

# Underwater Stereo

José P. Queiroz-Neto  
Centro Federal de Educação Tecnológica  
Departamento de Informática  
Av. Danilo Areosa, 138, Manaus, AM  
jppq@dcc.ufmg.br

Rodrigo Carceroni, Wagner Barros, Mário Campos  
Universidade Federal de Minas Gerais  
Departamento de Ciência da Computação  
Av. Antônio Carlos, 6627, Belo Horizonte, MG  
{carceron,wbarros,mario}@dcc.ufmg.br

## Abstract

*Images of underwater scenes suffer from poor contrast. Water-induced contrast decay varies across the scene and is exponential in the depths of scene points, which prevents standard computer vision algorithms from operating properly. In this paper we show how to overcome this problem by adapting an existing model of light propagation in the (foggy) atmosphere to describe the behavior of light in liquid media. By integrating the resulting model within a dense stereo algorithm, we recover disparity maps of scenes immersed in water from pairs of images of these scenes acquired from distinct viewpoints. Experiments performed with real underwater images of various degrees of turbidity show that the use of a physically-based light-propagation model allows one to reconstruct underwater scenes more accurately than with standard stereo algorithms alone.*

## 1. Introduction

Sensing, mapping and manipulation in underwater environments is a type of application of computer vision techniques that has great potential for economic and social impact. The ability to use vision as a sensing modality in this kind of environment is very helpful in tasks as diverse as maintenance of underwater structures, water quality monitoring, and water fauna identification and assessment [20, 14, 10]. Unfortunately, underwater images are subject to strong influence of particles in suspension within the liquid medium where light propagates and, as a consequence of this, they are typically much less sharp than most images acquired in the atmosphere [22], and in general not amenable to standard computer vision techniques.

More specifically, two factors that impair traditional computer vision algorithms when they are used in underwater scenes are *absorption* and *scattering* [1, 7]. Absorption, the physical process in which light interacts with matter and is converted in other forms of energy, makes bright objects have their apparent intensity reduced as they move

away from the camera used in the image-capture process. The effects of this phenomenon are combined with those of scattering, the physical process where photons “collide” with particles and change their propagation direction, in general either by a small angle (forward scattering) or by about 180 degrees (backscattering). As a result of this later phenomenon, dark objects tend to appear brighter as they move away from the camera. Combined, these two effects cause significant reduction in image contrast.

Even though light propagation in various media has been extensively studied in the areas of radiative transport [6, 3] and computer graphics [8], the models developed in these areas to explain the phenomena mentioned above have a computational cost that is prohibitively high.

This has led many researchers, especially in the discipline of Autonomous Underwater Robotics [21, 5, 15], to use *ad hoc* image enhancement methods that do not try to model exactly how underwater images are formed, or to use auxiliary sensors of other modalities such as sonars. While this kind of approach works fine for tasks such as autonomous navigation of underwater vehicles, it does not solve the problem of understanding what exactly is the content of the images acquired.

On the other hand, computationally-efficient models of absorption and scattering have appeared recently in the computer vision literature [13]. But, as pointed out by Narasimhan and Nayar [11], these models refer specifically to the propagation of light in the (foggy) atmosphere and do not describe in a precise way the behavior of light within liquid media. Moreover, except in very particular scenes, existing algorithms that use these models for recovery of 3D structure and removal of weather effects from images [12] require the availability of two images of the same scene on light-propagation media with significantly distinct properties (*i.e.*, under different weather conditions), which may take days to acquire. Thus, such algorithms are not useful in applications where scene contents change faster than the optical properties of the light-propagation medium, such as

fish identification and counting.

In this paper we introduce a methodology to reconstruct the geometry of 3D scenes immersed in liquid media that overcomes the limitations discussed above by:

1. adapting existing efficient models of light propagation in non-transparent media to the environmental conditions found in underwater applications;
2. using as input pairs of images of each scene acquired from distinct viewpoints under the same absorption and scattering conditions, rather than pairs of images acquired from a single viewpoint under different light-propagation conditions.

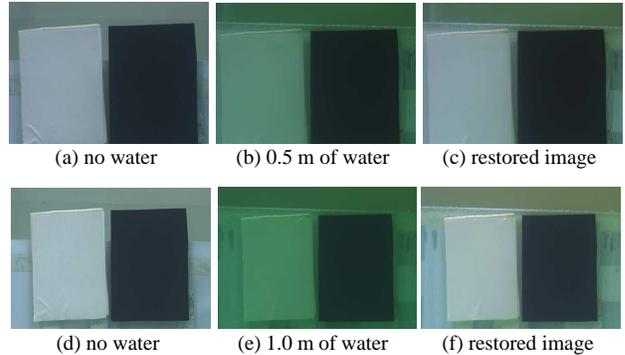
The second point above greatly increases the range of applications of the techniques developed because one can obviously acquire distinct-viewpoint images simultaneously, simply by using two cameras.

