LIGHT FIELD IMAGE RESTORATION FOR VISION IN SCATTERING MEDIA

Vigil Varghese[⋆]*, Donald G. Dansereau*[†]*, Mitch Bryson*[⋆]*, Oscar Pizarro*[⋆]*, and Stefan B. Williams*[⋆]

[⋆]Australian Centre for Field Robotics, The University of Sydney, Australia
[†]Department of Electrical Engineering, Stanford University, USA

ABSTRACT

Recovering information from contrast-limited, SNR-limited, color-attenuated images in a scattering media is of paramount importance for the autonomous functioning of robotic agents. The task is challenging due to the transient state of the medium, unknown medium parameters and in many cases the need for fully autonomous operation. This work presents a target-less, calibration-less method for restoring underwater light field images and requires no explicit model of the medium. The method adopts a light-field imaging approach to capture, model and compensate for backscatter in the scene leading to the recovery of high-fidelity images. The proposed method for backscatter compensation is validated against other state-of-the-art methods and is demonstrated to yield superior image quality.

***Index Terms*—** Image restoration, image enhancement, scattering media, light fields, underwater imaging.

1. INTRODUCTION

Recovering high fidelity images from a scattering medium such as rain, fog, haze, snow, etc., is a complex engineering task. Yet, it is one of timely importance owing to the increase in the autonomous operation of robotic agents on land (autonomous cars, trucks, delivery vehicles), air (drones, unmanned aerial vehicles) and water (autonomous underwater vehicles). These robotic agents are designed to function in "fair-weather" conditions and the systems they depend on to function (object-detection, collision avoidance, visual SLAM, visual odometry, etc.,) become unreliable under harsh environmental conditions.

Imaging in scattering media is a challenging proposition due to a reduction in SNR owing to the particulate matter in the medium along with a loss of image contrast - both of which are scene depth dependent, a point driven home by Figure 1(a). We consider highly turbid water as a scattering medium of interest and attempt to restore images captured in this medium (Figure 1(b)). Turbid water is an excellent choice of study owing to the fact that light not only undergoes absorption but also scattering, both on its way from the illuminant to the scene (backscatter) and from the scene to the sensor (forward scatter), imposing severe challenges to the restoration techniques.

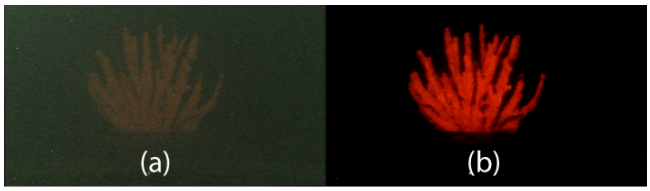

Fig. 1: (a) Representative image showing extent of degradation when imaged under turbid conditions [1]. (b) Restored image by using the techniques presented in this paper.

A recent trend in underwater imaging has been to adopt the dark channel prior (DCP) as proposed by He et al. [2] for aerial haze removal. In the presence of haze, application of the DCP prior results in elimination of the direct attenuation component enabling estimation of the transmission parameter, leading to a restored image. Drews et al. [3] introduce an underwater dark channel prior (UDCP) based on the green and blue channels by noting that the absorption in the red channel is significant over the imaging range and is an unreliable estimate of scene statistics.

Hardware based approaches such as [4], [5], [6] and [7] require additional optical components such as multiple illumination sources or polarization filters and need to capture multiple snap-shots of the scene in order to estimate the medium parameters. These techniques impose additional complexity on power-limited, real-time autonomous systems which has limited their widespread use in real-world applications.

In [1], Skinner et al. use light field cameras [8] to capture images in turbid media and restore those images using the dehazing model of Li et al. [9]. The compact form factor of light field cameras together with their sophisticated algorithms make them a potential candidate for use in real-world applications.

