Illumination and Attenuation Correction Techniques for Underwater Robotic Optical Imaging Platforms

Jeffrey W. Kaeli and Hanumant Singh

***Abstract*—We demonstrate a novel method of correcting illumination and attenuation artifacts in underwater optical imagery. These artifacts degrade imagery and hinder both human analysis as well as automated classification algorithms. Our approach estimates separately the attenuation coefficient of the water column and the beam pattern of the illumination source using sequences of overlapping color images and acoustic ranges from a Doppler Velocity Log (DVL). These values are then used in the correction step to remove color and intensity artifacts with the overarching goal of more consistent results for input into classification algorithms.**

***Index Terms*—color correction, computational photography, multi-sensor fusion, underwater imaging, underwater robotics**

I. Introduction

We have better maps of the surfaces of Venus, Mars, and our moon than we do of the seafloor beneath Earth's oceans [1], primarily because, in many respects, the imagery is easier to obtain. Water is a strong attenuator of electromagnetic radiation, so while satellites can map entire planets from space using cameras and laser ranging, underwater vehicles must be within tens of meters at best for optical sensors to be useful. While mechanical waves do travel well through water, there are practical tradeoffs between source strength, frequency, and propagation distance. Ship-based sonars use lower frequencies to reach the bottom, but these longer wavelengths come at the price of reduced resolution. To map fine-scale features relevant to many practical applications, both optical and acoustic imaging platforms must operate relatively close to the seafloor. We are particularly interested in optical imaging because it captures the color and texture information useful for distinguishing habitats and organisms.

An underwater photograph not only captures the scene of interest, but is an image of the water column as well. Attenuation of light underwater is caused by absorption, a thermodynamic process that varies nonlinearly with wavelength, and by scattering, a mechanical process whereby a photon's direction is changed [2], [3]. At increasing depths, ambient light is attenuated to where colors can no longer be distinguished and eventually to effective darkness. Artificial light sources must subsequently be used to illuminate the scene, but these sources contribute to scattering and can introduce beam pattern artifacts in the image. In summary, uncorrected underwater imagery is typically characterized by non-uniform illumination, reduced contrast, and colors that are saturated in the green and blue channels.

The authors are with the Woods Hole Oceanographic Institution, Woods Hole, MA, 02543 e-mail: (see http://www.whoi.edu/people/jkaeli).

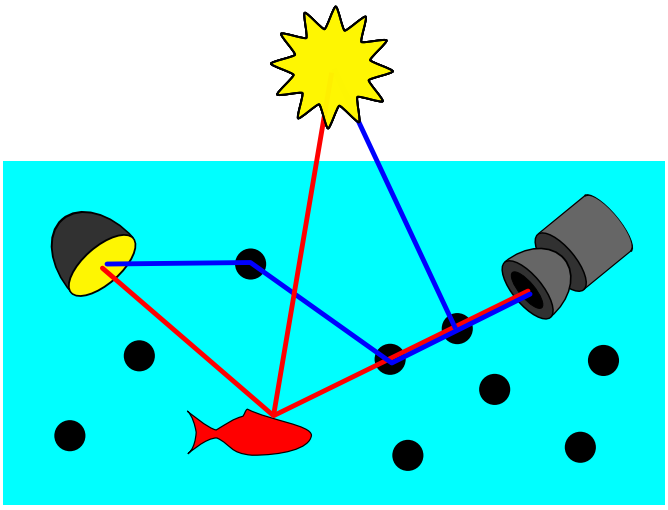

Fig. 1. Capturing an underwater image. Light originating from the surface and/or an artificial source reflects off of an object (the fish) and toward the camera along a direct path (red line) or is scattered by particles along infinite possible paths into the camera's field of view (blue lines).

It is often desirable for an underwater image to appear as if it were taken in air, either for aesthetics or as a pre-processing step for automated classification. Methods range from purely post-processing techniques to novel hardware configurations, and the choice depends heavily on the imaging system, the location, and the goals of the photographer. In this paper, we first build a model of underwater image formation, describing how various artifacts arise. Next, we discuss a variety of common methods used to correct for these artifacts in the context of different modes of underwater imaging. Lastly, we present a novel method of correction for robotic imaging platforms that estimates environmental and system parameters using multi-sensor fusion.

II. Underwater Image Formation

An underwater photograph not only captures the scene of interest, but is an image of the water column as well. Figure 1 diagrams a canonical underwater imaging setup. Light rays originating from the sun or an artificial source propagate through the water and reach the camera lens either by a direct path or by an indirect path through scattering. We deal with each of these effects in turn.

A. Attenuation

The power associated with a collimated beam of light is diminished exponentially as it passes through a medium in

accordance with the Beer-Lambert Law

$$P_\ell(\lambda) = P_0(\lambda)\ e^{-\alpha(\lambda)\ell} \tag{1}$$

where P_0 is the source power, P_ℓ is the power at a distance ℓ through the medium, λ is wavelength, and α is the wavelength-dependent attenuation coefficient of the medium [2]. Attenuation is caused by absorption, a thermodynamic process that varies with wavelength, and by scattering, a mechanical process whereby a photon's direction is changed.

$$\alpha(\lambda) = \alpha_a(\lambda) + \alpha_s \tag{2}$$

where $\alpha_a(\lambda)$ and α_s are the medium absorption and scattering coefficients, respectively. Scattering underwater is largely wavelength-independent because the scattering particle sizes are much larger than the wavelength of light. Underwater scenes generally appear bluish green as a direct result of water more strongly absorbing red light than other wavelengths. However, the attenuation properties of water vary greatly with location, depth, dissolved substances and organic matter [3].

B. Natural Lighting

Natural illumination E_n from sunlight S_n attenuates exponentially with depth z and can be characterized by $\bar{K}(\lambda)$, the average spectral diffuse attenuation coefficient for spectral downwelling plane irradiance.

$$E_n(\lambda, z) = S_n(\lambda)\ e^{-\bar{K}(\lambda)z} \tag{3}$$

While related to α, the diffuse attenuation coefficient represents the sum of all light arriving at a given depth that has been attenuated along infinitely many scattered paths. It is strongly correlated with phytoplankton chlorophyll concentrations and is often measured in remote sensing applications [3].

C. Artificial Lighting

At a certain depth, natural light is no longer sufficient for illumination, so artificial lights must be used. For robots that operate untethered from ship power, such as Autonomous Underwater Vehicles (AUVs), this limits the available energy for lighting and thus beam pattern artifacts are common. We can model the artificial illumination pattern E_a from a single source as

$$E_a(\lambda) = S_a(\lambda)\ BP_{\theta,\phi} \frac{e^{-\alpha(\lambda)\ell_a}}{\ell_a^2}\ \cos\gamma \tag{4}$$

where S_a is the source spectrum, $BP_{\theta,\phi}$ is the angularly-dependent beam pattern of the source, ℓ_a is the path length between the source and the scene, and γ is the angle between the source and surface normal assuming a Lambertian surface [4]. In practice, imaging platforms may carry one or multiple light sources, but in our model we assume a single source for simplicity.

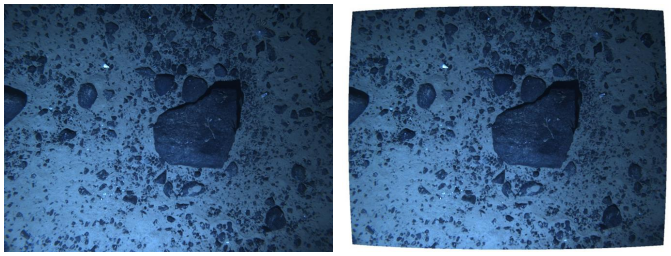

Fig. 2. A raw underwater image (left) has been rectified to correct for lens distortion underwater (right).

