Improving Color Correction for Underwater Image Surveys

Jeffrey W. Kaeli, Hanumant Singh, Chris Murphy, Clay Kunz
Deep Submergence Lab
Woods Hole Oceanographic Institution
Woods Hole, MA

***Abstract*—We propose and implement a novel method of estimating attenuation coefficients and a strobe beam pattern using sequences of overlapping underwater color images and acoustic ranges from a Dopper Velocity Log (DVL). These values are used to correct the images for color and illumination artifacts with the goal of more consistent color correction for input into classification algorithms.**

Keywords-underwater imaging; color correction; multisensor fusion.

I. INTRODUCTION

Optical imaging is an important mode of underwater sensing for many scientific, commercial, and military applications. This task often demands high resolution imagery of the seafloor with accurate color representation for benthic characterization. However, the physical range at which cameras can operate underwater is severely limited by high attenuation from scattering and absorption which varies by location [9]. Furthermore, since natural light does not effectively penetrate past the upper few hundred meters of water column, deep-diving imaging platforms must supply their own sources of illumination, and those operating untethered from ship resources are limited in the power they can provide for illumination. Thus, a broad class of underwater images suffers from uneven altitude-dependent lighting patterns that must be corrected for both in intensity as well as in color. An example image is shown in Fig. 1.

Robotic imaging platforms are capable of providing large datasets that exceed the practical limits of exhaustive human analysis. This has created a strong motivation for automatic processing algorithms capable of segmenting and classifying regions and objects in underwater imagery. However, such algorithms rely on accurate, and moreover consistent, contrast and color information for texture and feature extraction. [1].

Recent work in single image dehazing has shown impressive results for reducing the scattering component of shallow naturally-lit horizontal underwater imagery and obtaining a metric for range [2]. However, this method assumes the color of the haze is constant so no color correction is performed. Markov Random Fields have been used with statistical priors learnt from training images to restore color in underwater images [3]. Frame averaging is a simple and useful approach, but only applicable in flat, relatively uninteresting

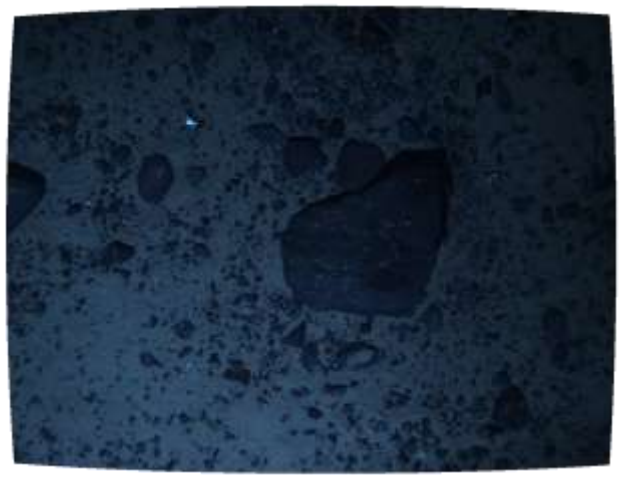

Figure 1. Example raw image after rectification for lens distortion.

areas [4]. Adaptive histogram equalization performs spatially varying histogram equalization over image subregions, and has been successfully applied to grayscale underwater images [5]. Applied in different color spaces, it can correct for some illumination, but haloing occurs around sharp intensity changes and processing color channels separately can lead to misrepresentations of the actual color. Homomorphic methods assume a raw image is the product of an illumination and a reflectance image, and that the illumination component varies spatially more slowly than the local contrast in the reflectance component [6]. This model becomes linear in the logarithmic domain, and the illumination component can be estimated through low-pass filtering or parametric surface fitting. White balancing can achieve simple useful results through adjusting relative color channel gains, but no correction is made for illumination patterns [7]. Example results of these methods are illustrated in Fig. 2.

