Dynamic Refraction Stereo

Nigel J.W. Morris and Kiriakos N. Kutulakos, *Member*, *IEEE*

Abstract—In this paper we consider the problem of reconstructing the 3D position and surface normal of points on an unknown, arbitrarily-shaped refractive surface. We show that two viewpoints are sufficient to solve this problem in the general case, even if the refractive index is unknown. The key requirements are 1) knowledge of a function that maps each point on the two image planes to a known 3D point that refracts to it, and 2) light is refracted only once. We apply this result to the problem of reconstructing the time-varying surface of a liquid from patterns placed below it. To do this, we introduce a novel "stereo matching" criterion called *refractive disparity*, appropriate for refractive scenes, and develop an optimization-based algorithm for individually reconstructing the position and normal of each point projecting to a pixel in the input views. Results on reconstructing a variety of complex, deforming liquid surfaces suggest that our technique can yield detailed reconstructions that capture the dynamic behavior of free-flowing liquids.

Index Terms—Stereo, time-varying imagery, shape-from-X, transparency, refractive index estimation.

✦

---

1 INTRODUCTION

MODELING the time-varying surface of a liquid has attracted the attention of many research fields, from computer graphics [10], [14], [36], [39] and fluid mechanics [20] to oceanography [11], [28], [32], [58]. While great strides have been achieved in the development of computer simulators that are physically accurate and visually correct [14], [36], capturing the time-varying behavior of a real liquid remains a challenging problem.

From the point of view of computer vision, analyzing the behavior of liquids from videos poses several difficulties compared to traditional 3D photography applications:

- *No prior scene model.* Spatiotemporal evolution is constrained only by the laws of fluid mechanics, making it difficult to assume a low-degree-of-freedom parametric model for such a scene [4], [57].
- *Nonlinear light path.* Liquid surfaces bend the incident light and, hence, a point below the surface will project along a nonlinear path to a viewpoint above it.
- *Shape-dependent appearance modulation.* Absorption, scattering, and Fresnel transmission cause the appearance of points below the surface to depend on the light's path and, hence, on the surface shape [18].
- *Turbulent behavior.* Liquid flow is an inherently volumetric phenomenon whose complete characterization requires capturing both its time-varying surface and a vector field describing internal motion [59].
- *Instantaneous 3D capture.* Since liquids are dynamic and can flow rapidly, shape recovery must rely on instantaneously-captured information.

- *N.J.W. Morris is with Morgan Solar Inc., Canada.*
  E-mail: nmorris@dgp.toronto.edu.
- *K.N. Kutulakos is with the University of Toronto, 10 King's College Road, Rm SF3302, Toronto, ON M5S 3G4, Canada.*
  E-mail: kyros@cs.toronto.edu.

As a first step, in this paper we consider the problem of reconstructing the time-varying 3D surface of an unknown liquid by exploiting its refractive properties. To do this, we place a known, textured pattern below the liquid's surface and capture image sequences of the pattern from two known viewpoints above the liquid (Fig. 1). Our focus is on imposing as few restrictions as possible on the scene—we assume that the liquid has a constant but unknown index of refraction and that its instantaneous 3D shape is arbitrary, as long as light coming from the pattern is refracted at most once before reaching the input viewpoints.

The reconstruction of refractive surfaces from photographs has a long history in photogrammetry [15], [24], [37], [42]. These techniques assume a low-parameter model for the surface (e.g., a plane) and solve a generalized structure from motion problem in which camera parameters, surface parameters, and 3D coordinates of feature points below the surface are estimated simultaneously. In related work, Treibitz et al. [51] show that a vision system observing a scene immersed under a planar refractive surface becomes multiperspective. They present a calibration technique to estimate the geometry involved and recover the 3D position of objects immersed in the refractive medium. In computer vision, the reconstruction of time-varying refractive surfaces was first studied by Murase [40]. Their work focused on water (whose refractive index is known) and followed a "shape-from-distortion" approach [2], [29], [40], [42], [56]. In their approach, 3D shape is recovered by analyzing one distorted image of a known pattern that is placed underwater. Unfortunately, it is impossible in general to reconstruct the 3D shape of a general refractive surface from one image, even if its refractive index is known. The inherently ill-posed nature of the problem has prompted a variety of assumptions, including statistical assumptions about the pattern's appearance over time [40], known average water height [29], [42], known points on the surface or surface integrability [50], and special optics [32], [58]. These assumptions break down when the refractive index is unknown or when the liquid undergoes significant deformations that cause changes in shape and height (e.g., pouring water in an empty tank). A different way to
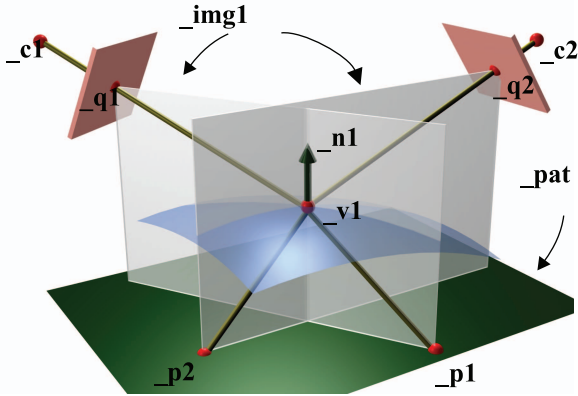
Fig. 1. Geometry of refraction stereo. For each pixel $\mathbf{q}$ in an input view, the goal is to reconstruct the 3D position and surface normal of point $\mathbf{p}$ on the refractive surface.

approach refractive distortion is to break up the observed surface into a triangulated mesh, where each triangle acts as a general linear camera (GLC) that warps the background [12], [13]. Solving for the parameters of these GLCs gives the Gaussian and mean curvature, but the 3D shape remains ambiguous. Another common technique for reconstructing water surfaces is known as the "optical wave-slope measurement" approach [29], [30], [42], [56]. This approach involves immersing optical components and a color or intensity gradient in the liquid so that distinct surface slopes map to distinct colors or intensities. This method is most appropriate for measuring small capillary or wind-driven waves and is unsuitable for the more general cases we consider. More recent work has focused on tomography-based approaches. Trifonov et al. [52] immerse target transparent objects in a transparent liquid with matching refractive index and measure the attenuation of a backlight from various views. In [25], transparent objects are immersed in a solution that fluoresces under laser illumination and are scanned with a sheet of laser light.

A closely related problem is the reconstruction of highly specular surfaces, such as mirrors [5], [19], [26], [43], [46], [47], [48], [50]. Mirrors interact with light in much the same way that refractive surfaces do—light incident at a point is reflected according to the point's surface normal, thereby tracing a nonlinear path. Blake [5] proposed using a moving observer to recover the differential properties of a smooth mirror surface from the observed motion of specularities. Sanderson et al. [45] were the first to analyze the ambiguities in single-view mirror reconstruction and to propose a stereo camera configuration for resolving them. Our work, which is based on a novel analysis of two-view ambiguities for refractive scenes, exploits some of the same basic insights. Recently, there have been several approaches for recovering the 3D shape of mirror surfaces [1], [6], [7], [31], [54], [55], as well as near-specular and glossy surfaces [9], [16].

Reconstructing transparent liquid surfaces is even more challenging than mirrors for three reasons. First, the interaction between light and a mirror does not depend on the mirror's material properties, but it does depend on a liquid's refractive index. When this index is unknown, it must be estimated along with 3D shape. Second, the

nonlinearity of light paths cannot be taken for granted in the case of fluctuating liquid surfaces, whose distance from a pattern below the surface may approach zero, diminishing the effect of refraction. To guarantee stable shape solutions, a reconstruction algorithm must be immune to such degeneracies. Third, establishing accurate pixel-wise correspondences between patterns and their distorted images is much easier in the case of a mirror. In liquids, the distortions are both geometric and radiometric (due to absorption, Fresnel effect, etc.) and can vary significantly from one instant to the next.

The starting point for our work is a novel geometrical result showing that two viewpoints are sufficient to compute both the shape and the refractive index of an unknown, generic refractive surface. The only requirements are 1) knowledge of a function that maps each point on the image plane to a known 3D point that refracts to it, and 2) light is refracted only once. Compared to mirrors, this is a stronger two-view result because it shows that the refractive index ambiguity, not present in mirror scenes, can be resolved without additional views.

On the practical side, our interest is in algorithms that can capture the detailed dynamic behavior of free-flowing liquids. To this end, our work has four contributions. First, we formulate a novel optimization criterion, called *refractive disparity*, appropriate for refractive scenes, that is designed to remain stable when refraction diminishes. Second, we develop an algorithm for individually reconstructing the position and normal of each point projecting to the input views. The algorithm is closer to traditional triangulation [23] and bundle adjustment [22], [53] than to voxel-based stereo [6] and imposes no constraints on the liquid's shape or its evolution. Third, we show that refraction stereo can produce a detailed, full-resolution depth map and a separate, full-resolution normal map for the unknown surface. Fourth, we present experimental results for a variety of complex deforming liquid surfaces. These results suggest that refraction stereo can yield detailed reconstructions that capture the complexity and dynamic behavior of liquids.

2 REFRACTION STEREO GEOMETRY

Consider an unknown, smooth, transparent surface that is viewed by two calibrated cameras under perspective projection (Fig. 1). We assume that the surface bounds a homogeneous transparent medium (e.g., water or alcohol) with an unknown refractive index. Our goal is to compute the refractive index of the medium and the 3D coordinates and surface normal at each point on the unknown surface. To do this, we place a known reference pattern below the surface and compute a pixel-to-pattern correspondence function, $\mathbf{C}(\mathbf{q}, t)$, that gives the 3D coordinates of the point on the pattern refracting to pixel $\mathbf{q}$ at time t. In the following, we assume that this function is known and concentrate on the instantaneous reconstruction problem at time t. We consider the problem of estimating the correspondence function in Section 5. To simplify notation, we omit the time parameter in the following discussion.