A motivation for using two views of each scene in our methodology is the recent surge of interest on dense stereo techniques within the computer vision community [17]. Dense stereo is a huge optimization problem. It amounts to finding what is the mapping between individual pixels of one image and corresponding pixels of a second image that minimizes some pre-defined global matching error. Research in this area has largely focused on finding restrictions on the set of possible mappings sufficiently strong to make the problem manageable under various degrees of computational-power limitation [19, 4, 16].

Dense stereo methods usually assume that the cameras used in image acquisition have been calibrated *a priori* both geometrically and radiometrically, so that the correspondences between epipolar lines are known and pixels that are images of the same 3D point should ideally have the same colors. This last assumption, in particular, prevents most existing dense stereo methods from working in underwater scenes because, as shown in Figure 1, in such scenes the color of each pixel depends not only on the illumination and on the radiometric properties of the scene’s visible surface, but also on the length of the water column that light needs to traverse between the scene and the camera. Thus, if the images in an underwater stereo pair are taken from different distances from the scene, corresponding pixels in them may have colors that are very different.

Nonetheless, if the absorption and scattering properties of a medium are uniform and known *a priori* and the distance between the camera and an object is also known, it is possible to take an actual image of the object in this medium and to adjust the pixel intensities in order to create an artificial image that looks like a real image acquired in an absorption- and scattering-free environment.

In this paper, we exploit this idea by (1) performing *a priori* radiometric calibration not only of the camera (as in traditional dense stereo) but also of the light-propagation



**Figure 1. Removal of attenuation and scattering effects from underwater images. (a) Image of a target out of water, (b) image of the same target under 0.5 m of water, (c) image generated by restoring (b), using the light-propagation model of Equation (5), (d)-(f) same as (a)-(c), but under 1.0 m of water.**

medium, and (2) using this information within a modified dense stereo algorithm in order to correct pixel colors when comparing pairs of pixels on corresponding epipolar lines. Note that such corrections *must* be done during the execution of dense stereo when depths of various objects in the scene with respect to each camera are *not* known *a priori*.

## 2. Light Propagation Model

Light absorption and scattering are complex phenomena whose exact effects depend on the type, orientation, size and spatial distribution of the particles that form the medium, as well as on the wavelength and polarization of incident light [14]. In this paper we use very simplified models of these phenomena that are, nonetheless, sufficient to approximate reasonably the process of image formation in scenes immersed in water.

More specifically, we use a the light propagation model that is largely inspired in that of Narasimhan and Nayar [12], but is adapted to describe propagation of light in water more accurately. As illustrated in Figure 2, this model assumes that the intensity of each pixel has two components: *attenuated scene intensity* and *waterlight*. Attenuated scene intensity is the effect of the energy that leaves the visible scene surfaces and travels in straight lines to the camera. Of course, because of both absorption and scattering, only a fraction of the energy that does leave the scene towards the camera actually reaches the camera’s lens. Thus the contribution of attenuated scene intensity to pixel intensities decreases with the length of the water column between the scene and the camera. Waterlight is the effect of energy

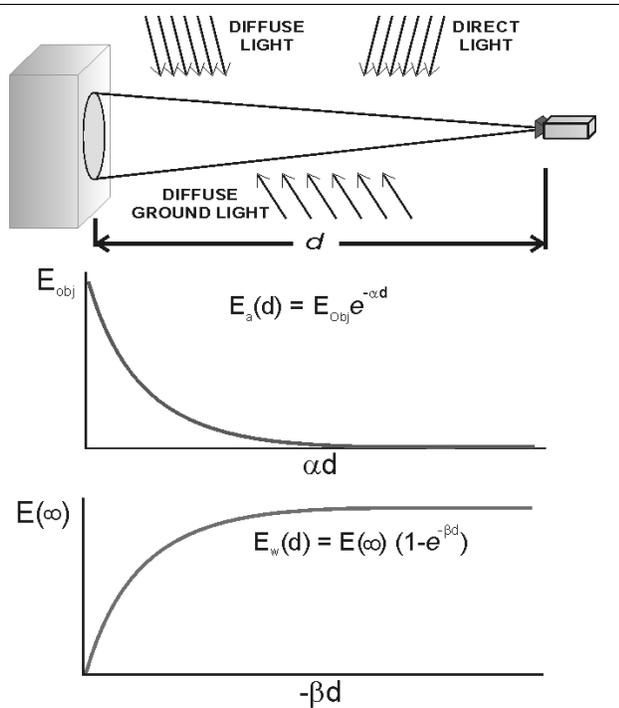


Figure 2. Light propagation model.

propagating along arbitrary directions within the water that gets scattered along the direction of the camera and is perceived as having come from visible surfaces of the scene, when in fact it did not. Waterlight results in an increase of pixel intensities and its contribution to the image formation process increases with the length of the water column.