Each of these techniques can achieve aesthetically pleasing results for a single image, but the color and texture consistency of these corrected images with respect to their input into a batch dataset-level classifier is questionable. These techniques also focus on the images themselves without taking advantage of the unique lighting constraints and additional sensors often present on underwater imaging platforms. The Doppler Velocity Log (DVL) is a ubiquitous oceanographic sensor used extensively for vehicle navigation [11]. It emits 4 beams of

sound, measuring the phase shift in the returns from the seafloor to determine both relative velocity and range to the bottom. We present a method to utilize the DVL ranges in conjunction with measured sensor offsets to estimate real physical parameters, such as attenuation coefficients and the beam pattern of the strobe, from the images themselves, and subsequently correct for color and illumination artifacts in the images.

II. METHODS

A. *Rectifying lens distortions*

Because the refractive index of water differs from that of air, we calibrate our camera lenses in water using the method described in [8] and rectify the images before subsequent processing.

B. *Color Attenuation Estimation*

We start with an illumination model that assumes the raw captured image is the product of some reflectance image we would like to isolate, and some illumination component that we would like to eliminate,

$$\boldsymbol{RAW}(p,\lambda,i) = \boldsymbol{REF}(p,\lambda,i)\cdot\boldsymbol{ILL}(p,\lambda,i) \quad (1)$$

where $p=(m,n)$ are pixel indices within the image, λ is the color channel, and i denotes the frame number. Many deep underwater imaging platforms use one or several strobes for illumination, so we assume a point light source with spectral content $s(\lambda)$, beam pattern $\boldsymbol{BP}(\phi,\psi)$, and photon path length $d(z)$ from source to object to camera, where z is the seafloor topography. The illumination component at each pixel can then be decomposed into

$$\boldsymbol{ILL}(p,\lambda,i) = K(i)\, s(\lambda)\, \boldsymbol{BP}(\phi,\psi)\, e^{-\boldsymbol{\alpha}(\lambda)d(z)}\, d(z)^{-q} \quad (2)$$

where $K(i)$ is a camera gain term and q is a power spreading term. The total attenuation coefficient $\boldsymbol{\alpha}(\lambda)=\{\alpha,\gamma,\beta\}$ (for the red, green, and blue channels, respectively) is comprised of the sum of the scattering $\boldsymbol{\alpha_s}$ and wavelength-dependent absorbsion $\boldsymbol{\alpha_a}(\lambda)$ coefficients [9].

We assume that the camera gain does not change between images, the spectral content of the strobe is white, and the beam pattern of the strobe is constant over the image (this last assumption is weak and explored further in the next section). Furthermore, we assume that there is overlap between adjacent images, i.e. two images share pixels whose actual color values are identical. We can then take the log of (1) and write, for a pair of images,

$$\log \boldsymbol{RAW}(p,\lambda,i) = \log \boldsymbol{REF}(p,\lambda) - d(z)\, \boldsymbol{\alpha}(\lambda) - \log d(z)\, q \quad (3)$$

Note that the reflectance image is no longer a function of the frame number i. If we assume that the seafloor is locally planar within the field of view, we can fit a plane to the DVL ranges and, using measured offsets between the strobe, camera, and DVL, obtain a measure of $d(z)$. The equation is now linear