Let $\mathbf{q}$ be a pixel in the input views and let $\mathbf{C}(\mathbf{q})$ be the point refracting to $\mathbf{q}$. Suppose that this refraction occurs at distance d from the image plane, at a point $\mathbf{p}(d)$ on the ray
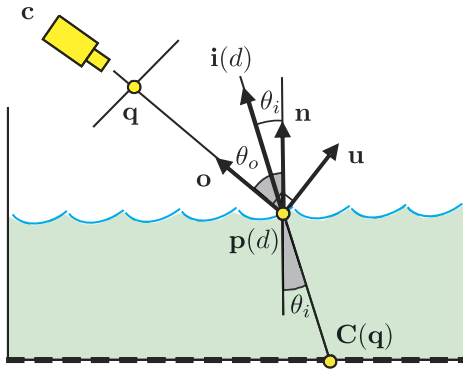
Fig. 2. Geometry of refraction. The figure shows a face-on view of the refraction plane defined by the incoming and outgoing light directions, $\mathbf{i}(d)$ and $\mathbf{o}$, respectively, from point $\mathbf{p}(d)$.

through pixel $\mathbf{q}$ (Fig. 1). The relation between pixel $\mathbf{q}$ and points $\mathbf{C}(\mathbf{q})$ and $\mathbf{p}(d)$ is governed by Snell's law, which describes how light is redirected at the boundary between two different media [18]. Snell's law can be expressed as two independent geometric constraints (Fig. 2):

- a *deflection constraint*, establishing a sinusoidal relation between incoming and outgoing light directions

$$\sin \theta_o = r \sin \theta_i, \qquad (1)$$

  where θ_i is the angle between the surface normal and the ray through $\mathbf{C}(\mathbf{q})$ and $\mathbf{p}(d)$; θ_o is the angle between the surface normal and the ray through pixel $\mathbf{q}$; and r is the refractive index;
- and a *planarity constraint*, forcing the surface normal at $\mathbf{p}(d)$ to lie on the plane defined by point $\mathbf{C}(\mathbf{q})$ and the ray through $\mathbf{q}$; we call this plane the *refraction plane* of pixel $\mathbf{q}$.

These two constraints give us a relation between the pixel, a known 3D point that refracts to it, and the unknown surface. Unfortunately, they are not sufficient to determine how far from the image plane the refraction occurs even when we do know the refractive index. This is because, for every hypothetical distance, there is a 1D set of possible normals that satisfy the planarity and deflection constraints. Each of these normals lies on the pixel's refraction plane and satisfies (1) for some value of the refractive index (Fig. 3). Hence, the unit surface normal that satisfies Snell's law for pixel $\mathbf{q}$ can be expressed as a two-parameter family, $\mathbf{n}(d, r)$, parameterized by the distance d and the unknown refractive index, r. A closed-form expression for this normal is[1]

$$\mathbf{n}(d, r) = \frac{\tilde{\mathbf{n}}(d, r)}{\|\tilde{\mathbf{n}}(d, r)\|}, \qquad (2)$$

$$\tilde{\mathbf{n}}(d, r) = r \, \| \, \mathbf{i}(d) \wedge \mathbf{o}\| \, \mathbf{u} + (r[\, \mathbf{i}(d) \cdot \mathbf{o}] - 1) \, \mathbf{o}, \qquad (3)$$

where $\wedge$ denotes vector product, $\mathbf{i}(d)$ is the unit vector in the direction of the incoming ray to $\mathbf{p}(d)$, $\mathbf{o}$ is the unit vector in the direction of the outgoing ray from surface point $\mathbf{p}(d)$ to

pixel $\mathbf{q}$, and $\mathbf{u}$ is the vector perpendicular to $\mathbf{o}$ on the refraction plane

$$\mathbf{i}(d) = \frac{\mathbf{c} - d \, \mathbf{o} - \mathbf{C}(\mathbf{q})}{\|\mathbf{c} - d \, \mathbf{o} - \mathbf{C}(\mathbf{q}))\|}, \quad \mathbf{o} = \frac{\mathbf{c} - \mathbf{q}}{\|\mathbf{c} - \mathbf{q}\|},$$
$$\mathbf{u} = \frac{\mathbf{o} \wedge (\mathbf{i}(d) \wedge \mathbf{o})}{\|\mathbf{o} \wedge (\mathbf{i}(d) \wedge \mathbf{o})\|}. \qquad (4)$$

When the refractive index has a known value r_0, there is only one consistent normal, $\mathbf{n}(d, r_0)$, for each distance d. Sanderson et al. [45] were the first to point out that this distance-normal ambiguity for a pixel $\mathbf{q}$ can be resolved with the help of a second viewpoint.[2] Intuitively, a second viewpoint allows us to "verify" whether or not a particular distance hypothesis d is correct (Fig. 1): Given such a hypothesis and given the projection $\mathbf{q}'$ of point $\mathbf{p}(d)$ in the second camera, we simply need to verify that point $\mathbf{C}(\mathbf{q}')$ on the reference pattern refracts to pixel $\mathbf{q}'$.

While this hypothesis-verification procedure leads directly to an algorithm when the surface has a known refractive index, it leaves open the question of how to reconstruct surfaces whose 3D shape *and* refractive index are unknown. In this case, the surface normal lies in the full, two-parameter family, $\mathcal{N} = \{\mathbf{n}(d, r) \mid d, r \in \mathbb{R}^+\}$. One approach would be to use a third viewpoint to verify that a hypothetical refractive index r and distance d are consistent with the pixel-pattern correspondences in the three views.

Rather than use a third viewpoint, we prove that two views are, in fact, sufficient to estimate the 3D shape and refractive index of an unknown, generic surface. Intuitively, generic surfaces embody the notion of *nondegeneracy*—they are smooth surfaces whose differential properties remain unchanged if we deform their surface by an infinitesimal amount [33]. As such, they are especially suitable for modeling the complex, unconstrained shape of a liquid. Theorem 1 tells us if the liquid's surface is generic, the family, $\mathcal{N}$, of ambiguous solutions is discrete.

Theorem 1. $\mathcal{N}$ *is a zero-dimensional manifold for almost all pixels in the projection of a generic surface.*

Theorem 1 holds for continuous $\mathbf{C}(\mathbf{q})$ and suggests that it might be possible to compute the refractive index of a surface by choosing a single pixel $\mathbf{q}$ and finding the distance and refractive index that are consistent with $\mathbf{C}(\mathbf{q})$ and the pixel-to-pattern correspondences in the second viewpoint. In practice, image noise and the possibility of multiple discrete solutions dictate an alternative strategy where measurements from multiple pixels contribute to the estimation of the refractive index. We consider the algorithmic implications of this result below.

3 DYNAMIC REFRACTION STEREO ALGORITHM

In order to reconstruct a liquid's surface at a time instant t, we need to answer three basic questions: 1) How do we compute the pixel-to-pattern correspondence function $\mathbf{C}(\mathbf{q}, t)$, 2) how do we compute the refractive index, and 3) how do we assign a distance and a normal to each pixel?

---

1. See the Appendix, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TPAMI.2011.24, for a derivation.

2. Sanderson et al. [45] made this observation in the context of reconstructing opaque specular, rather than refractive, surfaces. Their analysis applies equally well to the case of refractive surfaces with a known refractive index.
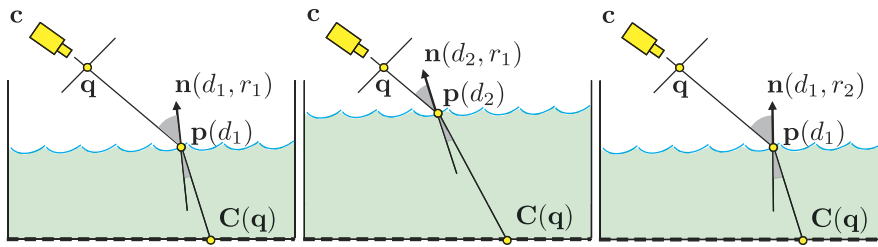
Fig. 3. Single-viewpoint ambiguities. (Left) A view of pixel $\mathbf{q}$'s refraction plane is shown. (Middle) Given a refractive index r_1, we can find, for each distance d_2 to the surface, a normal that refracts point $\mathbf{C}(\mathbf{q})$ to its corresponding pixel $\mathbf{q}$. (Right) Similarly, given a distance d_1, we can find, for each refractive index r_2, a normal that refracts point $\mathbf{C}(\mathbf{q})$ to pixel $\mathbf{q}$.

To compute $\mathbf{C}(\mathbf{q}, t)$ we rely on a procedure that computes the correspondences for time $t = 0$ and then propagates them through time using optical flow estimation.

Since the refractive index is the same for all pixels, we seek a value that most closely satisfies the refractive stereo geometry across all pixels and all frames. We perform a discrete 1D search in an interval of plausible refractive indices and, for each hypothesized refractive index value, attempt to reconstruct the scene for all pixels and frames. We then choose the value that produces the smallest reconstruction error. This leads to the following general algorithm whose steps are discussed in the following sections:

- Step 1. Initialize pixel-to-pattern correspondences, $\mathbf{C}(\mathbf{q}, 0)$.
- Step 2. For each frame $t > 0$, estimate 2D optical flow to compute $\mathbf{C}(\mathbf{q}, t)$ from $\mathbf{C}(\mathbf{q}, t-1)$.
- Step 3. For every refractive index $r \in \{r_1, \ldots, r_n\}$, every frame t, and every pixel $\mathbf{q}$

    - assuming refractive index r for the liquid, estimate the 3D position $\mathbf{p}$ and normal $\mathbf{n}$ of the surface point projecting to pixel $\mathbf{q}$ at time t.
    - estimate the reconstruction error (Section 4.2)

$$e(r, \ t, \ \mathbf{q}) = RE(\mathbf{p}, \mathbf{n}). \qquad (5)$$

- Step 4. Set r^* as

$$r^* = \arg \min_r \sum_{t, \mathbf{q}} e(r, \ t, \ \mathbf{q}), \qquad (6)$$

and return the distances and normals reconstructed with this index value.

4   PIXEL-WISE SHAPE ESTIMATION

The key step in refraction stereo is an optimization procedure that assigns a 3D point $\mathbf{p}$ and a surface normal $\mathbf{n}$ to each pixel. The procedure assumes that the refractive index has a known value r and computes the $\mathbf{p}, \mathbf{n}$ that are most consistent with Snell's law and the pixel-to-pattern correspondence function for the input views.