D. Diffuse Lighting

Light that is scattered back into the camera's line of sight is known as backscatter, a phenomenon similar to fog in the atmosphere [5]. If we denote $F(\lambda, z)$ to be the diffuse light field at any given point, we can recover the backscatter by integrating the attenuated field along a camera ray.

$$E_b(\lambda) = \int_0^{\ell_s} F(\lambda, z)\ e^{-\alpha(\lambda)\ell} d\ell \tag{5}$$

Under the assumption that $F(\lambda, z) \approx F(\lambda)$ is uniform over the scene depth ℓ_s, then

$$E_b(\lambda) = A(\lambda)\left(1 - e^{-\alpha(\lambda)\ell_s}\right) \tag{6}$$

where $A(\lambda) = \frac{F(\lambda)}{\alpha(\lambda)}$ is called the airlight. This additive light field reduces contrast and creates an ambiguity between scene depth and color saturation [6].

E. Camera Lens

The lens of the camera gathers light and focuses it onto the optical sensor. Larger lenses are preferable underwater because they are able to gather more light in an already light-limited environment. The lens effects L can be modeled as

$$L = \left(\frac{D_L}{2}\right)^2 \cos^4\theta_L\ T_L\ \left(\frac{Z_s - F_L}{Z_s F_L}\right)^2 \tag{7}$$

where D_L is the diameter of the lens, θ_L is the angle from lens center, T_L is the transmission of the lens, and Z_s and F_L are the distance to the scene and the focal length, respectively. Assuming there are no chromatic aberrations, the lens factors are wavelength-independent. A detailed treatment of this can be found in McGlamery and Jaffe's underwater imaging models [4], [7].

Subsequent computations may require that the projected ray direction in space is known for each pixel. Because the refractive index of water differs from that of air, the camera lens can be calibrated to account for distortion using the method described in [8]. An image can subsequently be warped such that each pixel is aligned with its appropriate location in space, as shown in Figure 2.

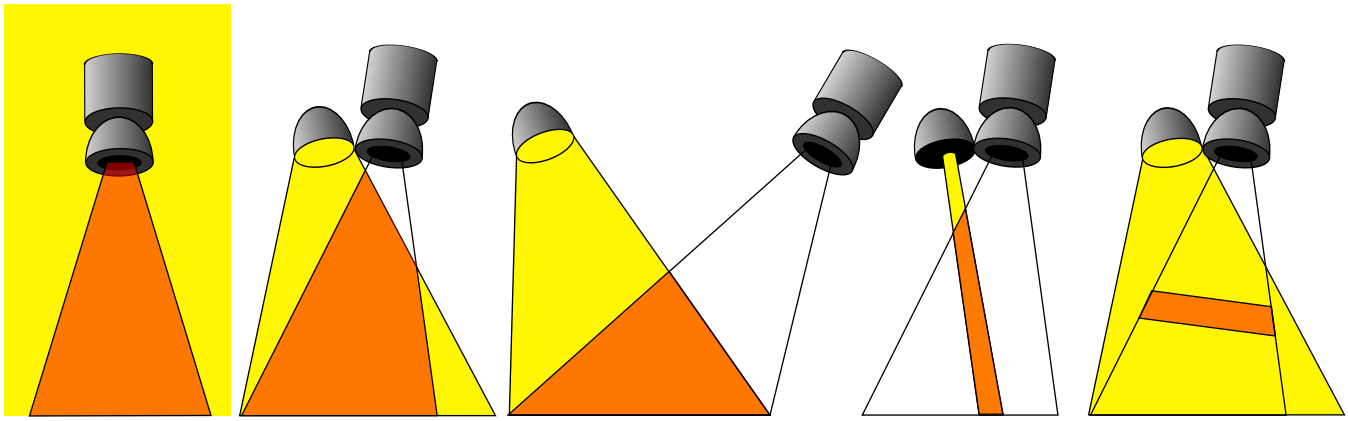

Fig. 3. Backscatter is a direct result of the intersection (shown in orange) between the illumination field (yellow) and the field of view of the camera. In ambient light (far left) and when using a coincident source and camera (second from left) the entire field of view is illuminated. Separating the source from the camera (middle) results in a reduction of backscattering volume. Structured illumination (second from right) and range gating (far right) drastically reduce the backscattering volume and are ideal for highly turbid environments.

F. Optical Sensor

Since most color contrast is lost after several attenuation lengths, early underwater cameras only captured grayscale images. The intensity of a single-channel monochrome image $\boldsymbol{c}$ can be modeled as the integrated product of the sensor's spectral response function ρ with the incoming light field

$$\boldsymbol{c} = \int E(\lambda)\boldsymbol{r}(\lambda)\rho(\lambda)\ d\lambda \tag{8}$$

where E is the illuminant and $\boldsymbol{r}$ is the reflectance of the scene. Bold variables denote pixel-dependent terms in the image. Creating a color image requires sampling over multiple discrete spectral bands Λ, each with spectral response function ρ_Λ. The human visual system does precisely this, using three types of cone-shaped cells in the retina that measure short (blue), medium (green), and long (red) wavelengths of light, known as the tristimulus response. Modern digital cameras have been modeled after human vision, with many employing a clever arrangement of red, green, and blue filters known as a Bayer pattern across the sensor pixels. This multiplexing of spatial information with spectral information must be dealt with in post-processing through demosaicking [9], [10].

A color image can be similarly modeled as

$$\boldsymbol{c}_\Lambda = \int E(\lambda)\boldsymbol{r}(\lambda)\rho_\Lambda(\lambda)\ d\lambda \approx E_\Lambda \boldsymbol{r}_\Lambda. \tag{9}$$

The illuminant and reflectance can be aproximated in terms of the camera's red, green, and blue channels $\Lambda = \{R, G, B\}$ with the understanding that they actually represent a spectrum [11]. By adjusting the relative gains of each channel, known as von Kries-Ives adaptation, one can transform any sensor's response into a common color space through simple linear algebra.

G. Imaging Model

Putting the pieces together, we arrive at a model with both multiplicative terms from the direct path and additive terms from the indirect scattered light field.

$$\boldsymbol{c}_\Lambda = G\left[\left(\boldsymbol{E}_{n,\Lambda} + \boldsymbol{E}_{a,\Lambda}\right)\boldsymbol{r}_\Lambda \frac{e^{-\alpha_\Lambda \boldsymbol{\ell}_s}}{\boldsymbol{\ell}_s^2} + \boldsymbol{E}_{b,\Lambda}\right)\right]\boldsymbol{L} \tag{10}$$

G is an arbitrary camera gain. We ignore forward scattering from our model because its contributions are insignificant for standard camera geometries [12].

III. Review of Correction Techniques

Removing the effects of the water column from underwater images is a challenging problem, and there is no single approach that will outperform all others in all cases. The choice of method depends heavily on the imaging system used, the goals of the photographer, and the location where they are shooting.

Imaging systems can range from a diver snapping tens of pictures with a handheld camera to robotic platforms capturing tens of thousands of images. Where divers often rely on natural light, robotic imaging platforms such as AUVs often dive deeper and carry artificial lighting. AUVs generally image the seafloor indiscriminately looking straight down from a predefined altitude, while divers are specifically advised to avoid taking downward photographs and get as close as possible [13]. One individual may find enhanced colors to be more beautiful, while a scientist's research demands accurate representation of those colors. Similarly, a human annotating a dataset might benefit from variable knobs that can enhance different parts of the images, while a computer annotating a dataset demands consistency between corrected frames.

