Depth and Image Restoration from Light Field in a Scattering Medium

Jiandong Tian[1] Zak Murez[2] Tong Cui[1,3] Zhen Zhang[1,3] David Kriegman[2] Ravi Ramamoorthi[2]
[1] State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences;
[2] Department of Computer Science and Engineering, University of California, San Diego;
[3] University of Chinese Academy of Sciences
{tianjd,cuitong,zhangzhen}@sia.cn, {zmurez,kriegman,ravir}@cs.ucsd.edu

Abstract

Traditional imaging methods and computer vision algorithms are often ineffective when images are acquired in scattering media, such as underwater, fog, and biological tissue. Here, we explore the use of light field imaging and algorithms for image restoration and depth estimation that address the image degradation from the medium. Towards this end, we make the following three contributions. First, we present a new single image restoration algorithm which removes backscatter and attenuation from images better than existing methods do, and apply it to each view in the light field. Second, we combine a novel transmission based depth cue with existing correspondence and defocus cues to improve light field depth estimation. In densely scattering media, our transmission depth cue is critical for depth estimation since the images have low signal to noise ratios which significantly degrades the performance of the correspondence and defocus cues. Finally, we propose shearing and refocusing multiple views of the light field to recover a single image of higher quality than what is possible from a single view. We demonstrate the benefits of our method through extensive experimental results in a water tank.

1. Introduction

Images captured in scattering media such as underwater or in fog are degraded by light absorption and scattering. This seriously affects the performance of standard computer vision algorithms which were developed to work well in clear air. As such, there has been a lot of effort to adapt these algorithms to handle scattering [18, 23, 1, 31, 15]. In this work, we introduce the use of light field cameras in scattering media.

Recently, light field cameras have become commercially available. These cameras use an array of micro lenses to capture enough spatial and angular information to allow post capture change of view point and refocusing. This enables the extraction of correspondence and defocus cues, for depth estimation, from a single shot [28, 36, 32, 14].

There has been a lot of previous work on restoring images affected by scattering. Most work on dehazing

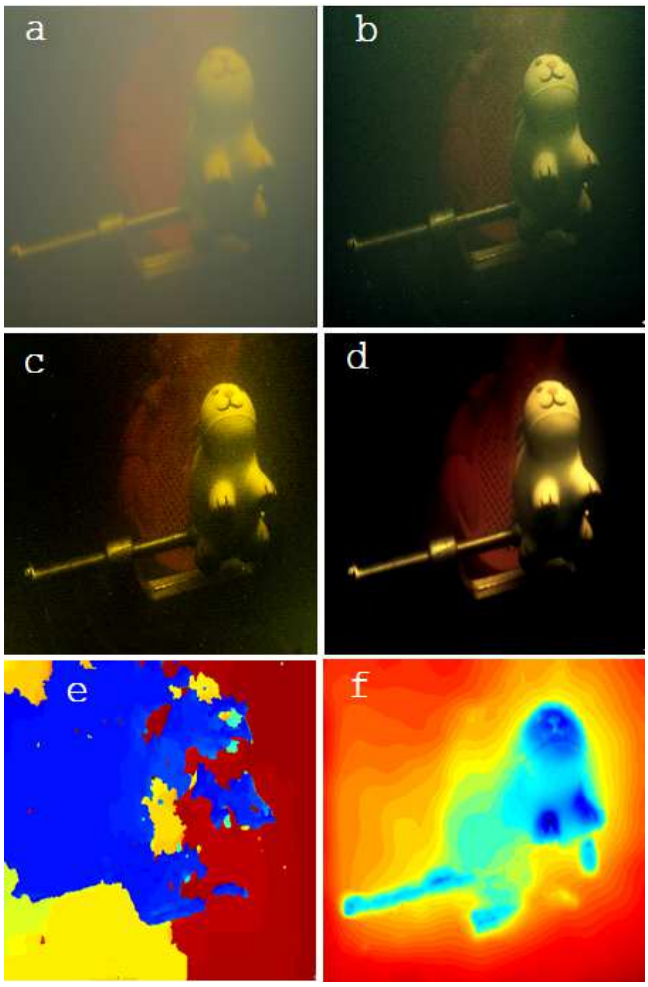

Figure 1. *Depth estimation and image restoration comparison. (a) The center view of an image captured by a light field camera in a densely scattering medium. (b) Restoration of the center view by the recent method of Drews et al. [6]. (c) Restoration of the center view by the recent method of Tsiotsios et al. [31]. Notice the noise on the orange lobster and the greenish tint of the entire image, especially on the white bunny. (d) Our restoration. (e) Depth estimation without accounting for scattering by the recent method of Wang et al. [32]. (f) Our depth.*

[35, 8, 7, 37, 12] focuses on removing the backscatter, also known as airlight. These algorithms are based on the scattering model proposed by Narasimhan and Nayar [17], which assumes distant illumination in which the light rays are all parallel and travel equal distances to scattering points. This model works well for sunlight, but fails on near-field illumination. More recently, Tsiotsios et al. [31] proposed an empirical algorithm which fits a quadratic function to remove backscatter from near-field illumination.

In this paper, we propose a novel physically based algorithm to remove backscatter as well as attenuation (Sec. 4). Our restoration requires knowing the scene depths, which are a priori unknown. As such, we first perform a rough restoration assuming a constant initialized depth. We apply this restoration to each view in the light field.

After our initial restoration, existing shape from light field algorithms can be used to estimate depth. However, in densely scattering media, the restorations are noisy and have low signal to noise ratio because the additive backscatter dominates the camera's dynamic range [30, 15]. As such the common defocus and correspondence cues do not provide reliable depth. To address this we introduce a new, non-uniform corrected, transmission based depth cue, which complements the light field cues (Sec. 5).

Once the depth is estimated, we can use it to perform a final image restoration. However as mentioned, the restored images are noisy. As such, we also propose shearing and averaging multiple views of the light field to produce one high quality image (Sec. 6). This produces much higher quality images than simply extracting the central view or simply averaging multiple views (Fig. 7 and Fig. 12). Furthermore this circumvents the need to properly focus the camera at capture time, which can be difficult in dense media (auto focus often fails).

