WHAT DOES A SINGLE LIGHT-RAY REVEAL ABOUT A TRANSPARENT OBJECT?

Chia-Yin Tsai[†], *Ashok Veeraraghavan*[‡] *and Aswin C. Sankaranarayanan*[†]

[†] Dept. of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA
[‡] Dept. of Electrical and Computer Engineering, Rice University, Houston, TX

ABSTRACT

We address the following problem in refractive shape estimation: given a single light-ray correspondence, what shape information of the transparent object is revealed along the path of the light-ray, assuming that the light-ray refracts twice. We answer this question in the form of two depth-normal ambiguities. First, specifying the surface normal at which refraction occurs constrains the depth to a unique value. Second, specifying the depth at which refraction occurs constrains the surface normal to lie on a 1D curve. These two depth-normal ambiguities are fundamental to shape estimation of transparent objects and can be used to derive additional properties. For example, we show that correspondences from three light-rays passing through a point are needed to correctly estimate its surface normal. Another contribution of this work is that we can reduce the number of views required to reconstruct an object by enforcing shape models. We demonstrate this property on real data where we reconstruct shape of an object, with light-rays observed from a single view, by enforcing a locally planar shape model.

Index Terms— 3D reconstruction, transparent objects, refraction, light-ray correspondence.

1. INTRODUCTION

A seminal result in shape estimation of transparent objects suggests the use of light-ray correspondences [1]. It is shown that shape of specular and transparent objects, in terms of depths and surface normals, can be estimated from the mapping of input to output light-rays. In particular, even when a set of light-rays refract twice upon intersection with a transparent object, estimating the shape of the object is tractable provided it is observed from at least three distinct viewpoints. We build upon this result.

We characterize the ambiguities in the depth and surface normal along a single light-path, i.e., given a single light-ray correspondence, we derive meaningful constraints on the depths and surface normals of the object along that light-ray (see Fig. 1). Our contributions can be summarized in the following two statements:

This work was supported in part by the NSF grant CCF-1117939.

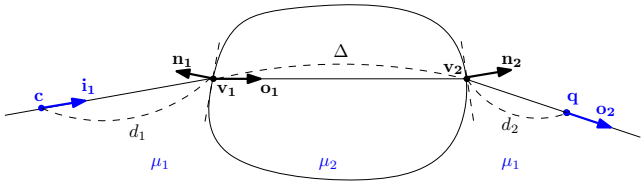


Fig. 1. **Depth-normal ambiguity of a transparent object.** A light-path originates from a point $\mathbf{c}$ with orientation $\mathbf{i}_1$, and, after refracting twice, ends up at $\mathbf{q}$ with orientation $\mathbf{o}_2$. The main results of this paper are two-fold: if $\mathbf{n}_1$ is specified, then d_1 is uniquely determined; and, if d_1 is specified, then $\mathbf{n}_1$ is constrained to lie on a 1D curve. The variables labeled in blue are known and form the inputs to the problem.

- Given the surface normal at which a refraction occurs, the depth is uniquely determined.
- Given knowledge of the depth at which a refraction occurs, the surface normal at that location is constrained to lie on a 1D curve.

These two statements provide a novel depth-normal ambiguity characterization for transparent object and are similar in spirit to the bas-relief ambiguity [2] for Lambertian objects and the depth-normal ambiguity [3] for specular objects. We use these statements to develop shape estimation algorithms that incorporate surface priors.

2. RELATED WORK

We look at some of the key ambiguities underlying shape estimation under various contexts.

Bas-relief ambiguity. The bas-relief ambiguitiy [2] specifies shape ambiguities when a Lambertian object is orthographically viewed under distant lighting. Specifically, there is an equivalence class of shapes that have the same appearance and hence, without additional knowledge (for example of the lighting), it is not possible to distinguish between them.