Lighting and camera geometry also play huge roles in the subsequent quality of underwater imagery. Figure 3 shows the effect that camera - light separation has on the additive backscatter component. Images captured over many attenuation lengths, such as a horizontally facing camera pointed towards the horizon, suffer more from backscatter than downward looking imagery captured from 1-2 attenuation lengths away. In many situations, the additive backscatter component can be ignored completely, while in highly turbid environments, more exotic lighting methods may be required.

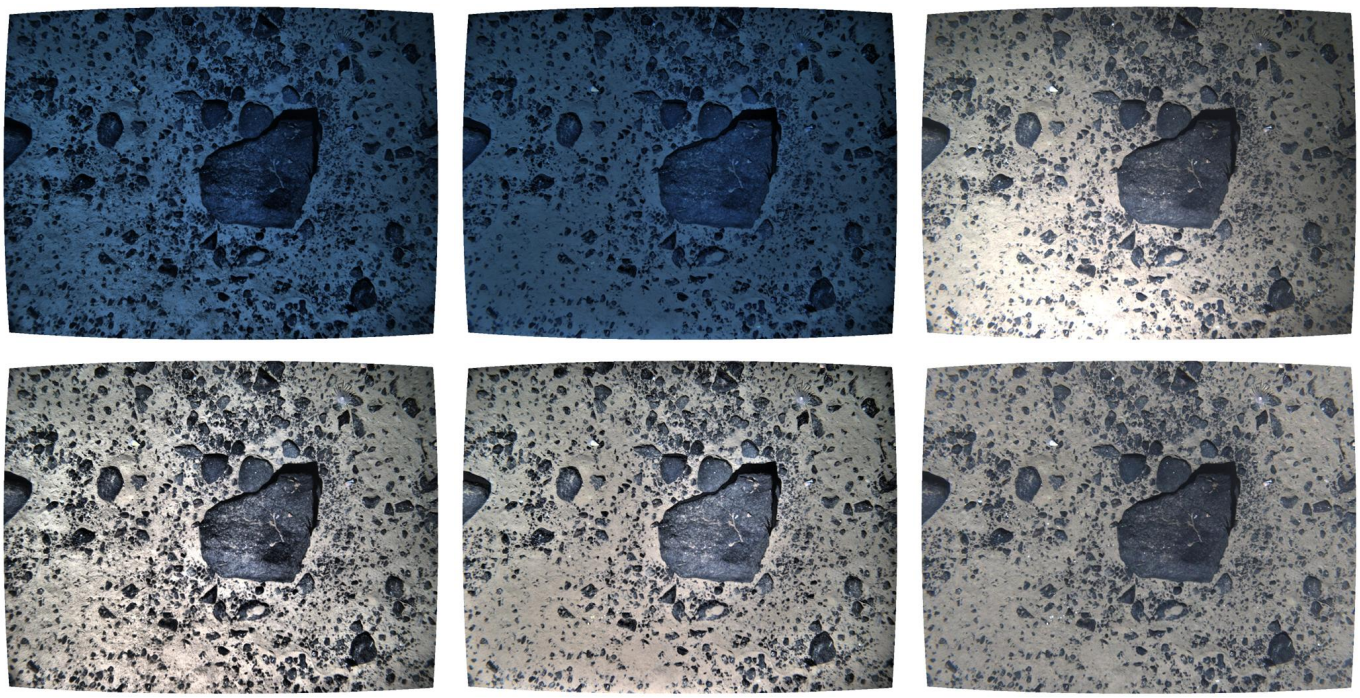

Fig. 4. Example methods of correcting illumination and attenuation artifacts in the underwater image from Figure 2. Adaptive histogram equalization (top left) and homomorphic filtering (top middle) attempt to remove the non-uniform illumination pattern, but do not correct for attenuation. White balancing (top right) attempts to correct for attenuation, but does not remove the non-uniform illumination artifacts. Applying white balancing to adaptive histogram equalization (bottom left) and homomorphic filtering (bottom middle) corrects for both illumination and attenuation but can distort colors and leave haloing artifacts around sharp gradients. Frame averaging (bottom right) also corrects for both illumination and attenuation. The bottom row represents several state-of-the-art methods currently used to batch-process large volumes of downward-looking underwater transect imagery captured from robotic platforms.

A. Shallow, Naturally Lit Imagery

Images captured in shallow water under natural illumination often contain a strong additive component. Assuming a pinhole camera model, image formation can be elegantly written as a matteing problem

$$\boldsymbol{c}_\Lambda = \boldsymbol{J}_\Lambda \; \boldsymbol{t} + (1 - \boldsymbol{t} \;) \, A_\Lambda \tag{11}$$

where $\boldsymbol{J}_\Lambda = \frac{\boldsymbol{E}_{n,\Lambda}}{\boldsymbol{\ell}_s^2} \boldsymbol{r}_\Lambda$ and $\boldsymbol{t} = e^{-\alpha_\Lambda \boldsymbol{\ell}_s}$ is the "transmission" through the water. Dehazing algorithms [6], [14] are able to both estimate the color of the airlight and provide a metric for range which are used to remove the effects of the airlight. Since the scattering leads to depolarization of incident light, other methods employ polarizing filters to remove airlight effects [15]. However, these methods do not attempt to correct for any attenuation effects.

B. Enhancing Contrast

Several methods simply aim at enhancing the contrast that is lost through attenuation and scattering. Adaptive histogram equalization performs spatially varying histogram equalization over image subregions to compensate for non-uniform illumination patterns in grayscale imagery [12]. Homomorphic methods work in the logarithmic domain, where multiplicative terms become linear. Examples of both are shown in Figure 4. Assuming that the illumination field $\boldsymbol{I}_\Lambda = (\boldsymbol{E}_{n,\Lambda} + \boldsymbol{E}_{a,\Lambda}) \, \frac{e^{-\alpha_\Lambda \boldsymbol{\ell}_s}}{\boldsymbol{\ell}_s^2}$ contains lower spatial frequencies than the reflectance image, and ignoring (or previously having corrected for) any additive components,

$$\log \boldsymbol{c}_\Lambda = \log \boldsymbol{I}_\Lambda + \log \boldsymbol{r}_\Lambda, \tag{12}$$

the illumination component can be estimated though low-pass filtering [16] or surface fitting [17] and removed to recover the reflectance image. These methods work well for grayscale imagery, can be applied to single images, and do not require any *a priori* knowledge of the imaging setup. However, they can sometimes induce haloing around sharp intensity changes, and processing color channels separately can lead to misrepresentations of actual colors. Other contrast enhancement methods model the point spread function of the scattering medium and recover reflectance using the inverse transform [18].

C. High-Turbidity Environments

Some underwater environments have such high turbidity or require an altitude of so many attenuation lengths that the signal is completely lost in the backscatter. Several more "exotic" methods utilizing unique hardware solutions are diagrammed in Figure 3. Light striping [19]–[22] and range gating [23] are both means of shrinking or eliminating, respectively, the volume of backscattering particles. Confocal imaging techniques have also been applied to see through foreground haze occlusions [24].

