Estimating Surface Reflectance Spectra for Underwater Color Vision

S. Skaff, J. J. Clark, and I. Rekleitis
Centre for Intelligent Machines, McGill University
sandra,clark,yiannis@cim.mcgill.ca

Abstract

This paper introduces a novel mathematical approach to surface spectral reflectance estimation in unknown underwater environments using uncalibrated color cameras. The approach derives surface spectral estimates without explicitly modeling the underwater medium characteristics such as light scattering and absorption. The latter two phenomena are dependent upon two parameters, which are the distance of the object from the camera and the depth of the object in water. The proposed approach does not require these parameters to be specified in advance. Spectral models are useful for underwater applications, where subtle differences in color need to be distinguished. Such models are also useful for fusing information from multiple images. We show that the proposed approach yields promising results.

1 Introduction

Enhancing images or video footage of underwater scenes is crucial to many applications. Underwater images have been used for mine detection and also for the inspection of underwater communication cables [4]. Marine biology [7] and archaeology [9] frequently employ vision techniques. Underwater images collected by recreational scuba divers help them share their experience with others and keep treasured mementos. Furthermore, in several cases, divers use the pictures for post-dive identification of the marine life they encountered [1]. Finally, robust computer vision techniques are crucial in the performance of autonomous underwater robotic vehicles [12]. Color is an essential cue in many computer vision algorithms, and color correction in the above scenarios is a useful tool to improve the quality of underwater images. The color can be represented either in 3D vector form such as RGB, CMY, YIQ, etc. or in spectral form. The latter constitutes the surface reflectance spectrum, that is the amount of light reflected off an object's surface at each wavelength, and is a physical property of the surface.

Underwater color imaging and analysis is increasing in popularity, however, the problem of estimating surface reflectance spectra in such a context has been barely addressed. Even though challenging to obtain, spectral model representations are useful when there is a need for an accurate representation of a color. For example, when inspecting metal underwater, the difference between the colors of the defective and good parts may be sufficiently small to require the extra information contained in the spectral models. Spectral data has also been used to discriminate between different corals [5]. Currently, methods relying on spectral data to distinguish between different plants, algae, or any material underwater usually employ expensive equipment to obtain such data. Therefore, introducing

a technique which can estimate these spectra relying *only* on a cheap, consumer digital camera is of great convenience.

In addition to providing an accurate representation of a surface color, spectral models allow for simple fusion of information obtained from different images. In this paper, images of a multi-patch color target at different distances from the camera and at different depths are used. We demonstrate the usefulness of spectral fusion by showing that the approach provides better spectral estimates when gathering sensor responses from multiple images rather than a single image for the same surface patch. Moreover, the proposed approach provides an ideal framework for fusing information as it does not require knowledge of the different parameters such as the target's depth or its distance from the camera.

The major contribution of this paper is to introduce a mathematical approach to estimate reflectance spectra of surfaces in underwater scenes given *only* responses from uncalibrated camera sensors. The paper is structured as follows. Underwater color image formation is explained and derived in Section 2, and the proposed approach is detailed in Section 3. Even though the paper is theoretical in nature, the approach proposed is applied on real underwater images. The measures used to evaluate the proposed approach in estimating surface reflectance spectra are presented in Section 4. Results show that spectral estimates improve upon fusion of information from different images in Section 5.

2 Underwater Color Image Formation

Most vision algorithms are devised under the assumptions that the camera is placed in open air at a close distance from the scene. However, devising vision algorithms that can operate in other scenarios such as water or foggy air is more challenging. In these environments, the light reaching the camera undergoes a wavelength-dependent attenuation. For example, images taken in open air at a far distance may appear yellow-reddish at sunset [14], while underwater images appear to be mostly of green-blue hue [13]. In these contexts, surface spectral estimation is important as the spectrum is a physical property of the surface. It therefore represents the surface color with a multi-dimensional vector that is characteristic of the surface itself and independent of the surrounding environment.

In a typical underwater image acquisition scenario, the light emitted from the sun hits the water surface, where some of this light is reflected off and some of it is refracted into the water. At the camera sensor, the light arrives via 3 different paths as shown in Fig. 1.

