Ray Tracing for Underwater Illumination Estimation

Jason Rock

May 16, 2011

1 Introduction

Oceanographic research often involves taking high resolution underwater images which can be analyzed in many ways including counting specific species such as scallops, identifying the spread of invasive species like didendum, and determining the effects of environment changes on ecosystems. In order for large scale deployment of imaging systems to be feasible, automatic techniques are required to assist in data analysis. Due to the optical properties of water, light is significantly attenuated due to both absorption and scattering in a manner dissimilar from air. Absorption affects light as a function of wavelength, attenuating the low more than high visible wavelengths. Also, since research vehicles are often operated at depths well below the euphotic depth, which by definition, is the depth at which less than 1% of sunlight penetrates, they must provide their own lights. These strobes cast a nonuniform illumination pattern on the seafloor which varies with height and orientation of the craft. Computer vision algorithm are often sensitive to uneven illumination so correcting the illumination is a very important preprocessing step.

This paper approaches illumination correction from a modeling perspective. We attempt to accurately model the camera, seafloor, lights and absorption of water in full spectrum for a specific sled vehicle called the HABCAM. The hope is that the HabCam can be accurately modeled through the use of ray tracing without photon mapping. Our primary global assumptions are (1) scattering effects of the water column are uniform, and can therefore be approximated alongside our absorption parameter, (2) the ocean floor is roughly planar, (3) the reflectance properties of the ocean floor can be modeled by a simplified Phong model.

1.1 HabCam

The HABitat mapping CAMera (HabCam) system is a towed vehicle. It flies within six meters above the ocean floor collecting images at 6 Hz. The camera is a 12 bit machine vision camera.[3] There are four strobes placed radially about the camera with a 1 meter diameter, the strobes contain xenon flash lamps often used in machine vision applications due to their production of full spectrum light.[4, 7] Altitude readings are taken with a benthos altimeter.

1.2 Previous Illumination Correction Work

Past illumination correction algorithms have been entirely empirical in nature, in general attempting to discover trends in or across images. Garcia et al. provide a nice overview and analysis.[5] The illumination-reflectance model, uses the fact that an image f is a function of gain g, illumination i, reflectance r and offset o such that $f(x,y) = g(x,y)i(x,y)r(x,y) + o(x,y)$.[6] Since the offset is only a small contribution, it is eliminated. The gain and illumination are both low order and can be combined to a single term $c(x,y)$. This leaves us with $f(x,y) \approx c(x,y)r(x,y)$. A convolution operation with a large Gaussian can eliminate the high frequency portion of the image, leaving us with $c(x,y)$ as long as the reflectance can be modeled as noise.[5]

Histogram equalization is often used in low contrast images, and it can be tweaked for underwater images.[12] Rather than considering and equalizing the whole image, local histograms are equalized.[10] This method works as long as the image is relatively uniform for the window size. However, it is undesirable because selecting a correct region size might need to be done per image. It also amplifies noise in low contrast areas which often make up the majority of the uninteresting images of the seafloor.[5]

Homomorphic filtering is also proposed as a possibility since illumination makes up much of the low frequency values in a Fourier transform.[9] It is formulated by returning to the illumination-reflectance model, and taking the log of the image to allow for the separation of reflectance and illumination leaving us with, $\log(f) = \log(c)+\log(r)$. Applying a Fourier transform to the log image, multiplying by a highpass filter, and then transforming back to image space eliminates the low frequency addition due to the illumination. However, it is obvious that any low frequency data in the reflectance will also be eliminated.[5]

More recently, Singh et al. proposed an improvement to homomorphic filtering, decompose an image F into illumination I and reflectance R such that $F(x, y, \lambda) = I(x, y, \lambda)R(x, y, \lambda)$. Assuming that reflectance is relatively high order, and illumination relatively low, fit a low order polynomial to the logarithm of the image which will effectively fit only the illumination. They achieve better results using this method than using a highpass filter, and their algorithm is faster. Unfortunately, they still rely on the assumption that reflectance is effectively high order noise, and that assumption is often violated.[13]

Leery et al. propose another improvement specifically for downward facing cameras which considers altitude as a parameter of illumination. Their basic algorithm is to sum images at height bins, and then compute a low order illumination polynomial $I(x, y, \lambda, h)$ where x, y are the image coordinates, λ is the color, and h is the image height using least squares. This algorithm only assumes that across multiple images at a height, reflectance is noise, a much weaker constraint than in any previous algorithm.

