Enhancing Images in Scattering Media Utilizing Stereovision and Polarization

Amin Sarafraz
CAE Department
University of Miami
Coral Gables, FL 33124
a.sarafraz@umiami.edu

Shahriar Negahdaripour
ECE Department
University of Miami
Coral Gables, FL 33124
shahriar@miami.edu

Yoav Y. Schechner
Dept. of Electrical Eng.
Technion-Israel Inst. of Technology
Haifa 32000, Israel
yoav@ee.technion.ac.il

Abstract

Consider photography in scattering media. One goal is to enhance the images and compensate for scattering effects. A second goal is to estimate a distance map of the scene. A prior method exists to achieve these goals. It is based on acquiring two images from a fixed position, using a single camera mounted with a polarizer at different settings. However, the shortcomings of this polarization-based method comprise having to acquire these images sequentially, reduced light level, and inapplicability at low backscatter degree of polarization. In this paper, a new technique is described to alleviate these issues by integrating polarization and stereo cues. More precisely, the earlier single-camera method is extended to a pair of cameras displaced by a finite baseline. Each camera utilizes polarizers at different settings. Stereo disparity and polarization analysis are fused to construct de-scattered left and right views. The binocular stereo cues provide additional geometric constraints for distance computation. Moreover, the proposed technique acquires the two raw images simultaneously. Thus it can be applied to dynamic scenes. Underwater experiments are presented.

1. Introduction

Imaging within scattering media, e.g., underwater, has become increasingly important for a large number of scientific and commercial applications [8, 9]. In such media, poor visibility due to back-scatter hinders both human assisted operations and computer vision systems. Additional complexities arise when operating in low-light environments, and thus the use of artificial sources is necessary. Visibility degrades quickly with increasing distance between the camera and objects. On the other hand, reducing the backscatter calls for careful arrangement of the source and camera. Hence, it becomes highly desirable to develop methods that can enhance visibility and improve
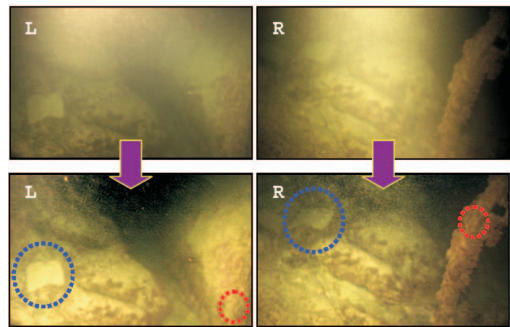

Figure 1. Polarization-based method for a single camera with two different polarization filter settings was applied to an uncalibrated stereo pair. Despite improvement, better results are expected with a model that accounts for the stereo baseline, thus utilizing both the stereo disparity cue and polarization analysis (courtesy of Tali Treibitz).

image quality in scattering media; e.g., [14].

Some of these issues are addressed by utilizing two images from a camera with a polarization filter, maintaining the same viewing position while varying the filter orientation [18, 21]. By applying a polarization-based mathematical model, these images are decoupled into backscatter and enhanced de-scattered components, while the scene distance map is recovered as an auxiliary product. The robustness of the method can be demonstrated by applying it to a stereo pair having a finite baseline; see Fig. 1. Yet, improved performance is expected by the generalization of polarization analysis to images acquired at multiple views.

We show that the integration of polarization analysis with stereovision leads to new methodology that provides several advantages over earlier methods. In the context of stereo vision several research has been done underwater irrespective of polarization [15, 6, 2, 16, 7, 20]. Those works, however, have focused on 3D reconstruction under reasonable visibility conditions, rather than enhancing the visibility itself. Furthermore, in [17], they account for light attenuation under uniform lighting while estimating the distance

map.

Here, we extend the earlier single-camera method [21] to a pair of cameras displaced by a finite baseline. Each camera utilizes a different polarization setting, and stereo disparity and polarization analysis are utilized to construct enhanced de-scattered views. The availability of binocular stereo cues to compute distance, enables the application of our method even when the images are obtained under low degrees of polarization. On the other hand, the use of polarization complements standard stereo, since it enables the removal of the backscatter (thus, improving stereo imagery), enhancement of backscatter modulation (by removing degeneracy), and the estimation of distances irrespective of the object texture or surface markings. Thus, our method can be applied even when only one of the two cues is dominant. A stereo setup can overcome temporal delays by acquiring images simultaneously, and thus can be used in dynamic scenes. While the proposed method is applicable to images acquired in a variety of scattering media, we present the results of experiments with underwater images acquired under controlled turbidity conditions.