Let us define the color signal as the wavelength-dependent light falling on the camera sensor. A color signal in vacuum is given by:

$$\Phi(\lambda) = S(\lambda)E(\lambda), \tag{1}$$

where $S(\lambda)$ is the surface reflectance spectrum, that is, the proportion of light reflected off a surface at each wavelength; $E(\lambda)$ is the illuminant spectrum; λ is the wavelength. Underwater, this color signal is a linear superposition of 3 color signals each corresponding to a different path for the light [6].

The first component of the color signal comes from light reflected off the object surface, and is attenuated as compared to the color signal which would be observed in open air at a close distance (Eq. 1):

$$\Phi_{uw_1}(\lambda) = \Phi(\lambda)e^{-\eta(\lambda)z}, \tag{2}$$

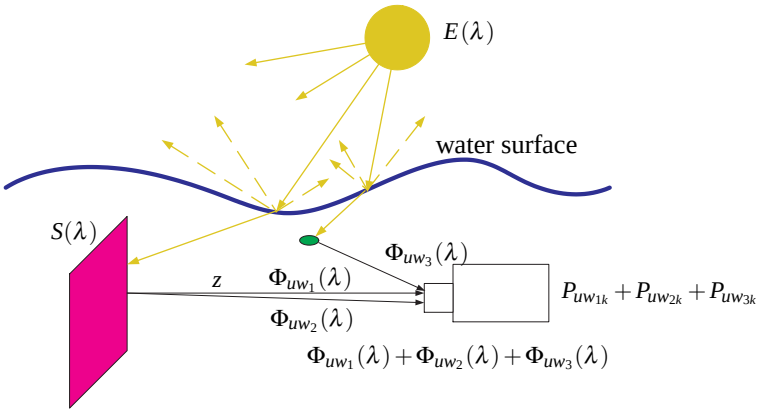


Figure 1: The components of light arriving at a camera sensor underwater. The green ellipse represents a particle underwater off which light is reflected, giving rise to a backscatter component. The variables are defined in the text of Section 2.

where $\eta(\lambda)$ is the attenuation coefficient given by $\eta(\lambda) = \alpha(\lambda) + \beta(\lambda)$; $\alpha(\lambda)$ is the absorption coefficient of the water; $\beta(\lambda)$ is the scattering coefficient, which represents how much light can be scattered by an infinitesimal volume of water; z is the distance between the object being viewed and the camera. For more detail please refer to [6].

The second component, $\Phi_{uw_2}(\lambda)$, arises from forward scattering of the light along the line of sight contributing to a blurring of the image. In our experiments, water clarity and camera-target distance were such that the blur was minimal. In addition, image blur would not affect the color of the image. Therefore $\Phi_{uw_2}(\lambda)$ is neglected.

The third component, $\Phi_{uw_3}(\lambda)$, arises from backscattered light. In simple terms, this is light that gets scattered backwards along the line of sight by particles suspended in the water between the object being viewed and the camera. The larger the distance between the object and the camera, the more light is backscattered.

Hence the color signal underwater that contributes to the image can be expressed as:

$$\Phi_{uw}(\lambda) = \Phi_{uw_1}(\lambda) + \Phi_{uw_2}(\lambda) + \Phi_{uw_3}(\lambda) \approx \Phi_{uw_1}(\lambda) + \Phi_{uw_3}(\lambda). \tag{3}$$

Projecting the color signals in this equation onto the sensor spectral sensitivities $R_k(\lambda)$ of a camera yields an equation in terms of the corresponding sensor responses:

$$P_{uw_k} = P_{uw_{1k}} + P_{uw_{3k}}, \quad k = 1, 2, ..., p, \tag{4}$$

where p is the number of sensor classes, each denoted by k. The sensor responses recorded from underwater images are P_{uw_k}. $P_{uw_{1k}}$ is the component of the responses corresponding to the attenuated light, while $P_{uw_{3k}}$ is the component corresponding to the backscattered light. Knowing that a black surface should have $P_{uw_k} = (0,0,0)$ where p=3, any offset from such a response is assumed to be due to backscattered light. Therefore, for a certain distance from the camera, subtracting the response of a black surface ($P_{uw_{3k}}$) from the measured response for a particular surface (P_{uw_k}) yields P_{uw_1k}, which given Eqs. 1 and 2, can be expressed as:

$$P_{uw_{1k}} = \sum_{\lambda=1}^{M} R_k(\lambda)\Phi(\lambda)e^{-\eta(\lambda)z} = \sum_{\lambda=1}^{M} R_k(\lambda)S(\lambda)E(\lambda)e^{-\eta(\lambda)z}, \tag{5}$$

where M is the dimension of the spectra.

