

Model effectiveness prediction and system adaptation for photometric stereo in murky water[☆]



C. Tsotsios^{a,*}, T.K. Kim^a, A.J. Davison^b, S.G. Narasimhan^c

^a Department of Electrical and Electronic Engineering, Imperial College London, SW72AZ, UK

^b Department of Computing, Imperial College London, SW72AZ, UK

^c Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213, USA

ARTICLE INFO

Article history:

Received 12 October 2015

Accepted 2 March 2016

Available online 9 March 2016

Keywords:

Underwater vision

Photometric stereo

Illumination

ABSTRACT

In murky water, the light interaction with the medium particles results in a complex image formation model that is hard to use effectively with a shape estimation framework like Photometric Stereo. All previous approaches have resorted to necessary model simplifications that were though used arbitrarily, without describing how their validity can be estimated in an unknown underwater situation. In this work, we evaluate the effectiveness of such simplified models and we show that this varies strongly with the imaging conditions. For this reason, we propose a novel framework that can predict the effectiveness of a photometric model when the scene is unknown. To achieve this we use a dynamic lighting framework where a robotic platform is able to probe the scene with varying light positions, and the respective change in estimated surface normals serves as a faithful proxy of the true reconstruction error. This creates important benefits over traditional Photometric Stereo frameworks, as our system can adapt some critical factors to an underwater scenario, such as the camera-scene distance and the light position or the photometric model, in order to minimize the reconstruction error. Our work is evaluated through both numerical simulations and real experiments for different distances, underwater visibilities and light source baselines.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Consider a robotic platform operating in a murky sub-sea environment, equipped with light sources and a camera to illuminate and image the scene in front. Imaging and scene understanding in this scenario is challenging for two reasons: (a) light from the sources is backscattered toward the camera reducing image contrast severely, and (b) light reaching the scene and reflected back to the camera is weak and results in dark (and noisy) scene appearance.

Feature-based methods such as Structure-from-Motion are effective in mapping large areas in clear water [14]. However, they fail to perform reliably in murky maritime environments due to the strong image degradation that de-features the captured images and dictates special post-processing [28].

Photometric approaches on the other hand attempt to model the cause of image degradation and develop algorithms for scene

reconstruction. However, the image formation model in murky water is complex and non-linear, making it hard to use effectively with a shape estimation approach such as Photometric Stereo (PS). For this reason all photometric approaches [15,21,23,38,39] have resorted to approximations to keep the problem tractable. For example, often the scene is assumed to be distant enough that inverse-square law can be ignored and that backscatter does not vary with distance, or the scene is assumed to be close enough that backscattering can be ignored.

The above photometric model simplifications are very effective when applied in the appropriate scenarios. But it is hard for a robotic platform exploring an unknown environment to know a priori which assumptions are valid. Blindly applying a model simplification is very likely to result in poor scene reconstruction. Fig. 1 shows the Photometric Stereo images of a barrel using an ROV in real murky port water, and the reconstruction results using the method of [38] as the vehicle was navigating toward the target. Being too far decreased the SNR severely as the backscatter dominated the dynamic range of the sensor and the reconstruction was poor. Being too close also yielded errors, since the photometric model neglected the strong non-uniform illumination on the scene. Since the scene was unknown, it was hard to predict which

[☆] This paper has been recommended for acceptance by Shahriar Negahdaripour.

* Corresponding author.

E-mail address: ctsotsios@imperial.ac.uk, chtsotsios@gmail.com (C. Tsotsios).

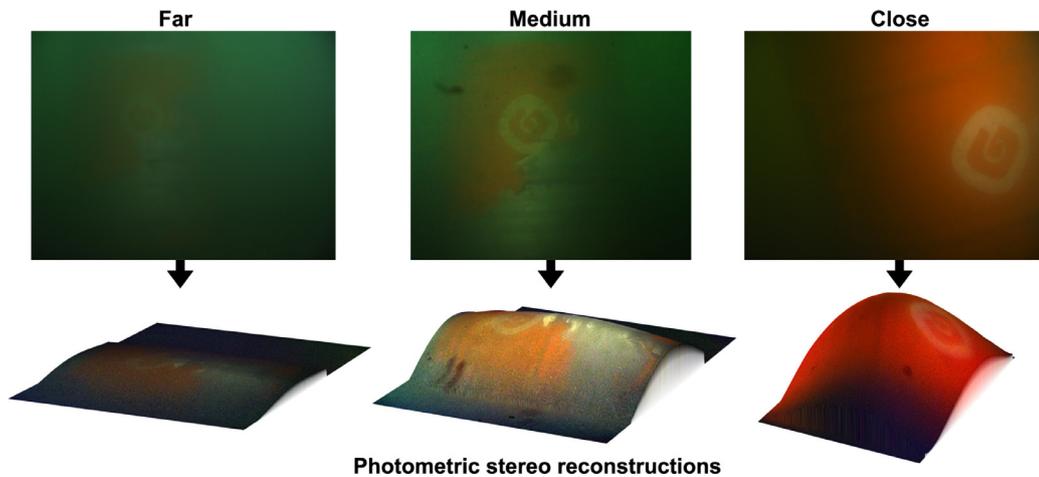


Fig. 1. The underwater images and the respective estimated Photometric Stereo reconstructions of a barrel object using an ROV in the port water of Leixões in Porto, Portugal. The reconstruction quality varied according to the distance. For large distances this was degraded due to the low camera SNR and for small distances it was also poor due to the photometric model invalidity [38]. Since the scene is unknown, it is hard to predict the optimal distance (middle reconstruction), or adopt automatically a more effective photometric model or light position.

distance was more effective, or if another photometric model or light source position could have been more successful.

In this work, we propose an effective approach for reasoning about the validity of such photometric models when the scene is unknown. To achieve this, we use a dynamic lighting framework where a robotic platform is able to probe the scene with varying light positions. This approach is based on a simple idea: if the photometric model is wrong for a particular scenario, the estimated surface normals will be erroneous and, more importantly, the error will vary significantly as the light source positions are varied with respect to the camera. On the other hand, if the photometric model is correct, the estimated surface normals will not vary as the source positions are varied.

In short, we obtain a faithful proxy for the true reconstruction error by estimating the change in surface normals under different light source positions. For example, when the source is close to the camera, backscatter is strong and any algorithm that ignores this produces worse shape estimates. But, as the source is moved away (even a short distance), the backscatter reduces [10] and the same algorithm produces better normal estimates. Our proposed dynamic lighting framework offers significant potentials to Photometric Stereo in murky water. The ability to approximate the reconstruction error can be used to adjust automatically: (a) the camera-scene distance, (b) the light position, and (c) the photometric model, in order to maximize the reconstruction quality.

We perform extensive numerical simulations where we mimic realistic scenarios underwater with different medium, distance, and system characteristics. Then, we present a real robotic platform in murky water navigating toward the scene of interest which can move the lighting fixtures along a mechanical arm. We demonstrate our system in the controlled environment of a big water tank, where the platform explores an unknown object for different distances, light positions and scattering levels, and we compare our results with the reconstruction from a depth sensor [24].

2. Related work

During the past years several approaches have shown how the effectiveness of Photometric Stereo can be extended by relaxing some of the limiting assumptions of the original method. [1,31] have shown how PS can be applied in uncontrolled environments where the scene is imaged by different cameras and variant

outdoor conditions. Specifically, internet images taken in different weather conditions were used to reconstruct tourism sites. In [25] and [18] the distant-lighting and orthographic projection assumptions were relaxed by modeling the effects of near-field illumination and perspective projection.

Various works have studied PS for the underwater case. The seminal work of [11] derived the image formation model in murky water, and since then many approaches showed how this can be simplified and optimized for the unknown orientation and albedo [15,20,21,23,38,39]. However, the resulting photometric models were used arbitrarily without examining their validity in different conditions. In our work we evaluate the effectiveness of such models considering various distances, scattering levels, light positions and sensor noise, and we propose a novel framework for predicting their effectiveness when the scene is unknown.

Some works emphasized the importance of the imaging system for the quality of the captured images in murky water. The impact of the vertical or horizontal displacement of the sources on image quality was investigated in [11,27], the use of polarizing filters [35,36] or the fusion of two images [37] were proposed to reduce the impact of backscatter, and the optimal separation between the camera and the sources in terms of image quality was calculated in [10] assuming that the scene is planar and the imaging conditions are known. Our work differs significantly from these works. First, because it tackles the problem of shape and albedo estimation using Photometric Stereo and not the improvement of visibility in murky water. Second, our work comprises an automatic approach that requires no prior knowledge about the scene.

In a sense our dynamic lighting system comprises an active approach. However, it could not be compared with active approaches in pure-air [5–7,17]. In these works, a specific photometric model was employed without further investigation. Our work provides the framework for evaluating such models automatically in murky water.