2. Imaging Model

2.1. Notation

Let two calibrated cameras view the scene (Fig. 2). The origin of the world coordinate system is the projection center of the left camera, the Z-axis is along the optical axis and the XY axes are parallel to the horizontal and vertical scan lines. We assume a calibrated stereo system where the coordinate systems of the two cameras are parallel, and that the baseline vector is $\mathbf{D} = (D, 0, 0)$ in the global coordinate system. Hence, the epipolar lines are parallel to the x axis. Let $\mathbf{X} = (X, Y, Z)$ be the world coordinates of a point in the water. The projection of $\mathbf{X}$ on the image plane is $\mathbf{x} = (x, y)$. In particular, an object point at $\mathbf{X}_{\mathrm{obj}} = (X_{\mathrm{obj}}, Y_{\mathrm{obj}}, Z_{\mathrm{obj}})$ corresponds to an image point $\mathbf{x}_{\mathrm{obj}}$. Variables associated with the left or right camera are denoted by L or R, respectively. These include the image coordinates $\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{R}}$, and the coordinates corresponding to the same scene object $\mathbf{x}^{\mathrm{L}}_{\mathrm{obj}}, \mathbf{x}^{\mathrm{R}}_{\mathrm{obj}}$. The same distinction applies to the scattering angle θ and the backscatter degree of polarization (DOP) p, as defined in the earlier method proposed in [18, 21]. Briefly, when two images are taken with different polarization filter setting, one image is typically brighter than the other. The DOP $p \in [0, 1]$ is the ratio of the difference between the two images to the backscatter component.

We define the line of sight of the left and right views, LOS$^{\mathrm{L}}$ and LOS$^{\mathrm{R}}$, as follows:

Figure 2. Imaging setup.

$$
\begin{aligned}
\mathrm{LOS}^{\mathrm{L}} \equiv \Big\{ \mathbf{X}: \ & Z \in [0, Z_{\mathrm{obj}}], \\
& X = (Z/f) x^{\mathrm{L}}_{\mathrm{obj}}, Y = (Z/f) y^{\mathrm{L}}_{\mathrm{obj}} \Big\} \\
\mathrm{LOS}^{\mathrm{R}} \equiv \Big\{ \mathbf{X}: \ & Z \in [0, Z_{\mathrm{obj}}], \\
& X = D + (Z/f) x^{\mathrm{R}}_{\mathrm{obj}}, Y = (Z/f) y^{\mathrm{R}}_{\mathrm{obj}} \Big\}
\end{aligned}
$$
(1)

where f is the focal length of the cameras. If the two cameras are enclosed in dome ports with equal radius r, then the distances from the dome to $\mathbf{X}$ for the left and right cameras, respectively, would be

$$R^{\mathrm{L}}_{\mathrm{cam}}(\mathbf{X}) = \|\mathbf{X}\| - r, \qquad (2)$$

$$R^{\mathrm{R}}_{\mathrm{cam}}(\mathbf{X}) = \|\mathbf{X} - \mathbf{D}\| - r. \qquad (3)$$

2.2. Stereo and Polarization

With no polarizer, the measured image for each camera is given by

$$I(\mathbf{x}_{\mathrm{obj}}) = S(\mathbf{x}_{\mathrm{obj}}) + B(\mathbf{x}_{\mathrm{obj}}) \ , \qquad (4)$$

where $S(\mathbf{x}_{\mathrm{obj}})$ is the object signal and $B(\mathbf{x}_{\mathrm{obj}})$ is the backscatter [5, 11, 12]. Before detailing these components, note that backscatter is the major cause of contrast deterioration [5], rather than signal blur. This was demonstrated in [18] using objective criteria. Interestingly, according to Ref. [23], human vision associates image quality mostly with contrast, rather than resolution. For these reasons, we do not focus here on image blur and deblurring, although the small-angle forward scattering in high turbidity can introduce a significant amount of image blur [5, 11, 12]. Rather, we consider the prime effects associated with turbidity to be backscatter and attenuation.

Define $L_{\mathrm{obj}}(\mathbf{x}_{\mathrm{obj}})$ as the object radiance we would have sensed had no disturbances been caused by the medium along the LOS, and under uniform illumination. Propagation of light to the object and then to the camera via the medium yields an attenuated signal [5, 11]:

$$S(\mathbf{x}_{\mathrm{obj}}) = L_{\mathrm{obj}}(\mathbf{x}_{\mathrm{obj}}) F(\mathbf{x}_{\mathrm{obj}}) \ , \qquad (5)$$

where F is the falloff function, described below.

Consider for the moment a single illumination point source. From this source, light propagates a distance R_{src} to $\mathbf{X}_{\mathrm{obj}}$. Free space propagation creates a $1/R_{\mathrm{src}}^2$ irradiance falloff. Yet, there is turbidity, characterized by an attenuation coefficient c. Hence the falloff functions for the left and right images are

$$F^{\mathrm{L}}(\mathbf{x}_{\mathrm{obj}}^{\mathrm{L}}) = \left(\frac{\exp\{-c[R_{\mathrm{src}}(\mathbf{X}_{\mathrm{obj}})-r]\}}{R_{\mathrm{src}}^2(\mathbf{X}_{\mathrm{obj}})} Q(\mathbf{X}_{\mathrm{obj}}) \right) \exp(-c\|\mathbf{X}_{\mathrm{obj}}\|) , \tag{6}$$

$$F^{\mathrm{R}}(\mathbf{x}_{\mathrm{obj}}^{\mathrm{R}}) = \left(\frac{\exp\{-c[R_{\mathrm{src}}(\mathbf{X}_{\mathrm{obj}})-r]\}}{R_{\mathrm{src}}^2(\mathbf{X}_{\mathrm{obj}})} Q(\mathbf{X}_{\mathrm{obj}}) \right) \exp(-c\|\mathbf{X}_{\mathrm{obj}} - \mathbf{D}\|) , \tag{7}$$

respectively. In these equations $Q(\mathbf{X}_{\mathrm{obj}})$ expresses the non-uniformity of the scene irradiance, solely due to the inhomogeneity of the illumination. It similarly exists if the water is clear, i.e., $c = 0$, and can thus be pre-calibrated in clear water. For multiple illumination sources, Eqs. (6) and (7) are derived for each source, and then all F's are summed up. This can be generalized to include illumination due to multiple scattering [19].