For a given pixel $\mathbf{q}$, the optimization works in two stages. In the first stage, we conduct a 1D optimization along the ray through pixel $\mathbf{q}$. The goal is to find the distance d that globally minimizes a novel criterion, called the *refractive disparity (RD)*. This criterion is specifically designed to

avoid instabilities due to degenerate refraction paths (e.g., when the liquid's surface is close to the reference pattern).

The optimal d-value gives us initial estimates, $\mathbf{p}(d)$ and $\mathbf{n}(d, r)$, for the 3D coordinates and surface normal of a point that projects to pixel $\mathbf{q}$. These estimates are further refined in a second, bundle adjustment stage in which all five parameters (two for the normal, three for the position) are optimized simultaneously.

4.1   Measuring Refractive Disparity

Each value of d defines an implicit correspondence between four known points (Fig. 1): pixel $\mathbf{q}$, point $\mathbf{C}(\mathbf{q})$ on the reference pattern that refracts to $\mathbf{q}$, the projection $\mathbf{q}'$ of $\mathbf{p}(d)$ in the second viewpoint, and point $\mathbf{C}(\mathbf{q}')$. This correspondence must be consistent with Snell's law.

In their work on reconstructing mirror-like surfaces, Bonfort and Sturm [6] noted that such a correspondence gives us two "candidate" normals for $\mathbf{p}(d)$ which must be identical when this hypothesis is correct. These normals are obtained by applying (2) twice, once for each viewpoint. Specifically, the first normal, $\mathbf{n}_1 = \mathbf{n}(d, r)$, ensures that point $\mathbf{C}(\mathbf{q})$ on the reference pattern refracts to pixel $\mathbf{q}$ via point $\mathbf{p}(d)$. The second normal, $\mathbf{n}_2$, enforces a similar condition for the second viewpoint, i.e., it ensures that point $\mathbf{C}(\mathbf{q}')$ refracts to pixel $\mathbf{q}'$ via point $\mathbf{p}(d)$. We obtain $\mathbf{n}_2$ by applying (2) to pixel $\mathbf{q}'$, using its distance from point $\mathbf{p}(d)$. Since points on a smooth surface have a unique normal, a necessary condition for $\mathbf{p}(d)$ being on the "true" surface is that $\mathbf{n}_1 = \mathbf{n}_2$.

Unfortunately, even though it is possible, in principle, to directly measure the alignment of vectors $\mathbf{n}_1$ and $\mathbf{n}_2$, such a measurement becomes unstable when the distance between the surface and the reference pattern approaches zero. This is because, as refraction diminishes, (2) becomes singular, normals cannot be estimated accurately, and the 3D reconstruction problem degenerates to standard stereo. In practice, this causes instability for low liquid heights, making direct comparison of normals uninformative and inappropriate for reconstruction.

Instead of measuring the alignment of the two normals $\mathbf{n}_1$ and $\mathbf{n}_2$ directly, we perform an indirect measurement that is not singular when refraction diminishes. The main idea is that if $\mathbf{n}_1$ and $\mathbf{n}_2$ were truly aligned, "swapping" them would still force point $\mathbf{C}(\mathbf{q})$ to refract to pixel $\mathbf{q}$ and point $\mathbf{C}(\mathbf{q}')$ to pixel $\mathbf{q}'$. We therefore define the criterion by asking two questions (Fig. 4a):
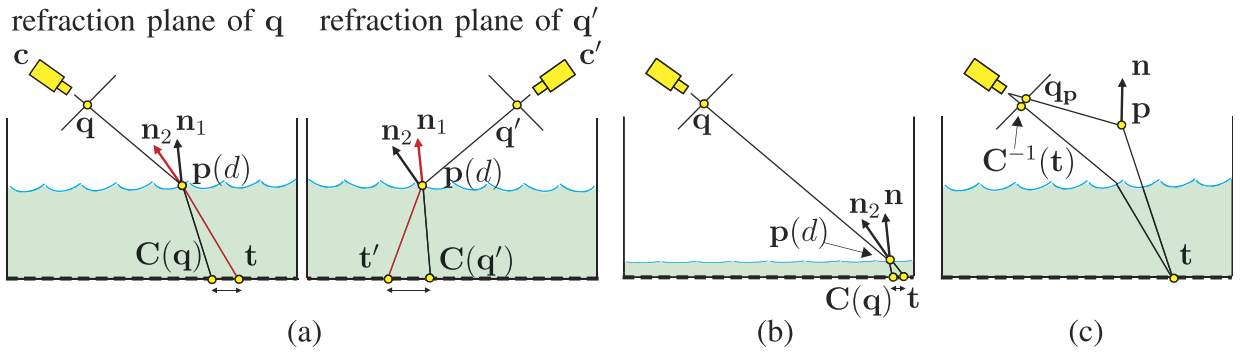
Fig. 4. Optimization criteria for refraction stereo. (a) Measuring refractive disparity. Normals are drawn according to the refractions they produce. (b) For small surface-to-pattern distances, swapping $\mathbf{n}_1$ and $\mathbf{n}_2$ does not influence the distance between $\mathbf{C}(\mathbf{q})$ and $\mathbf{t}$ significantly. (c) Measuring image reprojection error at one of the viewpoints.

- Suppose the normal at $\mathbf{p}(d)$ is $\mathbf{n}_2$; which point on the reference pattern will refract to $\mathbf{q}$?
- Suppose the normal at $\mathbf{p}(d)$ is $\mathbf{n}_1$; which point on the reference pattern will refract to $\mathbf{q}'$?

Now suppose that points $\mathbf{t}, \mathbf{t}'$ are the points that refract to pixels $\mathbf{q}, \mathbf{q}'$, respectively. The distance between $\mathbf{t}$ and $\mathbf{C}(\mathbf{q})$ and, similarly, the distance between $\mathbf{t}'$ and $\mathbf{C}(\mathbf{q}')$ can be thought of as a measure of disparity. Intuitively, this distance tells us how swapping the normals $\mathbf{n}_1, \mathbf{n}_2$ affects consistency with the available pixel-to-pattern correspondences. To evaluate a hypothesis d, we simply sum these distances:

Refractive Disparity

$$RD(d) = \|\mathbf{t} - \mathbf{C}(\mathbf{q})\|^2 + \|\mathbf{t}' - \mathbf{C}(\mathbf{q}')\|^2 . \quad (7)$$

When the distance between the true surface and the reference pattern is large, refractive disparity is equivalent to a direct measurement of the alignment between vectors $\mathbf{n}_1, \mathbf{n}_2$, i.e., it is zero if and only if $\mathbf{n}_1 = \mathbf{n}_2$. On the other hand, as the liquid's true surface approaches the reference pattern, refractive disparity diminishes. This is because the refractive effect of changing a point's surface orientation diminishes as well (Fig. 4b). As a result, the minimization can be applied to any image pixel for which $\mathbf{C}(\mathbf{q})$ is known, regardless of whether or not the ray through the pixel actually intersects the liquid's surface.

To compute point $\mathbf{t}$ for a given d value, we trace a ray from the first viewpoint through pixel $\mathbf{q}$, refract it at point $\mathbf{p}(d)$ according to normal $\mathbf{n}_2$, and intersect it with the (known) surface of the reference pattern. Point $\mathbf{t}'$ is computed in an identical manner. To find the distance d that globally minimizes refractive disparity along the ray, we use Matlab's `fminbnd()` function, which is based on golden-section search [44].

4.2 Computing 3D Position and Orientation

Even though refractive disparity minimization yields good reconstructions in practice, it has two shortcomings. First, it treats the cameras asymmetrically because optimization occurs along the ray through one pixel. Second, it only optimizes the distance along that ray, not the 3D coordinates and orientation of a surface point. We therefore use an additional step that adjusts all shape parameters ($\mathbf{p}$ and $\mathbf{n}$) in order to minimize a symmetric image reprojection error.

To evaluate the consistency of $\mathbf{p}$ and $\mathbf{n}$, we check whether the refractions caused by such a point are consistent with the refractions observed in the input views. In particular, let $\mathbf{q_p}, \mathbf{q'_p}$ be the point's projections in the two cameras and suppose that $\mathbf{t}, \mathbf{t}'$ are the points on the reference pattern that refract to $\mathbf{q_p}, \mathbf{q'_p}$, respectively, via point $\mathbf{p}$. To compute the reprojection error, we measure the distance between pixels $\mathbf{q_p}, \mathbf{q'_p}$ and the "true" refracted image of $\mathbf{t}, \mathbf{t}'$ (Fig. 4c):

$$RE(\mathbf{p}, \mathbf{n}) = \|\mathbf{q_p} - \mathbf{C}^{-1}(\mathbf{t})\|^2 + \|\mathbf{q'_p} - \mathbf{C}^{-1}(\mathbf{t}')\|^2 \\ + \beta G(\|\mathbf{p} - \mathbf{p}_0\|; \ \sigma)^{-1}, \quad (8)$$

where $\mathbf{C}^{-1}(.)$ denotes the inverse of the pixel-to-pattern correspondence function, $G(.; \sigma)$ is the Gaussian with standard deviation σ, and $\mathbf{p}_0$ is the starting point of the optimization. The Gaussian term ensures that the optimization will return a point $\mathbf{p}$ whose projection, $\mathbf{q_p}$, always remains in the neighborhood of the originally chosen pixel $\mathbf{q}$ in Steps 1 and 2 of the Dynamic Refraction Stereo algorithm (Section 3). We used $\sigma = 4$ and $\beta = 200$ for all of our experiments. To minimize the RE functional with respect to $\mathbf{p}$ and $\mathbf{n}$ we use the downhill simplex method [44].

5 IMPLEMENTATION DETAILS

5.1 Estimating Pixel-to-Pattern Correspondences

Accurate 3D shape recovery requires knowing the pixel-to-pattern correspondence function $\mathbf{C}(\mathbf{q}, t)$ with high accuracy. While color-based techniques have been used to estimate this function for image-based rendering applications [10], they are not appropriate for reconstruction for several reasons. First, different liquids absorb different wavelengths by different amounts [49], altering a pattern's appearance in a liquid-dependent way. Second, since light absorption depends on distance traveled within the liquid and since this distance depends on the liquid's instantaneous shape, the appearance of the same point on a pattern will change through time. Third, the intensity of light transmitted through the surface depends on the Fresnel effect [18] and varies with wavelength and the angle of incidence. This makes it difficult to use color as a means to localize points on a pattern with subpixel accuracy.

