

Refractive Height Fields from Single and Multiple Images

Qi Shan

University of Washington

shanqi@cs.washington.edu

Sameer Agarwal

Google Inc.

sameeragarwal@google.com

Brian Curless

University of Washington

curless@cs.washington.edu

Abstract

We propose a novel framework for reconstructing homogenous, transparent, refractive height-fields from a single viewpoint. The height-field is imaged against a known planar background, or sequence of backgrounds. Unlike existing approaches that do a point-by-point reconstruction – which is known to have intractable ambiguities – our method estimates and optimizes for the entire height-field at the same time. The formulation supports shape recovery from measured distortions (deflections) or directly from the images themselves, including from a single image. We report results for a variety of refractive height-fields showing significant improvement over prior art.

1. Introduction

Recovering surface geometry from images is a long-standing problem in computer vision. Transparent, refractive surfaces are among the most challenging to recover, in large part due to the non-linearity inherent in refraction.

An important class of refractive surfaces are height fields comprised of a single refractive material, which arise when modeling liquids [7], as well as man-made objects, such as many “obscure glasses” used to separate spaces with a distorting visual barrier [10]. We focus on reconstructing these surfaces, in particular differentiable height fields, imaged against a known background (or set of backgrounds) from a single viewpoint.

This problem, originally studied in computer vision by Murase [7], remains only partially solved. In that formulation, and others that have followed, the refractive surface is imaged against a planar background. Per-pixel deflections are estimated and then used to recover surface normals, which can in turn be integrated to yield a height field. However, it is not possible to accurately estimate normals without knowing the depth in the first place [6]. Thus, accurate surface reconstruction in this setting has remained elusive without resorting to multiple viewpoints or unusual background emitters.

In this paper, we address the fundamental shortcoming

of estimating normals and then the height field by instead formulating the problem directly in terms of the height field itself. As is common, we assume no interreflections (and for the direct image approach, minimal impact from Fresnel attenuation), and we also assume an orthographic camera. In our formulation, normals are simply derived properties, giving rise to coupled PDEs over the height field. Our formulation supports estimating the geometry from per-pixel deflections (“deflection maps”) or directly from the image measurements through a global optimization over the height field. In the deflection-map formulation, we derive a novel isocontour regularization term based on co-planarity of rays to constrain the optimization. We demonstrate that our approach is measurably more accurate than the normal-based approach with comparisons to ground truth laser scans. We present results derived from high quality deflection maps captured using multiple backgrounds. Finally, we show detailed object reconstruction from a single image of a known background by optimizing to match observed pixel values.

Our primary contributions are (1) a framework based on differentiable height fields that overcomes the pixel-wise, height-normal ambiguity and (2) a method for optimizing each height field from a single image rather than a deflection map. This second contribution is important because it avoids dependence on precisely recovering a deflection vector (two variables) per pixel, instead focusing on recovering the height field (one variable per pixel) that indirectly gives rise to deflections and explains the image measurement.

In the remainder of the paper, we discuss related work (Sec. 2), and then provide the background formulation for our problem (Sec. 3). We describe objectives and optimization strategies for recovering shape from a deflection map (Sec. 4) and directly from an image (Sec. 5). We then describe experimental procedures and results (Sec. 6) and conclude with discussion of limitations and future work (Sec. 7).

2. Related Work

Shape from refraction has been an active area of research. See [4] for an excellent survey of this area, as well as shape from reflection. Several efforts have focused on

height field surfaces, with the common assumption that interreflection is negligible. Murase’s pioneering work [7] on reconstructing water surfaces uses optical flow between a known background and each image to recover approximate per-pixel normals. Normal recovery is based on a given average surface height and index of refraction, from which the height field may be recovered by integration. As noted in the previous section, the normals cannot be recovered accurately without first knowing the per-pixel heights, a significant limitation of this approach. More recently, Tian and Narashimhan [13] recover time-varying distortions of an unknown background beneath a water tank, based on the behavior of water waves. They demonstrate surface reconstructions as well, but these are similarly based on normal estimation and integration. Morris and Kutulakos [6] overcome the height-normal ambiguity by using two views which are jointly matched to each other and a known background and then estimating height and normal per pixel, as well as the index of refraction. Most recently, Wetstein et al. [15] employ a novel background that includes a lenslet array to encode directions as well as vertical (1D) positions of emitted light rays, enabling direct recovery of normals and a sparse set of heights. This encoding implies a spatial vs. directional resolution trade-off, but enables single view recovery. The surface is reconstructed from integration of normals guided by sparsely recovered depths.

Refractive height fields fall into the category of single-interface refractive surfaces discussed by Kutulakos and Steger [5]. They prove that pointwise estimation of surface position is not possible without using two views or measuring the refracted rays. However, we show in this paper that refractive surface reconstruction is in fact possible from a single view without encoding refracted ray directions if we move from pointwise estimation to integrable height field surface estimation.

