, rock, sand and seagrass. As it is a strongly tidal region, the visibility can vary greatly due to resuspension of sediments, incoming clear oceanic and turbid flood waters.

Four different length transects were performed with the AUV and covered water depths from approximately 0.3 to 3 m with the AUV travelling towards and away from the sun. The terrain was relatively planar sloping gently up towards and island with protruding coral heads, seagrass (Sargassum) and small rocks. During the experimental campaign the Secchi depth was measured from 2.5 - 4 m. Weather conditions were fair with the sun at its maximum height in the sky. In total, 1494 images were collected from the four transects.

2) Heron Island Transects: This survey site was on a steep sided reef slope and consisted of hard coral going down to a sand floor with a depth range of 1 - 8 m. Although the site had considerably greater visibility than in Moreton Bay, strong winds exceeding 25 knots during the experimental campaign had reduced the visibility to a Secchi depth of 5 - 8 m. In these experiments the horizontally mounted AUV stereo camera system was used to track another AUV as it maintained constant altitude above the reef surface. A total of 870 images were collected during four transects which travelled up and down the reef slope.

C. Results

A comparison between image enhancement using our approach against established histogram equalisation methods to increase image contrast are shown for representative images from Peel and Heron Island in Figure 9. The methods used for comparison are the standard MATLAB™functions for spatially-invariant histogram equalisation (histeq), and the spatially-variant adaptive histogram equalisation (adapthisteq) as employed in [8]. It can be seen that our results provide greater image detail and overall quality compared to the histogram methods for both turbidity (visibility) scenarios.

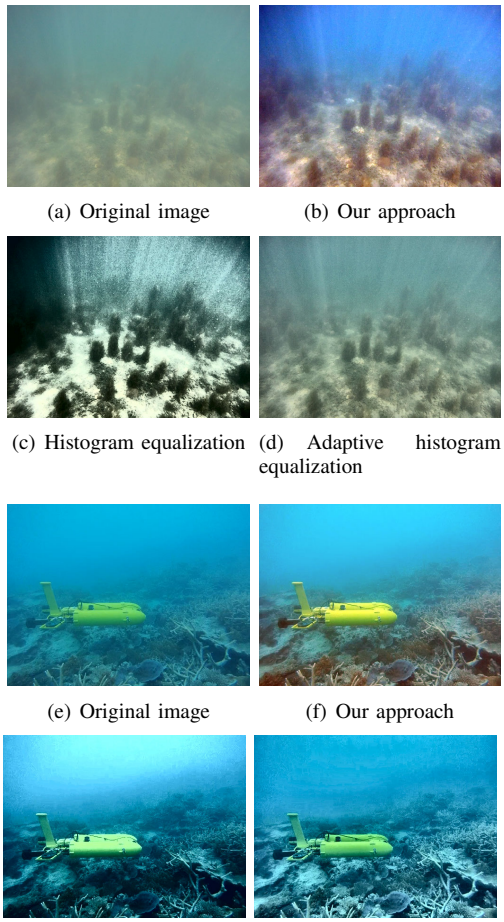
Figures 10(a) and 10(b) show the results for the estimated visibility coefficient, the visibility distance and the diving depth for both the Peel and Heron Island transects respectively. As seen the proposed approach yields stable and coherent estimation results for the range of depths and visibility conditions. However, there can be a slight over estimation of visibility distance when very close to the surface where the scene depth cannot be reliably estimated (surface occupies a significant portion of the image) and the visibility coefficient is underestimated.

The estimated depth profiles reflect the actual camera depth during transects, except for a constant offset of approximately 0.3 m in the Peel Island data. This is likely caused by the large amount of organic material present in the water column which may influence the ambient light chromaticity and the estimation performance as discussed in Section II-D.

The visibility coefficient estimate shows some outliers in Peel Island Sequence 4 that stem from low textured images captured in higher diving depths where no disparities could be reliably matched.

Figures 11(a) and 11(b) show representative images and their enhanced image and disparity maps for the Peel Island and Heron Island transects respectively. The enhanced images allow for automatic visual inspection and obtaining repeatable observations of benthic habitats. An empirical study by marine scientists has indicated that they are able to more reliably classify coral and seagrass species on the enhanced images over the original images, particularly for the Peel Island transects.

