

Photometric Stereo in a Scattering Medium

Zak Murez¹, Tali Treibitz², Ravi Ramamoorthi¹, and David Kriegman^{1,3}

¹University of California, San Diego, ²University of Haifa, ³Dropbox, {zmurez,ravir,kriegman}@cs.ucsd.edu, ttreibitz@univ.haifa.ac.il

Abstract

Photometric stereo is widely used for 3D reconstruction. However, its use in scattering media such as water, biological tissue and fog has been limited until now, because of forward scattered light from both the source and object, as well as light scattered back from the medium (backscatter). Here we make three contributions to address the key modes of light propagation, under the common single scattering assumption for dilute media. First, we show through extensive simulations that single-scattered light from a source can be approximated by a point light source with a single direction. This alleviates the need to handle light source blur explicitly. Next, we model the blur due to scattering of light from the object. We measure the object point-spread function and introduce a simple deconvolution method. Finally, we show how imaging fluorescence emission where available, eliminates the backscatter component and increases the signal-to-noise ratio. Experimental results in a water tank, with different concentrations of scattering media added, show that deconvolution produces higher-quality 3D reconstructions than previous techniques, and that when combined with fluorescence, can produce results similar to that in clear water even for highly turbid media.

1. Introduction

Obtaining 3D information about an object submersed in fog, haze, water, or biological tissue is difficult because of scattering [5, 15, 22]. In this paper, we focus on photometric stereo, which estimates surface normals from intensity changes under varying illumination. In air, photometric stereo produces high-quality geometry, even in textureless regions with small details, and is a widely used 3D reconstruction method.

In a scattering medium, however, light propagation is affected by scattering which degrades the performance of photometric algorithms unless accounted for (Fig. 1). Distance dependent attenuation caused by the medium has been dealt with in the past [16]. Here, our contributions lie in handling three scattering effects (Fig. 2), based on a single scatter model [25]: 1) light traveling from the source to the object is blurred due to forward scattering; 2) light traveling from the object to the camera is blurred due to forward

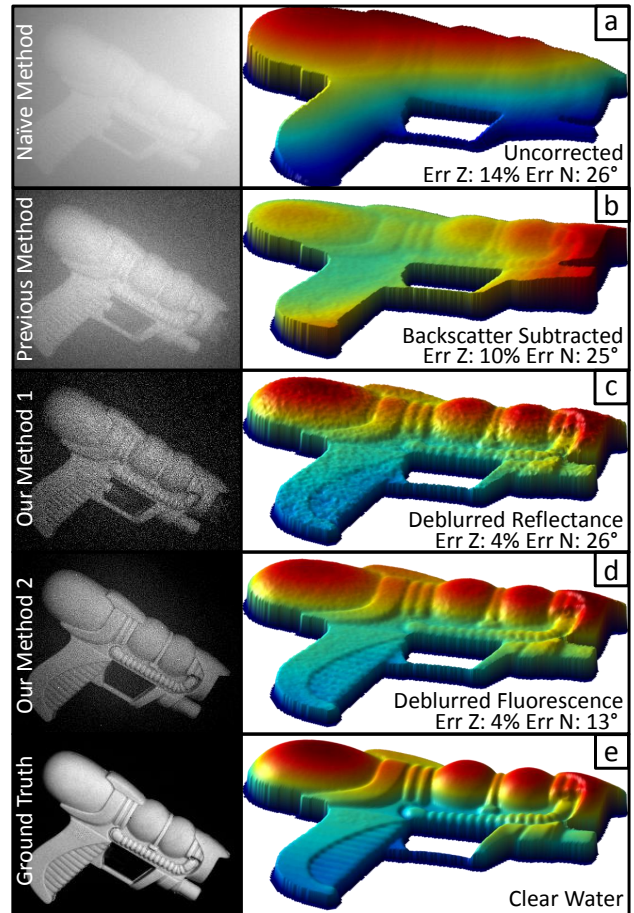


Figure 1. 3D shape recovery of a squirt gun using photometric stereo in a scattering medium. [Left] one of the input images, [Right] reconstruction results. Err Z is the mean absolute error in the heights and captures the overall shape error, while Err N is the mean angular error in the normals and captures the error caused by noise. a) Result of standard photometric stereo in air. It ignores the medium effects, and thus the shape is not well defined. b) Reconstruction result of a recent method [32]. c) By applying our method to account for blur due to scattering from the object, we improve the previous result and recover the correct shape. However the surface still exhibits noise due to a small dynamic range after backscatter subtraction. d) Using fluorescence to optically eliminate the backscatter prior to deblurring reduces the noise and produces the best result. e) Result in clear water.

scattering; 3) light traveling *from the source is scattered back towards the camera* without hitting the object. This is known as backscatter and is an additive component that veils the object. All these effects are distance dependent and thus depend on the object 3D surface: the property we aim to reconstruct. To handle this we introduce the *small surface variations* approximation for the object (Sec. 3), that assumes surface changes are small relative to the distance from the object (that is assumed to be known). This assumption removes the dependence on the unknown surface heights Z , but unlike the common distant light/camera approximations, it still allows for dependencies on spatial locations X and Y . One important consequence of this is the ability to model anisotropic light sources, which is not possible for distant lights.

Forward scatter was previously compensated for iteratively for both pathways (light to object and object to camera) simultaneously [21]. We analyze the paths separately. The resulting algorithm is simpler, requires fewer images and yields better results. First, consider the blurring of light traveling *to* the object from the source (Fig. 2b). The photometric stereo formulation assumes a point light source, illuminating from a single direction. However, if the source is scattered by the medium, this no longer holds: the point light source is spread, and the direction of light rays incident on the object changes. Nonetheless, for a Lambertian surface, illuminated from a variety of directions, we still get a linear equation between the image intensities and the surface normals [24]. Here, we show through simulations in a large variety of single scattering media, that a forward-scattered light source illuminating a Lambertian surface can be well approximated by a non-blurred light source in an effective purely absorbing medium (Sec. 6). This allows for much easier calibration in practice.

Next, we observe that the blur caused by scattering *from* the object to the camera (Fig. 2c) significantly affects the shape of the surface reconstructed by photometric stereo. This important effect has been neglected in many previous works. In general, the point-spread function (PSF) for an object is spatially varying and dependent on the unknown scene depths. However, we demonstrate that a spatially invariant approximation can still achieve good results, when calibrated for the desired medium and approximate object distance. We estimate the PSF and use it to deconvolve the images after backscatter has been removed (Sec. 7). These corrected images are used as input to a linear photometric stereo algorithm to recover the surface normals, which are then integrated. This results in much higher quality 3D surfaces across varying turbidity levels.