D. Restoring Color

The effects of attenuation can be modeled as a spatially varying linear transformation of the color coordinates

$$\boldsymbol{c}_\Lambda = \boldsymbol{I}_\Lambda \boldsymbol{r}_\Lambda \tag{13}$$

where $\boldsymbol{I}_\Lambda$ is the same illumination component defined in Equation 12. The reflectance image can be recovered simply by multiplying by the inverse of the illumination component. Assuming that the illumination and attenuation $\boldsymbol{I}_\Lambda \approx I_\Lambda$ are uniform across the scene, this reduces to a simple white balance via the von Kries-Ives adaptation [11]. The white point can be set as the image mean under the grey world assumption, a manually selected white patch, or as the color of one of the brightest points in the image [25], [26]. This method achieves good results for some underwater images, but performs poorly for scenes with high structure. Results can also be negatively affected when the grey world assumption is violated, for instance a large colored object which shifts the mean color of the image. Figure 4 shows the effects of white balancing both on a raw image and on contrast enhanced imagery. Recent work in spatially varying white balance [27] deals with correcting for multiple illumination sources and may have applications to underwater images as well.

More computationally involved methods include fusion-based approaches that combine the "best" result of multiple methods for color correction and contrast enhancement [28]. Markov Random Fields have been used with statistical priors learnt from training images to restore color [29]. A novel hardware solution to restoring color employs colored strobes to replace the wavelengths lost via attenuation [30].

E. Beyond a Single Image

Additional information beyond that contained in a single image can be useful for correcting a series of underwater images. The simplest method is to compute the average across many image frames

$$\frac{1}{K}\sum_k^K \boldsymbol{c}_{\Lambda,k} \approx \boldsymbol{I}_\Lambda \; \frac{1}{K}\sum_k^K \boldsymbol{r}_{\Lambda,k} = \boldsymbol{I}_\Lambda \bar{\boldsymbol{r}}_\Lambda \tag{14}$$

under the assumption that the illumination component does not vary between images. This assumption is valid for many types of robotic surveys where a constant altitude is maintained over a relatively uniform seafloor [31], [32]. Correction is then akin to that of a spatially varying white balance where the white point of each pixel is the mean over the dataset. An example is shown in Figure 4.

Robotic platforms often carry additional sensors other than a single camera and light source. An acoustic altimeter can provide information to perform range-dependent frame averaging useful for towed systems where a constant altitude is difficult to maintain [33]. However, this approach fails when the bottom is not flat relative to the imaging platform. Multiple acoustic ranges, such as those obtained from a Doppler Velocity Log (DVL), can be used under the assumption that the bottom is locally planar [34]. Stereo camera pairs [35] or a sheet laser in the camera's field of view [36] can similarly provide bathymetry information for modeling attenuation path lengths.

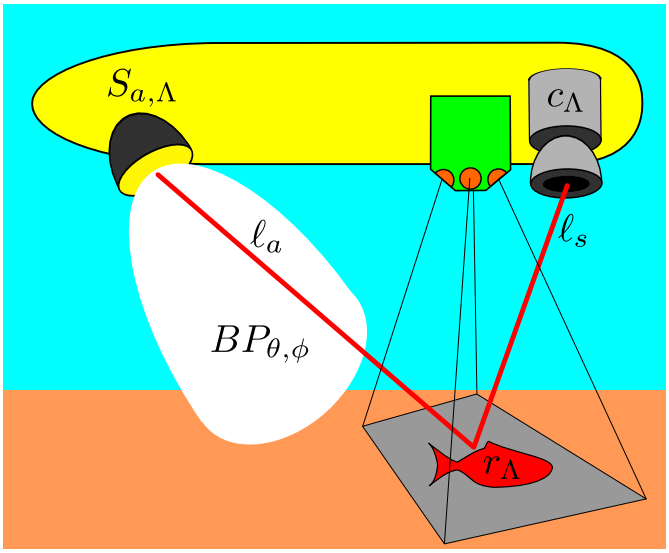

Fig. 5. Diagram of a typical underwater robotic imaging platform setup. The camera and light are separated to reduce backscatter, and a DVL (in green) mounted near the camera is used to estimate a bottom plane (in grey).

IV. Correction for Robotic Imaging Platforms

In addition to cameras and artificial light sources, robotic imaging platforms generally carry a suite of navigational sensors as well. One such sensor in widespread use is the Dopper Velocity Log (DVL) which measures both range and relative velocity to the seafloor using 4 acoustic beams [37]. Figure 5 diagrams a common configuration for many robotic imaging platforms. The camera and light source are separated to reduce backscatter, and the DVL is mounted adjacent to the camera so its beams encompass the field of view. Unlike many correction methods for single images that rely on assumptions such as low frequency illumination patterns, we exploit multiple images and additional sensor information to estimate the unknown parameters of the imaging model and use this to obtain more consistent image correction.

A. Assumptions

We first assume that our images are captured deep enough so that natural light is negligible relative to artificial lighting. We also assume there is a single strobe, and its spectrum $S_\Lambda \approx \{1, 1, 1\}$ is approximately white. This is an acceptable assumption because, while deviations in the strobe spectrum will induce a hue shift in the corrected reflectance image, this shift will be constant over all images in a dataset. Thus, even a strongly colored strobe would have no effect on automated classification results assuming the training and testing were both performed with corrected imagery.

Next, we assume that we can ignore the additive effects of scattered diffuse lighting. This is a valid assumption for images captured in relatively clear water within a few meters of the seafloor, as demonstrated in Figure 6. The log of each color channel mean for 3000 images has been plotted as a function of vehicle altitude over the course of a mission. At high altitudes, diffuse light from the scattered illumination field dominates, asymptotically approaching the airlight

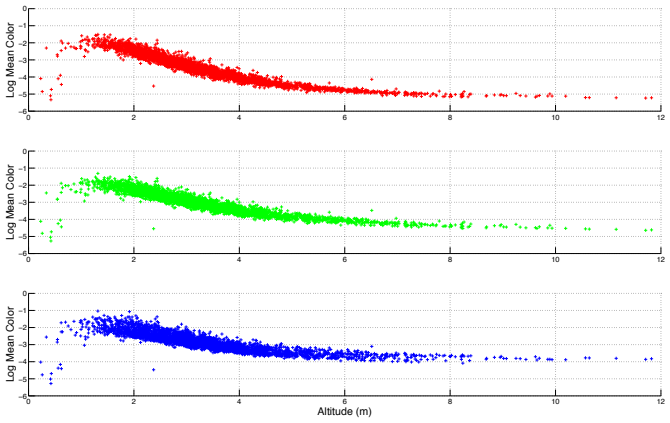

Fig. 6. Log color channel means (colored respectively) as a function of altitude for over 3000 images captured along a transect. Note the linearity within the first few meters, suggesting that additive effects can be ignored within this regime. Diffuse lighting dominates at higher altitudes, asymptotically approaching the airlight color. The falloff at very low altitudes is due to the strobe beam leaving the camera's field of view.

color. Within the first few meters, however, the relationship is roughly linear, indicating the relative absence of additive scattering effects.

Neglecting additive components allows us to work in the logarithmic domain, where our image formation model becomes a linear combination of terms. Approximating the seafloor as a locally planar surface, we can neglect the Lambertian term $\cos\gamma$ as it will vary little over the image. The gain G and lens $\boldsymbol{L}$ terms are constant between images and effect only the brightness but not the color. Omitting them as well, our model reduces to

$$\log \boldsymbol{c}_\Lambda = \log \boldsymbol{r}_\Lambda + \log \boldsymbol{BP}_{\theta,\phi} - \alpha_\Lambda(\boldsymbol{\ell}_a + \boldsymbol{\ell}_s) - 2\log \boldsymbol{\ell}_a\, \boldsymbol{\ell}_s. \quad (15)$$

From this we can clearly see three processes corrupting our underwater image. Firstly, the beam pattern of the strobe creates a non-uniform intensity pattern across the image as a function of beam angles θ and ϕ. Second is attenuation, which is wavelength-dependent and directly proportional to the total path length $\boldsymbol{\ell} = \boldsymbol{\ell}_a + \boldsymbol{\ell}_s$. Lastly, there is spherical spreading, which in practice we have found to be less significant than the exponential attenuation, supported by [38], and henceforth omit from our calculations.