Related to shape from refraction is shape from specular reflection. We note that this problem is somewhat easier, as reflection does not have the non-linearity introduced by refraction, though still fundamentally difficult. Tarini et al. [12] recover surfaces from reflection to a set of known backgrounds and estimate shape from recovered distortions, applying an alternating optimization over height fields and normals. We instead optimize for the height field directly, in the refractive setting with unknown index of refraction, and demonstrate single-image recovery. Shape recovery from specular flow has also gained increased attention in recent years [9, 1, 2, 14]. We note that Vasilyev et al. [14] in particular have begun exploring the use of surface integrability in shape recovery from specular flow induced by motion of an unknown, distant environment, demonstrating synthetic results for this scenario.

For a more complete view of shape from refraction and reflection that includes non-height-field surfaces, we again

refer the reader to Ihrke’s survey [4].

3. Background Formulation

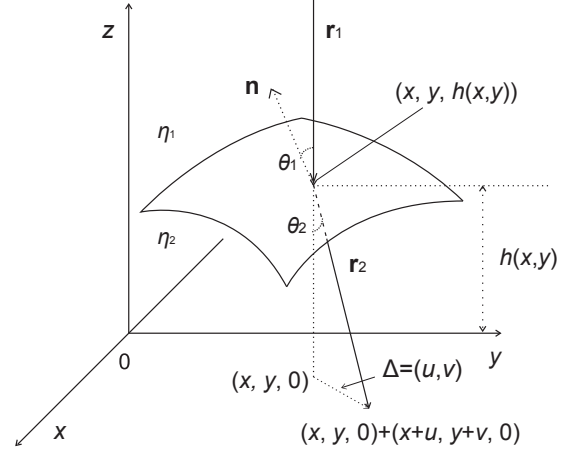


Figure 1. Geometry of shape from height field refraction.

In this section, we develop analytical expressions that describe the refraction of the camera rays from an orthographic camera looking down upon a refractive surface. These expressions will be used in the surface reconstruction approaches described in the sections that follow. Fig. 1 illustrates the geometric configuration.

The refracting surface is assumed to be a height field $[x, y, h(x, y)]$ that sits upon a background image $B(x, y)$ placed on the xy -plane $z = 0$. We assume no interreflections, that each camera ray gets refracted exactly once before hitting the background image. We also assume that the surface is G^1 continuous.

The image plane of the orthographic camera is placed at some height z_∞ . We will denote the observed image by $I(x, y)$. The observer is in a medium with refractive index η_1 , and the refracting surface h is made up of a material with an constant but *unknown* refractive index η_2 . We will assume that $\eta_2 > \eta_1$; i.e., the refracting surface is denser. We assume that η_1 is known, typically the index of refraction of air.

For every point $[x, y, z_\infty]$ in the observed image, there is a camera ray $[x, y, z_\infty] + \lambda[0, 0, -1]$ that gets refracted to the point $[x + u(x, y), y + v(x, y), 0]$ in the background image. We will refer to $\Delta(x, y) = [u(x, y), v(x, y)]$ as the deflection vector. Let us derive an analytic expression for $\Delta(x, y)$. Since we will be considering the same point (x, y) throughout, in the interest of brevity we will drop the explicit dependence on x and y from various quantities.

For a point $[x, y, h]$ on the refracting surface, the normal vector is given by

$$\mathbf{n} = (h_x^2 + h_y^2 + 1)^{-1/2} [-h_x, -h_y, 1] \quad (1)$$

where, $h_x = \partial_x h$ and $h_y = \partial_y h$. Further, let

$$\rho = (h_x^2 + h_y^2 + 1)^{-1/2} \quad (2)$$

which gives us $\mathbf{n} = [-\rho h_x, -\rho h_y, \rho]$.

If \mathbf{r}_1 is the direction of the incoming ray, \mathbf{n} is the normal at the point of refraction and \mathbf{r}_2 is the direction of the refracted ray, then Snell's law tells us that

$$\eta_2 \mathbf{r}_2 = \eta_1 \mathbf{r}_1 + (\eta_1 \cos \theta_1 - \eta_2 \cos \theta_2) \mathbf{n} \quad (3)$$

Where, θ_1 and θ_2 are the angles between the normal direction and the incident and refracted rays respectively. In our setting,

$$\mathbf{r}_1 = [0, 0, -1] \quad (4)$$

$$\cos \theta_1 = -\mathbf{n}^\top \mathbf{r}_1 = \rho \quad (5)$$

$$\cos \theta_2 = \sqrt{1 - \left(\frac{\eta_1}{\eta_2}\right)^2 (1 - \rho^2)} \quad (6)$$

which gives us

$$\eta_1 \cos \theta_1 - \eta_2 \cos \theta_2 = \eta_1 \rho - \sqrt{\eta_2^2 - \eta_1^2 (1 - \rho^2)} \quad (7)$$