We demonstrate our method through extensive experiments, in a water tank, with varying concentrations of scattering media added (Sec. 7). As shown in Fig. 1, our method significantly outperforms recent image restoration and depth estimation methods.

2. Related work

There are many existing techniques that restore the visibility of degraded underwater images. Schechner et al. [24] use two images with different degrees of polarization for underwater image restoration. Roser et al. [22], Swirski et al. [27], and Negahdaripour et al. [19] use a stereo pair of images to simultaneously solve for depth estimation and visibility enhancement. Murez et al. [15] and Tsiotsios et al. [31] use three or more images under varying illumination and solve for photometric stereo problem in underwater scattering condition. These methods all require multiple images of the same scene.

Recently, many single image underwater restoration methods have been proposed based on the Narasimhan-Nayar imaging model [17]. Most previous work has focused on improving the estimation of the transmission map. He et al. [8] propose the well-known dark channel prior (DCP) to estimate scene depths in hazy images. This prior assumes that most non-sky patches of haze-free outdoor images have low pixel intensities. In [11, 35, 3], the DCP methodology was applied to underwater image restoration. Serikawaa et al. [25] propose a variation of DCP to refine the medium transmission map by using a guided joint trilateral filter. Bianco et al. [2] proposed an improved underwater DCP that exploits the difference in attenuation between the three image color channels, i.e., water attenuates red light more than green and blue. Drews-Jr et al. [6] proposes the underwater dark channel prior (UDCP) and similarly assume that the blue and green channels contain most of the visual information.

These works all assume mild haze and tend to fail in the more challenging case of dense scattering. In this paper, we combine light field (LF) imaging and an improved near-field scattering model to tackle this problem.

Recently, LF cameras have become readily available for consumers. Because of their ability to capture multiple viewpoints in a single image, LF cameras provide both monocular and stereo depth cues in an easy-to-capture package. Tao et al. [28] propose a depth estimation method that combines correspondence and defocus cues in the 4D Epipolar Image (EPI). Wanner et al. [34] propose a globally consistent depth estimation framework by applying structure tensors to estimate the directions of feature pixels in the 2D EPI. Wang et al. [32] develop an occlusion-aware depth estimation algorithm from a LF camera, which can obtain more accurate depth even in the presence of occlusions. Mousnier et al. [14] describe a novel depth estimation approach to partially reconstruct high-resolution 4D light fields from a stack of differently focused photographs taken with a fixed camera. Dansereau et al. [5] propose the volumetric focus method to improve signal quality that maintains focus over a controllable range of depths.

Although light field cameras have proven advantageous over traditional cameras, their application to underwater imaging is still limited. In this paper we propose a novel application of LF cameras and extend LF based depth estimation to work in scattering media.

3. Underwater Image Formation Model

As shown in Fig. 2, 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 $\mathbf{x} = (x, y)$. We have the following relations, $\mathbf{x} = \left(f\frac{X}{Z}, f\frac{Y}{Z}\right)^t$ and $\mathbf{X} = \left(\frac{Z}{f}x, \frac{Z}{f}y, Z\right)^t$, where f is the focal length. We ignore the (u, v) angular coordinates of the light field in this section.

Let $S = (X_s, Y_s, Z_s)$ be the world coordinates of a near-field point light source, and define $\mathbf{D}(\mathbf{X}) = \mathbf{S} - \mathbf{X}$ as the vector from the object to the source. As in related

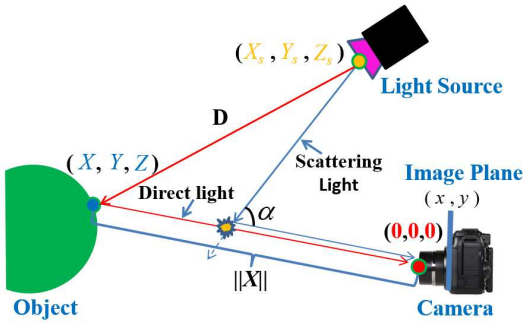

Figure 2. *Image formation and light propagation in a scattering medium. The radiance arriving at the camera is the sum of two components: the direct reflected light and the backscatter. The direct light travels distance D from the light to the object, and then distance $\|X\|$ to the camera. The backscatter is scattered through angle α directly into the camera.*

works [31, 18], we adopt the single scattering model, and ignore small angle forward scattering, and thus only consider backscatter from the source. Thus the radiance arriving at the camera can be expressed as the sum of two terms:

$$I(\mathbf{x}) = I_d(\mathbf{x}) + I_b(\mathbf{x}) \tag{1}$$

where I_d is the direct light reflected from the object and I_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. This term is known as backscatter.

3.1. Direct Radiance Term

As seen in Fig. 2, consider an isotropic point source with radiant intensity E. Light from it travels a distance $||\mathbf{D}(\mathbf{X})||$ to the object. Thereafter, the light is reflected by the surface with albedo $\rho(\mathbf{X})$ and travels a further distance $||\mathbf{X}||$ ($||\mathbf{X}|| = \sqrt{X^2 + Y^2 + Z^2}$) to the camera. The direct radiance can be written as,

$$I_d(\mathbf{x}) = \frac{E}{||\mathbf{D}(\mathbf{X})||^2} e^{-\sigma||\mathbf{D}(\mathbf{X})||} \rho(\mathbf{X}) e^{-\sigma||\mathbf{X}||} \tag{2}$$

where σ denotes extinction coefficient of the medium. Let $J(\mathbf{x}) = \frac{E}{||\mathbf{D}(\mathbf{X})||^2}\rho(\mathbf{X})$ be the clear image, and we have,

$$I_d(\mathbf{x}) = J(\mathbf{x}) e^{-\sigma(||\mathbf{D}(\mathbf{X})|| + ||\mathbf{X}||)} \tag{3}$$

3.2. Backscatter Term

Light which is scattered directly into the camera by the medium without reaching the object is termed backscatter. The fraction of light scattered to each direction is determined by scattering coefficient β and phase function $P(g, \alpha)$. We adopt the common Henyey-Greenstein phase function [9] in this paper.