As in the model used by Narasimhan and Nayar [12], the total radiance,  $E$ , captured by each sensor element, is a sum of two terms, as shown the following equation [6]:

$$E(d) = E_a(d) + E_w(d), \quad (1)$$

where  $d$  is the length of the water column between the sensor element and the scene point imaged by it, and  $E_a(d)$  and  $E_w(d)$  are, respectively, the attenuated scene intensity and the waterlight radiance.

As illustrated in Figure 2, at each differential water column length that light traverses, a fixed fraction of its energy is absorbed [1, 7], which results in an exponential decay of the radiance that exits the scene surface,  $E_{obj}$  as a function of the total length of the light path in the water,  $d$ :

$$E_a(d) = E_{obj}e^{-\alpha d}. \quad (2)$$

The scalar value  $\alpha$  in Equation (2), which we call *optical density*, measures the strength of absorption and scattering in the medium.

Similarly, scattering towards the camera also occurs at a fixed proportion per differential water column element.

However, part of the energy that enters the optical path that leads to the camera is absorbed or re-scattered later on, before it exits the water. The resulting behavior is modelled in the following equation [12]:

$$E_w(d) = K(1 - e^{-\beta d}), \quad (3)$$

where  $K$  is a scalar factor that depends on how the water is illuminated and on the water's scattering function, and  $\beta$  is a property of the medium that we call *waterlight coefficient*.

Our model differs from that of Narasimhan and Nayar [12] because we use separate coefficients  $\alpha$  and  $\beta$  to characterize the strength of attenuation and waterlight effects, respectively, while they use the simplifying assumption that  $\alpha = \beta$ , which is reasonable in the atmosphere, but not in a much optically denser medium such as the water.

In Equation (3), the radiance  $E_w(d)$  is maximized when  $d$  tends to infinity, *i.e.*, if the "object" being imaged is actually the scene background. Substituting the value  $d = \infty$  in Equations (2) and (3) and combining the results via Equation (1), we obtain  $K = E(\infty)$ , *i.e.*,  $K$  can be estimated from the *observed* intensity of points in the scene background. Using this fact, we can re-write Equation (3) as

$$E_w(d) = E(\infty)(1 - e^{-\beta d}). \quad (4)$$

Putting everything together, the complete model for the intensity  $E(d)$ , in a specific wavelength, of an arbitrary pixel in an underwater image is:

$$E(d) = E_{obj}e^{-\alpha d} + E(\infty)(1 - e^{-\beta d}). \quad (5)$$

Since absorption and scattering vary with wavelength in very complex ways, we deal with color images simply by having distinct  $\alpha$  and  $\beta$  coefficients for each of the red, green and blue channels.

Importantly, by a simple algebraic manipulation of Equation (5) we obtain a formula to estimate the radiance that leaves the scene's visible surfaces,  $E_{obj}$ , from the radiance that is measured by the camera,  $E(d)$ , after light traverses a water column of length  $d$ :

$$E_{obj} = [E(d) - E(\infty)(1 - e^{-\beta d})] e^{\alpha d}. \quad (6)$$

### 3. Underwater Stereo Methodology

A central problem in stereo vision, whose resolution yields the three-dimensional structure of any scene, is the so-called correspondence problem. It amounts to identifying every pair of pixels that are visible projections of the same 3D point in two distinct input images. The existence of ambiguities, *i.e.*, of multiple plausible matches for a given pixel, or on the other hand the absence of any feasible correspondence for a pixel, due to occlusion, complicate the problem. Fortunately, in geometrically- and

radiometrically-calibrated set-ups, there are a few constraints that greatly reduce the problem’s search space. A strong one is the epipolar constraint, which restricts the geometric locus where a pixel’s correspondence may be found to a line in the other image. But even with this type of constraint, which is universally employed in dense stereo techniques, in order to find a pixel’s correspondence it is still necessary to compare its color to those of its candidate matches, usually under the assumption that the closer are the colors of two pixels, the more likely they are of being stereo matches.

In underwater environments the correspondence problem is exacerbated by the medium, since each scene point will typically be located at distinct distances from the two cameras and hence, according to Equation (5), the measured intensities at its projections on the two images will differ, potentially by large amounts. Moreover, because these variations in intensities created by the immersion of the scene in water are non-linear both along each image and with respect to the unknown distances to the cameras, radiometric calibration techniques are not capable of compensating from them, especially in cases where the differences between water-column lengths from the scene to each of the cameras are large.