Let,

$$\xi = \rho \left(\rho - \sqrt{\eta^2 - 1 + \rho^2} \right), \quad (8)$$

where $\eta = \eta_2/\eta_1$ is the relative index of refraction. Then, after some straightforward substitutions, we have

$$\eta \mathbf{r}_2 = [-\xi h_x, -\xi h_y, \xi - 1] \quad (9)$$

Starting from (x, y, h) the refracted ray along the direction \mathbf{r}_2 and parameterized by distance along the ray, μ , intersects the plane $z = 0$ at $[x + u, y + v, 0]$, i.e.

$$[x, y, h] + \frac{\mu}{\eta} [-\xi h_x, -\xi h_y, \xi - 1] = [x + u, y + v, 0] \quad (10)$$

Eliminating μ from this system of equations gives us

$$\Delta = [u, v] = \frac{\xi h}{\xi - 1} [h_x, h_y] \quad (11)$$

Observe that

Lemma 1. $\xi < 1$

Proof. Suppose not. Then

$$1 \leq \rho \left(\rho - \sqrt{\eta^2 - (1 - \rho^2)} \right) \quad (12)$$

$$\frac{1}{\rho} - \rho \leq -\sqrt{\eta^2 - (1 - \rho^2)} \quad (13)$$

Since $0 < \rho \leq 1$, the left hand side is always non-negative. The only way this inequality could be satisfied is if both sides are zero. This would happen with $\rho = 1$ and $\eta = 0$, but the latter is not possible. \square

Thus (11) is always well defined. It is now straightforward to see that the following two lemma are true.

Lemma 2. $\eta_1 = \eta_2 \rightarrow \eta = 0 \rightarrow \Delta(x, y) = 0$

Lemma 3. $h(x, y) = c \rightarrow \Delta(x, y) = 0$.

It is also worth observing here that methods that start from $\Delta(x, y)$ and solve for depths and normals independently have a fundamental ambiguity [6]. Since the normal to the surface depends only on h_x and h_y , treating the normal independent of the surface height is equivalent to treating h_x and h_y as variables independent of h . Doing so would turn (11) from a system of non-linear PDEs to non-linear algebraic equations in three unknowns, which in general will have an infinite number of solutions.

4. Reconstruction from a Deflection Map

From (11) we have a coupled system of non-linear partial differential equations. Given a deflection map Δ , one way to solve for $h(x, y)$ would be to minimize the following functional (a summation over image pixels):

$$\sum_{x,y} \left(u - \frac{\xi h h_x}{\xi - 1} \right)^2 + \left(v - \frac{\xi h h_y}{\xi - 1} \right)^2 \quad (14)$$

We expect the input deflection map to be corrupted by errors and noise. To counter this, we introduce a novel regularization term based on a characterization of the isocontours of h . The following lemma gives us a constructive method for identifying the isocontours of h .

Lemma 4. Let $[x_0, y_0]$ be an arbitrary point in the domain and if $\Delta(\mathbf{x}, \mathbf{y}) = [\mathbf{u}, \mathbf{v}]$ is a noise free deflection map. Then the curve

$$x(t) = x_0 + \int_0^t v(x(s), y(s)) ds \quad (15)$$

$$y(t) = y_0 - \int_0^t u(x(s), y(s)) ds \quad (16)$$

is an isocontour of h .

Proof.

$$\frac{d}{dt} h(x(t), y(t)) = h_x \frac{d}{dt} x(t) + h_y \frac{d}{dt} y(t) \quad (17)$$

$$= h_x v - h_y u = \frac{\xi - \eta_1}{\xi h} (uv - vu) \quad (18)$$

$$= 0. \quad \square \quad (19)$$

The key fact used in this proof is that the deflection vector at each point $[x, y]$ is parallel to the gradient of h at that point, i.e., $h_x v - h_y u = 0$. A more geometrical way to see this is to observe that the gradient ∇h is collinear with the

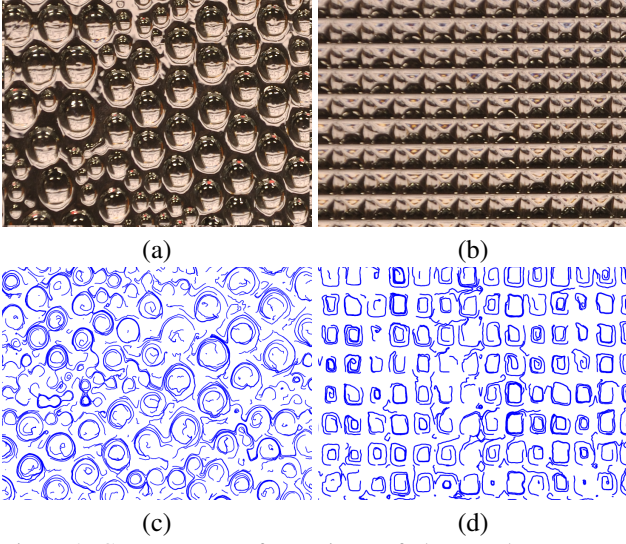


Figure 2. Contour maps of two pieces of obscure glass. We operate on a 768×512 image grid. (a) and (b) show the images of the glass examples under indoor illumination; both are flat on one side, textured on the other. (c) and (d) show their contour maps constructed from experimentally recovered deflection maps. Here we only show one out of every 100 contour curves.

projection of the normal to the surface onto the xy -plane. Similarly the deflection vector $\Delta(x, y)$ is aligned with the projection of the refracted ray. By Snell's law, the normal and the refracted ray are constrained to be co-planar with the viewing ray, thus their projections on the xy -plane must be collinear. As a result, the direction perpendicular to the deflection direction is also perpendicular to the height gradient and thus is tangent to a height isocontour. Fig. 2 shows sparse contour curves of two pieces of obscure glass.