Depth-normal ambiguity of specular objects. A fundamental ambiguity in specular surface reconstruction is the so-called depth-normal ambiguity [3]. Given a light-ray that is

the reflection of a known 3D reference point on a mirror, the surface normal and the depth at which reflection occurs are constrained. Specifically, for every possible depth along the light-ray, there is a corresponding normal leading to the same reference point. There are multiple approaches to resolve the depth-normal ambiguity for specular surfaces. Methods in [4–6] overcome the depth-normal ambiguity by using correspondences of a known reference target and by regularizing the depth using smoothness priors. Kutulakos and Steger [1] show that observing the specular object from two views provides dense 3D-2D correspondences. In [7], the shape of a mirror is recovered by obtaining images of a planar target, thus the need for knowledge of the reference point is relaxed. Shape from specular flow [8,9] recovers the shape of a mirror from scene motion requiring little effort in terms of calibration. In this setting, it is even possible to obtain invariant signatures of the shape of the object from multiple images [10].

Transparent object reconstruction. For transparent objects, it is common that light refracts at least twice — once each upon entering and exiting the object. This makes shape estimation a hard problem since we have to jointly reason about the shape of the object at multiple locations. Kutulakos and Steger [1] show that if we have light-ray correspondences, it is possible to reconstruct transparent objects even when the rays undergo refraction twice. Some other method on shape reconstruction of even more complex transparent objects includes tomography for reconstructing transparent objects [11, 12] and reconstructing fluid [13], Schlieren imaging for transparent object reconstruction [14] and thin gas flows [15], using radiometric cues to reduce the number of views needed [16], using light polarization to simultaneously recover shape and refractive index of transparent objects [17], and shape estimation using single refraction approximation [18–21].

3. SHAPE FROM RAY CORRESPONDENCES

The goal of this paper is to find what a single light-ray reveals about a transparent object.

Notation. We denote vectors in **bold**. Light-rays in 3D are denoted by a pair of vectors, $\{\mathbf{o}, \mathbf{p}\}$, where $\mathbf{o} \in \mathbb{R}^3$ is a unit-norm vector that denotes the orientation of the ray and $\mathbf{p} \in \mathbb{R}^3$ is a point in 3D space that the ray passes through. For $\lambda \in \mathbb{R}$, $\{\mathbf{o}, \mathbf{p}\}$ and $\{\mathbf{o}, \mathbf{p} + \lambda\mathbf{o}\}$ are identical.

Problem statement. Suppose that a ray $\{\mathbf{i}_1, \mathbf{c}\}$ is incident on a transparent object and after refracting twice, once each upon entering and exiting the object, becomes the ray $\{\mathbf{o}_2, \mathbf{q}\}$ (see Fig. 1). Given knowledge of the rays $\{\mathbf{i}_1, \mathbf{c}\}$ and $\{\mathbf{o}_2, \mathbf{q}\}$, the relative refractive indices of the medium μ_1 and the object μ_2 and their ratio $\rho = \mu_2/\mu_1$, what can we infer about the shape of the object as encoded in the locations of the refraction events, $\mathbf{v}_1$ and $\mathbf{v}_2$, and the surface normals, $\mathbf{n}_1$ and $\mathbf{n}_2$, of the object at these locations.

Observation 1. Since the refraction events occur on the rays, we can identify two depth values d_1 and d_2 such that $\mathbf{v}_1 = \mathbf{c} + d_1\mathbf{i}_1$ and $\mathbf{v}_2 = \mathbf{q} - d_2\mathbf{o}_2$. The problem is equivalent to estimating d_1, d_2, $\mathbf{n}_1$, and $\mathbf{n}_2$.

Observation 2. The light-path is fully determined from the depth and surface normals at $\mathbf{v}_1$ (or equivalently, $\mathbf{v}_2$). Given d_1 and $\mathbf{n}_1$, the outgoing ray $\mathbf{o}_1$ is fully specified from the laws of refraction. The intersection of this ray with $\{\mathbf{o}_2, \mathbf{q}\}$ provides both the surface normal $\mathbf{n}_2$ and the 3D point $\mathbf{v}_2$.

Observation 3. Any constraint that we derive on the shape at $\mathbf{v}_1$, the first refraction point, translates to a similar constraint on the shape at $\mathbf{v}_2$. This is simply a consequence of Helmholtz reciprocity.

