

Refractive Shape from Light Field Distortion

Gordon Wetzstein¹

David Roodnick¹

Wolfgang Heidrich¹

Ramesh Raskar²

¹University of British Columbia

²MIT Media Lab

Abstract

Acquiring transparent, refractive objects is challenging as these kinds of objects can only be observed by analyzing the distortion of reference background patterns. We present a new, single image approach to reconstructing thin transparent surfaces, such as thin solids or surfaces of fluids. Our method is based on observing the distortion of light field background illumination. Light field probes have the potential to encode up to four dimensions in varying colors and intensities: spatial and angular variation on the probe surface; commonly employed reference patterns are only two-dimensional by coding either position or angle on the probe. We show that the additional information can be used to reconstruct refractive surface normals and a sparse set of control points from a single photograph.

1. Introduction

The reconstruction of transparent, refractive, and specular surfaces from photographs has been a target for active investigation in computer vision, but also other areas, including computer graphics and fluid imaging.

One strategy for dealing with such surfaces is to alter the reflectance or transmission characteristics of the surface under investigation to simplify the scanning. This can be achieved through coating with diffuse materials [9] or immersion in special liquids [32, 13]. However, such intrusive methods are not always desirable or feasible, for example when the object under investigation is itself a liquid.

A popular alternative is the analysis of refractive distortions of diffuse background or illumination pattern [23, 28, 5, 1, 17, 21, 31, 6, 22]. Such approaches typically require multiple cameras, or multiple images from the same camera taken with varying illumination or background patterns.

In our work, we aim for a single camera, single image method more similar in spirit to photometric stereo [34], and especially to single-image variants using colored light sources [10]. We propose to reconstruct transparent surfaces from the observed distortion of higher-dimensional reference patterns, called light field probes [33]. These probes can encode the 2D spatial and the 2D angular do-

main on their surface; possible implementations include lenslet arrays, parallax-barriers, or holograms. The distortion of a light field emitted by such a probe allows us to simultaneously reconstruct the normals and a sparse set of absolute 3D points representing either a single refractive boundary surface or a thin refractive solid. Specifically, our method combines the following characteristics:

- Only a single image is required for reconstruction, making the method suitable for both static and dynamic surface reconstruction using a single, calibrated camera.
- Our reconstruction assumes a single refractive event along each camera ray. Thin solids can be reconstructed by applying an approximation similar to the thin lens model in optics.
- The acquisition setup is inexpensive due to the use of a single camera and a light field probe that is easily manufactured from off-the-shelf parts.

2. Related Work

Light fields are ray-based representations of the 4D spatio-angular light distribution [18]. Lenslet arrays and parallax-barriers in combination with interlaced 2D sensors or displays have been used for more than a century to record and synthesize light fields [27]. Light field illumination can be integrated into microscopes to produce exotic lighting effects on reflective, microscopic specimen [19]. Recently, light field probes have been used to qualitatively visualize refractions caused by macroscopic, transparent solids and liquids [33]. While these approaches are successful in visualizing complex reflective and refractive events, we are the first to use them for acquiring quantitative data to reconstruct the shapes of certain classes of refractive objects.

Schlieren and phase imaging are non-intrusive, optical approaches to visualize and quantify refraction in transparent media. These techniques have been developed in the fluid mechanics community over centuries [29]. Approaches to phase-contrast microscopy [24], such as Zernike phase contrast and differential interference contrast (DIC), also encode refractions caused by transparent specimen in changes of intensity and color. Traditional and

background-oriented Schlieren [7] techniques are successful in coding *two-dimensional* light ray deflections with a high precision [12]. Light field probes, as employed in this paper, can encode up to *four dimensions*, thereby allowing both normals and some positions of thin refractive surfaces to be reconstructed independently. As opposed to phase-contrast methods, light field probes do not require coherent illumination and are fabricated from off-the-shelf parts.