Thus, if $\Omega(x, y)$ is the set of points on the isocontour that starts from the point $[x, y]$, we can define a regularization

$$\phi(h) = \sum_{x,y} \sum_{x',y' \in \Omega(x,y)} (h - h')^2 \quad (20)$$

Here, $h' = h(x', y')$. Combining (14) and (20) we get the energy functional

$$E(h, \eta_2) = \sum_{x,y} \left(u - \frac{\xi h h_x}{\xi - \eta_1} \right)^2 + \left(v - \frac{\xi h h_y}{\xi - \eta_1} \right)^2 + w \sum_{x,y} \sum_{x',y' \in \Omega(x,y)} (h - h')^2 \quad (21)$$

where w is a regularization constant. We use $w = 100$ in our experiments.

4.1. Optimizing from a deflection map

Even if η_2 is known, minimizing the energy defined in Eq. (21) is a complicated nonlinear optimization problem, and some care is needed to avoid trivial local minima.

We use a multi-scale optimization approach where, at each scale, the optimization proceeds in two stages, similar to the work of [11]. (We initialize with white noise at the coarsest scale.) The advantage of the optimization process is three-fold. The multi-scale scheme avoids poor local minima. The two-stage approach, and relaxation scheme in Eq (23), accelerate convergence by breaking the non-linear optimization into sequences of convex optimizations, a common approach to such problems. The isocontour term adds a physics-based constraint, reducing sensitivity to noise and giving somewhat smoother results.

In the first stage, we lift the problem to a higher dimensional space, and reduce its non-linearity by relaxing the constraint

$$h_x = \partial_x h, \quad h_y = \partial_y h. \quad (22)$$

We do this by treating h_x and h_y as independent variables in their own right and adding a term to the optimization problem that penalizes deviations from the true gradients, giving us

$$E(h, h_x, h_y, \eta_2) = \sum_{x,y} \left(u - \frac{\xi h h_x}{\xi - \eta_1} \right)^2 + \left(v - \frac{\xi h h_y}{\xi - \eta_1} \right)^2 + w \sum_{x,y} \sum_{x',y' \in \Omega(x,y)} (h - h')^2 + \lambda \sum_{x,y} (h_x - \partial_x h)^2 + (h_y - \partial_y h)^2 \quad (23)$$

Here, λ is a constant whose value is increased by a factor of 2 at the end of each iteration. We now solve (23) using coordinate descent by optimizing over h , h_x and h_y independently.

At the end of this first stage we have a rough estimate of the surface. In the second stage, we use this rough estimate to initialize the full optimization of (21), which we solve using gradient descent, with an Armijo-Goldstein line search.

As the refractive index η_2 is unknown, we optimize for possible height fields using a range of values for η_2 (typically $\eta_2 \in [1.4, 1.8]$ in increments of 0.02 for glass and plastics) and pick the value of η_2 that gives the smallest residual error. As an acceleration, we estimate the refractive index by performing reconstructions at lower resolution with fewer optimization steps. In our implementation, we run the refractive index estimation on a lower resolution image with 1/4 of the original size, and the number of gradient descent steps are 1/10 of the number used for full reconstruction.

4.1.1 Computing isocontours from the deflection map

Considering a point (x_0, y_0) , the trajectory of the isocontour at that point is $(v(x_0, y_0), -u(x_0, y_0))$, according to

Eq. (15) and (16). By following the trajectory, the next point is

$$(x_1, y_1) = (x_0, y_0) + \frac{(v(x_0, y_0), -u(x_0, y_0))}{\sqrt{v^2(x_0, y_0) + u^2(x_0, y_0)}} s, \quad (24)$$

where $s = 0.01$ is a pre-defined distance between (x_1, y_1) and (x_0, y_0) . We then bilinearly interpolate the trajectory at (x_1, y_1) from the deflection vectors on the four nearest grid points. The process continues for 1000 steps, resulting in a contour of length 10. A longer contour provides smoothness constraints between farther away points and suffers from error accumulated during the curve tracking process. In Eq. (20), $\Omega(x, y)$ is a collection of points on the contour curve. In practice we define $\Omega(x, y)$ to be a set of pixels close enough to the contour curve. The threshold is defined to be 0.1 of pixel width in our experiments.