At the moment, the entire right hand side of Equation 15 consists of unknowns. However, using the 4 range values from the DVL, and with *a priori* knowledge of offsets between sensors, we can fit a least squares local plane to the seafloor and compute the values of $\boldsymbol{\ell}$, $\boldsymbol{\theta}$, and $\boldsymbol{\phi}$ for each pixel in the image. Although the vast majority of DVL pings result in 4 usable ranges, sometimes there are unreturned pings. In the case of three pings, a least squares fit reduces to the exact solution. For one or two returns, the bottom is simply assumed to be flat, although these cases are rare.

B. Attenuation Coefficient Estimation

For the moment, let us assume that the beam pattern $\boldsymbol{BP}_{\theta,\phi} \approx 1$ is uniform across the image. For each pixel in each color channel, we have one equation but 2 unknowns: the attenuation coefficient α_Λ and the reflectance value r_Λ that we are trying to recover. However, if that same point is imaged from another pose with a different path length, an equation is added and the system can be constrained. For the purposes of navigation and creating photomosaics, finding shared features between overlapping images is a common problem. Keypoints can be reliably detected and uniquely described using methods such as Harris corners and Zernike moments [32] or SIFT features [39]. An example of two overlapping images with matched keypoints is shown in Figure 7.

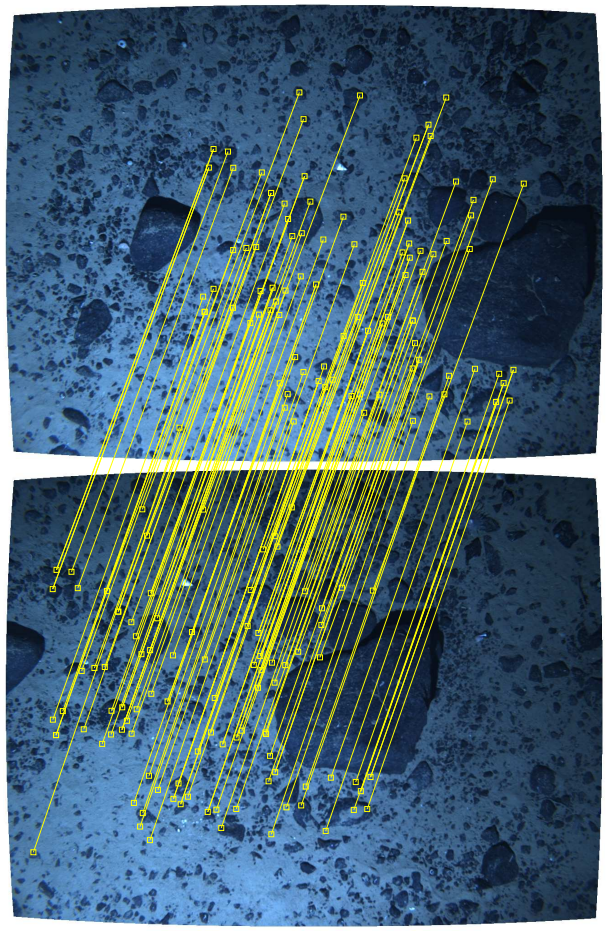

Fig. 7. A pair of overlapping images with matched keypoints.

For each pair of matched keypoints, the average local color value is computed using a Gaussian with standard deviation proportional to the scale of the keypoint. Assuming that the corrected values of both colors should be the same, we can explicitly solve for the attenuation coefficients

$$\alpha_\Lambda = \frac{\log c_{\Lambda,1} - \log c_{\Lambda,2}}{\ell_2 - \ell_1}. \quad (16)$$

The mean values of α_Λ were calculated for each of 100 images. Values less than 0.1 were considered unrealistic and omitted. This accounted for 20% of the images. The results are plotted at the top of Figure 8.

This method is feasible for as few as two images assuming that there is overlap between them and enough structure

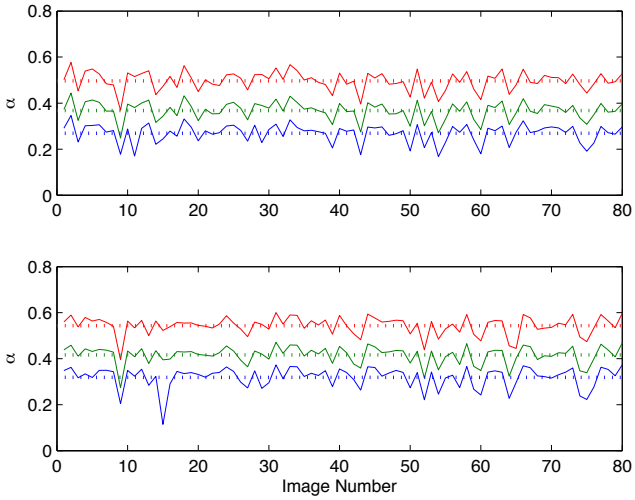

Fig. 8. Estimated α_Λ, color coded respectively, for uncorrected images (top) and beam pattern corrected images (bottom). Values less than 0.1 have been ignored. Dotted lines are mean values. Note how well the triplets correlate with each other, suggesting that the variation in the estimate originates from brightness variations between images.

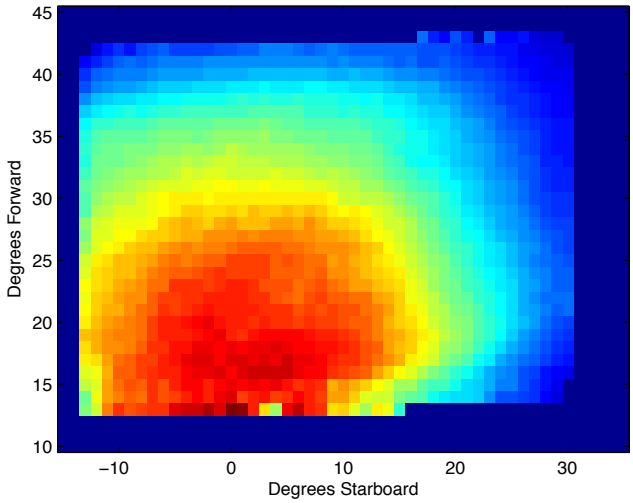

Fig. 9. Estimated beam pattern of the strobe in angular space. Warmer hues indicate higher intensities, while the dark blue border is outside the camera field of view. Axes units are in degrees, with (0,0) corresponding to the nadir of the strobe. The strobe was mounted facing forward with a downward angle of approximately 70 degrees from horizontal. The camera was mounted forward and starboard of the strobe. Note how the beam pattern is readily visible in figure 10.

present to ensure reliable keypoint detection, which is usually not an issue for AUV missions. For towed systems, however, altitude and speed are more difficult to control, so we propose a second simpler method for estimating the attenuation coefficients. Figure 6 plotted the log color channel means over a range of altitudes. For altitudes between 1-5 meters, the relationship is roughly linear, which we used to justify ignoring any additive scattering in our model. Assuming that the average path length $\bar{\ell} \approx 2a$ is approximately twice the altitude, then the attenuation coefficients can also be estimated as half the slope of this plot.