Transparent and specular object reconstruction has recently gained a lot of traction [16]. Kutulakos and Steger [17] analyze the space of these reconstructions based on acquisition setup and number of refractive events in the optical path of light rays. Generally, refractive object capture and reconstruction can be performed using a single camera but multiple images or, alternatively, using multiple cameras. Ben-Ezra and Nayar [5] reconstruct smooth, parameterized refractive objects from the distortions of a diffuse background in an image sequence from a fixed viewpoint. Optical flow can be formulated to account for refraction [1, 30] or reflection [8]. Miyazaki and Ikeuchi [20] and Huynh et al. [14] exploit the polarization of refracted light to estimate transparent surfaces. A tomographic reconstruction of transparent solids from multiple images was proposed by Trifonov et al. [32]. Ihrke et al. [15] compute the shape of flowing water by dying it with fluorescent chemicals. Range scanning can be used for the acquisition of refractive solids, if they are immersed in a fluorescent liquid [13]. Morris and Kutulakos show that the surface of complex refractive objects can be reconstructed from multiple photographs with changing illumination [22]. Furthermore, specular objects can be acquired using shape from distortion [31, 6]. Multiple cameras have been used for dynamic refractive stereo [21] and for the reconstruction of smooth gas flows [4]. As opposed to all of these approaches, our method only requires a single image.

Alternative **single image reconstruction** techniques include the seminal work by Murase [23], where a wavy water surface is reconstructed by observing the distortions of a diffuse probe under water with an orthographic camera. Zhang and Cox [35] also reconstruct a water surface with an orthographic camera by placing a big lens and a 2D screen at its focal length in the water. This allows the surface gradients to be measured, which can subsequently be integrated to compute the surface shape. For both approaches the mean water level needs to be known. Savarese et al. [28] present an analysis of single image reconstruction of smooth mirroring objects using shape from distortion. Compared to these techniques, our approach also assumes that there is only a single refractive or reflective event; however, no constraints are placed on the camera setup. Furthermore, we show how to reconstruct both surface points and normals simultaneously from a captured photograph.

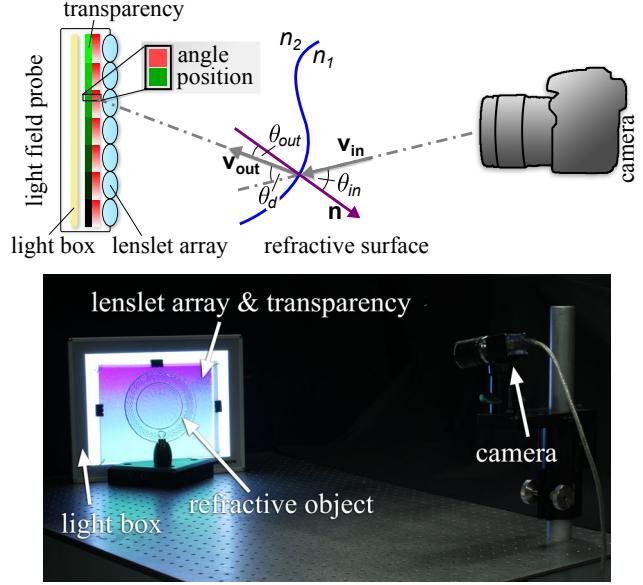


Figure 1. Schematic showing how both position and incident angle of a refracted ray are color coded by a light field probe (top). Our probe prototypes consist of a light box, transparencies, and a lenslet array positioned behind a refractive object.

3. Shape from Light Field Probes

3.1. Coding Light Field Illumination

Light field probes are capable of emitting 4D illumination by encoding the outgoing light ray positions and angles in varying intensities and colors [33]. Standard displays only emit 2D illumination, because the light at each pixel is uniformly displayed to all directions. 4D probes can, for instance, be implemented by mounting high-resolution transparencies on a light box behind a lenslet array (see Fig. 1). This approach does not increase the total number of display pixels, but distributes them between spatial and angular resolution. The number of pixels under each lenslet corresponds to the angular probe resolution, while the size of the lenslets determines the spatial resolution. Other hardware implementations, such as holograms, have the potential to overcome the resolution tradeoff of lenslet arrays.

For the purpose of single-shot transparent object reconstruction, the color and intensity codes emitted by a light field probe need to satisfy two important criteria. First, the patterns are required to uniquely encode position and angle on the probe surface, so that a camera measures this information in a single image. Second, in order to account for slight miscalibrations of the probe prototype, the color codes should be smooth in the 4D spatio-angular domain. We restrict our prototype to readily available hardware, as illustrated in Figure 1, and limit the feasible colors and intensities to the combined printer and camera gamut and dynamic range.

The most intuitive coding scheme satisfying the above requirements are color gradients. In our implementation, we use red, blue, and green gradients to code the 2D directions and a 1D vertical position, respectively. As demonstrated in Section 3.3, the missing second spatial dimension can be recovered through geometric constraints in post-processing. This encoding is illustrated for a 1D case in Figure 1 (top). Here, the incident angle is coded in a shade of red and the position on the probe surface is coded in green.

