Eyes for Relighting

Ko Nishino* Shree K. Nayar†
Department of Computer Science, Columbia University

Abstract

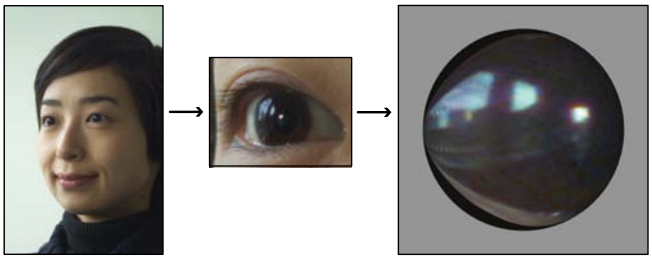
The combination of the cornea of an eye and a camera viewing the eye form a catadioptric (mirror + lens) imaging system with a very wide field of view. We present a detailed analysis of the characteristics of this corneal imaging system. Anatomical studies have shown that the shape of a normal cornea (without major defects) can be approximated with an ellipsoid of fixed eccentricity and size. Using this shape model, we can determine the geometric parameters of the corneal imaging system from the image. Then, an environment map of the scene with a large field of view can be computed from the image. The environment map represents the illumination of the scene with respect to the eye. This use of an eye as a natural light probe is advantageous in many relighting scenarios. For instance, it enables us to insert virtual objects into an image such that they appear consistent with the illumination of the scene. The eye is a particularly useful probe when relighting faces. It allows us to reconstruct the geometry of a face by simply waving a light source in front of the face. Finally, in the case of an already captured image, eyes could be the only direct means for obtaining illumination information. We show how illumination computed from eyes can be used to replace a face in an image with another one. We believe that the eye not only serves as a useful tool for relighting but also makes relighting possible in situations where current approaches are hard to use.

CR Categories: I.2.10 [Artificial Intelligence]: Vision and Scene Understanding—Intensity, color, photometry, and thresholding I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Color, shading, shadowing, and texture; I.4.1 [Image Processing and Computer Vision]: Digitization and Image Capture—Imaging geometry

Keywords: Eye, Illumination Estimation, Relighting

1 Introduction

In order to relight real objects and scenes, it is important to know their lighting conditions. This lighting information is estimated from images of the scene when the properties of the scene are known, or measured using a light probe when such a probe can be inserted into the scene. There are, however, many instances where it is not possible to estimate or measure the lighting of a scene. For instance, consider the task of inserting a virtual object into an arbitrary image. In this case, it is not possible to estimate the illumination as the precise geometry and reflectance of the scene are unknown. On the other hand, since the image has already been captured, we cannot measure the illumination using a probe. In short, obtaining the lighting of a scene from a single image remains a difficult and open problem. It turns out, however, that this problem can be solved when we have a face (and hence an eye) in the image, which is often the case with images and videos.



(a) (b) (c)

Figure 1: (a) An image of a face. (b) A magnified version of the image of the right eye of the person. (c) An environment map computed from (b). One can see the sky and buildings through the windows.

Tsumura et al.[2003] recently made the interesting observation that an eye in an image can potentially be used as a "mirrored ball" that conveys information about light sources in the environment. Based on this observation, they model the eye as a sphere and use the highlights on it to estimate the directions of three point sources. The estimated source directions are used to apply photometric stereo to the face and then relight the face. In this previous work, however, the extent to which the reflections from the eye is used is rather limited. The three point sources are located at known positions with respect to the camera and the three highlights on the eye are essentially used to determine where the face is located with respect to the camera. Despite its limited scope, the work of Tsumura et al.[2003] is important as it is the first to use eyes to recover information regarding lighting.

In this paper, we show how an image of an eye taken in a completely unstructured setting can be used to extract a dense illumination distribution. Figure 1(a) shows an image of a face. In the magnified image of the right eye of the person in Figure 1(b), one can clearly see the surrounding world reflected by the cornea of the eye. We show that an environment map of the scene with a large field of view can be computed from the eye, as shown in Figure 1(c). This environment map represents the illumination distribution of the scene with respect to the eye. In other words, the eye can be used as a natural light probe which gives us not just the directions of a few point sources but the complete distribution of the frontal illumination incident on the eye and hence the face. This distribution can be exploited in various relighting applications.

The key insight which enables us to use the eye as a light probe is that the combination of an eye and a camera viewing the eye form an uncalibrated catadioptric (mirror + lens) imaging system. We refer to this as the *corneal imaging system*. The mapping of the environment by the cornea onto the image plane is non-trivial as it depends on the geometry of the cornea and its pose with respect to the camera. Anatomical studies of the eye reveal that the shape of a normal cornea (without major defects) can be well approximated with an ellipsoid with fixed parameter values. Given an image, we first find the boundary of the cornea and then use the ellipsoidal shape model to compute its 3D coordinates and orientation with

*e-mail:kon@cs.columbia.edu
†e-mail:nayar@cs.columbia.edu

respect to the camera. This is equivalent to calibrating the corneal imaging system. Once this is done, we can compute a wide-angle view of the environment surrounding the eye.