Let $b(\theta)$ be the phase function. The backscatter for a small illumination source can be derived by integration along the LOS for each of the left and right images [5, 21]. Following Ref. [21], we assume that the degree of polarization p, as well $b(\theta)$ for all backscatter angles, are constant. Hence, we obtain

$$B^{\mathrm{L}}(\mathbf{x}_{\mathrm{obj}}^{\mathrm{L}}) = b\exp(cr)$$
$$\int_{R_{\mathrm{cam}}^{\mathrm{L}}=0}^{\|\mathbf{X}_{\mathrm{obj}}\|-r} I^{\mathrm{src}}(\mathbf{X}) \exp(-c\|\mathbf{X}\|)dR_{\mathrm{cam}}^{\mathrm{L}}, \quad \mathbf{X} \in \mathrm{LOS}^{\mathrm{L}} , \tag{8}$$

$$B^{\mathrm{R}}(\mathbf{x}_{\mathrm{obj}}^{\mathrm{R}}) = b\exp(cr)$$
$$\int_{R_{\mathrm{cam}}^{\mathrm{R}}=0}^{\|\mathbf{X}_{\mathrm{obj}}-\mathbf{D}\|-r} I^{\mathrm{src}}(\mathbf{X}) \exp(-c\|\mathbf{X}-\mathbf{D}\|)dR_{\mathrm{cam}}^{\mathrm{R}}, \quad \mathbf{X} \in \mathrm{LOS}^{\mathrm{R}}, \tag{9}$$

where $I^{\mathrm{src}}(\mathbf{X})$ is given by

$$I^{\mathrm{src}}(\mathbf{X}) = \frac{L^{\mathrm{src}}}{R_{\mathrm{src}}^2(\mathbf{X})} \exp\left(-cR_{\mathrm{src}}(\mathbf{X})\right) Q(\mathbf{X}). \tag{10}$$

Some variables are clearly independent of L and R. These include L^{src}, R_{src}, and $Q(\mathbf{X})$ (thus I^{src}), which are associated solely with the illumination. This also applies to $\mathbf{X}$ and $\mathbf{X}_{\mathrm{obj}}$, since they are defined in the global coordinate system.

Since $L_{\mathrm{obj}}(\mathbf{x}_{\mathrm{obj}})$ is defined as the object radiance under uniform illumination and in the absence of any disturbance from the medium along the LOS, we will adopt the brightness-constancy[1] assumption [4], as commonly done

---

[1]The simplified brightness constancy model can be readily relaxed with more general optical flow models, allowing us to generalize the analysis.

in stereo analysis. That is, L_{obj} is similar for L or R (at corresponding pixels). Hence,

$$S^{\mathrm{L}}(\mathbf{x}_{\mathrm{obj}}^{\mathrm{L}}) = L_{\mathrm{obj}}(\mathbf{x}_{\mathrm{obj}}^{\mathrm{L}})F^{\mathrm{L}}(\mathbf{x}_{\mathrm{obj}}^{\mathrm{L}}) \quad ,$$
$$S^{\mathrm{R}}(\mathbf{x}_{\mathrm{obj}}^{\mathrm{R}}) = L_{\mathrm{obj}}(\mathbf{x}_{\mathrm{obj}}^{\mathrm{R}})F^{\mathrm{R}}(\mathbf{x}_{\mathrm{obj}}^{\mathrm{R}}) . \tag{11}$$

In summary, substituting into Eq. (4) yields the final model of the stereo images:

$$I^{\mathrm{L}}(\mathbf{x}_{\mathrm{obj}}^{\mathrm{L}}) = S^{\mathrm{L}}(\mathbf{x}_{\mathrm{obj}}^{\mathrm{L}}) + B^{\mathrm{L}}(\mathbf{x}_{\mathrm{obj}}^{\mathrm{L}}) \tag{12}$$

$$I^{\mathrm{R}}(\mathbf{x}_{\mathrm{obj}}^{\mathrm{R}}) = S^{\mathrm{R}}(\mathbf{x}_{\mathrm{obj}}^{\mathrm{R}}) + B^{\mathrm{R}}(\mathbf{x}_{\mathrm{obj}}^{\mathrm{R}}) . \tag{13}$$