This simple, yet effective coding scheme allows both angle and position of light rays to be encoded in observed colors and intensities. Without refraction in the optical path, the measured colors at each pixel of a calibrated camera correspond to the information predicted by the calibration, but in the presence of refraction these differ. In the following subsections we show how to reconstruct refractive surfaces from such measurements. The employed color codes ignore the wavelength-dependency of refraction as well as attenuation and scattering caused by the medium.

3.2. Reconstructing Surface Normals

The normal of each surface point imaged by a camera pixel can be computed using Snell's law: $n_1 \sin \theta_{in} = n_2 \sin \theta_{out}$. In our application, we seek the unknown normals given the incoming normalized rays v_{in} , which are known from camera calibration, and the refracted ray directions v_{out} , which are extracted from the imaged probe color (Fig. 1, top). The absolute angles θ_{in} and θ_{out} are unknown, but we can compute the difference θ_d between the two as $\cos \theta_d = v_{in} \cdot v_{out}$. For known refractive indices of the two media n_1 and n_2 , the angle between incoming ray and surface normal is then given as

$$\theta_{in} = \tan^{-1} \left(\frac{n_2 \sin \theta_d}{n_2 \cos \theta_d - n_1} \right). \quad (1)$$

Therefore, the surface normal \mathbf{n} can be computed independently for each camera pixel by rotating v_{in} by the angle θ_{in} . The rotation is performed on the plane spanned by v_{in} and v_{out} , so

$$\mathbf{n} = \mathbf{R}(\theta_{in}, v_{in} \times v_{out}) (-v_{in}), \quad (2)$$

where $\mathbf{R}(\theta, \mathbf{v})$ is a rotation matrix defined by angle θ around an axis \mathbf{v} .

3.3. Point Cloud Estimation

In order to triangulate absolute 3D surface points for each camera pixel, we need to determine the intersection of the lines $c + tv_{in}$ and $p + sv_{out}$. The camera position c as well as the unrefracted ray directions v_{in} are known from camera calibration and uniquely define a line in 3D space. The direction v_{out} is estimated from the observed colors of the light field probe refracted by an object, however, only a single spatial coordinate is coded by the probe

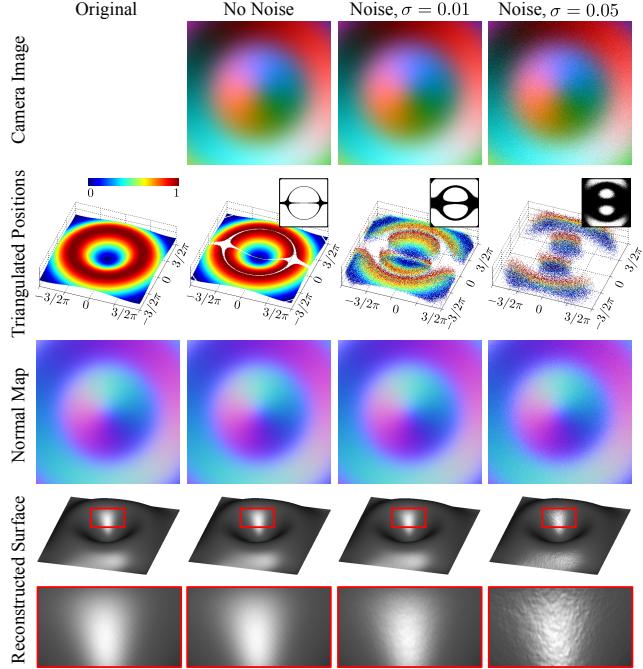


Figure 2. Synthetic results for a refractive sinusoidal object. Normals and positions are shown for the original object (left column), and for reconstructions (other columns) from simulated camera images with an increasing amount of noise (top row).

color, i.e. p^y . Nevertheless, the intersection problem for the two lines results in a linear system with three equations and three unknowns p^x , s , and t because the origin of the coordinate system is defined on the plane of the probe, i.e. $p^z = 0$. Therefore, we can uniquely triangulate a 3D point per camera pixel as

$$t = \frac{1}{v_{in}^y - \frac{v_{in}^z v_{out}^y}{v_{out}^z}} \left(p^y + \frac{c^z v_{out}^y}{v_{out}^z} - c^y \right). \quad (3)$$

The triangulated positions are only numerically robust when significant refraction occurs along a ray; otherwise v_{in} and v_{out} are co-linear. At the same time, all measured ray directions v_{out} will be noisy due to camera noise and possible color non-linearities of a fabricated probe. Therefore, we can only hope to robustly estimate a sparse set of 3D points from such measurements at camera pixels that observe a strong amount of refraction. The noise sensitivity of triangulated points is illustrated for a synthetic example in Figure 2.