3 Underwater Spectral Reflectance Estimation

The goal of this work is to estimate the surface reflectance spectrum of an object underwater given *only* camera sensor responses, denoted by $P_{uw_{1k}}$ as given in Eq. 5. The unknowns are the camera sensor spectral sensitivities $R_k(\lambda)$, the illuminant spectrum $E(\lambda)$, and the attenuation function $e^{-\eta(\lambda)z}$. Grouping the unknowns of Eq. 5 into one function $C_k(\lambda) = R_k(\lambda)E(\lambda)e^{-\eta(\lambda)z}$, the responses can be expressed as:

$$P_{uw_{1k}} = \sum_{\lambda=1}^{M} S(\lambda)C_k(\lambda). \tag{6}$$

We refer to $C_k(\lambda)$ as the product spectrum. In the following sections, we discuss how we model each of the components of Eq. 6 before discussing the proposed algorithm.

3.1 The Surface Reflectance Spectra

We represent surface reflectance spectra by maximum entropy models. Such models were successfully used to estimate Munsell patch reflectance spectra given only photoreceptor responses in [2]. The use of maximum entropy models was inspired by Jaynes, who stated that a physical quantity frequently observed in practice will tend to a value that can be produced in the largest number of ways [8]. In the case of physical processes representing spectra, many surfaces observed in our everyday-life surroundings have spectra of high entropy, as opposed to monochromatic surfaces which have low entropy spectra [2].

The surface spectra need to be represented by probability density functions (pdf) in order to compute their entropy. The spectrum of a collection of photons can be thought of as a histogram of photons over wavelength. An incident photon can be either absorbed by or reflected off the surface it hits. Denoting the event that a photon is absorbed by the surface by A and the event that a photon is reflected off the surface by R, we can write: $P(A)+P(R)=1$. Now $\Phi(\lambda)$ as given in Eq. 1 can be represented by the pdf $p(\lambda|R)$, the probability of the wavelength given that event R occurred, which means that a photon has been reflected off the surface. Given Bayes' rule:

$$p(\lambda|R) = \frac{p(R|\lambda)\ p(\lambda)}{p(R)}, \tag{7}$$

where $p(\lambda) \equiv p(\lambda|photon_0)$ denotes the pdf of a wavelength given an incident photon, $photon_0$, and represents the illuminant. Since it is assumed that there is always an incident photon, $p(\lambda|photon_0)$ is written as $p(\lambda)$. Given Eqs. 1 and 7, the surface reflectance spectrum can be written as such:

$$S(\lambda) = p(R|\lambda)/p(R). \tag{8}$$

This means that $S(\lambda)$ denotes the likelihood function of a particular wavelength given the event R. This function is normalized by the fraction of photons reflected off the surface over all wavelengths. The pdf representation of $S(\lambda)$, $p_S(\lambda)$, is obtained by the following:

$$p_S(\lambda) = S(\lambda)/\sum_{\lambda=1}^{M} S(\lambda), \quad \lambda = 1,...,M. \tag{9}$$

The entropy H of $p_S(\lambda)$ is defined as:

$$H = -\sum_{\lambda=1}^{M} p_S(\lambda) \log p_S(\lambda). \tag{10}$$

3.2 The Product Spectra

A typical consumer digital camera has 3 sensor classes, red, green, and blue ($p = 3$ in Eq. 4). Therefore, the corresponding product spectra of the sensor spectral sensitivities, the illuminant and the water attenuation spectra over wavelength are:

$$C_k(\lambda) = R_k(\lambda)E(\lambda)e^{-\eta(\lambda)z}, \quad k = 1,2,3. \tag{11}$$

