Retrographic sensing for the measurement of surface texture and shape

Micah K. Johnson
MIT
kimo@mit.edu

Edward H. Adelson
MIT
adelson@csail.mit.edu

Abstract

We describe a novel device that can be used as a 2.5D "scanner" for acquiring surface texture and shape. The device consists of a slab of clear elastomer covered with a reflective skin. When an object presses on the skin, the skin distorts to take on the shape of the object's surface. When viewed from behind (through the elastomer slab), the skin appears as a relief replica of the surface. A camera records an image of this relief, using illumination from red, green, and blue light sources at three different positions. A photometric stereo algorithm that is tailored to the device is then used to reconstruct the surface. There is no problem dealing with transparent or specular materials because the skin supplies its own BRDF. Complete information is recorded in a single frame; therefore we can record video of the changing deformation of the skin, and then generate an animation of the changing surface. Our sensor has no moving parts (other than the elastomer slab), uses inexpensive materials, and can be made into a portable device that can be used "in the field" to record surface shape and texture.

1. Introduction

We describe a new sensor that converts information about surface shape and pressure into images. In its simplest form, it consists of a slab of clear elastomer covered with a reflective elastic skin, along with a camera and lighting. Figure 1(a) and (b) shows an Oreo cookie being pressed against the elastomer slab. The reflective skin, which is made from opaque elastomer paint, takes on the shape of the Oreo's surface. Oblique illumination converts the deformation to a shaded image. We refer to this device as a retrographic sensor, because it senses the deformation of the device's surface by imaging from the reverse side.

In order to extract the shape we need to apply some shape-from-X method. Since we have control over the bidirectional reflectance distribution function (BRDF) and the lighting, we use photometric stereo. We capture a single frame illuminated by red, green, and blue lights at different positions; this provides three shaded images. Using a photometric stereo algorithm tailored to the sensor, we reconstruct the surface, which is rendered in Fig. 1(c).

The retrographic sensor is not a general purpose 3D scanner, but it has some unique properties. It can be built as a simple, compact device that can be used to acquire surface textures instantly "in the field" rather than in the lab. It has high spatial resolution. Since it provides its own BRDF, it has no problem dealing with specular or transparent materials, or materials with varying albedo. Since it captures 2.5D surface data in a single frame, we can also use it record video of how the surface changes in response to changing pressure, which offers some unusual possibilities that we will discuss.

2. Related Work

There are many methods for measuring the 3D topography of a surface. The oldest is a profilometer that laboriously runs a mechanical probe over the surface. More recently, optical techniques have become widely used because they have the advantage of not requiring physical contact and can be fast and convenient. Optical methods are classified as passive (*e.g.*, stereopsis or shape from defocus) or active (*e.g.*, structured light, time-of-flight, photometric stereo). The active systems are the most successful at this time, and they can be quite versatile and accurate [3, 8, 7].

Photometric stereo was first described by Woodham as a method for estimating surface orientation from intensity data [14]. In that work, the reflectance function was assumed to be known analytically, but later work has followed two general approaches: using either parametric models of surface reflectance or calibration objects. Differences between techniques in both classes include whether or not the lighting is known or if albedo is assumed to be constant over the surface.

Of the approaches that assume parametric models for surface reflectance, many assume Lambertian reflectance or that the reflectance function can be described by a mixture of diffuse and specular components [5, 13, 12, 2, 9, 3]. While these models allow for relatively simple equations, they are not able to model the reflectance functions of many materials [10].

978-1-4244-3991-1/09/$25.00 ©2009 IEEE

(a) (b) (c)

Figure 1. (a) A cookie is pressed against the skin of an elastomer block. (b) The skin is distorted, as shown in this view from beneath. (c) The cookie's shape can be measured using photometric stereo and rendered at a novel viewpoint.

In more recent work, Goldman et al. modeled surface reflectance as an unknown linear combination of basis reflectance functions [7]. They optimized for both the basis functions and the surface normals and showed impressive results on fairly complex materials, provided the materials were well-approximated by an isotropic Ward model. Another recent line of research has investigated surface reconstructions that do not assume parametric models for reflectance. Instead, they assume physical properties that apply to a wide variety of reflectance functions, such as energy conservation, non-negativity, and Helmholtz reciprocity [16, 1]. These techniques are data-driven and often require many images under controlled lighting conditions.

With Lambertian surfaces, it is possible to achieve accurate reconstructions with arbitrary unknown lighting [3], but with more complex reflectance functions, including specular surfaces, controlled illumination and often many separate images are usually necessary. For example, some techniques use tens of images [7, 3], while others use in the hundreds [1].