5. Reconstruction from an Image

In this section, we describe the formulation for optimizing the height field directly from a single image. Given (11), the observed image $I(x, y)$ can be expressed in terms of the background image $B(x, y)$ as

$$I(x, y) = F(h_x, h_y, \eta_1, \eta_2) B \left(x + \frac{\xi h h_x}{\xi - 1}, y + \frac{\xi h h_y}{\xi - 1} \right) \quad (25)$$

where F accounts for Fresnel attenuation of the refracted ray and B is assumed to emit light equally in all directions. In practice, F is close to one for a wide range of surface normals; thus, we currently drop this term.

We can now pose a new optimization problem written purely in terms of the two images and the refracting surface h without introducing an intermediate deflection map Δ .

$$\sum_{x, y} \left\| I(x, y) - B \left(x + \frac{\xi h h_x}{\xi - 1}, y + \frac{\xi h h_y}{\xi - 1} \right) \right\|^2 \quad (26)$$

This optimization problem opens up the possibility of recovering the shape of a refracting surface from a single image of a known background, for which dense, accurate deflection estimation can be challenging.

5.1. Optimizing from an image

To minimize the energy defined in (26), we adopt a similar multi-scale scheme as we did for minimizing (21). In each scale, we perform gradient descent, with an Armijo-Goldstein line search, to optimize h, h_x, h_y simultaneously.

Reasonable initialization of this optimization is important, as the connection between surface geometry and image observations is far less direct than it is for the surface geometry and deflections. As will be discussed in Section 6.2,

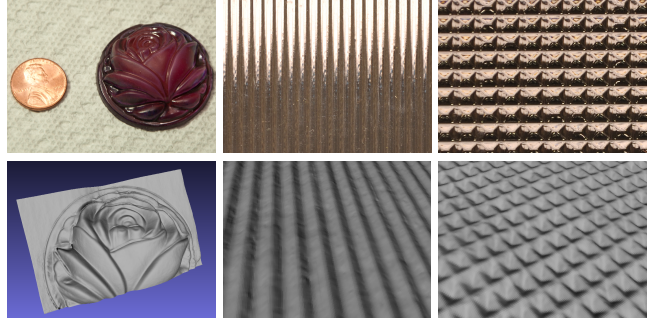


Figure 4. Photos of refractive objects and corresponding reconstructions using our deflection-based method. The flower object was painted for the photograph to better visualize its shape.

with suitable choice of background, we can get a rough estimate of deflections from a single image, from which we can recover an initial surface. This surface is then downsampled and used to initialize the coarsest scale of the image-based optimization.

6. Experiments and Evaluations

In this section, we demonstrate experimental results for reconstructing refractive height fields from deflection maps and directly from images.

Across all experiments we use a Canon DSLR camera with a 400mm lens to approximate orthography. The objects we reconstruct are made of glass or plastic of unknown index of refraction. Each is flat on one side with no overhanging surfaces on the opposite side, i.e., can be modeled as a height field. We place the flat side directly against the background.

6.1. Reconstruction from deflection maps

We first evaluate reconstruction from high quality deflection maps. Environment matting [3] can be used to recover such deflection maps; we employ the variant described by Shan et al. [10]. The procedure consists of displaying a sequence of colored Perlin noise [8] patterns on an LCD monitor and photographing the monitor both with and without the object. We recover per-pixel kernels and simply use the centroid of each kernel to set the deflections.

We compare our deflection-based technique with our implementation of Murase’s algorithm [7]. Murase’s method requires a known index of refraction and average height. As we have neither a priori, we first run our recovery algorithm and then provide Murase’s method with our recovered index and average height. Fig. 3 shows a surface reconstruction result of transparent dome glass. We show this example first due to its simple and regular geometry.

We also compare the results to ground truth models. We painted several objects and scanned them with a Cyberware Model 15 laser scanner to obtain reference shapes. We then

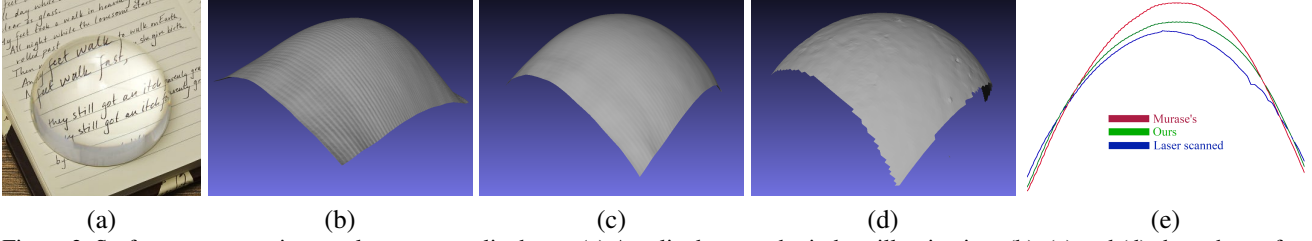


Figure 3. Surface reconstruction results on an acrylic dome. (a) Acrylic dome under indoor illumination. (b), (c) and (d) show the surfaces from Murase’s algorithm [7], our deflection based method, and a laser scanner. (e) shows a height profile extracted as a horizontal slice through the center of the height fields.

align our surfaces and Murase’s surfaces to the scanned models using a standard ICP algorithm and then compute RMS errors. For this example, we used objects that are relatively easy for Murase’s method as they are low relief. As shown in Fig. 5, the recovered shapes appear somewhat similar, but our method demonstrates measurable improvement as the height profiles reveal. Numerically, the mean absolute distance errors of our surfaces in Fig. 5(b) and (c) are 1.4 mm and 0.80 mm, respectively, an improvement over Murase’s result of 1.9 mm and 1.3 mm.