C. Beam Pattern Estimation

While the strobe's beam pattern induces non-uniform illumination patterns across images, this beam pattern will remain constant within the angular space of the strobe. Assuming a planar bottom, each image represents a slice through that space, and we are able to parameterize each pixel in terms of the beam angles θ and ϕ. If we consider only the pixels $\boldsymbol{p} \in [\theta_i, \phi_j]$ that fall within an angular bin, the average intensity value corrected for attenuation will be a relative estimate of the beam pattern in that direction.

$$\log BP(\theta_i, \phi_j) = \sum_{\Lambda} \frac{1}{|\boldsymbol{p}|} \sum_{\boldsymbol{p}} \log \boldsymbol{c}_\Lambda + \alpha_\Lambda \boldsymbol{\ell} \qquad (17)$$

Assuming that there is no spatial bias in image intensity (for instance the left half of the images always contain sand and the right half of the images only contain rocks) then the reflectance term only contributes a uniform gain. This gain is removed when the beam pattern is normalized over angular space. The resulting beam pattern is shown in Figure 9.

We also recompute the attenuation coefficients using color values for the beam pattern corrected imagery. The results are shown in the bottom of Figure 8. The triplets correlate quite well with each other, suggesting that variation in the estimates arises from intensity variation between images and not necessarily within images.

D. Image Correction

Each captured image can be corrected by multiplication with the inverse of the beam pattern and attenuation terms. Figure 10 shows different stages of correction performed on the upper left raw underwater image. The upper right image has been corrected for attenuation alone, and while its colors look more realistic there are strong illumination artifacts present. The lower left image has been corrected for beam pattern alone, and thus maintains a bluish hue from attenuation. The bottom right image has been corrected for both illumination and attenuation.

The artifacts of attenuation and illumination are sometimes hidden when photomosaics are created and images blurred together. Figure 11 shows photomosaics of the same area before and after correction. While much of the along-track variation in illumination has been blurred away, there is still a definitive difference in brightness in the across-track direction. This can create difficulties when attempting to mosaic adjacent track lines together.

Several more pairs of raw and corrected images are shown in Figures 12 and 13 and have been published in [40]. These images were captured at various benthic locations between the Marguerite Bay slope off the western Antarctic Peninsula and the Amundsen Sea polynia. While the same beam pattern estimates are used for all images, the value of the attenuation coefficients varied enough between locations that using mismatched coefficients produced unrealistic looking results. While correlating attenuation with parameters such as salinity or biological activity is beyond the scope of this thesis, it presents interesting topics for future research into measuring environmental variables using imagery.

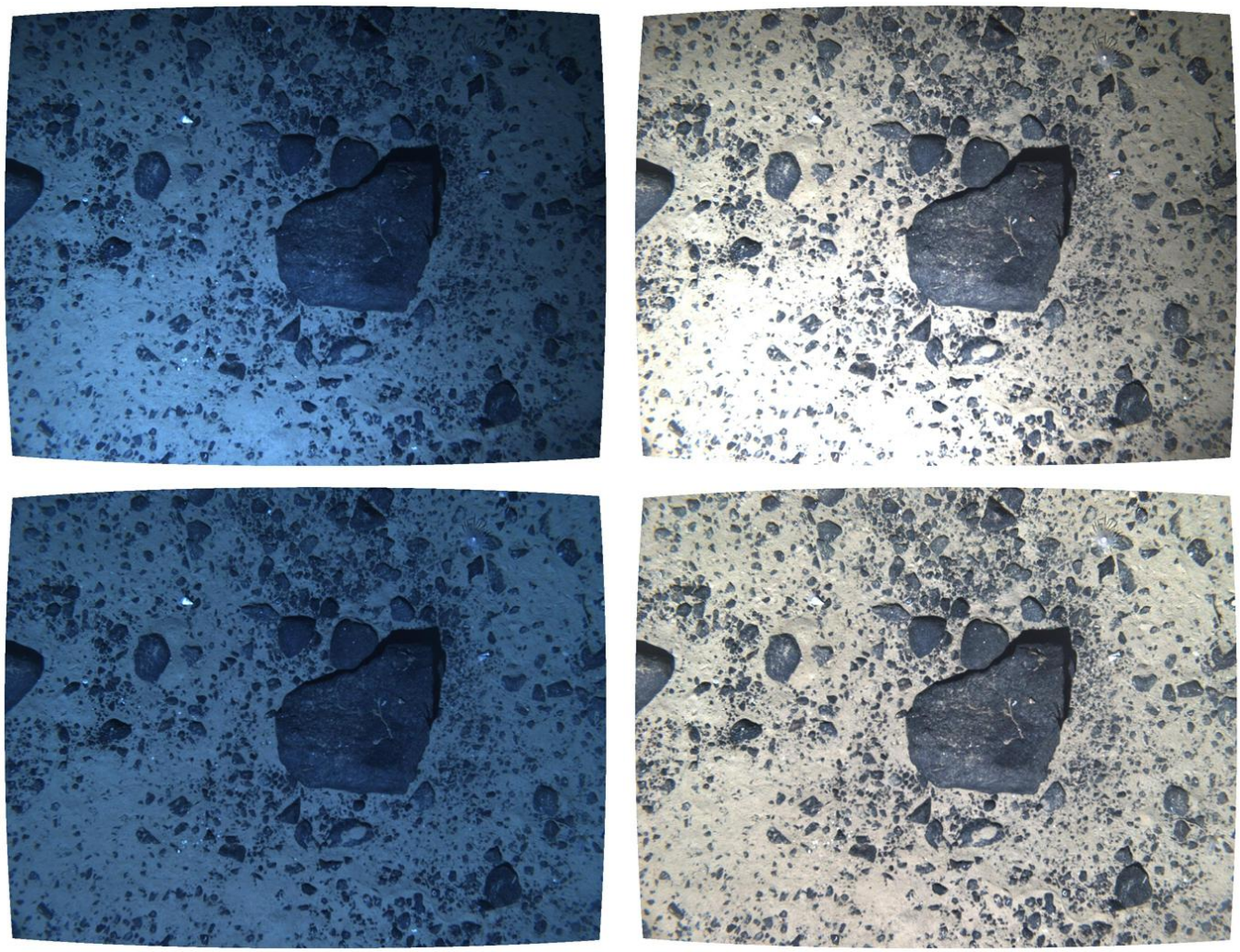

Fig. 10. Sample underwater image (upper left) corrected for attenuation alone (top right), beam pattern alone (bottom left), and both beam pattern and attenuation using the method presented in this paper.

V. Conclusions and Future Work

We have presented a detailed model of image formation underwater, discussed a diverse assortment of methods used to obtain and correct high-quality underwater images, and presented our own method for correcting underwater images captured from a broad family of robotic platforms. For this method, we use additional sensor information commonly available on underwater vehicles to constrain a strong physical model. Because an underwater image captures both water column and illumination processes in addition to the scene of interest, we are able to isolate and estimate the attenuation coefficient of the water as well as the beam pattern of the illumination source from the images themselves. Our goal was never to develop a unifying method of correction, but rather to emphasize a unified understanding in how correction methods should be applied in different imaging situations.