### 3.1. Stereo Using the Light Propagation Model

Removing attenuation and waterlight effects from underwater images of unknown scenes before stereo matching is an ill-posed problem, even if the parameters of the light propagation model ( $\alpha$ ,  $\beta$  and  $E(\infty)$ , for each color channel) are known. That’s because a fundamental piece of information necessary to compute the contribution of these effects to the images is missing: the depths of the scene points relative to the cameras. On the other hand, it is in principle necessary to remove the effects of these phenomena from the images in order for existing dense stereo techniques to be able to solve the correspondence problem.

The key observation that we use to overcome this difficulty is that, even though the light propagation model can not be applied *a priori* to a single image, it can be applied to any arbitrary pair of pixels in corresponding epipolar lines. That’s because the choice of one point in each epipolar line unambiguously defines the depths with respect to the two cameras. Thus, all that needs to be done to incorporate the light propagation model within any existing dense stereo technique that uses the epipolar constraint is to identify the points where the colors of candidate matches are compared and to perform a correction of these colors before the comparison, based on the positions of the pixels from which they came from, and on Equation (6). This idea leads directly to Algorithm 1.

The key feature of this algorithm is the fact that it tries to

---

#### Algorithm 1 General Underwater Stereo Algorithm

---

Calibrate the cameras and rectify the images geometrically and radiometrically;

Calibrate  $\alpha$ ,  $\beta$  and  $E(\infty)$ , for each image band;

**for** each point  $p_1 = (i, j)$  in image 1 (reference image):  
**do**

**for** each point  $p_2$  in row  $i$  of image 2: **do**

        Compute the water column lengths  $d_1$  and  $d_2$ , relative to the two cameras;

        Calculate  $E_{obj1}$  and  $E_{obj2}$ , by substituting, respectively,  $d_1$  and  $d_2$  in Equation (6);

        Compute the residual  $\varepsilon_{(d_1, d_2)} = |E_{obj1} - E_{obj2}|$ .

**end for**

**end for**

The optimal disparities for all pixels are then obtained by minimizing an energy functional composed of a term  $\varepsilon_{(d_1, d_2)}$  that measures how well each candidate pair of pixels  $p_1$  and  $p_2$  match, and potentially of other terms that reward desirable properties of the reconstruction such as spatial smoothness or preservation of discontinuities, and penalize undesirable results such as unmatched pixels.

---

remove the effects of interactions of light with water from the intensities of pairs of candidate matches before these intensities are compared. For pairs of pixels that are *not* actual matches, this compensation step will not do anything useful, because the water column lengths will be estimated incorrectly. But this does not matter, the important fact is that for pairs of pixels that are true matches, water column lengths will be correctly computed and by substituting them into Equation (6), the pixel colors will be corrected, generating a small residual  $\varepsilon_{(d_1, d_2)}$ .

### 3.2. Using Kolmogorov and Zabih’s Algorithm

The methodology described so far can be applied to modify essentially any existing dense stereo technique for use in underwater applications. In this paper, we chose to perform experiments with Kolmogorov and Zabih’s algorithm for energy minimization via graph cuts [9] due to its exceptional ability in dealing with occlusion.

Occlusions areas are those image regions that depict scene points that are not visible in the other image of the stereo pair. the algorithm proposed by Kolmogorov and Zabih [9] handles occlusions properly by using an energy functional that imposes smoothness constraints but at the same time tries to preserve discontinuities. More specifically, the algorithm’s energy functional is defined as a function of a disparity configuration,  $\mathcal{D}$ , which is a set of discrete values for the disparities of all pixels in an image:

$$\epsilon(\mathcal{D}) = \epsilon_{data}(\mathcal{D}) + \epsilon_{smoothness}(\mathcal{D}) + \epsilon_{visibility}(\mathcal{D}). \quad (7)$$

The definition above includes:

- a term  $\epsilon_{data}$  that captures the differences in intensity between pairs of pixels in the two images;
- a term  $\epsilon_{smoothness}$  that encourages smoothness, making neighbors pixels tend to have similar disparities;
- a term  $\epsilon_{visibility}$  that imposes a penalty for every pixel that is labelled as occluded.

This energy functional is then minimized efficiently with a maximum-flow algorithm developed specifically for the types of graphs that arise in energy-minimization applications in computer vision [2].

#### 4. Calibration of Model Parameters

The methodology described in the previous section assumes that the parameters of the light-propagation model – in particular, the values of  $\alpha$ ,  $\beta$  and  $E(\infty)$  for each chromatic band – have been estimated *a priori*. In this section we describe the procedure that we used to extract these parameters from actual underwater images.

