

Estimating Dual-scale Properties of Glossy Surfaces from Step-edge Lighting

Chun-Po Wang*
Cornell University

Noah Snavely†
Cornell University

Steve Marschner‡
Cornell University

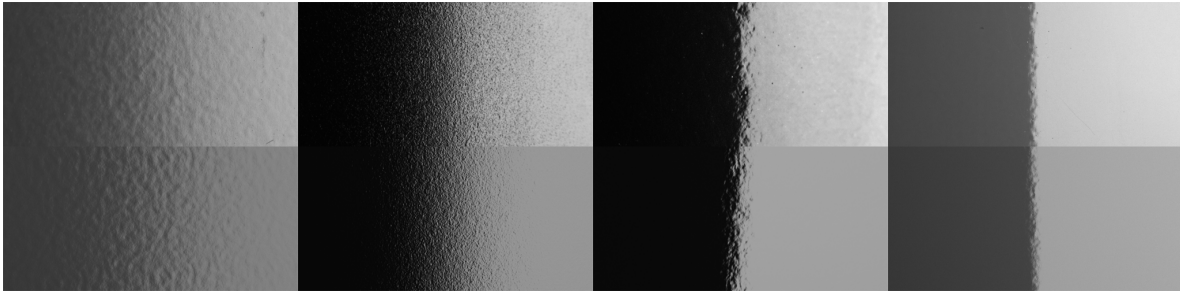


Figure 1: Example results of capturing four different surfaces under step-edge illumination. From left to right: painted bookend, metal bookend, shiny black folder, and whiteboard. The upper row shows real photos, and the bottom row shows images synthesized from the parameters acquired by our method.

Abstract

This paper introduces a rapid appearance capture method suited for a variety of common indoor surfaces, in which a single photograph of the reflection of a step edge is used to estimate both a BRDF and a statistical model for visible surface geometry, or mesostructure. It is applicable to surfaces with statistically stationary variation in surface height, even when these variations are large enough to produce visible texture in the image. Results are shown from a prototype system using a separate camera and LCD, demonstrating good visual matches for a range of man-made indoor materials.

Keywords: appearance capture, reflectance, rendering

Links: [DL](#) [PDF](#) [WEB](#) [DATA](#)

1 Introduction

Acquiring the appearance of real-world surfaces is important when real scenes need to be modeled faithfully by computer graphics. However, even the near-homogeneous manufactured surfaces—paints, metals, plastics—that are ubiquitous in man-made environments have evaded easy measurement. With the assumption of homogeneity, surfaces can be described by a single reflectance function that can be estimated from images [Yu et al. 1999] or rapidly acquired by handheld devices [Dong et al. 2010; X-Rite 2011], but the resulting surfaces always look too featureless, because surfaces with no visible texture are rare. Good, detailed appearance can be

achieved by measuring high-resolution parameter maps for particular samples [Gardner et al. 2003; Ghosh et al. 2009; Ghosh et al. 2010], but these methods are considerably less convenient—too much work to add subtle texture to a basically homogeneous surface.

In this paper we explore a middle path, suited for many common indoor surfaces, in which a rapid single-image measurement results in a statistical description of both visible and microscopic surface roughness that is sufficient to yield renderings of a surface that qualitatively *match* its appearance. Figure 2 shows some examples of surfaces that appear in a typical indoor scene. Note that these surfaces can be described not only in terms of reflectance—the diffuseness of the wall, the gloss evident in the reflection off the cabinet—but also by visible “bumps” with characteristic frequency content. Both of these phenomena are important to the appearance of surfaces (particularly in high-resolution imagery), and our approach explicitly handles both. However, our goal is *not* to measure exact surface properties—full BRDFs, accurate normal maps—but instead to capture enough information about the statistics of a surface to achieve a qualitative appearance match. To this end, we propose a dual-level model of surface appearance, with one level modeling microscopic surface geometry (described by the BRDF), and another modeling visible surface bumps (the mesostructure often represented in a normal or bump map). To estimate the parameters of our model, we propose an appearance capture system that uses a single image of the reflection of *step-edge* illumination from a planar sample of a surface to estimate both the BRDF and the statistics of meso-scale geometry (Figure 1). While we present a prototype capture system using a separate camera and LCD, our system could potentially be implemented on a consumer device with a display and a front-facing camera, such as an Apple iPad, and eventually could use natural illumination (which often includes step edges). Our appearance model and capture system are designed as a whole with the goal of enabling very simple, robust acquisition, while still handling many interesting real-world surfaces. We begin by briefly describing these two components.

Dual-level appearance model. Our reflectance model includes both diffuse and specular components; to represent specular reflection, we use a Cook-Torrance model (a microfacet-based model). The BRDF is then modeled by the slope distribution of the microfacets with a single parameter, the slope variance. Similarly, we observe that the visible bumps (the meso-scale structure) can be seen

*e-mail:cpwang@cs.cornell.edu

†e-mail:snavely@cs.cornell.edu

‡e-mail:srm@cs.cornell.edu



Figure 2: Examples of typical surfaces in an indoor environment, characterizing both by their BRDF (e.g., the gloss evident in the shiny metal cabinet) and texture created by visible bumps. (Please zoom into this figure to see the details.)

as visible “microfacets”; hence the roughness of the mesostructure can also be represented by the variance of the slopes. On the other hand, mesostructure also exhibits visible texture at different frequency scales, and thus requires a more detailed description than the invisible microfacets. Prior work has acquired complete normal maps to provide this detail [Chen et al. 2006; Francken et al. 2008a; Francken et al. 2008b; Ghosh et al. 2009; Ghosh et al. 2010], but many real-world surfaces—including certain paints, metals, plastics, and ceramics—have the following properties, which allow a simple statistical description to suffice:

- **Homogeneous:** the surface consists of single material with spatially invariant diffuse and specular albedoes.
- **Stochastic and stationary:** the mesostructure can be described as a stationary stochastic process (i.e., it doesn’t have structured patterns).
- **Isotropic:** the surface has isotropic appearance at both micro- and meso-scale.

Given these assumptions, we can capture the texture of the mesostructure by a 1D power spectrum over frequencies, the idea being that such surfaces can be approximated by a radially symmetric spectrum with random phase [Galerne et al. 2011], as the perception of many textures is well-characterized by their power spectra [Julesz 1962]; similar models are common in graphics [Perlin 1985]. In short, our surface model consists of two albedoes (one each for diffuse and specular reflection), two surface roughness parameters (for micro- and meso-scales), and the 1D power spectrum of the mesostructure.

Appearance capture. Given our appearance model, we propose a simple acquisition method: the surface parameters are estimated by analyzing an image of the surface under step-edge illumination (a half-black, half-white pattern displayed by an LCD). The core idea is that, roughly speaking, the reflection off of a surface can be approximated by a Gaussian filtered incident light [Ramamoorthi and Hanrahan 2001] (due to the BRDF), composed with a “scrambling” effect caused by the mesostructure (see Figure 4(c)), and that we can reason about the reflection in terms of these operations. With suitable approximation, this analysis can be done in the image domain, and the step-edge illumination makes this analysis particularly convenient. Our method requires only an off-the-shelf camera and LCD, and no polarization filters are required to separate diffuse

and specular components.¹

To summarize, our main contribution is an end-to-end system for simple appearance capture based on (1) a simple dual-scale surface model with a unified view of roughness in both scales, based on statistics rather than exact capture, and (2) a simple, fast, and robust acquisition method utilizing off-the-shelf hardware and simple image-based analysis. We apply our method to a variety of real surfaces, and evaluate results by comparing photographs of real reflections to synthesized images using the acquired parameters.

The rest of the paper is organized as follows. In Section 2, we briefly review related work in material and mesostructure acquisition. In Section 3, we describe our dual-scale appearance model. In Section 4, we describe our image-based acquisition method. Results are presented in Section 5.