Finally, consider the backscatter component (Fig. 2a). In a previous work (Tsotsios et al. [32]) backscatter was calibrated and subtracted from the input images. However, when backscatter is strong relative to the object signal, subtracting it after image formation leads to lack of dynamic range and lower signal-to-noise ratios (SNR) [30] that significantly degrades deblurring and reconstruction. Here we

show that if the object fluoresces, this can be leveraged to optically remove the backscatter prior to image formation (Sec. 8). Fluorescence is the re-emission of photons in wavelengths longer than the excitation light [6], and therefore the backscatter can be eliminated by optically blocking the excitation wavelengths and imaging only the fluorescence emission. This improves SNR, especially in high turbidity. This approach is feasible as many natural underwater objects such as corals and algae fluoresce naturally.

We demonstrate our method experimentally in a water tank (Sec. 10) with varying turbidity levels. Deblurring can be used separately or combined with fluorescence imaging to significantly improve the quality of photometric stereo reconstructions, as shown in Fig. 1.

2. Previous Work

The traditional setup for photometric stereo assumes a Lambertian surface, orthographic projection, distant light sources, and a non-participating medium [34]. However, underwater light is exponentially attenuated with distance, and thus the camera and lights must be placed close to the scene for proper illumination. This means that the orthographic camera model, and the distant light assumptions, are no longer valid. In addition, attenuation and scattering by the medium need to be accounted for. These effects were partially considered in previous works.

Near-Field Effects and Exponential Attenuation: Photometric stereo in air was solved with perspective cameras [17, 27], nearby light sources [14], or both [4, 11]. Kolagani et al. [16] uses a perspective camera, nearby light sources and includes exponential attenuation of light in a medium. Their formulation leads to nonlinear solutions for the normals and heights. We handle these near field effects but linearize the problem (Sec. 3).

Photometric Stereo with Backscatter: Narasimhan et al. [20] handles backscatter and attenuation, with the assumption of distant light sources and an orthographic camera. Tsotsios et al. [32] extends this to nearby point sources and assumes the backscatter saturates close to the camera and thus does not depend on the unknown surface height. Then, it can be calibrated and subtracted from the images. We use the method in [32] in one of our variants.

Backscatter Removal: Backscatter was previously removed for visibility enhancement, by structured light [7], range-gating [15], or using polarizers [29]. Nevertheless, these methods do not necessarily preserve photometric information. It is sometimes possible to reduce backscatter by increasing the camera light source separation [7, 12], but this often leads to more shadowed regions, creating problems for photometric stereo.

Fluorescence Imaging: Removing scatter using fluorescence is used in microscopy [33], where many objects of interest are artificially dyed to fluoresce. Hullin et al. [9] imaged objects immersed in a fluorescent liquid to reconstruct their 3D structure. It was recently shown that the fluorescence emission yields photometric stereo reconstruc-

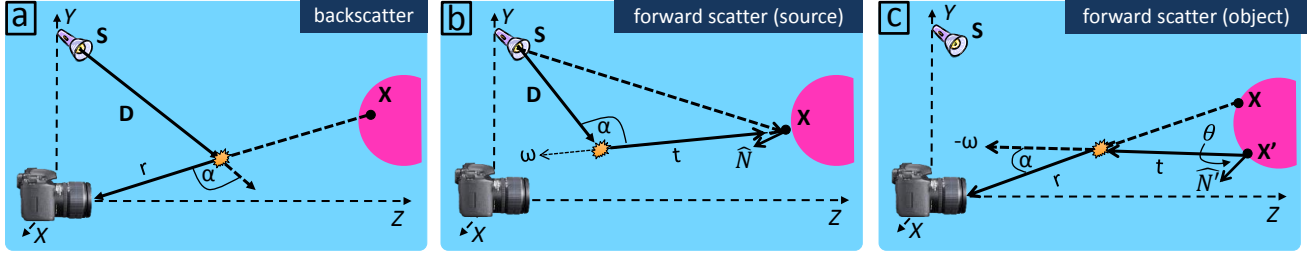


Figure 2. A perspective camera is imaging an object point at $\mathbf{X}$, with a normal $\mathbf{N}$, illuminated by a point light source at $\mathbf{S}$. The object is in a scattering medium, and thus light may be scattered in the three ways shown, detailed in Sec. 3.

tions [23, 28] in air that are superior to reflectance images as the fluorescence emission behaves like a Lambertian surface due to its isotropic emission.

Deblurring Forward Scatter: Zhang et al. [35] and Negahdaripour et al. [21] handle blur caused by forward scatter using the PSF derived in [12, 18]. Their PSF depends on the unknown distances, as well as three empirical parameters, and affects both the path from the light source to the object and from the object to the camera. They iteratively deconvolve and update the depths until a good result is achieved. Trucco et al. [31] simplify the PSF of [12, 18] to only depend on two parameters while assuming the depth is known. Our PSF is nonparametric, independent of the unknown depths and only affects the path to the camera, which allows for a direct solution without iteration. While we look at a Lambertian surface in a scattering medium, Inoshita et al. [10] and Dong et al. [3] consider the problem of photometric stereo in air on a surface that exhibits subsurface scattering, which blurs the radiance across the surface. They deconvolve the images to improve the quality of the normals recovered using linear photometric stereo. Tanaka et al. [26] also model forward scatter blur as a depth dependent PSF and combine it with multi (spatial) frequency illumination to recover the appearance of a small number of inner slices of a translucent material.

3. Overview and Assumptions

In this section we introduce the image formation model, considering each of the modes of light propagation in a single scattering medium, as shown in Fig. 2. We derive expressions for each component in the following sections.

Consider a perspective camera placed at the origin, with the image (x, y) coordinates parallel to the world's (X, Y) axes, and the Z -axis aligned with the camera's optical axis. Let the point $\mathbf{X} = (X, Y, Z)$ be the point on the object's surface along the line of sight of pixel $\mathbf{x} = (x, y)$. Let $\mathbf{S}$ be the world coordinates of a point light source, and define $\mathbf{D}(\mathbf{X}) = \mathbf{S} - \mathbf{X}$ as the vector from the object to the source.

We assume a single scattering medium which allows us to express the radiance L_o reflected by a surface point as the sum of two terms:

$$L_o(\mathbf{x}) = L_d(\mathbf{x}) + L_s(\mathbf{x}) \quad (1)$$

where L_d is the direct radiance from the source (Sec. 4), and

L_s is the radiance from the source which is scattered from other directions onto $\mathbf{X}$ (Sec. 6 and Fig. 2b).