A second general approach to photometric stereo uses calibration objects—objects with the same reflectance properties as the surface being reconstructed. This was the approach taken in the earliest works on photometric stereo [12, 15], and it was recently revisited by Hertzmann et al. where the calibration object idea was extended to a set instead of a single object [8]. The advantage of these approaches is that the reconstruction algorithm is simple—to reconstruct a region, find the closest matching region in the set of observed data and use its shape (*i.e.*, its gradients). The main disadvantage is the need for a calibration object with the same reflectance properties as the object being modeled—this requirement could be met with paint, but otherwise seemed too restrictive for modeling natural materials.

In the case of our retrographic sensor, however, any material pressed into it will inherit the sensor's reflectance characteristics. In other words, the sensor non-destructively changes the reflectance function of an object on contact. This property of our sensor removes the key disadvantage of the calibration-target approach and makes the use of calibration targets attractive. In addition, since we have control over the reflectance of the sensor, we can design it to be simple, avoiding the need for complicated lighting or capturing processes involving hundreds of images. Many of the assumptions of early photometric stereo work once considered restrictive (*e.g.*, uniform albedo, directional lighting, known reflectance function), are now in fact satisfied. Therefore, our reconstruction technique closely resembles some of the earliest work on photometric stereo, with improvements to allow for better reconstructions from our sensor.

3. Building the sensor

In building a retrographic device, there are many decisions to be made. First is the choice of a clear elastomeric material. We have experimented with polymers from various families including silicones, polyurethanes, and thermoplastic elastomers (TPEs) such as styrenic block copolymers. We have mainly used TPEs because they combine elasticity and strength. They can be formed into arbitrary shapes and are fairly robust, returning to their original shape under normal usage. They can also be dissolved in common solvents, which means they can be used as a base for an elastic paint.

A typical sensor is made as follows. A quantity of elastomer is melted in an oven into a block of the desired shape. A quantity of the same elastomer is dissolved in toluene and/or other solvents. A reflective pigment is added to this liquid to form a paint. The paint is applied to the surface of the elastomer by spraying with an airbrush.

The BRDF of the skin (as viewed from behind) is determined by the choice of pigment and the method of application. An ordinary pigment (such as titanium dioxide) yields a diffuse BRDF. We have found that diffuse BRDFs are useful for measuring relatively deep objects (depths on the order of 1 cm). However, to capture small variations in the

(a) (b)

Figure 2. (a) This decorative pin consists of a glass bas-relief portrait mounted in a shiny gold setting. (b) The RGB image provided by the retrographic sensor. The pin is pressed into the elastomer skin, and colored lights illuminate it from three directions.

surface normal, we need a more specular BRDF. Fine metal flakes (usually aluminum or bronze) produce a skin resembling brushed metal with a specular lobe that is strong but somewhat broad. The width of the lobe is determined by the flatness of the individual flakes and the randomness of their orientation. (The effective width of the specular lobe can be increased by using an extended source of illumination.) With metal flake skin, small changes in surface normal can yield large changes in the luminance of the reflected light, *i.e*., there is a high gain that allows small undulations to be readily visible. However, far from the specular lobe the gain is correspondingly low, which leads to poor resolution of other angles. Thus, the choice between diffuse or specular BRDFs is a tradeoff between depth and detail.

The thickness of the slab of clear elastomer, along with its elastomeric parameters, determines the maximum depth variation that can be measured. We have built sensors with thicknesses ranging from less than 1 mm to 4 cm. The hardness of an elastomer is commonly measured on the Shore A scale, where 5 is very soft and 95 is very hard. We typically use elastomers with Shore A values between 5 and 20.

4. Optical system

To capture images of the retrographic device, we secure it in a mounting stand and arrange the camera and lights. For the results in this paper we used a Canon digital SLR (EOS-1D Mark III) equipped with a 100 mm macro lens, aimed straight at the sensor from a distance of 40 cm. Our lights are red, green, and blue floodlights that utilize LED arrays. The lights are positioned 25 cm from the center of the sensor at an elevation angle of 30 degrees.

Figure 2 shows a decorative pin (a) and the RGB image (b) that is captured when the object is pressed against the sensor from behind. It is evident that each light is providing a shaded image, with the shading in a different direction. To a first approximation, we can see the three images by simply looking at the R, G, and B channels individually. In practice there is crosstalk, but it doesn't matter since we will use a lookup that is empirically derived for this sensor.

