

# Structured Light in Scattering Media\*

Srinivasa G. Narasimhan<sup>+</sup>, Shree K. Nayar<sup>++</sup>, Bo Sun<sup>++</sup> and Sanjeev J. Koppal<sup>+</sup>

<sup>+</sup>Robotics Institute, Carnegie Mellon University, Pittsburgh, USA

<sup>++</sup>Computer Science Department, Columbia University, New York, USA

## Abstract

Virtually all structured light methods assume that the scene and the sources are immersed in pure air and that light is neither scattered nor absorbed. Recently, however, structured lighting has found growing application in underwater and aerial imaging, where scattering effects cannot be ignored. In this paper, we present a comprehensive analysis of two representative methods - light stripe range scanning and photometric stereo - in the presence of scattering. For both methods, we derive physical models for the appearances of a surface immersed in a scattering medium. Based on these models, we present results on (a) the condition for object detectability in light striping and (b) the number of sources required for photometric stereo. In both cases, we demonstrate that while traditional methods fail when scattering is significant, our methods accurately recover the scene (depths, normals, albedos) as well as the properties of the medium. These results are in turn used to restore the appearances of scenes as if they were captured in clear air. Although we have focused on light striping and photometric stereo, our approach can also be extended to other methods such as grid coding, gated and active polarization imaging.

## 1 Introduction

Structured lighting has received wide attention since the early work of Will and Pennington [28]. Virtually all structured lighting techniques modulate the appearance of a surface (or volume) by projecting a particular pattern of light onto it [15]. Often this modulation makes subsequent processing of acquired images simple; the correspondence problem and issues related to calibration are often alleviated [6; 13; 30] and accurate 3D reconstruction is obtained irrespective of the complex surface texture [20; 21]. As a result, structured lighting has been a key enabling technology for several industrial applications such as assembly, alignment and inspection.

An implicit assumption made in most structured light methods is that light is neither scattered nor absorbed by the medium in which the scene and sources are immersed (as in pure air). This assumption, however, is violated in several media including atmospheric conditions (fog, haze, mist), fluids (murky water, milk) and smoke. An optical image taken in such a medium suffers from significant loss of contrast and attenuation of brightness. Further, this degradation increases exponentially with distance, making it hard to capture meaningful images of scenes that are not near the imaging system. Thus, it is critical to

take into account the effects of scattering while applying structured light methods in such media.

But why use structured light in scattering media at all? Consider, for example, underwater optical imaging. Because of scattering by impurities in water, natural sources such as daylight attenuate completely before reaching significant depths. So, deep underwater (semi-)autonomous vehicles rely on active illumination to explore wreckages<sup>1</sup>, find mines, and inspect vessels, docks and rigs, thus playing a key role in maintenance, construction and rescue scenarios. A variety of structured lighting techniques have been developed for these applications, ranging from using thin laser beams [14], to using a number of carefully spaced confocal sources [17], to more sophisticated time-gated [19] and synchronization-gated techniques [7]. While these methods enhance visibility (resulting in better detection of targets), they do not explicitly analyze the appearances of scenes under structured light<sup>2</sup>.

In this paper, we are interested in both enhancing visibility using structured light and also in analyzing the acquired images to recover properties of the scene and the medium. To achieve this, three relevant questions must be addressed. First, what are the scattering effects that result from the interaction of structured light with the medium and the scene? Second, how do we overcome these scattering effects to obtain the results that the structured light methods were traditionally designed for?<sup>3</sup> Third, is there additional information that one can extract from these scattering effects that is not possible to obtain using the traditional methods?

We address these questions specifically for two representative techniques - light stripe range scanning and photometric stereo. For each of these, we derive an analytic image formation model that is based on the physics of single scattering. These models describe the interactions of structured light with the medium and the scene. Using the image formation model for light striping, we develop a simple algorithm to reliably detect objects and obtain a 3D reconstruction of the scene in the presence of strong scattering. Based on the image formation model for photometric stereo, we conclude that at least five light source directions (instead of the usual three) are required to reconstruct surface normals and albedos of a lambertian

<sup>1</sup>Famously, the 1985 Franco-American expedition discovered and explored the remains of the *Titanic* that sank in 1912 [1].

<sup>2</sup>Work has also been done on a related but different problem of analyzing the appearances of scenes in scattering media (underwater or the atmosphere) using passive methods [5; 24; 26; 23; 18; 22] that rely on natural illumination external to the medium.

<sup>3</sup>Note that works that address this question (for instance, laser scanning to reconstruct sea floors [3; 16]) by ignoring scattering effects (first question) are severely limited in their applicability.

\*This work is supported by an ONR contract #N00014-05-1-0188. The authors also thank Estuardo Rodas for building the tank used for the experiments in this paper.

object. Interestingly, our method also yields a depth map of the scene, which is not possible using traditional photometric stereo. Further, in both techniques, the interaction of structured light with the medium allows us to estimate the properties of the medium. This result can in turn be used to remove the effects of scattering and compute the appearance of the scene as if seen in clear air.

To verify our methods using real experiments, we have constructed a setup that consists of a glass tank filled with a scattering medium (dilute milk), with a projector (source) and camera placed outside the tank. Note that calibration of this setup requires us to handle light refraction at the medium-glass-air interfaces. We present a calibration procedure that is similar in spirit to [11] and that does not require either explicit geometric calibration of the camera and the projector or the knowledge of refraction locations or refractive indices of media. Although we have focused on light striping and photometric stereo, our results can be used to extend several other techniques such as grid coding [28] and gated [7] and active polarization imaging [25; 10]. We believe that our results can significantly benefit a wide range of underwater [14], aerial and microscopic imaging [9] applications.

## 2 Single Scattering in Media

In order to keep our techniques tractable, we assume that the scattering medium is homogeneous and not highly dense (for example, murky water, light fog, mist, dilute milk). This allows us to develop simple models based on single scattering. We now define the properties of scattering media [4] and present the single scattering model.