Next, we express the radiance arriving at the camera as the sum of three terms:

$$L(\mathbf{x}) = L_o(\mathbf{x})e^{-\sigma\|\mathbf{X}\|} + L_b(\mathbf{x}) + L_c(\mathbf{x}) \quad (2)$$

where L_o is the light reflected by the surface point $\mathbf{X}$ which arrives at the camera without undergoing scattering. Note that it is attenuated by $e^{-\sigma\|\mathbf{X}\|}$ where σ is the extinction coefficient. L_b is composed of rays of light emitted by the source that are scattered into $\mathbf{x}$'s line of sight before hitting the surface (Sec. 5 and Fig. 2a). This term is known as backscatter. Finally, L_c is composed of rays of light reflected by other points on the surface that are scattered into pixel $\mathbf{x}$'s line of sight (Sec. 7 and Fig. 2c).

In order to write analytic expressions for these terms and derive a simple solution we make **two assumptions**. First, the surface is Lambertian with a spatially varying albedo $\rho(\mathbf{X})$. Second, we assume that surface variations in height are small compared to object distance from the camera. We call this the *small surface variations approximation* and note that it is weaker than the common distant light sources and orthographic projection approximations. Let $\bar{Z}$ be the average Z coordinate of the surface (assumed to be known). Then, the approximation claims that for every point on the surface: $|Z(\mathbf{X}) - \bar{Z}| \ll \bar{Z}, \forall \mathbf{X}$.

The approximation results in a weak perspective such that the projection $\mathbf{x}$ of $\mathbf{X}$ in the image plane is given by

$$\mathbf{x} = \left(f \frac{X}{\bar{Z}}, f \frac{Y}{\bar{Z}} \right)^t ; \quad \mathbf{X} = \left(\frac{\bar{Z}}{f} x, \frac{\bar{Z}}{f} y, \bar{Z} \right)^t, \quad (3)$$

where f is the known focal length. Note that for a given pixel, since we know its (x, y) coordinates and the average object distance $\bar{Z}$, the world coordinates $\mathbf{X}$ are known. Specifically, $\mathbf{D}(\mathbf{X})$ is independent of the unknown object height Z but still depends on X and Y , whereas in the distant light sources approximation $\mathbf{D}(\mathbf{X})$ is a constant.

Outline of Our Method

Given an input image L we eliminate the backscatter L_b by one of two methods. The first follows [32]: backscatter from each light source is measured by imaging it with no objects in the scene, and then the measured backscatter is

subtracted from the input images. In the second, backscatter is optically eliminated using fluorescence as we explain in Sec. 8. Once backscatter is removed, the resulting images are deblurred, using a calibrated PSF, to recover L_o (Sec. 7, Eq. 18). Next we write L_o as a linear equation between the unknown surface normals, albedo and an equivalent light source (Sec. 6, Eq. 9), which we approximate as an effective point source in a purely absorbing medium with effective extinction coefficient (Sec. 6, Eq. 11). With a minimum of 3 images under distinct light locations the normals can be solved for, as in conventional photometric stereo. The normals are then integrated to recover a smooth surface.

4. Direct Radiance

First, consider the direct reflected radiance from a Lambertian surface[34]:

$$L_d(\mathbf{x}) = I(\mathbf{X}) \frac{\rho(\mathbf{X})}{\pi} \hat{\mathbf{D}}(\mathbf{X}) \cdot \hat{\mathbf{N}} , \quad (4)$$

where $\hat{\mathbf{N}}$ is the unit surface normal and $\hat{\mathbf{D}}$ is the normalized source-to-object vector. The radiance on the object surface $I(\mathbf{X})$ depends on the radiant intensity I_0 of the source in direction¹ $(-\hat{\mathbf{D}})$:

$$I(\mathbf{X}) = \left(I_0(-\hat{\mathbf{D}}(\mathbf{X})) e^{-\sigma \|\mathbf{D}(\mathbf{X})\|} \right) / \|\mathbf{D}(\mathbf{X})\|^2 . \quad (5)$$

Eq. 5 accounts for nearby angularly-varying sources, exponential attenuation along the optical path length with extinction coefficient σ , and inverse-square distance falloff.

5. Backscatter

Light is scattered as it travels through a medium. The fraction of light scattered to each direction is determined by the phase function $P(\alpha)$, where $\alpha \in [0, 2\pi]$ is the angle between the original ray direction and scattered ray, and β is the scattering coefficient.

Light which is scattered directly into the camera by the medium without reaching the object is termed *backscatter* and is given by [19, 25] (Fig. 2a):

$$L_b(\mathbf{x}) = \beta \int_0^{\|\mathbf{X}\|} I(r\hat{\mathbf{X}}) P(\alpha) e^{-\sigma r} dr \quad (6)$$

The integration variable r is the distance from the camera to the imaged object point $\mathbf{X}$ along the line of sight (LOS), that is a unit direction $\hat{\mathbf{X}}$. The scattering angle α is given by $\cos(\alpha) = \hat{\mathbf{D}} \cdot \hat{\mathbf{X}}$ (recall that $\mathbf{D}$ is the direction to the light) and $I(r\hat{\mathbf{X}})$ is the direct radiance of the source at point $r\hat{\mathbf{X}}$ as defined in Eq. 5.

Note that for the *small surface variations approximation*, $\mathbf{X}$ and hence the limits of the integral are known for a given pixel. Therefore, the backscatter does not depend on the unknown height of the object and is a (different) constant for each pixel, similar to Tsitsios et al. [32].

¹the direction is negative as we consider outgoing rays from the source.

6. Single Scattered Source Radiance

Because of the medium, light rays that are not originally pointed at an object point may be scattered and reach it from the entire hemisphere of directions Ω (Fig. 2b), termed forward scattered radiance

$$L_s(\mathbf{x}) = \frac{\rho(\mathbf{X})}{\pi} \int_{\omega \in \Omega} L_i(\omega) (\omega \cdot \hat{\mathbf{N}}) d\omega . \quad (7)$$

where $L_i(\omega)$ is the total radiance scattered into the direction ω and is given by

$$L_i(\omega) = \beta \int_{t=0}^{\infty} I(\mathbf{X} + t\omega) P(\alpha) e^{-\sigma t} dt , \quad (8)$$

where t is the distance from the object, and the angle α is given by $\cos(\alpha) = \hat{\mathbf{D}}(\mathbf{X} + t\omega) \cdot \omega$. Note that $\mathbf{D}$ is the direction of the integration point to the light.