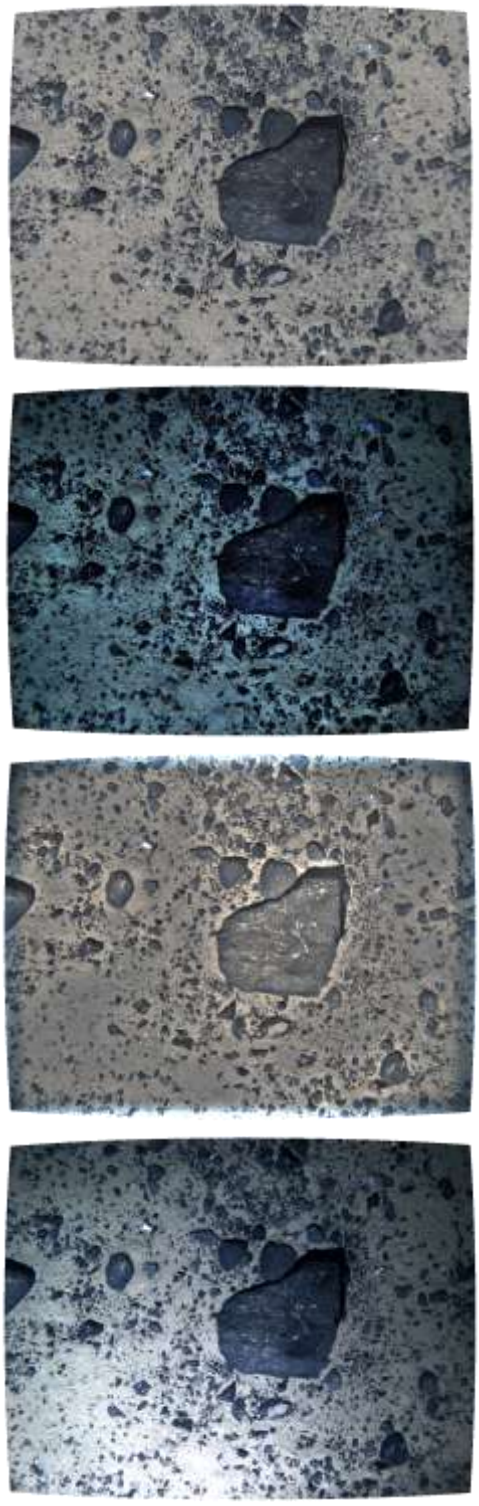

Figure 2. Example methods of correction. (from top to bottom) Frame averaging suffers when imagery is collected over sloping terrain or from varying altitudes. Adaptive histogram equalization, here performed in L*a*b color space, suffers from haloing and misrepresentation of color. Homomorphic filtering, here using a Gaussian low-pass filter, suffers severe haloing artifacts around high contrast boundaries. White balancing handles colors intuitively, but does not account for the illumination. Note how each image itself is sufficient for human visual inspection, but there is significant variation in color between each method.

with respect to unknowns log $\boldsymbol{REF}(p,\lambda)$, $\boldsymbol{\alpha}(\lambda)$, and q, which we can solve for with a constrained linear least-squares regression over multiple pixels. Using average values of $\boldsymbol{\alpha}(\lambda)$ and q over a set of images, the reflectance images can be easily recovered.

C. Beam Pattern Estimation

Contrary to our assumption in the previous section, the strobe beam pattern $\boldsymbol{BP}(\phi,\psi)$ is rarely uniform across the image. It does, however, remain constant in angular space, with each captured image approximating a 2-D slice through that space. We can use our approximation of $d(z)$ from the DVL ranges to warp each pixel into the angular space of the strobe. Noting that the recovered reflectance image in the previous section is actually the product of the beam pattern and the true reflectance image, we average a series of images in angular space to obtain an estimate of the beam pattern of the strobe. Because the strobe beam pattern is independent of wavelength, we can correct the raw images for beam pattern alone and then recalculate $\boldsymbol{\alpha}(\lambda)$. The true reflectance image can be recovered through

$$\boldsymbol{REF}(p,\lambda,i) = \boldsymbol{RAW}(p,\lambda,i)\, e^{\boldsymbol{\alpha}(\lambda)d(z)}\, d(z)^{q} / \mathrm{norm}(\, \boldsymbol{BP}(\phi,\psi)\,) \quad (4)$$

where norm($\boldsymbol{BP}(\phi,\psi)$) denotes normalization of the beam pattern in angular space.

D. Correlating Underwater Imagery

Our method of estimating the attenuation coefficients $\boldsymbol{\alpha}(\lambda)$ is dependent upon our ability to correlate two overlapping images. Correlating underwater imagery has been studied for the purpose of creating photomosaics of the seafloor, utilizing Fourier-based methods [5] and feature-based methods such as SIFT [10] and Zernike moments [4]. Because our goal is to find distinct pixel matches between images, we opt for feature-based methods, using a local pixel average around each feature as an input to our regression. Correlating features between images is done using the RANSAC algorithm. We have experimented with both SIFT and Zernike moments; the results presented in this paper use Zernike features. An example is shown in Fig. 3.

III. Results

Images were collected using the SeaSLED towed camera system during a cruise aboard the icebreakers N.B. Palmer and Oden, surveying depths ranging from 300-2600m over the continental shelf and slope of Antarctica. Relevant onboard sensors included a pair of Prosilica 1024x1360 CCD cameras with 12 bits of dynamic range, a custom strobe, and a 1.2 MHz RD Instruments DVL.