We present a comprehensive analysis on the characteristics of the corneal imaging system. We show that the field of view of this system exceeds a hemisphere for a very large range of orientations. Note that the resolution and the dynamic range of the environment map computed from an eye are limited by those of the camera used. At first glance, these limitations might seem critical. However, our experiments demonstrate that, even when an eye is captured with relatively low resolution and dynamic range, it conveys illumination information that is useful in many practical relighting scenarios. Also, the environment map computed from an eye includes the reflection from the iris beneath the cornea; it is not an exact representation of the illumination. The separation of the reflections by the cornea and the iris is a significant problem that is beyond the scope of this paper. In the relighting examples we show, we have used eyes that are relatively dark and hence have weak iris textures. This has allowed us to directly use the environment maps computed from eyes as illumination distributions.

We demonstrate the use of eyes for relighting in several settings. We use the eye to insert an object in the scene such that its appearance is consistent with the illumination of the scene. Next, we show how the eye can be used to sample the appearance of a face by simply waving a light source in front of the face. Finally, we demonstrate how eyes can be used to replace faces in an image with different faces that are lit the same way as the original ones. The results of this paper show that the eye not only serves as a useful tool for relighting, but also makes relighting possible in situations where current approaches are hard to use.

2 Related Work

Relighting real objects has recently attracted wide attention in computer graphics. In most of the proposed methods, relighting is accomplished in two steps.

In the first step, the appearance of an object is sampled under different lighting conditions. This is done by controlling the lighting. In model-based approaches, the reflectance properties of the object are estimated from the captured images [Sato et al. 1997; Boivin and Gagalowicz 2001; Yu et al. 1999; Lensch et al. 2001]. In image-based approaches, the captured images are combined to obtain an image of the object corresponding to a desired novel lighting [Lin et al. 2001; Debevec et al. 2000; Magda et al. 2001; Matusik et al. 2002]. Alternatively, appearance sampling can be done under unknown lighting conditions and the lighting can be estimated [Marschner and Greenberg 1997; Sato et al. 2003] along with the reflectance properties and the shape [Georghiades 2003; Ramamoorthi and Hanrahan 2001; Blanz and Vetter 1999]. However, because of inherent ambiguities that result when reflectance and lighting are unknown [Ramamoorthi and Hanrahan 2001], it is preferable to know the lighting conditions.

In the second step, the estimated parameters are used to synthesize an image of the object under a novel lighting condition. The novel lighting is usually the measured illumination distribution of another scene in which one wants to insert the relit object. Images of the environment of this new scene [Szeliski and Shum 1997; Teller et al. 2003] or images of a light probe (such as a mirrored sphere) placed in the scene [Debevec 1998] are commonly used for this purpose. It is important to note that finding lighting conditions is critical to both steps of the relighting process.

A large body of research has been done on various aspects of the eye in different fields. In computer graphics and human-computer interaction, research has mainly focused on using gaze direction as an interface (for examples, see [Bolt 1982; Hutchinson et al. 1989; Jacob 1990]), identify people from their iris textures [Flom and Safir 1987; Daugman 1993] and photorealistic synthesis of the appearance of an eye [Lefohn et al. 2003; Sagar et al. 1994] or a scene seen through an eye [Mostafawy et al. 1997; Barsky et al. 2002].

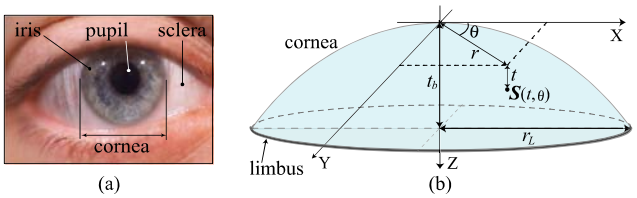


Figure 2: (a) An external view of the human eye. (b) A normal adult cornea can be modeled as an ellipsoid whose outer limit corresponds to the limbus. The eccentricity and radius of curvature at the apex can be assumed to be known.

Tsumura et al.[2003] model the eyeball as a sphere and use the highlights on it to estimate the directions of three point sources at known locations with respect to the camera. It is important to note that in our case the illumination can be arbitrary. We explore in detail the information regarding illumination that can be recovered from the image of an eye taken in a general setting. In addition, we demonstrate the use of an eye as a light probe in several relighting scenarios.

3 Corneal Imaging System

In this section, we show how the surrounding world is captured by an eye in an image. The combination of the cornea of an eye and a camera viewing it form a catadioptric (mirror + lens) imaging system. We study the characteristics of this imaging system, to understand how and when the eye can be used as a light probe.

3.1 Geometric Model of the Cornea

As shown in Figure 2(a), the external appearance of an eye is primarily due to the cornea and the sclera. The cornea consists of a lamellar structure of submicroscopic collagen fibrils arranged in a manner that makes it transparent and its surface very smooth [Kaufman and Alm 2003]. In addition, it has a thin film of tear fluid on it. As a result, the surface of the cornea behaves like a mirror. As a result, the combination of the cornea and a camera viewing it can be considered to be a catadioptric imaging system consisting of a mirror and a lens [Baker and Nayar 1999; Yagi and Yachida 1991]. We will refer to this as the *corneal imaging system*.