Our work also proposes a way for estimating the optimal light configuration for Photometric Stereo in murky water. Some works have examined the problem of finding the light position that is more robust to gaussian noise in pure air [3,8,16,32,33]. However, once again these works adopted a specific model (distant-lighting) without examining its validity. In our work, we take into account both the model validity and sensor noise in murky water, and we show that the optimal light position varies according to the scenario.

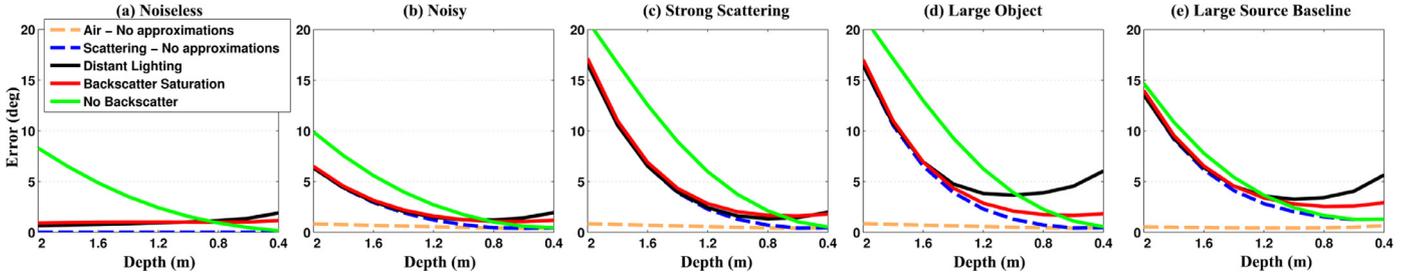


Fig. 3. Simulations showing the reconstruction error for various photometric models in different conditions. The error depends on the noise level, the scene depth, the scattering level, the size of the object and the light source baseline.

c of the medium. In this case, the direct component is non-linear and has more unknowns than its distant approximation. [15,23] solved the Near-Lighting problem using external hardware for estimating c and an iterative optimization algorithm. According to this, the Distant-Lighting approximation was initially considered in order to recover a first estimate of the normal map of the object. By integrating the normal map and having prior knowledge about the average camera-scene distance, the depth map of the object was estimated and the incident illumination on the object was recalculated taking into account the depth variation per pixel. Then, a refined version of the normal map was estimated. The algorithm was iterated until convergence.

Combining the photometric models for the direct and backscatter components yields potential optimization strategies. For example, for large camera-scene distances it was shown that Distant-Lighting & Backscatter Saturation model approximations yield a simple, linear solution that requires no prior knowledge about the medium coefficients [38]. For small camera-scene distances, Distant-Lighting might fail (depending on the object size) and the non-linear Near-Lighting model can be solved either assuming that the backscatter is saturated or neglected entirely.

In the next section, we investigate the factors that influence the effectiveness of such photometric models in murky water.

5. Photometric model effectiveness evaluation

Consider a Photometric Stereo system operating in murky water. Given the captured images, a version of the normal map is estimated by inverting the photometric model that is used to describe the image formation. We now examine how the validity of such models changes according to the imaging conditions.

In pure-air and controlled imaging conditions there are some rough rules for predicting when a photometric model is correct. For example, Distant-Lighting dictates that the scene depth is at least an order of magnitude larger than the object size. In murky water, the validity of Distant-Lighting approximation is harder to predict as it also depends on the total attenuation coefficient c of the medium. The camera Signal-to-Noise ratio (SNR) is another important factor in murky water since it is directly related with the effects of attenuation and backscattering [10]. We performed a large number of simulations in order to evaluate the effectiveness of different photometric approximations in murky water. In this case, the metric that indicates the model effectiveness is the difference between the estimated and the ground-truth normal map of the object.

We have numerically simulated the image formation underwater, a Photometric Stereo system with 4 sources, and a sphere object at different scene depths (the term depth in this work refers to the scene depth, i.e. distance, rather than the water depth). Fig. 3a shows the reconstruction error for the ideal case that the sensor is noiseless and has unlimited dynamic range. When no model approximations are made (only possible in simulations), the

reconstruction error is zero everywhere (blue dotted line) as would happen in pure air (orange dotted line). When a model approximation is used (red, green and black lines), some error appears in the estimated normals and this varies with distance according to the validity of the model's assumptions. At large distances for example, the Distant-Lighting model is very effective because the object size is small compared with the depth, while the No Backscatter model fails badly as it erroneously neglects the strong backscatter component. Thus, in the noiseless case the error varies only due to the level of model invalidity.

Consider now a realistic sensor with limited dynamic range and some noise. In this case the image formation model is rewritten as $E = D + B + N_S$, where N_S is additive noise. The Signal-to-Noise-Ratio for a captured image is

$$SNR \equiv \frac{D}{B + N_S}. \quad (5)$$

Unlike in pure air, the SNR is affected by the backscatter component which takes up part of the limited dynamic range of the sensor, and the strength of the direct component which is attenuated according to the scattering level of the medium.

Fig. 3b shows the error for the realistic-noisy case. The reconstruction in the scattering medium (blue dotted line) suffers from high error compared with pure air (orange dotted line) especially at large scene depths where the SNR is low. This affects all model approximations, the use of which introduces additional error. As some imaging characteristics change – for example the scattering level (Fig. 3c), the object size (Fig. 3d), or the light baseline (i.e. the distance of the sources from the camera – Fig. 3e) the reconstruction effectiveness for each photometric model changes as well.

Overall, the system effectiveness depends on several factors and hence using a photometric model arbitrarily can yield significant reconstruction error. At the same time, it is hard to predict automatically when the reconstruction is effective since the scene and the environment are normally unknown. In the next sections we describe how this can be achieved using a Photometric Stereo system with dynamic lighting.

6. Model effectiveness prediction using dynamic lighting

In a real underwater scenario the ground-truth normal map of the object is unknown, making the evaluation of a photometric model's effectiveness hard. We tackle this problem using a Photometric Stereo system with dynamic lighting. Consider a PS light baseline where a number of sources (at least 3) are placed around the camera. Combining the model approximations of Section 4, we can estimate a version of the normal vector \mathbf{n}_r at every scene point, where r denotes the baseline distance of the sources from the camera. Whether the ground-truth normal map \mathbf{n} was known, we could estimate the true reconstruction error (p-norm) as

$$\epsilon_r = \|\mathbf{n}_r - \mathbf{n}\|_p. \quad (6)$$

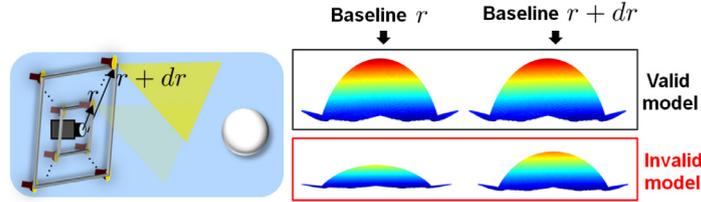


Fig. 4. Consider a Photometric Stereo system which can move the light sources by a small incremental step dr . This way, two versions of the normal map are estimated: one using the light baseline r , and one using $r + dr$. When the photometric model is valid, the estimated normal maps coincide as they match the ground-truth. Otherwise the estimated normals differ from the ground-truth as they absorb some error due to the model invalidity.

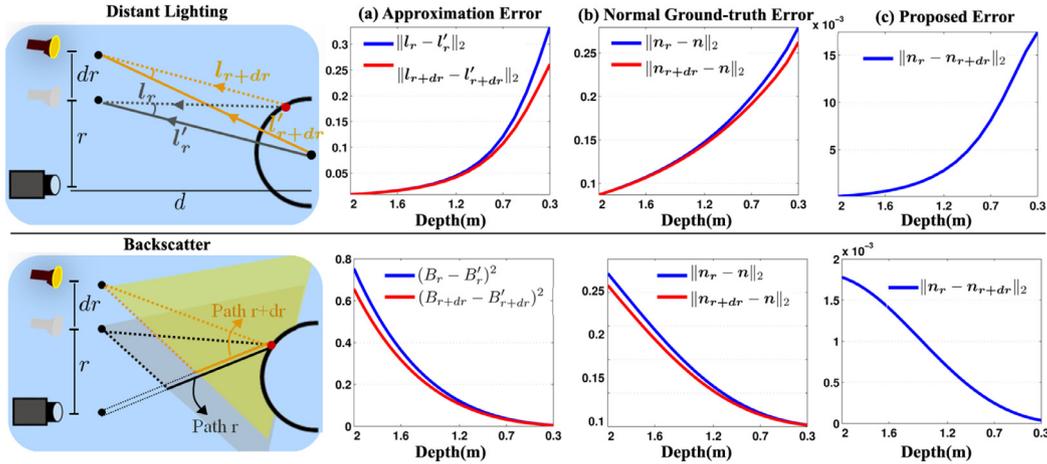


Fig. 5. Given that the source-object distance is large compared with the object size, Distant Lighting (top-row) assumes that the illumination vector \mathbf{l} can be approximated by the constant distant vector \mathbf{l}' . The level of agreement between these two (a) is reflected in the respective approximation error in the recovered normal (b). When we change the light position, the source-object distance is changed. This affects the distant lighting assumption and its respective approximation error. For a valid model, the error is small at both source positions and the recovered normals coincide. For an invalid model, the error differs significantly, and we see a similar difference in the estimated normals (c). Similarly, when we change the light position the backscatter component is changed (bottom-row), affecting the validity of the respective approximations.