The scattering coefficient  $\beta$  is defined as the fraction of the incident flux scattered by a unit volume of the medium in all directions<sup>4</sup>. The phase function  $\mathcal{P}(\alpha)$  defines the angular scattering distribution of the incident flux, where  $\alpha$  is the angle between incident and scattered directions. In general, the phase function is smooth and may be represented by a low-order polynomial of  $\cos \alpha$  [4]. We use the first-order approximation as given in [4],

$$\mathcal{P}(g, \alpha) = (1/4\pi) (1 + g \cos \alpha), \quad (1)$$

where,  $g \in (-1, 1)$  is a parameter that controls the shape of the phase function. Now consider a light ray with radiance  $L_0$  that travels a distance  $x$ , gets scattered by a particle at an angle  $\alpha$ , before it further travels a distance  $y$  to reach the viewer. The intensity of this light ray is attenuated exponentially according to the total distance traveled. Then, the single scattering irradiance at the viewer is given by [4],

$$E_{med} = L_0 \beta \mathcal{P}(g, \alpha) e^{-\beta(x+y)}. \quad (2)$$

For an isotropic point source with radiant intensity  $I_0$ , we may further write  $L_0 = I_0/x^2$ , while for a collimated beam,  $L_0$  is constant with respect to  $x$ . We build upon equations 1 and 2 to derive image formation models for light striping and photometric stereo.

<sup>4</sup>When absorption is present, the scattering coefficient is replaced by the extinction coefficient defined as the fraction of incident flux scattered and absorbed by a unit volume.

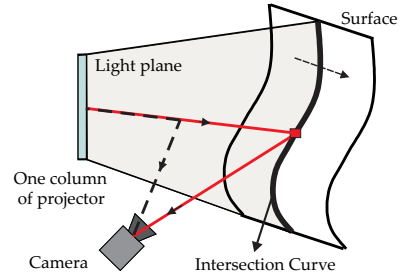


Figure 1: **Light striping in media.** A sheet of light swept across a surface produces a brightness discontinuity (black curve). When there is no scattering, the pixel brightness is only due to this discontinuity (red ray). In the presence of scattering, the light plane itself becomes visible (dashed ray) making surface detection hard.

## 3 Light Striping in Scattering Media

Light stripe range scanning is a technique where a plane (or sheet) of light is swept across an object (figure 1) to obtain its 3D reconstruction. The key observation is that the plane of light intersects the object surface at a curve, producing a large brightness discontinuity. Then, the 3D coordinates of each point on this curve is computed by intersecting the camera ray and the light plane. A critical requirement here is that the intersection curve be detected reliably, which is usually done by thresholding the acquired image. Unfortunately, in the presence of scattering, the entire light plane becomes visible and detecting this intersection is not possible by simple thresholding. In this section, we derive the model for image formation when the light plane and the surface are immersed in a scattering medium and develop algorithms for reliable scene detection, and 3D reconstruction and for obtaining a *clear-air* appearance of the scene.

### 3.1 Image Formation Model

Imagine a light plane sweeping across a surface in a scattering medium. The camera not only receives light reflected by the surface, but also from the medium after scattering (see figure 2). The dashed lines indicate light rays that reach the camera after attenuation and scattering in the medium, but without reaching the surface. Then, the irradiance  $E_{med}$  at the camera is exactly given by equation 2. The red line indicates the path traveled by a light ray from the source to the surface and then reflected by the surface toward the camera. The intensity of this ray is exponentially attenuated according to the total distance traveled. Hence, the irradiance  $E_{surf}$  at the camera due to this ray is written as<sup>5</sup>,

$$E_{surf} = L_0 e^{-\beta(d_s+d_v)} R, \quad (3)$$

where,  $R$  is the radiance (normalized by source intensity) in the absence of scattering. Thus, the image formation model may be compactly written using the Dirac delta function  $\delta$  as,

$$E = E_{surf} \delta(x = d_s) + E_{med} \delta(x < d_s). \quad (4)$$

<sup>5</sup>Single scattering of the exponentially attenuated surface radiance towards the camera is a minor effect compared to the scattering  $E_{med}$  from bright sources, and hence can be safely ignored.

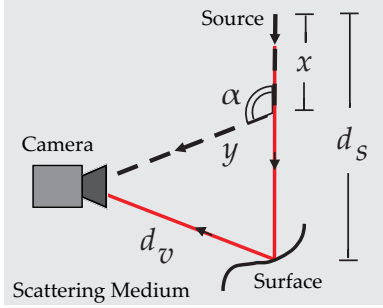


Figure 2: **Image formation in light striping.** The irradiance at the camera is produced by either the light rays that reach the camera after being scattered once by the medium (dashed) or by light rays that are reflected by the surface (solid red). In both cases, the intensities of the rays are attenuated exponentially according to the distance traveled.

### 3.2 Intersection of Surface and Light Plane

Figure 3 illustrates the profile of the camera irradiance  $E$  as a function of the distance  $x$  of the source from the surface, according to equation 4. The brightness profile resembles an exponential fall-off followed by a discontinuity at the surface. When there is no scattering ( $\beta = 0$ ), we have  $E_{med} = 0$ ,  $E_{surf} = L_0 R$  and hence  $E_{surf} \gg E_{med}$ . In this case, the brightness profile is a delta function and it is easy to detect the intersection using a threshold, as is done traditionally. For thresholding to work in the presence of scattering, we must have

$$R \gg \beta \mathcal{P}(g, \alpha) e^{\beta(d_s - x + d_v - y)}. \quad (5)$$

However, when scattering is significant (large  $\beta$ ), it is mostly the opposite case,  $E_{med} \geq E_{surf}$ , as shown by the green and blue profiles in figure 3. Thus, the light plane itself becomes brightly visible (see second column in figure 6). In order to detect the intersection of the light plane and the surface, we simply use the brightness profile as a template until a brightness discontinuity is reached at the end. Even for the hard case where the density of the medium is high, this simple scheme performs well.