In order to avoid these complications, we use a monochrome checkered pattern and rely on corners to
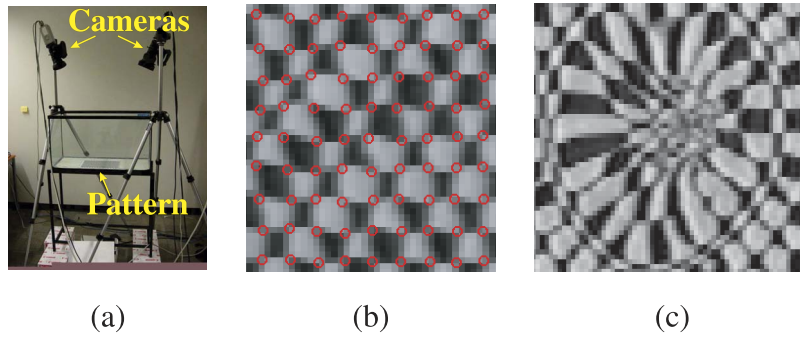
Fig. 5. (a) Experimental setup. (b) Typical close-up view of pattern, seen through water surface. (c) Distorted view, corresponding to tracking failure at the central corners.

establish and maintain pixel-to-pattern correspondences (Figs. 5a and 5b). We assume that the liquid's surface is undisturbed at time $t = 0$ and use the Harris corner detector [21] to detect corners at subpixel resolution. This gives us the initial pixel-to-pattern correspondences. To track the location of individual corners in subsequent frames while avoiding drift, we estimate flow between the current frame and the frame at time $t = 0$, using the flow estimates from the previous frame as an initial guess. We compute flow with a translation-only version of the Lucas-Kanade inverse-compositional algorithm [3] and use Levenberg-Marquardt minimization to obtain subpixel registration. This algorithm is applied to an 11×11 pixel neighborhood around each corner. We use the registration error returned by the algorithm as a means to detect failed localization attempts. In the case of failure, the flow computed for that corner is not used and the corner's previously known location is propagated instead. This allows our tracker to overcome temporary obscurations due to blur, splashes, or extreme refractive distortions (Fig. 5c). The above procedure gives values of the correspondence function $\mathbf{C}(\mathbf{q}, t)$ for a subset of the pixels. To evaluate the function for every pixel, we use bilinear interpolation.

5.2 Fusing 3D Positions and Orientations

Refraction stereo yields a separate 3D position and 3D normal for each pixel. While this is a richer shape descriptor, the problem of reconstructing a single surface that is consistent with both types of data is still open. A key difficulty is that point and normal measurements have different noise properties, and hence a surface computed

via normal integration and a surface computed by fitting a mesh to the 3D points will not necessarily agree. As a first step, we used simulations and ground-truth experiments to estimate the reliability of each data source as a function of surface height, i.e., distance from the plane of the reference pattern (Fig. 6). Since reconstructed normals are highly reliable for large heights, we used this analysis to set a height threshold below which normals are deemed less reliable than positions. That portion of the surface is reconstructed from positional data. For the remaining pixels, we reconstruct the surface via normal integration using the Ikeuchi-Horn algorithm [27] and merge the results. In cases where all reconstructed positions are above the height threshold, we rely on normal integration to compute 3D shape and use the average 3D position to eliminate the integrated surface's height ambiguity.

6 EXPERIMENTAL RESULTS

6.1 Experimental Setup

Fig. 5a shows our setup. The checkered pattern at the bottom of the tank was in direct contact with the water to avoid secondary refractions. During our experiments, the pattern was brightly lit from below to avoid specular reflections and to enable use of a small aperture for the cameras (and hence, a large depth of field). Images were acquired at a rate of 60 Hz with a pair of synchronized Sony DXC-9000 progressive-scan cameras, whose electronic shutter was set to 1/500 sec to avoid motion blur. Both cameras were approximately 1 meter above the tank
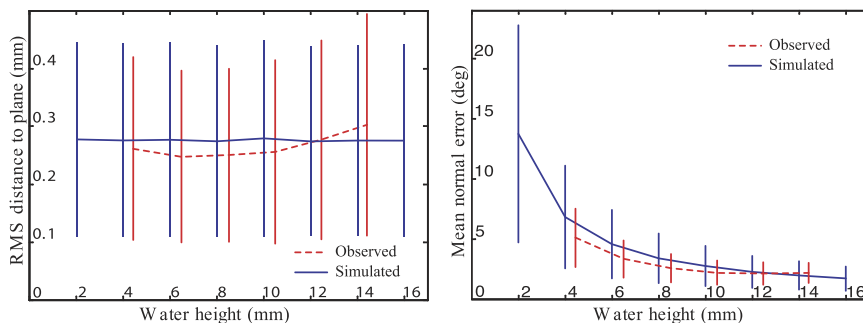

Fig. 6. Reconstruction accuracy as a function of water height, for real (dotted line) and simulated (solid line) flat water surfaces. Bars indicate standard deviation. Simulations are for a 0.08 pixel localization error; in real flat water experiments, corner localization precision was measured to be ~0.1 pixels.
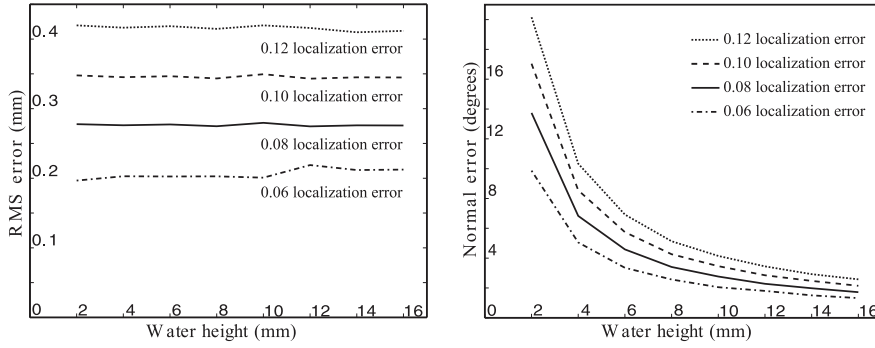
Fig. 7. (Left) Reconstruction accuracy as a function of water height for several values of pixel localization error. (Right) Normal reconstruction error for the same pixel localization error values.

bottom and were calibrated using the Matlab Calibration Toolbox [8].

6.2 Simulations

To evaluate the stability of our algorithms, we performed simulations that closely matched the experimental conditions in the lab (e.g., relative position of cameras, pattern-to-camera distances, feature localization errors, etc.). We simulated the reconstruction error for planar water surfaces as a function of the surface-to-pattern height, and for various levels of error in corner localization and camera calibration. We modeled the localization error by perturbing the image coordinates of the projected corners by a Gaussian with a fixed standard deviation. For each height, we reconstructed 10,000 individual points and measured their deviation from the ground-truth plane (Fig. 6). These simulations confirm that the accuracy of reconstructed normals degrades quickly for water heights less than 4 mm. Importantly, the accuracy of distance computations is not sensitive to variations in water height, confirming the stability of our optimization-based framework for refractive stereo (Section 4.1). We also compared the effect of localization error on the results (Fig. 7).

In addition, we ran simulations to test the stability of refractive index estimation. We simulated a stationary sinusoidal surface whose average height was 40 mm and whose amplitude was 2 mm. We then computed the total reconstruction error for various combinations of true and hypothesized refractive index values (Fig. 8). We used a localization error of 0.1 pixels for these simulations to reflect our actual experimental conditions. These simulations show that our objective function has a minimum very close to the expected refractive index. Also note that the valley around the minimum becomes more shallow as the refractive index increases.
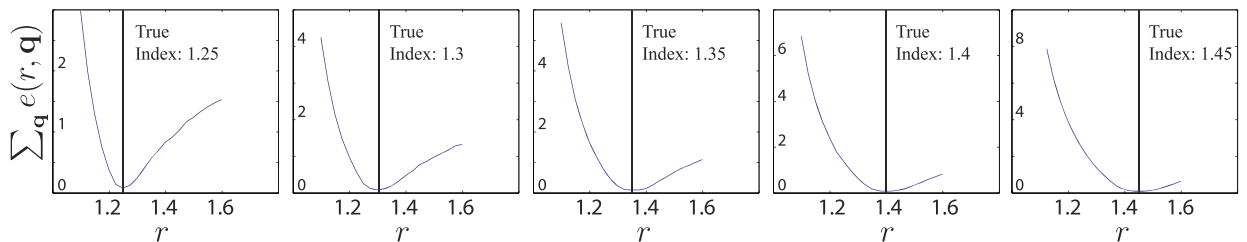
6.3 Accuracy Experiments

Since ground truth was not available, we assessed our algorithm's accuracy by applying it to the reconstruction of flat water surfaces whose height from the tank bottom ranged from 4 to 15 mm. For each water height, we reconstructed a point **p** and a normal **n** independently for each of 1,836 pixels in the two image planes, giving rise to as many 3D points and normals. No smoothing or postprocessing was performed. To assess the accuracy of the reconstructed points, we fit a plane using least squares and measured the points' RMS distance from this plane. To assess accuracy in the reconstructed normals, we computed the average normal and measured the mean distance of each reconstructed normal from the average normal. These results, also shown in Fig. 6, closely match the behavior predicted by our simulations. They suggest that reconstructions are highly precise, with distance variations around 0.25 mm (i.e., within 99.97 percent of the surface-to-camera distance) and normal variations on the order of 2 degrees for water heights above 8 mm.

6.4 Experiments with Dynamic Surfaces

Figs. 11, 12, and 13 show reconstructions for several dynamic water surfaces. The experiments test our algorithm's capabilities under a variety of conditions—from rapidly fluctuating water that is high above the tank bottom to water that is being poured into an empty tank where the water height is very small and refraction is degenerate or near-degenerate for many pixels.