2 Related Work

Capturing the appearance of a surface is a fundamental problem in computer graphics and vision, and many different approaches to this problem have been developed. On one hand are techniques that directly measure detailed BRDFs through specialized hardware, such as gonireflectometers [Ward 1992], camera and light source setups [Marschner et al. 1999], and specially designed optical equipment [Dong et al. 2010]. Such techniques often involve expensive, complex hardware or time-consuming capture, and are thus geared towards experts. At the other extreme are techniques that estimate reflectance from images under natural conditions. Ramamoorthi and Hanrahan develop a signal processing framework that allows for the recovery of BRDFs from images of a known shape in a range of scenarios, including when both the lighting and BRDF are unknown [2001]. Zickler *et al.* estimate spatially varying BRDFs (SVBRDFs) of known shapes from images, using smoothness priors to reduce the number of images needed for acquisition [2006]. Romeiro and Zickler estimate homogeneous reflectance from single images under natural illumination using priors on lighting environments [2010]. These techniques often involve challenging inverse problems, and assume known geometry. Our work lies between these two extremes: we use off-the-shelf hardware along with a simple parametric appearance model to quickly estimate surface parameters (including meso-scale geometry), while avoiding difficult inverse problems.

Parametric models are often used in BRDF capture techniques to simplify acquisition, or for fitting sampled BRDF measurements [Ngan et al. 2005]. For instance, Gardner *et al.* [2003] use a linear light source capture setup to find the parameters of a Ward reflectance model [Ward 1992], while Georgiades uses a Torrance-Sparrow model to fit the shape and BRDF of surfaces from a small number of images [Georgiades 2003]. Ghosh *et al.* use second order statistics of surface reflections to capture rough specular surfaces using spherical gradient illumination from a relatively small number of images [Ghosh et al. 2009].

In addition to material properties, the geometry of a surface also plays a critical role in its appearance. In this work, we are primarily concerned with mesoscale surface geometry (often described with a bump or normal map), rather than macro-scale object geometry. Recovering detailed mesoscale geometry has been achieved with a variety of techniques, many related to shape-from-shading or photometric stereo [Woodham 1980]. While many such methods assume Lambertian reflectance, recent work has extended photometric stereo to specular surfaces; e.g., Wang and Dana recover

¹The sample surface needs to be planar, which is usually true in practice since we only need a small surface; also, we assume that the input illumination and mesostructure do not generate attached shadows.

detailed relief from specularities [2006], and Chen *et al.* [2006] reconstruct normals based on finding peaks of specular highlights in large numbers of images under varying illumination. In subsequent work, the number of required images has been reduced using coded illumination [Francken *et al.* 2008b] or linear gradient illumination [Francken *et al.* 2008a]. Another way to measure surfaces with unknown BRDF is using a contact sensor [Johnson and Adelson 2009]. However, these techniques all recover normal maps but not material properties. Paterson *et al.* recover shape and heterogeneous BRDF using flash photography, but require multiple views [2005]; in contrast, our system is designed for homogeneous textured surfaces and needs less input data. Bidirectional texture functions (BTFs) have also been used to model appearance, but involving a much heavier data-driven approach than ours [Dana *et al.* 1999].

Our work fits into what we believe is a “sweet spot” in the space of appearance capture techniques: our method recovers both BRDF and mesostructure (as in prior work [Gardner *et al.* 2003; Georgiades 2003; Ghosh *et al.* 2009; Ghosh *et al.* 2010]), but adopts a simplified frequency spectrum model for the latter and a simple capture system using a single reflected image of a step-edge illuminant displayed on an LCD screen (along with a couple of images to calibrate the setup beforehand). Without an exact measurement of the mesostructure, we use a statistical approach to recover micro-scale properties. This idea has connections to the work of Dror *et al.* [2001]: the image statistics reveal surface properties when the illumination is constrained (or known). While the method is simple, our results show that for many man-made surfaces it is sufficient to render surfaces with visually convincing appearance. Our step-edge pattern allows us to separate the diffuse and specular components of the reflection by data fitting (as shown in Section 4.2). Hence, unlike other techniques [Francken *et al.* 2008b; Francken *et al.* 2008a; Ghosh *et al.* 2009], we do not require polarized filters, further decreasing the complexity of our setup.

3 Surface Reflectance Model

One of our key goals is to develop a model of surfaces that is simple enough to reliably estimate in practical settings, but which also describes a range of real-world surfaces. Our reflectance model includes both diffuse (Lambertian) and specular components. We assume the meso-scale bumps are small, so the bumpiness observed in an image is mostly due to specular (and not diffuse) reflection; however, diffuse reflectance must still be modeled, since our analysis requires separating the two components in the input image.

We use a microfacet model for specular reflectance, but further categorize the facets into micro- and meso-scale based whether they are visible in the image; micro-scale surface roughness corresponds to bumps that are microscopic, while meso-scale roughness corresponds to larger bumps that are resolvable in an image.² In this section we describe our models for each scale in detail, and show how they mix to give an overall roughness of the surface in terms of distribution of slopes on the surface. We relate this distribution to the blur observed in an image.

3.1 Micro-scale

Micro-scale facets are so tiny (i.e., much smaller than a pixel) that each pixel of the reflected image only captures the aggregate reflection from multiple facets. The reflectance therefore depends on the fraction of the facets that reflect the incident light to the

²Of course, whether a bump can be resolved depends on the resolution of the image; as the resolution decreases, the meso-scale structure starts to mix with the micro-scale structure [Han *et al.* 2007], as described below.

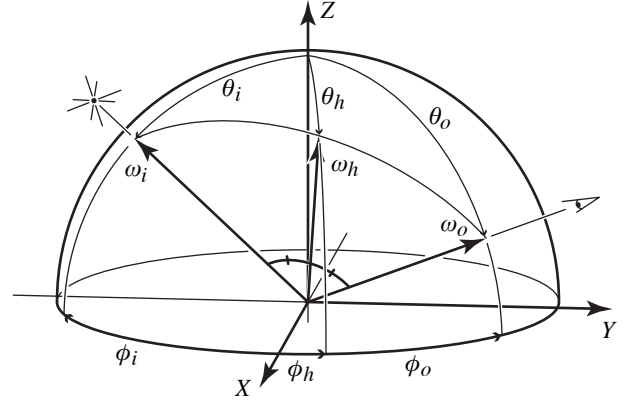


Figure 3: Geometry of reflection [Torrance and Sparrow 1967]. ω_h is the half-vector of the viewing direction ω_o and incident light direction ω_i .

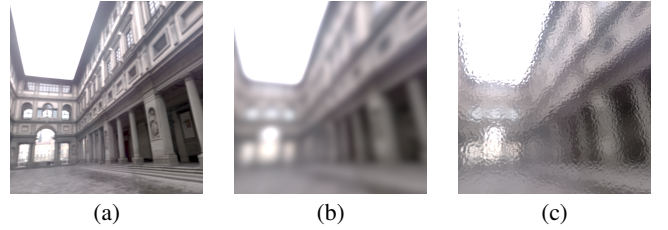


Figure 4: Synthesized reflection of environment map illumination off of (a) a perfect mirror; (b) a surface with only micro-scale roughness; (c) a surface with both micro- and meso-scale roughness. Notice that when viewed from a distance, the micro- and meso-scales of (c) blend together to form a “mixed”(overall) roughness.

observer, which can be characterized by the Cook-Torrance reflectance model [Cook and Torrance 1981] via the following reflection equation [Pharr and Humphreys 2010]:

$$dL_o(\omega_o) = \frac{F_r(\omega_o) G(\omega_o, \omega_i) L_i(\omega_i) D(\omega_h) d\omega_i}{4 \cos \theta_o} \quad (1)$$

(see figure 3) where $dL_o(\omega_o)$ is the differential radiance observed from viewing direction ω_o due to the incident radiance $L_i(\omega_i)$ over the solid angle $d\omega_i$. $F_r(\omega_o)$ is the Fresnel reflectance, $G(\omega_o, \omega_i)$ is the geometrical attenuation factor due to self-shadowing, and $D(\omega_h)$ is the slope probability density at the half-vector direction ω_h . For example, for a perfect mirror, $D(\omega_h)$ is a Dirac delta distribution that is non-zero only at the direction of the surface normal (resulting in perfectly sharp reflections). In our model, we assume that $D(\omega_h)$ for the micro-scale roughness is a Gaussian distribution, as in [Torrance and Sparrow 1967], with roughness a function of a single parameter, $\sigma_{D,\mu}$, the standard deviation (where μ stands for “micro”). This represents roughness at the micro-scale in terms of the facet slope distribution.