3. Efficient Polarization in a Stereo System

The methods in Refs [18, 21] were implemented by sequentially taking two frames, each with a different polarization setting of the imaging system. However, in a stereo system, such two images can be taken simultaneously. Hence, stereo complements this earlier approach by overcoming its temporal delays and its inaccuracy of distance estimation. Moreover, stereo can also be performed if $p = 0$ (zero DOP). On the other hand, the use of polarization complements standard stereo, since it decreases backscatter (improving the image to begin with) [3]. Polarization enhances the backscatter modulation (removing degeneracy), and thus enables the estimation of distances irrespective of the object's texture and surface markings. The pathological case, where the approach fails, corresponds to a texture-less object in a $p = 0$ medium. In Refs [18, 21], the polarizer modulates the backscatter, but not the signal. The latter simply divides its energy equally between the polarization components. Suppose that the L camera is set with a polarizer (termed analyzer), such that the backscatter is minimal, while the R camera is set with an orthogonal analyzer. Then, the raw images would be

$$I_{\mathrm{pol}}^{\mathrm{L}}(\mathbf{x}_{\mathrm{obj}}^{\mathrm{L}}) = \tfrac{1}{2}S^{\mathrm{L}}(\mathbf{x}_{\mathrm{obj}}^{\mathrm{L}}) + \tfrac{1-p}{2}B^{\mathrm{L}}(\mathbf{x}_{\mathrm{obj}}^{\mathrm{L}})$$
$$I_{\mathrm{pol}}^{\mathrm{R}}(\mathbf{x}_{\mathrm{obj}}^{\mathrm{R}}) = \tfrac{1}{2}S^{\mathrm{R}}(\mathbf{x}_{\mathrm{obj}}^{\mathrm{R}}) + \tfrac{1+p}{2}B^{\mathrm{R}}(\mathbf{x}_{\mathrm{obj}}^{\mathrm{R}}) \tag{14}$$

since p is insensitive to the backscatter ray direction. Note that $I_{\mathrm{pol}}^{\mathrm{R}}$ is worse than I^{R} defined in Eq. (13): the signal in $I_{\mathrm{pol}}^{\mathrm{R}}$ is attenuated by half, while the backscatter is amplified relative to the signal by $(1 + p) > 1$. Thus, we may utilize an analyzer only on the L camera, yielding $I_{\mathrm{pol}}^{\mathrm{L}}$ according to Eq. (14). The right camera has no polarizer, hence it simply outputs I^{R} as defined in Eq. (13). This increases the light throughput, and makes the formulation applicable even if p may change from the L view to the R view (since polarization has no impact on the R image). Anyway, there are essentially two unknowns per pixel $\mathbf{x}^{\mathrm{L}}$: the object distance Z_{obj} and the object radiance under uniform lighting L_{obj}. To estimate them, we have at our disposal two independent raw images $I_{\mathrm{pol,raw}}^{\mathrm{L}}$ and $I_{\mathrm{raw}}^{\mathrm{R}}$. The unknowns should be consistent with these images.

4. Implementation

The analytical framework in the last sections is the foundation of an algorithm for reconstructing de-scattered images by decoupling the backscattering fields of the left and right stereo pairs. Both c and b may be measured in a-priori experiments or by utilizing standard equipment. In addition, we make use of the known source illumination field, determined through calibration with the images of a uniform planar target.

The operations are carried out at each pixel $\mathbf{x}_{obj}^L = \mathbf{x}^L$. With calibrated cameras, the coordinate $\mathbf{x}^L$ uniquely defines LOS^L, up to an unknown distance Z_{obj}. This unknown could be estimated using correspondences. The estimated $\hat{Z}_{obj}$ and $\mathbf{x}^L$ uniquely define

$$\begin{aligned}
\hat{\mathbf{X}}_{obj} = \hat{\mathbf{X}}_{obj}(\hat{Z}_{obj} \mid \mathbf{x}^L) &= [\hat{X}_{obj}, \hat{Y}_{obj}, \hat{Z}_{obj}] \\
&= \left[(\hat{Z}_{obj}/f)x^L, (\hat{Z}_{obj}/f)y^L, \hat{Z}_{obj} \right] \;,
\end{aligned} \tag{15}$$

and the correspondence

$$\hat{\mathbf{x}}^R = \hat{\mathbf{x}}^R(\hat{Z}_{obj} \mid \mathbf{x}^L) = \left[(\hat{X}_{obj} - D)f/\hat{Z}_{obj}, \hat{Y}_{obj}f/\hat{Z}_{obj} \right], \tag{16}$$

defines LOS^R according to Eq. (1). Then, utilizing the measured values of c and b, Eqs. (6)-(9) directly yield the *modeled* falloff and backscatter fields. Recall that for each pixel $\mathbf{x}^L$, these fields depend only on a *single* unknown degree of freedom, the distance $\hat{Z}_{obj}$, thus denoted $F_{model}^L(\hat{Z}_{obj} \mid \mathbf{x}^L)$, $F_{model}^R(\hat{Z}_{obj} \mid \mathbf{x}^L)$, $B_{model}^L(\hat{Z}_{obj} \mid \mathbf{x}^L)$, and $B_{model}^R(\hat{Z}_{obj} \mid \mathbf{x}^L)$. With two equations for each pair of corresponding points in the stereo pair, the object reflectance can be estimated. Having the estimated backscatter field for the left and right images, the left and right signals can be computed by manipulating Eqs. (14). The entire process can be repeated, by feeding back the de-scattered stereo images to the disparity estimation module, for computing a better range map. The block diagram in Fig. 3 depicts the iterative process.