Several observations can be made from these experiments. First, our tracking-based framework allows us to maintain accurate pixel-to-pattern correspondences for 100s of frames, enabling dynamic reconstructions that last several seconds. Second, the reconstructed distances remain

Fig. 8. Total reconstruction error as a function of hypothesized refractive index r. Simulations are shown for five different ground-truth refractive indices.
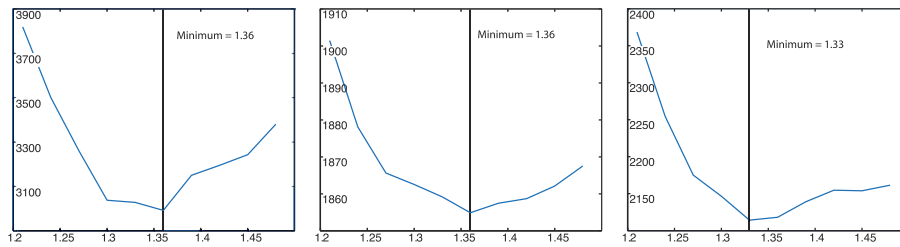
Fig. 9. Total reconstruction error as a function of refractive index. (Left) Combined error of several frames from the RIPPLE sequence. (Middle) Combined error of several frames from the POUR sequence. (Right) Combined error of several frames from the WAVES sequence.

stable despite large variations in water height and are accurate enough to show fine surface effects, even in cases where the total water height never exceeds 6 mm (e.g., the "pour" sequence). Third, the reconstructed normal maps, as predicted, show fine surface fluctuations at larger heights but degrade to noise level for water heights near zero. Qualitatively, when there is sufficient water in the tank, they appear to contain less noise than depth maps. Fourth, our normal integration algorithm seems to oversmooth fine surface details that are clearly present in the depth and normal maps. This suggests that a more sophisticated approach is needed here (e.g., [41]). The following paragraphs discuss our results in more detail.

Ripple sequence (Fig. 11). The ripple sequence shows a drop of water dripping into a tank with approximately 25 mm of water in it.[3] The drop causes layers of circular waves to emanate from the point of contact. There are a combination of large and fine-scale waves in the sequence, making for an interesting reconstruction problem. We achieved a good reconstruction of depth and normals since the surface depth provided sufficient refraction for reliable readings in both categories. The normals provide better resolution of the fine ripples. The initial splash caused when the drop first hits the water produced strong distortions of the underlying pattern which our system was unable to track. We thus reinitialized the system immediately after the worst distortions had elapsed and were able to capture the expanding circular waves.

In the depth maps, errors are more common at points of greatest distortion of the ripple, causing sharp peaks (bright or dark spots in Fig. 11). Despite these depth errors, the corresponding normals appear to be correct and the normal maps are very smooth, allowing tiny leading ripples to be easily identified.

Pour sequence (Fig. 12). This sequence shows water being poured into an empty tank, spreading across the pattern from left to right. This sequence tests our system's ability to cope with water heights that are zero or near zero. The sequence also exhibits large shape variations due to ripples, as well as bubbles forming on the surface. Our reconstruction successfully handled the height variations using refractive disparity: Although reconstructed normals exhibited increased noise in the shallow portions, the depths were reconstructed faithfully. Note that the large regions of noise in the normal maps correspond to areas with no water.

Also note that the reconstructed height maps are smooth at the water edges, suggesting robustness to shallow water.

In areas of nonzero water height, the normal maps provide fine detail of the surface, lacking much of the noise evident in the height maps. For instance, fine waves and ripples can be seen in the normal maps at all three time instants shown in Fig. 12. By comparison, only the larger-scale waves appear in the height map because noise tends to obscure finer details. Note that while our system is not designed to handle bubbles, the reconstruction actually captured indentations corresponding to the bubbles as well as the ripples forming when they burst (e.g., see the middle of the normal map in the second time instant).

Waves sequence (Fig. 13). This sequence shows waves propagating from the left to the right on a water surface at a height of approximately 24 mm. There are several interleaving wavefronts of various scales. This sequence provides the greatest variation in water height and overall roughness. Reconstructions from this data set suffered from more calibration error and thus exhibit stronger noise. This is most noticeable in the moiré patterns appearing in both the depth and normal maps. Despite this, we did obtain interesting reconstructions of the rough water surface. We recover both the large and small scale wave fronts, which can be visually tracked across the sequence in the normal maps.

6.5 Refractive Index Determination

In addition to reconstructing both surface and normal data, our system was able to obtain an estimate of the refractive index of the transparent medium. We ran our algorithm on all three sequences of Figs. 11, 12, and 13, over a range of seven frames for each. Fig. 9 shows the total reconstruction error corresponding to specific values of the refractive index, for the each data set. The curves exhibit a minimum near the correct refractive index for water, 1.33, confirming the predictions of Theorem 1. This is because Step 4 of our algorithm enforces a very strong *global* constraint: The light path of every pixel at every time instant and at every viewpoint must be consistent with the *same* refractive index value.

7 DISCUSSION

7.1 Ambiguous Surfaces

Although our experimental results suggest that it is possible to estimate refractive indices from stereo sequences, a general question is whether this problem can be ambiguous, i.e., whether two surfaces with different refractive indices can produce the same observed stereo image pair. Our
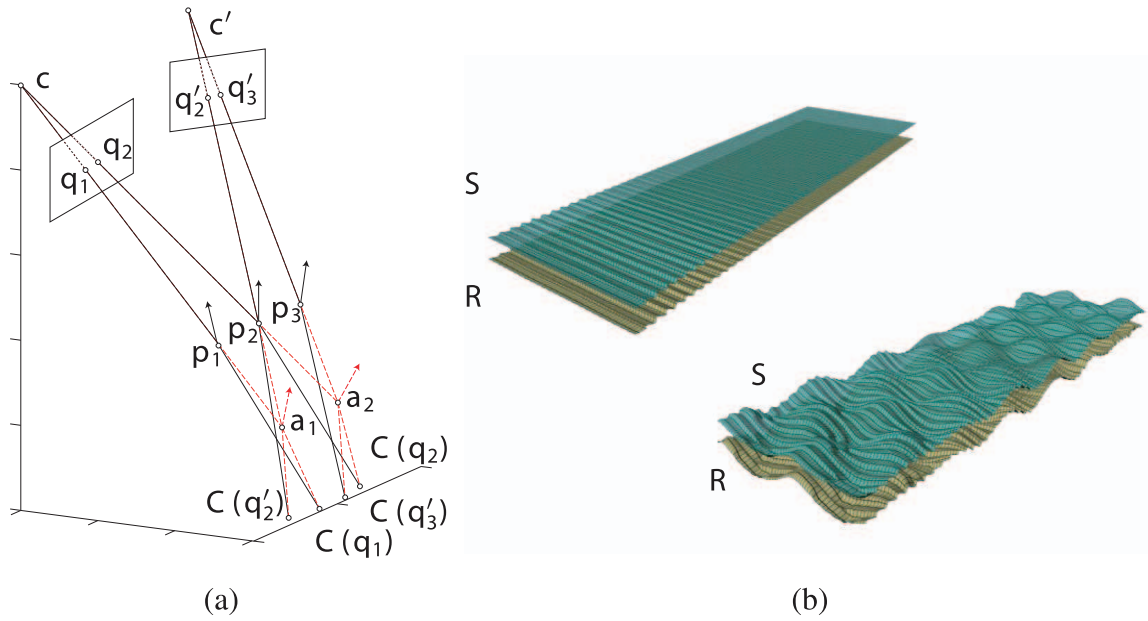
---

3. See the supplemental videos, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TPAMI.2011.24.

Fig. 10. (a) Ambiguous surface construction: Points $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$ lie on surface S, points $\mathbf{a}_1$ and $\mathbf{a}_2$ lie on the ambiguous surface R. Note that correspondence functions $\mathbf{C}(\mathbf{q}_2')$, $\mathbf{C}(\mathbf{q}_3')$, $\mathbf{C}(\mathbf{q}_1)$, and $\mathbf{C}(\mathbf{q}_2)$ all agree with both the true and ambiguous surface points and their normals. (b) Ambiguous surface pairs constructed according to Section 7.1. (Left) The depths of S were chosen to fit a plane. (Right) The depths of S were chosen to fit a sinusoidal surface.

initial analysis, described below, indicates that such pairs of surfaces do exist but they do not have a "simple" shape: Indeed, the points on a pair of ambiguous surfaces must satisfy a joint system of very complex trigonometric equations for such an ambiguity to exist.

We use a numerical approach to construct point samples on such surface pairs as follows: Given the refractive index r of the true surface S and the refractive index r_a of an ambiguous surface R, we construct points and normals that lie on both S and R. First, we begin with a seed point $\mathbf{p}_1$ on S which is imaged at $\mathbf{q}_1$ in $\mathbf{c}$ (Fig. 10). Then, we find a second point $\mathbf{a}_1$ that lies along the ray from $\mathbf{q}_1$ to $\mathbf{p}_1$ and on the ambiguous surface R. The depth of $\mathbf{a}_1$ and its normal $\mathbf{m}_1$ must be chosen so that the pixel ray refracts to point $\mathbf{C}(\mathbf{q}_1)$ on the tank bottom. Next, we project $\mathbf{a}_1$ into view $\mathbf{c}'$ and intersect this ray with S to obtain another point $\mathbf{p}_2$. The normal of $\mathbf{p_2}$ must cause the ray through pixel $\mathbf{q}_2'$ to be refracted toward $\mathbf{C}(\mathbf{q}_2')$. From our initial seed point, this process gives us two new points: $\mathbf{a}_1$ on R and $\mathbf{p}_2$ on S. We can repeat this process either by adding new seed points or by repeating the above steps with $\mathbf{p}_2$ as the seed. Fig. 10a shows this construction repeated twice. Since we have only one degree of freedom in the depth/normal of the points $\mathbf{p}_i$ and $\mathbf{a}_i$, any ambiguous surface pairs must be highly constrained. For instance, if we force the points $\mathbf{p}_i$ and $\mathbf{a}_i$ to lie on parallel planes, the normals associated with these points cannot both satisfy these constraints and agree with the global planar normal. Fig. 10b shows two surface pairs constructed from a sparse set of seeds as described. Intermediate surface points were interpolated according to the sparse points and their normals.