3.2 Meso-scale

At the meso-scale, the facets are large enough to be seen in the image, so the facet slope distribution alone is not enough to describe the surface appearance; instead, 2D information about the surface height or normal variation is needed, which can (for instance) be encoded into a 2D bump map or normal map. However, for random, isotropic mesostructure with random phase we can represent this

2D information in frequency space with a 1D power spectrum, since the 2D spectrum will be radially symmetric.

In practice, we found the slope distribution of a random-phase bump map to be Gaussian or very close to it, and thus we can also characterize the meso-scale roughness in terms of the standard deviation, as in Section 3.1. We call this parameter $\sigma_{D,m}$ (m stands for “meso”).

3.3 Overall roughness

While the two different scales considered above each have very different effects on surface appearance, the roughness at each scale can be characterized by its facet slope distribution, both modeled as Gaussian distributions (with parameters $\sigma_{D,\mu}$ and $\sigma_{D,m}$ respectively). When a surface has both micro- and meso-scale roughness, we can compute the overall slope distribution of the surface by combining these two Gaussian distributions. Intuitively, if the surface is viewed from a large enough distance (or captured with low enough resolution) so that the meso-scale bumps are no longer resolved, then the meso-scale roughness blends with the micro-scale roughness to form a combined roughness, as illustrated in Figure 8. The standard deviation of this “mixed” Gaussian distribution is simply:

$$\sigma_{D,M} = \sqrt{\sigma_{D,\mu}^2 + \sigma_{D,m}^2} \quad (2)$$

This blending of roughness parameters can be considered a simplified case of the normal map filtering proposed by Han *et al.* [2007], where both the normal map and the BRDF are characterized by Gaussian slope distributions.

3.4 Surface roughness in the image domain

We consider an image formation model that describes the effects of micro-scale and meso-scale roughness on a reflected image, both of which can be seen as “scrambling” a perfect mirror reflection at different scales. Imagine a surface lit with an environment map, as in Figure 4. If we were to replace the surface with a mirror, then an image of a reflection (assuming a pinhole camera with infinite depth of field) would simply be a warped version of the environment map (Figure 4(a)). If the surface exhibited only micro-scale roughness (i.e., had no visible bumps), then the effect on the reflected image is a “blur” operation (Figure 4(b)); the incoming illumination is “scrambled” by the microfacets, but at such a small scale that the result is a smoothly blurred version of the perfect reflection. This intuition is also used by work on pre-filtered environment maps [Kautz *et al.* 2000], and is used to estimate reflectance in the angular domain by Ramamoorthi and Hanrahan [2001]. Finally, the meso-scale acts to scramble the illumination once more, but at a visible scale where the scrambling can actually be seen as a permutation of pixels in the reflected image (Figure 4(c)). To summarize, the action of the surface roughness is to take a perfect mirror reflection, blur it (via micro-scale roughness), and then scramble the pixels (meso-scale roughness).

In this subsection, we want to derive the relationship between observed image blurriness and underlying surface roughness. Since the difference between micro- and meso-scale is just a matter of scale (i.e., unresolvable v.s. resolvable bumps), we first consider micro-scale roughness, where the blurred reflection is smooth. In order to work in the image domain (rather than the angular domain, as in [Ramamoorthi and Hanrahan 2001]), our setup uses a camera with a narrow field of view (FOV), but places the camera and light source close to the surface (so their arrangement must be considered); see the camera/light/surface setup shown in Figure 6.

For simplicity, suppose the image is the reflection of a point light

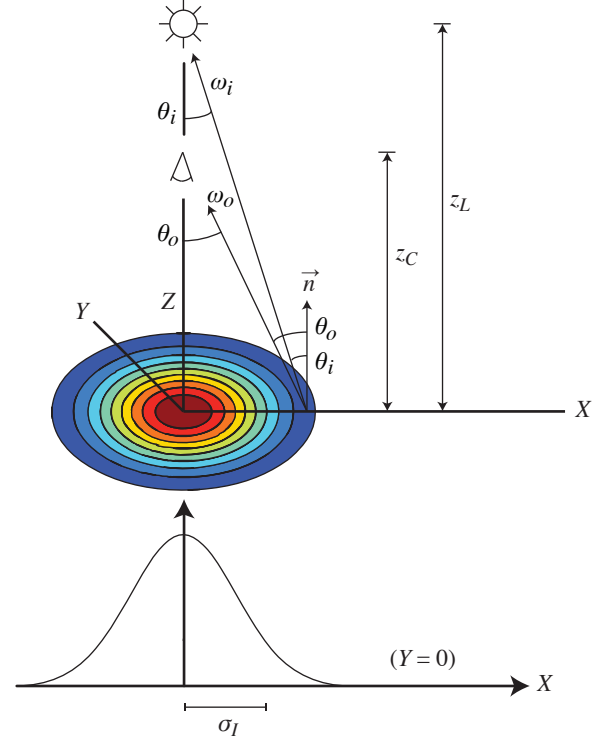


Figure 5: A point light source setup to find the relationship between σ_I (blurriness in image domain) and σ_D (slope distribution). The figure at bottom is the intensity of a 1D slice of the reflection at $y = 0$.

source; by using Equation 1 and the small angle approximation, the resulting image turns out to be a Gaussian function, say with standard deviation σ_I . Given a 1D slice of the captured image (say $y = 0$), we can relate the observed σ_I to the slope distribution parameters σ_D above (the details are presented in Appendix I):

$$I(x) \propto \exp \left\{ -\frac{(z_L + z_C)^2 x^2}{8\sigma_D^2 z_L^2 z_C^2} \right\} \quad (3)$$

where z_L, z_C are the distances from the light and camera to the surface in pixels (as measured on the object surface), respectively (Figure 5). This is simply a Gaussian with standard deviation $\frac{2\sigma_D z_L z_C}{z_L + z_C}$, so we have:

$$\sigma_D = \frac{z_L + z_C}{2z_L z_C} \sigma_I$$

Note that z_C (the distance from camera to surface in pixels) is exactly the focal length f_C (in pixels) of the camera. We can rewrite this in terms of the distance ratio $\frac{z_C}{z_L}$:

$$\sigma_D = \frac{1 + \frac{z_C}{z_L}}{2z_C} \sigma_I = \frac{1 + \frac{d_C}{d_L}}{2f_C} \sigma_I \quad (4)$$

where d_C and d_L are the distances from the camera and light to the surface, respectively (in any units, e.g., centimeters; note that the distance ratio is independent of the units). Therefore, in order to find σ_D given σ_I , we only need to know the focal length of the camera f_C (in pixels) and the distance ratio $\frac{d_C}{d_L}$.

The same relationship also holds for step-edge lighting, since it can be thought as an aggregation of point lights. In Section 4.2, we will show that resolved roughness (i.e., meso-scale and overall) can also

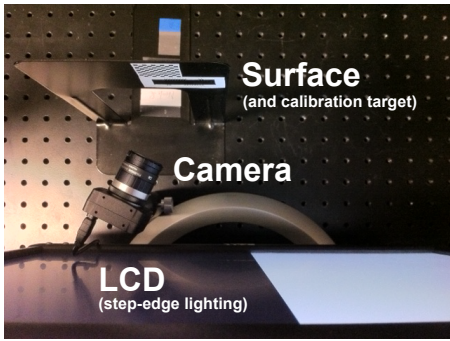


Figure 6: The setup of our acquisition system.

be analyzed in the same way by averaging, or finding a least-squares fit; thus, Equation 4 applies to all scales of surface roughness, resolvable or not.