In the fields of physiology and anatomy, extensive measurements of the shape and dimensions of the cornea have been conducted [Kaufman and Alm 2003; von Helmholtz 1909]. It has been found that a normal adult cornea is very close to an ellipsoid and its parameters do not vary much across people. In Cartesian coordinates (x,y,z), an ellipsoid can be written as [Baker 1943]

$$pz^2 - 2Rz + r^2 = 0\,, \tag{1}$$

where $r = \sqrt{x^2+y^2}$, $p = 1-e^2$ where e is the eccentricity, and R is the radius of curvature at the apex of the ellipsoid. Now, a point $\mathbf{S}$ on the corneal surface can be expressed as

$$\mathbf{S}(t,\theta) \quad = \quad (\sqrt{-pt^2+2Rt}cos\theta, \sqrt{-pt^2+2Rt}sin\theta, t)\,, \tag{2}$$

where $0 \le \theta < 2\pi$ (see Figure 2(b)). It is known that on average, the eccentricity e is 0.5 ($p = 0.75$) and the radius of curvature R at the apex is 7.8 mm [Kaufman and Alm 2003].

In our work, we assume normal adult corneas and approximate their shape with the above ellipsoidal model and parameter values. This approximation is widely used and suffices our purposes. However, when the shape of a cornea significantly deviates from a normal one, for instance due to diseases such as keratoconus, it will be necessary to measure its shape. Corneal shape measurement including its precise curvature can be done by using structured light [Barsky 2003; Halstead et al. 1996]. Note that such a measured shape can also be directly used in our method.

The boundary between the cornea and the sclera is called the *limbus*. The sclera is not as highly reflective as the cornea. As a result, the limbus defines the outer limit of the reflector of our imaging system. From a physiological perspective, the cornea "dissolves" into the sclera. In the case of an adult eye, the limbus has been found to be close to circular with radius r_L of approximately 5.5 mm. Therefore, in equation (2), the parameter t ranges from 0 to t_b where t_b is determined from equation (1) to be 2.18 mm.

3.2 Self-Calibration

Using the geometric model of the cornea, we compute the 3D coordinates and orientation of the cornea in the camera's coordinate frame from a single image. This is equivalent to calibrating the corneal imaging system. It turns out that this can be done by locating the limbus in the image.

The limbus in an image is the projection of a circle in 3D space. Even in the extreme case when the gaze direction is perpendicular to the optical axis of the camera, the depth of this circle is only $2r_L = 11$ mm. Hence, the camera projection model can be assumed to be weak-perspective; orthographic projection followed by scaling. Then, the limbus is imaged as an ellipse.

Let (u,v) denote the horizontal and vertical image coordinates, respectively. An ellipse in an image can be described using five parameters which we denote by a vector $\mathbf{e}$. These parameters are the center (c_u, c_v), the major and minor radii r_{max} and r_{min} and the rotation angle ϕ of the ellipse in the image plane. We localize the limbus in an image $I(u,v)$ by searching for the ellipse parameters $\mathbf{e}$ that maximize the response to an integro-differential operator applied to the image smoothed with a Gaussian g_σ along the radial directions. This operator can be expressed as

$$|g_\sigma(r_{max}) * \frac{\partial}{\partial r_{max}} \oint_{C_\mathbf{e}} I(u,v)ds + g_\sigma(r_{min}) * \frac{\partial}{\partial r_{min}} \oint_{C_\mathbf{e}} I(u,v)ds|, \tag{3}$$

where the integrals are computed along the perimeter of the ellipse ($C_\mathbf{e}$). This is an extension of the integro-differential operator introduced by Daugman [1993] to detect the limbus and the pupil as circles in an image of a forward-looking eye. In contrast, our algorithm detects the limbus as an ellipse, which is necessary when the gaze direction of the eye is unknown and arbitrary.

In order to estimate the ellipse parameters ($\mathbf{e}$) that maximize equation (3), we first provide an initial estimate by drawing an ellipse in the image. This is accomplished with a simple user interface – clicking to locate and dragging to change the size of a predefined ellipse. We also roughly specify the arc range of this initial ellipse to be used to evaluate equation (3). This avoids misinterpreting eyelids as the limbus. We found that this initial estimate and arc range can be quite rough for most cases since the limbus almost always causes a large intensity gradient. Given this initial estimate, we use simplex method to search for the ellipse parameters that maximize the response of equation (3). Instead of the rather brute force simplex method, other gradient based optimization methods can be also used to efficiently estimate the limbus parameters.

Under weak-perspective projection, the major axis of the ellipse detected in the image corresponds to the diameter of the circular limbus in 3D space. Therefore, the (average) distance d of the limbus from the camera can be found as

$$d = r_L \frac{f}{r_{max}}, \tag{4}$$

where $r_L = 5.5$mm, f is the focal length in pixels and r_{max} is known from the detected limbus. The depth d, the ellipse center (c_u, c_v) in the image and the focal length f determine the 3D coordinates of the center of the limbus.

From the image parameters of the limbus we can also compute the 3D orientation of the cornea. The 3D orientation is represented using two angles (ϕ, τ). ϕ is the rotation of the limbus in the image plane, which we have already estimated. Consider the plane in 3D on which the limbus lies. τ is the angle by which this plane is tilted with respect to the image plane and can be determined from the major and minor radii of the detected ellipse:

$$\tau = \arccos \frac{r_{min}}{r_{max}} . \tag{5}$$

Note, however, that there is an inherent two-way ambiguity in the estimate of τ; for instance, an eye looking downward and upward by the same amount will produce the same ellipse in the image. In our current implementation, we manually break the ambiguity.