This work proposes an image restoration pipeline based on estimating the backscatter component by utilizing the depth selectivity afforded by light field imaging to overcome occluders and noise. Our main contributions are:

1. A light field image restoration pipeline that is based on backscatter elimination using adaptive depth selective

This work was supported by the Australian Research Council Discovery Project, DP150104440.

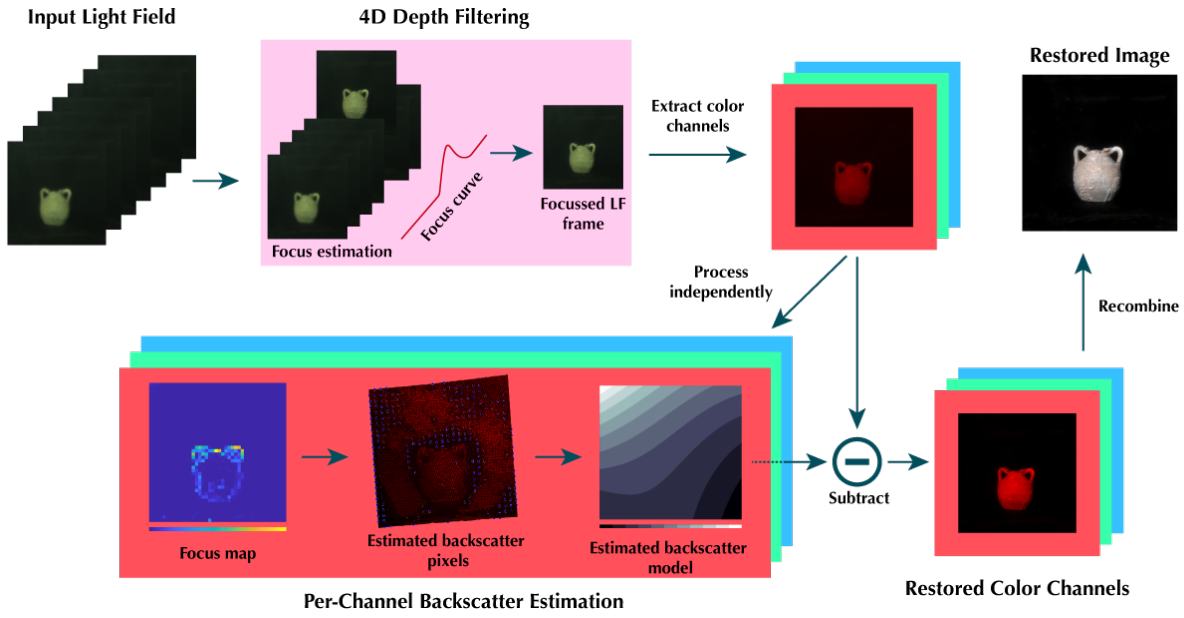

Fig. 2: The major stages of the light field image restoration pipeline. The input light field is filtered based on the depth range of the objects in the scene generating a focussed light field frame. Each color channel of the focussed frame is processed independently to recover the backscatter-eliminated image.

hyperfan filters.

2. A robust technique that handles high turbidity in underwater images for a variety of object depth ranges.

Unlike existing techniques that require *in-situ* calibration targets, additional components such as transmissometers or multiple illumination sources, the proposed model is a targetless, calibration-less method. However, the method works best for scenes for which at least some pixels are backscatter dominated, i.e. some parts of the frame see distant objects. It also does not handle color dependent attenuation, which is of lesser concern in highly turbid medium as backscatter components dominate the image degradation process. The system requires no additional moving components, power sources or external optics, making it highly robust and feasible for extended operation in challenging conditions. Further, the model is generic enough to be extensible for aerial applications with minor modifications.

2. LIGHT FIELD IMAGE RESTORATION MODEL

The input to the model is a light field and the output is a backscatter-eliminated, color-corrected image. Figure 2 highlights the main stages in the light field image restoration pipeline. The input light field is swept through its depth range to identify the depths at which one or more objects might exist. Once the depth range of an object in the scene is known, a focussed frame is generated by using a hyperfan filter [10]. The resulting focussed image is then processed independently per color channel to recover a backscatter-eliminated image and finally recombined to generate a restored image.