The level of agreement between the two vectors depends on the validity of the photometric model. When the model is correct we expect the vectors to be equal or to vary insignificantly. Otherwise they will differ, as the error in modeling will be absorbed by the estimated normal \mathbf{n}_r .

Consider now a framework that allows the baseline to change by moving all sources by a small displacement dr (Fig. 4). Using the new baseline $r + dr$, we can recover a new version of the normal vector \mathbf{n}_{r+dr} , that corresponds to ground-truth error

$$\epsilon_{r+dr} = \|\mathbf{n}_{r+dr} - \mathbf{n}\|_p, \quad (7)$$

as in Eq. (6).

Since we are imaging the same object, the estimated normals using the two baselines should coincide and equal the ground-truth when the photometric model is valid. In this case, $\mathbf{n}_r \approx \mathbf{n}_{r+dr} \approx \mathbf{n}$. Nonetheless, when the model is invalid, the modeling error is evident in an estimated normal vector which differs from the ground-truth one. Our key observation is that this error differs with baseline, and thus we expect $\epsilon_r \neq \epsilon_{r+dr} \neq 0$ which leads to the estimation of a different normal for the same scene point $\mathbf{n}_r \neq \mathbf{n}_{r+dr} \neq \mathbf{n}$. This gives us the insight that the level of disagreement between the estimated normals for the same scene point using proximate baselines indicates the model validity:

$$\epsilon'_r = \|\mathbf{n}_{r+dr} - \mathbf{n}_r\|_p. \quad (8)$$

This requires no prior knowledge about the ground-truth object normals, but only two separate Photometric Stereo sessions using different light source positions.

6.1. Model approximation error

We now give the physical explanation of our proposed method considering different model approximations and noise due to low SNR. As described in the previous section we assume that all light sources are displaced by dr away from the camera (Fig. 5 shows only one source only for clarity).

Distant-Lighting assumes that the illumination vector from a light source is constant. Consider the case of Fig. 5-top left. When the source is at position r , the real illumination vector at a scene point will be \mathbf{l}_r . For Distant-Lighting, this is approximated by the vector \mathbf{l}'_r between the light source and the centroid of the object. This will cause some error to the estimated vector \mathbf{n}_r , according to how valid the approximation of \mathbf{l}_r is with \mathbf{l}'_r . Similarly, when the sources are moved by dr , \mathbf{l}_{r+dr} is approximated with \mathbf{l}'_{r+dr} , which causes error to \mathbf{n}_{r+dr} .

When the object size X is small compared with the source-object distance $\sqrt{r^2 + d^2} \gg X$, the Distant-Lighting model is valid. Then, the difference between the real illumination vector and the distant one is negligible $\mathbf{l}'_r \approx \mathbf{l}_r$, and $\mathbf{n}_r \approx \mathbf{n}$. If $\sqrt{r^2 + d^2} \gg X$, then $\sqrt{(r+dr)^2 + d^2} \gg X$ also, meaning that Distant-Lighting is also valid for the baseline $r + dr$. Thus $\mathbf{l}'_{r+dr} \approx \mathbf{l}_{r+dr}$ and $\mathbf{n}_{r+dr} \approx \mathbf{n}$. Then, the proposed error of Eq. (8) will be $\epsilon'_r = \|\mathbf{n}_r - \mathbf{n}_{r+dr}\|_p \approx 0$, indicating that the Distant-Lighting model is correct.

As the ratio between the source-object distance and the object size is decreased, the Distant-Lighting approximation becomes invalid. In this case, $\sqrt{r^2 + d^2} \sim X$, $\mathbf{l}'_r \neq \mathbf{l}_r$, and $\mathbf{n}_r \neq \mathbf{n}$. Since the source-object distance and the object size are now comparable, the displacement in source position dr is important. Specifically, $\frac{\sqrt{(r+dr)^2 + d^2}}{X} > \frac{\sqrt{r^2 + d^2}}{X}$, meaning that the error in approximating

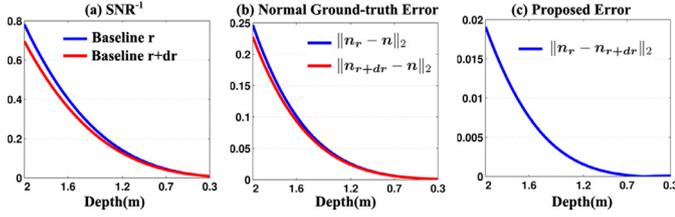


Fig. 6. The SNR of the captured images depends on the light position. Thus in the presence of SNR error the estimated normal maps differ from one light position to the other. When the SNR error is low (small scene depths where the direct component is strong) the estimated normals coincide.

\mathbf{l}_{r+dr} with \mathbf{l}_{r+dr} , will be smaller than approximating \mathbf{l}_r with \mathbf{l}_r , leading to $\mathbf{n}_r \neq \mathbf{n}_{r+dr}$.

Fig. 5a-top row shows the difference between the approximated and the real illumination vectors for light positions r and $r + dr$. At large scene depths, the vectors are well approximated with their distant versions for both baselines, and the estimated normals in both cases are close to ground-truth (Fig. 5b). As the scene depth decreases, a small change in light baseline dr has a big impact on the Distant-Lighting approximation, and the estimated normals using each baseline differ. Thus our proposed error metric (Fig. 5c) can be used for predicting the validity of the model approximation.

Backscatter Component: The change in source position similarly affects the validity of the backscatter component approximations. In this case though, the reflected error in the estimated normals is due to the erroneous approximation of the backscatter component B with B' , where $B' = 0$ for the No Backscatter assumption, or $B' = E_\infty$ for the Backscatter Saturation. Fig. 5-bottom left shows the effect of the light position on the backscatter. The backscatter integration path for the source at position r will be always greater than the respective one at $r + dr$, leading to $B_r > B_{r+dr}$ [10].

Consider the No Backscatter model approximation for a baseline r . This assumes that the backscatter component is negligible with respect to the strong direct component, and approximates B_r with $B'_r \simeq 0$. In this case, the estimated normal will absorb some error according to how valid the approximation of B_r is with B'_r (Fig. 5a-bottom row). This is valid for small scene depths. If $B_r \simeq B'_r \simeq 0$ though, then $B_{r+dr} \simeq 0$ as well, since $B_{r+dr} < B_r$. Hence when the model is valid the error is negligible for both baselines which recover normals close to ground-truth $\mathbf{n}_r \simeq \mathbf{n}_{r+dr} \simeq \mathbf{n}$ (Fig. 5b). For large scene depths where backscatter is not negligible, using the No Backscatter approximation will reflect significant error in the estimated normals. Since $B_{r+dr} < B_r$, the error that will affect \mathbf{n}_{r+dr} will be smaller than the error to \mathbf{n}_r . This leads to $\mathbf{n}_r \neq \mathbf{n}_{r+dr}$ (Fig. 5c).

SNR Error: Even when a model approximation is correct, the reconstruction might suffer from error due to low SNR (Section 5). In [10] it was described that this also varies with the light position since the position affects the direct and backscatter components and therefore the image SNR (Eq. (5)). Therefore, similarly with the model approximation errors, the error e_r in Eq. (6) for baseline r will be different than e_{r+dr} when SNR error is present, which leads to $\mathbf{n}_r \neq \mathbf{n}_{r+dr}$. In the absence of SNR error, the two normal maps would coincide.

Fig. 6a shows the impact of the noise to the useful direct component (expressed as SNR^{-1}). This is strong in large scene depths where the direct component is low. The fact that the SNR error depends on the light position reflects a respective change to the estimated normal maps that differ from the ground-truth (Fig. 6b) and they also differ from each other (Fig. 6c).

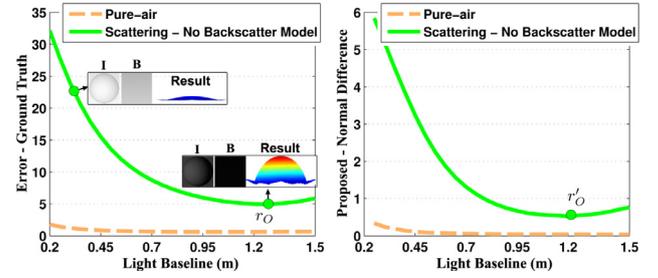


Fig. 7. As the light baseline is increased, the No Backscatter assumption becomes valid as the backscatter is decreased and thus the ground-truth reconstruction error is also decreasing (left graph). After the optimal light baseline the true reconstruction error is increasing again due to the low SNR. Our proposed error metric (right graph) approximates the ground-truth error and predicts the optimal baseline.