$$P(g, \alpha) = \frac{1}{4\pi} \cdot \frac{1 - g^2}{[1 + g^2 - 2g\cos\alpha]^{3/2}} \tag{4}$$

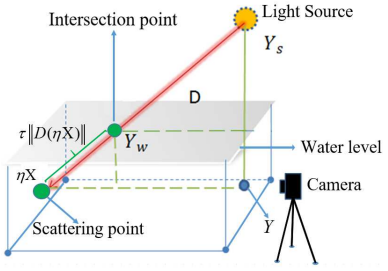

Figure 3. *Illustration of the geometry of a near field source outside the medium.*

where scattering angle α is given by $\cos(\alpha) = \hat{\mathbf{D}} \cdot \hat{\mathbf{X}}$ ($\hat{\mathbf{D}}$ and $\hat{\mathbf{X}}$ denote normalized vectors) and satisfy $\alpha \in [0, \pi]$. The parameter $g \in (-1, 1)$ controls the relative amounts of forward and backward scattering.

Based on previous work [26, 16], the backscatter is given by,

$$I_b(\mathbf{x}) = F L_b \tag{5}$$

where

$$L_b(\mathbf{x}) = \int_0^\infty \frac{e^{-\sigma||D(r\hat{\mathbf{X}})||}}{||D(r\hat{\mathbf{X}})||^2} P(\alpha) e^{-\sigma r} dr \tag{6}$$

and $F = E\beta$. Note that β is the scattering coefficient and is related to extinction coefficient σ by $\sigma = \beta + \epsilon$, where ϵ is the absorption coefficient. We have absorbed the radiant intensity of the source E and the scattering coefficient β into one effective constant F. Note that since backscatter from a near-field source saturates close to the camera [31], it is not a problem to integrate to infinity even if the ray hits an object at a finite distance.

3.3. Extension to Light Sources Outside the Medium

So far, we have assumed the light source is in the medium. Here we extend the scattering model presented in the previous sections to the case where the light is near-field but outside the medium. Consider the geometry in Fig. 3. Light travels from the source and enters the medium at point $\mathbf{X}_w$ without undergoing any scattering. Then the light continues into the medium where it undergoes scattering as before. The direct radiance becomes

$$I_d(\mathbf{x}) = J(\mathbf{x}) e^{-\sigma(||\tau\mathbf{D}(\mathbf{X})|| + ||\mathbf{X}||)} \tag{7}$$

and the backscatter becomes

$$L_b(\mathbf{x}) = \int_0^\infty \frac{e^{-\sigma||\tau D(r\hat{\mathbf{X}})||}}{||D(r\hat{\mathbf{X}})||^2} P(\alpha) e^{-\sigma r} dr \tag{8}$$

where the only difference from Eq. 6 is that the attenuation of the source is scaled by $\tau \in (0, 1]$. Note that when $\tau = 1$ the source is in the water.

τ is given by

$$\tau(\mathbf{X}) = \frac{Y_w - Y}{Y_s - Y} \tag{9}$$

where Y_w, Y_s, and Y denote water surface, source location, and scattering point location respectively in the vertical direction. It is worth noting that τ is a function of $\mathbf{X}$, which means we do not assume all points are equidistant from the source, as in [18].

Note that we have not included terms here to account for refraction nor the Fresnel effect, as they could be safely neglected due to the geometry of our experimental setup, although they could easily be added and do not affect the derivations in the rest of the paper. In fact, for the rest of our analysis we just assume the light is in the medium for notational clarity.

4. Single Image Restoration

Substituting Eq. 3 into Eq. 1 we get

$$I(\mathbf{x}) = J(\mathbf{x})e^{-\sigma(||\mathbf{D}(\mathbf{X})||+||\mathbf{X}||)} + I_b(\mathbf{x}) \tag{10}$$

To recover the restored image we need to solve Eq.10 for J which requires estimating the unknown medium parameters F, σ, and g, as well as the depth for each pixel. To make the optimization simpler, we first estimate the medium parameters by examining pixels which only contain backscatter. Once the medium parameters are known, an initial restoration is computed. This is done using Eqs (2), (3) and (7), assuming a constant known depth Z_{ref}. These restored images are used to estimate depth as described in the following section, and then the estimated depth is used to compute a final restoration. Our experiments show that further iteration does not improve the results, and are robust to the initialization Z_{ref}.

4.1. Estimating Medium Parameters

For pixels that only contain backscatter, $J(\mathbf{x}) = 0$ and Eq. 10 reduces to

$$I(\mathbf{x}) = I_b(\mathbf{x}) \tag{11}$$

Let V be a set of points that only contain backscatter (the assumption of the existence of backscatter only pixels is common in the literature [31] and often satisfied in underwater imaging conditions, we will describe how to find such a set of points in the next paragraphs). We estimate F, σ, and g by minimizing

$$\min_{g} \min_{\sigma,F} \sum_{\mathbf{x}\in V} \|I(\mathbf{x}) - I_b(\mathbf{x})\| \tag{12}$$

while assuming $Z = Z_{\text{ref}}$. The inner optimization is solved using the simplex method [10], while the outer optimization is solved by brute force search with a step size of 0.01 over the limited range $g \in [0.7 - 1.0]$ valid for water [16].

In general, the medium parameters, as well as the light source intensity, depend on wavelength. Although we can solve Eq. 12 for each color channel independently, we found it more robust to assume σ and g are wavelength independent, while allowing for wavelength dependent sources E^c and scattering coefficients β^c, where $c \in \{R, G, B\}$ color channels. Note that $F^c = E^c\beta^c$ absorbs both wavelength dependent parameters into a single one per color channel. We solve the optimization in Eq. 12 for the blue channel, and then compute F^R and F^G by

$$F^c = \frac{1}{|V|}\sum_{\mathbf{x}\in V}\left(\frac{I^c(\mathbf{x})}{I^B(\mathbf{x})}\right)F^B, F^c = E^c\beta, c \in \{R, G\} \tag{13}$$

where $|V|$ is the number of pixels in set V.