### 3.3 Experimental Setup and Calibration

The experimental setup consists of a  $20'' \times 20'' \times 10''$  glass tank filled with water (see figure 4(a)). Different quantities of milk are mixed to emulate scattering media with different densities ( $\beta$ ). The glass faces are anti-reflection coated to avoid reflections. We used an 8-bit Canon XLIS 3-CCD video camera and an Infocus LP120 1000 ANSI Lumens DLP projector in our experiments. To keep the size of the tank small, the camera and the projector are placed outside the tank. Hence, we need to handle light refractions at the *air-glass-medium* interfaces. Our calibration method is similar in spirit to techniques in [11]. Figure 4(b) illustrates a light plane from the projector shining into the glass tank after refraction. Calibration involves sweeping the light plane across two vertical planar surfaces - the  $(u, v)$ - and the  $(s, t)$ -planes - placed in the medium. The 3D world coordinates of a few points on these planes are measured *a priori* (the remaining points are interpolated). Then, the equation of each light plane

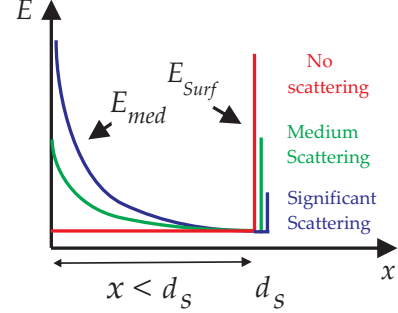


Figure 3: **Brightness profile for detecting the surface and light plane intersection.** When there is no scattering (red), the profile is a delta function which can be thresholded to detect the intersection. As the density of the medium ( $\beta$ ) increases (green and blue), the brightness of the discontinuity ( $E_{surf}$ ) decreases and the light plane becomes brighter ( $E_{med}$ ).

is obtained using its line intersections with the  $(u, v)$ - and  $(s, t)$ -planes. Let this be represented by,

$$Ax + By + Cz + D = 0. \quad (6)$$

Next, we associate with each incoming camera ray (pixel  $(i, j)$ ), its intersections  $P(u, v, r)$  and  $Q(s, t, 0)$  with the  $(u, v)$ - and the  $(s, t)$ -planes respectively (blue line in figure 4(c)). This yields a parametric equation for each camera ray, which is represented by:

$$[x, y, z] = [s, t, 0] + k[u - s, v - t, r - 0], \quad (7)$$

where,  $k$  is a scalar parameter. We calibrated our setup with the two planes placed at  $z = 0$  inches and  $z = 6.0$  inches. To verify calibration accuracy, we reconstructed (as described in Section 3.4) a plane placed at  $z = 4.18$  inches with a low RMS error of 0.21 inch (figure 5). *In summary, our method does not require explicit geometric calibration of either the camera or the projector and does not require the position/orientation of the glass face or the refractive indices of media.*

### 3.4 Scene and Medium Recovery

Once calibrated the setup may be used to recover the 3D structure and *clear-air* appearance of any object in the medium as well as the properties of the medium itself.

**3D surface reconstruction:** Figure 4(c) shows a top-view (2D) illustration of the light striping setup and the profile of an object's surface. Since a point on the surface lies at the intersection of the reflected ray (blue) and the light plane (red), we may substitute  $(x, y, z)$  from equation 7 into equation 6, to solve for the parameter  $k$ :

$$k = \frac{As + Bt + D}{A(s - u) + B(t - v) - Cr}. \quad (8)$$

The value of  $k$  is then substituted back into equation 7 to obtain the 3D coordinates  $(x, y, z)$  of the surface point.

**Medium properties:** The properties of the medium can be obtained by observing the brightness decay of the light plane without the surface (see profile of  $E_{med}$  in figure 3). The distances  $x$  and  $y$  can be computed using the 3D coordinates of points on the light plane and the dimensions

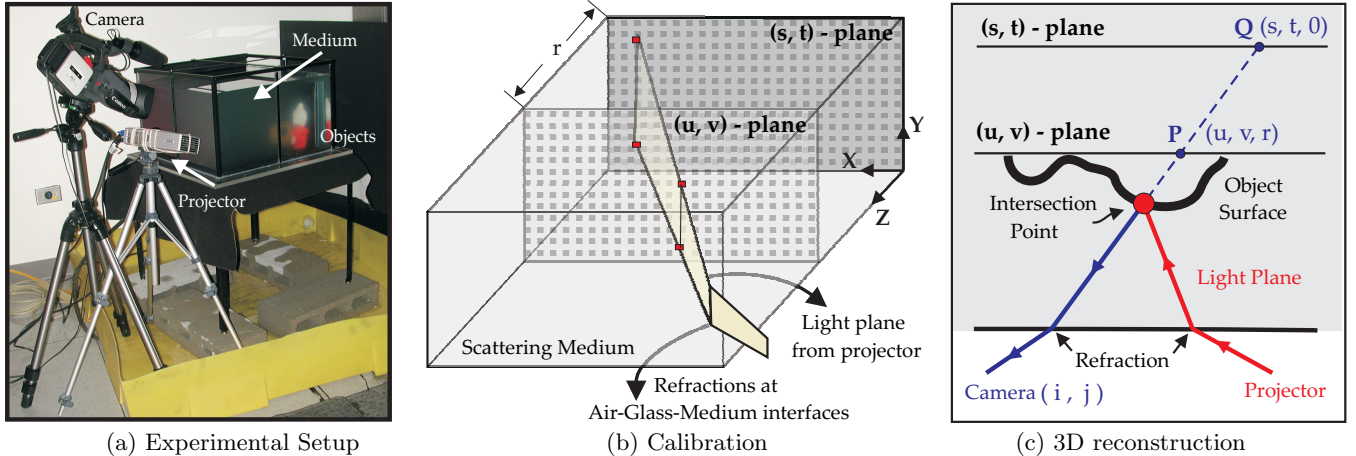


Figure 4: **Light striping experimental setup and calibration.** (a) The setup consists of a glass tank filled with a scattering medium (dilute milk). The scene of interest is immersed in the medium. A projector illuminates the medium and the scene with planes of light and a video camera views the scene with the effects of scattering. (b) The light plane sweeps (one at a time) two planar surfaces placed vertically in the tank at known distances ( $z = 0$  and  $z = r$ ), called the  $(u, v)$ - and the  $(s, t)$ -planes. The discrete mappings between the light plane and the  $(u, v)$ - and  $(s, t)$ -planes, and between the camera ray and the  $(u, v)$ - and  $(s, t)$ -planes constitute calibration. Note that no knowledge of the refraction locations or indices is required. (c) The top view of the setup illustrating the intersection of the light plane and the camera ray to yield the 3D coordinates of a surface point.