1.3 Related Graphics Work

Much work has been done into accounting for absorption and refraction in materials such as water. Watt presents a photon tracing algorithm which correctly accounts for caustics caused by water.[15] Sun et al. present an extension to any raytracer for dealing with absorption, and demonstrate the necessity of full spectrum over three channel data.[14] Yapo et al, Rendering Lunar Eclipses provides a base motivation for a simplified algorithm which solves a topical problem extremely well. It also provides a base for much of the absorption approximations used for this paper.[16]

2 Model

2.1 Lights

The HabCam uses four VIGI-Lux Machine Vision Strobes mounted radially about fifty centimeters from the camera. The strobes produce diffuse light 63.5 by 30.2 millimeters in size through the use of a diffuser. This diffuse light is then passed through a circular Fresnel lens to focus and colimate the light such that the four strobe beams provide a illumination pattern approximately one meter squared at an altitude of three meters. The fore and aft strobes are angled at approximately 26 degrees, while the port and starboard are approximately 28 degrees from vertical.

We wish to create an algorithm which doesn't employ photon mapping, therefore it is necessary to approximate the light output from a diffuse surface through a Fresnel lens without explicitly modeling that interaction. Since a Fresnel lens is an approximation of a plano-convex lens, we deal with the later for simplicity. From figure 1 we can see a pattern emerging, mainly that a a portion of the light is colimated, a portion is focused, and the remaining is distributed diffusely. First, since the F number for Fresnel lenses are normally approximately 1, we can guess that the focal length is probably on the order of a few centimeters, which makes any focused contribution from the lights on the seafloor impossible. Therefore, we can approximate the light as three directional light sources with increasing angles of influence. The percentage of light incident on a point is thus

$$P_I(n, d) = \alpha_{\text{colimate}} J(n, d; colimate) + \alpha_{\text{focus}} J(n, d; focus) + \alpha_{\text{diffuse}} J(n, d; diffuse) \quad (1)$$

Where α is a weighting factor decided by the percentage of light from the original diffuse source which translates to the type of light, n is the vector normal of the light, d is the direction to the light from the point, and $J(n, d; y)$ is a cosine function which equals zero any time the angle formed by n and d falls outside an acceptable range for the type of light.

2.2 Modeled Alpha Values

We set the alpha values based on model approximation. Assuming the diffuse light is at the focal point of the lens, all of the colimated light comes from the center of the light. Since the sensor is 6.3 cm by 3.02 cm, the area is approximately 19 cm squared. To compute the percentage of light which is colimated, we have to assume that anything within some radius of the focal point will be colimated. Intuitively we set the radius value to one centimeter, which means approximately 16.5 percent of original light is output as directional light, which we model as output which is within 10 degrees of the light normal. For focused light, we assume that any incident light within twenty five degrees of perpendicular focuses. Since the hemisphere equation is $2\pi hr$, with $h = 1 - \cos(\theta)$, we get that 9.4 percent of the original light is output as focused light. Now, we know that our focal length must be approximately the diameter of the lens which means that the angle of light emitted once we are past the focal point is, $\arctan(.5) \approx 26.5°$ from the normal. Now the remaining 74.1 percent of light is emitted as truly diffuse light.