While the above construction suggests the existence of ambiguous surfaces, these surfaces are not as important in practice because the scene is dynamic: In our algorithm, a single refractive index value must account for the refractions produced by the *entire sequence* of 3D surfaces of the liquid, not just the 3D surface in a single instant. In fact, refractive index estimation exploits the dynamic/statistical nature of liquids in three ways: 1) The surface is highly variable and hence we observe many different, complex surfaces with the *same* refractive index during image acquisition, 2) their deforming surface is unlikely to globally match one of the "special" ambiguous shapes, and 3) even if it does, it is unlikely that such a "special" shape will occur for many time instants. In practice, this allows us to side-step the issue of ambiguities by enforcing refractive index consistency with all available data (multiple time instants, even multiple acquisition experiments) using our search-based refractive index estimation algorithm. Experimentally, the lack of shape-index ambiguities over an acquired data set is confirmed by error curves that have only one (global) minimum, as in Fig. 9.

7.2 Pixel-to-Pattern Function

In our description of the algorithm and its implementation, we assumed that $\mathbf{C}()$ was invertible. We note, however, that $\mathbf{C}()$ can be many-to-one and not generally invertible. The conditions that cause $\mathbf{C}()$ to be invertible were noted by Murase [40], deemed reasonable, and used in his work on liquid reconstruction. From a technical standpoint, however, our analysis does not require $\mathbf{C}()$ to be globally invertible: All we need is that it is *locally* invertible, i.e., for almost all pixels $\mathbf{q}$ (in a measure-theoretic sense), the restriction of $\mathbf{C}()$ to some open neighborhood of $\mathbf{q}$ is an invertible function. This does permit the occurrence of isolated singularities (i.e., pixels or image curves where $\mathbf{C}()$ is not invertible for any neighborhood).

For example, if we were to take a simple 2D scene such as a surface defined by $y = cos(x)$ with the camera looking down the $-y$ axis, $\mathbf{C}()$ is not globally invertible. The scene is, however, locally invertible for all values of x except two: For
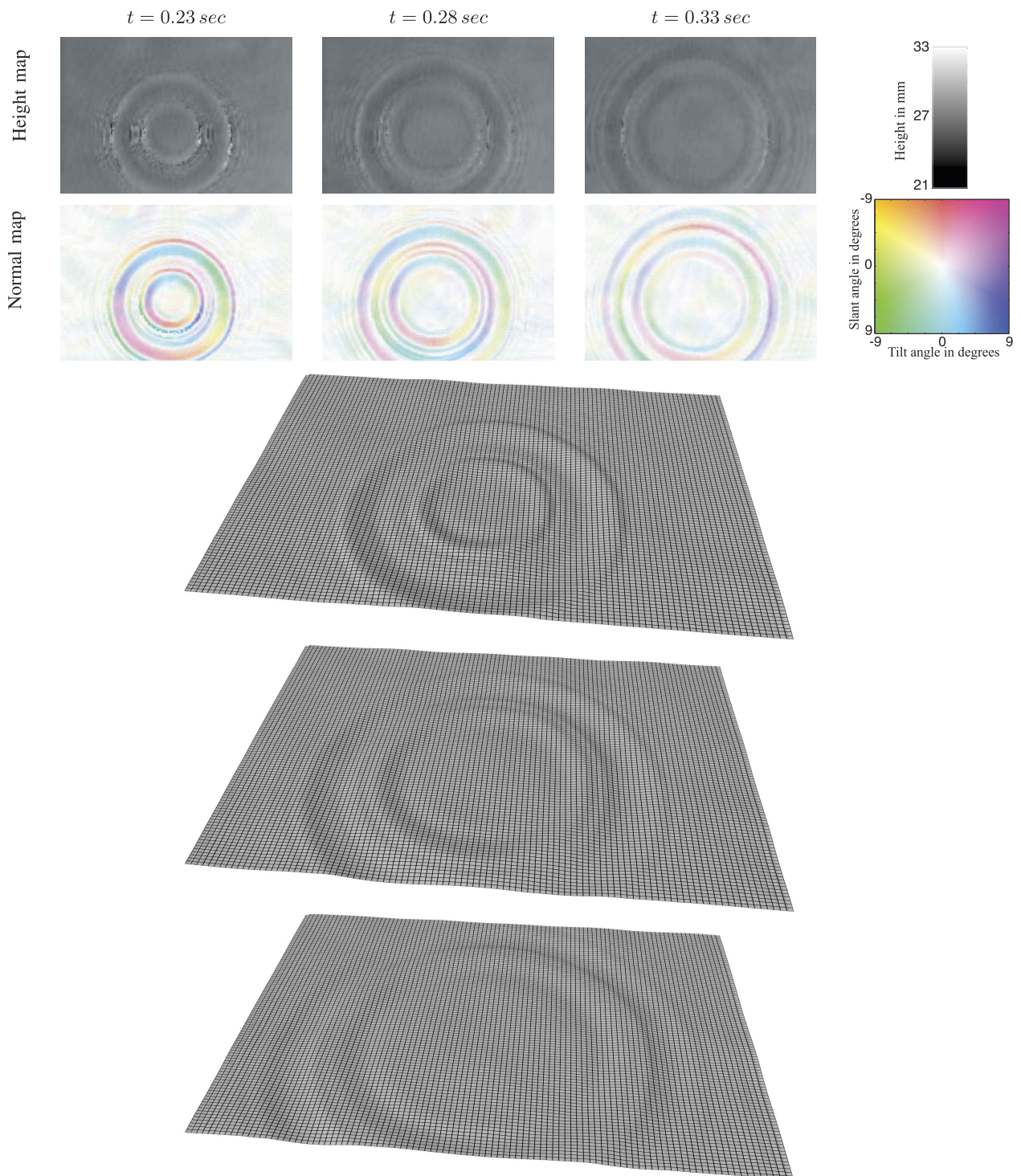
Fig. 11. RIPPLE sequence. All maps correspond to a top view of the tank and show raw, per-pixel data. The mesh images show a surface fit to the data scaled by 5 in the vertical axis.

a given refractive index value and a camera located at infinity, there are only two incoming rays/pixels where local invertibility breaks down: These rays hit the $cos(x)$ curve near its inflection points, where the mapping from incoming rays to points on the x-axis "folds" onto itself. More generally, the singularities where local invertibility breaks down have properties analogous to the singularities of the Gauss map where, generically, the mapping from surface points to their normals is singular either on parabolic curves,

corresponding to "folds" of the Gauss map, or on isolated points, corresponding to "cusps" of the map.

We also examine in further detail how the flow propagates from $\mathbf{C}(t-1)$ to $\mathbf{C}(t)$. There are two cases: 1) The Lucas-Kanade algorithm is able to localize in frame t a corner that was also localized in frame $t-1$. In this case, the flow vector assigned to the corner at time $t-1$ is its displacement between the two frames. This process is completely local and is well defined wherever $\mathbf{C}()$ is locally

$t = 0.6\,sec$ $t = 0.92\,sec$ $t = 1.32\,sec$
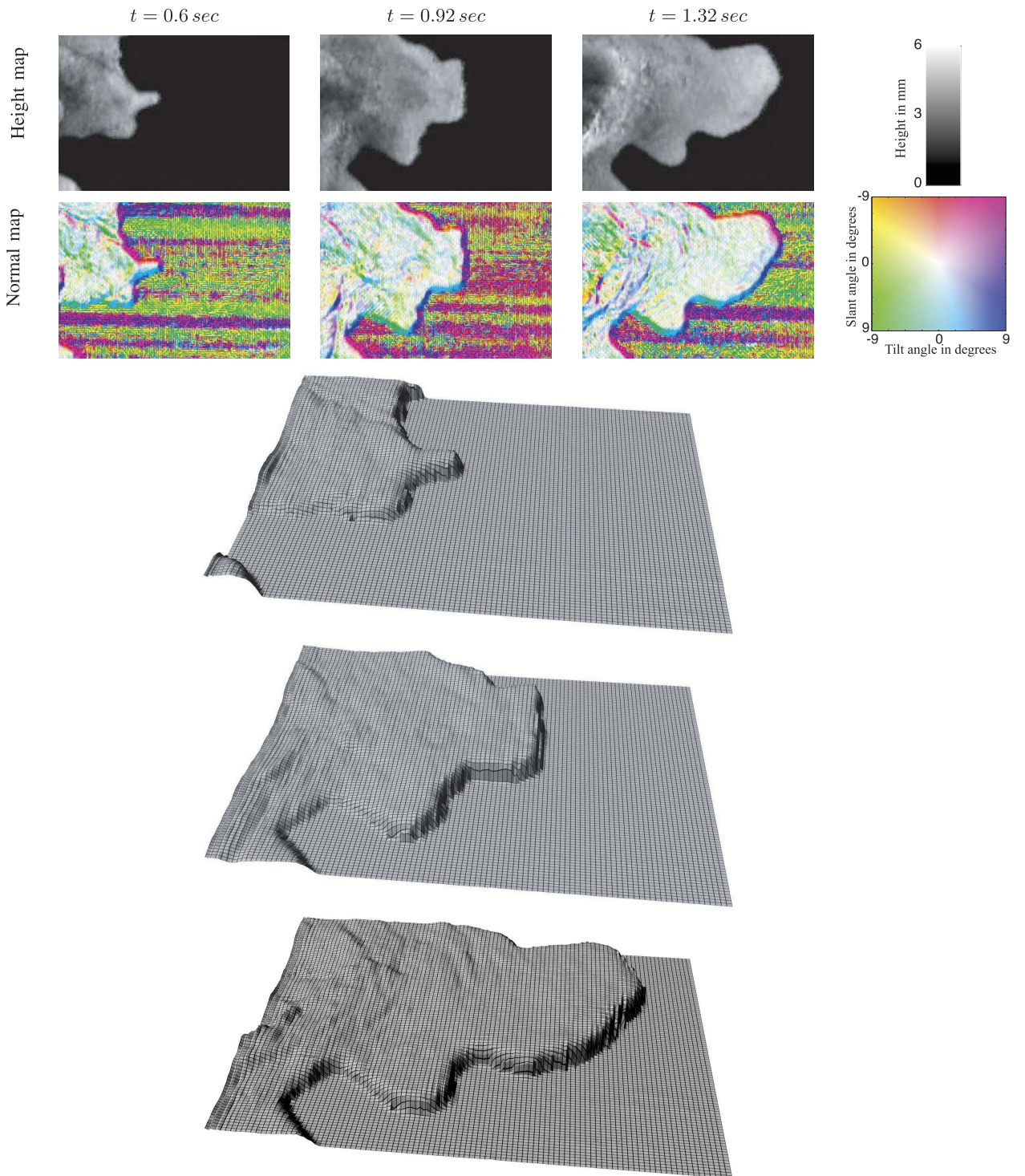

Fig. 12. POUR sequence. All maps correspond to a top view of the tank and show raw, per-pixel data. The mesh images show a surface fit to the data scaled by 5 in the vertical axis.