The basic idea of our calibration procedure is to place objects with homogeneous colors at regularly-spaced, known distances inside the medium and then to measure average intensities within their visible projections in the images. In particular, to obtain the model parameters used in each experiment reported in Section 5, we placed a half-black, half-white planar target within a  $120 \times 60 \times 30$  cm fish tank and then followed the steps below:

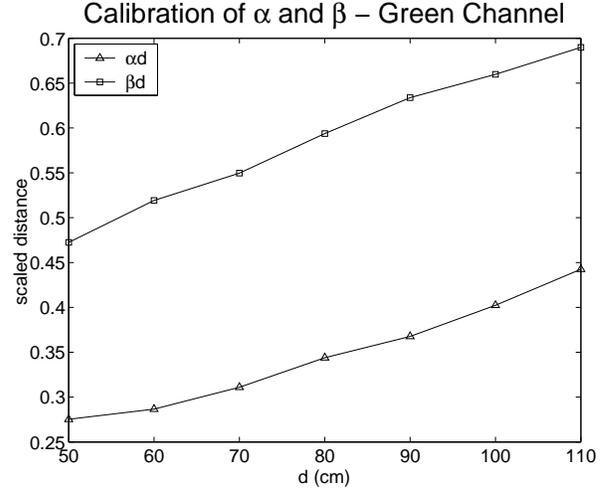
1. Take a picture of the target outside the water;
2. Place the target at a set of pre-defined distances within the water and take a picture of it at each distance;
3. Measure average radiance values for the black and white regions;
4. Use the light propagation model to compute initial values for  $\alpha$ ,  $\beta$  and  $E(\infty)$  from the measurements;
5. Refine the values computed in the previous step, using an optimization algorithm to minimize the differences between the out-of-water images and the underwater images corrected according to model.

##### 4.1. Initializing the Parameters

Initially, we measure the value of  $E(\infty)$  by marking pixels that correspond to the longest water column in the scene (about 120 cm), after which increasing the distance  $d$  practically does not affect the radiance collected by the camera.

Then, in order to obtain an initial value for  $\beta$ , we use the average intensity of the pixels in the target's black patch, with is regarded as having  $E_{obj} = 0$ . Substituting this value in Equation (5), we obtain

$$\beta d = -\ln \left( \frac{E(\infty) - E(d)}{E(\infty)} \right). \quad (8)$$



**Figure 3. Scaled distances  $\alpha d$  and  $\beta d$ , as a function of the known metric distances  $d$ .**

At each known value of  $d$ , we calculate the average value of the scaled distance  $\beta d$  using Equation (8). Now, we obtain an initial value for  $\beta$  simply by fitting a line to the points in the upper curve of Figure 3.

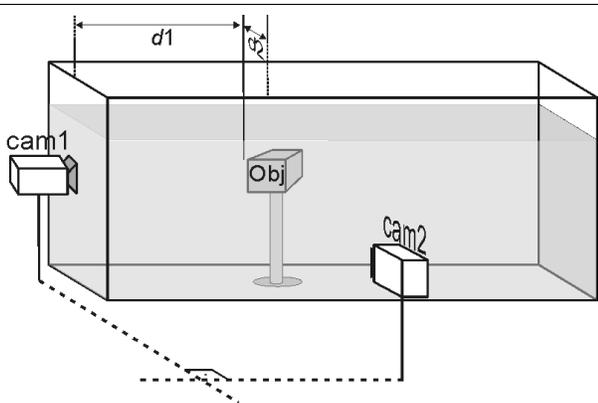
To estimate an initial value for  $\alpha$ , we use the average intensities of the white patch,  $E_{obj}$  and  $E(d)$ , respectively, in the pictures taken with no water in the fish tank and in the pictures taken with various values of the water column length,  $d$ , between the target and the camera. Substituting these values on Equation (5), we obtain

$$\alpha d = -\ln \left( \frac{E(d) - E(\infty)(1 - e^{-\beta d})}{E_{obj}} \right). \quad (9)$$

Similarly to what is done in the initialization of  $\beta$ , at each known distance  $d$  the scaled distance  $\alpha d$  is calculated using Equation 9. The initial value of  $\alpha$  is then obtained by fitting a line to the points in the lower curve of Figure 3.

##### 4.2. Refining the Model

Because the initial values for the parameters are calculated based on the simplifying assumption that the black patch reflects no light, it is necessary to refine these values, in order to obtain a reliable calibration of the light-propagation model. Such refinement is performed using Equation (6) in order to recover the value of  $E_{obj}$  from the underwater images. More specifically, the Residual Mean Squares (RMS) error between the restored image and an actual image of the same target obtained without water is measured. This error functional is fed to a non-linear optimization system that tries to minimize it by applying corrections to the model parameters. In our implementation of this step, we used the minimization function *fminunc* from MatLab



**Figure 4. Experimental setup: Cameras 1 (left) and 2 (right) were positioned so that the water column lengths between them and each object in the fish tank ( $d_1$  and  $d_2$ , respectively) were typically very different.**

6.1. Figure 1 displays the calibration target used, both before and after its underwater images are restored with the parameter values computed in the calibration.