3.4. Surface Estimation from Normals and Points

While a normal field can be efficiently integrated to reconstruct surfaces (see e.g., [2]), including an additional set of sparse 3D control points can remove ambiguities in these integration schemes [11, 25]. For all of our reconstructions, we employ the integration method proposed by

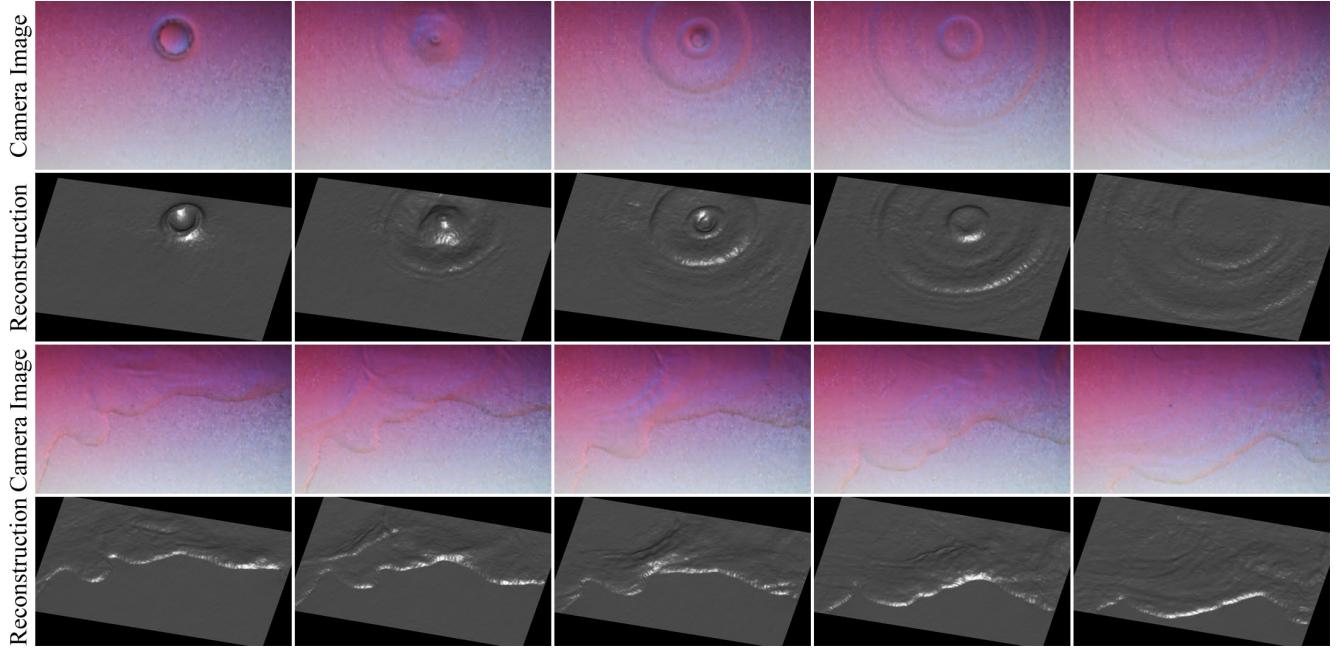


Figure 3. Camera images and reconstructed surfaces of dynamic water surfaces. The upper rows shows a drop falling into the water, whereas the lower rows depict water being poured into the tank.

Wu et al. [26], which uses an optimization with kernel basis functions.

We show synthetic results in Figure 2. Here, a sinusoidal function acts as the original surface with a refractive index corresponding to water; 3D positions and normals of the original surface are shown in the left column. We simulated photographs of an orthogonal camera that show the surface in front of a light field probe with the color coding scheme discussed in Section 3.1 along with estimated normals, triangulated control points, and final reconstructions. While the extracted normals are relatively resilient to an increasing amount of camera noise, the triangulated positions quickly become less reliable. We mask out triangulated points that correspond to small angles between incoming and refracted rays for each pixel; the masks are shown in the insets of the second row.