6.2. Optimal light baseline

We have explained that the change in the estimated normals under two source positions reflects the validity of the photometric model. Given that the sources can move to more than two positions, our proposed metric ϵ' (Eq. (8)) can also be used to compare the reconstruction effectiveness at different light baselines. Since the error changes according to the light position, there is an optimal light baseline r_O that minimizes the ground-truth error:

$$r_O = \arg \min_r \|\mathbf{n}_r - \mathbf{n}\|_p. \quad (9)$$

Consider Fig. 7, which corresponds to a system that can move the light sources by many small steps, creating an equal number of baselines (x-axis). Left graph indicates the ground-truth reconstruction error at every light source baseline. In the absence of any model approximation error in pure air, all baselines yield an insignificant error that changes due to sensor noise only (orange dotted line). Within scattering and when a model approximation is used, the baseline is crucial as it affects the reconstruction error. Here, the No Backscatter model is subject to large error for small baselines, since backscatter is strong there and neglecting it reflects large error to the estimated normals (Section 6.1). Increasing the baseline reduces the backscatter component and hence the error is decreased. After the optimal baseline r_O , the reconstruction error is increased again because then the distance between the sources and the object is too big and SNR is low [10].

Right graph corresponds to our error metric, which compares the estimated normal map of every two successive baselines ($r, r + dr$). In the pure air case and when no model approximation is used, the normal map changes insignificantly. When the model approximation is used, the reconstruction changes a lot when the model is erroneous and the noise is strong (small baselines), and exhibits the minimum change when the reconstruction is least subject to errors. Thus, we predict the optimal baseline by estimating

$$r'_O = \arg \min_r \|\mathbf{n}_r - \mathbf{n}_{r+dr}\|_p. \quad (10)$$

It is important to note here that the optimal light baseline for PS differs according to the scene depth, the photometric model and the murkiness level as Fig. 8 indicates. Hence, a Photometric Stereo system with fixed light sources has a decreased effectiveness compared with a dynamic system.

7. Results

7.1. Simulations

We have conducted a large number of numerical simulations considering different object distances and underwater visibilities in order to evaluate the effectiveness of our approach. Figs. 9 and

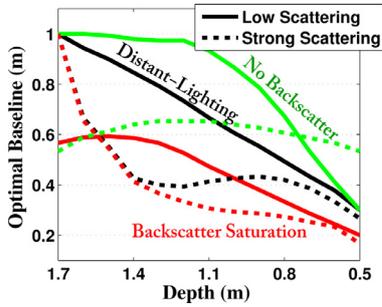


Fig. 8. The optimal light baseline differs according to the scene depth, the photometric model, and the murkiness level. In general, at bigger scene depths larger baselines are preferred (regardless of the photometric model) in order to minimize the backscatter that affects the SNR. At small scene depths, the backscatter is low anyway and smaller baselines are preferred to maximize the direct component. Then, each photometric model dictates a different baseline so that its assumptions become valid.

10 outline our framework, considering a scenario where a robotic platform navigates toward a sphere object.

At this part we considered three photometric models by combining different model approximations for the direct and backscatter components. These are the following: (a) Distant-Lighting and Backscatter Saturation, (b) Near-Lighting and Backscatter Saturation, and (c) Near-Lighting and No Backscatter. Section 4 contains further details about each photometric model.

We considered a large sphere object with a diameter of 40 cm so that the Distant-Lighting assumption is violated at small scene depths and the reconstruction error between the photometric models is significantly different. We used the values of $c = 0.1/b = 0.04$, $c = 0.4/b = 0.2$, and $c = 2/b = 1.8$ for the cases of low, medium and strong scattering respectively, according to [12]. At every scene depth the system varies the light baseline by moving all sources by small displacements of $dr = 2$ cm, within a range from 0.2–1.5 m from the camera. Then, using Eq. (8) it predicts the error for each model.

Fig. 9a shows the ground-truth reconstruction error for every model over the varying light baselines and 9b shows our proposed metric that can be estimated without ground-truth

Table 1
Normalized processing time per model.

Photometric model	Time
Distant+Saturation	1
Near+No Backscatter	125
Near+Saturation	170

information. We can notice that our method approximates effectively the ground-truth reconstruction error at all scene depths. It also indicates the optimal baseline r'_0 for every model, i.e. the baseline where the error is minimized.

Then, considering that each photometric model is characterized by the reconstruction error at its optimal baseline $\epsilon'_{r'_0}$ (i.e. that the system adapts a different light baseline for each model at every scene depth), we can predict how the effectiveness of every model changes with scene depth. Fig. 10 corresponds to three different scattering levels and outline the several advantages of our method. The dotted line in (a) corresponds to the ground-truth error per model and scene depth, after selecting the true optimal baseline r_0 using Eq. (9). The solid line, corresponds to the ground-truth error when our predicted optimal baseline r'_0 is selected using Eq. (10). These two coincide in almost all the different scattering, scene depth and model cases. Thus, our proposed metric can be used to adapt the light position to the scenario. The graphs at (b) show the final predicted model effectiveness per scene depth. We can see that this successfully approximates the ground-truth error of (a).

Our framework can also be used for comparing the effectiveness of different models and forming an improved optimization strategy in robotic underwater missions. Recall here that the processing capability of underwater vehicles is constrained by the limited power and computational resources [4]. Each photometric model has a different complexity and computational cost. For example Distant-Lighting+Backscatter Saturation corresponds to a linear solution, while the Near-Lighting models are non-linear and comprise additional unknowns. Table 1 shows an estimate of the processing time (normalized by the time of the fastest method) when each model is optimized using our simulations.

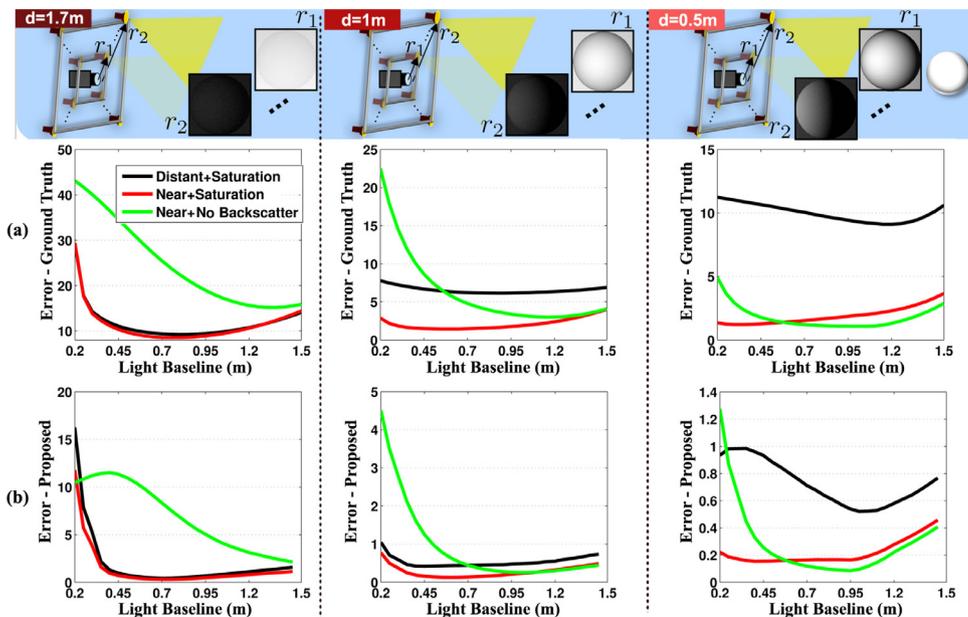


Fig. 9. Simulations of a realistic underwater scenario, where a robotic platform is carrying a dynamic lighting system, and it estimates several reconstructions of the sphere at every scene depth under the different light positions. (a) shows the ground-truth reconstruction error of every model. (b) shows the average difference in the estimated normal map as the light positions change.