5. Photometric Stereo

We model the surface of the sensor with a height function $z = f(x, y)$. The height function represents the displacement of the sensor from its resting state; when nothing is touching the sensor, the height is zero. We assume that image projection is orthographic—the position (x, y) in the image corresponds to the location (x, y) on the sensor. Under this assumption, the gradient (p, q) at position (x, y) is given by

$$p = \frac{\partial f}{\partial x}, \qquad q = \frac{\partial f}{\partial y}, \tag{1}$$

and the surface normal is $N(x, y) = \begin{pmatrix} p & q & -1 \end{pmatrix}^T$. We assume that the shading at a point on the surface depends only on its surface normal; that is, there are no cast shadows or interreflections. Under this assumption, the intensity at a point (x, y) can be modeled as $I(x, y) = R(p, q)$ where (p, q) is the gradient at (x, y).

The reflectance function models both the lighting environment and the BDRF of the sensor surface. We assume the albedo is constant, which is reasonable since the sensor is painted with a uniform pigment. Note that we do not assume Lambertian reflectance or point light sources—the specific reflectance function for the sensor is learned from calibration targets.

The reflectance function $R(p, q)$ maps values from a two-dimensional space into a one dimensional space of intensities. In general, there are many sets of p and q that map to the same intensity value, and thus the reflectance function is not trivially invertible. To reduce ambiguities, we use a photometric stereo approach: multiple images under different illumination conditions. With three images, the problem is theoretically overconstrained—three measurements per pixel are used to estimate two gradient values:

$$\vec{I}(x, y) = \vec{R}(p(x, y), q(x, y)), \tag{2}$$

where $\vec{I}(x, y) = \begin{pmatrix} I_1(x, y) & I_2(x, y) & I_3(x, y) \end{pmatrix}^T$ and $\vec{R}(p, q) = \begin{pmatrix} R_1(p, q) & R_2(p, q) & R_3(p, q) \end{pmatrix}^T$. The reflectance functions are, in general, non-linear functions of the gradients p and q and we are interested in the inverse function—a function that maps observed intensity triples to gradients. For this, we build a lookup table.

5.1. Lookup tables

Similar to [12, 15, 8], we use a lookup table to learn the correspondence between intensity triples and gradients. Our algorithm is closest to that presented by Woodham [15], but

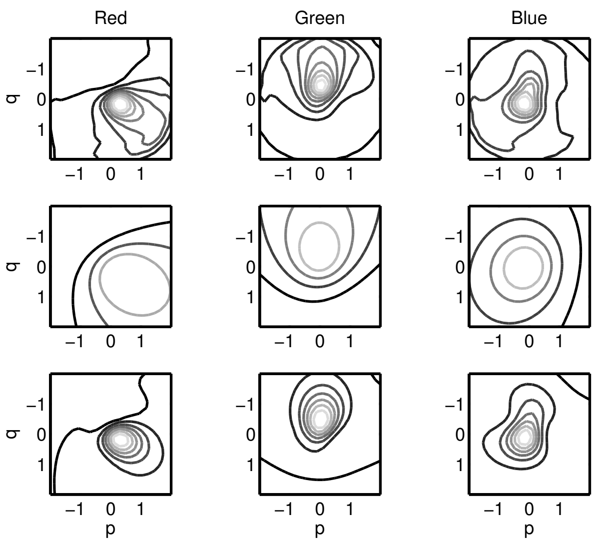


Figure 3. Comparison of the retrographic sensor reflectance function (top) to a Lambertian approximation (middle), and higher-order spherical harmonic approximation (bottom) for the three color channels. The axes are gradients, (p, q), and the isophotes are drawn at the same intensities in each plot. The specular lobe of the sensor reflectance function is shown by the relative closeness of the isophotes.

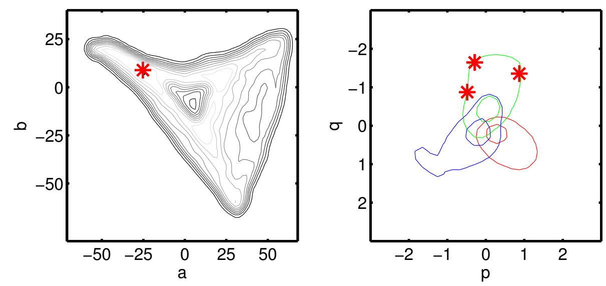


Figure 4. Ambiguous mappings. On the left are isodensity contours of the color distribution of the calibration sphere in (a^*, b^*) space. On the right is (p, q) gradient space. The red stars denote ambiguities between color and gradients—multiple disjoint pairs of gradients can produce the same color, within a small tolerance.

we extrapolate the data in the table using a low-order approximation of the reflectance function, refine the gradient estimates within each bin, and handle collisions, *i.e.*, multiple gradient pairs that produce the same intensity triple. Our lookup table is three-dimensional and each bin contains a gradient and a first-order approximation of the reflectance functions in the neighborhood near the gradient.