We have conducted extensive experiments on the accuracy of the estimated corneal coordinates and orientations [Nishino and Nayar 2004]. The results show that the depth is estimated with 1.9% RMS error and the orientation angles (ϕ, τ) are estimated with 3.9° and 4.5° RMS errors, respectively. These errors are small given that the limbus is inherently a fuzzy boundary.

3.3 Field of View and Resolution

We now analyze the geometric characteristics of the corneal imaging system. As shown in Figure 3, let us denote the surface normal of a point on the cornea $\mathbf{S}(t,\theta)$ with $\mathbf{N}(t,\theta)$ and the camera pupil with $\mathbf{P}$. Then, the incident light ray direction $\mathbf{V}^i(t,\theta)$ at the point $\mathbf{S}(t,\theta)$ can be derived using the law of reflection as

$$\mathbf{V}^i = \mathbf{V}^r - 2\mathbf{N}(\mathbf{N}\cdot\mathbf{V}^r) , \quad \mathbf{V}^r = \frac{\mathbf{P}-\mathbf{S}}{|\mathbf{P}-\mathbf{S}|} . \tag{6}$$

Since the cornea is a convex ellipsoid, the field of view (FOV) of the corneal imaging system is bounded by the incident light rays that are reflected by the outer limit (limbus) of the cornea. Let us define this FOV on a unit sphere. As we traverse the circular limbus, $\mathbf{V}^i(t,\theta)$ forms a closed loop on the unit sphere. The surface area of the unit sphere within this closed loop is the FOV (in steradians) and can be determined as:

$$\int_0^{2\pi}\int_0^{\arccos V_z^i(t_b,\theta)} sin\varphi d\varphi d\theta = \int_0^{2\pi} (-V_z^i(t_b,\theta)+1)d\theta.$$

Here, V_z^i is the Z-coordinate of $\mathbf{V}^i$ and θ and φ are azimuth and polar angles defined in the coordinate frame of the unit sphere.

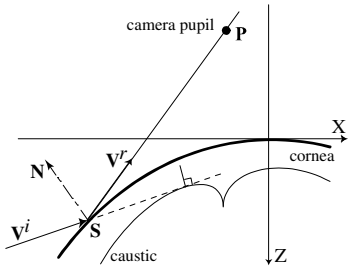


Figure 3: Geometry of the corneal imaging system.

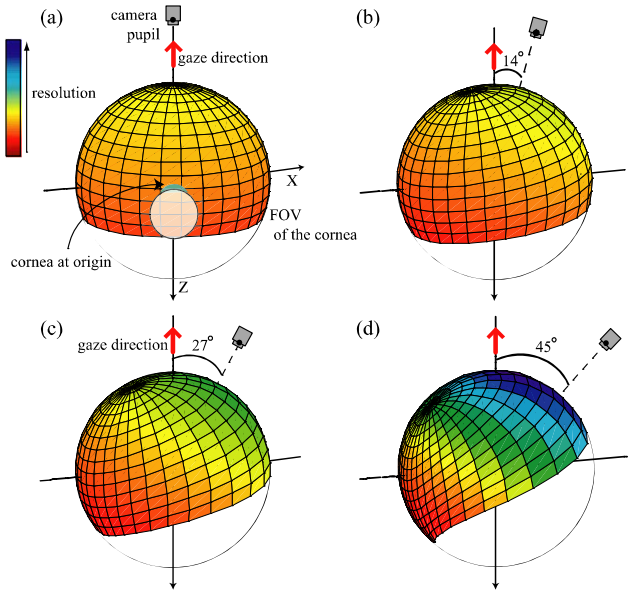


Figure 4: The field of view and the spatial resolution of the corneal imaging system for different eye-camera configurations. The shaded regions on the spheres represent the field of view. The colors within the shaded regions represent the spatial resolution (increases from red to blue). The corneal system captures a wide-angle view of the environment with highest resolution around the gaze direction.

The fields of view for different eye-camera configurations are shown in Figure 4. Note that the FOV is larger than 2π in all these cases. These results show that we generally obtain more than a hemispherical FOV of the environment from an eye in an image. Note that the FOV is more or less centered around the gaze direction. In other words, we can always recover most of the *frontal* illumination surrounding the eye. In practice, however, other facial parts (such as the nose, eyelids and eyelashes) may partially occlude the FOV and hence the actual FOV will be less than the ones depicted in Figure 4.

For gaze directions that are extreme with respect to the camera, the camera may not see the entire extent of the cornea. We are neglecting such self-occlusions here which occurs when the angle between the gaze direction and the camera exceeds roughly 40 degrees. Also because of this self-occlusion, limbus localization has to be accomplished relying on a single arc. As a result, self-calibration can become unstable when the the angle significantly exceeds 40 degrees. We found that, in many cases, the limbus can be reliably localized to the extent where its orientation exceeds this angle for 5 to 10 degrees.