(but perhaps not globally) invertible. 2) The LK algorithm fails to localize the corner at frame t. This does *not* cause a breakdown of flow estimation for subsequent frames. In this case, the algorithm interpolates the flow vectors computed at four neighboring corners in frame $t - 1$ in order to assign a flow vector to the corner that was lost in frame t. Bilinear interpolation is used, with weights determined by the corners' distance from each other. We then use the position in frame $t - 1$ of the lost corner and the interpolated flow vector to assign it a "virtual" position in frame t. This position is used to initialize the LK algorithm in frame $t + 1$ in an attempt to relocalize the lost corner. In case of failure at $t + 1$, propagation is repeated until the corner is reacquired. Since these distortions are local and persist for just few frames, we have found that the strategy works well in practice and has enabled propagation of pixel-to-pattern correspondences for 100s of frames.

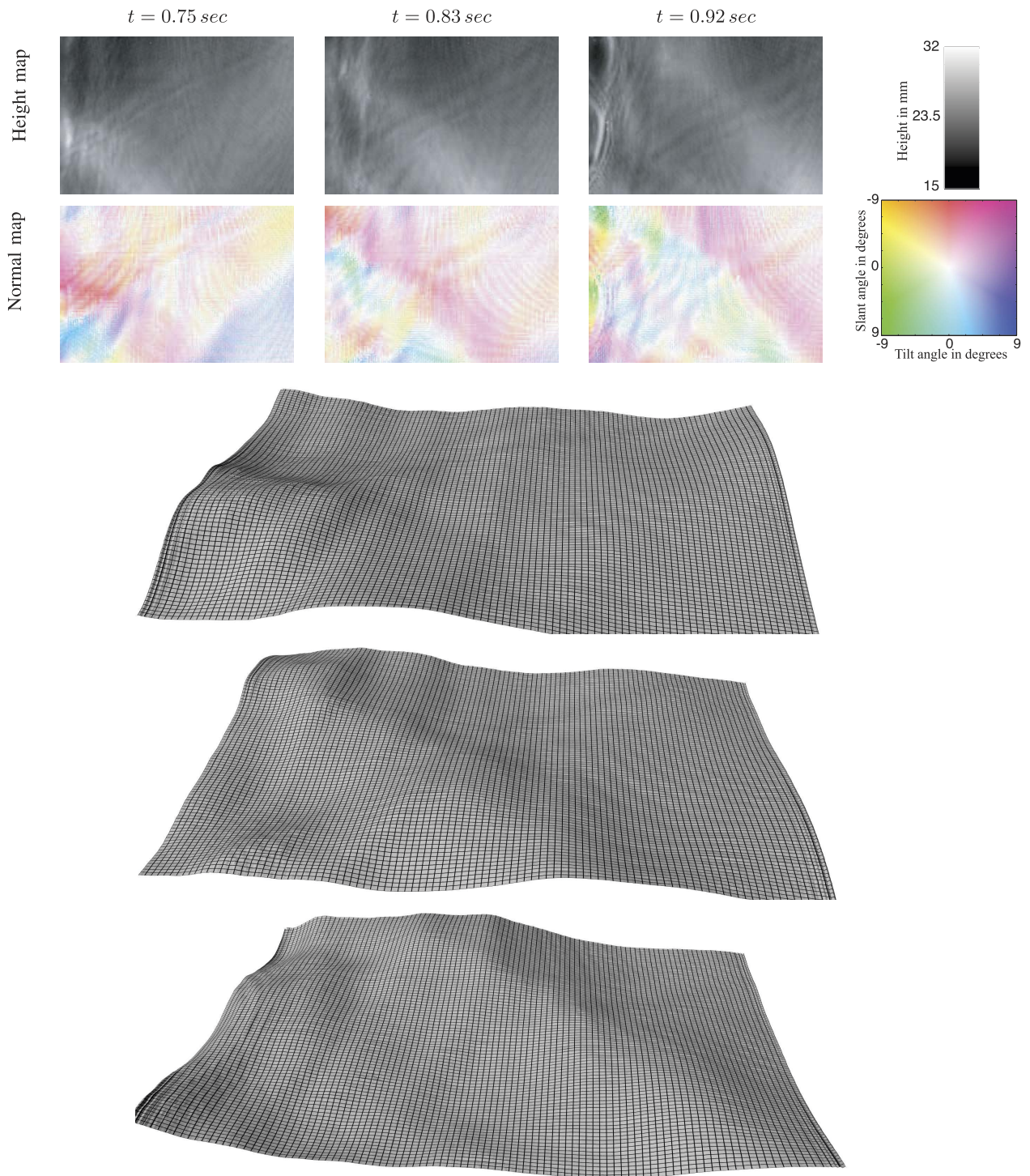$t = 0.75\,sec$ $t = 0.83\,sec$ $t = 0.92\,sec$

Fig. 13. WAVES sequence. All maps correspond to a top view of the tank and show raw, per-pixel data. The mesh images show a surface fit to the data scaled by 5 in the vertical axis.

Our current implementation does not check for the possibility that after a tracking failure (i.e., singularity of $\mathbf{C}()$), a corner at frame $t-1$ appears in more than one location in frame t (or vice versa). While it is certainly possible to do so, the sequences we have acquired suggest that such events are very transient and cause significant distortions in the local neighborhood of a corner, making it very hard to localize it, let alone identify multiple images of

it. In such cases, flow propagation allows the corner to be reacquired when distortions are reduced.

7.3 Optimization

Our focus in this paper has been on the design of optimization functionals for performing pixel-wise refraction stereo calculations ((8) and (5)), not on the optimization process itself. In particular, our exhaustive search over discretized refraction indices could potentially be replaced

by a golden-section search [17] or by a nonlinear optimization procedure that optimizes all points, normals, and refractive index simultaneously. Similarly, the surface reconstruction procedure in Section 5.2 is largely heuristic —a more rigorous approach would be to formulate it as a global procedure that solves for the entire surface, in the spirit of recent approaches to multiview stereo [34]. In that case, the functional $RE(\mathbf{p}, \mathbf{n})$ in (8) would replace the traditional measures of photo consistency in multiview stereo.

8 CONCLUDING REMARKS

Liquids can generate extremely complex surface phenomena, including breaking waves, bubbles, and extreme surface distortions. While our refraction stereo results are promising, they are just an initial attempt to model liquid flow in relatively simple cases. Our ongoing work includes 1) reconstructing surfaces that produce multiple refractions [35], 2) reconstructing liquids by exploiting their refractive and reflective properties (e.g., by also treating them as mirrors), and 3) "reusing" captured 3D data to create new, realistic fluid simulations.