For these product spectra, a model that does not require knowledge of specific parameters such as those of the water absorption is sought. We resort to a linear model using Fourier bases. Unlike other basis functions, Fourier bases do not require a database of spectra to be specified in advance in order to perform principal components analysis on. Each of the product spectra is modeled as such:

$$C_k(\lambda) = \sum_{j=1}^{J} b_{jk} B_j(\lambda), \tag{12}$$

where J is the number of basis functions used for each of the product spectra models, $B_j(\lambda)$ denotes the basis functions used, and b_{jk} denotes the weight for the j^{th} basis function for the k^{th} sensor sensitivity curve. We choose to use 9 Fourier bases ($J = 9$) as testing the algorithm on more than 9 Fourier bases did not provide a significant improvement in modeling the surface spectra. Moreover, Finalyson *et al.* [3] found that 9-15 Fourier bases are sufficient to model the camera sensor spectral sensitivities, which have the same shapes as the product spectra.

3.3 The Algorithm

The input to the algorithm is a set of sensor responses from an underwater image, a vector of dimension M for the surface spectrum, initialized to $1/M$, and 3 initial weight vectors, each initialized to $1/J$ ($J = 9$), corresponding to each of the product spectra. A cost function comprising of 2 main components is formed. The first term represents the entropy of the pdf representation of the surface spectrum and the second term represents the closeness of the computed sensor responses to the measured sensor responses, such that:

$$\begin{aligned}[\hat{S(\lambda)}, \hat{\mathbf{b_1}}, \hat{\mathbf{b_2}}, \hat{\mathbf{b_3}}] &= \underset{S(\lambda), \mathbf{b_1}, \mathbf{b_2}, \mathbf{b_3}}{\operatorname{argmin}} \sum_{\lambda=1}^{M} p_S(\lambda) \log p_S(\lambda) + \left[P_{uw_{11}} - S(\lambda) \sum_{j=1}^{J} b_{j1} B_j(\lambda) \right]^2 \\ &+ \left[P_{uw_{12}} - S(\lambda) \sum_{j=1}^{J} b_{j2} B_j(\lambda) \right]^2 + \left[P_{uw_{13}} - S(\lambda) \sum_{j=1}^{J} b_{j3} B_j(\lambda) \right]^2 \end{aligned} \tag{13}$$

where $\mathbf{b_1} = [b_{11} ... b_{J1}]$, $\mathbf{b_2} = [b_{12} ... b_{J2}]$, and $\mathbf{b_3} = [b_{13} ... b_{J3}]$. To minimize the cost function, *fmincon* for nonlinear constrained optimization from the Matlab Optimization Toolbox is used. The constraints are those of positivity on the surface spectrum and each of

the product spectra. The wavelength range considered for the spectra is 400-700 nm discretized into 10 nm intervals. Therefore the dimensionality of each spectrum is $M = 31$.

In the above derivations, it is assumed that there is only one surface patch in the scene. Upon adding a second surface patch, the new cost function would comprise of the terms in Eq. 13 in addition to similar ones pertaining to $S_2(\lambda)$, the sought spectrum for the second surface patch. We should note that the proposed approach seeks one of the multiple combinations of maximum entropy surface and product spectral models that can give rise to the same sensor responses. Using sensor responses from multiple surface patches or multiple images of the same patch constrains the possible combinations, thus resulting in more accurate surface spectral estimates. The fusion is facilitated by the use of spectral models, which are invariant representations of surface colors.

4 Underwater Spectral Reflectance Estimation Evaluation

The results of the proposed approach are validated by measuring the root mean square error and by displaying color patches constructed from the estimated spectra. These patches give insight as to what color an underwater robot would perceive once it has a system in place to estimate spectra.

4.1 Root Mean Square Error

The root mean square error (RMSE) between the actual spectrum $S_A(\lambda)$ of a surface and the spectrum obtained from the proposed algorithm $S_M(\lambda)$, referred to as a model spectrum, is computed. The Munsell patch spectra (*matte*) are those measured by Parkkinen *et al.* [11]. The RMSE for a surface spectrum is $E_{RMSE} = \sqrt{\frac{1}{M}\sum_{\lambda=1}^{M}(S_A(\lambda) - S_M(\lambda))^2}$.