To find a good set V of pixels that only contain backscatter, we modify the DCP method proposed by He et al. [8]. The dark channel prior states that in most natural image patches, at least once color channel has some very low intensity pixels. To make our method more robust, we first convert the input image into HSV color space and extract its saturation and intensity channels. We then extract regions, whose intensity and saturation are both low, using Otsu's method [21]. In these regions, we further sample points at twenty pixel intervals in both the x and y directions. Finally we take the minimum over five pixel neighborhoods for these sample points to generate set V. Note that we do not extract the minimum over color channels, as in the original DCP, because we allow wavelength dependent light sources.

5. Depth Estimation

After the initial image restoration, assuming a constant depth Z_{ref}, we estimate the true depth using shape from light field [28, 29]. However, for densely scattering media, where the restored images have a low signal to noise ratio, the defocus and correspondence cues are not enough to recover accurate depth. As such we introduce our new transmission based depth cue which can be combined with the defocus and correspondence cues to recover better depth.

5.1. Defocus and Correspondence Cues

We first use the following equation from Ng et al. [20] to shear the LF data to various depths.

$$J_\kappa(\mathbf{x}, \mathbf{u}) = J(\mathbf{x} + \mathbf{u}(1 - \frac{1}{\kappa}), \mathbf{u}) \tag{14}$$

where J is the initially haze removed 2D LF input image, $J_\kappa(\mathbf{x}, \mathbf{u})$ is the 4D sheared LF images at relative depth κ, $\mathbf{x} = (x, y)$ denotes the spatial coordinates and $\mathbf{u} = (u, v)$ denotes the angular coordinates. For each pixel, the refocused image $\overline{J_\kappa}$ for the shear value κ is calculated by,

$$\overline{J_\kappa}(\mathbf{x}) = \frac{1}{N}\sum_{\mathbf{u}} J_\kappa(\mathbf{x}, \mathbf{u}) \tag{15}$$

where N is the number of angular pixels. In our implementation, we shear κ from 0.2 to 2 with 256 steps. Therefore, we have 256 refocused $\overline{J_\kappa}(\mathbf{x})$.

The defocus DE and correspondence CO cues [29] are given by

$$DE_\kappa(\mathbf{x}) = \frac{1}{|W|}\sum_{\mathbf{x'}\in W} |\overline{J_\kappa}(\mathbf{x'}) - J(\mathbf{x'}, \mathbf{0})| \tag{16}$$

$$CO_\kappa(\mathbf{x}) = \frac{1}{N}\sum_{\mathbf{u}} |J_\kappa(\mathbf{x},\mathbf{u}) - J(\mathbf{x},\mathbf{0})| \tag{17}$$

When a patch is sheared to its correct depth, it will exhibit small variance and defocus. Therefore, we choose the corresponding depth responses by,

$$\begin{aligned} Z_D(\mathbf{x}) &= \underset{\kappa}{\operatorname{argmin}}\, DE_\kappa(\mathbf{x}) \\ Z_C(\mathbf{x}) &= \underset{\kappa}{\operatorname{argmin}}\, CO_\kappa(\mathbf{x}) \end{aligned} \tag{18}$$

5.2. Transmission Depth Cue

Our transmission depth cue is derived from the depth dependent backscatter intensity after the object's reflected intensity has been removed using the DCP prior.

Taking the minimum over color channels and spatial neighborhoods $\Omega(\mathbf{x})$ of Eq. 10 gives

$$\begin{aligned} &\min_c \min_{\mathbf{y}\in\Omega(\mathbf{x})} \left(I^c(\mathbf{y})\right) = \\ &\min_c \min_{\mathbf{y}\in\Omega(\mathbf{x})} \left(J^c(\mathbf{y}) \cdot e^{-\sigma(||\mathbf{D}(\mathbf{X})||+||\mathbf{X}||)} + I_b^c(\mathbf{y})\right) \end{aligned} \tag{19}$$

According to the DCP prior [8], $\min_c \min_{\mathbf{y}\in\Omega(\mathbf{x})} \left(J^c(\mathbf{y})\right) = 0$. Let $I^\dagger(\mathbf{x}) = \min_c \min_{\mathbf{y}\in\Omega(\mathbf{x})} \left(I^c(\mathbf{y})\right)$. Then Eq. 19 reduces to (after substituting Eqs. 5,6)

$$I^\dagger(\mathbf{x}) = F^c \int_0^{\|\mathbf{X}\|} \frac{e^{-\sigma||D(r\hat{\mathbf{X}})||}}{||D(r\hat{\mathbf{X}})||^2} P(\alpha) e^{-\sigma r} dr \tag{20}$$

Eq. 20 is a nonlinear equation for the depth $\|\mathbf{X}\|$, and could be solved using nonlinear optimization. However, we found this to be very time consuming and often did not converge to a good solution. As such we made the following manipulations, followed by a linear approximation which solved both of these problems.

First we change the variable of integration in Eq. 20 to $dr = \|X\|ds$ giving

$$I^\dagger(\mathbf{x}) = \|\mathbf{X}\|F^c \int_0^1 \frac{e^{-\sigma||D(\|\mathbf{X}\|s\hat{\mathbf{X}})||}}{||D(\|\mathbf{X}\|s\hat{\mathbf{X}})||^2} P(\alpha') e^{-\sigma\|\mathbf{X}\|s} ds \tag{21}$$

where α' is a function of s. Next we let $I_b^\dagger$ be defined as the integral in Eq. 21, and remove the global parameter F^c by normalization, giving

$$I^\dagger(\mathbf{x}) = \|\mathbf{X}\| I_b^\dagger(\mathbf{x}) \tag{22}$$

Finally, we solve Eq. 22 for $\|\mathbf{X}\|$ by making the approximation that $I_b^\dagger(\mathbf{x})$ depends on the constant depth Z_{ref} instead of the unknown depth Z. Let $Z_s(\mathbf{x}) = ||\mathbf{X}||$ denote our transmission depth cue.