4. Experimental Results

Our prototype (see Fig. 1, bottom) is composed of a light box, two stacked transparencies, a lenslet array, and a camera. The light box is LED-based, as opposed to fluorescent, in order to maintain consistent lighting throughout the capture process even when using a short exposure time, such as in video. The lenslet sheet is a FresnelTech hexagonal lenslet array with a focal length of 0.12" and a lenslet diameter of 0.09". The transparencies are printed with an Epson Stylus Photo 2200 printer at 1440 dpi, which, in combination with the lenslets, results in a theoretical angular reso-

lution of 0.32° . This printer has six ink-based primaries; for improved contrast we stack two transparencies on top of each other. For still photographs we use a Canon D5 Mark II and for the videos a Prosilica EC1350C camera.

Intrinsic and extrinsic camera parameters are estimated in a pre-processing step [3]. The gamma curves of the printer are also estimated as a pre-processing step and compensated in the measurements.

Reconstructions of water surfaces are shown in Figure 3. Here, we positioned the probe underneath a rectangular water tank and filmed the scene from above (Fig. 3, rows 1 and 3). Secondary refractions from the glass tank bottom are negligible in this case. The results show a water drop falling into the tank in rows 1 and 2; rows 3 and 4 depict water being poured into the tank. Some high-frequency noise is visible in the reconstruction, which is due to the printer half-toning patterns on the transparencies that become visible as noise on the probe when the camera is focused on it. Alternative printing technologies, such as light valve technology (www.bowhaus.com), could alleviate this problem.

Figure 4 shows reconstructions of three thin solid objects from a single photograph each. Although two refractive events occur for each camera ray, one at the air-glass interface toward the camera and another one at the glass-air boundary on the other side, the objects are thin enough that ray displacements within the glass are negligible. This is a common assumption for thin lenses. The reconstructed normals (Fig. 4, column 3) for these examples therefore show