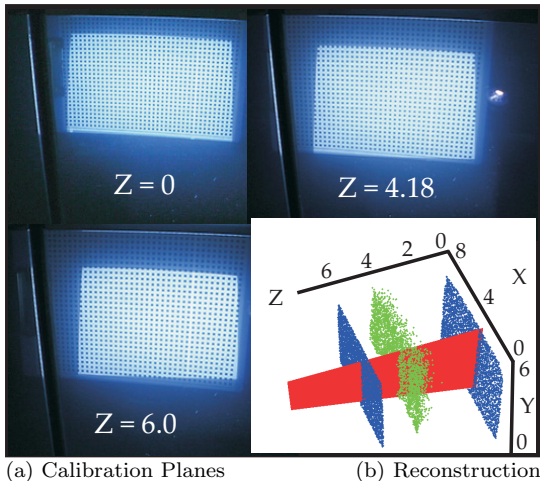


Figure 5: **Verification of light striping calibration.** (a) Two planes at  $z = 0$  and  $z = 6.0$  inches are used for calibration. (b) The computed equations of light planes and camera rays are then used to reconstruct a third plane at  $z = 4.18$  inches (with RMS error 0.21 inch). The 3D view shows the three vertical planes and a light plane (red) for illustration.

of the tank. Then, equation 2 is nonlinear in the two unknown medium parameters,  $\beta$  and  $g$ . Thus, by observing the irradiances  $E_{med}$  along a profile on the light plane, we can estimate the two parameters  $\beta$  and  $g$  using a nonlinear optimization method (“fminsearch” in Matlab<sup>TM</sup>).

**Scene appearance without scattering:** Once the scattering coefficient  $\beta$  is estimated and the 3D surface is reconstructed, the scene appearance without scattering can be computed for each object intersection strip, from equation 3 as,

$$L_0 R = E_{surf} e^{+\beta(d_s + d_v)}, \quad (9)$$

where,  $E_{surf}$  is the observed brightness of the object in the presence of scattering. Then, all the intersection strips

are mosaiced to create the appearance of the entire scene as if captured in clear air.

The results of applying the scene and medium recovery algorithms are shown using real experiments in figure 6. The detection of the object intersections and hence the 3D reconstruction obtained under different densities of scattering compare well with the ground truth. Despite the strong effects of scattering, we are able to remove them completely to restore the original scene contrast. Also a comparison to the floodlit images demonstrates that simply using bright sources does not enhance visibility in scattering media, and that structured lighting methods that are designed to focus light on the scene to alleviate blurring and backscattering must be used.

## 4 Photometric Stereo in Scattering Media

In situations where light stripe scanning takes too long to be practical (for example, dynamic scenes), photometric stereo [29] provides an attractive alternative. Traditionally, photometric stereo is a technique for scene reconstruction (surface normal and albedo) from a small number of images of the scene acquired under different lighting directions. Many variants of this problem exist in vision literature [12; 2], but none of the proposed solutions are effective in scattering media.

In this section, we show how photometric stereo can be extended to scattering media. We choose the simplest version of the problem that assumes the surfaces to be lambertian, the sources distant, interreflections negligible and the camera to be orthographic. In the absence of scattering, it is known that three images of a scene illuminated from different but known directions are sufficient to uniquely determine the surface normals and albedos. We will first determine how many sources are needed in the presence of scattering and then show how scene properties can be recovered from the corresponding images.