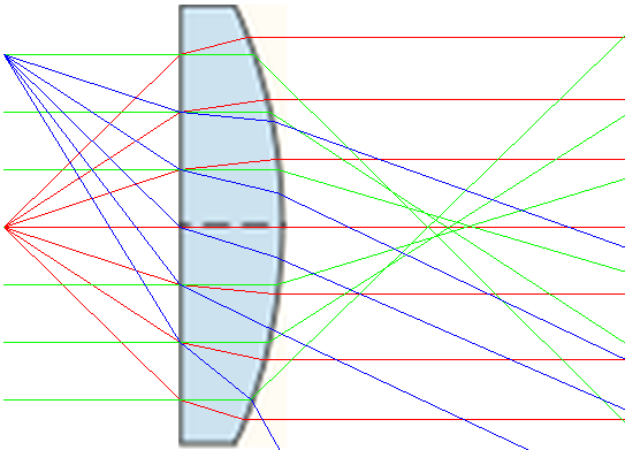


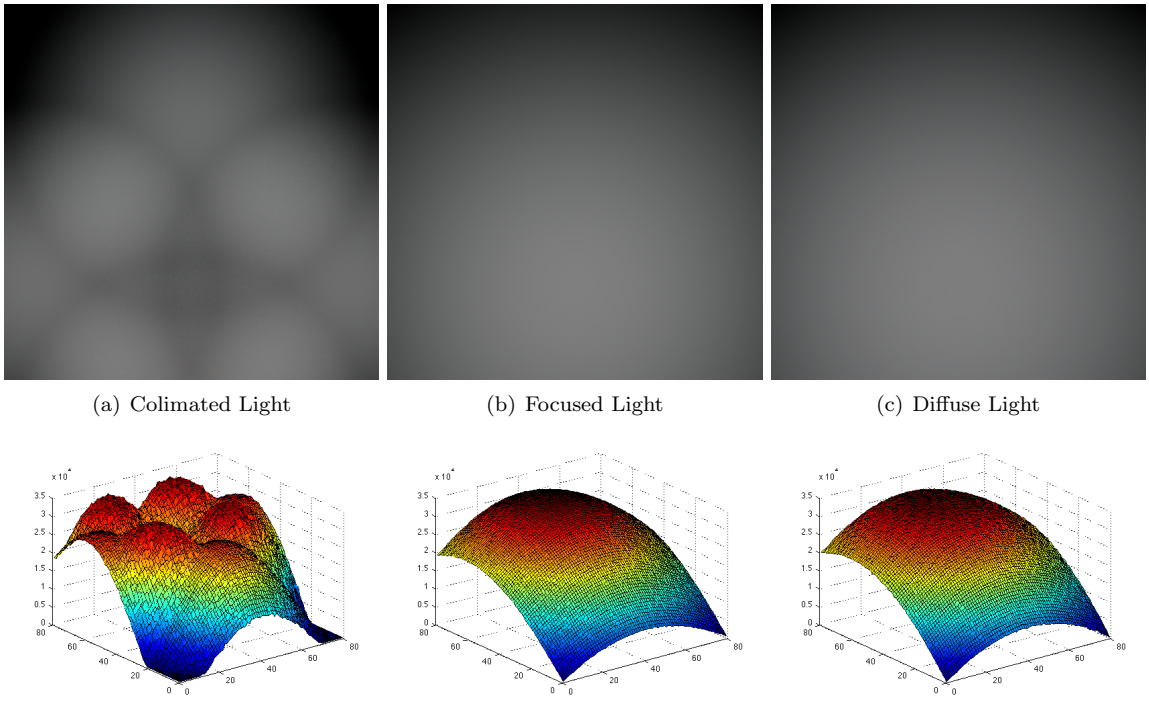
Figure 1: Focusing and colimation caused by a plano convex lens. In reality we would see point light from every point on the diffuse surface. For simplicity we only draw parallel light, light which passes through the focal point, and one non focal point light source (which can be seen to diffuse).