5. Experiments and Results

Experiments were carried out with stereo pairs acquired in a 6'(W)× 12'(L)× 6'(H) indoor water tank, in different turbidity levels, but the same positions relative to the scene. The water turbidity and scattering level were varied by adding a known volume of low-fat milk to clear water, namely quarter, half and one liter. Hereafter, we call experiments with quarter, half and one liter contamination of milk as low, medium and high concentrations, respectively.

The stereo baseline[2] is about 17.2 [cm]. The light source position is (8.5,9,0.0) [cm] (located between the two cameras). The backscatter decreases by increasing the baseline

---

[2]Recall that the left-camera projection center is chosen as the origin of the world coordinate system.
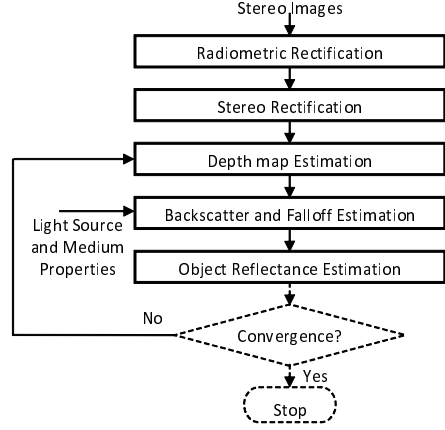

Figure 3. Block-diagram of de-scattering algorithm.

between each camera and the source. However we deliberately used a small camera-source distance to challenge our method.

To estimate the illumination field, first we acquired a high-dynamic-range image of the field over a narrow spatial band to determine the source pattern for one cross section. Then we assumed the same distribution for other cross sections at different orientations. Measurement of inherent optical properties of ocean waters, including the attenuation coefficient can be made with off-the-shelf spectrometers. Furthermore, ocean optics models of these coefficients exist as functions of environmental conditions, e.g., size and concentration of scatterers that can be readily estimated.

For our experiments, we determined the attenuation coefficient a priori as a linear function of the milk concentration, using images at different source-to-camera propagation paths [13]. This linear model was applied at various milk concentrations. Also, the scattering coefficient b was estimated from the published data in [13]; the backscatter coefficient for low-fat milk is approximately the same as the attenuation coefficient.

There are three main possible states for combining a camera and a polarizer. The polarizer could be aligned such that it removes maximum amount of backscatter. Therefore, the resulting image has minimum amount of backscatter. Conversely, the polarizer could be aligned in such a way that the resulting image has maximum backscatter. Finally, the camera could have no polarizer. In a stereo setup, both left and right cameras be in any one of these states. Among all possible combinations, we refer to *minimum scattering setting* when polarizers on both cameras remove maximum amount of backscatter and consequently the raw images then have minimum amount of backscatter. Accordingly, *maximum scattering setting*

(a)   (b)

Figure 4. Raw images under the *maximum scattering setting* with (a) low and (b) medium concentration of low fat milk. The bottom row shows the magnified cylindrical object. Its enhanced result is shown at the right side at each image.

refers to the setting with maximum amount of backscatter in raw images and *unpolarized setting* refers to the setting with no polarizer on both cameras.

In the first data set, the background is the tank wall at one end at about 2 [m] distance from the stereo system, covered with black Neoprene sheet. Images were recorded at low and medium concentrations, corresponding to average attenuation coefficients of roughly $\{0.37\}$ [m^{-1}] and $\{0.74\}$ [m^{-1}], respectively. The targets comprise two planar objects at 0.8 [m] and 1.2 [m], and a cylindrical object at 1.5 [m] from the stereo cameras.

Fig. 4 depicts stereo pairs of the same scene recorded under the *maximum scattering setting*, at low and medium concentrations of low fat milk. The cylindrical object is magnified in Fig. 4. It is shown before and after removing backscatter in both left and right images, at different concentrations. In medium concentration, the object is almost veiled and the enhanced images show a great deal of improvement. The recovered signal in medium concentration is darker than the recovered signal in low concentration, due to stronger attenuation. Fig. 5 contrasts the recovered signal and raw images.

For quantitative assessment, we compare the ability of the SIFT point detector [10] to find corresponding points in raw and enhanced stereo pairs using identical parameter settings. After determining corresponding points using SIFT, a denser range map was obtained by a match propagation technique [1] and interpolation. A ground-truth range map was estimated using images taken by the same setup in clear water.

Fig. 6 compares the initial and final estimated range maps that are obtained from the raw data and the descattered images for the first data set. In all cases, the initial seeds obtained from SIFT are depicted by black circles. It is noted that the choice of our range estimation method

is mainly for demonstration purposes, and in general, one can apply other algorithms. However, match propagation is based on color (intensity) constancy assumption. Thus the denseness of the estimated map is strongly coupled to the similarity of the left and right views in the raw and descattered images. Only a small number of initial matches were used to explore the effectiveness of match propagation. As can be seen, in both concentrations, the final estimated range map is closer to the ground truth than the initial range map. Also, notice that at medium concentration, SIFT was not able to find any match on farther objects (Welcome mat and cylindrical object). In contrast, in the recovered signal, several corresponding points were detected on these objects.

Fig. 7 compares the results of the proposed method (stereo method) with the single camera method [21] at different milk concentrations. For comparison purposes, we illustrate the results of stereo method where the left polarizer is set to minimize the backscatter and the right polarizer is set to maximize the backscatter. In all images in Fig. 7, the results were superimposed on image with minimum backscatter. As can be seen, in all cases, the results show improvement over the image with minimum backscatter. Since both methods use the same physical basis for backscatter modeling, the results of two methods are consistent with each other.