Our main results are in the form of ambiguities in the values of the depth d_1 given knowledge of the surface normal $\mathbf{n}_1$, and vice versa. These can be succinctly summarized in the following statements.

Theorem 1 (Depth ambiguity given normal)**.** *Given the surface normal $\mathbf{n}_1$, the depth d_1 at which the refraction occurs is unique, provided the light-path does not entirely lie in a plane.*

Proof. By Snell's law, the refracted ray $\mathbf{o}_1$ is uniquely determined given both the incident ray $\mathbf{i}_1$ and the normal $\mathbf{n}_1$ by

$$\mathbf{o}_1 = \frac{\mathbf{i}_1}{\rho} - \left(\langle \mathbf{i}_1, \mathbf{n}_1 \rangle + \sqrt{\left(\langle \mathbf{i}_1, \mathbf{n}_1 \rangle^2 - (1 - \rho^2) \right)} \right) \frac{\mathbf{n}_1}{\rho}. \quad (1)$$

Since a light-path is connected, the relationship between d_1 and d_2 can be characterized as

$$\mathbf{c} + d_1\mathbf{i}_1 + \Delta\mathbf{o}_1 + d_2\mathbf{o}_2 = \mathbf{q}.$$

This can be further simplified to

$$d_1 \langle \mathbf{i}_1 \times \mathbf{o}_2, \mathbf{o}_1 \rangle = \langle (\mathbf{q} - \mathbf{c}) \times \mathbf{o}_2, \mathbf{o}_1 \rangle .$$

When $\langle \mathbf{i}_1 \times \mathbf{o}_2, \mathbf{o}_1 \rangle \neq 0$,

$$d_1 = \frac{\langle (\mathbf{q} - \mathbf{c}) \times \mathbf{o}_2, \mathbf{o}_1 \rangle}{\langle \mathbf{i}_1 \times \mathbf{o}_2, \mathbf{o}_1 \rangle}. \quad (2)$$

In (2), $\mathbf{i}_1$, $\mathbf{o}_2$, $\mathbf{c}$ and $\mathbf{q}$ are known. Further, the value of $\mathbf{o}_1$ is fully-determined from (1) when $\mathbf{n}_1$ is given. Therefore, provided $\langle \mathbf{i}_1 \times \mathbf{o}_2, \mathbf{o}_1 \rangle \neq 0$, d_1 is uniquely determined when the surface normal $\mathbf{n}_1$ is known. Finally, $\langle \mathbf{i}_1 \times \mathbf{o}_2, \mathbf{o}_1 \rangle = 0$ if and only if $\mathbf{o}_1$ lies in the plane spanned by $\mathbf{i}_1$ and $\mathbf{o}_2$ or, equivalently, the entire light-path lies in a plane. □

We provide a geometric interpretation of Theorem 1 in Fig. 2. Given $\mathbf{i}_1$ and $\mathbf{n}_1$, $\mathbf{o}_1$ is determined. All the possible rays after first refraction will form a plane. The light-ray $\{\mathbf{o}_2, \mathbf{q}\}$ will intersect the plane at one point. Therefore, we can find a unique depth d_1.

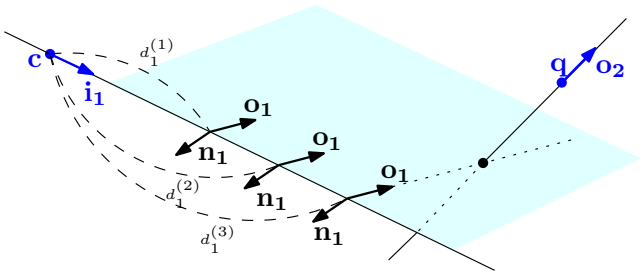


Fig. 2. **Illustration of Theorem 1.** Given normal $\mathbf{n}_1$, $\mathbf{o}_1$ is fully determined. By finding the intersection of the light-ray $\{\mathbf{o}_2, \mathbf{q}\}$ and the plane formed by all possible $\mathbf{v}_1$ and $\mathbf{o}_1$, we can find the corresponding depth $d_1^{(3)}$.