2.3 Empirical consideration of alpha values

Using the output angles for colimated, focused and diffuse light from above, three test cases were run setting one alpha value to 100% and the rest to zero for the habcam model at two meters. See figure 2. This gives us intuition about how the types of light affect the illumination pattern. The contribution from individual lights is also noticeable in the directional light pattern. It becomes immediately apparent that while the true illumination needs to be a linear combination of the three types of light, a linear combination doesn't exist which creates the dark band seen in the true illuminations in figure 8. This means that there must be something wrong with our model of the Fresnel lens, and at this time deriving a completely new model isn't within the time constraints of this project. It should be noted that removing the port and starboard lights produces a dark center band, however it is fairly faint compared to the true images. Also, it is highly unlikely that the side strobes aren't contributing to the illumination of the image. Further discussion of this is given in the future work section.

2.4 Flashbulbs

The flashbulbs used, are Xenon lamps which produce full spectrum light. We approximate the spectrum output based on the measured output given in a Xenon lamp datasheet, see figure 3. Let $I(\lambda)$ be the intensity of the light for a specific wavelength, therefore our illumination equation given a light's normal, direction to



(a) Colimated Light (b) Focused Light (c) Diffuse Light

(d) Colimated Light (e) Focused Light (f) Diffuse Light

Figure 2: Illumination caused by a single type of light, images have been scaled to accentuate illumination changes.

the light, and specific wavelength.

$$F(n, d, \lambda) = P_I(n, d)I(\lambda) \tag{2}$$

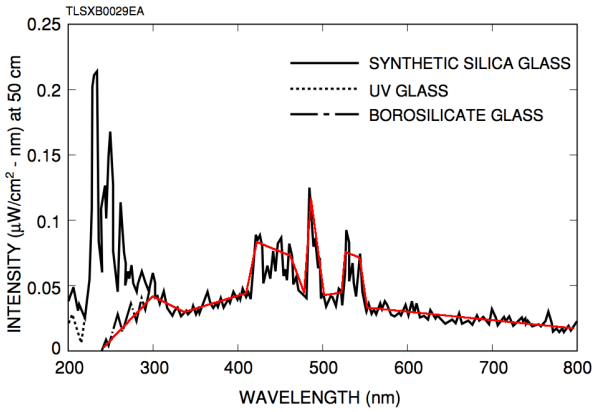


Figure 3: Spectral distribution of light from a Xenon flash bulb, from HAMATSU.[4] Red line indicates the approximation Used for this paper.

2.5 Water

Water absorption is taken from Pope et al. in which absorption in pure water was measured.[11] While salt water probably has different absorptions, this is the most complete measurement of water's absorption. We use their values linearly interpolated over the full spectrum.

2.6 Seafloor

The seafloor is approximated as a planar surface, however, since the HabCam is being towed, it almost always has a pitch of a few degrees. We take this into consideration and tilt the ground plane five degrees. Reflectance of the seafloor is dependent on the type of substrate, for ease of implementation, we chose sand because any specular component is limited and it has a flat shape. The reflectance properties from Barille et al for sand were used, and are interpolated to full spectrum. [1] See figure 4 for a plot of reflectance vs wavelength.

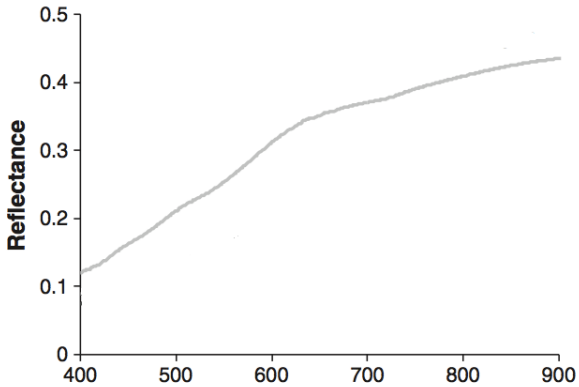


Figure 4: Percentage of reflectance of wavelength light by sand

2.7 Camera

While there are many color spaces which change spectrum light into a trismus space such as the commonly used CIE 1931 standard, those color spaces are modeled to accurately simulate human vision in a manner which color filters and ccd sensors can't. In a real camera, a postprocessing step is necessary to translate the raw camera output to RGB value final output. Since we are attempting to correct the attenuation in images while the images are still in a raw format, it makes sense that our camera model should attempt to model filter and ccd values rather than human vision. I was unable to find the datasheet for the sensor the UNIQ camera on the HabCam employs, so a best estimate taken from srgb filter response combined with a specific ccd efficiency is used.[2] See figure 5 for a graph of the percentage of light absorbed per color channel. The images produced appear for the most part to have reasonable colors, but improving the camera model might improve results.