Substituting Eqs. 4,7 into Eq. 1 and rearranging yields:

$$L_o(\mathbf{X}) = L_d(\mathbf{X}) + L_s(\mathbf{X}) = \frac{\rho(\mathbf{X})}{\pi} L^{\text{eq}}(\mathbf{X}) \cdot \hat{\mathbf{N}} , \quad (9)$$

where

$$L^{\text{eq}}(\mathbf{X}) = I(\mathbf{X}) \hat{\mathbf{D}}(\mathbf{X}) + \int_{\omega \in \Omega} L_i(\omega) \omega d\omega . \quad (10)$$

Here, the direct light as well as the integrated scattered contributions can be thought of as an equivalent distant light source not in a medium, although, this equivalent source may be different (in direction and magnitude) for each surface point, and thus is not a true distant source. Nevertheless, Eq. 9 gives a linear equation for the unknown normals and albedo, which can be solved for given a minimum of 3 images under distinct light locations.

Unfortunately, evaluating Eq. 10 requires careful calibration of the scattering parameters β , σ and $P(\alpha)$ which can be difficult. Instead, we next show through simulations, that for a wide variety of media, $L^{\text{eq}}(\mathbf{X})$ can be approximated as an effective point source in a purely absorbing medium with effective extinction coefficient. This allows for simple calibration.

Effective Point Source Simulations

We approximate $L^{\text{eq}}(\mathbf{X})$ as:

$$\tilde{L}^{\text{eq}}(\mathbf{X}) \approx \frac{\kappa I_0(-\hat{\mathbf{D}}(\mathbf{X})) e^{-\tilde{\sigma} \|\mathbf{D}(\mathbf{X})\|}}{\|\mathbf{D}(\mathbf{X})\|^2} \hat{\mathbf{D}}(\mathbf{X}) , \quad (11)$$

where the effective source has the same position $\mathbf{S}$ and intensity distribution I_0 as the real source, but is scaled by κ , and the effective medium has extinction coefficient $\tilde{\sigma}$. Note that κ is a global brightness scale, which is the same for all the lights, and thus does not need to be explicitly calibrated, as it cancels out in the normal estimation.

For our simulations we used an isotropic point source at a distance d from a Lambertian surface patch with an angle ϕ

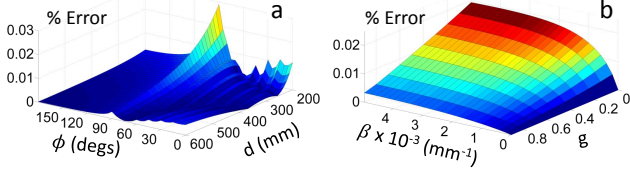


Figure 3. a) The relative error between $L_o(d, \phi)$ and $\tilde{L}_o(d, \phi)$ for $g = 0.8$ and $\beta = 0.0026$. Note that the spike only reaches 3% and is located at $\phi = 90^\circ$ where $\tilde{L}_o = 0$ due to shadowing. ϕ near 90° and above is not usually relevant for photometric stereo. b) The mean relative error between L_o and $\tilde{L}_o$ for $\beta \in [0, 0.005]\text{mm}^{-1}$ and $g \in [0, 0.9]$. The approximation errors are small.

between the surface normal and light direction. For the scattering function, we used the common Henyey-Greenstein phase function [8] that has a single parameter of the medium $g \in [-1, 1]$. In water, g is usually between $0.7 - 0.9$ [19].

We compute $L_o(d, \phi)$ for $d \in [200, 600]\text{mm}$, $\phi \in [0, \pi]$, for a variety of media given by $\beta \in [0, 0.005]\text{mm}^{-1}$ and $g \in [0, 0.9]$. Note that we choose I_0 such that $L_o(200, 0)$ is normalized to 1. To reduce the number of parameters we set $\sigma = \beta$, which does not influence the analysis². For each parameter pair g, β we compute the approximation parameters κ and $\tilde{\sigma}$ by minimizing:

$$\min_{\kappa, \tilde{\sigma}} \sum_{d, \phi} |L_o(d, \phi) - \tilde{L}_o(d, \phi)|^2. \quad (12)$$

As an example, Fig. 3a depicts the relative error, $\text{RE}(d, \phi) = |L_o(d, \phi) - \tilde{L}_o(d, \phi)|$, for $g = 0.8, \beta = 0.0026$, that fit our setup. Figure 3b depicts the mean relative error (MRE) for a range of values $\beta \in [0, 0.005]\text{mm}^{-1}$ and $g \in [0, 0.9]$, where the mean is taken over d and ϕ . The MRE was less than 2% across all medium conditions, justifying our approximation.

7. Single Scatter Object Blur

Similar to the light source blur, radiance from the object is also blurred while it propagates to the camera (Fig. 2c). As we demonstrate, this effect deteriorates the performance of photometric stereo, although it has been neglected in previous works [25].

The contribution of object blur to the pixel intensity is computed by integrating light scattered into the LOS of $\mathbf{X}$ from all other points on the surface:

$$L_c(\mathbf{x}) = \beta \int_{r=0}^{\|\mathbf{x}\|} \int_{\omega \in \Omega} L_o(\mathbf{X}') P(\alpha) e^{-\sigma(t+r)} d\omega dr. \quad (13)$$

Here r is the distance along the LOS, $\mathbf{X}'$ is the object surface point intersected by the ray starting at point $r\mathbf{X}$ in direction ω . Its radiance is $L_o(\mathbf{X}')$ and its distance to the scatter point in the LOS is given by $t = \|r\hat{\mathbf{X}} - \mathbf{X}'\|$ with scattering angle $\cos \alpha = \omega \cdot (-\hat{\mathbf{X}})$.

²In general $\beta \leq \sigma$, but since β purely scales L_s , a smaller value of beta would make L_o closer to L_d and thus $\tilde{L}_o$ would be an even better fit than we calculated.

We now show that L_o can be recovered from $L_o e^{-\sigma\|\mathbf{x}\|} + L_c$ by deconvolution with a constant PSF.