2.1. Hyperfan volume filtering

The first stage in the pipeline achieves 4D depth filtering by using a hyperfan volume filter along with a grey level local variance filter as the focus measure. Unlike planar focus filters, hyperfan filters allow focusing on a range of depths while improving SNR [10]. The true benefit of employing a volumetric filter like the hyperfan is its ability to *see through* occluders that exist over the small camera baseline.

The hyperfan (HF) filter is a volumetric depth selective filter formed by the intersection of a dual-fan (DF) and a hypercone (HC) filter [10]. The passband of a hyperfan filter is given by:

$$H_{HF}(\omega,\theta) = H_{HC}(\omega)H_{DF}(\omega,\theta) \tag{1}$$

where $H_{HC}(\omega)$ is the passband of the hypercone filter, and is given by:

$$H_{HC}(\omega) = \exp(-[\frac{(\omega_s\omega_v - \omega_t\omega_u)}{\beta_{HC}^2/\sqrt{2ln2}}]^2) \tag{2}$$

and $H_{DF}(\omega,\theta)$ is the passband of the dual-fan filter which in turn is formed by the superposition two 2D fan filters [11]:

$$H_{DF}(\omega,\theta) = H_{FAN}^{2D}(\omega_s,\omega_u,\theta_1,\theta_2)H_{FAN}^{2D}(\omega_t,\omega_v,\theta_1,\theta_2) \tag{3}$$

$\omega_{u,v,s,t}$ are the continuous-domain frequency space variables corresponding to the two-plane parameterization model. θ_1 and θ_2 determine the angular range of the epipolar lines and help to select a depth-volume. β_{HC} is the 3-dB bandwidth of

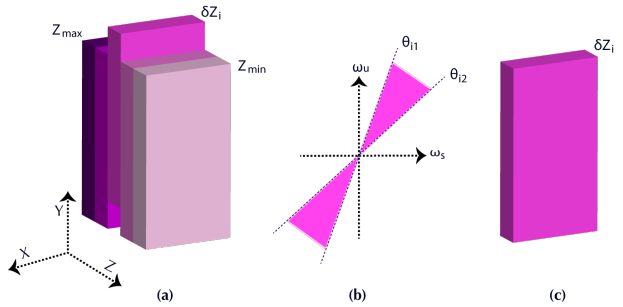

Fig. 3: (a) The input light field exists over a total depth volume from Z_{min} to Z_{max}. (b) The frequency range over which a focussed depth volume, δZ_i, exists. (c) Hyperfan filtered focussed depth volume, δZ_i.

the hypercone passband, and determines the filter's selectivity to noise.

2.2. Adaptive 4D depth filtering

The object space can be considered to lie within a depth range of Z_{min} and Z_{max} that corresponds to the angular range of θ_{min} and θ_{max} in the frequency space (Figure 3). We consider the entire depth range to be made up of incremental depth-volumes each with a range δZ or equivalently $\delta\theta$. Figure 3 highlights one such depth-volume, δZ_i (Figure 3(a)) and its range in the frequency space ω_s, ω_u (or ω_t, ω_v), which is, θ_{i1}, θ_{i2} (Figure 3(b)).

We sweep the total light field depth volume from θ_{min} to θ_{max} and find the grey level local variance for each incremental depth volume, $\delta\theta$. Grey level local variance is a focus measure that estimates the variability in the grey level values over a local region. This is based on the principle that when a scene is in perfect focus, the local grey level variance will be high as opposed to when it is out of focus [12].