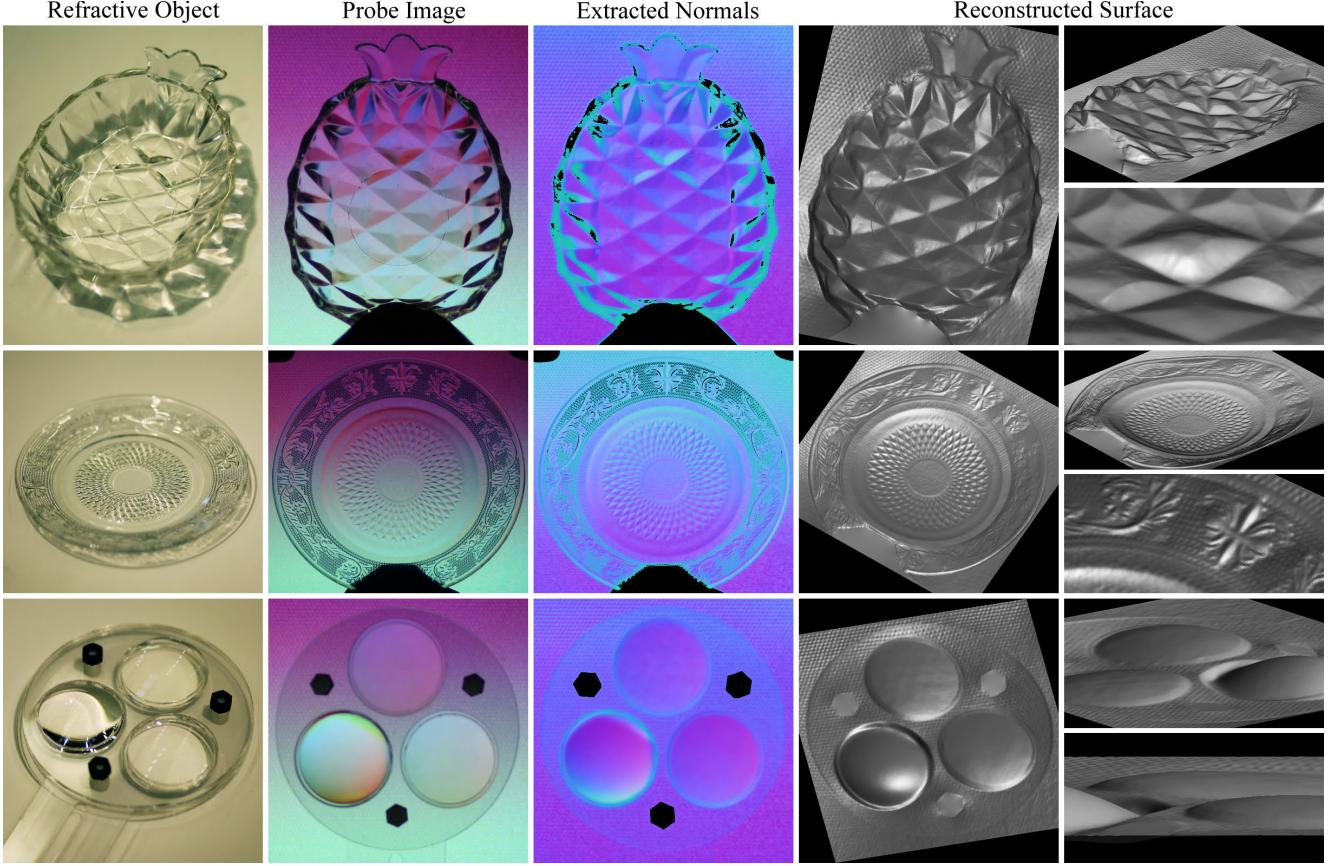


Figure 4. Three thin refractive objects under room illumination (left column) and in front of a light field probe (center left column). The distorted colors of the probe allow us to estimate refractive surface normals from a single image (center row), which can be integrated to reconstruct thin shapes that approximate the geometry of transparent, refractive solids (right).

the difference between front and back normal of the surface; for the plate and the pineapple, the front side is flat and parallel to the fine details on the rear side. The reconstructed surfaces (Fig. 4, right) only contain a flat triangle mesh and corresponding normals.

5. Evaluation

The acquisition of ground truth data for refractive, transparent objects is difficult. We qualitatively evaluate reconstructions of our prototype by comparing a rendering of the three lenses (see Fig. 4) with analytic descriptions of the same lenses in Figure 5. The diameters and focal lengths of these lenses are known and used to simulate them as bi-convex refractive surfaces in front of a textured background with POV-Ray (www.povray.org). The same procedure is used to simulate the reconstructed lens surfaces in front of the background. Slight differences in the lower left lens are mainly due to a violation of the thin lens model.

A quantitative evaluation of the proposed reconstruction algorithm with respect to camera noise and refractive



Figure 5. Reconstructed and synthetic lenses from Figure 4 rendered as a refractive mesh in front of an image.

index mismatches is shown in Figure 6. In this experiment, we simulate the acquisition and reconstruction of a 1D parabolic surface. An orthographic camera observes the scene from above with a light field probe illuminating it from the bottom. The surface represents the boundary between two media, the upper one is air and the lower one has a refractive index of $n = 1.5$. We add zero-mean Gaussian noise to the simulated sensor measurements and evaluate reconstruction quality for different refractive index

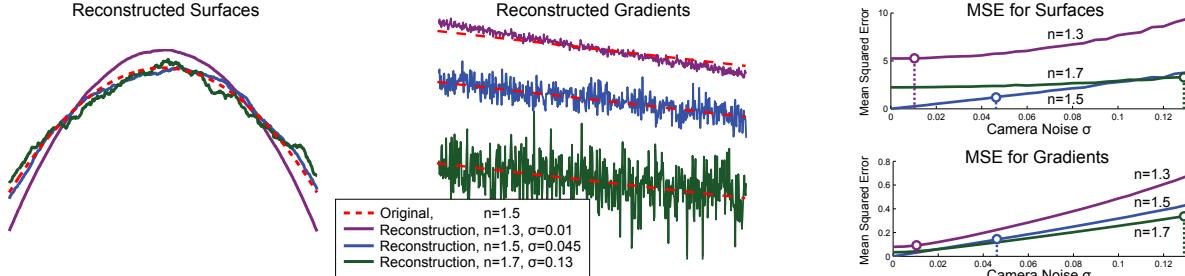


Figure 6. Evaluation of reconstruction with respect to noise and refractive index mismatch. A 1D parabolic surface (left, dotted red) is simulated to be captured with a light field probe and reconstructed with different amounts of camera noise and mismatches in the refractive index of the medium (left). While noise results in high frequency artifacts, a mismatch in the refractive index causes low frequency distortions. We show the mean squared error of surfaces (top right) and gradients (bottom right) for an increasing amount of sensor noise.

mismatches. Surface gradients (Fig. 6, center) are directly computed from the noisy sensor measurements and subsequently integrated to yield the actual surfaces (Fig. 6, left).