Deblurring Object Scatter

First we rewrite Eq. 13 to integrate over the area of the object surface $dA = d\omega \cdot t^2 / \cos \theta$ instead of solid angle $d\omega$, where t is the distance from $\mathbf{X}'$ to the scattering event, and θ is the angle between the normal at $\mathbf{X}'$ and the ray of light before scattering. Eq. 13 now becomes

$$L_c(\mathbf{x}) = \beta \int_{\mathbf{X}'} L_o(\mathbf{X}') \int_0^{\|\mathbf{x}\|} P(\alpha) e^{-\sigma(t+r)} \frac{\cos \theta}{t^2} dr dA. \quad (14)$$

Now we define the scattering kernel

$$K(\mathbf{X}, \mathbf{X}') = \delta(\mathbf{X} - \mathbf{X}') e^{-\sigma\|\mathbf{x}\|} + \beta \int_0^{\|\mathbf{x}\|} P(\alpha) e^{-\sigma(t+r)} \frac{\cos \theta}{t^2} dr, \quad (15)$$

where $\delta(\mathbf{X} - \mathbf{X}')$ is the Dirac delta function. Now,

$$L_o(\mathbf{x}) e^{-\sigma\|\mathbf{x}\|} + L_c(\mathbf{x}) = \int_{\mathbf{X}'} K(\mathbf{X}, \mathbf{X}') L_o(\mathbf{X}') dA(\mathbf{X}') . \quad (16)$$

In general, the kernel K , depends on $\mathbf{X}, \mathbf{X}'$ and the unknown normals $\hat{\mathbf{N}}'$. For an orthographic camera viewing a plane at constant depth, K is shift invariant and Eq. 16 can be written as a convolution with a PSF. Motivated by this, we found empirically that for a given $\tilde{Z}$ it is approximately shift invariant (and rotationally symmetric).

Denoting the PSF as h , we get

$$L_o(\mathbf{x}) e^{-\sigma\|\mathbf{x}\|} + L_c(\mathbf{x}) \approx h * L_o. \quad (17)$$

We emphasize here that we have shown that under a single scattering model, the forward scatter from the object can be written as an integral transform with kernel K . This justifies approximating the forward scattering as a PSF which is not obvious in the form of Eq. 13.

We solve Eq. 17 by writing the image as a column vector and representing the convolution as a matrix operation

$$L - L_b = H L_o \quad (18)$$

where we have substituted the backscatter compensated image $L - L_b$ for $L_o e^{-\sigma\|\mathbf{x}\|} + L_c$, and H is the matrix representation of h . Here H is a large nonsparse matrix and thus storing it in memory and directly inverting it is infeasible. Instead we solve the linear system of Eq. 18 using conjugate gradient descent. This requires only the matrix vector operation which can be computed as a convolution and implemented using a Fast Fourier Transform (FFT).

8. Backscatter Removal Using Fluorescence

While we are able to subtract the backscatter component, it is an additive component that effectively reduces

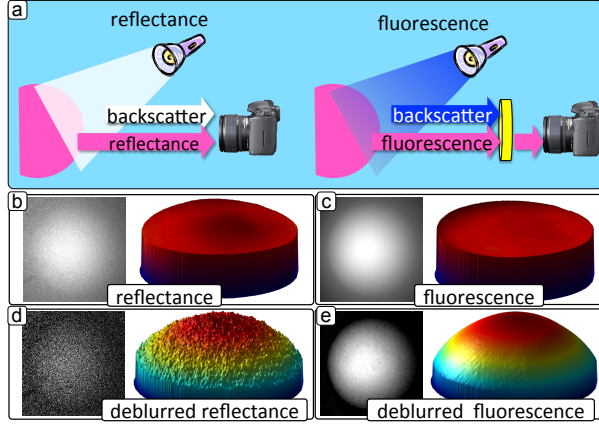


Figure 4. a) Backscatter is caused by light that is scattered into the camera by a medium, before it reaches the object and has the same color as the illumination. Thus, the barrier filter used to block fluorescence excitation also blocks backscatter, while imaging the signal from the object. We use this property to remove backscatter in input images. b) Photometric stereo reconstruction of a fluorescent sphere using backscatter subtracted reflectance images. One of the input images is shown on the left, with visible noise and blur. Blur in the input images flattens the reconstruction. c) Looking at fluorescence images as an input, the backscatter is eliminated while maintaining a higher SNR. However the blur still flattens the reconstruction. d) Deblurring the backscatter subtracted images recovers the general shape but suffers from noise as seen by the spiky surface. e) Deblurring the fluorescence results in the correct shape with much less noise.

the dynamic range of the signal from the object, degrades the image quality and reduces SNR [30]. As such it is beneficial to optically remove it when imaging. Here, we use the observation that for fluorescence images taken with non-overlapping excitation and emission filters, there is no backscatter in the image (Fig. 4a). In fluorescence imaging, the signal of interest is composed of wavelengths that are longer than that of the illumination, and a barrier filter on the camera is used to block the reflected light. The backscatter is composed of light scattered by the medium *before* it reaches the object. Thus, the backscatter has the same spectral distribution as the light source, which is blocked by the barrier filter on the camera. This insight enables imaging without loss of dynamic range even in highly turbid media. Compared to a backscatter subtracted reflectance image, a fluorescence image has less noise (Fig. 4b,c). This difference becomes even more apparent after deconvolution (Fig. 4d,e).

In addition, in [23, 28] it was shown that the fluorescence emission acts as a Lambertian surface in photometric reconstructions. Thus, imaging fluorescence has an additional advantage as it relaxes the need for a Lambertian surface.

In the development of our algorithm we assumed a single set of medium parameters β , σ and $P(\alpha)$. However these quantities are in general wavelength dependent. In reflectance imaging, the wavelength of the light is the same on both pathways: light to object, and object to camera. How-

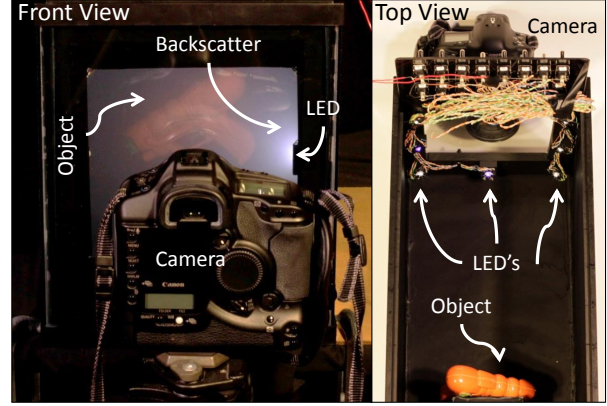


Figure 5. [Left] Our experimental setup consists of a camera looking through a glass port into a tank. [Right] 8 LEDs are mounted inside the tank around the camera port illuminating the object placed at the back of the tank.

ever, in fluorescence imaging they are different. Nevertheless, the only parameters that require calibration in our solution are the effective extinction coefficient $\tilde{\sigma}$ and the PSF. The parameter $\tilde{\sigma}$ is estimated for the excitation wavelength and the PSF is estimated for the emission wavelength, and as such we do not need to calibrate any extra parameters in the case of fluorescence imaging.