The focus measure is applied to a windowed region $\Omega(x, y)$ of size $N \times N$ pixels (15×15 here) centered at (i, j). μ is the mean over the local region $\Omega(x, y)$.

$$\phi(x,y) = \frac{1}{N^2} \sum_{(i,j)\in\Omega(x,y)} (I(i,j) - \mu)^2 \qquad (4)$$

After the sweep, the depth-volume with maximum variance is considered to be a potential candidate for the final depth volume. For the simplest case of a planar object, the focus curve peaks for a single $\delta\theta$ and the object is considered to lie within this $\delta\theta$. When the object spans multiple $\delta\theta$'s the focus curve reaches its peak over multiple $\delta\theta$'s and all these $\delta\theta$'s are considered to be potential candidates (Figure 2). Valid candidates are considered by thresholding the focus curve, which gives the depth volume in which the object lies.

Once the depth volume is known, the hyperfan filter is applied to the input light field with the known values of θ_{i1} and θ_{i2} to render a focussed frame.

2.3. Backscatter model estimation

Once we have the focussed frame we apply the grey level local variance focus measure over the windowed regions of the frame and derive a focus map (Figure 2). Higher values of this measure indicate regions with objects in sharp focus and lower values indicate regions with slowly varying quantities. The illumination and backscatter components are slowly varying [13], and therefore can be detected by thresholding the focus measure $\phi(x, y)$ by a constant value ρ_ϕ ($1 \times e^{10}$ here).

The success of the restoration framework depends on accurately estimating the backscatter pixels and for real scenes either a backscatter saturation assumption [7] or a dark pixel assumption [3] can be used.

The mean of the windowed region is the estimated backscatter value and the center coordinate is its location.

The estimated backscatter pixels are used to fit a polynomial surface with second degree terms in x and third degree terms in y (Equation 5), relative to the cameras' frame of reference. This two dimensional polynomial function has 9 unknowns ($\alpha_0, \alpha_1, \ldots, \alpha_8$) and depends on the accurate estimation of at least 9 backscatter pixels. For a robust estimate we need as many backscatter pixels as can be reliably determined and a least squares solution is employed to fit the data to the model.

$$\mathcal{S}(x,y) = \alpha_0 + \alpha_1 x + \alpha_2 y + \alpha_3 x^2 + \alpha_4 xy + \alpha_5 y^2 + \alpha_6 x^2 y + \alpha_7 x y^2 + \alpha_8 y^3 \qquad (5)$$

The choice of the polynomial model depends on the system setup, particularly on the position and quantity of the illumination sources used. In principle, the backscatter term is a slowly varying function [7] and a quadratic surface fits typical empirical data. Tsiotsios et al. [7] use a second degree term in x and y to fit the surface model and the choice suits well for their imaging setup.

2.4. Image restoration

The estimated backscatter function can be used to eliminate the backscatter contribution by subtracting it from the captured image. In highly turbid media backscatter dominates the image degradation process and even without any knowledge about the attenuation coefficients lead to a good restored image (Figure 4).

3. RESULTS AND DISCUSSION

Figure 4 compares the proposed technique with other state-of-the-art techniques. Although DCP was not meant for underwater image restoration, it has inspired a number of "prior"-based restoration techniques (such as UDCP [3]) and is shown here for completeness. Backscatter elimination by Tsiotsios et al. [7] eliminates backscatter components