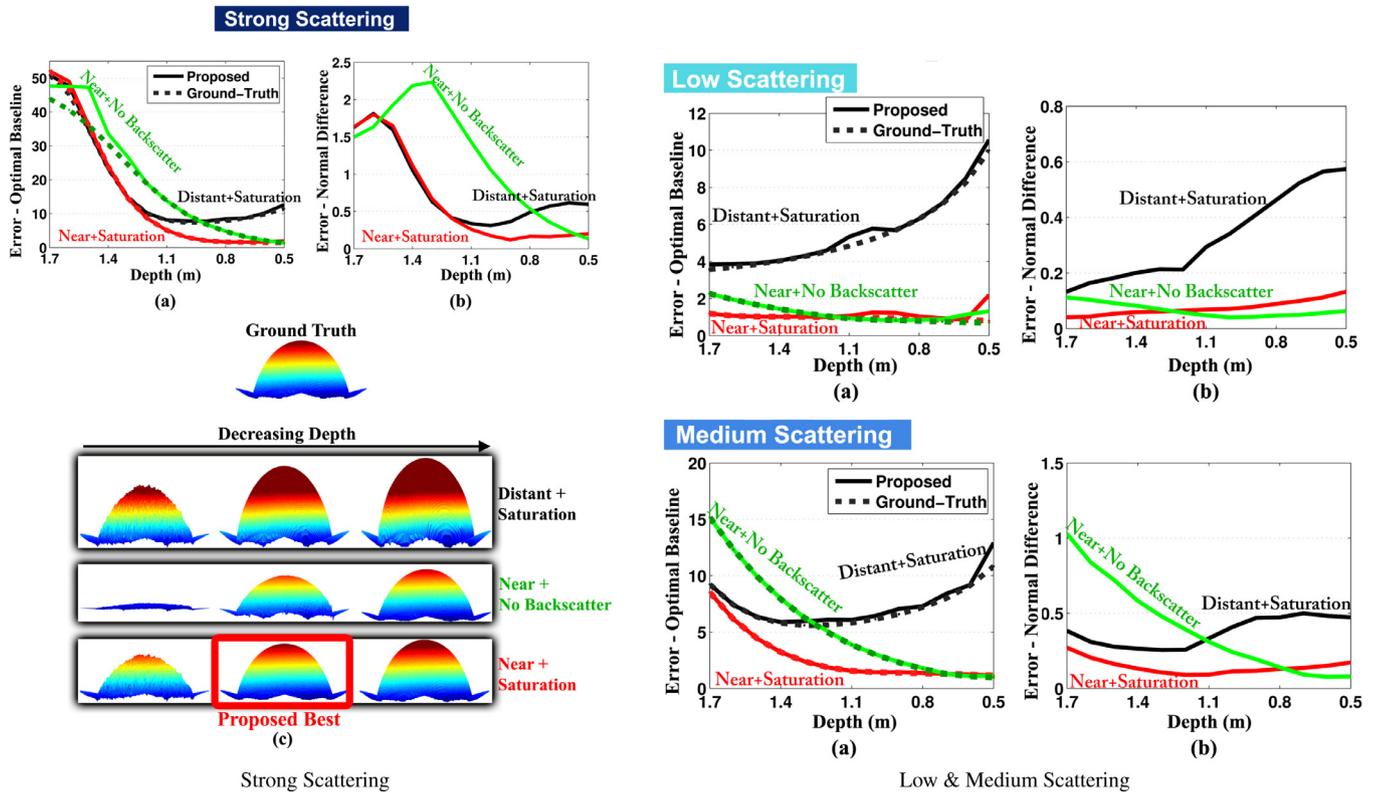


Fig. 10. Simulations showing the effectiveness of our proposed approach. (a) shows the reconstruction error per method, assuming that we have selected the ground-truth optimal light baseline per scene depth (dotted line) and the respective reconstruction error when we select the optimal baseline according to our proposed method (solid line). (b) shows our automatic estimation of the error per model and scene depth. (c) shows how the reconstructed sphere shape changes with scene depth, directly reflecting the effectiveness of the respective model. Finally, we can predict the best reconstruction, as the one that minimizes the change in normal map for all scene depths and models (red box). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

Approximating the reconstruction error per model can be used in order to adapt such models automatically according to the scenario. Consider for example that the Distant+Saturation model is used initially, due to its simplicity/low computational cost and its effectiveness in large camera-scene distances where all models are affected by the SNR error (Fig. 10). The robotic platform could predict that navigating towards the object decreases the reconstruction error until the scene depth where it starts increasing again because of the invalidity of the Distant-Lighting model. Then it could compare whether the more complex Near-Light models perform significantly or insignificantly better and adapt the one with the minimum error.

In Fig. 10(c) shows some instances of the reconstructed sphere shapes using different models per scene depth in the case of strong scattering. Notice how the reconstructed shapes change, in accordance with the respective ground-truth and approximated errors. As our framework can be used for approximating the level of error for every model at every scene depth, all possible reconstructions can be compared and the one that corresponds to the minimum predicted error can be selected as optimal (highlighted with red).

7.2. Discussion of other photometric factors

Our method is based on the observation that the change in estimated surface normals under different source positions is a faithful proxy of the true reconstruction error per model. For robustness, we estimate the average change in a large number of pixels. It is important to investigate whether other effects can influence the effectiveness of our method.

Ambient light might be present, however it is an additive signal that is constant for all source positions. Thus, it can be measured

by capturing an image with all lights off and subtracted from the original images. This will lead to a reduced SNR which we already investigated in our work.

Forward-scattering can take place in small camera-scene distances, causing resolution loss [11,20]. In our work, we did not model forward-scattering. In [13] it was described that the contrast loss due to attenuation and backscatter in underwater images is dominant compared with the resolution loss caused by forward-scattering. This was demonstrated with objective criteria in [29,30], and a wide series of real experiments in a water tank in [19]. In [39] forward-scattering was included, however the additional model parameters were assigned arbitrary values, as the authors described that these are hard to estimate without tedious, high-precision calibration that is subject to errors. For this reason, in a later work a detailed sensitivity analysis was performed in order to determine the importance of the various model variables and it was concluded that the impact of the forward-scattering parameters is negligible compared with attenuation and backscatter [39]. In every case, our work provides the framework for evaluating such photometric models in murky water that can be investigated in future work.

Specularities are generally weak underwater and otherwise they can be mitigated using polarizers on the source and camera, respectively [30]. Similarly, inter-reflections can be compensated using a high-frequency spatial light modulator on the source [22]. We believe that modeling these effects mathematically adds unnecessary complexity when they can be mitigated optically, especially since we are designing an underwater robot. Self-shadowing may influence our approach for objects near the camera. Specifically, as a light source is moving it can create a self-shadow at some point that will change the pixel's intensity, causing a

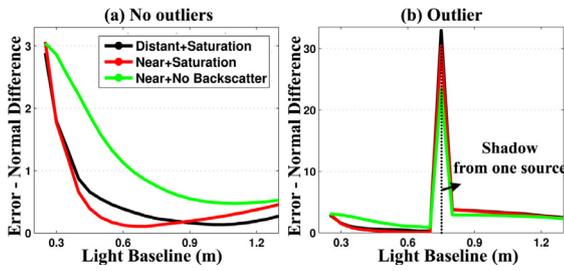


Fig. 11. As the sources move, the estimated normal changes smoothly (left). When an outlier is created from one of the sources, the respective change in the estimated normal is abrupt.

respective change to the estimated normal vector. However, contrary to the change in a normal vector due to erroneous model approximations or SNR error, the change due to such outliers is very abrupt. Fig. 11 demonstrates this effect. Left graph corresponds to a scene point that is not subject to outliers. The estimated normal vector in this case changes smoothly for all photometric models as we change the baseline. When one of the sources creates an outlier (shadow or specularity) in the right graph, the estimated normal will change abruptly at that point no matter what model is used. Thus, we can easily detect pixels that exhibit changes above some threshold, and omit them from our Photometric Stereo framework.

7.3. Discussion of lighting setup

We denoted the change in light source position by dr . Theoretically, if dr tends to zero the estimated normal maps will coincide regardless of the model validity. In reality dr corresponds to non-zero values that are feasible to achieve using a mechanical system or an array of light sources.

Selecting the value of dr depends on the sensing capabilities of the sensor that should be able to measure the respective change in image brightness. In our work, we tested different values of dr ($dr = 1$ cm in Section 6, $dr = 2$ cm is Section 7.1, and $dr = 5$ cm in the real experiments of Section 7.4), all of which in the order of cm, considering the physical limitations of the setup underwater where the light sources position varies usually within cm on the robotic platform and the scene distance varies from cm to 1–2 m in turbid water [11].

An interesting future direction would be to automatically adjust the displacement dr as well according to the accuracy needed and the sensing properties of the system. For example, starting from a large dr the system would predict an optimal light position, and then refine this using smaller displacements.

The number of sources is irrelevant with the effectiveness of our method, which requires at least 3 sources as any traditional PS system. We have tested our method using 3, 4 and 8 sources with no change in performance.

In the simulations of Section 7.1 we considered a symmetrical lighting system where all sources move in a square around the camera, while in the real experiments we used a system where the sources move in horizontal direction that is easier to implement. The displacement pattern does not affect the performance of our method, as soon as all of the sources move either further away or toward the camera. In our case, we considered only symmetrical setups that are coherent with the previous works of [21,38].