Corollary 1 (Planar light-path)**.** *If the vectors* $\mathbf{i}_1$, $\mathbf{o}_2$ *and* $(\mathbf{q}-\mathbf{c})$ *are co-planar, then specifying* $\mathbf{n}_1$ *does not constrain* d_1.

Theorem 2 (Normal ambiguity given depth)**.** *Given the depth* d_1*, the surface normal* $\mathbf{n}_1$ *is constrained to lie on a 1D curve, which is the intersection of an oblique cone with a unit sphere.*

Proof. Recall that the first refraction happens at $\mathbf{v}_1$, which is d_1 units away from the camera center,

$$\mathbf{v}_1 = \mathbf{c} + d_1\mathbf{i}_1.$$

The second refraction occurs at $\mathbf{v}_2$ which lies on the line defined by a point $\mathbf{q}$ and the vector $\mathbf{o}_2$,

$$\mathbf{v}_2 \in \{\mathbf{q} - \lambda\mathbf{o}_2, \quad \lambda \in \mathbb{R}, \lambda \geq 0\}.$$

From Fig. 3, we observe that $\mathbf{o}_1$ belongs to the plane that $\mathbf{q} - \mathbf{v}_1$ and $\mathbf{o}_2$ are on. Therefore, we can represent $\mathbf{o}_1$ by:

$$\mathbf{o}_1 = \mathcal{B}(d_1)\begin{bmatrix}\cos(\psi)\\ \sin(\psi)\end{bmatrix}, \quad \psi \in [0, 2\pi),$$

where $\mathcal{B}(d_1)$ is an orthonormal basis for the column span of $\mathbf{q} - \mathbf{v}_1$ and $\mathbf{o}_2$.[1] Since $\mathbf{c}$, $\mathbf{i}_1$, $\mathbf{o}_2$ and $\mathbf{q}$ are known and $\mathbf{v}_1 = \mathbf{c} + d_1\mathbf{i}_1$, $\mathcal{B}$ is dependent only on d_1.

Define $\mathbf{n}_\perp$ as a unit-norm vector, co-planar to $\mathbf{i}_1$ and $\mathbf{n}_1$, and orthogonal to $\mathbf{n}_1$. From Snell's law, we know $\mu_1 \sin\theta_1 = \mu_2 \sin\theta_2$, where θ_1 and θ_2 are the angle formed by $\mathbf{i}_1$ and $\mathbf{o}_1$, respectively, to $\mathbf{n}_1$.

$$\begin{aligned}\mu_1 \langle \mathbf{i}_1, \mathbf{n}_\perp\rangle &= \mu_2 \langle \mathbf{o}_1, \mathbf{n}_\perp\rangle \\ \mu_1 \left\langle \mathbf{i_1} - \frac{\mu_2}{\mu_1}\mathbf{o}_1, \mathbf{n}_\perp \right\rangle &= 0 \\ \langle \mathbf{i}_1 - \rho\mathbf{o}_1, \mathbf{n}_\perp\rangle &= 0\end{aligned}$$

[1] Another way to parameterize is $\mathbf{o}_1 = \frac{\mathbf{q}-\lambda\mathbf{o}_2-\mathbf{v}_1}{\|\mathbf{q}-\lambda\mathbf{o}_2-\mathbf{v}_1\|}$, where $\lambda > 0$. While this constraints the surface normal to a smaller set, the resulting expressions are harder to analysis due to their complex dependence on λ.