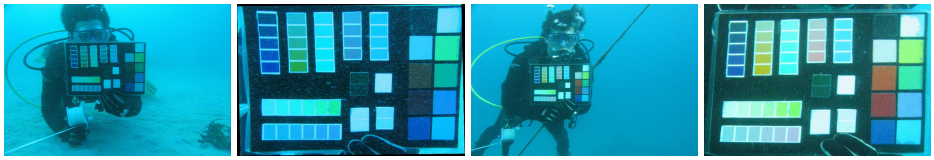
4.2 Qualitative Evaluation

The CIE XYZ and the CIE xy chromaticity coordinates of each of the actual and the model spectra are also computed. We assume white light as we want to measure the error in chromaticity values for the surface spectra without having it affected by illumination. The color matching functions used are those of the CIE 1931 standard observer (2°) as introduced by Judd (1951) and modified by Vos (1978). Denoting the coordinates corresponding to each of the actual and model spectra by $x_A y_A$ and $x_M y_M$ respectively, the CIE chromaticity error is the Euclidean distance between these coordinates: $E_{xy} = \sqrt{(x_A - x_M)^2 + (y_A - y_M)^2}$.

In order to visualize the patches constructed from the model spectra, we transform the CIE XYZ coordinates as computed above to coordinates in the RGB (Red, Green, Blue) color space using the CIECAM02 model [10]. Given these RGB coordinates, we display the corresponding patches, which are referred to as reconstructed patches. The patches obtained from both the model and the actual spectra are displayed in order to account for any discrepancies that may have arisen upon conversion to RGB or upon display.

5 Experimental Results

The images of a color patch target were taken by 2 scuba divers in the Caribbean Sea in Barbados. The target has a set of 41 Munsell patches glued on black cardboard. We choose to use Munsell surface patches in our experiments as these have high entropy spectra [2] and it was shown in [15] that maximum entropy modeling provides reasonable estimates of these spectra in the case of real images. All the images were taken with a Canon PowerShot A85 camera. The images were taken at 3 depths: about 7.9 m, 5.2 m, and 4.9 m, and at 2 distances from the camera: about 1.5 m (close) and 3.5 m (far). Fig. 2 shows the images of the target for the 7.9 m and 4.9 m depth cases at the far and close distances.



(a) (b) (c) (d)

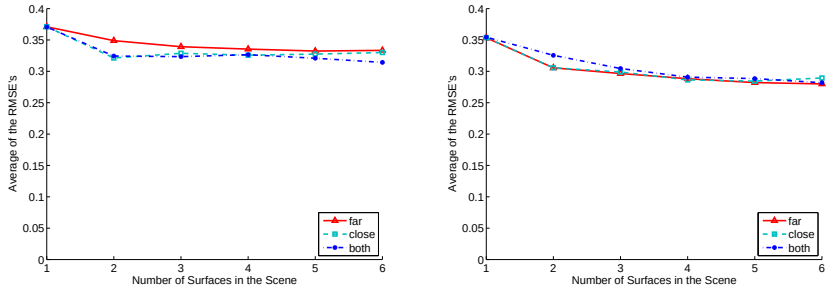
Figure 2: The target at a (a) depth of 7.9 m and distance of 3.5 m, (b) depth of 7.9 m and distance of 1.5 m, (c) depth of 4.9 m and distance of 3.5 m, (d) depth of 4.9 m and distance of 1.5 m; where distance denotes the distance from the camera.

Upon segmenting the images with a semi-automated algorithm, 10x20 pixel rectangles in the middle of each patch were obtained. The sensor responses over these pixels were averaged to obtain one RGB response per patch. The spectral estimate for each of the 41 surface patches was obtained by minimizing the cost function given in Eq. 13 using 0 to 5 additional patches each time. The latter patches were chosen at random from the remaining 40 Munsell patches on the target. Therefore a total of 41 cases, which we refer to as scenes, were constructed for each number of surface patches per scene (1 to 6). The average of the RMSE's, which are computed as explained in Section 4.1, is taken for the repeated surface patch over all 6 scenes. Two cases are investigated: (1) the effect of changing the distance of the target from the camera at each of 2 depths in Section 5.1 and (2) the effect of changing the depth of the target for the close and far distances from the camera in Section 5.2.