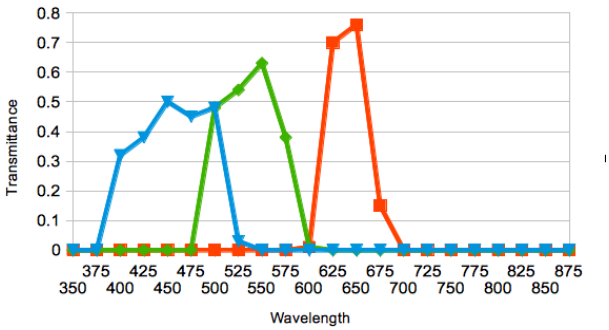


Figure 5: Percentage of light absorbed at wavelengths by the a ccd sensor with an scientific srgb filter per color channel.

3 Results

3.1 The Effect of Water

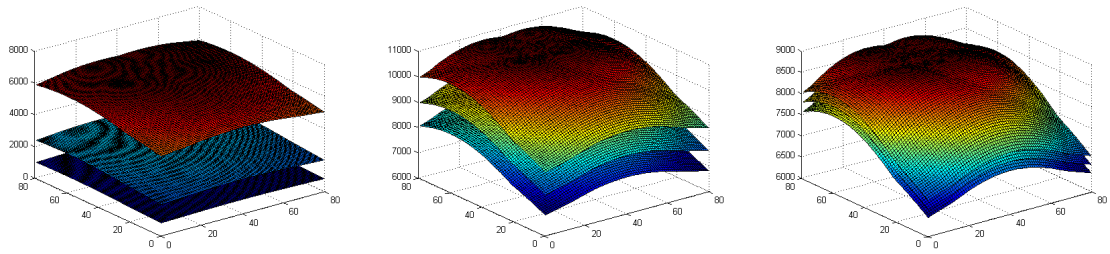
To get a sense for whether absorption is implemented correctly, three scenes were rendered with the same geometry approximating the habcam at an altitude of two meters, but with a different starting medium. The first is air with no absorption , any light falloff is due only to the decrease in intensity over area. The second scene has a medium which has half the absorption value of water, and the final scene is using the absorption values for water. Attenuation flattens and decreases the illumination values as can be seen in figure 6.

3.2 Comparison to Real Illumination

Returning to the illumination-reflectance model of an image $f(x, y) = I(x, y)R(x, y)$, due to the multiplication of illumination and reflectance there is a confounding scaling factor which which we are unable to extract. For the results images, we chose a scaling factor which allows for comparison between the lightmaps generated from real images, and the lightmaps generated from our model. Also, taking the average of a large number of images eliminates noise which in this case is the reflectance values.

Thanks to the incredibly large amount of data provided by WHOI, a set of images taken over a sandy substrate were identified and used for comparison. To compute the illumination at a height h, all images in this set within a epsilon (.1 meter) of h were averaged. A sample image from the set can be seen in figure 7. A side by side comparison of the illumination images can be seen in figure 8

Comparing the images, it is apparent that the modeled illumination is fairly different than the real illumination. Mainly the generated illuminations do not have the saddle curve in the middle. The saddle



(a) Red (b) Green (c) Blue

Figure 6: A comparison of the habcam illumination displayed as a surface in three different medium. Water is the bottom surface, air the top per channel.