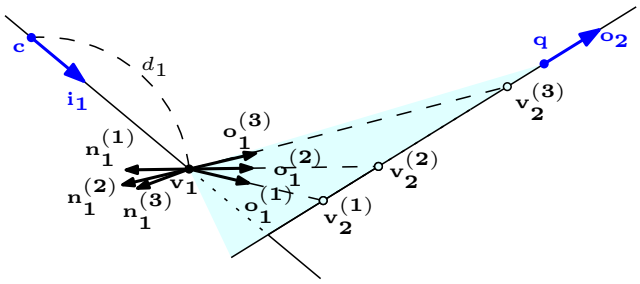


Fig. 3. **Illustration of Theorem 2.** The refracted ray $\mathbf{o}_1$ lies in the plane spanned by $\mathbf{q} - \mathbf{v}_1$ and $\mathbf{o}_2$. By choosing different $\mathbf{v}_2$ locations, we will have different corresponding $\mathbf{o}_1$ and $\mathbf{n_1}$. In this figure, we show the normal $\mathbf{n}_1^{(i)}$ and $\mathbf{o}_1^{(i)}$ corresponding to different $\mathbf{v}_2^{(i)}$ locations.

Recall that $\mathbf{n}_1$ is co-planar to $\mathbf{i}_1$ and $\mathbf{o}_1$ and, by definition, $\langle \mathbf{n}_1, \mathbf{n}_\perp\rangle = 0$. Hence, we can conclude that $\mathbf{n}_1$ is parallel to $\mathbf{i}_1 - \rho\mathbf{o}_1$.

$$\mathbf{n}_1 \propto \mathbf{i}_1 - \rho\mathcal{B}(d_1)\begin{bmatrix}\cos(\psi)\\ \sin(\psi)\end{bmatrix}, \quad \psi \in [0, 2\pi). \tag{3}$$

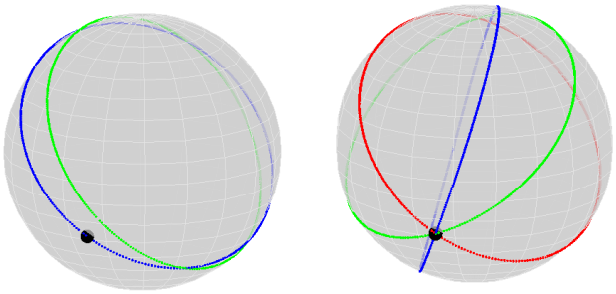
The RHS of (3) traces a circle in $\mathbb{R}^3$ as ψ varies. Since $\|\mathbf{n}_1\| = 1$, we can recover $\mathbf{n}_1$ by computing the intersection of an oblique cone and a unit sphere. Therefore, normal $\mathbf{n}_1$ lies on a closed 1D curve on a 2D unit sphere. □

Corollary 2 (Proposition 1a of [1])**.** *To uniquely identify the surface normal at a point, we need at least three light-ray correspondences.*

Proof. Given the point under consideration, we know the value of d_1 for each ray correspondence. From Theorem 2, each ray-correspondence restricts the surface normal to lie on a closed 1D curve. However, any two arbitrary 1D curves on the unit-sphere can potentially intersect. Hence, we need a third correspondence to verify the intersection produced by the first two correspondences. (See Fig. 4) □

Remark. The depth-normal ambiguities described in Theorems 1 and 2 are fundamental to studying the shape of transparent objects, where we can expect a majority of ray-correspondences to be from double-refraction events. These are similar in spirit to two well-known ambiguities in computer vision: the depth-normal ambiguity for mirror object, and the bas-relief ambiguity in Lambertian shape-from-shading. An understanding of these fundamental ambiguities is important to the design of techniques for shape estimation.

Relationship to [1]. In Kutulakos and Steger [1], the algorithmic development as well as analysis is performed using d_1 and d_2 — variables pertaining to two distinct points, $\mathbf{v}_1$ and $\mathbf{v}_2$. In contrast, we only use d_1 and $\mathbf{n}_1$, which are local to $\mathbf{v}_1$. This leads to a simpler explanation of the underlying ambiguities that we state in Theorems 1 and 2.



(a) (b)

Fig. 4. **Illustration of Corollary 2.** (a) Any two 1D curves on the sphere can intersect. (b) We need a third curve to validate if the intersect is indeed correct.