The spatially varying but depth independent term $I_b^\dagger(\mathbf{x})$ in the approximation can be seen as a non-uniform correction factor to the standard DCP depth algorithm. Although this correction can be neglected in mild scattering, in dense scattering, the backscatter varies spatially and cannot be ignored (see Fig. 4). As such, our transmission depth can achieve better results than both the classical DCP and the improved underwater DCP.

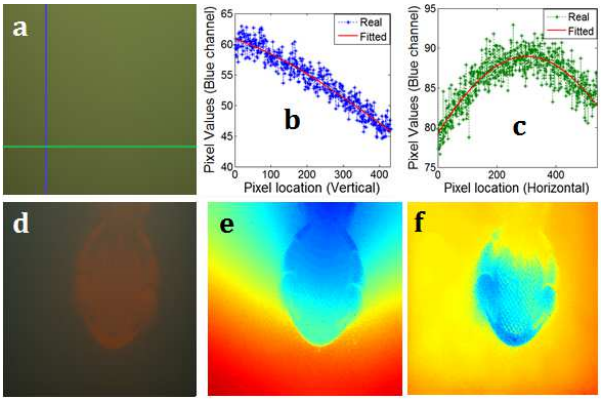

Figure 4. *(a) A scene with no object yields a backscatter only image. (b,c) Plots of the blue and green cross sections show that the backscatter is not spatially uniform. (d) An image of a toy lobster in a dense scattering media. (e) Depth from UDCP [6]. (f) Depth from our new non-uniform corrected transmission depth cue. Note the incorrect low frequency height variations from top to bottom and center out in (e), caused by the non-uniform backscatter, which are absent in ours (f)*

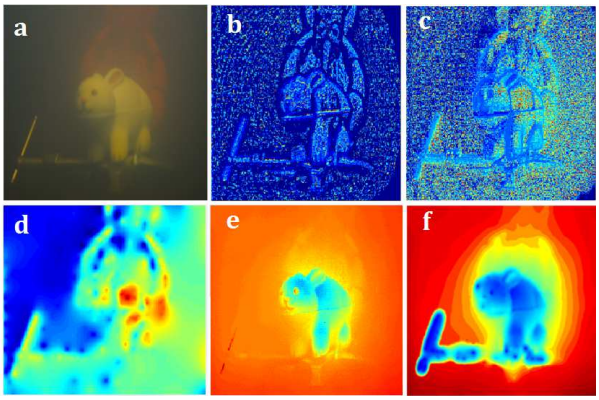

Figure 5. *Combining depth cues. We see that any one depth cue is not enough, and only after combing all three we get good results. (a) Input image. (b) Defocus depth. (c) Correspondence depth. (d) Combined LF defocus and correspondence depth. (e) Our new transmission depth. (f) All three depth cues combined.*

5.3. Depth Fusion and Propagation

As can be seen in Fig. 5, any one depth cue on its own does not produce a reliable depth estimate. However, by combining all three complementary cues, we can recover good depth estimates.

To combine the defocus, correspondence, and transmission depth cues, we need to determine the confidence weights $\Gamma(\mathbf{x})$ of each component. $\Gamma_D(\mathbf{x})$ and $\Gamma_V(\mathbf{x})$ are

given in [29]. We define our transmission confidence as,

$$\Gamma_S(\mathbf{x}) = \sum_{\substack{c_1=\{R,R,G\} \\ c_2=\{G,B,B\}}} |J^{c_1}(\mathbf{x},\mathbf{0}) - J^{c_2}(\mathbf{x},\mathbf{0})| \quad (23)$$

The intuition for this is that properly restored images will be less white than images with backscatter. Thus if the color channels are very similar then there is likely to still be backscatter.

The goal now is to combine different depth cues and to propagate information to regions with low confidence (using a smoothness term). The final depth is obtained by,

$$\min_Z \left(\lambda_1 \sum_{j=D,V,S} \Gamma_j ||Z - Z_j||^2 + \lambda_2 ||Z \otimes \Delta||^2 \right) \quad (24)$$

where $\otimes$ is the convolution operator and Δ is the discrete Laplacian. We used $\lambda_1 = \lambda_2 = 1$ in our implementation, and solved (Eq. 24) using the Trust-Region-Reflective Algorithm [4].

Now we can use the estimated depths to obtain an improved image restoration.

6. High Quality Images from Shearing and Refocusing LF

Although the backscatter and attenuation are removed by restoration, the resulting images are often noisy (see Fig. 6 and Fig. 7 (b)), especially for densely scattering media. This is because the backscatter takes up most of the dynamic range of the camera, and the attenuated signal from the scene is weak. To recover a single high quality image, we propose shearing and averaging multiple views of the light field.

Without scattering, a point on a Lambertian surface will appear the same from different views of an in focus LF image. On the other hand, in a scattering medium, these different views will have slightly different scattering paths, and thus the point may appear different. In densely scattering media, after our restoration, the effect becomes very noticeable, as shown in Fig. 6. Our restoration algorithm does not explicitly handle this variation of scattering with respect to LF view. However, by combining the different information from multiple views, we can produce a single, all in focus, high quality image with less noise than what is possible from a single view (Fig. 7).

For comparison, we also show that simply averaging the different views into a single 2D image reduces the noise, but adds blur since the rays of the LF have not been properly aligned (Fig. 7 (c)). In section 7 we will show another comparison that simulates noise reduction from a single large aperture, as in a standard DSLR.