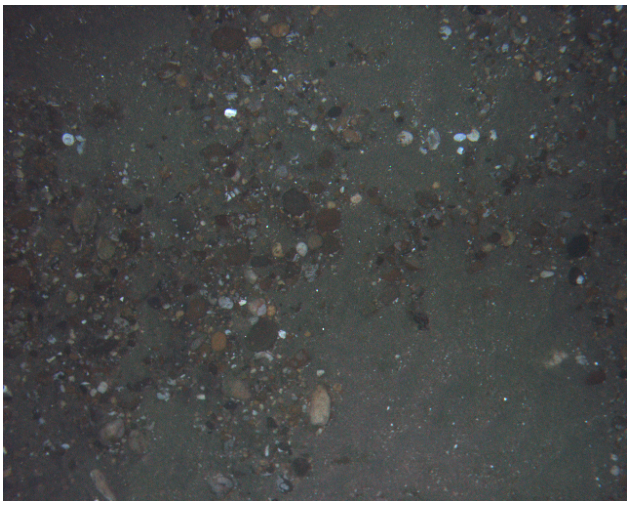


Figure 7: Example image from the specific set used to test the model against.

curve appears to be generated because the port and starboard strobes only illuminate one side of the image rather than the center as would be expected.

4 Conclusions

While a completely accurate simulation of the HabCam wasn't achieved, I believe this project was successful. All of the implementation for a full spectrum ray tracer with absorption and refraction was completed. A simulation of the HabCam illumination with illumination patterns which are at least somewhat reasonable was achieved. Also, given limited information on the hardware, the results are promising that an accurate model could be produced.

5 Future Work

In the future, it is necessary to work with the hardware to get accurate readings for everything from light output to sensitivity of the camera to specific wavelengths. Exact measurements for the strobes placement and angles is also important for accurate output. Without accurate measurements, it might also be possible to attempt to tweak the angle, size and shape of the lights in the scene to get a model which fits better. Full raytracing is also really expensive, so figuring out a way to make it faster is fairly important. Given

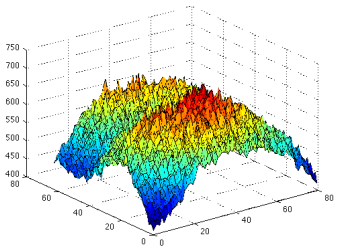
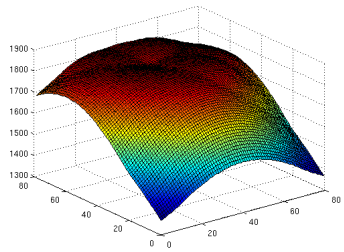
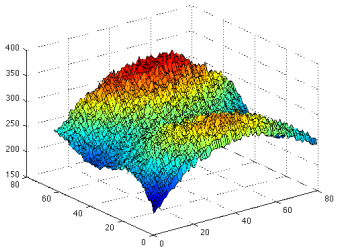
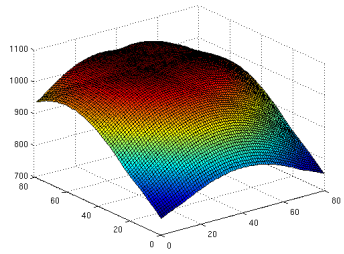
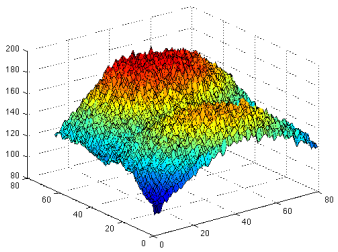
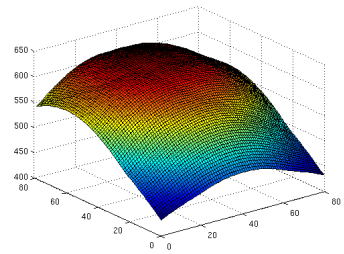
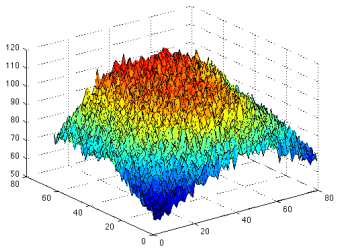
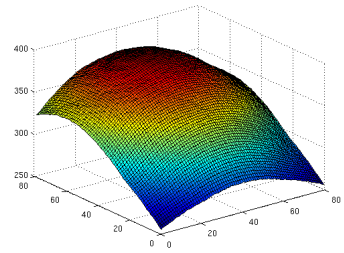
Height	Real	Modeled
1.5		
2		
2.5		
3		