In the second data set, the tank wall is at about 2.5 [m] distance from the stereo system. The images were obtained medium and high concentrations. Furthermore, we acquired images in clear water to serve as the benchmark to compare with our results. For lack of space, we show and compare the left views of the processed data. Also, we depict an estimation of L_{obj}, instead of the estimated signal.

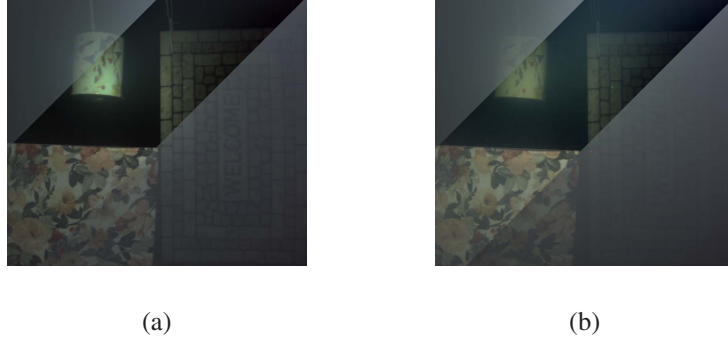In this experiment, we explore a different combinations of a stereo vision system with polarizers. Fig. 8

(a)   (b)

Figure 5. Comparison between the raw images and the recovered signals, (a) low and (b) medium concentration.
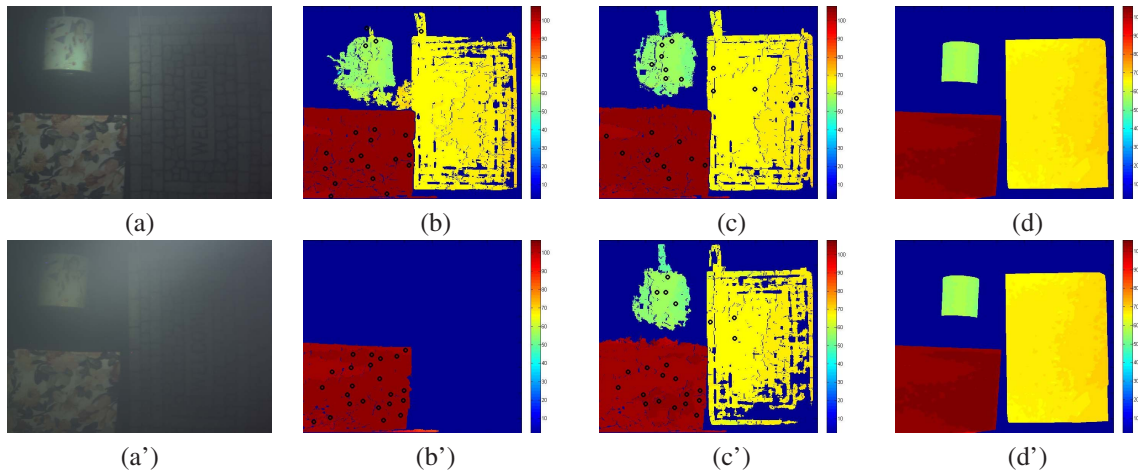

(a)   (b)   (c)   (d)

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

Figure 6. Comparison of range map estimation at different concentrations. First row: low concentration results (a) left view of raw image, (b) initial range map using raw images, (c) final range map using final enhanced images, (d) ground-truth range map. Second row: medium concentration results (a') left view of raw image, (b') initial range map using raw images, (c') final range map using final enhanced images, (d') ground-truth range map.

illustrates the experiment with medium milk concentration. First, the raw images in (a,a') are recorded under *minimum scattering setting*. In (b,b'), the polarizer on left camera is set to minimize the backscatter, while the right camera utilizes no polarizer. The images in (c,c') were taken without any polarizer. The single camera method cannot be applied in any of these cases, since either both cameras have the same polarizer settings, or one camera employs no polarizer. In (d), we show the left image in clear water, used here as the ground truth. The images in (e-g) were obtained by applying our method to the data in (a-c), respectively.

We obtain the best results in the case where both the polarization and stereo cues are significant. As we remove the polarizer from one or both cameras, the contrast of the reconstructed reflectance map deteriorates. With no polarizer, the stereo disparity is the sole source of range cue in order

to estimate the backscatter field. Note that by computing the reflectance map, we also compensate for non-uniform illumination; compare the processed data and the ground truth image taken in clear water.