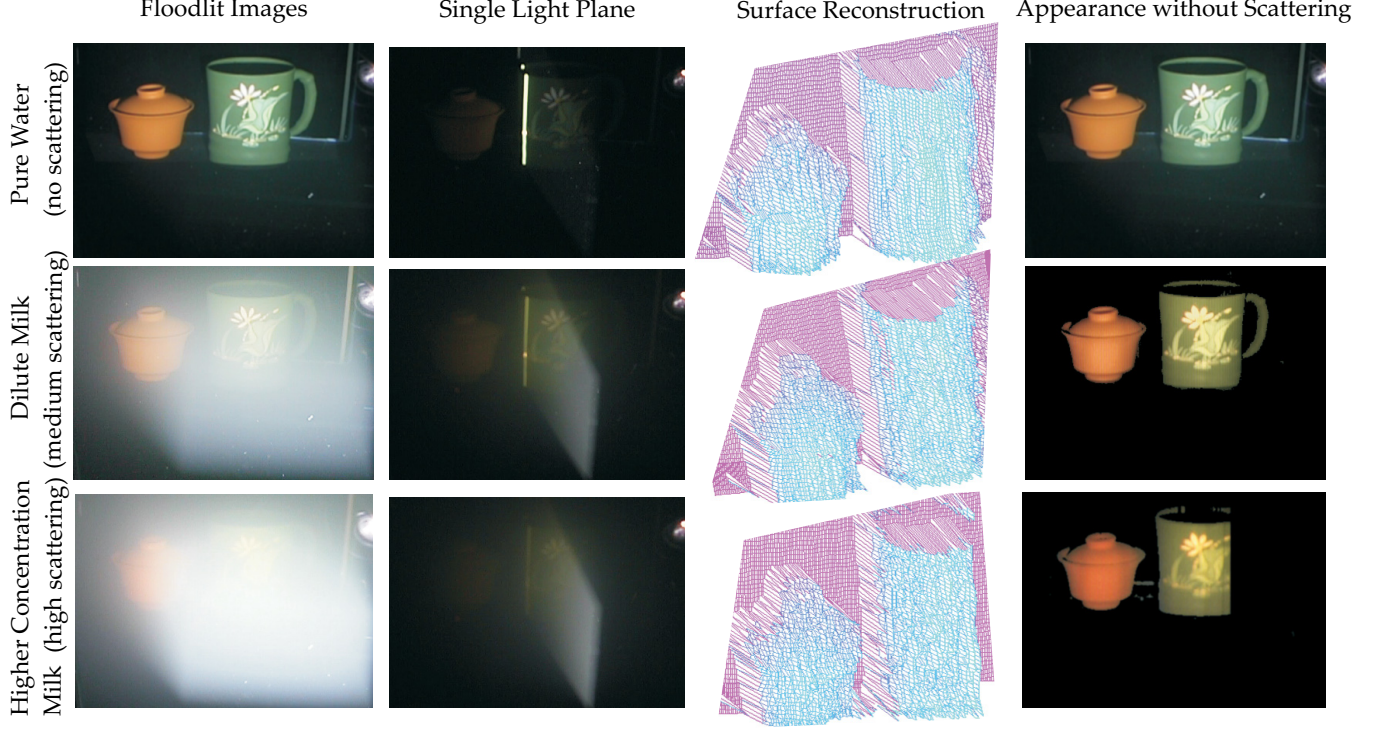


Figure 6: **Experimental results for light striping in scattering media.** The scene consists of two objects immersed in pure water (no scattering, ground truth), water mixed with 6 ml milk (medium scattering) and 15 ml milk (high scattering). The floodlit images (obtained by turning on all pixels in the projector) illustrate the adverse effects due to scattering by the medium. The brightness profile of a single light plane focused on the object confirms the template of the profile model in figure 3. For the two concentrations, our methods estimated  $\beta = 0.07 \text{ in}^{-1}$  and  $0.16 \text{ in}^{-1}$  and  $g = 0.9$ . In the medium scattering case, our results (3D reconstruction and scene appearance without scattering) are nearly identical to the ground truth (percentage RMS error = 2.1%). In the 15 ml milk case, the green cup is barely visible (especially since its albedo is low) and yet the result is close to the ground truth. The handle of the cup is completely invisible and is hence missed (else, percentage RMS error = 5.5%). Note that our algorithm must be applied to separately to individual color channels. The color difference between the pure water and the other cases is due to white balancing differences between different experiments. (**Please see a video on our website [27] for better visualization and for other examples.**)

#### 4.1 Image Formation Model

Consider the illumination and observation geometry in figure 7. A distant source (direction  $\mathbf{s}$ ) illuminates a surface point  $\mathbf{P}$  with unit normal  $\mathbf{n}$  and albedo  $\rho$ . A camera observing the surface receives irradiance  $E_{surf}$  due to the light reflected by the surface (solid red lines) and irradiance  $E_{med}$  due to light scattered by the medium (dashed lines) in the viewing direction. The irradiance  $E_{surf}$  is the same as for light striping (see equation 3),

$$E_{surf} = L_0 e^{-\beta d_s} (\rho \mathbf{n} \cdot \mathbf{s}) e^{-\beta d_v}. \quad (10)$$

Here, we have replaced the normalized radiance  $R$  by  $(\rho \mathbf{n} \cdot \mathbf{s})$  for a lambertian surface. The irradiance  $E_{med}$  at the camera due to single scattering by the medium is obtained by integrating the brightness along the viewing direction (see equation 2),

$$E_{med} = \int_0^{d_v} L_0 e^{-\beta x} \beta \mathcal{P}(g, \alpha) e^{-\beta y} dy. \quad (11)$$

Note that  $\alpha$ ,  $\mathcal{P}(g, \alpha)$ ,  $\beta$  and  $L_0$  are all independent of the integration variable  $y$ . Further, we shall also assume

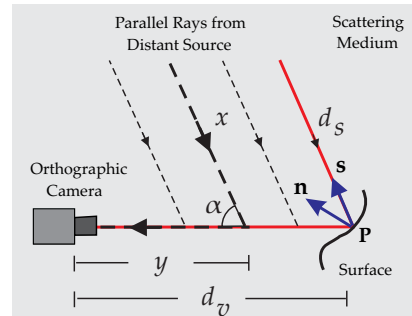


Figure 7: **Image formation for photometric stereo in scattering media.** The sources, viewer and the surface of interest are immersed in the scattering medium. The sources are distant and thus illuminate the surface and the viewing ray in a collimated fashion. The brightness at a pixel is the sum of the contributions from the solid red and the dashed rays.

the source uniformly illuminates the viewing distance  $d_v$ . In other words,  $x = d_s$  is constant with respect to  $y$  (this assumption will be relaxed when we discuss our specific setup). This allows us to simplify equation 11 as,

$$E_{med} = L_0 \mathcal{P}(\alpha) e^{-\beta d_s} (1 - e^{-\beta d_v}). \quad (12)$$

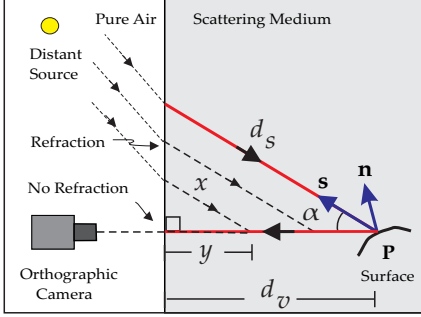


Figure 8: **Refraction of rays in the photometric stereo model.** The sources and camera are outside the scattering medium. The viewing direction of the orthographic camera is normal to the air-medium interface to avoid refractions of incoming camera rays. However, refraction of light rays from the source must be modeled.

Then, the total irradiance  $E$  at the camera can be written as the sum of the irradiances  $E_{med}$  and  $E_{surf}$ :

$$E = L_0 [e^{-\beta(d_s+d_v)} \rho \mathbf{n} \cdot \mathbf{s} + \mathcal{P}(g, \alpha) e^{-\beta d_s} (1 - e^{-\beta d_v})]. \quad (13)$$

For an isotropic point source,  $L_0 = I_0/d_s^2$ . Equation 13 represents the image formation model for one distant source. Similar equations can be written for each distant source that illuminates the scene.

## 4.2 Experimental Setup and Calibration

The glass tank described in Section 3 is again used in these experiments and, as before, we place the camera and the sources outside the tank. A 12-bit per channel Canon EOS-20D camera with a 70 – 300 mm zoom lens is placed 20 feet away from the tank and observes the front face of the tank normally (perpendicularly). The field of view occupied by the tank in the image is 2.0 degrees and is hence approximately orthographic.