A sequence of 100 sample images was rectified for lens distortion, features calculated, and then correspondences computed between adjacent images. If the vehicle altitude changed significantly between frames, it became necessary to adjust the intensity of a pair of images so that features would be recognized similarly in each. Once correspondences were determined, the average color within a 3x3 neighborhood surrounding each pair of features was used as input to the regression in (3) to obtain values for log $\boldsymbol{REF}(p,\lambda)$, $\boldsymbol{\alpha}(\lambda)$, and q.

This work was funded by the NSF Center for Subsurface Sensing and Imaging Systems (CenSISS) Engineering Research Center (ENC) grant no. EEC-99868321.

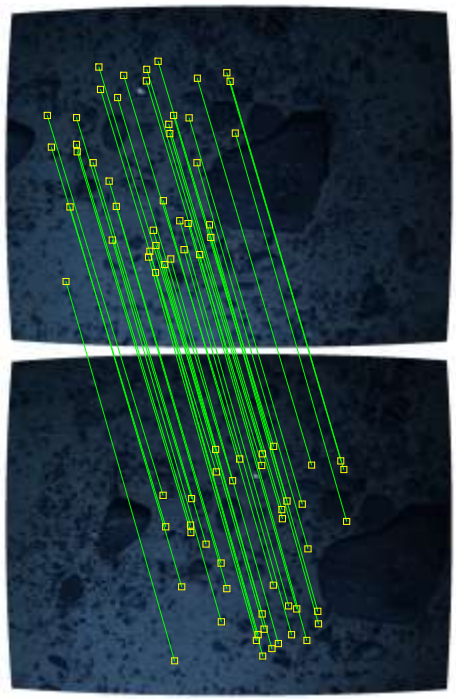

Figure 3. A pair of images with corresponding Zernike features drawn.

We found q to consistently be approximately zero, so we ignored the spreading term in our subsequent calculations. Estimated values of $\boldsymbol{\alpha}(\lambda)$ are shown at the top of Fig. 4. Approximately 20% of the values were less than 0.1, considered outliers, and ignored.

Images were then corrected for attenuation using the mean values of $\boldsymbol{\alpha}(\lambda)$, Fig. 6 at top. These images were then averaged in angular space in 1° bins to estimate the strobe beam pattern, shown in Fig. 5. The beam pattern was normalized in angular space, interpolated over the angular domain of each image, and used to correct the raw images for illumination alone, Fig. 6 middle. From these images, the values of $\boldsymbol{\alpha}(\lambda)$ were recalculated, Fig. 4 bottom. Finally, the raw images were corrected for both attenuation and illumination, Fig. 6 bottom. Photomosaics comparing raw and fully corrected imagery are shown in Fig. 7.

IV. Discussion

We propose and implement a novel method of estimating attenuation coefficients and a strobe beam pattern using sequences of overlapping underwater color images and ranges from a DVL mounted in a known configuration with the strobe and camera. These values are used to correct the images for color and illumination artifacts. An overarching assumption we make is that the seafloor is locally planar, and the 4 ranges returned from the DVL adequately capture that plane. Our approach is advantageous over simple frame averaging in that

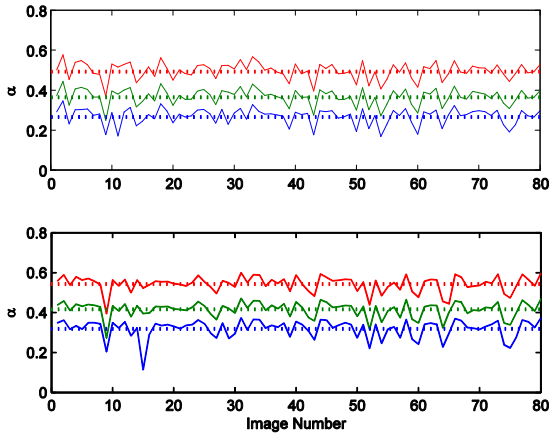

Figure 4. Estimated alpha, color coded appropriately, for raw images (top) and beam pattern corrected images (bottom). Values less than 0.1 have been ignored. Dotted lines are mean values. Note how well the alpha triplets correlate with each other, suggesting that the variation in our estimates originates from intensity variation between images.