Based on these experiments, we can see that a mismatch in the refractive index results in a vertical shear of the gradients (Fig. 6, center, purple line), which corresponds to low frequency distortions of the actual surface (Fig. 6, left, purple line). The mean squared error (MSE) between original surface and reconstruction is particularly high when the assumed refractive index is lower than that of the medium (Fig. 6, top right, purple line). Furthermore, there is an approximately linear relationship between sensor noise and the noise observed in both reconstructed gradients and surfaces (Fig. 6, right). The mean squared error plots on the right of Figure 6 are averaged over 500 experiments, each exhibiting random noise.

6. Discussion and Conclusions

In summary, we have presented a single image approach to thin refractive surface acquisition placing no constraints on the camera setup. Instead of analyzing the distortion of purely diffuse or purely angular reference background patterns, as done in previous work, we encode the angular and spatial dimensions of a light field probe simultaneously. The observed distortion of high-dimensional light fields allows us to reconstruct surface normals and triangulate a sparse set of control points from a single photograph. While the normals are relatively resilient to sensor noise and allow high-quality reconstructions, the triangulated control points are very sensitive to noise, but allow low-frequency ambiguities of the surface normals to be corrected.

6.1. Limitations

Our approach is currently mostly limited by the employed off-the-shelf hardware. Instead of using lenslet arrays and printed transparencies as light field probes, we expect much better results with alternative light field display technologies, such as holograms (www.zebraimaging.com).

Furthermore, the lenslets have a limited field of view and introduce intensity variations over the probe surface, which become visible in the reconstructions; holograms could resolve this problem as well. Color non-linearities and cross-talk introduced by the printing process also affect the accuracy of the reconstructions.

Disregarding the prototype, our approach is fundamentally limited by the light field coding scheme and the reconstruction algorithm. Although the employed color codes are optimized for single image reconstructions, attenuation and scattering within the medium as well as wavelength-dependency of refraction are assumed to be negligible. Alternative, dynamic codes can overcome these limitations at the cost of requiring multiple photographs. The proposed reconstruction algorithm requires the refractive index of the medium to be known and restricts light rays to refract only once in the scene. In combination with advanced coding schemes, novel algorithms could overcome these limitations as well.

6.2. Future Work

In the future, we would like to experiment with alternative technologies for fabricating light field probes, such as holograms, and test more sophisticated light field coding schemes. Applying temporal multiplexing with dynamic probes could lift current limitations; multi-spectral displays and cameras could improve the amount of coded information as well. We would like to explicitly separate attenuation and refraction caused by the medium and test our approach with multi-camera, multi-probe configurations.

References

- [1] S. Agarwal, S. P. Mallick, D. Kriegman, and S. Beilongie. On Refractive Optical Flow. In *Proc. ECCV*, pages 483–494, 2004.
- [2] A. Agrawal, R. Raskar, and R. Chellappa. What is the Range of Surface Reconstructions from a Gradient Field? In *Proc. ECCV*, pages 483–494, 2006.