4 Reflectance and Bump Map Measurement

This section describes our method for acquiring the parameters of the surface reflectance model proposed in the previous section. Our setup, described in Section 4.1, uses an LCD as an area light source to illuminate a subject surface with a step-edge pattern (a half-black, half-white image with a vertical edge), and a camera to capture the reflected image, from which we estimate parameters.

Intuitively, given a surface with both micro- and meso-scale roughness, the overall roughness can be found by observing the surface from a distance (so the mesostructure merges into micro-scale), then fitting a Gaussian filter that blurs the step-edge image to produce the observed image. Note that a Gaussian filtered step-edge image $I(x, y)$ can be written as a normalized Gauss error function:

$$I(x, y) = \frac{1}{2} \left[1 + \text{Erf} \left(\frac{x}{\sqrt{2}\sigma_I} \right) \right] \quad (5)$$

In practice, we cannot always observe a surface at arbitrarily long range. However, the same result can be achieved by least-squares fitting Equation 5 to the reflected image, effectively averaging over a sufficiently large image area, thus estimating the overall roughness. The details are described in Section 4.2.

The next step is to separate the micro- and meso-roughness from the overall roughness. In Section 4.3, we estimate micro-scale roughness by a method involving sorting pixels by intensity, which “undoes” the scrambling caused by mesostructure; in Section 4.4, we use this estimation to simulate a reflected image from a surface with only micro-scale roughness, and compare it to the original image to estimate meso-scale roughness. All three roughness estimates are then adjusted to satisfy Equation 2.

Finally, we recover the texture of the mesostructure, represented as a 1D frequency spectrum (as described in Section 3.2). In Section 4.4 we show this spectrum can be derived from the step-edge reflected image by using a first-order approximation to the scrambling behavior in the image domain.

4.1 Setup

Our current setup uses a regular LCD and a vision camera (Point Grey Chameleon). We use a lens with narrow FOV (about 20 degrees horizontally) to conform with the small angle approximation made in Appendix I; however, the viewing direction is not required

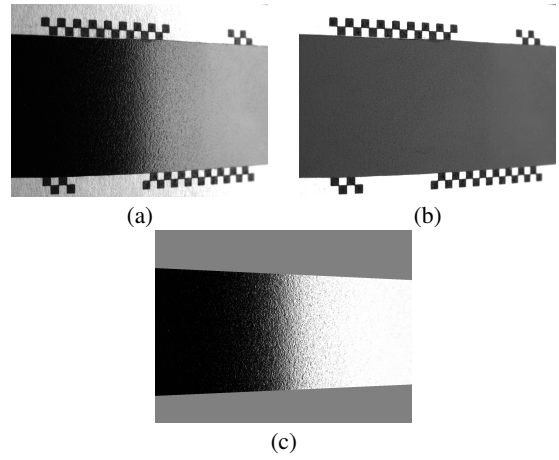


Figure 7: Normalizing the step-edge reflection (a) by the gray reflection (b). (c) is the normalized result, where 0 corresponds to black and white corresponds to 1. The calibration target is also removed.

to be close to surface normal, as Figure 6 shows. A calibration target is put on the surface to find the pose of the surface with respect to the camera. Given this setup, a total of three images are needed: the reflection of the step-edge image, the reflection of a gray image (used for brightness calibration, see Section 4.2), and a calibration image for the LCD, which is taken by attaching an mirror to the surface. In a practical realization of this system, the last two images could be replaced by a one-time calibration of the relative position of the LCD to the camera, and the brightness variation of the LCD.

4.2 Overall roughness estimation

In Section 3.4, we established the relationship (Equation 4) between the slope-domain roughness parameter σ_D and the image-domain roughness σ_I for unresolved roughness (e.g., when we observe the surface from a distance). When the bumps are resolvable (i.e., for meso-scale or overall roughness), then the relationship still holds as long as we average over a large enough window in the image (so that the bumps “average out”). In fact, for a step-edge-illuminated image, averaging over the columns of the image achieves the same effect (see Figure 8). Hence, if we use least squares to fit an error function to the overall roughness, Equation 4 still holds.

In practice, before fitting Equation 5 to the reflected image, two issues need to be considered: first, the reflected image is a combination of diffuse and specular components, while only the latter is used to estimate surface roughness. To separate them, we observe that in our setup (Figure 6) the diffuse component changes very slowly compared to the specular component, and can be approximated by a linear gradient. The result is not as accurate compared to other techniques (e.g., using polarized filters [Francken et al. 2008b; Francken et al. 2008a; Ghosh et al. 2009]), but it requires no special devices.

Second, the LCD is not a perfectly diffuse area light source (i.e., the radiance is dependent on the viewing angle); this breaks our model since the reflection intensities depend on the viewing angles to the LCD, and the image can no longer be approximated by a Gauss error function. This intensity variation, however, can be calibrated by capturing another surface reflection while displaying a constant gray image on the LCD; the specular component of the gray reflection encodes the variation due to the changing viewing angle to the LCD. Hence, we can normalize the specular component of the step-

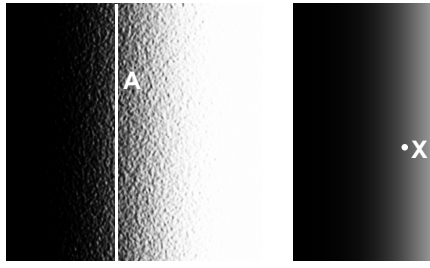


Figure 8: Two synthesized surfaces with roughness in different scales but the same overall slope distribution: the pixel value at X is the same as the average over the column A .

edge reflection by dividing it by the specular component of the gray reflection, as shown in Figure 7. Despite being a rough calibration, the variation due to viewing angle is often small enough for this method to work well.

Given these issues, we use the following model for the captured image:

$$I = c(I_G - D_G)E + D$$

where I is the captured step-edge reflected image we want to fit, I_G is the captured gray image, E is a Gauss error function, and c is a real-value factor (close to 2 if middle gray is used for I_G). D and D_G are the diffuse components of both reflections, respectively; both are modeled as linear functions. In this step, intensity variations due to meso-scale roughness are essentially noisy measurements that are averaged. We use MATLAB's `lsqnonlin` function for non-linear fitting to estimate the parameters of this model. We initialize D and D_G to zero, and the standard deviation of E to a sixth of the image width (so it increases smoothly across the image). After the fitting converges, the overall surface roughness $\sigma_{I,M}$ is given by the standard deviation of E .

4.3 Micro-scale estimation

To estimate micro-scale roughness, we observe that the mesostructure randomly “scrambles” the image that would be reflected by a surface with only micro-scale roughness, where pixels are monotonically increasing from side to side when a step-edge lighting is used. Thus, by sorting the pixels in column-major order (that is, sorting all pixels by intensity and reshaping them back to the same size as the input image in column-major order, so the pixels with lowest intensities are in the first column, and those with largest intensities are in the last), we can restore the scrambled reflection and isolate the micro-scale roughness. This is shown by the first two curves in Figure 9: they match each other almost perfectly despite very different appearances before sorting. However, due to the image noise and outliers such as the scratches or dust on the surface, what we get from real photos is more similar to the third curve, where the outliers, which usually have extreme pixel values, gather to the left and right end of the image (A and B), and the noise is sorted by its magnitude which yields the inclined lines. (Notice that we may have negative or greater-than-one pixels since we normalize the image) Therefore, a least-squares fitting method similar to Section 4.2 can not be used here.