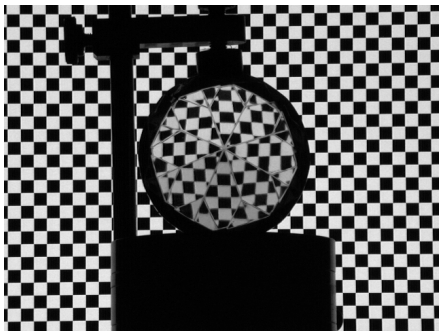


Fig. 5. **Experiment setup.** The dataset is courtesy of Kutulakos and Steger [1].

4. EXPERIMENTS

Single-view reconstruction algorithm using surface prior. If the first refraction happens on a plane, $\mathbf{n}_1$ is the same for all the light-path correspondences. Thus, from Theorem 1, for every candidate $\mathbf{n}_1$, we can calculate d_1 of m light-ray correspondences. Since we impose a planar model on the collection of $\mathbf{v}_1$, we use the variance of

$$[\langle \mathbf{n}_1, \mathbf{v}_1^{(1)} \rangle, \langle \mathbf{n}_1, \mathbf{v}_1^{(2)} \rangle, ..., \langle \mathbf{n}_1, \mathbf{v}_1^{(m)} \rangle]$$

to determine the goodness the candidate $\mathbf{n}_1$. The normal that gives smallest variance is the normal estimation. Once the normal is recovered, both $\mathbf{v}_1$ and $\mathbf{v}_2$ are determined.

Real data. We use real world datasets from [1] to verify our single-view reconstruction algorithms. The data is collected by placing the object of interest between a camera and a movable LCD, as shown in Fig. 5. Image pixel to LCD pixel correspondences are collected at two LCD positions to give light-ray correspondences.

We observe that the back of the diamond scene is planar, therefore, we impose a planar model on the back of the diamond. The refractive index is set to 1.55. As shown in Fig. 6, by enforcing planar model on one side, we can recover a complex object. Without enforcing additional constraint, different facets of the diamond meet at the same position in space. To evaluate the correctness of the normal estimation, we use different camera views. The normal estimation should be consistent in all views. The results are shown in Table 1.

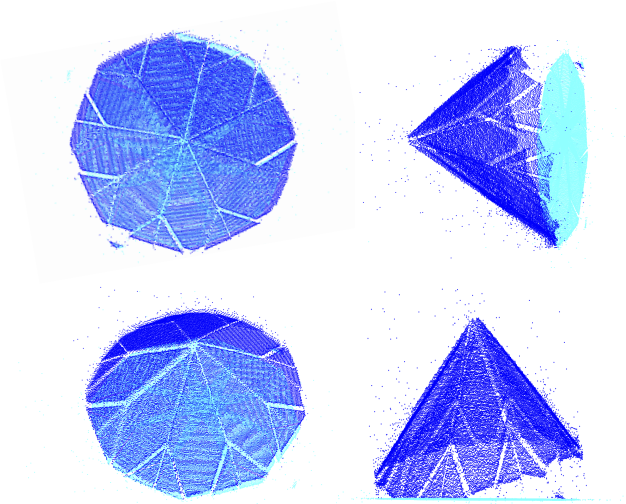


Fig. 6. **Single-view reconstruction with a planarity constraint.** We exploited the planarity of the "back" of the diamond in Figure 5 to recover its shape from light-rays observed at a single camera. Specifically, we randomly chose 40 light-rays (out of a total of 33,737) entering the camera to estimate the normal and intercept of the plane. Given the parameters of the plane, we can now estimate densely both the front surface (show in blue) and the back (shown in cyan).

View number	1	2	3	4	5	6	7
Elevation (°)	84.0	84.0	82.6	79.4	83.5	84.2	83.5
Azimuth (°)	82.7	82.9	92.5	68.9	63.2	69.8	61.4
Standard deviation (°)	3.9	0.7	1.9	1.9	3.8	0.9	6.3

Table 1. **Normal estimation using different camera positions.** We randomly select 50 light-ray correspondences to estimate each normal. Shown are average normal estimates, in terms of azimuth and elevation angles, as well as standard deviation across 50 trials.