The disparity maps in these experimental conditions show only moderate improvements with up to 6.3% in unmatched pixels and range. There are four possible causes of this; (1) imprecise camera calibration (particularly for distant objects due to the short baseline and lack of a global shutter on the



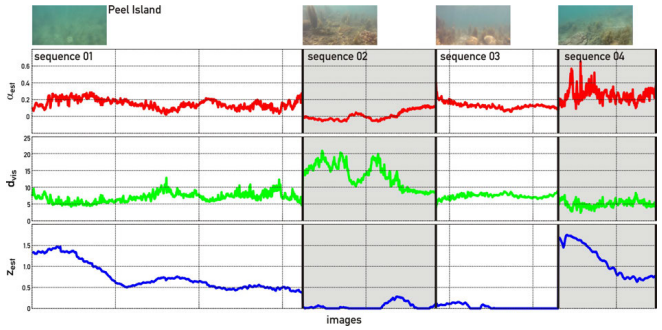
(a) Original image (b) Our approach

(c) Histogram equalization (d) Adaptive histogram equalization

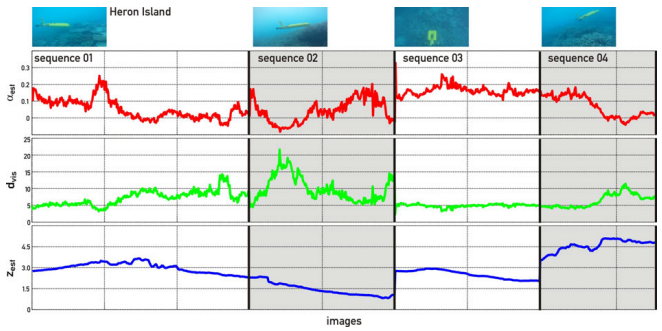
(e) Original image (f) Our approach

(g) Histogram equalization (h) Adaptive histogram equalization

Fig. 9. Comparison of our approach against histogram equalization methods for representative images from Peel Island (a)-(d) and Heron Island (e)-(h).



(a) Peel Island Transects



(b) Heron Island Transects

Fig. 10. Estimation results for the visibility coefficient (top trace), the visibility distance (center trace) and the estimated diving depth (bottom trace) using four underwater sequences acquired from different diving transects at (a) the high turbidity water at Peel Island, and (b) the moderate turbidity water at Heron Island.

Algorithm	Mean (s)	Std Deviation (s)
Sparse disparity mesh computation	0.1171	0.0149
Visibility estimation (step 1)	0.6139	0.1589
Image enhancement L (step 1)	0.3149	0.0123
Improved disparity computaton (including image enhancement R)	0.7705	0.0202
Visibility estimation (step 2)	0.5669	0.0502
Image enhancement (step 2)	0.1822	0.0179
TOTAL	2.5655	

TABLE I

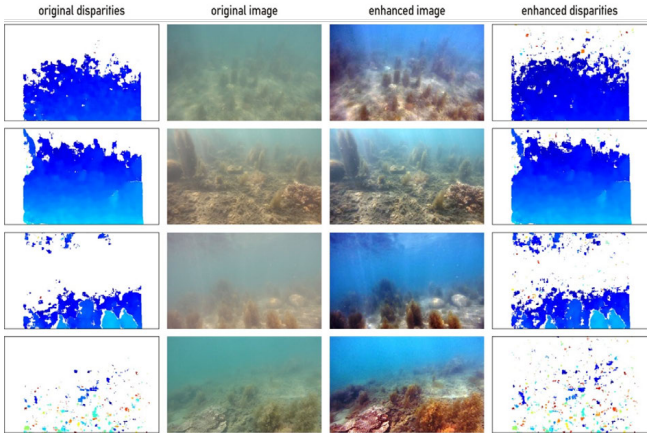
MEAN AND STANDARD DEVIATION COMPUTATION TIMES FOR EACH STEP IN THE PROPOSED IMAGE ENHANCEMENT METHOD.

stereo camera rig), (2) aliasing due to small-scale repetitive structures, (3) the planarity of the scene with the AUV travelling parallel to the seafloor with only the lower half containing features, and (4) more dominantly due to particles in the water and inhomogenously scattered light (*e.g.* from sun beams in the shallow water).