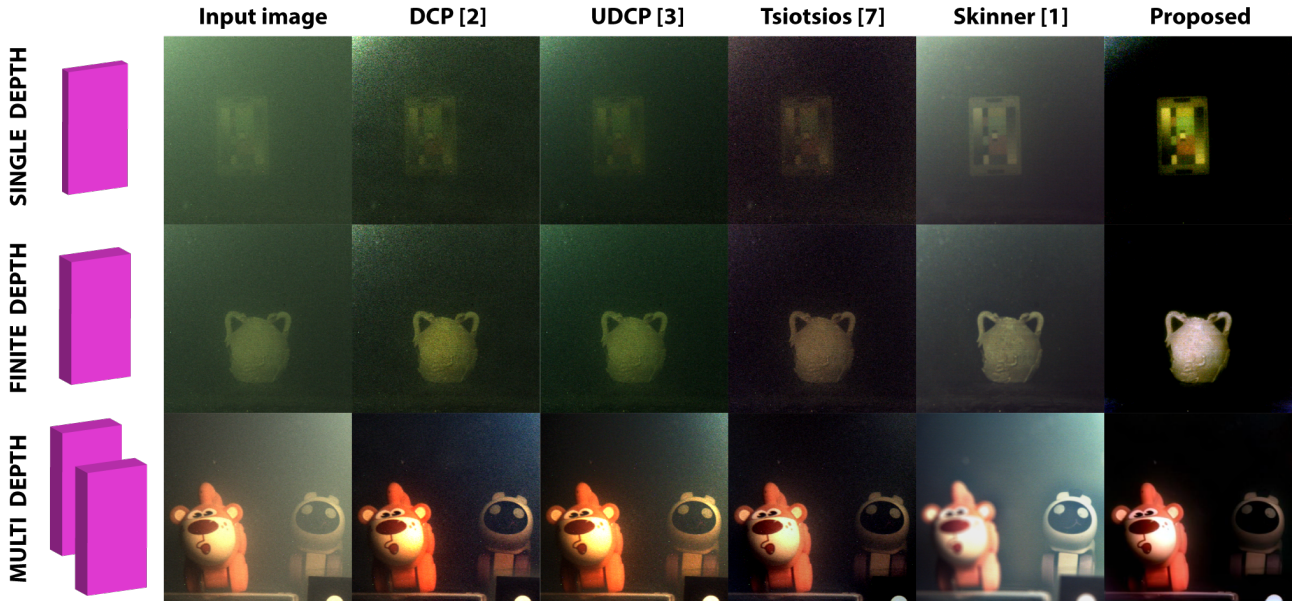

Fig. 4: Comparison of image restoration techniques. On the left are the depth ranges corresponding to the objects in the image. Unlike other techniques that leave large backscatter residuals, the proposed technique removes most of the backscatter component from the images and handles objects at different depth ranges.

Table 1: Spearman correlation coefficient for the red, green and blue channels of the color board.

	Red	**Green**	**Blue**
Input image	0.8039	0.7295	0.2341
DCP [2]	0.8526	0.7866	0.3348
UDCP [3]	0.8460	0.7535	0.1603
Tsiotsios [7]	0.8854	0.8059	0.3411
Skinner [1]	0.8971	0.8036	0.5058
Proposed	**0.9407**	**0.8799**	**0.6259**

that closely match the estimated second degree polynomial model, but leaves behind residuals that do not perfectly match the model. The light field image restoration technique from Skinner et al. [1] reduces backscatter, but is ultimately limited by the poor performance of its core haze removal algorithm [9] in the presence of dense scattering.

Unlike other techniques that leave behind some backscatter residuals, the proposed technique removes most of the backscatter components from the images and handles objects at different depth ranges.

We use the Spearman's rank ordered correlation coefficient [14] as an objective measure of restoration quality which compares the correlation between reconstructed intensities and reference intensities measured in air, which are not affected by scattering or other degradations. We compare the RGB values of different color swatches (red, green, blue, black and white) of a color board imaged in air with that of restored images from different techniques. This eliminates the need for a perfectly co-registerd reference image which is difficult to obtain under measurement conditions making other metrics such as MSE, PSNR or SSIM unreliable.

For each of the color swatches in the restored color board image we randomly select 20 pixel patches (of size 3×3) and determine their Spearman's correlation coefficient with the color swatches captured in air. Table 1 shows the Spearman's correlation coefficient for the different techniques for red, green and blue color channels. As can be observed, our technique outperforms others and is consistent with the subjective results of Figure 4. Blue channel shows weak correlation compared to the red and green channels and is a result of the blue component getting scattered more compared to the red and green components in highly turbid media [15], as is observed consistently across all reconstructed techniques.