The spatial resolution of the corneal imaging system can be defined as the ratio of the solid angle subtended by an infinitesimal area on the image plane from the camera's pupil to the solid angle subtended by the corresponding scene area from the virtual viewpoint inside the cornea. The derivation of the resolution is fairly lengthy as it involves computing the caustics [Cornbleet 1984; Burkhard and Shealy 1973] of the corneal imaging system (details are given in [Nishino and Nayar 2004]). In Figure 4, the colors shown within each field of view represent the spatial resolution. Here, resolution increases from red to blue. Notice how the resolution increases around the gaze direction (the optical axis of the cornea[1]). Interestingly, regardless of the eye-camera configuration, the corneal imaging system captures the surrounding environment

[1]Strictly speaking, the gaze direction is slightly offset from the optical axis of the eye (the center of the fovea does not lie along the optical axis) [Kaufman and Alm 2003].

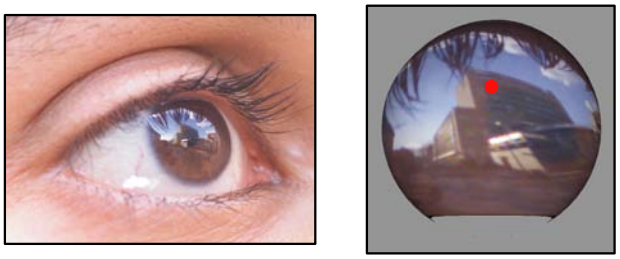


Figure 5: From the eye in an image (left), we can compute a wide-angle environment map of the scene (right). The red dot depicts the gaze direction.

with a resolution function that is similar in form to that of the retina of an eye; highest around where the person is looking and lower towards the periphery.

4 Illumination from the Eye

Once the cornea's location and orientation are computed in the camera coordinate frame as described in Section 3.2, we can trace back each light ray $\mathbf{V}^r$ that enters the camera pupil via the corneal surface using equation (6). As a result, a spherical environment map can be computed. Figure 5 shows an image of an eye and the computed environment map (see also Figure 1). As predicted by the analysis, the environment map captures more than half of the entire illumination distribution of the scene with respect to the eye.

We used a 6 mega-pixel Kodak DCS 760 camera to capture the eye image in Figure 5[2]. Roughly speaking, in order to obtain a 1° angular resolution in the frontal half of the environment map computed from an eye, we need $180 \times 180 = 32,400$ pixels within the eye. This is roughly 0.5% of the entire image. Therefore, currently available high-end cameras already provide us with the spatial resolution needed for our purposes. The dynamic range of the recovered environment map is also limited by that of the camera. To obtain precise radiance values of the illumination, one can enhance the effective dynamic range of the camera by combining multiple images of the eye taken with different exposures [Debevec and Malik 1997; Mitsunaga and Nayar 1999; Nayar and Mitsunaga 2000]. In our experiments, we have not used this approach but rather relied on single images. Since the cornea has a very low reflectivity (less than 1% [Kaufman and Alm 2003]), after reflection it compresses the dynamic range of the illumination in the scene. This can actually help us capture both objects (such as the face) and the bright illumination of the scene in a single image with a limited dynamic range[3]. As a result, the images we used gave us illumination distributions with adequate quality for our relighting scenarios[4]. However, whenever possible, we advocate using images with higher spatial resolution and dynamic range.

Note that the pixel values of the computed environment map include reflections from the iris as well. Strictly speaking, this must be subtracted from the environment map to obtain the correct illumination colors and brightnesses. The problem of separating reflections from the iris and the cornea is, however, significant by itself and will be pursued in future work. In our relighting examples, we used dark colored eyes and directly used the environment maps computed from them as the illumination distributions.

[2]The focal lengths of this camera and the digital video camera used in the later examples were estimated using the method in [Stein 1995].

[3]Note that this is not the case with high quality mirrors which have almost 100% reflectivity.

[4]We do not attempt to estimate accurate radiance values of the illumination. The unknown scaling factor due to the reflectivity of the cornea will simply, for instance, scale the albedo values of the face in Section 5.2

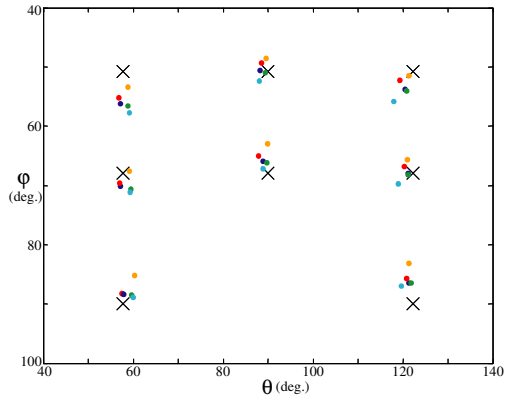


Figure 6: The estimated point source directions in spherical coordinates. The 8 ground truth directions are shown as black crosses. The colored dots are the estimates for the 5 subjects. The errors are less than 3° on average.

We evaluated the directional accuracy of the illumination distributions computed from eyes. We used a projector to project small bright patches on a screen at 8 known positions to simulate 8 different point source directions. Five subjects were asked to sit in front of the screen and face it. An image of an eye of each of the subjects was taken. The source locations (highlights) were manually located in each of the eye images and the source directions were computed. Figure 6 shows the estimated source directions (a different color is used for each person). The black crosses depict the ground truth directions. The RMS errors in the azimuth (θ) and polar (φ) angles were 1.56° and 3.13°, respectively. These results show that we can estimate the illumination distribution from an eye with reasonable accuracy.

5 Relighting using Eyes

We now demonstrate the use of the eye as a light probe for relighting. We present results for three different relighting scenarios.