However, we observe that the slope at the middle of the curve (which is the same as the gradient of the sorted image in the X direction) seems unaltered by the noise. This can be explained by considering the inverse function, i.e., position as a function of intensity. Observe that the position can then be interpreted as cumulative probability; thus the inverse function is exactly the cumulative distribution function (CDF). For the CDF, the effect of

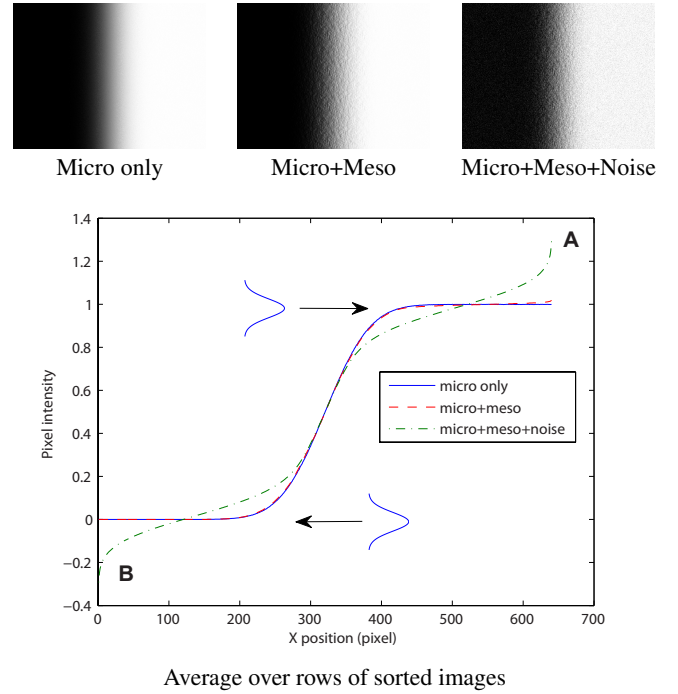


Figure 9: The scrambling behavior can be “undone” by sorting the pixels; this figure shows the average curves over the rows of three sorted images (this can be seen as the inverse CDF of the image intensities). The mesostructure has almost no effect on the curve, and the effect of noise can be approximated by a Gaussian filter along the Y -axis. Notice how the slopes at the middle of the curves stay the same.

the image noise can be seen as a convolution of the CDF and the noise distribution function: for example, the third curve in Figure 9 represents a reflected image with both scales of roughness plus zero-mean Gaussian noise, whose CDF is the Gaussian filtered version of the second curve, as shown by the rotated Gaussian kernel. Therefore, we can see that the slope of the middle segment of these curves is not changed by the existence of noise, as long as the noise is small compared to the image contrast.

Given this property, we show that the micro-scale roughness can be estimated by measuring this slope: the partial derivative in the X direction of the Gauss error function image (Equation 5) is

$$I_x(x, y) = \frac{\partial}{\partial x} I(x, y) = \frac{1}{\sqrt{2\pi}\sigma_I} \exp\left\{-\frac{x^2}{2\sigma_I^2}\right\}$$

Hence the slope s at the middle is given by $I_x(0, y) = \frac{1}{\sqrt{2\pi}\sigma_I}$, and the micro-scale surface roughness $\sigma_{I,\mu}$ in the image domain can be computed by $\sigma_{I,\mu} = \frac{1}{\sqrt{2\pi}s}$.

4.4 Meso-scale estimation

There are two steps in the meso-scale estimation: computing roughness (slope distribution) and texture modeling.

Roughness computation. In theory, it is possible to compute meso-scale roughness from micro-scale and overall roughness using Equation 2. However, when meso-scale roughness is small (so micro-scale roughness is almost equal to overall roughness), even a small error in the two estimates leads to a large relative error in

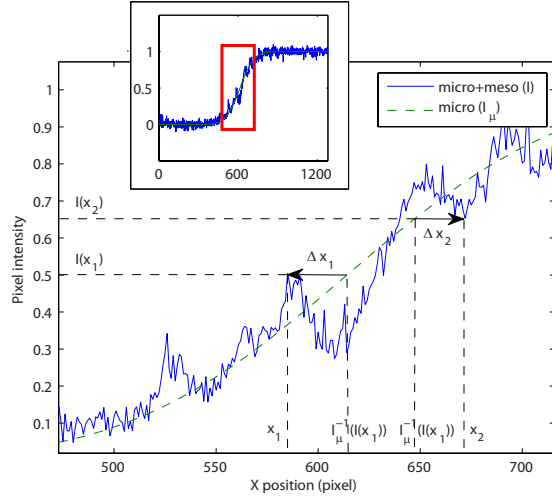


Figure 10: Pixel perturbation caused by mesostructure: this figure shows a close look at a slice of the transition region of a scrambled reflected image I (blue curve), from a surface with both micro and meso-scale roughness, and its unscrambled version I_μ (green dashed curve) from the same surface without mesostructure. $I(x)$ denotes the pixel intensity at position x , and its corresponding position in I_μ is computed by the inverse function $I_\mu^{-1}(I(x))$.

the computed meso-scale roughness. Therefore, we instead estimate meso-scale roughness directly by using an image perturbation model, and then adjust all three estimates to conform with Equation 2.

Consider a surface with only micro-scale roughness: by adding mesostructure to the surface, the pixels are scrambled to their new positions; however, since we already know the micro-scale roughness $\sigma_{I,\mu}$ (estimated in Section 4.3), the position of a given pixel intensity in the unscrambled image can be computed by applying the inverse function of Equation 5, where $\sigma_I = \sigma_{I,\mu}$. A one-dimensional illustration is shown in Figure 10: the amount of shift $\Delta x = x - I_\mu^{-1}(I(x))$ shows the perturbation in the image domain caused by the mesostructure, and by the argument in Section 3.4 we know this perturbation follows a Gaussian distribution, whose standard deviation is the meso-scale roughness $\sigma_{I,m}$. Therefore, we can compute $\sigma_{I,m}$ by estimating the variance of Δx over all pixels:

$$\sigma_{I,m} = \sqrt{\text{Var}(\Delta x)}$$

In practice, the inverse function is very sensitive in the dark and bright regions (where its slope is close to infinity), so we only compute Δx in the transition region. We set the width of the region to $\sigma_{I,\mu}$, since we know it is proportional to micro-scale roughness.

After having all three scales of roughness, we need to adjust them to satisfy Equation 2. We find the adjustments $\Delta M, \Delta m, \Delta \mu$ for each scale as the solution to the following constrained minimization problem:

$$\begin{aligned} \arg \min_{\Delta M, \Delta m, \Delta \mu} & \left(\frac{\Delta M}{\sigma_{I,M}} \right)^2 + \left(\frac{\Delta m}{\sigma_{I,m}} \right)^2 + \left(\frac{\Delta \mu}{\sigma_{I,\mu}} \right)^2 \\ \text{s.t.} & (\sigma_{I,M} + \Delta M)^2 = (\sigma_{I,m} + \Delta m)^2 + (\sigma_{I,\mu} + \Delta \mu)^2 \end{aligned}$$

The solution to this problem represents the minimum adjustments (relative to the magnitude of the initial estimates) which satisfy the constraint.

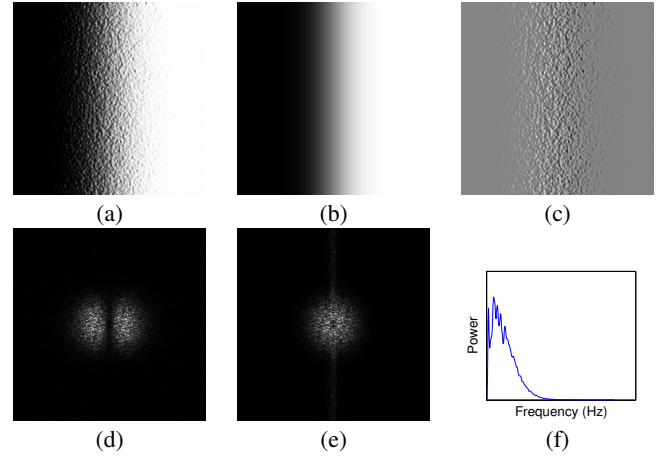


Figure 11: Acquiring mesostructure texture properties. (a) Step-edge reflection. (b) Step-edge reflection of the same surface with only micro-scale roughness. (c) Intensity change due to mesostructure (i.e., (a) minus (b)). (d) Fourier transform of (c). (e) Reconstructed Fourier transform by dividing (d) with $i\omega$. (f) 1D power spectrum computed by averaging over concentric rings of (e).