ACKNOWLEDGMENTS

REFERENCES

[1] Y. Adato, Y. Vasilyev, O. Ben-Shahar, and T. Zickler, "Toward a Theory of Shape from Specular Flow," *Proc. IEEE 11th Int'l Conf. Computer Vision,* pp. 1-8, 2007.
[2] S. Agarwal, S.P. Mallick, D. Kriegman, and S. Belongie, "On Refractive Optical Flow," *Proc. Eighth European Conf. Computer Vision,* pp. 483-494, 2004.
[3] S. Baker and I. Matthews, "Lucas-Kanade 20 Years On: A Unifying Framework," *Int'l J. Computer Vision,* vol. 56, no. 3, pp. 221-255, 2004.
[4] M. Ben-Ezra and S. Nayar, "What Does Motion Reveal about Transparency?" *Proc. Ninth Int'l Conf. Computer Vision,* pp. 1025-1032, 2003.
[5] A. Blake, "Specular Stereo," *Proc. Int'l Joint Conf. Artificial Intelligence,* pp. 973-976, 1985.
[6] T. Bonfort and P. Sturm, "Voxel Carving for Specular Surfaces," *Proc. Ninth Int'l Conf. Computer Vision,* pp. 591-596, 2003.
[7] T. Bonfort, P. Sturm, and P. Gargallo, "General Specular Surface Triangulation," *Proc. Asian Conf. Computer Vision,* pp. 872-881, 2006.
[8] J.-Y. Bouguet, "MATLAB Camera Calibration Toolbox," http://www.vision.caltech.edu/bouguetj/calib_doc/, 2010.
[9] T. Chen, M. Goesele, and H.-P. Seidel, "Mesostructure from Specularity," *Proc. 2006 IEEE CS Conf. Computer Vision and Pattern Recognition,* vol. 2, pp. 1825-1832, 2006.
[10] Y.-Y. Chuang, D.E. Zonkger, J. Hirdorff, B. Curless, and D.H. Salesin, "Environment Matting Extensions: Toward Higher Accuracy and Real-Time Capture," *Proc. ACM SIGGRAPH,* pp. 121-130, 2000.
[11] J.M. Daida, D. Lund, C. Wolf, G.A. Meadows, K. Schroeder, J. Vesecky, D.R. Lyzenga, B.C. Hannan, and R.R. Bertram, "Measuring Topography of Small-Scale Water Surface Waves," *Proc. Geoscience and Remote Sensing Symp. Conf.,* vol. 3, pp. 1881-1883, 1995.
[12] Y. Ding and J. Yu, "Recovering Shape Characteristics on Near-Flat Specular Surfaces," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 1-8, 2008.
[13] Y. Ding, J. Yu, and P. Sturm, "Recovering Specular Surfaces Using Curved Line Images," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* June 2009.
[14] D. Enright, S. Marschner, and R. Fedkiw, "Animation and Rendering of Complex Water Surfaces," *Proc. ACM SIGGRAPH,* pp. 736-744, 2002.
[15] P. Flach and H.-G. Maas, "Vision-Based Techniques for Refraction Analysis in Applications of Terrestrial Geodesy," *Proc. Int'l Archives of Photogrammetry and Remote Sensing,* pp. 195-201, 2000.
[16] Y. Francken, T. Cuypers, T. Mertens, J. Gielis, and P. Bekaert, "High Quality Mesostructure Acquisition Using Specularities," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 1-7, June 2008.
[17] C.B. Moler, G.E. Forsyth, and M.A. Malcom, *Computer Methods for Mathematical Computations.* Prentice Hall, 1976.
[18] A.S. Glassner, *Principles of Digital Image Synthesis.* Morgan Kaufmann, 1995.
[19] M.A. Halstead, B.A. Barsky, S.A. Klein, and R.B. Mandell, "Reconstructing Curved Surfaces from Specular Reflection Patterns Using Spline Surface Fitting of Normals," *Proc. ACM SIGGRAPH,* pp. 335-342, 1996.
[20] F.H. Harlow and J.E. Welch, "Numerical Calculation of Time-Dependent Viscous Incompressible Flow," *Physics of Fluids,* vol. 8, pp. 2182-2189, 1965.
[21] C. Harris and M. Stephens, "A Combined Edge and Corner Detector," *Proc. Fourth Alvey Vision Conf.,* pp. 189-192, 1988.
[22] R.I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision.* Cambridge Univ. Press, 2000.
[23] R.I. Hartley and P. Sturm, "Triangulation," *Computer Vision and Image Understanding,* vol. 68, no. 2, pp. 146-157, 1997.
[24] J. Höhle, "Reconstruction of the Underwater Object," *Proc. Photogrammetric Eng.,* pp. 948-954, 1971.
[25] M.B. Hullin, M. Fuchs, I. Ihrke, H.-P. Seidel, and H.P.A. Lensch, "Fluorescent Immersion Range Scanning," *ACM Trans. Graphics,* vol. 27, no. 3, pp. 87-1-87-10, Aug. 2008.
[26] K. Ikeuchi, "Determining the Surface Orientations of Specular Surfaces by Using the Photometric Stereo Method," *IEEE Trans. Pattern Analysis Machine Intelligence,* vol. 3, no. 6, pp. 661-669, Nov. 1981.
[27] K. Ikeuchi and B.K.P. Horn, "Numerical Shape from Shading and Occluding Boundaries," *Artificial Intelligence,* vol. 17, pp. 141-184, 1981.
[28] B. Jähne, J. Klinke, P. Geissler, and F. Hering, "Image Sequence Analysis of Ocean Wind Waves," *Proc. Int'l Seminar on Imaging in Transport Processes,* 1992.
[29] B. Jähne, J. Klinke, and S. Waas, "Imaging of Short Ocean Wind Waves: A Critical Theoretical Review," *J. Optical Soc. Am. A,* vol. 11, no. 8, pp. 2197-2209, 1994.
[30] B. Jähne, M. Schmidt, and R. Rocholz, "Combined Optical Slope/Height Measurements of Short Wind Waves: Principle and Calibration," *Measurement Science and Technology,* vol. 16, no. 10, pp. 1937-1944, 2005.
[31] J. Kaminski, S. Lowitzsch, M.C. Knauer, and G. Häusler, "Full-Field Shape Measurement of Specular Surfaces," *Proc. Fifth Int'l Workshop Automatic Processing of Fringe Patterns,* 2005.
[32] W.C. Keller and B.L. Gotwols, "Two-Dimensional Optical Measurement of Wave Slope," *Applied Optics,* vol. 22, no. 22, pp. 3476-3478, 1983.
[33] J.J. Koenderink and A.J. van Doorn, "The Structure of Two-Dimensional Scalar Fields with Applications to Vision," *Biological Cybernetics,* vol. 33, pp. 151-158, 1979.
[34] K. Kolev, M. Klodt, T. Brox, and D. Cremers, "Continuous Global Optimization in Multiview 3d Reconstruction," *Int'l J. Computer Vision,* vol. 84, no. 1, pp. 80-96, 2009.
[35] K.N. Kutulakos and E. Steger, "A Theory of and Refractive and Specular 3d Shape by Light-Path Triangulation," *Proc. 10th Int'l Conf. Computer Vision,* pp. 1448-1455, 2005.
[36] F. Losasso, F. Gibou, and R. Fedkiw, "Simulating Water and Smoke with an Octree Data Structure," *ACM Trans. Graphics,* vol. 23, no. 3, pp. 457-462, 2004.
[37] H.-G. Maas, "New Developments in Multimedia Photogrammetry," *Optical 3D Measurement Techniques III.* Wichmann Verlag, 1995.

[38] N. Morris and K.N. Kutulakos, "Dynamic Refraction Stereo," *Proc. 10th Int'l Conf. Computer Vision,* pp. 1573-1580, 2005.

[39] M. Müller, D. Charypar, and M. Gross, "Particle-Based Fluid Simulation for Interactive Applications," *Proc. 2003 ACM SIG-GRAPH/Eurographics Symp. Computer Animation,* pp. 154-159, 2003.

[40] H. Murase, "Surface Shape Reconstruction of an Undulating Transparent Object," *Proc. Third Int'l Conf. Computer Vision,* pp. 313-317, 1990.

[41] D. Nehab, S. Rusinkiewicz, J. Davis, and R. Ramamoorthi, "Efficiently Combining Positions and Normals for Precise 3D Geometry," *Proc. ACM SIGGRAPH,* pp. 536-543, 2005.

[42] A. Okamoto, "Orientation Problem of Two-Media Photographs with Curved Boundary Surfaces," *Proc. Photogrammetric Eng. and Remote Sensing,* pp. 303-316, 1984.

[43] M. Oren and S.K. Nayar, "A Theory of Specular Surface Geometry," *Proc. Fifth Int'l Conf. Computer Vision,* pp. 740-747, 1995.

[44] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, *Numerical Recipies in C.* Cambridge Univ. Press, 1988.

[45] A.C. Sanderson, L.E. Weiss, and S.K. Nayar, "Structured High-light Inspection of Specular Surfaces," *IEEE Trans. Pattern Analysis Machine Intelligence,* vol. 10, no. 1, pp. 44-55, Jan. 1988.

[46] S. Savarese and P. Perona, "Local Analysis for 3D Reconstruction of Specular Surfaces—Part II," *Proc. Seventh European Conf. Computer Vision,* pp. 759-774, 2002.

[47] H. Schultz, "Retrieving Shape Information from Multiple Images of a Specular Surface," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 16, no. 2, pp. 195-201, Feb. 1994.

[48] O.H. Shemdin, "Measurement of Short Surface Waves with Stereophotography," *Proc. Eng. in the Ocean Environment Conf.,* pp. 568-571, 1990.

[49] S.A. Sullivan, "Experimental Study of the Absorption in Distilled Water, Artificial Sea Water, and Heavy Water in the Visible Region of the Spectrum," *J. Optical Soc. Am.,* vol. 53, pp. 962-968, 1963.

[50] M. Tarini, H.P.A. Lensch, M. Goesele, and H.-P. Seidel, "3D Acquisition of Mirroring Objects," Technical Report MPI-I-2003-4-001, Max Planck Institut für Informatik, 2003.

[51] T. Treibitz, Y.Y. Schechner, and H. Singh, "Flat Refractive Geometry," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 1-8, 2008.

[52] B. Trifonov, D. Bradley, and W. Heidrich, "Tomographic Reconstruction of Transparent Objects," *Proc. 17th Eurographics Symp. Rendering,* pp. 51-60, 2006.

[53] B. Triggs, P.F. McLauchlan, R.I. Hartley, and A.W. Fitzgibbon, "Bundle Adjustment—A Modern Synthesis," *Proc. Int'l Workshop Vision Algorithms,* pp. 298-372, 2000.

[54] Y. Vasilyev, Y. Adato, T. Zickler, and O. Ben-Shahar, "Dense Specular Shape from Multiple Specular Flows," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 1-8, 2008.

[55] J. Wang and K.J. Dana, "Relief Texture from Specularities," *IEEE Trans. Pattern Analysis Machine Intelligence,* vol. 28, no. 3, pp. 446-457, Mar. 2006.

[56] Z. Wu and G.A. Meadows, "2-D Surface Reconstruction of Water Waves," *Proc. Eng. in the Ocean Environment. Conf.,* pp. 416-421, 1990.

[57] L. Zhang, N. Snavely, B. Curless, and S.M. Seitz, "Spacetime Faces: High Resolution Capture for Modeling and Animation," *ACM Trans. Graphics,* vol. 23, no. 3, pp. 548-558, 2004.

[58] X. Zhang and C.S. Cox, "Measuring the Two-Dimensional Structure of a Wavy Water Surface Optically: A Surface Gradient Detector," *Experiments in Fluids,* vol. 17, pp. 225-237, 1994.

[59] L. Zhou, C. Kambhamettu, and D.B. Goldgof, "Fluid Structure and Motion Analysis from Multi-Spectrum 2D Cloud Image Sequences," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* vol. 2, pp. 744-751, 2000.

Nigel J.W. Morris received the BScH degree in computing and information science from Queen's University, Canada, in 2002, and the MSc and PhD degrees in computer science from the University of Toronto in 2004 and 2010, respectively. He has worked at Mitsubishi Electric Research Labs and is currently lead computer scientist at Morgan Solar Inc.

Kiriakos N. Kutulakos received the BA degree in computer science from the University of Crete, Greece, in 1988, and the MS and PhD degrees in computer science from the University of Wisconsin-Madison in 1990 and 1994, respectively. Following his dissertation work, he joined the University of Rochester, where he was a US National Science Foundation (NSF) postdoctoral fellow and later an assistant professor until 2001. He is currently a professor of computer science at the University of Toronto. He won the Best Student Paper Award at CVPR '94, the David Marr Prize in 1999, a David Marr Prize Honorable Mention in 2005, and a Best Paper Honorable Mention at ECCV '06. He is the recipient of a CAREER award from the US National Science Foundation, a Premier's Research Excellence Award from the government of Ontario, and an Alfred P. Sloan Research Fellowship. He was an associate editor of the *IEEE Transactions on Pattern Analysis and Machine Intelligence* from 2005 to 2010, a program cochair of CVPR '03 and ICCP '10, and will serve as program cochair of ICCV '13. He is a member of the IEEE.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.