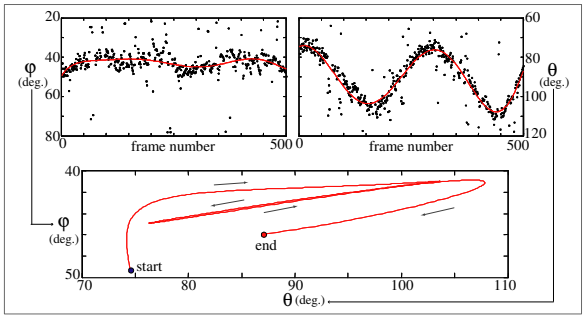
5.1 Inserting Virtual Objects

Using the illumination distribution computed from an eye, we can superimpose a virtual object into the scene such that it appears consistent with the light of the scene. We are of course assuming here that the lighting around the object's location is similar to the light around the eye. We captured a video of a person by moving a point source manually. We used a 720x480 resolution DV camera in this case. We located the limbus of the left eye by minimizing equation (3)[5], computed the coordinates (equation (4)) and orientation (equation (5)) of the cornea and computed the environment map as described in Section 4 for each frame. Then, from each environment map, we computed the direction of the brightest point and used it as the point source direction. The black points in Figure 7(b) are the computed point source directions in spherical angles. The outliers in the trajectory are mainly due to drifts in estimated limbus positions caused by blinking. We fit a 7th order polynomial to the computed source directions along the azimuth (θ) and polar (φ) angles, separately, and obtained the smooth trajectory which is shown as the red curve. One can see that the source was moved back and forth roughly along a line.

Using the estimated point source directions, we rendered a teacup and inserted it into the video (frame by frame). The table was modeled as a plane and, for each frame, the shadow of the virtual teacup was rendered on the table. The positions and the orientations of the teacup and the table were manually selected. Figure 7(a) shows three frames from the resulting video. Notice how the shading and shadows of the cup appear consistent with the illumination of the rest of the scene. It is worth noting that, in this example, the left eye of the person only occupied about 15x15 pixels in each frame. Even with such a low resolution, we are able to estimate the lighting condition with adequate resolution to insert virtual objects into the scene.

[5] We manually specified the initial estimate only for the first frame. Each new estimate was used as the initial estimate for the subsequent frame.

(a) object inserted images (3/500)



(b) point source trajectory

Figure 7: (a) Three out of 500 images of a talking person captured under varying illumination (a moving point source). The teacup is a rendered (virtual) one. The source directions estimated from the left eye were used to render the teacup. (b) The trajectory of the moving point source computed from the left eye. The black points are the computed source directions. The outliers are mainly due to blinking. The red curve is the estimated source trajectory, obtained by fitting a polynomial to the discrete computed source directions. One can see that the source moved back and forth, roughly along a line.

5.2 Relighting Faces

The use of an eye as a light probe becomes particularly relevant when we want to relight faces. We have a natural light probe exactly where (on the face) we want to measure the incident illumination. In the previous work of Tsumura et al.[2003], the eye was used to compute the directions of three point sources rigidly located at known positions with respect to the camera. The computed source directions were used to apply photometric stereo [Woodham 1978] and then render the face from novel viewpoints and under novel lighting conditions. In contrast to [Tsumura et al. 2003], our results enable us to sample the appearance of the face by freely waving a point source in front of the face along an unknown trajectory. As a result, we can recover the shape and albedo map of the face in a very unstructured and convenient manner.

We captured a 1 second video of a person by waving a point source in front of the person. Figure 8(a) shows 3 out of the 30 input images. As in the previous example, we located the limbus,

computed the environment map and estimated the point source direction for each image. In this case, the person did not blink and we directly used the computed source directions. Figure 8(b) shows the computed source trajectory. One can see that the point source was roughly moved along a circle.

In this case, we manually segmented the skin region. We also assumed that the intensity of the point source was the same for all images. Using the captured images and their computed source directions, we first fit the Lambertian reflection model to RGB values at each pixel [Woodham 1978]. This gave us the RGB albedo values and the surface normal for each pixel. Nearly saturated (specular) and dark pixel values (shadows) were thresholded and excluded from the fitting. Erroneous albedo values on the sides and tip of the nose caused by a few remaining shadows and specularities were thresholded and filled in with neighboring albedo values. Figure 9(a) shows the estimated albedo map and (b) shows the estimated surface normals encoded in RGB. Once the surface normals are computed, a height map can be computed by integrating the surface normals [Frankot and Chellappa 1988]. The height map can then be triangulated to obtain a mesh model of the face, as shown in Figure 9(c).

As shown in Figure 9(d), we can render the face under arbitrary lighting and viewing conditions using the estimated face model and albedo map. Here, we added specularities using the Torrance-Sparrow reflection model [1967] with manually selected parameter values. Note that other reflection models can be used instead. We can also render the cornea under the novel lighting condition by using the geometric model introduced in Section 3.1. For the iris reflection, we simply used a constant color taken from the input images. If the images are rendered in high resolution, it will be necessary to render the iris with a photorealistic texture map [Lefohn et al. 2003].

As this example demonstrates, using the eyes to detect light sources provides a convenient means for relighting a face. All we did to sample the appearance of the face was to wave a lamp in front of the person; no special equipment was needed. The eyes in the captured images were at relatively low resolution (approximately 20x20 pixels).