5. CONCLUSION

We characterize the information pertaining to the shape of a transparent object that is encoded in a single light-ray. Our contributions are in the form of two novel depth-normal ambiguities which characterizes the fundamental limitations of shape estimation and provides a foundation for future techniques. To this end, we outline how simple surface priors, such as a planar model, can be incorporated into a shape estimation framework for transparent objects.

6. REFERENCES

[1] K. N. Kutulakos and E. Steger, "A theory of refractive and specular 3d shape by light-path triangulation," *Intl. J. of Computer Vision*, vol. 76, no. 1, pp. 13–29, 2008.

[2] P. N. Belhumeur, D. J. Kriegman, and A. L. Yuille, "The bas-relief ambiguity," *Intl. J. of Computer Vision*, vol. 35, no. 1, pp. 33–44, 1999.

[3] J. Kaminski, S. Lowitzsch, M. C. Knauer, and G. Husler, "Full-field shape measurement of specular surfaces," in *Intl. Workshop on Automatic Processing of Fringe Patterns*, 2005.

[4] M. Tarini, H. PA Lensch, M. Goesele, and H.-P. Seidel, "3d acquisition of mirroring objects using striped patterns," *Graphical Models*, vol. 67, no. 4, pp. 233–259, 2005.

[5] M. Liu, R. Hartley, and M. Salzmann, "Mirror surface reconstruction from a single image," in *CVPR*, 2013.

[6] M. Weinmann, A. Osep, R. Ruiters, and R. Klein, "Multi-view normal field integration for 3d reconstruction of mirroring objects," in *ICCV*, 2013.

[7] T. Bonfort, P. Sturm, and P. Gargallo, "General specular surface triangulation," in *ACCV*, 2006.

[8] Y. Adato, Y. Vasilyev, O. Ben-Shahar, and T. Zickler, "Toward a theory of shape from specular flow," in *ICCV*, 2007.

[9] A. C. Sankaranarayanan, A. Veeraraghavan, O. Tuzel, and A. Agrawal, "Specular surface reconstruction from sparse reflection correspondences," in *CVPR*, 2010.

[10] A. C. Sankaranarayanan, A. Veeraraghavan, O. Tuzel, and A. Agrawal, "Image invariants for smooth reflective surfaces," in *ECCV*, 2010.

[11] C. Ma, X. Lin, J. Suo, Q. Dai, and G. Wetzstein, "Transparent object reconstruction via coded transport of intensity," in *CVPR*, 2014.

[12] Y. Ji, J. Ye, and J. Yu, "Reconstructing gas flows using light-path approximation," in *CVPR*, 2013.

[13] J. Gregson, M. Krimerman, M. B Hullin, and W. Heidrich, "Stochastic tomography and its applications in 3d imaging of mixing fluids.," *ACM Trans. Graph.*, vol. 31, no. 4, pp. 52, 2012.

[14] G. Wetzstein, R. Raskar, and W. Heidrich, "Hand-held schlieren photography with light field probes," in *ICCP*, 2011.

[15] G. S. Settles, *Schlieren and shadowgraph techniques*, Springer, 2001.

[16] V. Chari and P. Sturm, "A theory of refractive photo-light-path triangulation," in *CVPR*, 2013.

[17] C. P. Huynh, A. Robles-Kelly, and E. Hancock, "Shape and refractive index recovery from single-view polarisation images," in *CVPR*, 2010.

[18] G. Wetzstein, D. Roodnick, W. Heidrich, and R. Raskar, "Refractive shape from light field distortion," in *ICCV*, 2011.

[19] D. Liu, X. Chen, and Y. H. Yang, "Frequency-based 3d reconstruction of transparent and specular objects," in *CVPR*, 2014.

[20] T. Xue, M. Rubinstein, N. Wadhwa, A. Levin, F. Durand, and W. T. Freeman, "Refraction wiggles for measuring fluid depth and velocity from video," in *ECCV*, 2014.

[21] Y. Ding, F. Li, Y. Ji, and J. Yu, "Dynamic fluid surface acquisition using a camera array," in *ICCV*, 2011.