- [3] B. Atcheson, F. Heide, and W. Heidrich. CALTag: High Precision Fiducial Markers for Camera Calibration. In *Proc. VMV*, 2010.
- [4] B. Atcheson, I. Ihrke, W. Heidrich, A. Tevs, D. Bradley, M. Magnor, and H.-P. Seidel. Time-resolved 3D Capture of Non-stationary Gas Flows. *ACM Trans. Graph. (Siggraph Asia)*, 27(5):132, 2008.
- [5] M. Ben-Ezra and S. K. Nayar. What Does Motion Reveal About Transparency? In *Proc. ICCV*, pages 1025–1032, 2003.
- [6] T. Bonfort, P. Sturm, and P. Gargallo. General Specular Surface Triangulation. In *Proc. ACCV*, pages 872–881, 2006.
- [7] S. Dalziel, G. Hughes, and B. Sutherland. Whole-Field Density Measurements by Synthetic Schlieren. *Experiments in Fluids*, 28(4):322–335, 2000.
- [8] C. Gao and N. Ahuja. A Refractive Camera for Acquiring Stereo and Super-resolution Images. In *Proc. CVPR*, pages 2316–2323, 2006.
- [9] M. Goesele, H. P. A. Lensch, J. Lang, C. Fuchs, and H.-P. Seidel. DISCO: Acquisition of Translucent Objects. *Proc. Siggraph*, 23:835–844, 2004.
- [10] C. Hernandez, G. Vogiatzis, G. Brostow, B. Stenger, and R. Cipolla. Non-Rigid Photometric Stereo with Colored Lights. In *Proc. ICCV*, 2007.
- [11] I. Horovitz and N. Kiryati. Depth from Gradient Fields and Control Points: Bias Correction in Photometric Stereo. In *Proc. ICIVC*, pages 681–694, 2004.
- [12] W. L. Howes. Rainbow Schlieren and its Applications. *Applied Optics*, 23(4):2449–2460, 1984.
- [13] M. B. Hullin, M. Fuchs, I. Ihrke, H.-P. Seidel, and H. P. A. Lensch. Fluorescent Immersion Range Scanning. In *ACM Trans. Graph. (Siggraph)*, pages 87:1–87:10, 2008.
- [14] C. P. Huynh, A. Robles-Kelly, and E. Hancock. Shape and Refractive Index Recovery from Single-View Polarisation Images. In *Proc. CVPR*, 2010.
- [15] I. Ihrke, B. Goldluecke, and M. Magnor. Reconstructing the Geometry of Flowing Water. In *Proc. ICCV*, pages 1055–1060, 2005.
- [16] I. Ihrke, K. N. Kutulakos, H. P. A. Lensch, M. Magnor, and W. Heidrich. Transparent and Specular Object Reconstruction. *CGF*, 29(8):2400–2426, 2010.
- [17] K. N. Kutulakos and E. Steger. A Theory of Refractive and Specular 3D Shape by Light-Path Triangulation. In *Proc. ICCV*, pages 1448–1455, 2005.
- [18] M. Levoy and P. Hanrahan. Light Field Rendering. In *Proc. Siggraph*, pages 31–42, 1996.
- [19] M. Levoy, Z. Zhang, and I. McDowall. Recording and Controlling the 4D Light Field in a Microscope. *Journal of Microscopy*, 235(2):144–162, 2009.
- [20] D. Miyazaki and K. Ikeuchi. Inverse Polarization Raytracing: Estimating Surface Shapes of Transparent Objects. In *Proc. CVPR*, pages 910–917, 2005.
- [21] N. J. W. Morris and K. N. Kutulakos. Dynamic Refraction Stereo. In *Proc. ICCV*, 2005.
- [22] N. J. W. Morris and K. N. Kutulakos. Reconstructing the Surface of Inhomogeneous Transparent Scenes by Scatter Trace Photography. In *Proc. ICCV*, 2007.
- [23] H. Murase. Surface Shape Reconstruction of an Undulating Transparent Object. In *Proc. ICCV*, pages 313–317, 1990.
- [24] D. B. Murphy. *Fundamentals of Light Microscopy and Electronic Imaging*. Wiley-Liss, 2001.
- [25] D. Nehab, S. Rusinkiewicz, J. Davis, and R. Ramamoorthi. Efficiently Combining Positions and Normals for Precise 3D Geometry. *ACM Trans. Graph. (Siggraph)*, 24(3), 2005.
- [26] H.-S. Ng, T.-P. Wu, and C.-K. Tang. Surface-from-Gradients without Discrete Integrability Enforcement: A Gaussian Kernel Approach. *IEEE Trans. PAMI*, 32:2085–2099, 2010.
- [27] T. Okoshi. *Three-Dimensional Imaging Techniques*. Academic Press, 1976.
- [28] S. Savarese and P. Perona. Local Analysis for 3D Reconstruction of Specular Surfaces - Part II. In *Proc. ECCV*, pages 759–774, 2002.
- [29] G. S. Settles. *Schlieren and Shadowgraph Techniques*. Cambridge University Press, 2001.
- [30] M. Shimizu and M. Okutomi. Reflection Stereo - Novel Monocular Stereo using a Transparent Plate. *Proc. CRV*, pages 1–14, 2006.
- [31] M. Tarini, H. P. Lensch, M. Goesele, and H.-P. Seidel. 3D Acquisition of Mirroring Objects using Striped Patterns. *Graphical Models*, 67(4):233–259, 2005.
- [32] B. Trifonov, D. Bradley, and W. Heidrich. Tomographic Reconstruction of Transparent Objects. In *Proc. EGSR*, pages 51–60, 2006.
- [33] G. Wetzstein, R. Raskar, and W. Heidrich. Hand-Held Schlieren Photography with Light Field Probes. In *Proc. ICCP*, pages 1–8, 2011.
- [34] R. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 1980.
- [35] X. Zhang and C. S. Cox. Measuring the Two-dimensional Structure of a Wavy Water Surface Optically: A Surface Gradient Detector. *Exp. in Fluids*, 17:225–237, 1994.