7.4. Real experiments

In order to evaluate our proposed system in real murky water, we performed experiments in a big water tank (approximately 5000l). We constructed a mechanical arm that allows the light fixtures to move along a horizontal line (Fig. 12). To the best of our

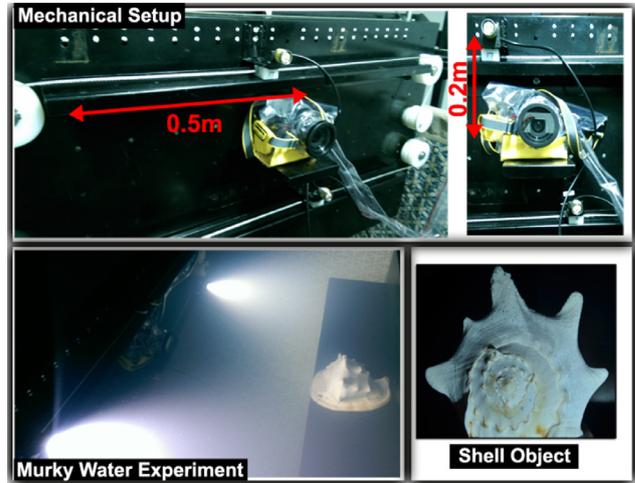


Fig. 12. Top: Our mechanical system for moving the light sources inside water. Two sources are placed above and two sources below the camera, that can move along a horizontal direction. Bottom: The imaging setup for one of the murky water experiments using a shell object.

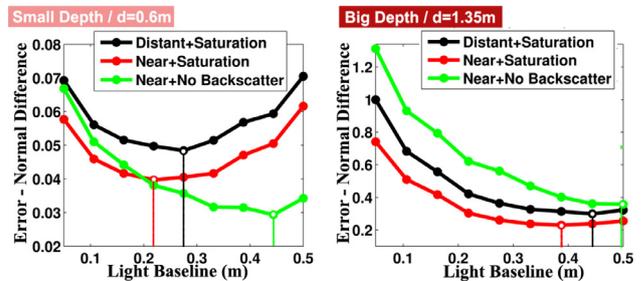


Fig. 13. At every scene depth we estimate the average change in normal map per model, and we select the baseline that exhibits the smallest change. In this way the system adapts the light baseline to the scenario, instead of using a fixed baseline.

knowledge this is the first dynamic Photometric Stereo system in murky water. Our platform was also able to move at different distances from the object. A NIKON D7000 camera with a NIKKOR 35 mm lens, and 4 sources were all immersed into the water. Two different levels of scattering were created by diluting milk into clean water as in [21,38], and a real shell object was imaged at four different scene depths at each scattering level. At every scene depth, the lighting setup around the camera was varied by moving 2 sources above and 2 below the camera along a horizontal line, creating 10 different baselines separated by a step of $dr = 5$ cm.

At every scene depth, we varied the light sources baseline and we estimated a different normal map for every photometric model. We estimated the change in the non-unit normal vectors since the reconstruction error can be reflected both to the normal direction and magnitude (which corresponds to the albedo). Fig. 13 shows our predicted error of the change in normals per model for two different scene depths in one of the scattering levels. This allowed us to select the optimal baseline per method, i.e. the one that minimizes the predicted error using Eq. (10). As in the ground-truth simulations of Fig. 8, for big depths the estimated optimal baseline was larger for all models compared with the respective ones at smaller depths, and also larger for the No Backscatter model compared with the other two models.

After selecting the optimal baseline for each photometric model at every scene depth, the final predicted error can be estimated (Fig. 14). This is very high for all models at large depths, where the SNR of the captured images was low due to the attenuated direct and the strong backscatter components. Then, the error was

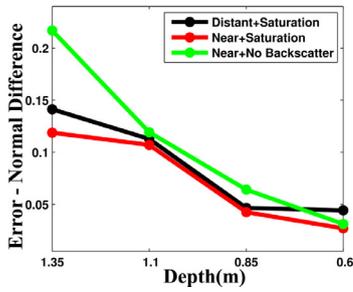


Fig. 14. The predicted error at a specific scattering level, per photometric model and scene depth (after selecting the optimal baseline in every case). All models exhibit high error at large scene depths, which is decreased rapidly as the platform moves closer to the object and SNR gets higher. For smaller scene depths the improvement in reconstruction is weaker, and in some cases (Distant+Backscatter Saturation model) the reconstruction error is increased, indicating the model invalidity.

decreased for smaller distances where the visibility was better. Fig. 15 shows the reconstructions for the Distant+Saturation model (for the optimal light positions at every depth) that correspond to the predicted error of Fig. 14. The reconstruction error was significantly decreasing initially as the scene depth was decreasing and the SNR was getting higher. Then at the closest depth where the near-light effect was strong, the predicted reconstruction error increased again indicating that the model becomes invalid. This is supported by the optical comparison between the reconstructed shapes and the reconstruction obtained using a depth sensor in pure air. For large scene depths the reconstruction is dominated by noise, for the smallest depth it has a strong bias in the middle (affected by the near-field illumination), and at the predicted optimal depth (marked by red) it is closest to the depth sensor reconstruction. For the smallest depth, our predicted error indicates that the two other models are more valid (Fig. 14), and this can be optically confirmed by the reconstructions in Fig. 15 (third row). Among the two, the error for the Near+Saturation method was predicted to be lower (marked by red).

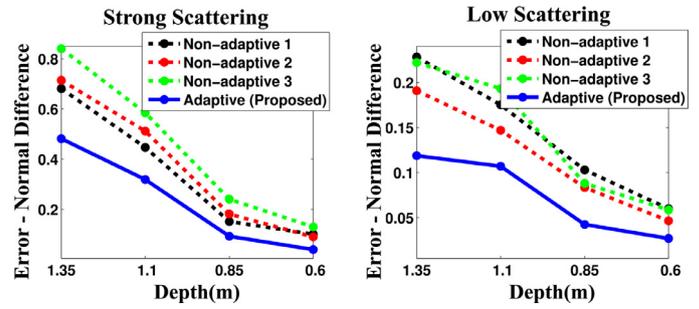


Fig. 16. The blue line corresponds to the predicted error of a system that adapts a different photometric model and light baseline at every scene depth (minimum change in surface normals). The dotted lines show the predicted error of different systems that use a fixed model and light baseline in all cases (black line: Distant+Saturation with a baseline of 40 cm, red line: Near+Saturation with a baseline of 30 cm, green line: Near+No Backscatter with a baseline of 20 cm). The respective reconstructions for every system can be seen in Figs. 17 and 18. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

Our method yields an adaptive Photometric Stereo system which can adjust the light position and the photometric model according to the scenario. The graphs in Fig. 16 show the predicted error of using such an adaptive system that selects a different model and light position at every scene depth (the ones that correspond to the minimum change in the surface normals) and different traditional, non-adaptive systems that use a constant light baseline and model regardless of the scenario. Figs. 17 and 18 show the respective optical results for each system at two scattering levels. First, it can be noticed how significantly different the reconstruction quality can be as the depth, scattering level, photometric model and light baseline vary. Using such factors arbitrarily can yield large reconstruction errors. Our adaptive system yields significantly better shape and albedo estimations for all distances in both scattering levels, without any prior knowledge about the true object shape. The reconstruction marked by red corresponds to our predicted optimal solution among all potential reconstructions (for every scattering level there were 10 light positions \times 3

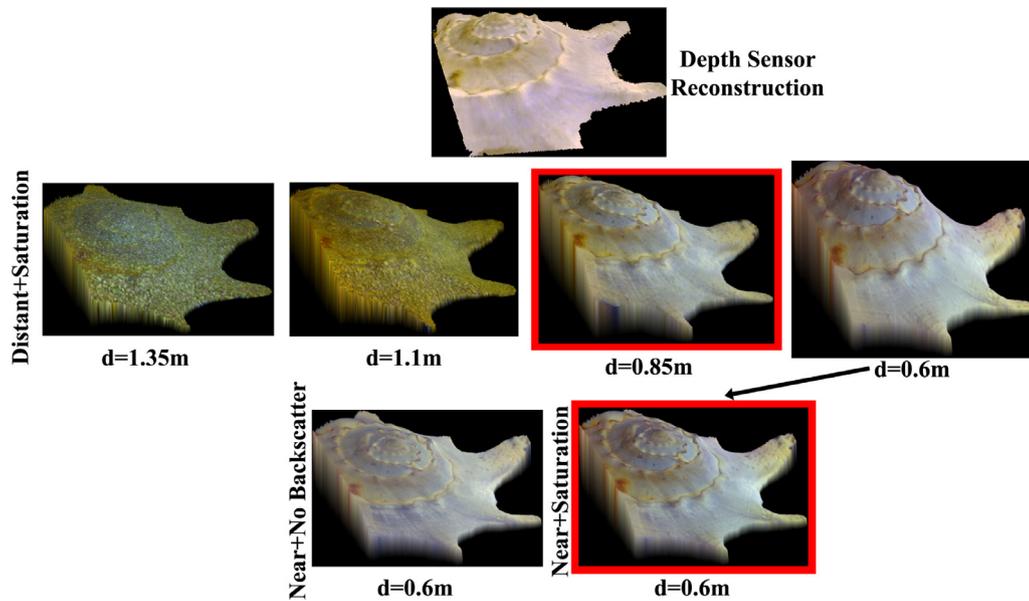


Fig. 15. The reconstructed object shape when the Distant-Lighting+Backscatter Saturation photometric model is used, for decreasing camera-scene distance (second row). Our method can automatically predict that the reconstruction quality is initially increasing as the distance is getting smaller. Then, for the smallest distance the reconstruction quality is getting worse due to the model invalidity. Thus our method predicts the optimal camera-scene distance (denoted by red) for the specific model. Then, for the smallest scene depth it is predicted that the other two models are more valid, and specifically that the reconstruction for the Near+Saturation model is optimal (marked by red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