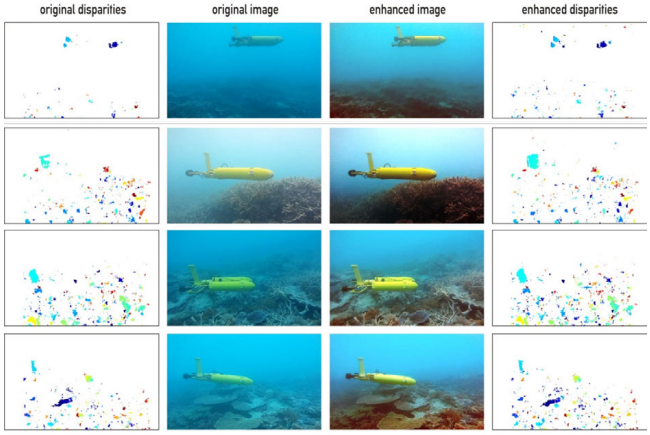
Table I illustrates the mean and standard deviation computation times for 640 $\times$ 480 pixel images using a single 2.4 GHz CPU running a non-optimized MATLAB^{TM}implementation of the proposed algorithms. Whilst not quite at framerate the results suggest the proposed methodology could be implemented onboard slower moving AUVs for improving navigation and obstacle avoidance.

V. CONCLUSIONS

This paper has described a novel method that simultaneously performs underwater image quality assessment, visibility enhancement and improved stereo estimation. We firstly propose the development of a coarse 3D map from the degraded input stereo images to allow visibility estimation and image enhancement using a physical model for atmospheric scattering adapted for use underwater in naturally lit environments. In a second step, the improved stereo map is used for a precise visibility estimation, image enhancement and 3D scene reconstruction. This approach was evaluated in simulation and using experimental data collected by an AUV in reduced visibility, naturally lit coastal waters. It is shown to improve stereo disparity estimation by up to 6.3% in the harsh experimental conditions experienced. This technique has proven robust and its performance has promise for implementation onboard an AUV and for image color correction to allow improved species identification by marine scientists. Future work will focus on investigating the robustness of stereo matching algorithms to improve the disparity image particularly in shallow water where surface waves cause rapidly varying lighting conditions.



(a) Peel Island



(b) Heron Island

Fig. 11. Example enhanced underwater images and disparity maps from (a) the high turbidity Peel Island Transects, and (b) the moderate turbidity Heron Island Transects.

REFERENCES

[1] R. Schettini and S. Corchs, "Underwater image processing: State of the art of restoration and image enhancement methods," *EURASIP Journal on Advances in Signal Processing*, p. pages, 2010.

[2] S. Duntley, "Light in the sea," *J. Opt. Soc. Amer.*, vol. 53, no. 2, pp. 214–233, 1963.

[3] R. C. Smith and K. S. Baker, "Optical properties of the clearest natural waters (200-800 nm)," *Applied Optics*, vol. 20, pp. 177–184, 1981.

[4] K. Iqbal, A. S.R., M. Osman, and A. Talib, "Underwater image enhancement using an integrated colour model," *International Journal Of Computer Science*, vol. 32, no. 2, pp. 239–244, 2007.

[5] J. Ahlen, D. Sundgren, and E. Bengtsson, "Application of underwater hyperspectral data for color correction purposes," *Pattern Recognition and Image Analysis*, vol. 17, pp. 170–173, 2007.

[6] J. Chiang and Y.-C. Chen, "Underwater image enhancement by wavelength compensation and dehazing," *Image Processing, IEEE Transactions on*, vol. 21, no. 4, pp. 1756 –1769, april 2012.

[7] J. Kaeli, H. Singh, C. Murphy, and C. Kunz, "Improving color correction for underwater image surveys," in *OCEANS 2011*, sept. 2011, pp. 1 –6.

[8] N. Carlevaris-Bianco, A. Mohan, and R. Eustice, "Initial results in underwater single image dehazing," in *OCEANS 2010*, 2010, pp. 1–8.

[9] D. Kocak, F. Dalgleish, F. Caimi, and Y. Schechner, "A focus on recent developments and trends in underwater imaging," *Marine Technology Society Journal*, vol. 42, no. 1, pp. 52–67, 2008.

[10] Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar, "Instant dehazing of images using polarization," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR '01)*, 2001, pp. 325–332.