Texture modeling. We describe the meso-scale texture by a 1D frequency spectrum computed by averaging over concentric rings in the 2D spectrum of the mesostructure height map. This 1D spectrum records the frequency content without the orientation content; the idea is related to the ring filters used in texture analysis [Tuceryan and Jain 1993; Randen and Husoy 1999].

As with the roughness computation, we consider the image reflected from a surface with only micro-scale roughness; the intensity change, which is caused by adding the mesostructure, can then be found by subtracting it from the original reflected image, as shown in Figure 11(c). Now, rather than using the inverse function to find the perturbation Δx in the image domain, we use a first-order Taylor approximation:

$$\begin{aligned} \Delta I(x, y) &= I(x + \Delta x, y) - I(x, y) \\ &\approx \Delta x \cdot \frac{\partial I}{\partial x}(x, y) \\ &= \Delta x \cdot G_{\sigma_{I,\mu}}(x) \end{aligned}$$

where $G_{\sigma_{I,\mu}}(x)$ is a Gaussian with standard deviation $\sigma_{I,\mu}$ (i.e., the derivative of the error function representing the micro-scale reflected image). Next, by the result in Appendix I, the horizontal half-vector angle θ_h is approximately proportional to the image position x on the X-axis. Therefore, the perturbation of this half-vector angle (due to mesostructure), $\Delta \theta_h$, is also proportional to the perturbation in image position Δx . In addition, by using the small angle approximation again, $\Delta \theta_h$ is equal to the partial derivative of the height map h in the x -direction. Combining all of these, we have

$$\Delta I(x, y) \propto \frac{\partial h}{\partial x}(x, y) \cdot G_{\sigma_{I,\mu}}(x) \quad (6)$$

Using the rule for the Fourier transform of the derivative of a function, the Fourier transform of the intensity change image (Figure 11(d)) is

$$\hat{\Delta I}(\omega, \phi) \propto [i\omega \hat{h} * \hat{G}_{\sigma_{I,\mu}}](\omega, \phi)$$

where $\hat{\Delta I}$, \hat{h} , and $\hat{G}_{\sigma_{I,\mu}}$ are the Fourier transforms of the corre-

Test case	$\sigma_{I,M}$	$\sigma_{I,\mu}$	$\sigma_{I,m}$
Metal bookend	109.83	91.31	61.02
Synthesized image	113.44	94.69	62.48
Beige bookend	170.88	158.93	62.78
Synthesized image	172.87	154.48	77.61
Black folder	21.35	15.97	14.17
Synthesized image	22.22	17.05	14.26

Table 1: Correctness test. *For each surface, the first row shows initial recovered parameters for that captured surface, and the second row shows result of rerunning the estimation process on images synthesized with recovered parameters.*

sponding functions³, and $*$ denotes convolution.

This relationship shows that we can compute (up to a scale factor) a Gaussian filtered Fourier transform of the mesostructure (the height map) by dividing the Fourier transform of the intensity change image by $i\omega$. The multiplication with the horizontal Gaussian function in Equation 6 becomes a convolution in the Fourier domain, which suggests the horizontal resolution of the power spectrum we can compute is limited by the micro-scale roughness $\sigma_{I,\mu}$; as the micro-scale roughness decreases, the support of the Gaussian function becomes larger in the frequency domain. In our tests, however, this is not a significant problem since we average over concentric rings in the power spectrum, which utilizes both vertical and horizontal frequencies. Our first-order approximation in Equation 6 also breaks down when the micro-scale roughness is small. In practice, this introduces high-frequency noise into our estimation of meso-scale texture; however, important characteristics of the texture are still captured, and we are still able to acquire visually reasonable matches (see Figure 12). Figure 11 illustrates our meso-scale texture acquisition procedure.

5 Results

We now demonstrate the performance of our acquisition method, through (1) a correctness test, and by (2) comparing our captured and synthesized results to real-world photographs of surfaces. All renderings described in this section were created with PBRT v2 [Pharr and Humphreys 2010]; the synthesized bump maps were generated by using the 1D power spectrum to rebuild a radially symmetric 2D Fourier transform and then randomizing its phase. The system then adjusts the magnitude of the bump map according to the estimated meso-scale roughness (i.e., the estimated slope distribution).

Correctness test. We first evaluate the correctness of our method by iteratively running our algorithm; we synthesize a surface using estimated roughness parameters, render this surface in PBRT under step-edge illumination, then run our estimation algorithm again on the rendered image. We apply this test on the step-edge-illuminated images of three real surfaces with various overall roughness. Table 1 shows the results, which demonstrate that our method can acquire parameters from synthesized images with relatively high accuracy: most of the estimates in all scales have less than 5% error. The large error in the estimated meso-scale roughness of the beige bookend is because of a less accurate estimate of the diffuse component due to high overall roughness; such high roughness makes the appearance of the specular component more similar to the diffuse component. Nevertheless, the meso-scale roughness is relatively small

³Here we compute the 2D Fourier transform of the 1D Gaussian function by defining $G_{\sigma_{I,\mu}}(x, y) = G_{\sigma_{I,\mu}}(x)$.

compared to its overall and micro-scale roughness, and thus has less effect on the surface appearance.

Comparison with real photos. In order to judge how well our system can recover parameters that yield an appearance match to a target surface, we have used it to capture a number of surfaces, and rerendered synthesized versions under two different conditions; for each condition, we qualitatively compare to a ground truth captured photo. The first condition is the same viewpoint and illumination used for acquisition (i.e., step-edge lighting, see columns 1 and 2 of Figure 12). The second is from a novel, more oblique viewpoint, illuminated by the screen of an Apple iPhone placed perpendicular to the surface (columns 3 and 4 of Figure 12). This novel viewpoint observes the surface from a larger distance, and shows the varying blurriness of the reflection across the image due to varying surface-to-light distance. In all tests, the position of the camera (and the iPhone, when present) was calibrated using the Bouguet toolkit [Bouguet 2010], so that the scene configuration could be reproduced in our renderer. Note that our acquisition system currently uses a grayscale camera, so no color information is estimated; for visualization, we manually tuned the color of the renderings in the novel view.⁴ In addition, we tune the absolute albedo of the surface manually (this could be recovered by using a calibration target with known albedo).

In general, the results show that our method often achieves a good appearance match. The overall acquired roughness can be judged by the widths of the dark-to-bright transition regions (under step-edge illumination), and by the level of blurriness of the iPhone LCD reflection (in the novel viewpoint) and is qualitatively similar in the captured and rendered views in each case, from the mirror-like whiteboard to the fairly rough beige bookend. The reconstructed mesostructures (the “bumpiness”) qualitatively capture the appearances of real surfaces for a range of bump size distributions (compare the characteristic scrambling effect in the transition regions for the step-edge images, and the ragged edges of the reflected iPhone display). Some differences between the two sets of images are apparent, however. In some cases (e.g., the two bookends), there is more contrast evident in the meso-scale bumps in the rendered images, which could partly be attributed to the perfect optics of the synthetic camera used for rendering, but also suggests that the energy in the meso-scale roughness is slightly overestimated. There are two causes for the bias: first, in some cases the diffuse component is over-estimated, which increases the contrast of the normalized image (Figure 7(c)) and leads to higher roughness; second, pixel-level noise from the sensor also tends to increase the estimated roughness. Both causes, however, have a larger effect on the meso-scale estimation than on the micro-scale and overall roughness. This is because the estimations later two estimates are less susceptible to the increased pixel variance due to their statistical nature.

There are also differences in the diffuse reflections, which can be observed in Figure 1, or the uncropped high-resolution images in the supplementary material (available using the links at the start of this paper). While these again suggest that the diffuse reflectance may be slightly misestimated, this effect is exacerbated by the fact that real LCDs are not perfectly diffuse area light sources (i.e., radiance is dependent on viewing angle)⁵. Despite these differences, our approach often achieves a good match, especially considering the small amount of input data we require.

Figure 13 shows two test cases where our method failed (a metal

⁴In the renderings, we only tuned the diffuse color; the specular color is assumed to be white. Estimating colored reflection would involve using a color camera as well as some simple modifications to our system.

⁵In our renderings, we used a cosine model to approximate this behavior.