During calibration, refraction of the light rays from sources at the *air-medium* boundary must be accounted for. Figure 8 shows a schematic of the side view of the setup. The distances  $d_s$  and  $d_v$  are related using trigonometry,

$$d_v = d_s \cos \alpha. \quad (14)$$

Notice that the light rays that illuminate the viewing ray and the surface travel different distances in the medium (compare the lengths of the dashed parallel rays in figures 7 and 8). Hence, the assumption in simplifying equation 11 that  $x$  is constant with respect to  $y$ , becomes invalid for our experimental setup. So, an appropriate correction is derived for  $E_{med}$  using equation 14 to obtain the irradiance (see appendix A):

$$E = L_0 e^{-\beta d_v(1+1/\cos \alpha)} \rho \mathbf{n} \cdot \mathbf{s} + \frac{L_0 \mathcal{P}(g, \alpha) \cos \alpha}{1 + \cos \alpha} (1 - e^{-\beta d_v(1+1/\cos \alpha)}). \quad (15)$$

We will henceforth call equation 15 as the image formation model. We calibrate our setup using images of a white lambertian sphere in pure water (scattering is minimal). The brightest point on the sphere yields the refracted direction  $\mathbf{s}$  (and  $\alpha$ ) and intensity  $L_0$  of the source.

## 4.3 Scene and Medium Recovery

Consider a set of images taken of an object under different source directions. In order to find out how many source directions are required to recover the scene and the medium, let us count the number of knowns and unknowns in equation 15. Recall that as part of calibration, the angle  $\alpha$ , the source direction  $\mathbf{s}$  and intensity  $L_0$  are all estimated *a priori*. Then, the unknowns for each scene point are the surface albedo  $\rho$ , unit normal  $\mathbf{n}$ , and optical thickness  $T_v = \beta d_v$ . The medium parameter  $g$  in the expression for  $\mathcal{P}(g, \alpha)$  (see equation 1) is constant and hence is a global unknown. Thus, there are four unknowns for each scene point and one global unknown. If there are  $P$  scene points and  $L$  light source directions, the number of unknowns  $4P + 1$  must be less than the number of equations  $PL$ . So, simple variable counting suggests that a minimum of  $L = 5$  is required<sup>6</sup>.

To empirically verify that indeed  $L = 5$  suffices (assuming the sources are not in degenerate positions), we performed numerical simulations on 4000 randomly generated combinations of source directions  $\mathbf{s}_i$ , surface normals  $\mathbf{n}$ , albedos  $\rho \in (0, 1)$ , optical thicknesses  $T_v \in (0, 2)$  and forward scattering parameters  $g \in (-1, 1)$ , for a single scene point. The Matlab<sup>TM</sup> function “fminsearch” was used to recover the unknowns by minimizing the sum of squared differences between the simulated values and the model in equation 15. In all trials, the search was initialized with random values for the unknowns. In all cases, the search algorithm converged to the global optimum solution within few seconds. This suggests the presence of a single global minimum of the error function<sup>7</sup>. As a test of robustness, we added uniform random noise (up to 5% of the simulated values) and found that the errors in recovered unknowns were low, as evidenced by the error histograms in figure 10. We also ran the above simulations using only 4 sources, but the global error minimum corresponded to several parameter sets, suggesting that 4 sources are insufficient for unique estimation. *Thus, we conclude that five non-degenerate light source directions are required and sufficient to uniquely estimate the properties of the scene and the medium. In practice, however, more source directions may be used for robustness.*

The experiments performed with our setup are shown in figure 9. Images of a teapot captured in the presence of scattering (by dilute milk) have poor contrast and colors. As expected, applying traditional photometric stereo results in poor results. On the other hand, the surface normals and the albedos obtained using our method<sup>8</sup> are

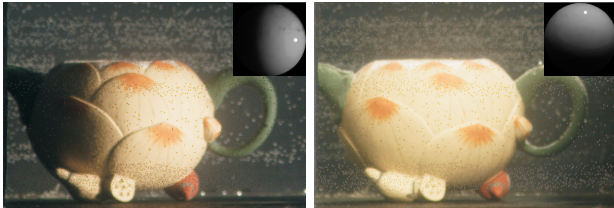
<sup>6</sup>In appendix B, we present an interesting but practically limited case where a unique linear solution with four sources is possible.

<sup>7</sup>However, the error function does contain local minima and the search was conducted starting from several (typically 100) initial guesses and the minimum of all the solutions was used.

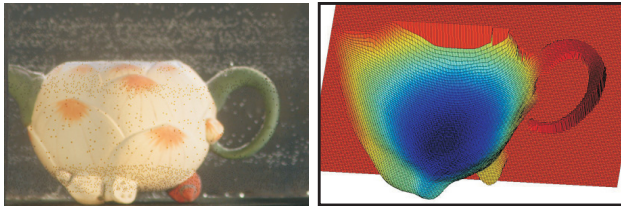
<sup>8</sup>The non-linear optimization can be executed independently for each pixel. But to speedup execution time, we masked the region where the object is not present ( $E_{surf} = 0$ ) to first estimate the global parameter  $g$ , before estimating the 4 parameters for each pixel. As a side note, from our experience, it is critical to use high quality (12 bits per channel HDR) radiometrically linear input images in order to obtain good reconstructions shown.



Pure Water (No scattering) - Ground Truth



(a) Images (2 out of 8) captured in pure water.