Because robotic platforms generate datasets that are often too large for exhaustive human analysis, the emphasis on their correction should involve batch methods that provide consistent, if not perfectly accurate, representations of color and texture. Available information from other onboard sensors can and should be utilized to improve results. For instance, [33] uses a towed camera system's sonar altimeter for range-dependent frame averaging, while [35] uses stereo camera pairs to estimate photon path lengths at the individual pixel scale. While an unfortunate side effect of such a mentality is that correction techniques can become somewhat platform-specific, it is not surprising that more realistic correction is obtained when all available information is taken into account.

Furthermore, we re-emphasize that a corrected image alongside a raw image contains information regarding the water column properties, the bottom topography, and the illumination source. Given this residual, any correction scheme will naturally provide insight to some projection of these values in parameter space. In the absence of ground truthing, which is often unrealistic to obtain during real-world mission scenarios, one possible metric is to compare methods based on their residual, or what they estimate the artifacts to be, for which approach most closely approximates the physical imaging situation. While this metric is somewhat contrived, it is apparent that the approach which best estimates this residual will also provide superior results.

Fig. 11. Raw (left) and corrected (right) photomosaics from a sequence of 10 images. Note how the non-uniform illumination pattern is blurred between frames in the left image.

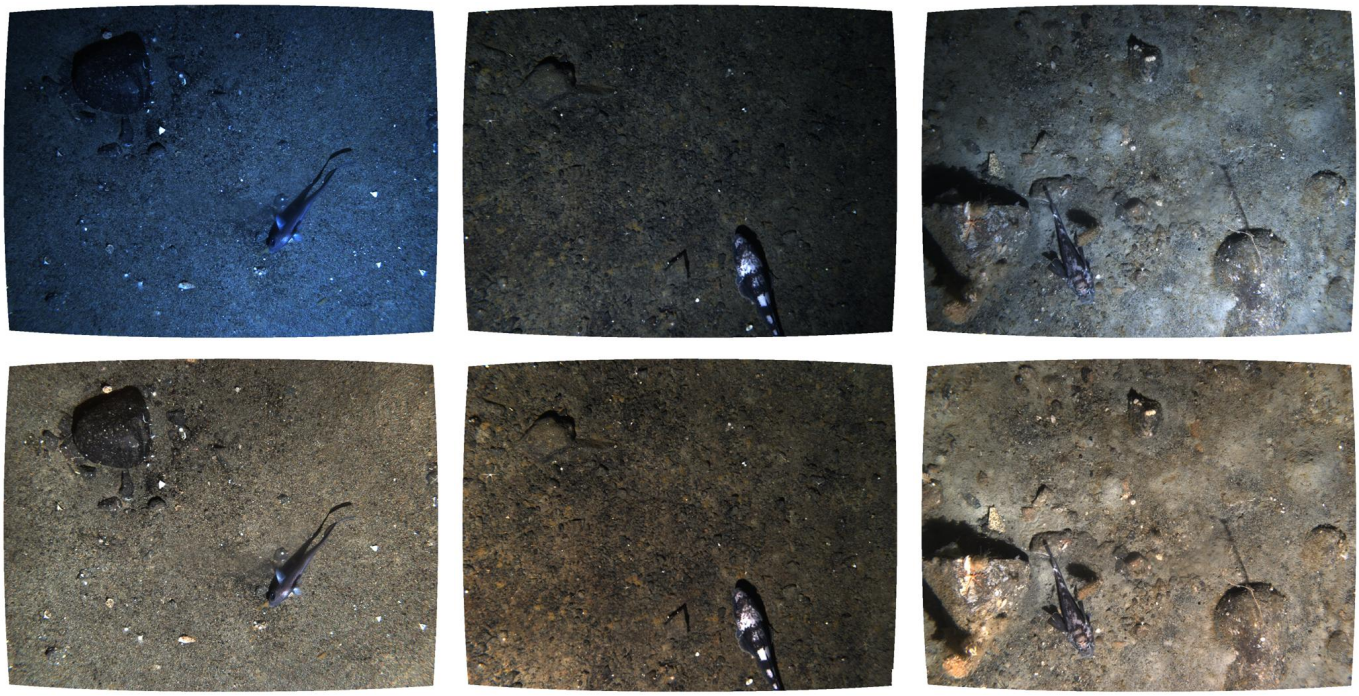

Fig. 12. Example raw (top) and corrected (bottom) images using the method presented in this paper.

REFERENCES

[1] W. H. Smith, "Introduction to this special issue on bathymetry from space," *Oceanography*, vol. 17, no. 1, pp. 6–7, 2004.

[2] S. Q. Duntley, "Light in the sea," *Journal of the Optical Society of America*, vol. 53, pp. 214–233, 1963.

[3] C. D. Mobley, "The optical properties of water," *Handbook of optics*, vol. 2, 1995.

[4] B. L. McGlamery, "Computer analysis and simulation of underwater camera system performance," *SIO ref*, vol. 75, p. 2, 1975.

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

[6] R. Fattal, "Single image dehazing," in *ACM SIGGRAPH 2008 papers*, ser. SIGGRAPH '08. New York, NY, USA: ACM, 2008, pp. 72:1–72:9. [Online]. Available: http://doi.acm.org/10.1145/1399504.1360671

[7] J. S. Jaffe, "Computer modeling and the design of optimal underwater imaging systems," *Oceanic Engineering, IEEE Journal of*, vol. 15, no. 2, pp. 101–111, 1990.

[8] J. Heikkila and O. Silven, "A four-step camera calibration procedure with implicit image correction," in *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, jun 1997, pp. 1106 –1112.

[9] X. Li, B. Gunturk, and L. Zhang, "Image demosaicing: A systematic survey," in *Proc. SPIE*, vol. 6822, 2008, p. 68221J.

[10] H. Malvar, L. wei He, and R. Cutler, "High-quality linear interpolation for demosaicing of bayer-patterned color images," in *Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04). IEEE International Conference on*, vol. 3, may 2004, pp. iii – 485–8 vol.3.

[11] B. Jähne, *Image processing for scientific applications*. CRC press Boca Raton, 1997.

[12] H. Singh, J. Howland, and O. Pizarro, "Advances in large-area photomosaicking underwater," *Oceanic Engineering, IEEE Journal of*, vol. 29, no. 3, pp. 872 – 886, july 2004.

[13] M. Edge, *The underwater photographer*. Focal Press, 2012.

[14] N. Carlevaris-Bianco, A. Mohan, and R. Eustice, "Initial results in underwater single image dehazing," in *proc. OCEANS*, sept. 2010.

[15] Y. Schechner and N. Karpel, "Clear underwater vision," in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, vol. 1, june-2 july 2004, pp. I–536 – I–543 Vol.1.

[16] R. Garcia, T. Nicosevici, and X. Cufi, "On the way to solve lighting problems in underwater imaging," in *proc. OCEANS*, vol. 2, oct. 2002.

[17] H. Singh, C. Roman, O. Pizarro, R. Eustice, and A. Can, "Towards high-resolution imaging from underwater vehicles," *The International Journal of Robotics Research*, vol. 26, no. 1, pp. 55 – 74, january 2007.

[18] W. Hou, D. J. Gray, A. D. Weidemann, G. R. Fournier, and J. Forand, "Automated underwater image restoration and retrieval of related optical properties," in *Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International*. IEEE, 2007, pp. 1889–1892.

[19] G. Gorman, "Field deployable dynamic lighting system for turbid water imaging," Master's thesis, MIT/WHOI Joint Program in Oceanography / Applied Oceans Science and Engineering, 2011.

[20] M. Gupta, S. G. Narasimhan, and Y. Y. Schechner, "On controlling light transport in poor visibility environments," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. IEEE, 2008, pp. 1–8.