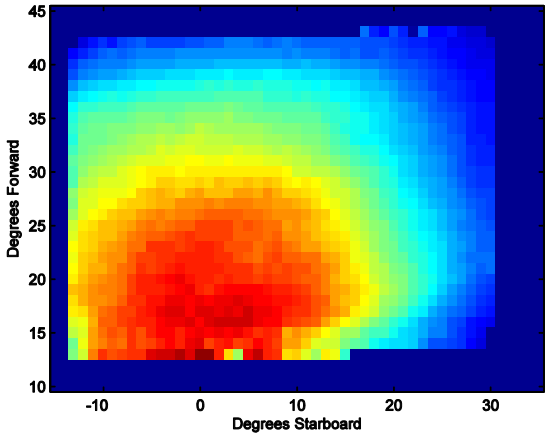

Figure 5. Estimated beam pattern of the strobe in angular space. Warmer hues indicate higher intensities; the dark blue border indicates angles outside the camera field of view. Axes units are in degrees, with (0,0) corresponding to the nadir of the strobe. The strobe was mounted facing forward with a downward angle of about 70 degrees; the camera was mounted forward and starboard of the strobe. Note how this beam pattern is readily visible in Fig. 6 top.

the seafloor need not be flat (i.e. it can be sloped) and that the imaging platform can undergo changes in pitch and roll, useful for towed sleds with limited control capability. Color and intensity errors occur when the planar assumption is violated, such as a large boulder, or when one of the DVL beams returns off of a similar object. This latter situation could be avoided by incorporating navigational information and creating a very low-resolution map of the seafloor topography. Future implementations of this method could use stereo camera pairs, structure from motion (SFM) algorithms, or multibeam sonar to obtain higher-resolution estimates of the bottom elevation.

Another overarching assumption we make is that there is sufficient overlap between images to compute feature correspondences. In the sequence of images shown here there