Our method demonstrates significant improvement on objects with larger height variation. One example is the dome glass shown in Fig. 3. Our surface has a absolute distance error of 0.51 mm, improving over Murase’s result of 1.12 mm.

We demonstrate additional reconstructions using our method in Figure 4.

6.2. Reconstruction from a single image

Here we describe our choice of background pattern and results for height field reconstruction from a single image against a known background using the image-based formulation.

To recover a height field from a single image taken against a known background using our image-based formulation (Section 5), we take some care in choosing the experimental setup and background. First, we require a background that emits light isotropically. LCD monitors exhibit significant directional variation, especially vertically, and were not suitable for these experiments. Instead we employed a light box, commonly available in art supply stores, that is comprised of several bright bulbs under a diffusing surface. We then printed a pattern onto photo paper and placed it over the light box; though the paper absorbed light and thus required longer exposure times (0.5 seconds at F/24, ISO 100 in our experiments), it served the additional purpose of more evenly diffusing the light.

Second, we must choose a background pattern. As noted in Section 5, we initialize the image-based optimization using the results of the deflection-based optimization. Thus, we desire a pattern from which approximate deflections

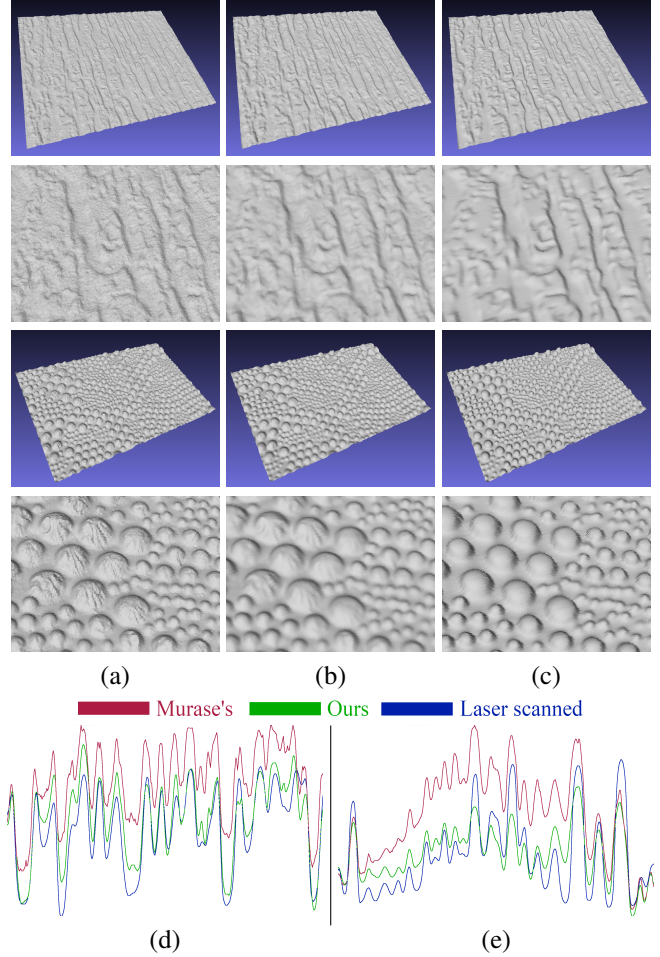


Figure 5. Quantitative evaluation with laser scanned surfaces. Column (a), (b) and (c) show surfaces from Murase’s algorithm, our deflection map based algorithm, and laser scanner, respectively, along with their close-up views. (d) and (e) show a height profile, each extracted as a horizontal slice through the center of the height fields.

may be recovered. Chuang et al. [3] employ color ramps to recover deflections from a single image in environment matting. These rough deflections may be suitable as a starting point, but we also desire some structure for the image-

based approach to match during optimization. We take a hybrid approach. We create a horizontal ramp from 0 to 255 in the red channel, a similar vertical ramp in the blue channel. To provide structure for image-based optimization, we put a high-pass filtered Perlin noise [8] pattern in the green channel.