[21] J. Jaffe, "Enhanced extended range underwater imaging via structured illumination," *Optics Express*, vol. 18, pp. 12 328–12 340, 2010.

[22] S. G. Narasimhan and S. K. Nayar, "Structured light methods for underwater imaging: light stripe scanning and photometric stereo," in *OCEANS, 2005. Proceedings of MTS/IEEE*. IEEE, 2005, pp. 2610–2617.

[23] J. S. Jaffe, K. D. Moore, J. McLean, and M. Strand, "Underwater optical imaging: status and prospects," *Oceanography*, vol. 14, no. 3, pp. 66–76, 2001.

[24] M. Levoy and H. Singh, "Improving underwater vision using confocal imaging," Stanford University, Tech. Rep., 2009.

[25] M. Chambah, D. Semani, A. Renouf, P. Courtellemont, and A. Rizzi, "Underwater color constancy: enhancement of automatic live fish recognition," in *Electronic Imaging 2004*. International Society for Optics and Photonics, 2003, pp. 157–168.

[26] C. Murphy, personal communication, 2011.

[27] E. Hsu, T. Mertens, S. Paris, S. Avidan, and F. Durand, "Light mixture estimation for spatially varying white balance," in *ACM Transactions on Graphics (TOG)*, vol. 27, no. 3. ACM, 2008, p. 70.

[28] C. Ancuti, C. O. Ancuti, T. Haber, and P. Bekaert, "Enhancing underwater images and videos by fusion," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE, 2012, pp. 81–88.

[29] L. Torres-Mendez and G. Dudek, "Color correction of underwater images for aquatic robot inspection," in *Energy Minimization Methods in Computer Vision and Pattern Recognition*, ser. Lecture Notes in Computer Science, A. Rangarajan, B. Vemuri, and A. Yuille, Eds. Springer Berlin / Heidelberg, 2005, vol. 3757, pp. 60–73.

[30] I. Vasilescu, C. Detweiler, and D. Rus, "Color-accurate underwater

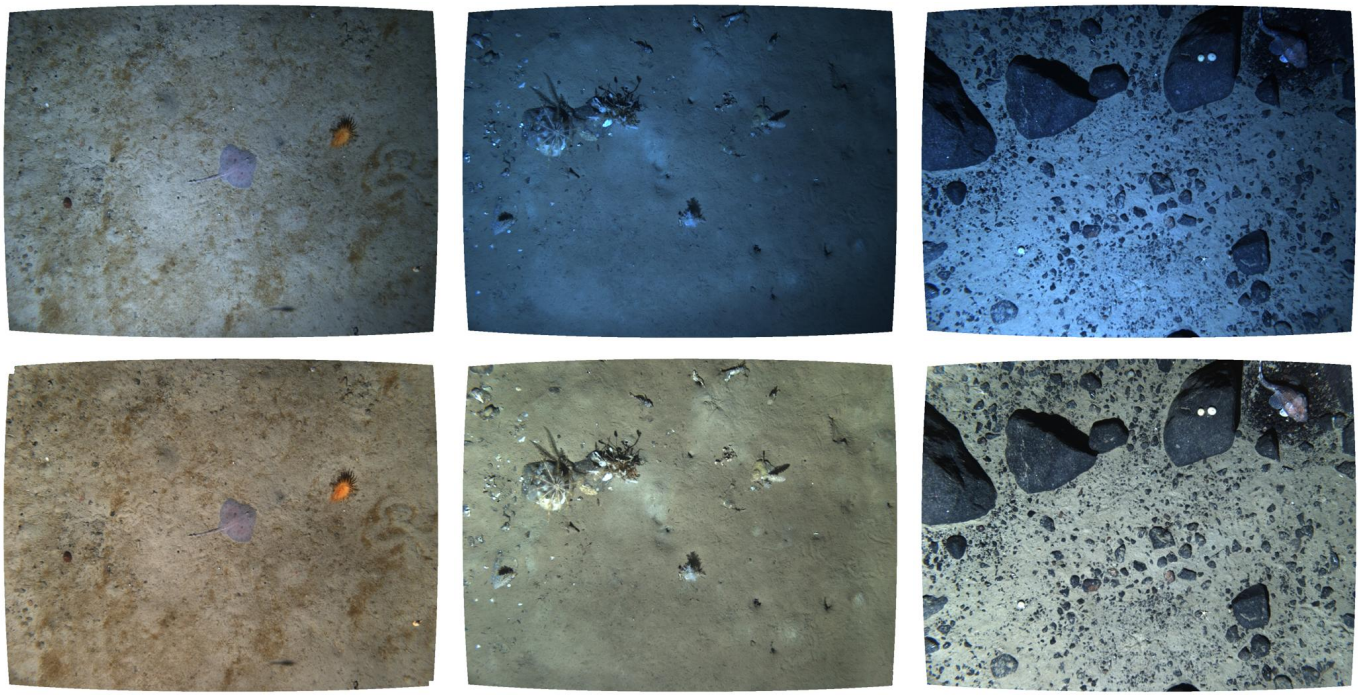

Fig. 13. Example raw (top) and corrected (bottom) images using the method presented in this paper.

imaging using perceptual adaptive illumination," *Autonomous Robots*, vol. 31, no. 2-3, pp. 285–296, 2011.

[31] M. Johnson-Roberson, O. Pizarro, S. B. Williams, and I. Mahon, "Generation and visualization of large-scale three-dimensional reconstructions from underwater robotic surveys," *Journal of Field Robotics*, vol. 27, no. 1, pp. 21–51, 2010. [Online]. Available: http://dx.doi.org/10.1002/rob.20324

[32] O. Pizarro and H. Singh, "Toward large-area mosaicing for underwater scientific applications," *Oceanic Engineering, IEEE Journal of*, vol. 28, no. 4, pp. 651 – 672, oct. 2003.

[33] J. Rock, P. Honig, C. Stewart, S. Gallager, and A. York, "Illumination correction for habcam imagery," unpublished.

[34] J. Kaeli, H. Singh, C. Murphy, and C. Kunz, "Improving color correction for underwater image surveys," in *proc. OCEANS*, 2011.

[35] M. Bryson, M. Johnson-Roberson, O. Pizarro, and S. Williams, "Colour-consistent structure-from-motion models using underwater imagery," in *Proceedings of the 2012 Robotics: Science and Systems Conference*, 2012, p. 8.

[36] A. Bodenmann, B. Thornton, T. Nakatani, and T. Ura, "3d colour reconstruction of a hydrothermally active area using an underwater robot," in *OCEANS 2011*. IEEE, 2011, pp. 1–6.

[37] RD Instruments. http://www.rdinstruments.com.

[38] G. Foresti and S. Gentili, "A vision based sytem for object detection in underwater images," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 14, pp. 167–188, 2000.

[39] T. Nicosevici, N. Gracias, S. Negahdaripour, and R. Garcia, "Efficient three-dimensional scene modeling and mosaicing," *Journal of Field Robotics*, vol. 26, no. 10, pp. 759–788, 2009. [Online]. Available: http://dx.doi.org/10.1002/rob.20305

[40] J. T. Eastman, M. O. Amsler, R. B. Aronson, S. Thatje, J. B. McClintock, S. C. Vos, J. W. Kaeli, H. Singh, and M. La Mesa, "Photographic survey of benthos provides insights into the antarctic fish fauna from the marguerite bay slope and the amundsen sea," *Antarctic Science*, vol. 1, no. 1, pp. 1–13, 2012.