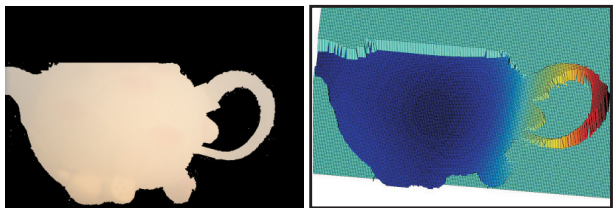


(b) Albedo and shape computed using traditional method

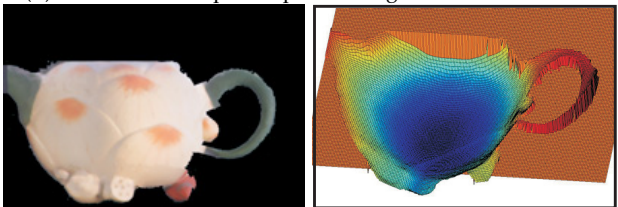
Dilute Milk (medium scattering)



(c) Input Images (2 out of 8) captured in dilute milk



(d) Albedo and shape computed using traditional method



(e) Albedo and shape computed using our method

**Figure 9: Experimental results of Photometric Stereo in Scattering Media.** (a) Two (out of eight) images of a teapot acquired under different lighting directions (depicted in the insets). (b) Results from applying traditional photometric stereo to images in (a) serve as ground truth. (c) The images acquired in dilute milk. Notice the significant loss of contrast. (d) If traditional photometric stereo applied to images in (c), the 3D shape obtained is very flat and the scattering effects are absorbed by the albedos. (e) The results obtained using our algorithm. The percentage RMS error in reconstructing the shape was 6.3%. In addition to surface normals and albedos, our method also yields a depth map, which is not possible using traditional photometric stereo (see figure 10.) The 3D shapes were computed from the surface normals using [8].

very accurate (with only a percentage RMS error of 6.3% in computed shape). *In addition, our method also yields a depth map of the scene  $T_v = \beta d_v$ , which is not possible using traditional photometric stereo.*

## 5 Discussion: Real Underwater Scenarios

Since all the experiments in this paper were done using milk as the scattering medium, it is important to identify issues that may arise in real underwater scenarios (lakes, oceans, seas). In initial experiments, we collected water samples from 4 different locations near a Pacific Ocean beach in San Diego, USA. These samples were collected close to the water surface where the impurity levels are generally high. By matching images of these water samples with those of milk, we found that the low to moderate concentration range of milk used in this paper corresponds to the concentrations observed in the ocean waters. Hence, our algorithms can be applied to many underwater scenarios as well. In the future, to make our techniques broadly applicable, we wish to develop underwater deployable systems and improve our algorithms to handle non-homogeneous and dynamic underwater media.

## References

- [1] R. D. Ballard. *The Discovery of the Titanic*. Warner Books, 1988.
- [2] R. Basri and D.W. Jacobs. Photometric stereo with general, unknown lighting. In *CVPR*, 2001.
- [3] F. M. Caimi, D. M. Kocak, and V. L. Asper. Developments in laser-line scanned undersea surface mapping and image analysis systems for scientific applications. In *Proc. MTS/IEEE Oceans*, 1996.
- [4] S. Chandrasekhar. *Radiative Transfer*. Dover Publications, Inc., 1960.
- [5] P. C. Y. Chang, J. C. Flitton, K. I. Hopcraft, E. Jakeman, D. L. Jordan, and J. G. Walker. Improving visibility depth in passive underwater imaging by use of polarization. *App. Opt.*, 42 (15), 2003.
- [6] S. Y. Chen and Y. F. Li. Self-recalibration of a colour-encoded light system for automated 3-d measurements. *MeasureSciTech*, 14(1), 2003.
- [7] G.R. Fournier, D. Bonnier, J.L. Forand, and P.W. Pace. Ranged underwater laser imaging system. *Opt. Eng.*, 32 (9), 1993.
- [8] R.T. Frankot and R. Chellappa. A method for enforcing integrability in shape from shading algorithms. *PAMI*, 10(4), 1988.
- [9] E. Fuchs and J. S. Jaffe. Thin laser light sheet microscope for microbial oceanography. *OPTICS EXPRESS*, 10 (2), 2002.
- [10] 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.
- [11] M.D. Grossberg and S.K. Nayar. The raxel imaging model and ray-based calibration. *IJCV*, 61(2), 2005.
- [12] A. Hertzmann and S.M. Seitz. Shape and materials by example: a photometric stereo approach. In *CVPR*, 2003.
- [13] D.Q. Huynh, R.A. Owens, and P.E. Hartmann. Calibrating a structured light stripe system: A novel approach. *IJCV*, 33(1), 1999.
- [14] J. S. Jaffe, J. McLean, M. P. Strand, and K. D. Moore. Underwater optical imaging: Status and prospects. *Tech. Report, Scripps Institution of Oceanography, La Jolla*, 2002.
- [15] C. Je, S.W. Lee, and R.H. Park. High-contrast color-stripe pattern for rapid structured-light range imaging. In *ECCV*, 2004.
- [16] D. M. Kocak, F. M. Caimi, T. H. Jagiello, and J. Kloske. Laser projection photogrammetry and video system for quantification and mensuration. *MTS/IEEE Oceans*, 2002.

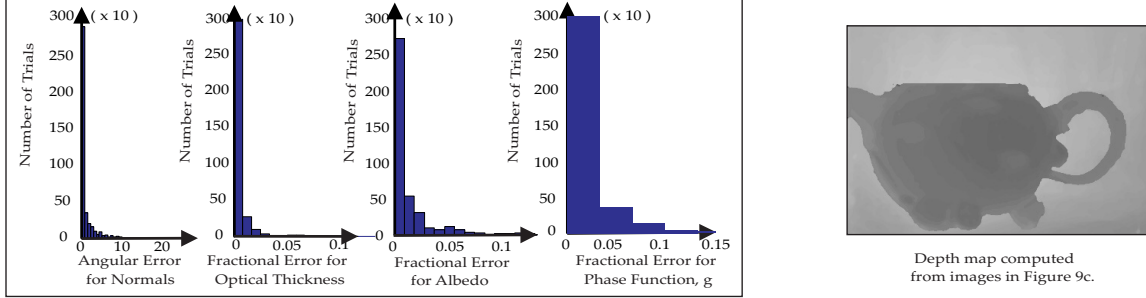


Figure 10: [Left] Simulations show performance of our algorithm for photometric stereo, in the presence of small amounts of noise (uniform random noise up to 5% of the simulated values). The peaks near zero values in the error histograms corresponding to the recovered parameters indicate robustness. [Right] The presence of scattering allows us to compute the scaled depth  $T_v$  of every scene point, which is not possible in traditional photometric stereo. The depth map was filtered to remove some noise.