In an ideal situation, we would be able to recover a rough deflection map from the red and blue channels, and then proceed with the two-step recovery, using the full image for image-based optimization. In practice, the red and blue ramps are piecewise constant when spread over a large area; thus, we compose a pattern by repeating these ramps over smaller areas, as shown in the left column of Figure 6. Further, the camera does not observe these color channels directly; a shade of red, for instance, will appear as a mix of red and green in the camera. To reduce this color cross-talk effect, we compute a color look-up table that maps the observed colors to the idealized colors and use the look-ups to approximately decouple the channels, generating a “color-corrected” image. Note that we only use the color-corrected image for locating local windows. The closest color matching is performed on the non-color-corrected image pair.

Finally, to compute a rough per-pixel deflection estimation, we use the red and blue values of a pixel to locate a local window (11×11) on the no-glass background. In our experiments, we assume that the glass is relatively thin and do not search over all squares with color ramps but choose the closest deflection. This is not a theoretical limitation, but suggests an area of future work, e.g., in pattern design.

Note that we photograph the background pattern only once in advance for calibration and to provide the known background pattern. After calibration, for each glass example, only one image is needed to reconstruct its surface.

We show a variety of reconstructions in Fig. 6. The deflection maps are quite noisy and generally do not provide strong enough cues for either Murase’s method or the deflection-based optimization. However, when then optimizing directly against the images, the structure and relief clearly come forward. Note in particular the intricate details revealed in the complex flower relief.

Running time We solve for the height field by solving a large non-linear optimization problem. The solver runs on a workstation with two Intel Xeon X5680 (3.33GHz) processors each with 6 cores. For the deflection-based method, it takes about 3 hours for a height field of 600×400 . The image-based approach takes about 5 hours for a 600×400 height field.

7. Conclusion

We have presented a method for reconstructing refractive height fields imaged against a known background or

set of backgrounds from a single viewpoint. Our formulation overcomes the standard height-normal ambiguity by optimizing for the height field itself. Typical reconstruction algorithms for this problem start from a deflection map. We develop an optimization framework, with a novel contour constraint, that can recover high quality surfaces from high quality deflection maps and demonstrate improvement over recovering normals and then integrating to get the height field. Moreover, we show that it is possible to optimize the height map by comparing a single image against the known background. This makes it possible to recover detailed shape even when deflection map estimation is difficult.

Our method has notable limitations and thus a number of areas for future work. We have not shown the conditions under which a unique solution exists, an area for future work. One extreme example is the case that all deflection values are zero, the surface can be a flat surface of any (constant) height. The method, while more accurate than the normal-based method, is computationally expensive. We have started exploring optimization strategies to accelerate the reconstruction. Our image-based formulation also neglects Fresnel attenuation as well as significant filtering due to demagnification of the background, both of which we hope to model in the future. We also hope to lift dependence on an orthographic viewer. Our experimental setup currently requires long exposures to recover surfaces; in the single image setting this limitation precludes recovering dynamic surfaces such as water. In addition, we currently use a repeating ramp pattern within the background that can cause ambiguities that are difficult to overcome during optimization. We plan to explore better background patterns and brighter sources to overcome these limitations.

Acknowledgments

This work was supported by funding from National Science Foundation grant IIS-0963657, Google, Adobe, Microsoft Research, Pixar, and University of Washington Animation Research Labs.

References

- [1] Y. Adato, Y. Vasilyev, O. Ben-Shahar, and T. Zickler. Towards a theory of shape from specular flow. In *ICCV*, 2007.
- [2] G. D. Canas, Y. Vasilyev, Y. Adato, T. Zickler, S. Gortler, and O. Ben-Shahar. A linear formulation of shape from specular flow. In *ICCV*, 2009.
- [3] Y.-Y. Chuang, D. E. Zongker, J. Hindorff, B. Curless, D. Salesin, and R. Szeliski. Environment matting extensions: towards higher accuracy and real-time capture. In *SIGGRAPH*, pages 121–130, 2000.
- [4] I. Ihrke, K. N. Kutulakos, H. P. A. Lensch, M. Magnor, and W. Heidrich. State of the art in transparent and specular ob-