4. CONCLUSION

In this paper we have proposed an image restoration pipeline for restoring light field images captured in scattering media. The technique is based on hyperfan volume filtering and polynomial backscatter elimination. By adaptively filtering the light field volume to restrict its depth range around the object of interest, we significantly reduce the effects of noise and restrict backscatter components to lie within a narrow depth range. The proposed method works well under highly turbid conditions and comparisons with other techniques show superior results. We anticipate that the proposed technique will alleviate issues related to imaging under challenging conditions and enable robust operation of autonomous vehicles.

5. REFERENCES

[1] Katherine A Skinner and Matthew Johnson-Roberson, "Underwater Image Dehazing with a Light Field Camera," in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 2017, pp. 1775–1782, IEEE.

[2] Kaiming He, Jian Sun, and Xiaoou Tang, "Single image haze removal using dark channel prior," in *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR Workshops)*. 2009, pp. 1956–1963, IEEE.

[3] P Drews Jr, E do Nascimento, F Moraes, S Botelho, and M Campos, "Transmission Estimation in Underwater Single Images," in *2013 IEEE International Conference on Computer Vision Workshops (ICCVW)*. Jan. 2014, pp. 825–830, IEEE.

[4] O Spier, T Treibitz, and G Gilboa, "In Situ Target-Less Calibration of Turbid Media," *Proc IEEE ICCP*, 2017.

[5] Y Y Schechner and N Karpel, "Recovery of Underwater Visibility and Structure by Polarization Analysis," *Ieee Journal of Oceanic Engineering*, vol. 30, no. 3, pp. 570–587, July 2005.

[6] T Treibitz and Y Y Schechner, "Turbid Scene Enhancement Using Multi-Directional Illumination Fusion," *Ieee Transactions on Image Processing*, vol. 21, no. 11, pp. 4662–4667, Oct. 2012.

[7] Chourmouzios Tsiotsios, Maria E Angelopoulou, Tae Kyun Kim, and Andrew J Davison, "Backscatter compensated photometric stereo with 3 sources," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Imperial College London, London, United Kingdom, Jan. 2014, pp. 2259–2266.

[8] Edmund Y Lam, "Computational photography with plenoptic camera and light field capture: tutorial," *Journal of the Optical Society of America. A, Optics, image science, and vision*, vol. 32, no. 11, pp. 2021–2032, 2015.

[9] Yu Li, Robby T Tan, and Michael S Brown, "Nighttime haze removal with glow and multiple light colors," in *Proceedings of the IEEE International Conference on Computer Vision*. National University of Singapore, Singapore City, Singapore, Feb. 2016, pp. 226–234.

[10] Donald G Dansereau, Oscar Pizarro, and Stefan B Williams, "Linear Volumetric Focus for Light Field Cameras," *ACM Transactions on Graphics (TOG)*, vol. 34, no. 2, Feb. 2015.

[11] Donald Dansereau and Len T Bruton, "A 4-D dual-fan filter bank for depth filtering in light fields," *Ieee Transactions on Signal Processing*, vol. 55, no. 2, pp. 542–549, Feb. 2007.

[12] J L Pech-Pacheco, G Cristobal, J Chamorro-Martinez, and J Fernandez-Valdivia, "Diatom autofocusing in brightfield microscopy: a comparative study," in *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*. pp. 314–317, IEEE Comput. Soc.

[13] Hanumant Singh, Chris Roman, Oscar Pizarro, Ryan Eustice, and Ali Can, "Towards high-resolution imaging from underwater vehicles," *International Journal of Robotics Research*, vol. 26, no. 1, pp. 55–74, Jan. 2007.

[14] J.D. Gibbons and S. Chakraborti, *Nonparametric Statistical Inference*, Statistics, textbooks and monographs. Marcel Dekker Incorporated, 2003.

[15] Linda Mullen, Brandon Cochenour, Alan Laux, Derek Alley, "Optical modulation techniques for underwater detection, ranging and imaging," in *Proc.SPIE*, 2011, vol. 8030, pp. 8030 – 8030 – 9.