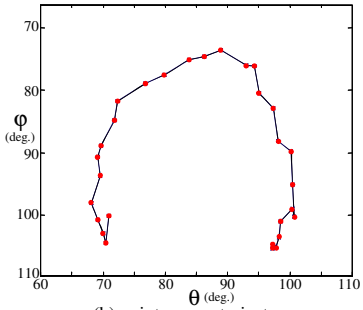
5.3 Replacing Faces

When we have no physical access to the actual location at which an image was taken, the eye may be the only source of illumination information. Here we demonstrate how faces in an already captured image can be replaced with different ones, while preserving realism.

Figure 10 shows two examples of replacing faces in images taken from commercial movies. Figure 10(a1) shows a frame from "Amelie" (courtesy of Miramax Film Corp. ©2001 Miramax Film Corp.). From the left eye of Amelie (a2), we computed the environment map of the scene shown in (a3)[6]. One can clearly see an area light source (probably a window) in the lower right side of the environment map. We subsampled this environment map and used 70 point sources to approximate the lighting condition [Agarwal et al. 2003]. Figure 10(a4) shows the result of replacing Amelie's face with another person (the person in Figure 1). We sampled this person's face using the procedure described in Section 5.2 and rendered her face under the illumination distribution computed from Amelie's eye. Figure 10(b) shows another example where the original image was taken from the movie "Roman Holiday" (courtesy of Paramount Pictures. ©1953 Paramount Pictures). In this case, Gregory Peck's face was replaced with the face sampled in Section 5.2. In the computed environment map in (b3) one can see several light sources. Again, we subsampled this map and obtained 30

[6]The focal lengths of the cameras in these two examples were roughly estimated from the approximate field of views of the cameras.

(a) input images (3/30)



(b) point source trajectory

Figure 8: (a) Three out of 30 images captured by manually waving a point source in front of a face. Images were captured with a DV camera with 480x720 resolution. (b) The trajectory of the point source computed from the left eye. One can see that the point source moved roughly along a circle.

point sources that approximate the lighting condition. In this example, the orientation of the new face was roughly matched with the original one by rotating the 3D model in Figure 9.

Note that we are not trying to solve all problems involved in replacing faces in images; we only focus on our contribution of using the eye to relight a face such that it is consistent with the original illumination. We manually specify the location and orientation of the face. Since we are not separating out the underlying iris texture from the imaged eye, we do not fully rely on the colors and intensity values of the recovered environment map. We manually match the color and the overall gamma of the replaced face. Since the eyes in Figure 9 have a dark iris texture, these manual adjustments were easily achieved. However, for eyes that have, for instance, a bright bluish iris texture, this can become a tedious procedure. For such cases, it would be much desirable to subtract the reflection by the iris from the eye image before computing the environment map. We will discuss this reflection separation problem in Section 6.

In both of these examples, only the face region was replaced. Other parts of the head, including the ears and hair, were retained from the original image by manually segmenting the image. Despite some artifacts due to the difference in the sizes of the faces, the shading of the replaced faces appear consistent with the original faces. These examples demonstrate the significance of using eyes for relighting.

6 Discussion

We proposed the use of eyes for relighting. We analyzed in detail the characteristics of the imaging system formed by the cornea of an eye and a camera viewing it. We showed that a large field of view of the illumination distribution of the environment surrounding a person can be computed from his/her eye. Despite the resolution and dynamic range limitations of currently available cameras, we showed that the eye can be used as a light probe in practical relighting scenarios. In particular, we used eyes to sample the appearance of a face, to include virtual objects in a scene and replace faces in