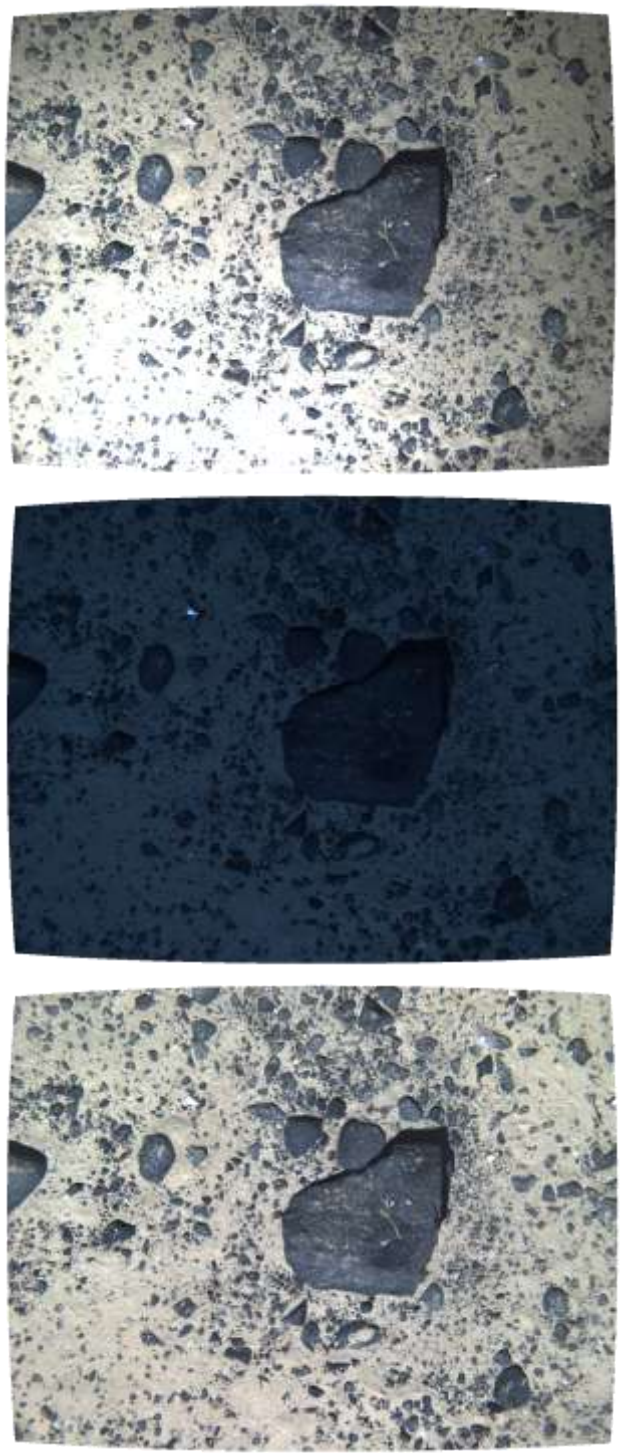

Figure 6. Correction for attenuation alone (top), beam pattern alone (middle), and both (bottom).

was consistently >50% overlap due to the slow towing speed, and the bottom was relatively full of features. Even on unfeatured seafloor with less overlap, our approach for calculating $\boldsymbol{\alpha}(\lambda)$ only needs a single common pixel between a

pair of images, and for a featureless bottom the average color over a larger region can be used. A potential problem with using features for calculating $\boldsymbol{\alpha}(\lambda)$ is that they may have more relief and are generally different intensities from the background, which could bias the estimates. Incorporating the average value of low-gradient areas between features could improve our estimates of $\boldsymbol{\alpha}(\lambda)$ in future work. This bias could also explain some of the bad estimates that were rejected as outliers, as the outliers were the same in both calculations of $\boldsymbol{\alpha}(\lambda)$. The outliers could also be due to features that violate the planar bottom assumption in those particular images.

We intend to implement this method across other imagery captured around the world to compare attenuation coefficients. We also recognize the need to capture images with objects of known color so that we can ground truth our results. This work will hopefully lead towards providing more consistent color correction for input into classification algorithms.

ACKNOWLEDGMENT

We would like to thank the crew and science parties aboard the N.B. Palmer and Oden, as well as other members of the Deep Submergence Lab who helped make this work possible.

REFERENCES

[1] J.W. Kaeli, H. Singh, and R.A. Armstrong, "An Automated Morphological Image Processing Based Methodology for Quantifying Coral Cover in Deeper-Reef Zones," OCEANS 2006.

[2] N. Carlevaris-Bianco, A. Mohan, and R.M. Eustice, "Initial Results in Underwater Single Image Dehazing," OCEANS, 2010.

[3] L.A. Torres-Mendez and G. Dudek, "Color Correction of Underwater Images for Aquatic Robot Inspection," Com. Vis. Patt. Recog., 2010.

[4] O. Pizarro, "Toward Large-Area Mosaicing for Underwater Scientific Applications," IEEE J. Ocean. Eng., vol. 28, no. 4, 2003.

[5] H. Singh, J Howland, and O. Pizarro, "Advances in Large-Area Photomosaicking Underwater," IEEE J. Ocean. Eng., vol. 29, no. 3, 2004.

[6] H. Singh, C. Roman, O. Pizarro, R. Eustice, and A. Can, "Towards High-resolution Imaging from Underwater Vehicles," Int. J. Rob. Res., vol. 26, no. 1, 2007.

[7] C. Murphy, personal communication.

[8] J. Heikkila and O. Silven, "A Four-step Camera Calibration Procedure with Implicit Image Correction," proc. Comp. Vis. Patt. Rec., 1997.

[9] S.Q. Duntley, "Light in the Sea," J. Opt. Soc. Am., Vol. 53, Iss. 2, 1963.

[10] T. Nicosevici, N. Gracias, S. Nagahdaripour, and R. Garcia, "Efficient Three-Dimensional Scene Modeling and Mosaicing," J. Field Rob., vol. 26, no. 10, 2009.

[11] http://www.rdinstruments.com

Figure 7. Raw (left) and corrected (right) photomosaics of a sequence of 10 images.