Fig. 9 shows the results at high milk concentration. The reconstructed image is considerably enhanced over the banner, where the SIFT method with disparity propagation produces a dense range map. However, SIFT does not give any seed match over the circular object, because the signal is too weak. Thus, we have applied in this region the estimated range from the less turbid data. This allows us to establish the upper bound on the performance in estimating the reflectance map of the circular object.
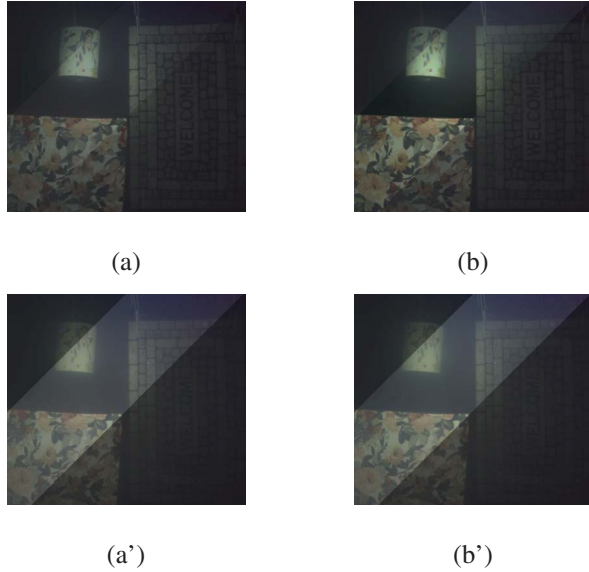
(a)　　　　　　　　　　　　(b)

(a')　　　　　　　　　　　　(b')

Figure 7. Comparison of stereo method with single camera method [21], results of (a) single camera method at low concentration, (b) stereo camera method at low concentration, (a') single camera method at medium concentration, (b') stereo camera method at medium concentration

(a)　　　　　　　　　　　　(a')

(b)　　　　　　　　　　　　(b')

(c)　　　　　　　　　　　　(c')
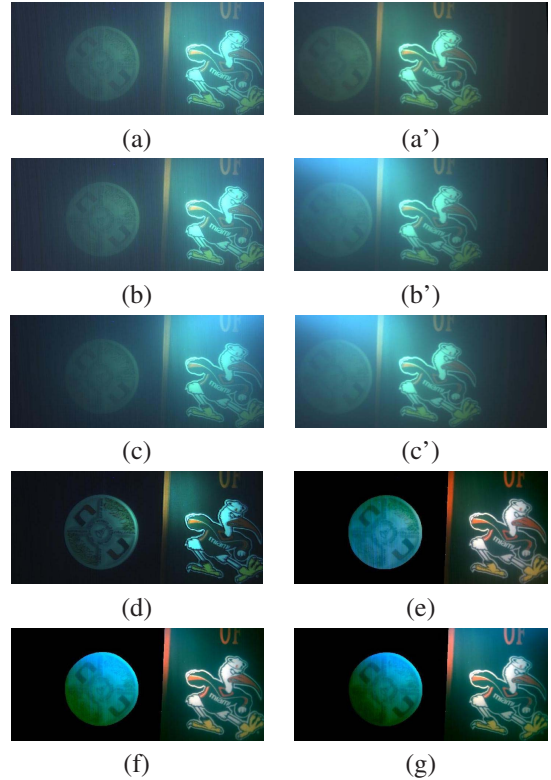
(d)　　　　　　　　　　　　(e)

(f)　　　　　　　　　　　　(g)

Figure 8. Results of data with medium concentration. Data of raw left and right stereo views (a,a') with polarizers on both cameras, whose orientation minimizes the backscatter, (b,b') polarizer filter only on left camera, (c,c') neither camera utilizing a polarizer. The left image in clear water (d) is compared with the reflectance maps estimated by our method. The images in (e-g) are de-scattered results that correspond to the data in (a-c), respectively.

6. Conclusions

We generalized an earlier polarization-based method to estimate and eliminate backscattering from underwater images. Instead of acquiring two images with different polarization settings on a single camera, we work with stereo cameras having suitable polarization filters. We considered different combinations of polarizers with stereo cameras. In all cases, the active light source employs a polarizer. The best results were achieved when the two cameras had polarizers aligned to yield images with minimum amount of backscatter. This may be related to conclusions made in ref. [22].

Among several advantages, the binocular stereo cue provides additional geometric constraint, and can thus work in the absence of a high DOP. Another advantage of our method is that the two images can be acquired simultaneously, avoiding temporal delays.

In the absence of both polarization and binocular cues, namely low DOP with a textureless object/scene, our method becomes ill-posed. Two sets of experiments each with two different turbidity levels verify that the proposed mathematical models lead to improved image quality. However, to achieve the best results, the system components, including the active illumination source, have to be calibrated. Generalization to deal with multiple scattering and nonuniform back-scatter are worth future explorations.

References

[1] Q. Chen and G. Medioni. A volumetric stereo matching method: application to image-based modeling. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 29–34, 1999.