The lookup table is populated using a calibration target with known geometry, such as a sphere, or grid of spheres. The target is manually located in the calibration image, though this step could be automated. Using the calibration target, we first construct the reflectance maps, Fig. 3. These functions encode the mapping from gradients to intensity and we approximate them by sampling the observed colors on a fixed grid of p and q values.

To fill in missing data, we approximate the reflectance function with a low-order spherical harmonic model. This model is motivated by the observation that Lambertian objects under arbitrarily complex lighting environments can be well-approximated using the first three orders of spherical harmonics (orders zero through two) [4, 11]. Specifically, at a surface normal $\vec{N}$, the reflectance can be modeled as a linear combination of spherical harmonic basis functions:

$$R(\vec{N}) = \sum_{n=0}^{k} \sum_{m=-n}^{n} \alpha_{n,m} Y_{n,m}(\vec{N}) \,, \qquad (3)$$

where $k = 2$ for the Lambertian case. To model the reflectance function of our sensor, we use the first six orders of spherical harmonics, which we find produces a better fit, Fig. 3. The entire calibration process is repeated when we switch sensors or change light positions to ensure an accurate mapping from intensities to surface normals.

We use quantization as the mapping from intensity to lookup table index since quantization is fast and for our specific conditions, it produces a minimal number of collisions. The quantization parameters are derived from the observed color distributions of the calibration target. For each observed color in the sampled reflectance maps, we store the associated gradient values in the lookup table. To detect collisions, for each non-empty bin in the table we use mean-shift clustering to group the observed gradient values into clusters [6]. The cluster bandwidth, δ, is set proportional to the average norm of the gradients in the bin—we find that this heuristic effectively groups gradients with normal vectors that are close in angle without having to introduce angular measurements into the clustering algorithm. The center of each detected cluster is used as a coarse gradient estimate for the bin.

Since the surface of our sensor is not Lambertian, the three color measurements alone are not always sufficient to uniquely determine the gradients at each pixel. For example, Fig. 4 shows the mapping stored in one of the bins of our lookup table. On the left of the figure is a representation of the color space of the lookup table—the contours show the density of colors captured from the calibration sphere in (a^*, b^*) space and the red star denotes the color associated with the bin. The right side of the figure is gradient space, with p and q on the horizontal and vertical axes. The red stars denote three pairs of gradient values that produced the observed color. Of the possible coarse estimates in each bin, we choose the estimate that best approximates the nearby intensity gradients, as described below. For our sensor, typically only 5% of the bins contain ambiguous mappings; the exact amount varies with the BRDF of the sensor and the lighting configuration. In addition, the surface of the calibration target is inherently 2D and therefore cannot produce the full 3D space of colors that the lookup table can store. Bins that are empty after calibration are filled with the index of the bin containing the closest observed color.

Table lookup returns a coarse estimate of the gradient, which we denote as (p_0, q_0). To refine the estimate, we approximate the reflectance function in the neighborhood of (p_0, q_0) using a first-order Taylor series expansion:

$$
\begin{aligned}
\vec{R}(p,q) &\approx \vec{R}(p_0,q_0) + J(p_0,q_0)\begin{pmatrix} p-p_0 \\ q-q_0 \end{pmatrix}, \\
J &= \begin{pmatrix} \frac{\partial R_1}{\partial p} & \frac{\partial R_1}{\partial q} \\ \frac{\partial R_2}{\partial p} & \frac{\partial R_2}{\partial q} \\ \frac{\partial R_3}{\partial p} & \frac{\partial R_3}{\partial q} \end{pmatrix}. \qquad (4)
\end{aligned}
$$

The Jacobian J is the matrix of first partial derivatives of the reflectance functions with respect to the gradients. This approximation can be inverted to estimate the gradient (p, q) that produced the observed reflectance $R(p, q)$:

$$
\begin{pmatrix} p \\ q \end{pmatrix} = J^{+}\left(\vec{R}(p,q) - \vec{R}(p_0,q_0)\right) + \begin{pmatrix} p_0 \\ q_0 \end{pmatrix}, \qquad (5)
$$

where J^{+} denotes the pseudo-inverse of the Jacobian. The rank of the Jacobian is at most two, but if the partial derivatives of the reflectance functions are small near (p_0, q_0) it may be ill-conditioned. To improve conditioning, we regularize the pseudo-inverse using Tikhonov regularization, $J^{+} = (J^T J + \lambda I)^{-1} J^T$, for a small parameter λ.