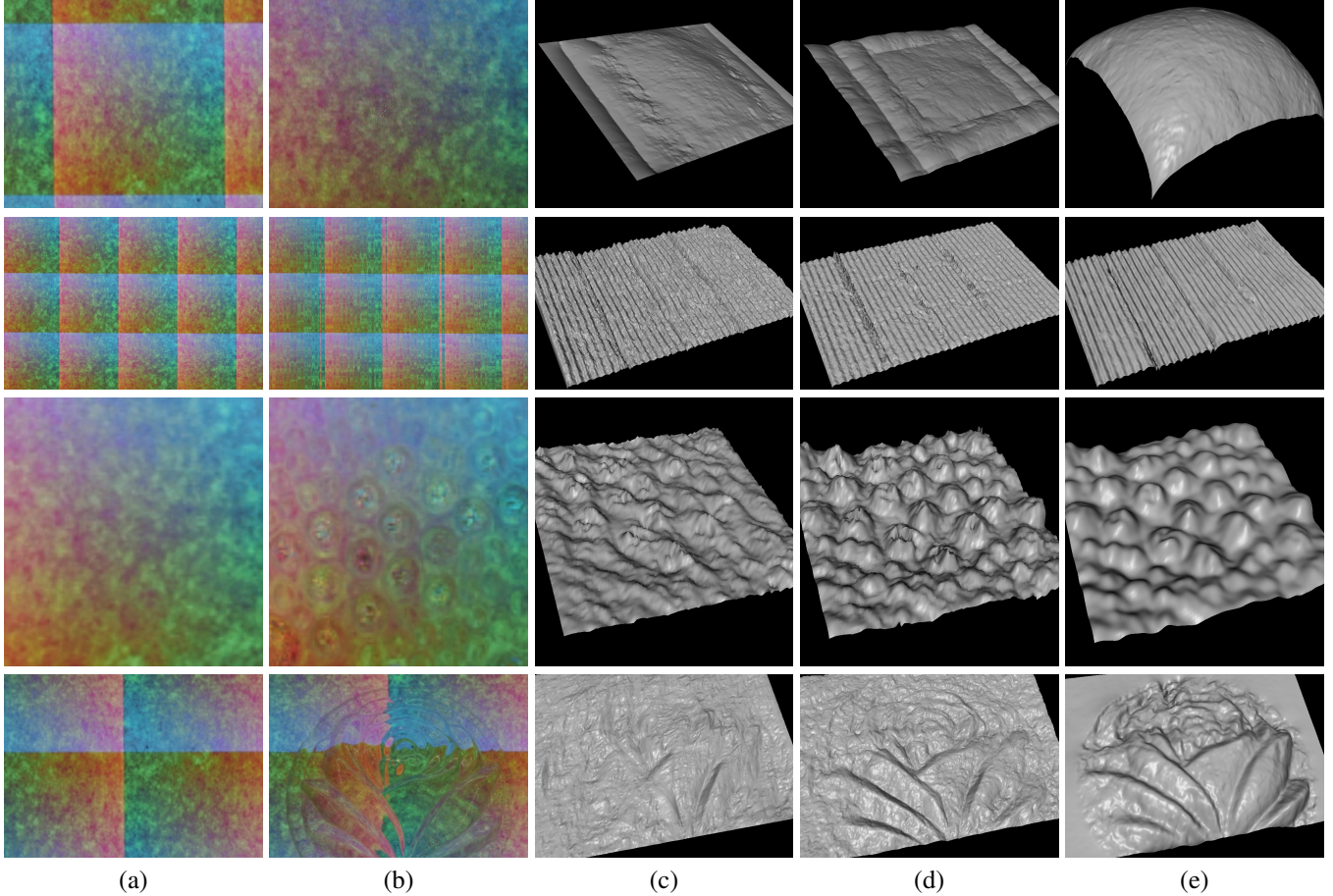


Figure 6. Surface reconstruction result from a single image pair. Top row: an acrylic dome shown in Fig. 3, second row: a wedge glass whose surface we shown in Fig. 4; third row: a glass with bubble texture (Fig. 5 (a)); bottom row: a flower glass shown in Fig. 4. (a) The background color image captured without the glass. (b) The image captured with glass. These images, (a) and (b), are cropped to the extent of the object, excluding tape applied to keep the object affixed the background. (c) Result from Murase’s algorithm after estimating deflections. (f) Result from our deflection-map approach. (e) Result from image-based approach. Note that the line artifacts is due to ramp boundaries.

- ject reconstruction. In *STAR Proceedings of Eurographics*, 2008.
- [5] K. N. Kutulakos and E. Steger. A theory of refractive and specular 3D shape by light-path triangulation. In *ICCV*, 2005.
- [6] N. J. Morris and K. N. Kutulakos. Dynamic refraction stereo. In *ICCV*, 2005.
- [7] H. Murase. Surface shape reconstruction of a nonrigid transparent object using refraction and motion. *TPAMI*, pages 1045–1052, 1992.
- [8] K. Perlin. An image synthesizer. *SIGGRAPH Comput. Graph.*, 19(3):287–296, 1985.
- [9] S. Roth and M. J. Black. Specular flow and the recovery of surface structure. In *CVPR*, 2006.
- [10] Q. Shan, B. Curless, and T. Kohno. Seeing through obscure glass. In *ECCV*, 2010.
- [11] Q. Shan, J. Jia, and A. Agarwala. High-quality motion deblurring from a single image. *SIGGRAPH*, 2008.
- [12] M. Tarini, H. Lensch, M. Goesele, and H.-P. Seidel. 3D acquisition of mirroring objects. *Graphical Models*, 64(4):233–259, 2005.
- [13] Y. Tian and S. Narasimhan. A globally optimal data-driven approach for image distortion estimation. In *CVPR*, 2010.
- [14] Y. Vasilyev, T. Zickler, S. Gortler, and O. Ben-Shahar. Shape from specular flow: Is one flow enough? In *CVPR*, 2011.
- [15] G. Wetzstein, D. Roodnick, W. Heidrich, and R. Raskar. Refractive shape from light field distortion. In *ICCV*, 2011.