[2] G. Dudek, P. Giguere, C. Prahacs, S. Saunderson, J. Sattar, L.-A. Torres-Mendez, M. Jenkin, A. German, A. Hogue, A. Ripsman, J. Zacher, E. Milios, H. Liu, P. Zhang,
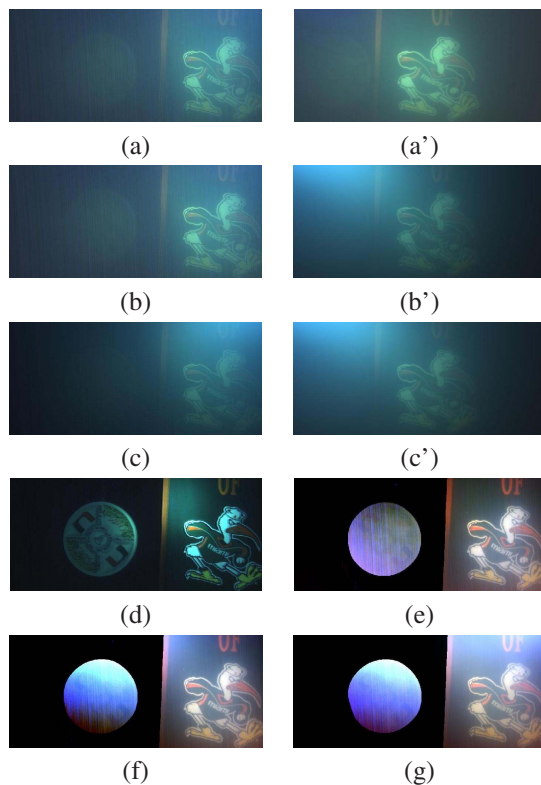
Figure 9. Results of data with high concentration. Data of raw left and right stereo views (a,a') with polarizers on both cameras, whose orientation minimizes the backscatter, (b,b') polarizer filter only on left camera, (c,c') neither camera utilizing a polarizer. The left image in clear water (d) is compared with the reflectance maps estimated by our method. The images in (e-g) are de-scattered results that correspond to the data in (a-c), respectively.

M. Buehler, and C. Georgiades. Aqua: An amphibious autonomous robot. *Computer*, 40(1):46–53, Jan. 2007.

[3] G. D. Gilbert and J. C. Pernicka. Improvement of underwater visibility by reduction of backscatter with a circular polarization technique. *Applied Optics*, 6(4):741–746, 1967.

[4] B. Horn and B. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

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

[6] M. Johnson-Roberson, S. Kumar, O. Pizarro, and S. Willams. Stereoscopic imaging for coral segmentation and classification. In *MTS/IEEE OCEANS conference*, pages 1–6, Sept. 2006.

[7] A. Khamene, H. Madjdi, and S. Negahdaripour. 3-d mapping of sea floor scenes by stereo imaging. In *MTS/IEEE OCEANS Conference*, volume 4, pages 2577–2583, 2001.

[8] D. M. Kocak and F. M. Caimi. The current art of underwater imaging with a glimpse of the past and vision of the future. *Marine Technology Society Journal*, 39(3), Oct. 2005.

[9] D. M. Kocak, F. R. Dalgleish, F. M. Caimi, and Y. Y. Schechner. A focus on recent developments and trends in underwater imaging. *Marine Technology Society Journal, Special issue on State of the Technology*, 42(1):52–67, 2008.

[10] D. Lowe. Object recognition from local scale-invariant features. *IEEE International Conference on Computer Vision*, pages 1150–1157, 1999.

[11] B. L. McGlamery. A computer model for underwater camera system. In *Proc. SPIE*, pages 221–231, 1979.

[12] C. D. Mobley. Light and water: Radiative transfer in natural waters. *Academic Press, San-Diego*, Ch. 3,5, 1994.

[13] S. G. Narasimhan, M. Gupta, C. Donner, R. Ramamoorthi, S. K. Nayar, and H. W. Jensen. Acquiring scattering properties of participating media by dilution. *ACM Trans. on Graphics*, 25(3):1003–1012, 2006.

[14] S. G. Narasimhan, S. K. Nayar, B. Sun, and S. J. Koppal. Structured light in scattering media. In *Proc. ICCV05, Beijing, China*, pages 420–427, 2005.

[15] S. Negahdaripour and P. Firoozfam. An rov stereovision system for ship hull inspection. *IEEE Journal of Oceanic Engineering*, 31(3):551–564, 2006.

[16] S. Negahdaripour and H. Madjidi. Stereovision imaging on submersible platforms for 3-d mapping of benthic habitats and sea-floor structures. *Oceanic Engineering, IEEE Journal of*, 28(4):625–650, Oct. 2003.

[17] J. P. Queiroz-Neto, R. Carceroni, W. Barros, and M. Campos. Underwater stereo. *Computer Graphics and Image Processing, Brazilian Symposium on*, 0:170–177, 2004.

[18] Y. Y. Schechner and N. Karpel. Recovery of underwater visibility and structure by polarization analysis. *IEEE Journal of Oceanic Engineering*, 30(3):570–587, 2005.

[19] B. Sun, R. Ramamoorthi, S. G. Narasimhan, and S. K. Nayar. A practical analytic single scattering model for real time rendering. *ACM Trans. on Graphics*, 24(3):1040–1049, 2005.

[20] Y. Swirski, Y. Y. Schechner, B. Herzberg, and S. Negahdaripour. Stereo from flickering caustics. In *IEEE International Conference on Computer Vision*, 2009.

[21] T. Treibitz and Y. Y. Schechner. Instant *3D*escatter. In *Proc. IEEE Computer Soc. Conf. on Computer Vision and Pattern Recognition*, pages 1861–1868, 2006.

[22] T. Treibitz and Y. Y. Schechner. Polarization: Beneficial for visibility enhancement? In *Proc. IEEE Computer Soc. Conf. on Computer Vision and Pattern Recognition*, 2009.

[23] B. Wells. MTF provides an image-quality metric. *Laser Focus World*, 41(10), 2005.