In regions of the reflectance function with large curvature (*e.g.*, near the specular lobe), the first-order Taylor series approximation may not be accurate. In this case, it is better to use the coarse estimate of the gradient (p_0, q_0). We can judge how well the approximation models the local neighborhood by considering the intensity gradients at (x, y). As described in [15], the Hessian matrix of the height function can be estimated from the intensity gradients by differentiating Equation 2:

$$
\begin{aligned}
\begin{pmatrix} \frac{\partial \vec{I}}{\partial x} & \frac{\partial \vec{I}}{\partial y} \end{pmatrix} &= JH, \\
H &= J^{+}\begin{pmatrix} \frac{\partial \vec{I}}{\partial x} & \frac{\partial \vec{I}}{\partial y} \end{pmatrix}. \qquad (6)
\end{aligned}
$$

We symmetrize the Hessian, $\hat{H} = (H + H^T)/2$, and then measure the error in approximating the gradients:

$$
\epsilon^2 = \left\| \begin{pmatrix} \frac{\partial \vec{I}}{\partial x} & \frac{\partial \vec{I}}{\partial y} \end{pmatrix} - J\hat{H} \right\|_F^2. \qquad (7)
$$

This error is used to weight the interpolation in Equation 5–large errors in the neighborhood cause the interpolation to use the coarse estimate (p_0, q_0), small errors cause the interpolation to use the linear refinement. The final interpolation scheme is:

$$
\begin{aligned}
\begin{pmatrix} p \\ q \end{pmatrix} &= \omega J^{+}\left(\vec{R}(p,q) - \vec{R}(p_0,q_0)\right) + \begin{pmatrix} p_0 \\ q_0 \end{pmatrix}, \\
\omega &= \exp(\epsilon^2/\sigma_\epsilon^2). \qquad (8)
\end{aligned}
$$

The error measure in Equation 7 is also used to select a coarse gradient approximation when there are ambiguous mappings in the lookup table.

With an estimate of the gradient (p, q) at every pixel, we reconstruct the height function $z = f(x, y)$ by minimizing a quadratic error function:

$$
E(z) = \sum_{x,y} \left(\frac{\partial f}{\partial x} - p\right)^2 + \left(\frac{\partial f}{\partial y} - q\right)^2. \qquad (9)
$$

We implemented a multigrid solver that uses the approximations:

$$
\begin{aligned}
\frac{\partial f}{\partial x} &= \frac{f(x+1,y) - f(x,y)}{h}, \\
\frac{\partial f}{\partial y} &= \frac{f(x,y+1) - f(x,y)}{h}, \qquad (10)
\end{aligned}
$$

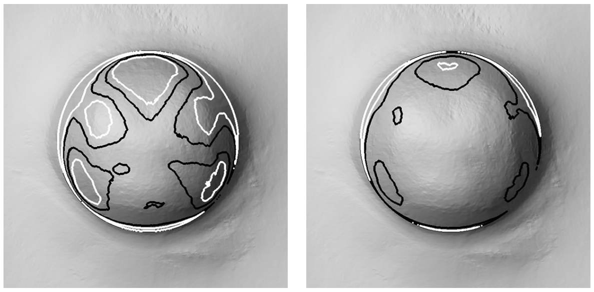
where h depends on the multigrid level. We find that the solver converges in approximately fifteen iterations.

6. Results

To determine the capabilities of our sensor, we "scanned" an object with known geometry (a sphere) and several novel objects that would be challenging for traditional photometric stereo techniques due to assumptions about the BRDF or scale. The sphere was used to assess reconstruction errors since we do not know the ground truth geometry of any of the other objects. The performance of the sensor for the novel objects was judged qualitatively by rendering the geometry with varying illumination and at various viewpoints.

The sensor was initially calibrated by imaging a grid of spheres with known radii. We constructed a lookup table, as described in Section 5.1, and used the same table for all of the results presented in this section. For all examples, a heightmap was estimated from a single RGB image (similar to Fig. 2(b)). The heightmap was rendered under orthographic projection as a diffuse material with directional lighting. The objects were also photographed separately under normal illumination with a 100 mm macro lens positioned 35 cm above the object.