## 5. Experimental Results

In order to verify if Algorithm 1 does yield more accurate disparity maps of underwater scenes than those obtained with dense stereo alone, we performed three experiments where multiple objects with very simple shapes were immersed in a  $120 \times 60 \times 30$  cm fish tank containing water with different levels of turbidity.

The experimental set-up is illustrated in Figure 4. Two cameras Sony DFW-X900 were positioned on different sides of the fish tank, with their optical axes oriented along perpendicular directions. More specifically, the cameras on the left and on the right were set to face maximum water column lengths of about 120 cm and 30 cm, respectively.

The two cameras were calibrated in unique global geometric and radiometric reference systems, with RMS geometric reprojection errors smaller than 0.5 pixels and RMS radiometric errors of 5 to 10 gray levels. Moreover, coordinates of the planes where each camera’s optical axis first intersects the fish tank (in the same global reference system) were estimated during the cameras’ geometric calibration, so that the length of the water columns between the cameras and the objects could be computed without the need to place the cameras in contact with the fish tank walls.

### 5.1. Scene in Clear Water (Low Turbidity)

In a first experiment, the fish tank was filled with clear water and three objects were immersed in it. A stereo pair of

images of this scene – rectified so that corresponding epipolar lines became identically-numbered scanlines – is shown in Figure 5. By comparing the two images, it can be observed that the three objects are at very different depths from the left camera (with water columns of about 54, 75 and 98 cm), and were at closer depths from the right camera (about 19, 13 and 4 cm underwater, respectively).

Figure 6 shows the disparity maps obtained by applying to these images (a) the unmodified stereo algorithm of Kolmogorov and Zabih [9], which we will call *Regular Stereo* from this point on, and (b) Algorithm 1 of this paper, which we will call *Underwater Stereo*. In these and all other disparity maps presented in this paper, 100% white pixels correspond to regions labelled as half-occluded (*i.e.*, visible in the left image, but occluded in the right one). Other graylevels correspond to various estimated disparities. More specifically, darker graylevels represent larger estimated depths with respect to the left camera.

Regular Stereo reconstructed the two objects closer to the left camera relatively well, because the water column lengths between each such object and the two cameras are not excessively different. However, it missed the deeper object completely, making it evident that when the underwater distances to the cameras are very distinct, using a light-propagation model to correct effects of attenuation and scattering is essential in order to get good reconstructions.

Underwater Stereo, on the other hand, reconstructed the deeper objects well, but failed to match the lighter stripes of the closer object. This happened because the radiometric calibration of that particular color was particularly noisy (which can be verified by comparing the two images in Figure 5). Since the corrections dictated by the light-propagation model resulted in an increase of the pixel intensities in this particular image area, the radiometric noise was amplified in the Underwater Stereo Algorithm, which could not find the correct correspondences. In future work, we intend to deal with this problem by weighting the residuals  $\varepsilon_{(d_1, d_2)} = |E_{obj1} - E_{obj2}|$  in a statistically optimal way before they get used in the last step of Algorithm 1.

### 5.2. Scene in Water + Milk (Medium Turbidity)

In a second experiment, 5 ml of milk were added to the about 200 liters of water in the fish tank, in order to increase the water turbidity. Three objects slightly different from those of the first experiment were placed at slightly different positions within the fish tank, so that their underwater distances with respect to the left and right cameras were, respectively,  $\{59, 80, 105\}$  cm and  $\{17, 11, 2\}$  cm. The rectified stereo pair of this scene is shown in Figure 7.

The disparity maps recovered by Regular and Underwater Stereo are shown in Figure 8. In this scene, because the radiometric calibration was more precise, both meth-

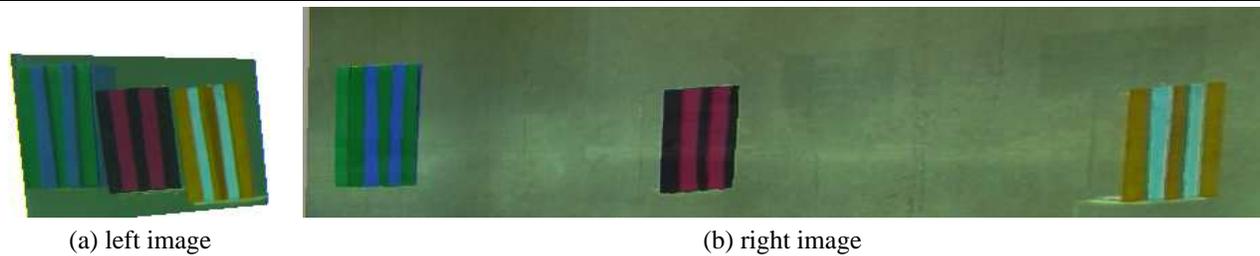


Figure 5. Rectified stereo pair used in the low-turbidity experiment.