9. Implementation

9.1. Experimental Setup

Our setup is shown in Fig. 5. We used a Canon 1D camera with a 28mm lens placed 2cm away from a 10 gallon glass aquarium. All sides except the front (where the camera looks in) were painted black to reduce reflection. In addition, a black panel was suspended just below the surface of the water to remove reflection from the air-water interface. The objects were placed at an average distance of 40cm from the front of the tank. For point illumination we used Cree XML - RGBW Star LEDs. The LEDs were water proofed by coating the electrical terminals with epoxy. Reflection images were taken under white illumination while fluorescent images were taken under blue illumination with a Tiffen #12 emission filter on the camera. We used tap water, and the turbidity was increased using a mixture of whole milk and grape juice (milk is nearly purely scattering, while grape juice is nearly purely absorbing and thus by mixing them we can achieve a variety of scattering conditions [19]). The LEDs were mounted inside the tank on a square around the camera, four on the corners and four on the edges. Their positions were measured. Images were acquired in Raw mode which is linear and the normal integration was done using the method of [1].

9.2. Geometric and Radiometric Calibration

Images of a checkerboard (in clear water) were used to calibrate the intrinsic camera parameters (implicitly accounting for refraction) [2]. The location of each light was

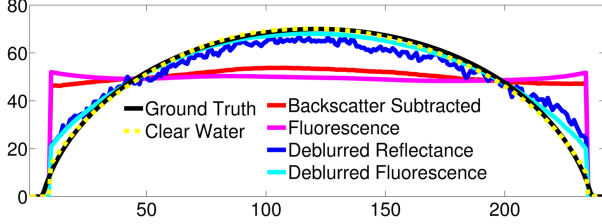


Figure 6. Cross-sections of the spherical cap reconstruction in turbid medium using various methods compared to ground truth. The clear water reconstruction resembles the ground truth. Only correcting for the backscatter (by subtraction or fluorescence) yields flattened results. Deblurring the backscatter subtracted images recovers the shape but is degraded by noise (the surface is jagged). Deblurring the fluorescence images produces the best results.

measured using a ruler, and transformed to the camera reference frame. To calibrate each light’s angular intensity distribution we imaged a matte painted (assumed to be Lambertian) plane at a known position under illumination from each light in clear water. Using Eq. 4, the known geometry, and $\sigma = 0$, we compute $I_0(-\hat{\mathbf{D}})$, the angular dependence of the light source.

9.3. Calibration of Medium Parameters

The backscatter component is measured using the calibration method of [32]. For each light an image is captured with no object in the field-of-view and subsequently subtracted from future reflectance images. This is not used when imaging fluorescence.

The PSF is estimated using a calibration target similar to [13]. We use a matte painted checkerboard which is imaged with its axis aligned to the image plane at the approximate depth of the objects we plan to reconstruct. As the PSF is rotationally symmetric its parameters are the values along a radius $[h_0, \dots, h_s]$, where h_0 is the center value and h_s is the value on the support radius s . The PSF and the effective extinction coefficient $\tilde{\sigma}$ are estimated by optimizing

$$\min_{\tilde{\sigma}} \min_{h_0 \dots h_s} \sum_{\mathbf{x}} \|h * L_o(\mathbf{x}, \tilde{\sigma}) - (L - L_b)\| \quad (19)$$

Note that the PSF is not normalized due to loss of energy (attenuation) from the object to the camera. L_o is computed using the calibrated lights, known geometry, and registering the checkerboard albedo, measured in clear water to the image in turbid water. The inner optimization is an over-determined linear system holding $\tilde{\sigma}$ fixed. We sweep over the values of $\tilde{\sigma}$ and choose the one with the minimum error.

10. Results

We imaged four objects: a spherical cap, a plastic toy squirt gun, a plastic toy lobster, and a fluorescent painted mask in clear water as well as four increasing turbidities (Figs. 1, 4, 6-8). Each turbidity level corresponded to adding 1.25ml of milk to the tank. For the spherical cap and

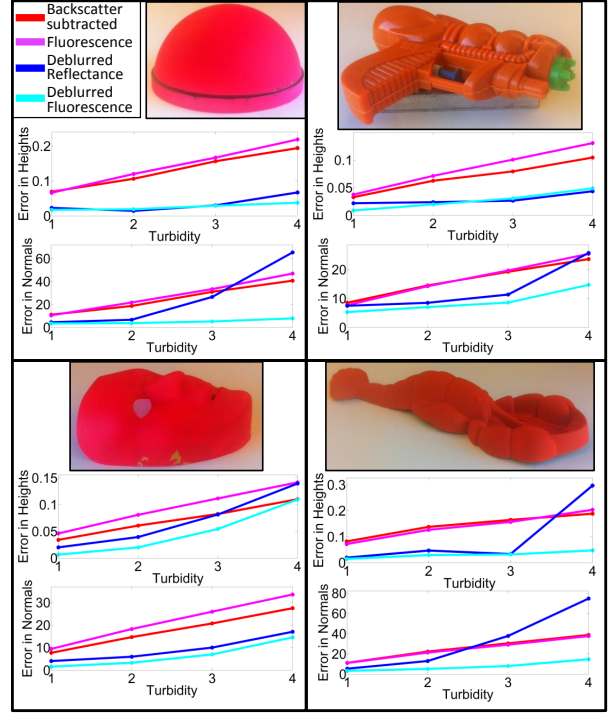


Figure 7. Errors in the reconstructions of four objects as a function of turbidity, compared to clear water reconstruction. Top rows are average percent errors in heights and bottom rows are average angular errors in normals. Removing backscatter by either subtraction or using fluorescence performs similarly. Deblurring the backscatter compensated images significantly improves the reconstructions. In high turbidity where the backscatter is strong compared to the object signal deblurring the backscatter subtracted images degrades due to noise, while deblurring the fluorescence suffers less, as the fluorescence images have a higher SNR.

the lobster, we also added 15ml of grape juice per turbidity level to increase absorption.

We employ two error metrics to evaluate the quality of our reconstructions: The mean absolute difference in heights ($\text{Err } Z = \text{mean}(Z - Z_{gt})$) and the mean angular error in the normals ($\text{Err } N = \text{mean}(\text{acos}(N \cdot N_{gt}))$), where Z_{gt} and N_{gt} are the ground truth heights and normals. Note that during integration, random noise in the normals cancels out locally, resulting in reconstructions with the correct overall shape, but with rough surfaces. As such $\text{Err } Z$ captures systematic errors that affect the overall shape, but is less sensitive to noise in the normals.