7. Experimental Results

First we did a simple simulation experiment to verify our backscatter removal under a wide variety of scattering con-

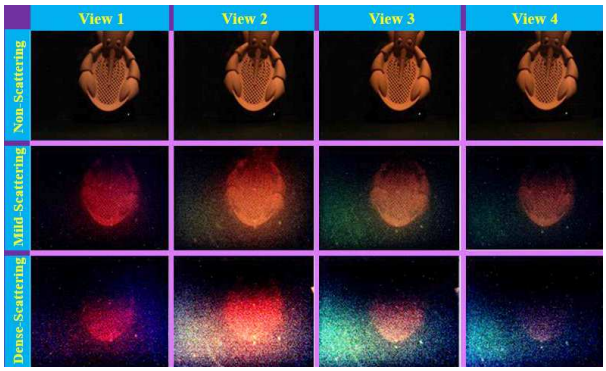

Figure 6. *Different views of LF data under different degree scatterings. We can see that without scattering, different views are similar and with good image quality. With scattering, the restored images from different views are different and affected by noise, and this phenomenon becomes more severe with more scattering.*

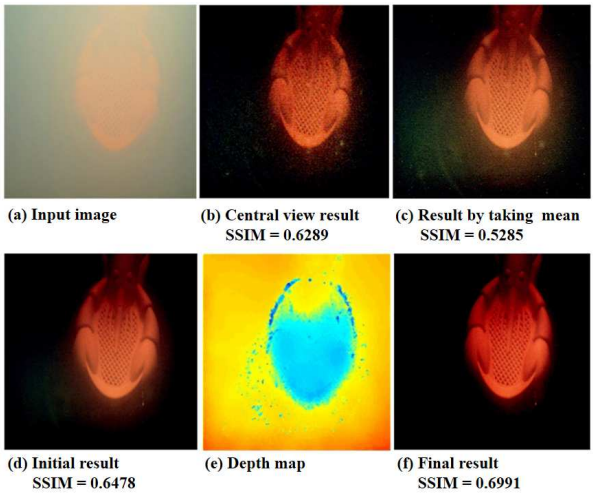

Figure 7. *Restoration (d) by shearing and refocusing on initially haze removed LF data can produce a better result than restoration only using the center view image (b) or simply taking the mean of different view directions (c). The final restoration result (f) by repeating the same shear and refocus procedure and incorporating the estimated depth (e) can produce an even clearer result.*

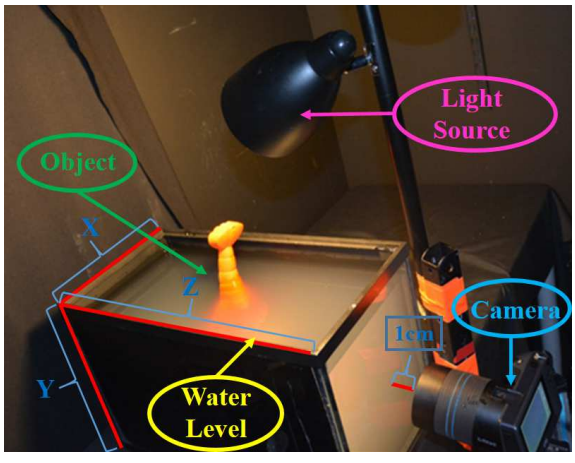

Figure 8. *Our experimental setup.*

Figure 9. *Image restoration comparison. Results for all methods except our final results are obtained from the central LF view only. We can see that our central view results are better than previous methods. None of the other methods can recover the uniform black background. Our final results are even better than our central view results because they contain less noise.*

ditions. We rendered 2100 images of near-field backscatter using the single scattering model [26] and randomly selecting the three parameters in the ranges $F \in (100, 350)$, $\sigma \in (0.001, 0.1)$, and $g \in (0.5, 1)$. We then evaluated how well our restoration method, He'method [8], and Tsiotsios's method [31] were able to recover the pure black background. The mean squared error (MSE) are 2.45, 36.71, and 79.18 respectively for pixels in the range [0,255]. Furthermore, our method recovers the true scattering parameters to within 1% in the simulation. We can see that our method out performs existing methods, and is robust in a wide variety of media.

Second, we demonstrate our method using real experiments conducted in a glass tank filled with tap water (see Fig. 8). A Lytro Illum light field camera with an $18mm$ lens was placed $1cm$ away from the tank. The light source was an incandescent light bulb placed at $X_s = 5cm, Y_s = 65cm$, and $Z_s = 10cm$. The turbidity was increased by adding varying amounts of milk.

We compare our restoration on real captured images with four single-image based methods: He's method [8], Meng's method [13], UDCP method [6], and Tsiotsios's method [31]. We also compare our depth results with two transmission based methods [8, 6] and two LF based methods: Tao's method [28] and Wang's method [32]. Since we are the first to do LF depth estimation in a scattering media, we also compare to a simple "combined" method that first removes the backscatter using [31] and then does depth estimation using [28]. We also note that our method roughly recovers the scattering parameters as predicted by [16] (within 10%) which is reasonable since our method is not optimized for this task.

Fig. 9 shows comparisons of our restoration with existing methods. We can see that the three methods that are based on Narasimhan–Nayar model [8, 13, 6] cannot account for the non-uniform backscatter, and thus don't improve the image quality that much. Tsiotsios's method [31] is better able to remove the scattering, but fails in dense scattering where their quadratic approximation is no longer valid. Our method consistently produces the best results. And after shearing and refocusing multiple views our results look even better.

Fig. 10 shows comparisons of our depth with existing methods. Similar to the restoration shown in Fig. 9, [8, 6] cannot handle the non-uniform scattering and thus do not give a uniform depth to the background. Also note that the same non-uniformity can be seen as a residual height gradient across the objects (bluer towards the top). The LF methods [28, 32] don't take scattering into account and thus produce bad depth. We also did a simple combined experiment for depth estimation (shown as "combined" in Fig. 10), which is haze removed by [31] followed by depth estimation by [28]. From the comparison, we can see that our method produces by far the best depth estimates.