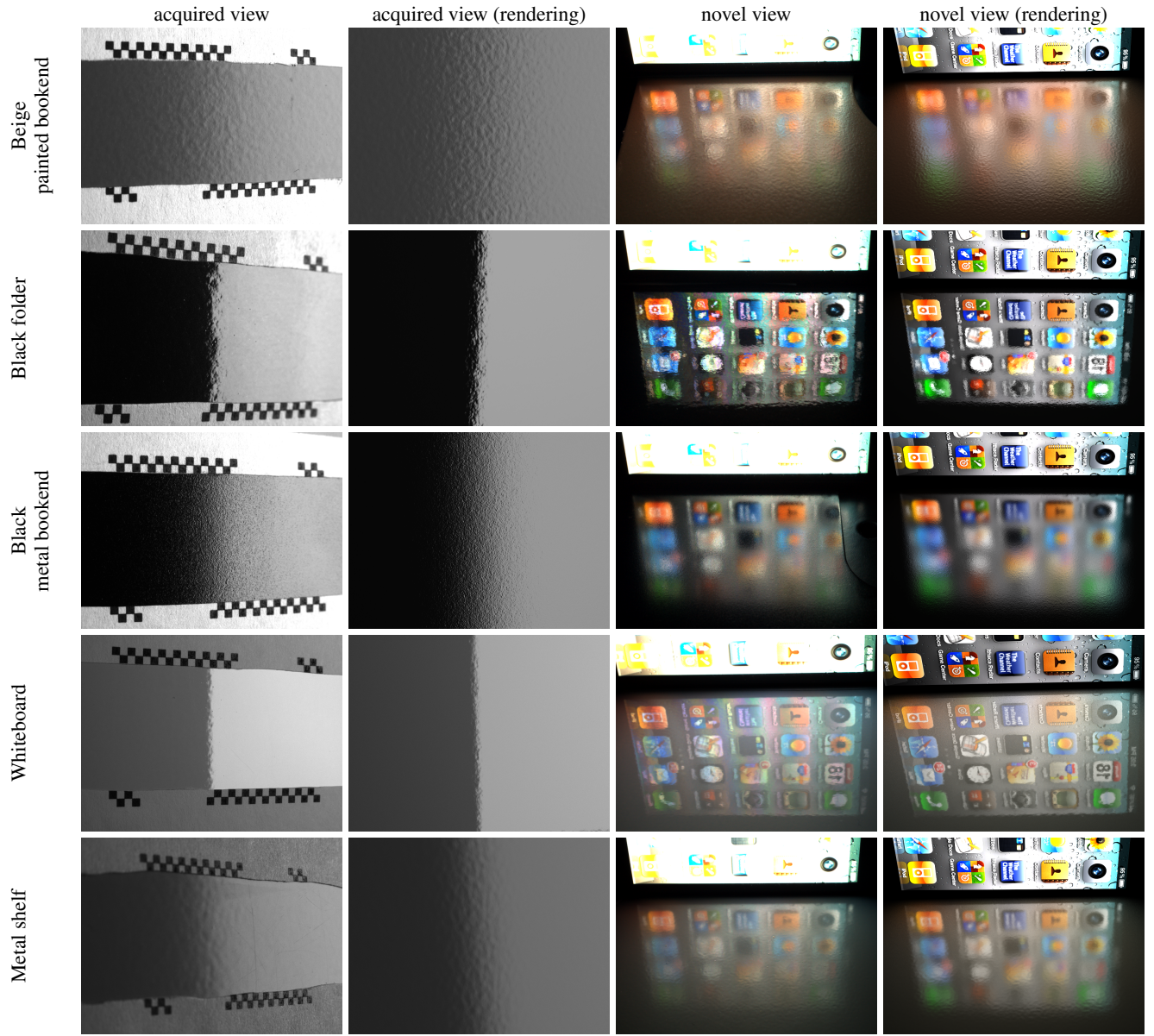


Figure 12: Results. Each row shows a different acquired surface. The first column shows the captured image under step-edge lighting used for parameter estimation. The second column shows a rendering of a patch of the surface synthesized using the acquired parameters. Column three shows a photograph of the surface from a new viewpoint, illuminated by an iPhone LCD. Column four shows the synthesized surface rendered under the same conditions. Please zoom into this figure to see the detail of the mesostructure, or see the supplemental material (available using the links at the start of this paper) for high-resolution images.

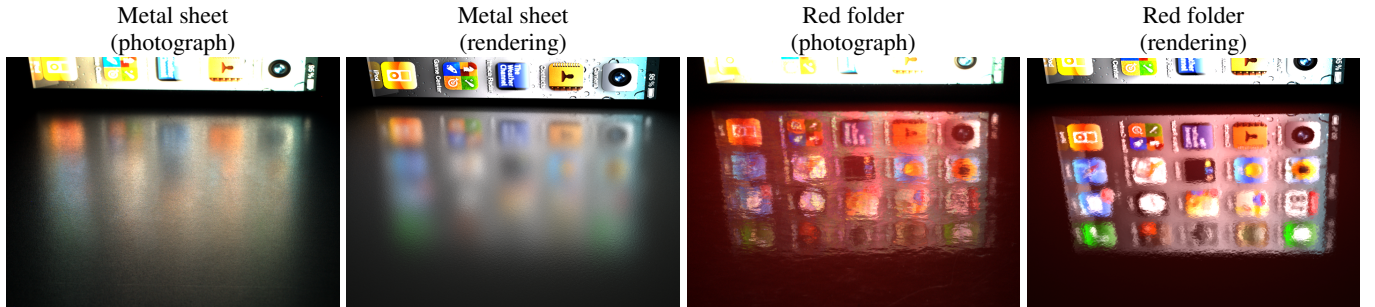


Figure 13: Failure cases. Left columns: an extremely rough metal sheet (photograph and rendering). Right columns: a red folder with anisotropic mesostructure (photograph and rendering).

sheet, and a red folder). For the metal sheet, the mesostructure is extremely dense and strong, and the resulting large slope variation breaks the small angle approximation in Equation 6 (and the potential for attached shadows from the meso-scale bumps could also be a problem). The strong meso-scale roughness yields the elongated reflection towards the bottom of the image, not present in our rendering. In the case of the red folder, the mesostructure is anisotropic: many horizontal stripes are evident in the captured image. Because this violates our assumption of isotropy, our surface model cannot produce a good fit.

6 Conclusions and Future Work

We have presented a simple end-to-end capture system for achieving a qualitative appearance match with a range of glossy, bumpy surfaces. Our main contributions are a dual-level statistical model of surface appearance, and a lightweight capture system using step-edge illumination that goes hand-in-hand with this model. We believe our work represents a promising advance towards lightweight appearance capture; while we present a prototype system, we believe that it could also be implemented on a more portable setup, such as an iPad with a front-facing camera, where the geometry and LCD emission can be precalibrated.⁶ There are also a number of other avenues for future work. Currently, absolute albedoes are not estimated; the estimation can be incorporated into surface calibration (e.g., using a square marker with known albedo). Separating diffuse and specular reflection can be difficult for very rough objects, and so a more robust method would be desirable. To handle materials that are anisotropic (at micro-scale, meso-scale, or both), one could capture a second step edge perpendicular to the first (or display both at once with a 2×2 checkerboard pattern). Finally, step edges (or near step edges) at a variety of orientations are common in natural illumination (e.g., due to occlusions or hard shadows), and an interesting extension would be to estimate the appearance of a planar patch from everyday lighting alone.

Acknowledgements

We would like to thank Kevin Matzen for assistance in data acquisition and creating figures in this paper, and all reviewers for their helpful comments and suggestions. This work was supported by the National Science Foundation (grants IIS-0964027 and IIS-1011919), MIT Lincoln Labs, and the Intel Science and Technology Center for Visual Computing.

References

BOUGUET, J., 2010. Camera calibration toolbox for Matlab. http://www.vision.caltech.edu/bouguetj/calib_doc/.

CHEN, T., GOESEL, M., AND SEIDEL, H. P. 2006. Mesostructure from specularity. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1825–1832.