To obtain an initial estimate of reconstruction errors, we "scanned" a half-inch diameter chrome ball bearing,



(a) (b)

Figure 5. Height and angular errors to an object with known geometry. (a) Rendered heightmap with level curves at 2% (black) and 4% (white) height error relative to the known radius. (b) Rendered heightmap with level curves at 2 degrees (black) and 4 degrees (white) of angular error between the estimated and actual surface normals.

which we assume is perfectly spherical. We estimated the heightmap from the RGB image and rendered it as a diffuse surface with directional lighting, Fig. 5. In Fig. 5(a), the level curves denote percent error in height relative to the known radius. The black curve shows points with 2% height error and the white curve shows points with 4% height error. In Fig. 5(b), the level curves denote angular error between the estimated and actual surface normals: the black curve shows points with 2 degrees of error and the white curve shows points with 4 degrees of error. Note that in general, the sensor has difficulty towards the edges of objects where the surface normals can reach steep angles compared to the normal of the sensor plane.

Figure 2(a) shows a decorative pin and Fig. 6 shows two views of the estimated surface. The pin is a challenging object for traditional techniques (including photometric stereo) because it has a transparent glass-like material in the center surrounded by gold-plated metal.

Figure 7 shows a drawer pull and two views of the estimated surface. The drawer pull is metal with varying albedo due to the pattern on the top. The BRDF of the pull would be difficult to model due to various degrees of wear and imperfections. The sensor is able to capture much of the detail in the pattern as well as the detail around the rim.

6.1. High-resolution sensor

Retrographic sensors can be designed with properties tailored to specific objects. The surface of a twenty dollar bill, shown in Fig. 8, was acquired with a sensor that gives higher resolution, but lower depth, than the sensor used for the previous results. The sensor is thinner and is made from three layers of 3M clear VHB mounting tape. The top surface is coated with a single layer of fine aluminum flakes.

Figure 8(b) shows a reconstruction of the twenty dollar bill. The estimated heightmap has been filtered to enhance the details. The text and patterns on the bill, as well as the security strip, are clearly visible.

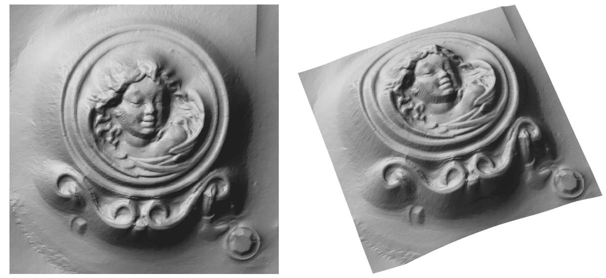


Figure 6. Two renderings of the acquired shape of a 1 cm decorative glass pin (see Fig. 2(a)).

6.2. Video

The retrographic sensor captures full information in a single frame, and therefore can be used to record video. However, it cannot capture the motions of freely moving objects, and cannot be used, for example, to capture hand gestures or facial movements. Instead it occupies a unique niche: it can capture its own deformations in response to changing pressure. Thus it can be used to study events that happen at the interface between two surfaces, where at least one of the surfaces (the sensor) is compliant.

As a test of our reconstruction technique for video, we recorded a finger rubbing the surface of the sensor both with and without lotion. Figure 9(a) shows a rendering of the estimated surface for a single frame of the video. Since the sensor has a soft flesh-like consistency, the finger presses into its surface, which deforms to meet the finger's shape. The resolution is high enough that the fingerprint is visible. Since the fingertip itself is soft, its overall shape is flattened compared to a free finger. Figure 9(b) shows a frame the video of the finger with St. Ive's Apricot Scrub, which is an exfoliating lotion that contains coarse scrubbing particles. The particles are clearly visible. As the finger moves, it is evident that the particles are not sliding, but rather are rolling between the finger and the sensor skin. They translate at half the speed of the finger, and they rotate as they move, as expected from rolling particles. In addition, the smooth part of the lotion acts as a lubricating cushion between the finger and the sensor skin, which causes the fingerprint to be invisible. Over time, as the finger continues to slide over the surface, the lotion is squeezed out, and the fingerprint becomes gradually visible again.