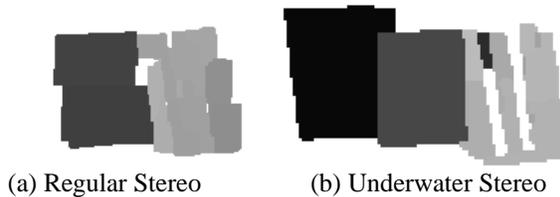


Figure 6. Disparity maps for pair of Figure 5.

ods managed to reconstruct all objects reasonably, but contrary to Underwater Stereo, Regular Stereo failed to identify matches for the two areas that appear as white patches in the central region of its depth map (Figure 8).

### 5.3. Scene in Water + Sand (High Turbidity)

Finally, we performed an experiment where about 150 g of sand were mixed in about 200 liters of clear water, in order to create a high-turbidity medium. The objects and their positions within the fish tank were almost identical to those of the previous experiment. In Figure 9, we show real images of this scene acquired by the left camera both with the fish tank empty and with it filled with water and sand.

Because of the high turbidity, even Underwater Stereo failed to identify the upper part of the deeper object, but it recovered its lower part reasonably, as well as most of the two objects closer to the left camera (Figure 10). Regular Stereo not only missed the deeper object completely: it missed large chunks of the middle object as well.

## 6. Conclusions

In this paper we introduced a methodology that uses a physically-based model of underwater light propagation within dense stereo algorithms in order to reconstruct scenes immersed in liquid media. In experiments with real underwater scenes, we have found that the three-dimensional reconstructions obtained with this methodology are in general more accurate than those ob-

tained by applying dense stereo alone to the same images, especially when the lengths of the water columns between the scene and the two cameras are very distinct and the medium has high turbidity.

We are currently working towards deploying the developed methodology in a real mechanism of fish transposition, located in the Grande River, MG/SP, Brazil, where we are doing experimental evaluation of various algorithms for fish tracking and classification.

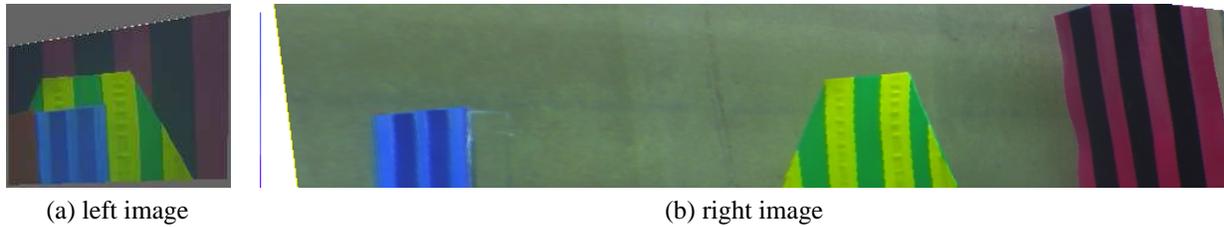
Unfortunately, the proposed methodology has a few drawbacks. It requires very careful radiometric calibration of both cameras and of the light-propagation medium, which may be hard to achieve outside a lab. As future steps, we are planning to ameliorate these difficulties through the development of radiometrically-uncalibrated underwater stereo techniques. We are also planning on doing a rigorous empirical comparison between our methodology and the underwater reconstruction algorithm based on light polarization that was presented by Schechner and Karpel [18] at the latest *IEEE CVPR* (after the date in which this paper was submitted for publication).

## Acknowledgments

The authors gratefully acknowledge the financial support of **CNPq** (Procs. 521259/2001-0, 140995/2002-1, 300592/2001-9, 132810/2003-4, 350750/94-7, 478859/2003-1), of **FAPEMIG** (Proc. CEX 491/02), and of **PRPq-UFMG** (Fundo Fundep RD).

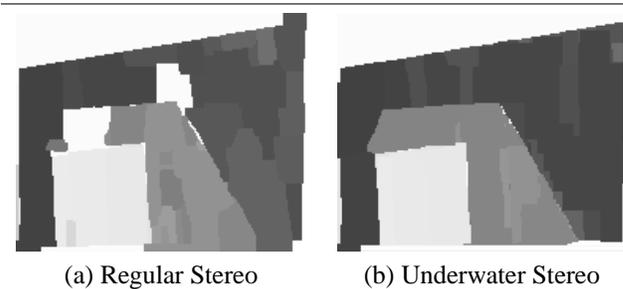
## References

- [1] M. Born. *Principles of optics*. Cambridge U. Press, 1999.
- [2] Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. In *Proc. Int. Workshop on Energy Min. Methods in Comp. Vis. and Pat. Recognition*, pages 359–374, 2001.
- [3] S. Chandrasekhar. *Radiative Transfer*. Dover Pub., 1960.
- [4] Q. Chen and G. Medioni. A volumetric stereo matching method: Application to image-based modeling. In *Proc. CVPR*, pages 29–34, 1999.