5.1 Variation of Distance from the Camera

For each of the 7.9 m and 4.9 m depth cases, we plot the average of the RMSE's versus the number of patches in the scene in Fig. 3 when the target is (1) far from, (2) close to the camera, and (3) when the sensor responses from both cases are combined.

Fig. 3 shows that the spectral estimates improve upon combining the sensor responses from images at both distances. The improvement is more evident at 4.9 m than at 7.9 m. In the case where images at both distances were used, the average RMSE drops from 0.37 to 0.31 for the 7.9 m deep images and from 0.35 to 0.28 for the 4.9 m deep images as the number of surfaces in a scene increases from 1 to 6. This shows that the performance of the approach is better at shallower levels where the daylight illumination is less attenuated, due to scattering and absorption, than at deeper levels. Also the spectral estimates



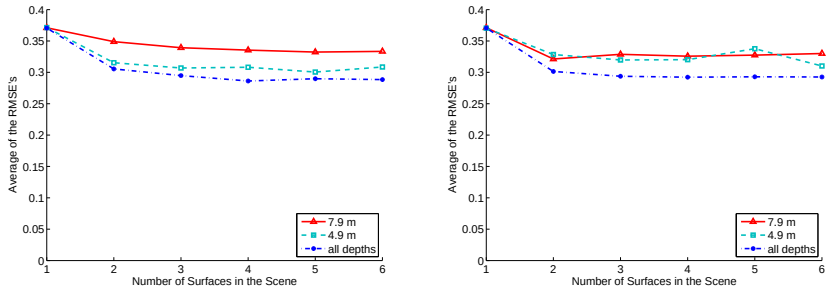
(a) 7.9 m deep images (b) 4.9 m deep images

Figure 3: The average of the RMSE's over all the scenes as a function of the number of patches in a scene when the target is far from the camera, close to the camera, and when the responses from the far and close images are combined at a depth of (a) 7.9m, (b) 4.9m.

generally improve upon the introduction of more surface patches into the scene, hence the usefulness of spectral models in fusing information.

5.2 Variation of Depth

For each of the far and close distance cases, we plot the average of the RMSE's versus the number of patches in the scene in Fig. 4 when the target is (1) 7.9 m deep, (2) 4.9 m deep, and (3) when the sensor responses from images at all depths are combined.

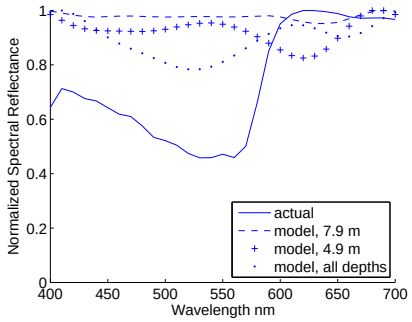


(a) far (b) close

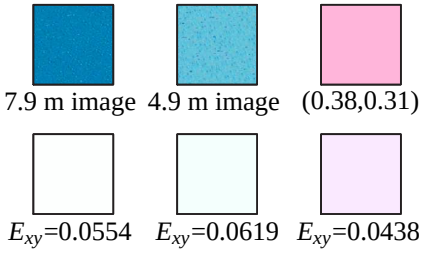
Figure 4: The average of the RMSE's over all the scenes as a function of the number of patches in a scene when the target is at a depth of 7.9m, 4.9m and when the responses from the images at all depths are combined at the (a) far, (b) close distance to the camera.

These plots show that the spectral estimates improve upon the introduction of more surface patches into the scene as in Section 5.1. Moreover, fusing spectral estimates across different depths considerably improves upon the average RMSE. This average drops from 0.37 to 0.29 as the number of surfaces in a scene increases from 1 to 6 in each of the far and close distance cases respectively.