As far as we know, this is the first time such an interaction has been directly recorded. In order to better understand the interaction between a finger and, for example, a cheek, it would be necessary to build a sensor that more accurately captured the physical properties of the cheek, including the texture and frictional properties of the skin, and

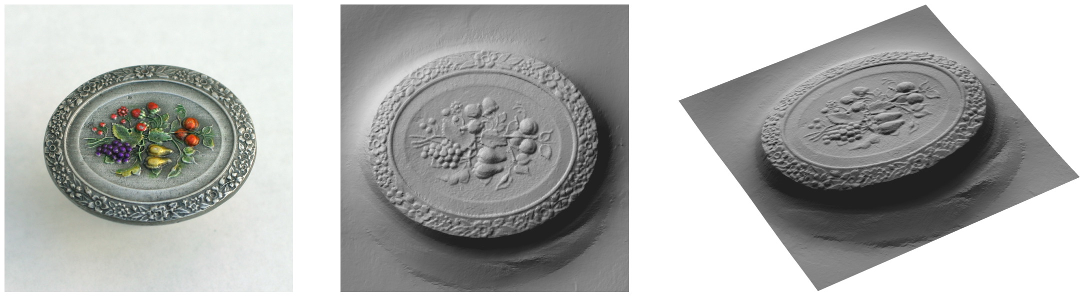


Figure 7. A drawer pull, about 2 cm across, along with two renderings of the acquired shape.

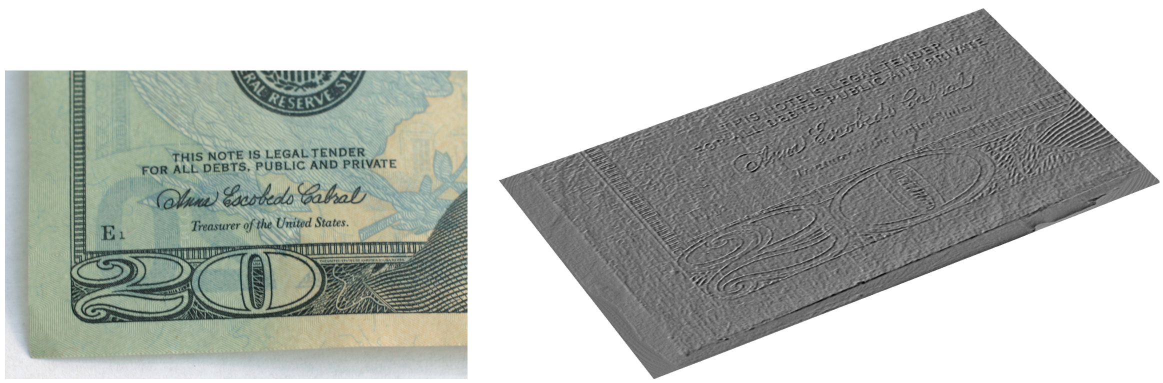


Figure 8. The corner of a twenty dollar bill. The ink is slightly raised, and the sensor is able to resolve this small difference in height, as shown by the reconstruction on the right.

the mechanical properties of the multiple layers of underlying flesh. We are moving in that direction, but even with the simple sensor we are now using, we are able to observe events that were previously inaccessible.

The deformation of human skin and flesh in response to external forces is a domain with many applications. The comfort of clothing against the skin, the texture of food as it is chewed and swallowed, and the feeling of cosmetic products as they are applied to the skin, are everyday examples of events that occur at the interface between skin and the outside world. These problems are of great practical and commercial importance, and their study could benefit from appropriately constructed retrographic sensors.

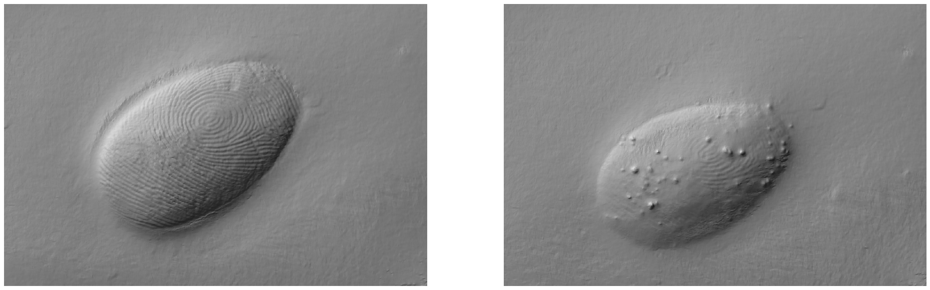
The retrographic sensor can also be considered a tactile sensor. It could provide a robot with a compliant fingertip with sensitivity exceeding human skin. It could also be used to fashion novel input devices such as high resolution touch pads.