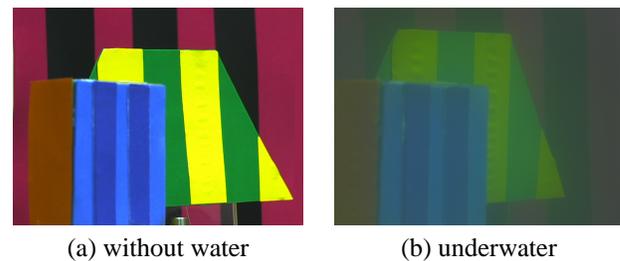


**Figure 7. Rectified stereo pair used in the medium-turbidity experiment.**

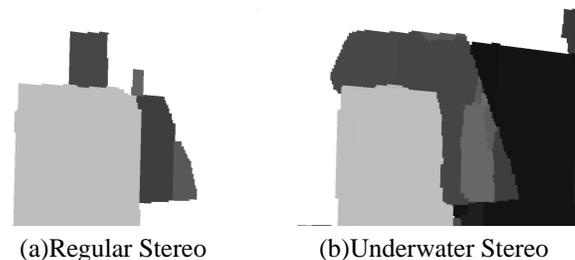
- [5] M. Chyba, N. E. Leonard, and E. D. Sontag. Optimality for underwater vehicles. In *Proc. IEEE Conference on Decision and Control*, pages 4204–4209, 2001.
- [6] A. Ishimaru. *Wave propagation and scattering in random media*. Wiley-IEEE Press, 1999.
- [7] F. A. Jenkins and H. E. White. *Fundamentals of optics*. McGraw-Hill, 1997.
- [8] H. W. Jensen, S. R. Marschner, M. Levoy, and P. Hanrahan. A practical model for subsurface light transport. In *Proc. SIGGRAPH*, pages 511–518, 2001.
- [9] V. Kolmogorov and R. Zabih. Multi-camera scene reconstruction via graph cuts. In *Proc. ECCV*, pages 82–96, 2002.
- [10] S. M. Luria and J. A. Kinney. Underwater vision. *Science*, 167:1454–1461, 1970.
- [11] S. G. Narasimhan and S. K. Nayar. Removing weather effects from monochrome images. In *Proc. CVPR*, volume 2, pages 186–193, 2001.
- [12] S. G. Narasimhan and S. K. Nayar. Vision and the atmosphere. *Int. J. Computer Vision*, 48(3):233–254, 2002.
- [13] S. K. Nayar and S. G. Narasimhan. Vision in bad weather. In *Proc. ICCV*, volume 2, pages 820–827, 1999.
- [14] S. Negahdaripour and C. Yu. *Underwater Robotic vehicles: Design and Control*. TSI Press, 1995.
- [15] M. Purcell, C. von Alt, B. Allen, T. Austin, N. Forrester, R. Goldsborough, and R. Stokey. New capabilities of the remus autonomous underwater vehicle. In *Proc. Oceans’ 2000*, pages 147–151, 2000.
- [16] S. Roy and I. J. Cox. A maximum-flow formulation of the N-camera stereo correspondence problem. In *Proc. ICCV*, pages 492–499, 1998.
- [17] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int. J. Comp. Vis.*, 47(1):7–42, 2002.
- [18] Y. Y. Schechner and N. Karpel. Clear underwater vision. In *Proc. CVPR*, volume 1, pages 536–543, 2004.
- [19] C. Silva and J. Santos-Victor. Intrinsic images for dense stereo matching with occlusions. In *Proc. ECCV*, volume 1, pages 100–114, 2000.
- [20] A. Trucco and V. Murino. Guest editors’ introduction. *Special Issue on Underwater Comp. Vis. and Pattern Recognition*, 79(1), July 2000.
- [21] C. Woosey and N. E. Leonard. Moving mass control for underwater vehicles. In *Proc. American Control Conference*, pages 2824–2829, 2002.



**Figure 8. Disparity maps for pair of Figure 7.**



**Figure 9. Actual images of the scene used in the high-turbidity experiment.**



**Figure 10. Disparity maps for image of Fig. 9.**

- [22] I. D. Y. Yitzhaky and N. S. Kopeika. Restoration of atmospheric blurred images according to weather-predicted atmospheric modulation transfer function. *Optical Engineering*, 36, 1998.