Fig. 11 shows a plot of turbidity level vs image quality (measured using SSIM) for each of the restoration algorithms. SSIM [33] is a widely used metric that is more consistent with human perception than other metrics like mean squared error (MSE) and peak signal-to-noise ratio (PSNR). We use an image captured in clear water as ground truth. We can see that our restored central view is better than all previous restoration methods and degrades more gracefully with increasing turbidity. By shearing and refocusing multiple views we get even better results across all turbidity

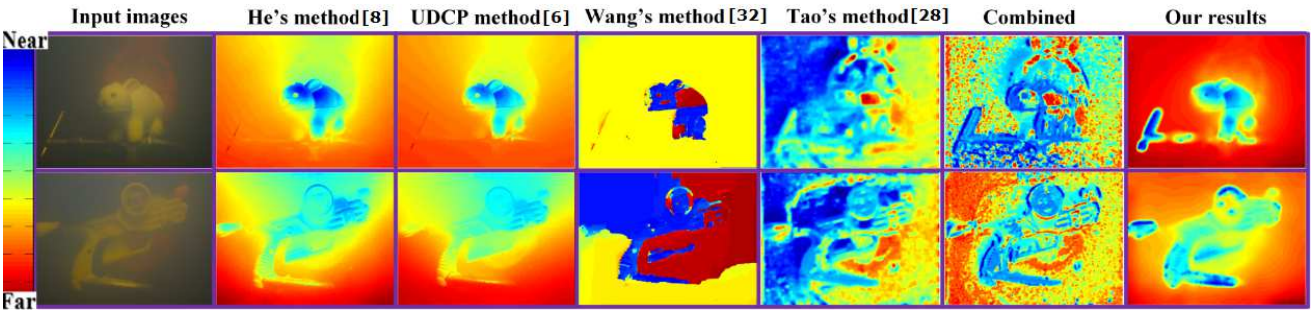

Figure 10. *Depth estimation comparison. We can see that our method produces better results than recent related methods as well as the simple "combined" method both in details and uniformity of the background. Note that the depths are relative and thus the colormaps do not correspond to the same absolute depths.*

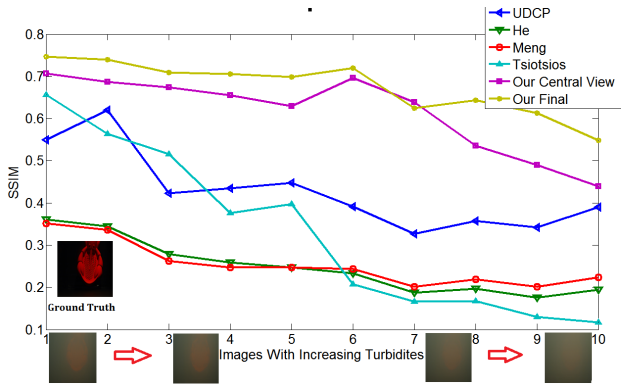

Figure 11. *Plot of image quality vs increasing turbidity for various restoration methods. The thumbnails give sample input images for turbidity levels 0,1,3,7 and 9. It is clear that our restorations are better than previous methods across a broad range of turbidities.*

levels.

Shearing and refocusing reduces noise by averaging over many small apertures. For fair comparison we compare our method with a baseline that simulates a single large aperture with less noise (Fig. 12). We shear and refocus the light field to a single 2D image, which we then remove the backscatter from using our backscatter subtraction. Our method clearly out performs the large aperture baseline showing that combining information from many sub-apertures of a descattered LF is better than using a single large aperture.

In summary, our transmission depth cue can effectively be combined with the standard LF depth cues to produce superior depth estimates than existing methods. These better depth estimates combined with our near-field illumination scattering model allow for better haze removal than existing methods. Finally, by shearing and refocusing our dehazed LF's we can achieve better restorations than is possible from a single large aperture image as well as each single view.

8. Conclusion

We have proposed the use of light field imaging in a scattering medium. Towards this end we have made three main contributions: an image restoration algorithm for near-field illumination, a novel transmission depth cue which we combine with existing light field cues, and a multi view image

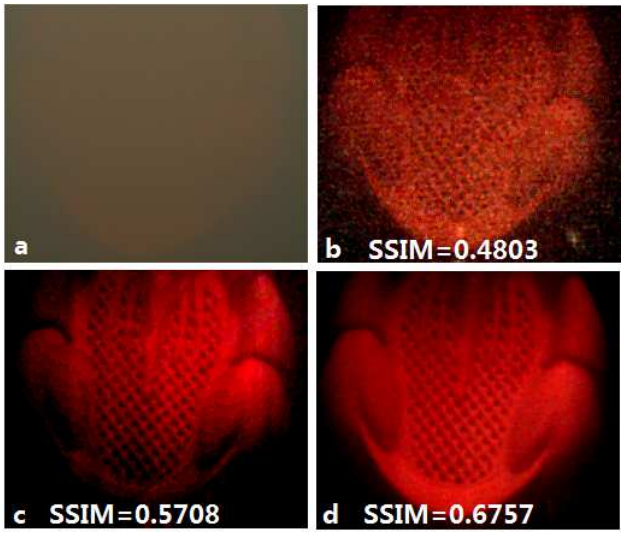

Figure 12. *Comparison of LF dehazing with effective large aperture single image restoration. (a) central view of input LF (b) our descattering applied to the central view (c) our descattering applied to the sheared and refocused effective large aperture 2D single image (d) our proposed method of descattering the LF followed by shearing and refocusing. We can see that our method produces the best results.*

fusion procedure for improved signal to noise ratio. Our image restoration algorithm outperforms existing methods, particularly for densely scattering media where prior methods often fail. Our combined depth estimation produces better depth estimates than previous methods. And our sheared and averaged final image has less noise and has higher quality than other methods for restoring single view images.

Acknowledgments

This work was supported by the Natural Science Foundation of China under Grant Nos. 61473280 and 91648118, the US Office of Naval Research grants N000141512013, N000141712687, the UC San Diego Center for Visual Computing, and a gift from Qualcomm. The authors also thank the support from Youth Innovation Promotion Association CAS.