We see that the reconstructed spherical cap in clear water nearly perfectly matches the ground truth (Fig. 6) with an $\text{Err } Z$ of 1.4% and $\text{Err } N$ of 3° . This justifies our use of clear water reconstructions as ground truth for the other objects where true ground truth is not available.

Fig. 1 depicts the input images and resulting reconstructions for the toy gun, and Figs. 4, 6 depict results on the sphere, both in the highest turbidity level. Fig. 8 depicts images and reconstructions for the toy lobster for turbidity levels 2-4 and the mask in turbidity level 4. In all

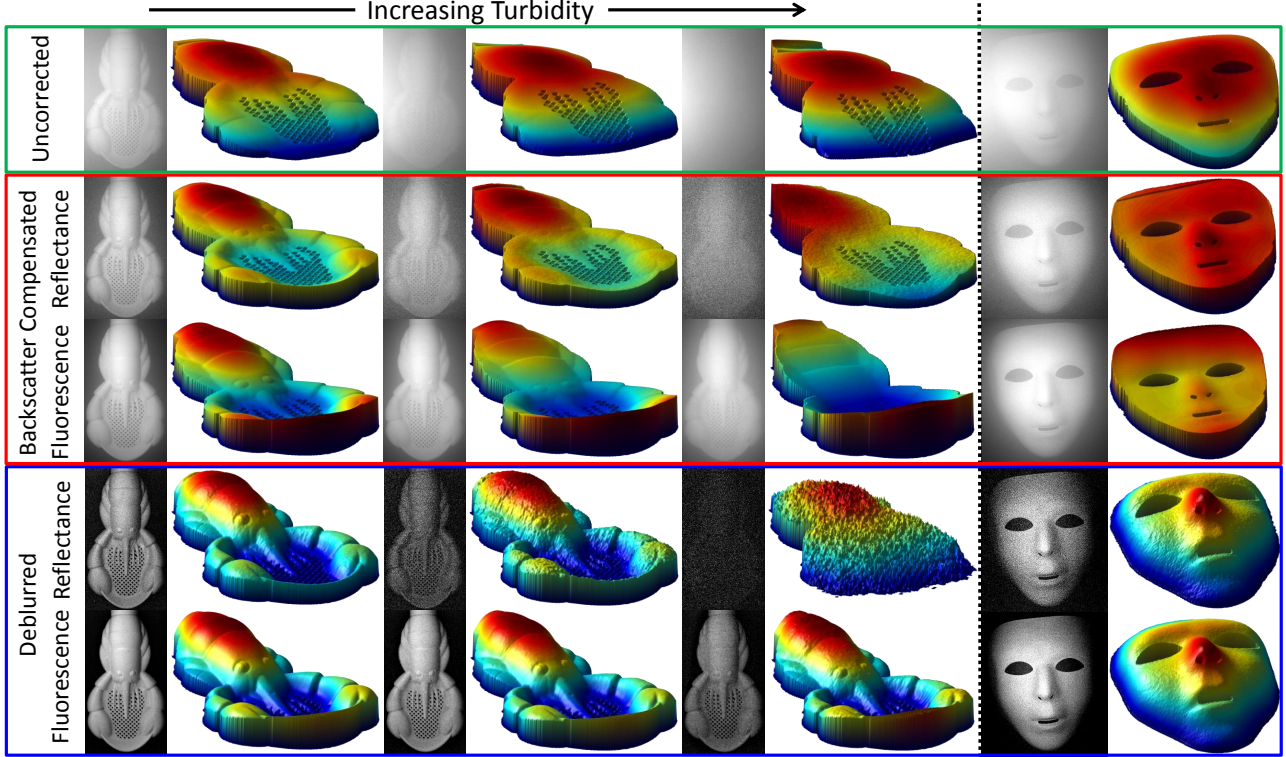


Figure 8. Surface reconstruction of the toy lobster and mask. The first three columns depict three levels of increasing turbidity from left to right. The fourth column depicts the mask in turbidity 4. [1st row] result of standard photometric stereo (scattering is ignored). The shape is not reconstructed correctly. [2nd row] Result of removing the backscatter as in [32]. The reconstruction is improved but still unsatisfactory. [3rd row] Using fluorescence to remove backscatter. The result is slightly better than backscatter subtraction. [4th row] result of deblurring the backscatter subtracted images. This recovers the shape quite well when the SNR is not too low. However this is not the case in high turbidity. [5th row] result of deblurring the fluorescence images. Here the SNR remains high even in high turbidity and thus we continue to get excellent quality reconstructions. Note the roughness on the fourth row, second column due to noise.

results, reconstructions from uncorrected images are flattened. Removing backscatter, either by backscatter subtraction (current state-of-the-art [32]), or using fluorescence, but without handling blur, also produces flattened results. For lower turbidities deblurring backscatter subtracted images produces excellent results, but in the highest turbidity, where the backscatter dominates the signal, using fluorescence reduces the noise and results in a smoother surface.

Quality of results as a function of turbidity level is demonstrated in Figs. 7, 8. The plots in Fig. 7 show how Err Z and Err N increase for each method as the turbidity increases, where the lowest error is achieved using the deblurred fluorescence images. In the highest turbidity level, the deblurred reflectance image often performs worse than all other methods, as the deblurring degrades with noise.

11. Conclusion

In this paper, we have developed a comprehensive and novel solution for photometric stereo in a scattering medium. We address each of the three key modes of single scattering, showing how a scattered light source can be modeled as an unscattered point light source, accounting for blur due to scattering from the object through a novel

deconvolution framework, and demonstrating how fluorescence imaging can optically eliminate backscatter, increasing SNR in high turbidity. With the simple *small surface variations approximation*, we reduce the problem to a linear system for the surface normals, almost identical to conventional photometric stereo. Our practical methods for deconvolution and fluorescence can be combined to produce reconstructions almost as accurate as those obtained in air, and significantly better than previous methods.

Future work includes removing the need to know the average object distance, the need to calibrate the PSF at that distance, as well as relaxing the *small surface variations approximation*. Although our theory only applies to a single scattering medium, in practice, our calibrated PSF may be taking multiple scattering effects into account. Extending our theory to multiple scattering would provide further insight.

Acknowledgments This work was supported in part by NSF grant ATM-0941760, ONR grant N00014-15-1-2013, W.M. Keck Foundation, and by the UC San Diego Center for Visual Computing. Tali Treibitz was an Awardee of the Weizmann Institute of Science – National Postdoctoral Award Program for Advancing Women in Science.