For each of the different depths cases, we plot the model spectra obtained for Munsell patch 5RP 6/6 for the deep images in Fig. 5. We also display the corresponding color



(a) E_{RMSE}=0.3266 (7.9 m), 0.3009 (4.9 m), 0.2371 (all depths)



(b)

Figure 5: Munsell Patch 5RP6/6 for the far distance case. (a) The actual and model spectra for the 7.9m, 4.9m, and all depths cases, (b) in left-right, top-bottom order: the patches cropped from the target images taken at 7.9m and 4.9m depths, the reconstructed patches from the actual spectrum with its xy chromaticity coordinates, and the model spectra in the 7.9m, 4.9m, and all depths cases with the corresponding CIE chromaticity errors.

patches. It is interesting to see that despite the fact that the patches cropped from both the deep and shallow images are bluish with no visible pinkish tint, the proposed approach is able to retrieve enough color information upon combining the sensor responses from all depths and without assuming any prior information. Moreover, even though the chromaticity errors are small, there is still some discrepancy in the perceived colors of the patches. This discrepancy arises from the difference between the model and the actual spectra, which contain essential and more accurate information than chromaticities.

6 Concluding Remarks

We introduced a new mathematical approach that solves for surface reflectance spectra using images from an uncalibrated, consumer-level, digital color camera in the absence of a detailed model of absorption and scattering effects of an underwater medium. Therefore, the approach provides the flexibility of being able to operate in a medium other than in open air at close distances. By estimating reflectance spectra, which are physical representations of a surface color, the approach provides the flexibility of fusing information from multiple surfaces of one image or from multiple images.

Acknowledgments

The authors would like to thank Eric Bourque who assisted in taking the underwater images. Financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC) is gratefully acknowledged.

References

[1] M. Chambah, D. Semani, A. Renouf, P. Courtellemont, and A. Rizzi. Underwater color constancy: Enhancement of automatic live fish recognition. *SPIE*, 5293:157–168, 2003.

[2] J. J. Clark and S. Skaff. Maximum entropy models of surface reflectance spectra. In *IMTC*, volume 2, pages 1557–1560, 2005.

[3] G. D. Finlayson, S. Hordley, and P. M. Hubel. Recovering device sensitivities with quadratic programming. In *CIC*, pages 90–95, 1998.

[4] G. L. Foresti. Visual inspection of sea bottom structures by an autonomous underwater vehicle. *IEEE Trans. SMC*, 31(5):691–705, 2001.

[5] H. Holden, E. LeDrew, and C. Melsheimer. Simulated spot-equivalent reflectance characteristics of common coral reef features. In *ACRS*, pages 22–25, 1999.

[6] J. S. Jaffe. Computer modeling and the design of optimal underwater imaging systems. *IEEE J. Oceanic Eng.*, 15(2):101–111, 1990.

[7] J.Ahlen, D. Sundgren, T. Lindell, and E. Bengtsson. Dissolved organic matters impact on colour reconstruction in underwater images. In *SCIA*, volume 3540, pages 1148–1156, 2005.

[8] E. T. Jaynes. Prior probabilities. *IEEE Trans. SSC*, 4:227–241, 1968.

[9] Y. Kahanov and J. Royal. Analysis of hull remains of the Dor D Vessel, Tantura Lagoon, Israel. *IJNA*, 30(2):257–265, 2001.

[10] N. Moroney, M. D. Fairchild, R. W. G. Hunt, C. J. Li, M. R. Luo, and T. Newman. The CIECAM02 color appearance model. In *CIC*, pages 23–27, 2002.

[11] J. P. S. Parkkinen, J. Hallikainen, and T. Jaaskelainen. Characteristic spectra of Munsell colors. *JOSA A*, 6(2):318–322, 1989.

[12] J. Sattar and G. Dudek. On the performance of color tracking algorithms for underwater robots under varying lighting and visibility. In *ICRA*, pages 3550–3555, 2006.

[13] Y. Y. Schechner and N. Karpel. Clear underwater vision. In *CVPR*, volume 1, pages 536–543, 2004.

[14] S. Shwartz, E. Namer, and Y. Y. Schechner. Blind haze separation. In *CVPR*, volume 2, pages 1984–1991, 2006.

[15] S. Skaff and J. J. Clark. Maximum entropy spectral models for color constancy. In *CIC*, pages 100–105, 2007.