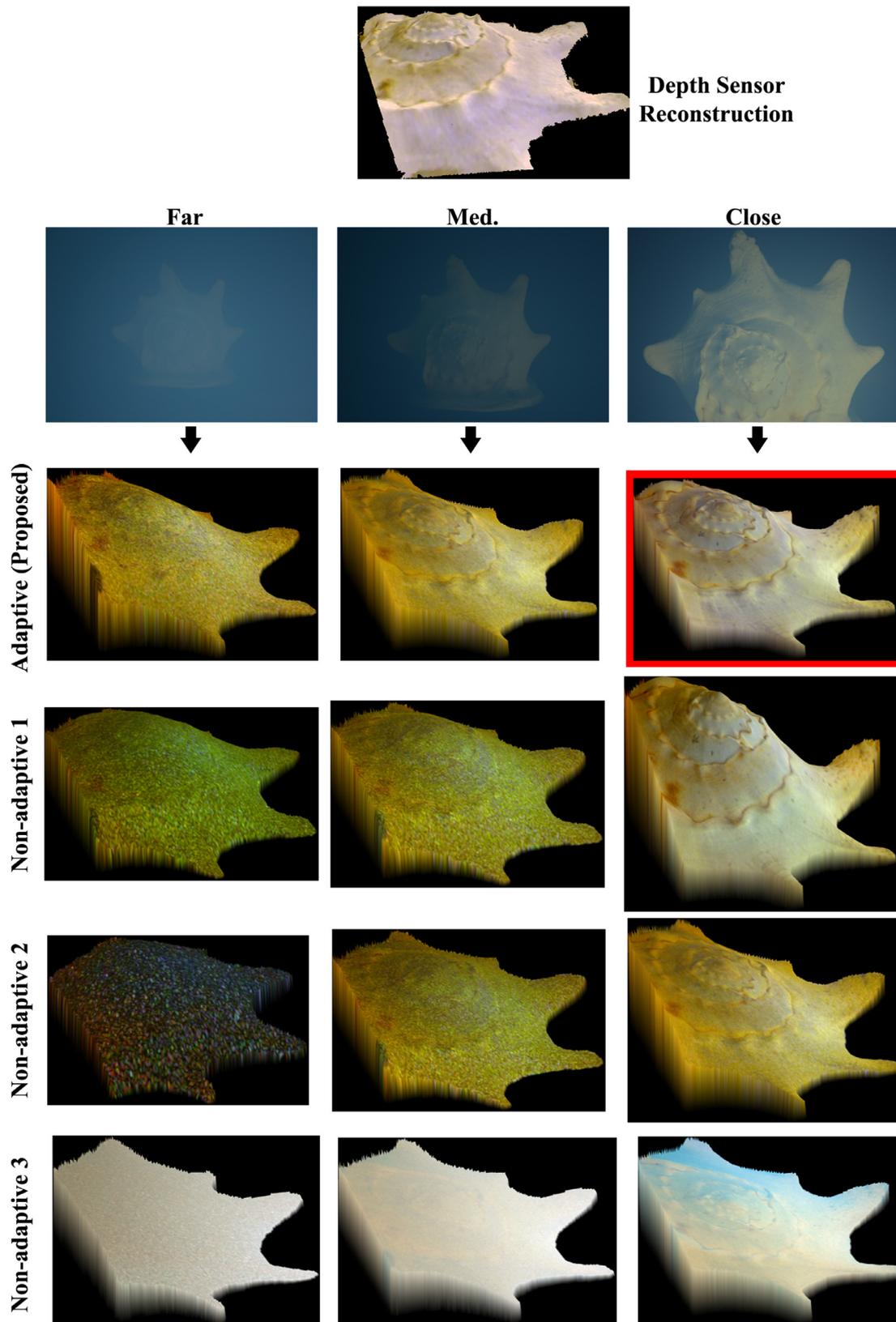


Fig. 17. The reconstruction results of an adaptive and different non-adaptive systems at different scene depths. The predicted optimal reconstruction (among all scene depths & models) is marked by red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

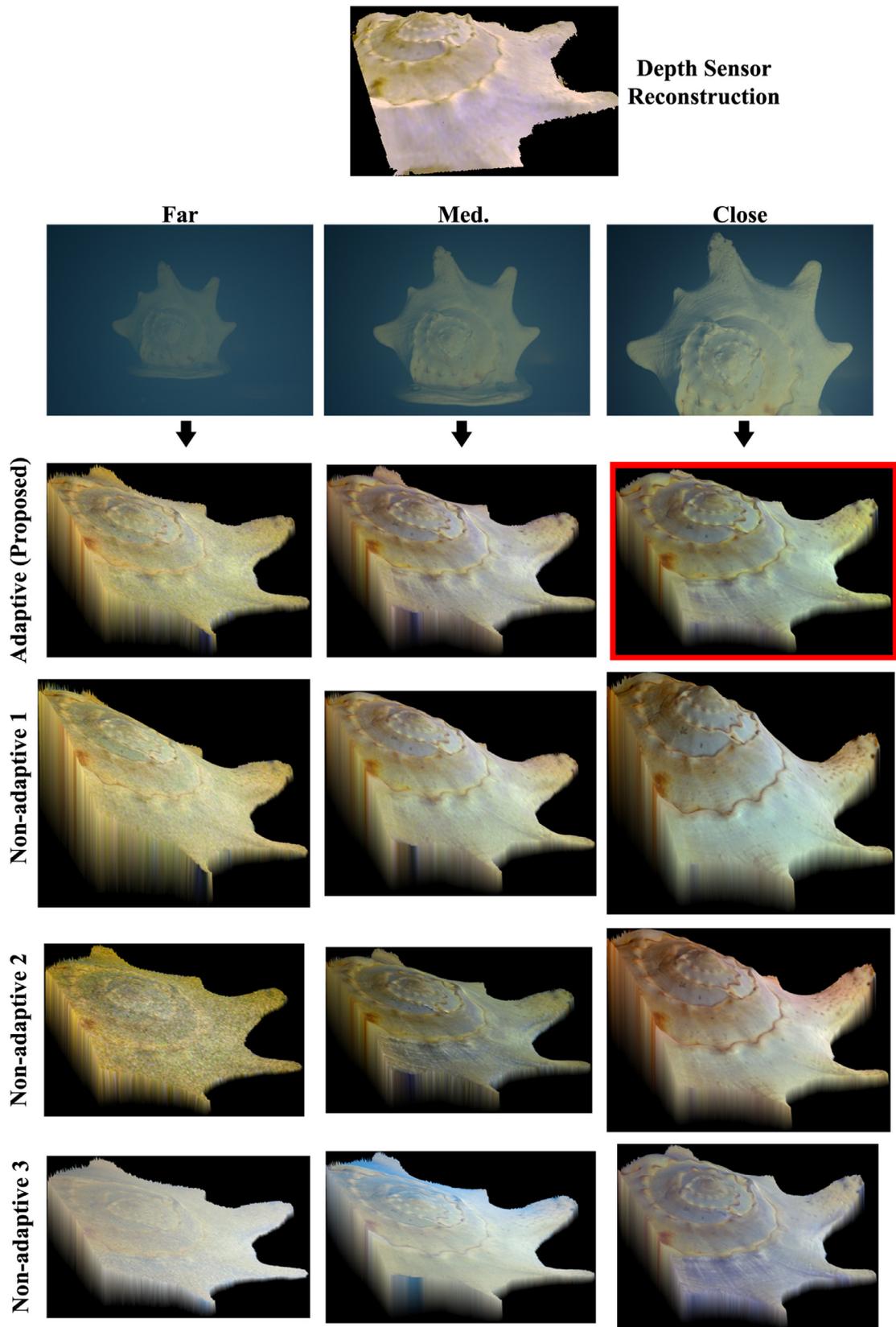


Fig. 18. The reconstruction results for a different scattering level.

models \times 4 scene depths = 120 potential reconstructions that can be seen in our supplementary material).

8. Discussion and conclusions

In this work, we evaluated the effectiveness of different photometric models in murky water considering several factors that are critical in realistic imaging conditions (model validity, camera SNR, light source baseline, distance, scattering etc.). We showed that the reconstruction error depends strongly on the imaging conditions and for this reason we presented a simple way for predicting the effectiveness of a photometric model when the scene and the environment are unknown. The effectiveness of our method lies on the observation that the change in estimated surface normals under different light source positions reflects the true error due to an invalid photometric model or noise due to low SNR.