Figure 8: Side by side comparison of illuminations for the red channel.

changing water conditions, it is important that we are able to generate illumination corrections on the fly. So the current 15 plus minutes per image isn't acceptable, since speed wasn't a focus, it is probably quite possible to speed up the current code substantially. The most important outcome of this work might be to potentially influence the design of future craft as far as angles and positions of strobes.

References

[1] L. Barille et al. "Spectral response of benthic diatoms with different sediment backgrounds". In: *Remote Sensing of Environment* (2011). ISSN: 0034-4257.

[2] Dick Berg. *The Response Function of a Filter-CCD Combination*. URL: `http://www.brightskies.us/filtCCD.html`.

[3] Uniq Vision Cameras. *UNIQ Vision, Inc. UP-1800-CL data sheet*. URL: `www.uniqvision.com/spec/up1800ds&1800dscl.pdf`.

[4] Perkin Elmer. *Perkin Elmer MVS-5000 data sheet*. URL: `www.perkinelmer.com/CMSResources/Images/44-3527DTS_mvs5000.pdf`.

[5] R. Garcia, T. Nicosevici, and X. Cufí. "On the way to solve lighting problems in underwater imaging". In: *OCEANS'02 MTS/IEEE*. Vol. 2. IEEE. 2002, pp. 1018–1024. ISBN: 0780375343.

[6] R.C. Gonzalez and R.E. Woods. "Digital Image Processing, Addison-Wesley". In: *Reading, Ma* 7 (1992).

[7] HAMAMATSU. *HAMAMATSU xenon flash lamp data sheet*. URL: `sales.hamamatsu.com/assets/applications/ETD/Xe-F_TLSX9001E05.pdf`.

[8] H. Narasimha-Iyer et al. "Automated analysis of longitudinal changes in color retinal fundus images for monitoring diabetic retinopathy". In: *IEEE Transactions on Biomedical Engineering* (2005).

[9] A.V. Oppenheim, R.W. Schafer, and T.G. Stockham Jr. "Nonlinear filtering of multiplied and convolved signals". In: *Proceedings of the IEEE* 56.8 (1968), pp. 1264–1291. ISSN: 0018-9219.

[10] S.M. Pizer et al. "Adaptive histogram equalization and its variations". In: *Computer vision, graphics, and image processing* 39.3 (1987), pp. 355–368. ISSN: 0734-189X.

[11] R.M. Pope and E.S. Fry. "Absorption spectrum (380–700 nm) of pure water. II. Integrating cavity measurements". In: *Applied Optics* 36.33 (1997), pp. 8710–8723. ISSN: 1539-4522.

[12] H. Singh et al. "Quantitative photomosaicking of underwater imagery". In: *OCEANS'98 Conference Proceedings*. Vol. 1. IEEE. 1998, pp. 263–266. ISBN: 0780350456.

[13] H. Singh et al. "Towards high-resolution imaging from underwater vehicles". In: *International Journal of Robotics Research* 26.1 (2007), pp. 55–74. ISSN: 0278-3649.

[14] Y. Sun, F.D. Fracchia, and M.S. Drew. "Rendering the phenomena of volume absorption in homogeneous transparent materials". In: *the 2nd Annual IASTED International Conference on Computer Graphics and Imaging (CGIM'99*. Citeseer. 1999, pp. 283–288.

[15] M. Watt. "Light-water interaction using backward beam tracing". In: *Proceedings of the 17th annual conference on Computer graphics and interactive techniques*. ACM. 1990, pp. 377–385. ISBN: 0897913442.

[16] Theodore C. Yapo and Barbara Cutler. "Rendering Lunar Eclipses". In: *Proc. Graphics Interface*. Kelowna, British Columbia, May 2009, pp. 63–69.