References

[1] M. Bryson, M. Johnson-Roberson, O. Pizarro, and S. B. Williams. Colour-consistent structure-from-motion models using underwater imagery. *Robotics: Science and Systems VIII*, page 33, 2013. 1

[2] N. Carlevaris-Bianco, A. Mohan, and R. M. Eustice. Initial results in underwater single image dehazing. In *OCEANS 2010 MTS/IEEE SEATTLE*, pages 1–8. IEEE, 2010. 2

[3] J. Y. Chiang and Y.-C. Chen. Underwater image enhancement by wavelength compensation and dehazing. *IEEE Transactions on Image Processing*, 21(4):1756–1769, 2012. 2

[4] T. F. Coleman and Y. Li. A reflective newton method for minimizing a quadratic function subject to bounds on some of the variables. *SIAM Journal on Optimization*, 6(4):1040–1058, 1996. 6

[5] D. G. Dansereau, O. Pizarro, and S. B. Williams. Linear volumetric focus for light field cameras. *Acm Transactions on Graphics*, 34(2):1–20, 2015. 2

[6] P. Drews, E. Nascimento, F. Moraes, S. Botelho, and M. Campos. Transmission estimation in underwater single images. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pages 825–830, 2013. 1, 2, 5, 7

[7] R. Fattal. Dehazing using color-lines. *ACM Transactions on Graphics (TOG)*, 34(1):13, 2014. 2

[8] K. He, J. Sun, and X. Tang. Single image haze removal using dark channel prior. *IEEE transactions on pattern analysis and machine intelligence*, 33(12):2341–2353, 2011. 2, 4, 5, 7

[9] L. G. Henyey and J. L. Greenstein. Diffuse radiation in the galaxy. *The Astrophysical Journal*, 93:70–83, 1941. 3

[10] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright. Convergence properties of the nelder–mead simplex method in low dimensions. *SIAM Journal on optimization*, 9(1):112–147, 1998. 4

[11] C. Liu and W. Meng. Removal of water scattering. *International Conference on Computer Engineering and Technology (ICCET), IEEE Transactions on*, 2:35–39, 2010. 2

[12] H. Lu, Y. Li, L. Zhang, and S. Serikawa. Contrast enhancement for images in turbid water. *JOSA A*, 32(5):886–893, 2015. 2

[13] G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan. Efficient image dehazing with boundary constraint and contextual regularization. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 617–624, 2013. 7

[14] A. Mousnier, E. Vural, and C. Guillemot. Partial light field tomographic reconstruction from a fixed-camera focal stack. *arXiv preprint arXiv:1503.01903*, 2015. 1, 2

[15] Z. Murez, T. Treibitz, R. Ramamoorthi, and D. Kriegman. Photometric stereo in a scattering medium. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3415–3423, 2015. 1, 2

[16] S. G. Narasimhan, M. Gupta, C. Donner, R. Ramamoorthi, S. K. Nayar, and H. W. Jensen. Acquiring scattering properties of participating media by dilution. In *ACM Transactions on Graphics (TOG)*, volume 25, pages 1003–1012. ACM, 2006. 3, 4, 7

[17] S. G. Narasimhan and S. K. Nayar. Contrast restoration of weather degraded images. *IEEE transactions on pattern analysis and machine intelligence*, 25(6):713–724, 2003. 2

[18] S. G. Narasimhan, S. K. Nayar, B. Sun, and S. J. Koppal. Structured light in scattering media. In *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, volume 1, pages 420–427. IEEE, 2005. 1, 3, 4

[19] S. Negahdaripour and A. Sarafraz. Improved stereo matching in scattering media by incorporating a backscatter cue. *IEEE Transactions on Image Processing*, 23(12):5743–5755, 2014. 2

[20] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan. Light field photography with a hand-held plenoptic camera. *Computer Science Technical Report CSTR*, 2(11):1–11, 2005. 4

[21] N. Otsu. A threshold selection method from gray-level histograms. *Automatica*, 11(285-296):23–27, 1975. 4

[22] M. Roser, M. Dunbabin, and A. Geiger. Simultaneous underwater visibility assessment, enhancement and improved stereo. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3840–3847. IEEE, 2014. 2

[23] A. Sarafraz, S. Negahdaripour, and Y. Y. Schechner. Enhancing images in scattering media utilizing stereovision and polarization. In *Applications of Computer Vision (WACV), 2009 Workshop on*, pages 1–8. IEEE, 2009. 1

[24] Y. Y. Schechner and N. Karpel. Clear underwater vision. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 1, pages I–536. IEEE, 2004. 2

[25] S. Serikawa and H. Lu. Underwater image dehazing using joint trilateral filter. *Computers & Electrical Engineering*, 40(1):41–50, 2014. 2

[26] 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(3):1040–1049, 2005. 3, 7

[27] Y. Swirski and Y. Y. Schechner. 3deflicker from motion. In *Computational Photography (ICCP), 2013 IEEE International Conference on*, pages 1–9. IEEE, 2013. 2

[28] M. W. Tao, S. Hadap, J. Malik, and R. Ramamoorthi. Depth from combining defocus and correspondence using light-field cameras. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 673–680, 2013. 1, 2, 4, 7

[29] M. W. Tao, P. P. Srinivasan, J. Malik, S. Rusinkiewicz, and R. Ramamoorthi. Depth from shading, defocus, and correspondence using light-field angular coherence. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1940–1948. IEEE, 2015. 4, 6

[30] T. Treibitz and Y. Y. Schechner. Active polarization descattering. *IEEE transactions on pattern analysis and machine intelligence*, 31(3):385–399, 2009. 2

[31] C. Tsiotsios, M. E. Angelopoulou, T.-K. Kim, and A. J. Davison. Backscatter compensated photometric stereo with 3 sources. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2251–2258, 2014. 1, 2, 3, 4, 7

[32] T.-C. Wang, A. A. Efros, and R. Ramamoorthi. Occlusion-aware depth estimation using light-field cameras. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3487–3495, 2015. 1, 2, 7

[33] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to

structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004. 7

[34] S. Wanner and B. Goldluecke. Globally consistent depth labeling of 4d light fields. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 41–48. IEEE, 2012. 2

[35] H.-Y. Yang, P.-Y. Chen, C.-C. Huang, Y.-Z. Zhuang, and Y.-H. Shiau. Low complexity underwater image enhancement based on dark channel prior. In *Innovations in Bio-inspired Computing and Applications (IBICA), 2011 Second International Conference on*, pages 17–20. IEEE, 2011. 2

[36] Z. Yu, X. Guo, H. Lin, A. Lumsdaine, and J. Yu. Line assisted light field triangulation and stereo matching. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2792–2799, 2013. 1

[37] Q. Zhu, J. Mai, and L. Shao. A fast single image haze removal algorithm using color attenuation prior. *IEEE Transactions on Image Processing*, 24(11):3522–3533, 2015. 2