- [17] M. Levoy, B. Chen, V. Vaish, M. Horowitz, I. McDowall, and M. Bolas. Synthetic aperture confocal imaging. In *SIG-GRAPH*, 2004.
- [18] S.G. Narasimhan and S.K. Nayar. Vision and the atmosphere. *IJCV*, 48(3).
- [19] P. Naulleau and D. Dilworth. Motion-resolved imaging of moving objects embedded within scattering media by the use of time-gated speckle analysis. *App. Opt.*, 35 (26), 1996.
- [20] Y. Sato, H. Kitagawa, and H. Fujita. Shape measurement of curved objects using multiple slit-ray projections. *PAMI*, 4(6), 1982.
- [21] D. Scharstein and R. Szeliski. High-accuracy stereo depth maps using structured light. In *CVPR*, 2003.
- [22] Y. Y. Schechner and N. Karpel. Attenuating natural flicker patterns. In *Proc. MTS/IEEE Oceans*, 2004.
- [23] Y. Y. Schechner and N. Karpel. Clear underwater vision. In *Proc. CVPR*, 2004.
- [24] J. S. Tyo, M. P. Rowe, Jr. E. N. Pugh, and N. Engheta. Target detection in optically scattering media by polarization-difference imaging. *App. Opt.*, 35 (11), 1996.
- [25] J. G. Walker, P. C. Y. Chang, and K. I. Hopcraft. Visibility depth improvement in active polarization imaging in scattering media. *App. Opt.*, 39 (27), 2000.
- [26] D. Walther, D. R. Edgington, and C. Koch. Detection and tracking of objects in underwater video. In *Proc. CVPR*, 2004.
- [27] Authors Website. <http://www.cs.cmu.edu/~srinivas/>.
- [28] P.M. Will and K.S. Pennington. Grid coding: A preprocessing technique for robot and machine vision. *AI*, 2, 1971.
- [29] R.J. Woodham. Photometric method for determining surface orientation from multiple images. *OptEng*, 19(1), 1980.
- [30] L. Zhang, B. Curless, and S. M. Seitz. Rapid shape acquisition using color structured light and multi-pass dynamic programming. In *The 1st IEEE International Symposium on 3D Data Processing, Visualization, and Transmission*, 2002.

## A: Refractions in Photometric Stereo

The expression for  $E_{surf}$  is the same as equation 10 with  $d_s$  being substituted from equation 14. The irradiance  $E_{med}$  is derived by observing that  $x = y \cos \alpha$  and  $d_v = d_s \cos \alpha$ :

$$E_{med} = \int_0^{d_v} L_0 e^{-\beta x} \beta \mathcal{P}(g, \alpha) e^{-\beta y} dy = L_0 \beta \mathcal{P}(g, \alpha) \int_0^{d_v} e^{-\beta(1+1/\cos \alpha)y} dy$$

$$= \frac{L_0 \mathcal{P}(g, \alpha) \cos \alpha}{1 + \cos \alpha} (1 - e^{-\beta(1+1/\cos \alpha)d_v})$$

which is rewritten in equation 15. Note that we have assumed here that the actual source outside the medium

is still distant and does not show an inverse square fall-off within the medium since the distance light traveled within the medium is much less compared to the distance of the source from the medium boundary. Thus,  $L_0$  is constant with respect to  $d_s$ .

## B: Linear Solution with Four Sources

We assume that (a) the sources and the camera are both immersed in an isotropic medium ( $\mathcal{P}(\alpha) = \frac{1}{4\pi}$ ), and (b) the sources are equidistant from any given surface point ( $d_s$  same for all sources, however,  $d_s$  may vary for different scene points). Although these assumptions limit practicality, this case serves as an illustration in comparison to the technique without any scattering. Let the intensities observed at a pixel under 4 different sources be  $E_1$ ,  $E_2$ ,  $E_3$ , and  $E_4$ . From equation 13:

$$E_i = L_0 e^{-\beta(d_{s_i}+d_v)} \rho \mathbf{n} \cdot \mathbf{s}_i + \frac{L_0}{4\pi} e^{-\beta d_{s_i}} (1 - e^{-\beta d_v}), \quad (16)$$

where, the subscript  $i$  denotes the source. For this special case,  $d_{s_1} = d_{s_2} = d_{s_3} = d_{s_4} = d_s$ . The second term in equation 16 can be eliminated by considering pair-wise differences between intensities to obtain a set of equations that are linear in the surface normal  $\mathbf{n}$ :

$$\begin{pmatrix} E_1 - E_2 \\ E_1 - E_3 \\ E_1 - E_4 \end{pmatrix} = L_0 e^{-\beta(d_s+d_v)} \rho \begin{pmatrix} \mathbf{s}_1 - \mathbf{s}_2 \\ \mathbf{s}_1 - \mathbf{s}_3 \\ \mathbf{s}_1 - \mathbf{s}_4 \end{pmatrix} \mathbf{n}. \quad (17)$$

In matrix and vector notation, we rewrite equation 17,

$$\Delta \mathbf{E} = L_0 e^{-\beta(d_s+d_v)} \rho \Delta \mathbf{S} \mathbf{n}, \quad (18)$$

where,  $\Delta \mathbf{E}$  and  $\mathbf{n}$  are  $3 \times 1$  column vectors and  $\Delta \mathbf{S}$  is a  $3 \times 3$  matrix. By applying  $\Delta \mathbf{S}^{-1}$ , we obtain the surface normal  $\mathbf{n}$  and the attenuated appearance  $L_0 e^{-\beta(d_s+d_v)} \rho$ . Thus, all the terms except for  $\beta d_s$  and  $\beta d_v$  are estimated or known in equation 16. If a relationship between  $d_s$  and  $d_v$  is known (such as equation 14), then  $\beta d_s$  and  $\beta d_v$  can be estimated using equation 16 and in turn the albedo  $\rho$  can be estimated from the product  $L_0 e^{-\beta(d_s+d_v)} \rho$ . Thus, in the case of isotropic scattering and equidistant sources, four light source directions are required to compute surface normals, albedos and optical thickness  $\beta d_v$  (or scaled distance).