Our methodology still assumed that some information is a-priori known or calibrated, such as the incident illumination vector on the object centroid. This can be easily estimated using a sonar or laser beam sensor [23], or using uncalibrated methods that have been suggested for pure air [2,9,34]. The goal of our work was to focus on the photometric model validity in murky water, rather than on its final optimization. In future work, our framework can be used to evaluate totally uncalibrated methods.

Approximating the model's effectiveness using our dynamic lighting system offers significant potentials to Photometric Stereo in murky water. Specifically, some critical parameters that were chosen arbitrarily in previous works, such as the scene depth, the light baseline and the photometric model, can be adapted automatically to an unknown imaging scenario. Additional photometric models could be evaluated through our framework, such as forward or multiple scattering [10,11,20,23], or models which assume different illumination profiles for the light sources such as [21] which neglects shadow volumes in the medium or [26] which models non-isotropic sources.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at [10.1016/j.cviu.2016.03.002](https://doi.org/10.1016/j.cviu.2016.03.002).

References

- [1] A. Abrams, C. Hawley, R. Pless, Heliometric stereo: shape from sun position, in: *Computer Vision-ECCV 2012*, Springer, 2012, pp. 357–370.
- [2] N.G. Aldrin, S.P. Mallick, D. Kriegman, Resolving the generalized bas-relief ambiguity by entropy minimization, in: *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on, IEEE, 2007*, pp. 1–7.
- [3] V. Argyriou, S. Zafeiriou, B. Villarini, M. Petrou, A sparse representation method for determining the optimal illumination directions in photometric stereo, *Signal Process.* 93 (11) (2013) 3027–3038.
- [4] A.M. Bradley, M.D. Feezor, H. Singh, F.Y. Sorrell, Power systems for autonomous underwater vehicles, *Ocean. Eng. IEEE J.* 26 (4) (2001) 526–538.
- [5] M. Chandraker, J. Bai, R. Ramamoorthi, On differential photometric reconstruction for unknown, isotropic brdfs, *Pattern Anal. Mach. Intell. IEEE Trans.* 35 (12) (2013) 2941–2955.
- [6] S. Chen, Y. Li, N.M. Kwok, Active vision in robotic systems: a survey of recent developments, *Int. J. Robot. Res.* 30 (11) (2011) 1343–1377.
- [7] C. Collewet, E. Marchand, Photometric visual servoing, *Robot. IEEE Trans.* 27 (4) (2011) 828–834.
- [8] O. Drbohlav, M. Chantler, On optimal light configurations in photometric stereo, in: *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on, vol. 2, IEEE, 2005*, pp. 1707–1712.
- [9] P. Favaro, T. Papadimitri, A closed-form solution to uncalibrated photometric stereo via diffuse maxima, in: *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, IEEE, 2012*, pp. 821–828.
- [10] M. Gupta, S. Narasimhan, Y. Schechner, On controlling light transport in poor visibility environments, in: *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on, IEEE, 2008*, pp. 1–8.
- [11] J.S. Jaffe, Computer modeling and the design of optimal underwater imaging systems, *Ocean. Eng. IEEE J.* 15 (2) (1990) 101–111.
- [12] J.S. Jaffe, Multi autonomous underwater vehicle optical imaging for extended performance, in: *OCEANS 2007-Europe, IEEE, 2007*, pp. 1–4.
- [13] W. Jagger, W. Muntz, Aquatic vision and the modulation transfer properties of unlighted and diffusely lighted natural waters, *Vis. Res.* 33 (13) (1993) 1755–1763.
- [14] M. Johnson-Roberson, O. Pizarro, S.B. Williams, I. Mahon, Generation and visualization of large-scale three-dimensional reconstructions from underwater robotic surveys, *J. Field Robot.* 27 (1) (2010) 21–51.
- [15] N. Kolagani, J.S. Fox, D.R. Blidberg, Photometric stereo using point light sources, in: *Robotics and Automation, 1992. Proceedings., 1992 IEEE International Conference on, IEEE, 1992*, pp. 1759–1764.
- [16] T. Kuparinen, V. Kyrki, Optimal reconstruction of approximate planar surfaces using photometric stereo, *Pattern Anal. Mach. Intell. IEEE Trans.* 31 (12) (2009) 2282–2289.
- [17] R. Mecca, A. Tankus, A. Wetzler, A.M. Bruckstein, A direct differential approach to photometric stereo with perspective viewing, *SIAM J. Imaging Sci.* 7 (2) (2014a) 579–612.
- [18] R. Mecca, A. Wetzler, A.M. Bruckstein, R. Kimmel, Near field photometric stereo with point light sources, *SIAM J. Imaging Sci.* 7 (4) (2014b) 2732–2770.
- [19] H. Mortazavi, J. Oakley, U. of Manchester. School of Electrical, E. Engineering, Mitigation of Contrast Loss in Underwater Images, University of Manchester, 2010.
- [20] Z. Murez, T. Treibitz, R. Ramamoorthi, D. Kriegman, Photometric stereo in a scattering medium, *Computer Vision, 2015. ICCV 2015. IEEE International Conference on (2015)*.
- [21] S. Narasimhan, S. Nayar, B. Sun, S. Koppal, Structured light in scattering media, in: *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on, vol. 1, 2005*, pp. 420–427. Vol. 1
- [22] S.K. Nayar, G. Krishnan, M.D. Grossberg, R. Raskar, Fast separation of direct and global components of a scene using high frequency illumination, *ACM Trans Graph. (TOG)* 25 (3) (2006) 935–944.
- [23] S. Negahdaripour, H. Zhang, X. Han, Investigation of photometric stereo method for 3-d shape recovery from underwater imagery, in: *OCEANS'02 MTS/IEEE, vol. 2, 2002*, pp. 1010–1017.
- [24] R.A. Newcombe, A.J. Davison, S. Izadi, P. Kohli, O. Hilliges, J. Shotton, D. Molyneux, S. Hodges, D. Kim, A. Fitzgibbon, Kinectfusion: Real-time dense surface mapping and tracking, in: *Mixed and augmented reality (ISMAR), 2011 10th IEEE international symposium on, IEEE, 2011*, pp. 127–136.
- [25] T. Papadimitri, P. Favaro, Uncalibrated near-light photometric stereo, in: *In Proceedings of British Machine Vision Conference, 2014*.
- [26] J. Park, S.N. Sinha, Y. Matsushita, Y.-W. Tai, I.S. Kweon, Calibrating a non-isotropic near point light source using a plane, in: *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on, IEEE, 2014*, pp. 2267–2274.
- [27] R. Patterson, Backscatter reduction for artificially illuminated in-water camera systems, *Opt. Eng.* 14 (4) (1975) 144357–144357
- [28] A. Sarafraz, S. Negahdaripour, Y. Schechner, Performance assessment in solving the correspondence problem in underwater stereo imagery, in: *OCEANS 2010, 2010*, pp. 1–7.
- [29] Y.Y. Schechner, N. Karpel, Clear underwater vision, in: *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on, vol. 1, IEEE, 2004*, pp. 1–536
- [30] Y.Y. Schechner, N. Karpel, Recovery of underwater visibility and structure by polarization analysis, *Ocean. Eng. IEEE J.* 30 (3) (2005) 570–587.
- [31] L. Shen, P. Tan, Photometric stereo and weather estimation using internet images, in: *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, IEEE, 2009*, pp. 1850–1857.
- [32] A. Spence, M. Chantler, Optimal illumination for three-image photometric stereo using sensitivity analysis, *IEE Proc. Vis. Image Signal Process.* 153 (2) (2006) 149–159.
- [33] J. Sun, M. Smith, L. Smith, A. Farooq, Examining the uncertainty of the recovered surface normal in three light photometric stereo, *Image Vis. Comput.* 25 (7) (2007) 1073–1079.
- [34] P. Tan, S.P. Mallick, L. Quan, D. Kriegman, T. Zickler, Isotropy, reciprocity and the generalized bas-relief ambiguity, in: *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on, IEEE, 2007*, pp. 1–8.
- [35] T. Treibitz, Y. Schechner, Instant 3descatter, in: *Computer Vision and Pattern Recognition, 2006 IEEE Conference on, vol. 2, 2006*, pp. 1861–1868.
- [36] T. Treibitz, Y. Schechner, Active polarization descattering, *Pattern Anal. Mach. Intell. IEEE Trans.* 31 (3) (2009) 385–399.
- [37] T. Treibitz, Y. Schechner, Turbid scene enhancement using multi-directional illumination fusion, *Image Process. IEEE Trans.* 21 (11) (2012) 4662–4667.
- [38] C. Tsotsios, M.E. Angelopoulou, T.-K. Kim, A.J. Davison, Backscatter compensated photometric stereo with 3 sources, in: *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on, IEEE, 2014*, pp. 2259–2266.
- [39] S. Zhang, S. Negahdaripour, 3-d shape recovery of planar and curved surfaces from shading cues in underwater images, *Ocean. Eng. IEEE J.* 27 (1) (2002) 100–116.