[11] Y. Schechner and N. Karpel, "Recovery of underwater visibility and structure by polarization analysis," *Oceanic Engineering, IEEE Journal of*, vol. 30, no. 3, pp. 570 –587, 2005.

[12] S. Narasimhan and S. Nayar, "Vision and the atmosphere," *International Journal of Computer Vision*, vol. 48, no. 3, pp. 233–254, 2002.

[13] S. G. Narasimhan and S. K. Nayar, "Contrast restoration of weather degraded images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 713–724, 2003.

[14] F. Cozman and E. Krotkov, "Depth from scattering," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR '97)*, 1997.

[15] S. Narasimhan and S. Nayar, "Interactive deweathering of an image using physical models," in *IEEE Workshop on Color and Photometric Methods in Computer Vision (in conjunction with ICCV '03)*, 2003.

[16] N. Hautière, J.-P. Tarel, and D. Aubert, "Towards fog-free in-vehicle vision systems through contrast restoration," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR '07)*, Minneapolis, Minnesota, USA, 2007, pp. 1–8.

[17] J. Kopf, B. Neubert, B. Chen, M. F. Cohen, D. Cohen-Or, O. Deussen, M. Uyttendaele, and D. Lischinski, "Deep photo: Model-based photograph enhancement and viewing," *ACM Transactions on Graphics (Proc. of SIGGRAPH Asia 2008)*, vol. 27, no. 5, pp. 116:1–116:10, 2008.

[18] R. Fattal, "Single image dehazing," *ACM Transactions on Graphics*, vol. 27, pp. 72:1–72:9, August 2008.

[19] R. Tan, "Visibility in bad weather from a single image," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR '08)*, 2008, pp. 1 –8.

[20] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior." in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR '09)*. IEEE, 2009, pp. 1956–1963.

[21] L. A. Torres-Méndez and G. Dudek, "Color correction of underwater images for aquatic robot inspection," in *EMMCVPR*, 2005, pp. 60–73.

[22] H.-Y. Yang, P.-Y. Chen, C.-C. Huang, Y.-Z. Zhuang, and Y.-H. Shiau, "Low complexity underwater image enhancement based on dark channel prior," in *Innovations in Bio-inspired Computing and Applications, 2011 Second International Conference on*, 2011, pp. 17–20.

[23] H. Koschmieder, *Dynamische Meteorologie*, 2nd ed., ser. Physik der Atmosphaere ; 2. Leipzig: Akad. Verl.-Ges., 1941.

[24] Y. Schechner and N. Karpel, "Clear underwater vision," in *Proc. 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, 2004, pp. I–536 – I–543 Vol.1.

[25] T. Coleman and Y. Li, "An interior, trust region approach for nonlinear minimization subject to bounds," *SIAM Journal on Optimization*, vol. 6, pp. 418–445, 1996.

[26] CIE17.4-1987, "International lighting vocabulary," Commission Internationale de L'Eclairage, Tech. Rep., 1987.

[27] A. Geiger, M. Roser, and R. Urtasun, "Efficient large-scale stereo matching," in *Asian Conference on Computer Vision (ACCV '10)*, Queenstown, New Zealand, 2010.

[28] J. R. Shewchuk, "Triangle: Engineering a 2D Quality Mesh Generator and Delaunay Triangulator," in *Applied Computational Geometry: Towards Geometric Engineering*. Springer-Verlag, May 1996, vol. 1148, pp. 203–222.

[29] A. Levin, D. Lischinski, and Y. Weiss, "A closed-form solution to natural image matting," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 30, no. 2, pp. 228 –242, 2008.

[30] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, June 1981.

[31] H. Hirschmueller, "Stereo processing by semiglobal matching and mutual information," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 30, no. 2, pp. 328 –341, 2008.

[32] A. Telea, "An image inpainting technique based on the fast marching method," *Journal of Graphics, GPU, and Game Tools*, vol. 9, no. 1, pp. 23–34, 2004.

[33] A. K. Jain, *Fundamentals of digital image processing*, ser. Prentice Hall information and system sciences series. Englewood Cliffs, N.J.: Prentice Hall, 1989.

[34] M. Dunbabin, J. Roberts, K. Usher, G. Winstanley, and P. Corke, "A hybrid AUV design for shallow water reef navigation," in *Proc. IEEE International Conference on Robotics and Automation*, Barcelona, Spain, 18-22 April 2005, pp. 2105–2110.