7. Discussion

There many circumstances in which it is useful to acquire information about surface shape, and different technologies are well suited to different domains. The retrographic sensor offers a novel approach that has its own strengths and weaknesses. It is a contact based sensor, because it must be pressed against the object of interest. It is a light based sensor, because the surface deformation is converted to an image and captured with a camera. While various methods could be used to convert the image data into 2.5D data, we have chosen to use photometric stereo based on three colored lights coming from different directions. The sensor skin provides a known BRDF, and when combined with the geometry of the lights and camera, this provides a specific mapping between RGB triples and surface normals.

In our current implementation, there are some remaining ambiguities in the mapping from color to surface normal. We could reduce these ambiguities by using a large number of images, even going to the dozens of images used by some investigators. However, by using three images in different colors, we can get capture all of the data instantly in a single RGB frame, which greatly simplifies the acquisition process.

Single frame capture allows us to use video recordings, and here the retrographic sensor offers unique possibilities. It is a compliant surface, similar to human skin, and can be used to study the events that occur at the interface between the skin and the world. We have shown the example of a finger pressing on the skin and rubbing it with an exfoliating lotion containing scrubbing particles. The entire pro-



(a) (b)

Figure 9. (a) One frame from a video showing the reconstructed shape of a finger sliding over the sensor skin. The fingerprint is well resolved. (b) One frame from a video showing the finger applying Apricot Scrub lotion to the sensor skin. The particles, which roll between the finger and the skin, are clearly resolved.

cess can be viewed, in high resolution 3D, "from the inside looking out." A more accurate assessment of this interaction could be achieved by building a sensor that more accurately mimics the properties of human skin and its underlying tissue. Similar techniques could be useful in studying other interactions, such as those between the skin and clothing. In another example, one could build sensors into crash test dummies, and then study, in great detail, the forces that occur during a crash with given seat belt.

At present, we have demonstrated the utility of the sensor in more straightforward applications: capturing the shapes and textures of moderately sized surfaces in the world around us. When compared to other devices, the physical sensor is quite simple and low-tech. It consists of a camera, a slab of painted elastomer and three lights. We convert the raw imagery into 2.5D maps through the use of a photometric stereo algorithm that is matched to the sensor parameters. Because the retrographic sensor is different from other common devices for 3D acquisition, it is hard to predict its range of applications. However, it holds intriguing possibilities, and in some applications offers unique advantages.

References

[1] N. Alldrin, T. Zickler, and D. Kriegman. Photometric stereo with non-parametric and spatially-varying reflectance. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2008.

[2] S. Barsky and M. Petrou. The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(10):1239–1252, 2003.

[3] R. Basri, D. Jacobs, and I. Kemelmacher. Photometric stereo with general, unknown lighting. *International Journal of Computer Vision*, 72(3):239–257, 2007.

[4] R. Basri and D. W. Jacobs. Lambertian reflectance and linear subspaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(2):218–233, 2003.

[5] E. N. Coleman, Jr. and R. Jain. Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry. *Computer Graphics and Image Processing*, 18(4):309–328, 1982.

[6] K. Fukunaga and L. Hostetler. The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Transactions on Information Theory*, 21(1):32–40, 1975.

[7] D. B. Goldman, B. Curless, A. Hertzmann, and S. M. Seitz. Shape and spatially varying BRDFs from photometric stereo. In *Proceedings of ICCV 2005*, pages 341–348, 2005.

[8] A. Hertzmann and S. M. Seitz. Example-based photometric stereo: Shape reconstruction with general, varying brdfs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1254–1264, 2005.

[9] S. Nayar, K. Ikeuchi, and T. Kanade. Determining shape and reflectance of hybrid surfaces by photometric sampling. *IEEE Transactions on Robotics and Automation*, 6(4):418–431, 1990.

[10] A. Ngan, F. Durand, and W. Matusik. Experimental analysis of BRDF models. In *Proceedings of the Eurographics Symposium on Rendering*, pages 117–226, 2005.

[11] R. Ramamoorthi. Analytic PCA construction for theoretical analysis of lighting variability in images of a lambertian object. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(10):1322–1333, 2002.

[12] W. M. Silver. Determining shape and reflectance using multiple images. Master's thesis, Massachusetts Institute of Technology, 1980.

[13] H. Tagare and R. J. de Figueiredo. A theory of photometric stereo for a class of diffuse non-lambertian surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(2):133–152, 1991.

[14] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):139–144, 1980.

[15] R. J. Woodham. Gradient and curvature from photometric stereo including local confidence estimation. *Journal of the Optical Society of America A*, 11:3050–3068, 1994.

[16] T. Zickler, P. N. Belhumeur, and D. J. Kriegman. Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction. *International Journal of Computer Vision*, 49(2–3):215–227, 2002.