COOK, R. L., AND TORRANCE, K. E. 1981. A reflectance model for computer graphics. *Computer Graphics (Proc. SIGGRAPH)* 15, 3, 307–316.

DANA, K. J., VAN GINNEKEN, B., NAYAR, S. K., AND KOENDERINK, J. J. 1999. Reflectance and texture of real-world surfaces. *ACM Trans. Graphics* 18, 1, 1–34.

DONG, Y., WANG, J., TONG, X., SNYDER, J., LAN, Y., BEN-EZRA, M., AND GUO, B. 2010. Manifold bootstrapping for SVBRDF capture. *ACM Trans. Graphics* 29, 4.

DROR, R. O., ADELSON, E. H., AND WILLSKY, A. S. 2001. Recognition of surface reflectance properties from a single image under unknown real-world illumination. In *Proc. IEEE Workshop on Identifying Objects across Variations in Lighting*.

FRANCKEN, Y., CUYPERS, T., AND BEKAERT, P. 2008. Mesostructure from specularity using gradient illumination. In *Proc. Int. Workshop on Projector Camera Systems*.

FRANCKEN, Y., CUYPERS, T., MERTENS, T., GIELIS, J., AND BEKAERT, P. 2008. High quality mesostructure acquisition using specularities. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*.

GALERNE, B., GOUSSEAU, Y., AND MOREL, J. 2011. Random phase textures: Theory and synthesis. *IEEE Trans. Image Processing* 20, 1, 257–267.

GARDNER, A., TCHOU, C., HAWKINS, T., AND DEBEVEC, P. 2003. Linear light source reflectometry. In *Proc. SIGGRAPH*, 749–758.

GEORGHIADES, A. S. 2003. Recovering 3-D shape and reflectance from a small number of photographs. In *Proc. Eurographics Workshop on Rendering*, 230–240.

GHOSH, A., CHEN, T., PEERS, P., WILSON, C. A., AND DEBEVEC, P. 2009. Estimating specular roughness and anisotropy from second order spherical gradient illumination. *Computer Graphics Forum* 28, 4, 1161–1170.

GHOSH, A., CHEN, T., PEERS, P., WILSON, C. A., AND DEBEVEC, P. 2010. Circularly polarized spherical illumination reflectometry. *ACM Trans. Graphics* 29, 6, 162–173.

HAN, C., SUN, B., RAMAMOORTHY, R., AND GRINSPUN, E. 2007. Frequency domain normal map filtering. *ACM Trans. Graph.* 26, 3.

JOHNSON, M. K., AND ADELSON, E. H. 2009. Retrographic sensing for the measurement of surface texture and shape. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1070–1077.

JULESZ, B. 1962. Visual pattern discrimination. *IRE Trans. Information Theory* 8, 2, 84–92.

KAUTZ, J., VÁZQUEZ, P.-P., HEIDRICH, W., AND SEIDEL, H.-P. 2000. A unified approach to prefiltered environment maps. In *Proc. Eurographics Workshop on Rendering*, 185–196.

MARSCHNER, S. R., WESTIN, S. H., LAFORTUNE, E. P., TORRANCE, K. E., AND GREENBERG, D. P. 1999. Image-based BRDF measurement including human skin. In *Proc. Eurographics Workshop on Rendering*, 139–152.

NGAN, A., DURAND, F., AND MATUSIK, W. 2005. Experimental analysis of BRDF models. In *Proc. Eurographics Symposium on Rendering*, 117–126.

PATERSON, J. A., CLAUS, D., AND FITZGIBBON, A. W. 2005. BRDF and geometry capture from extended inhomogeneous samples using flash photography. *Computer Graphics Forum* 24, 3, 383–391.

PERLIN, K. 1985. An image synthesizer. *Computer Graphics (Proc. SIGGRAPH)* 19, 3, 287–296.

⁶Some limitations apply to the camera: it needs to be able to focus on short distance, and must have acceptable image quality.

- PHARR, M., AND HUMPHREYS, G. 2010. *Physically Based Rendering: From Theory to Implementation*, 2nd ed. Morgan Kaufmann.
- RAMAMOORTHY, R., AND HANRAHAN, P. 2001. A signal-processing framework for inverse rendering. In *Proc. SIGGRAPH*, 117–128.
- RANDEN, T., AND HUSOY, J. H. 1999. Filtering for texture classification: A comparative study. *IEEE Trans. Pattern Analysis and Machine Intelligence* 21, 4, 291–310.
- ROMEIRO, F., AND ZICKLER, T. 2010. Blind reflectometry. In *Proc. European Conf. Computer Vision*. Springer, 45–58.
- TORRANCE, K. E., AND SPARROW, E. M. 1967. Theory for off-specular reflection from roughened surfaces. *Journal of the Optical Society of America* 57, 9, 1105–1112.
- TUCERYAN, M., AND JAIN, A. K. 1993. Texture analysis. In *Handbook of Pattern Recognition and Computer Vision*. World Scientific, ch. 2.1.
- WANG, J., AND DANA, K. J. 2006. Relief texture from specularities. *IEEE Trans. Pattern Analysis and Machine Intelligence* 28, 3, 446–457.
- WARD, G. J. 1992. Measuring and modeling anisotropic reflection. *Computer Graphics (Proc. SIGGRAPH)* 26, 2, 265–272.
- WOODHAM, R. J. 1980. Photometric method for determining surface orientation from multiple images. *Optical Engineering* 19, 1, 139–144.
- X-RITE, I., 2011. MA98 portable multi-angle spectrophotometer. www.xrite.com/product_overview.aspx?ID=1148.
- YU, Y., DEBEVEC, P., MALIK, J., AND HAWKINS, T. 1999. Inverse global illumination: Recovering reflectance models of real scenes from photographs. In *Proc. SIGGRAPH*, 215–224.
- ZICKLER, T., RAMAMOORTHY, R., ENRIQUE, S., AND BELHUMEUR, P. N. 2006. Reflectance sharing: Predicting appearance from a sparse set of images of a known shape. *IEEE Trans. Pattern Analysis and Machine Intelligence* 28, 8, 1287–1302.

Appendix I

By using Equation 1 we can write the reflected image I by

$$I(x, y) = \int_{\mathcal{H}^2(\mathbf{n})} \frac{F_r(\omega_o) G(\omega_o, \omega_i) L_i(\omega_i, x, y) D(\omega_h)}{4 \cos \theta_o} d\omega_i$$

Notice that the incident illumination L_i depends on the image/surface position (x, y) since the light is not infinitely far away. Next, because the camera we use has a small FOV, we can assume the Fresnel reflectance F_r , geometrical attenuation term G , and the cosine factor $\cos \theta_o$ to be constant, hence

$$I(x, y) \propto \int_{\mathcal{H}^2(\mathbf{n})} L_i(\omega_i, x, y) D(\omega_h) d\omega_i$$

Now we look at the center row of the image, where $y = 0$. Let $I(x) = I(x, 0)$ and $L_i(\omega_i, x) = L_i(\omega_i, x, 0)$. In this row, the incident light L_i due to the point source is a delta function

$$L_i(\omega_i, x) = \delta\left(\theta_i = \tan^{-1} \frac{x}{z_L}\right) = \delta\left(\theta_i \approx \frac{x}{z_L}\right)$$

where z_L is the distance from the light to the surface in pixels. The last approximation comes from the fact that FOV is small, so we

can use the small angle approximation. Also, we know that the slope distribution function D only depends on θ_h since the surface is isotropic. To find θ_h , we need to know the viewing angle θ_o , which can again be approximated by using the small angle approximation:

$$\theta_o = \tan^{-1} \frac{x}{z_C} \approx \frac{x}{z_C}$$

where z_C is the distance from the camera to the surface in pixels. Combining the equations above, we get

$$I(x) \propto D(\theta_h) \approx D\left(\frac{\frac{x}{z_L} + \frac{x}{z_C}}{2}\right)$$

Since D is assumed to be a Gaussian, we get Equation 3 in Section 3.4.