References

- [1] A. Agrawal, R. Raskar, and R. Chellappa. What is the range of surface reconstructions from a gradient field? In *Proc. ECCV*, pages 578–591. Springer, 2006. 6
- [2] J. Y. Bouguet. Camera calibration toolbox for matlab. vision.caltech.edu/bouguetj/calib_doc. 6
- [3] B. Dong, K. D. Moore, W. Zhang, and P. Peers. Scattering parameters and surface normals from homogeneous translucent materials using photometric stereo. In *Proc. IEEE CVPR*, pages 2299–2306, 2014. 3
- [4] M. Galo and C. L. Tozzi. Surface reconstruction using multiple light sources and perspective projection. In *Proc. IEEE ICIP*, volume 1, pages 309–312, 1996. 2
- [5] N. Gracias, P. Ridao, R. Garcia, J. Escartin, M. L’Hour, F. Cibecchini, R. Campos, M. Carreras, D. Ribas, N. Palomeras, et al. Mapping the moon: Using a lightweight auv to survey the site of the 17th century ship la lune. In *Proc. MTS/IEEE OCEANS*. IEEE, 2013. 1
- [6] G. Guilbault. *Practical fluorescence*, volume 3. CRC, 1990. 2
- [7] M. Gupta, S. Narasimhan, and Y. Schechner. On controlling light transport in poor visibility environments. In *Proc. IEEE CVPR*, 2008. 2
- [8] L. G. Henyey and J. L. Greenstein. Diffuse radiation in the galaxy. *The Astrophysical Journal*, 93:70–83, 1941. 5
- [9] M. Hullin, M. Fuchs, I. Ihrke, H. Seidel, and H. Lensch. Fluorescent immersion range scanning. *ACM Transactions on Graphics (TOG)*, 27(3):87–87, 2008. 2
- [10] C. Inoshita, Y. Mukaigawa, Y. Matsushita, and Y. Yagi. Surface normal deconvolution: Photometric stereo for optically thick translucent objects. In *Proc. ECCV*, pages 346–359. Springer, 2014. 3
- [11] Y. Iwahori, H. Sugie, and N. Ishii. Reconstructing shape from shading images under point light source illumination. In *Proc. IEEE ICPR*, volume 1, pages 83–87, 1990. 2
- [12] J. S. Jaffe. Computer modeling and the design of optimal underwater imaging systems. *IEEE J. Oceanic Engineering*, 15(2):101–111, 1990. 2, 3
- [13] N. Joshi, R. Szeliski, and D. Kriegman. PSF estimation using sharp edge prediction. In *Proc. IEEE CVPR*, pages 1–8. IEEE, 2008. 7
- [14] B. Kim and P. Burger. Depth and shape from shading using the photometric stereo method. *CVGIP: Image Understanding*, 54(3):416–427, 1991. 2
- [15] D. Kocak, F. Dalgleish, F. Caimi, and Y. Schechner. A focus on recent developments and trends in underwater imaging. *Marine Technology Society Journal*, 42(1):52–67, 2008. 1, 2
- [16] N. Kolagani, J. Fox, and D. Blidberg. Photometric stereo using point light sources. In *Proc. IEEE Int. Conf. Robotics and Automation*, pages 1759–1764, 1992. 1, 2
- [17] K. M. Lee, C. J. Kuo, et al. Shape from photometric ratio and stereo. *J. Vis. Comm. & Image Rep.*, 7(2):155–162, 1996. 2
- [18] B. McGlamery. A computer model for underwater camera systems. In *Ocean Optics VI*, pages 221–231. International Society for Optics and Photonics, 1980. 3
- [19] 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 Transactions on Graphics (TOG)*, 25(3):1003–1012, 2006. 4, 5, 6
- [20] S. G. Narasimhan, S. K. Nayar, B. Sun, and S. J. Koppal. Structured light in scattering media. In *Proc. IEEE ICCV*, pages 420–427, 2005. 2
- [21] S. Negahdaripour, H. Zhang, and X. Han. Investigation of photometric stereo method for 3-D shape recovery from underwater imagery. In *Proc. MTS/IEEE OCEANS*, volume 2, pages 1010–1017, 2002. 2, 3
- [22] R. Pintus, S. Podda, and M. Vanzi. An automatic alignment procedure for a four-source photometric stereo technique applied to scanning electron microscopy. *IEEE Trans, Instrumentation and Measurement*, 57(5):989–996, 2008. 1
- [23] I. Sato, T. Okabe, and Y. Sato. Bispectral photometric stereo based on fluorescence. In *Proc. IEEE CVPR*, 2012. 3, 6
- [24] A. Shashua. *Geometry and photometry in 3D visual recognition*. PhD thesis, Massachusetts Institute of Technology, 1992. 2
- [25] B. Sun, R. Ramamoorthi, S. G. Narasimhan, and S. K. Nayar. A practical analytic single scattering model for real time rendering. *ACM Transactions on Graphics (TOG)*, 24:1040–1049, 2005. 1, 4, 5
- [26] K. Tanaka, Y. Mukaigawa, H. Kubo, Y. Matsushita, and Y. Yagi. Recovering inner slices of translucent objects by multi-frequency illumination. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5464–5472, 2015. 3
- [27] A. Tankus and N. Kiryati. Photometric stereo under perspective projection. In *Proc. IEEE ICCV*, 2005. 2
- [28] T. Treibitz, Z. Murez, B. G. Mitchell, and D. Kriegman. Shape from fluorescence. In *Proc. ECCV*, 2012. 3, 6
- [29] T. Treibitz and Y. Y. Schechner. Active polarization descattering. *IEEE Trans. PAMI*, 31:385–399, 2009. 2
- [30] T. Treibitz and Y. Y. Schechner. Resolution loss without imaging blur. *JOSA A*, 29(8):1516–1528, 2012. 2, 6
- [31] E. Trucco and A. T. Olmos-Antillon. Self-tuning underwater image restoration. *IEEE J. Oceanic Engineering*, 31(2):511–519, 2006. 3
- [32] C. Tsotsios, M. E. Angelopoulou, T.-K. Kim, and A. J. Davison. Backscatter compensated photometric stereo with 3 sources. In *Proc. IEEE CVPR*, pages 2259–2266, 2014. 1, 2, 3, 4, 7, 8
- [33] V. V. Tuchin and V. Tuchin. *Tissue optics: light scattering methods and instruments for medical diagnosis*, volume 13. SPIE press Bellingham, 2007. 2
- [34] R. Woodham. Photometric method for determining surface orientation from multiple images. *Opt. Eng.*, 19(1):139–144, January 1980. 2, 4
- [35] S. Zhang and S. Negahdaripour. 3-D shape recovery of planar and curved surfaces from shading cues in underwater images. *IEEE J. Oceanic Engineering*, 27(1):100–116, 2002. 3