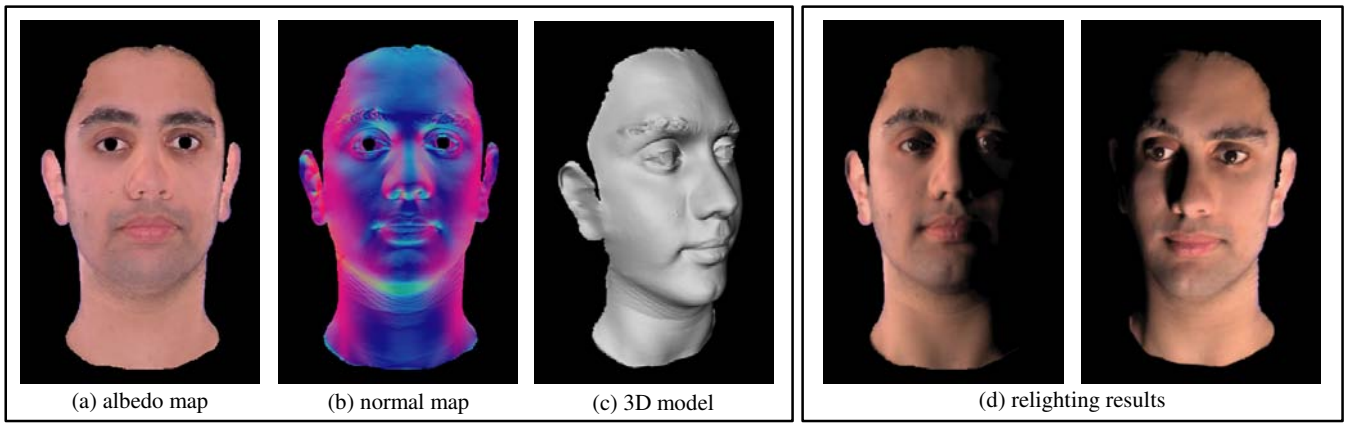
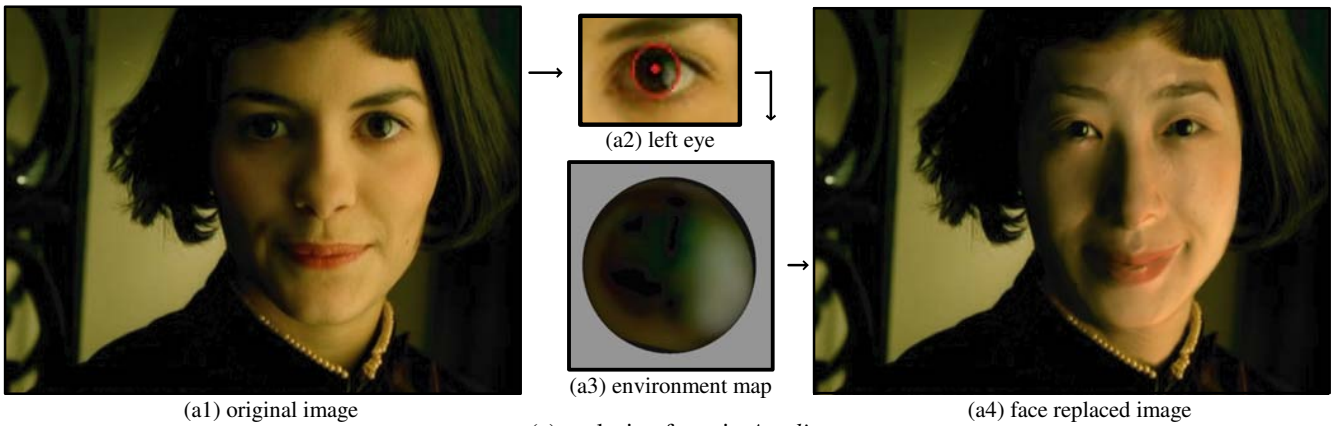
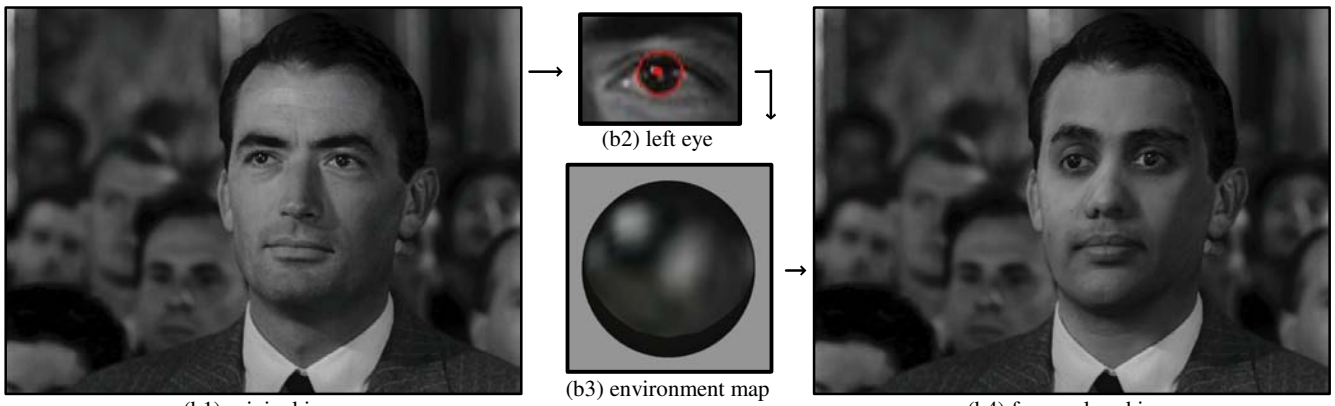


Figure 9: (a) Estimated albedo map and (b) surface normal map computed by fitting the Lambertian reflection model to every pixel of the skin region in the captured image sequence in Figure 8(a). The source directions computed from the eye shown in Figure 8(b) were used for the fitting. (a) The normal map is integrated to obtain a 3D model of the face. (d) Using the computed albedo map and 3D model, the face can be relit under novel lighting conditions.



(a) replacing faces in *Amelie*



(b) replacing faces in *Roman Holiday*

Figure 10: **(a) Amelie:** (a1) A frame from "Amelie" (courtesy of Miramax Film Corp. ©2001 Miramax Film Corp.). (a2) The left eye of Amelie and (a3) the environment map computed from it. By subsampling this environment map, we obtain a set of point sources that approximates the lighting condition. (a4) Using the computed point sources, a different face sampled with the procedure described in Section 5.2 is relit and seamlessly integrated into the scene. **(b) Roman Holiday:** (b1) A frame from "Roman Holiday" (courtesy of Paramount Pictures. ©1953 Paramount Pictures). (b2) The left eye of Gregory Peck and (b3) the computed environment map. (b4) The face used in Section 5.2 is relit using the estimated illumination and integrated into the scene. Hairs and ears are retained from the original images in both cases. The eye enables us to estimate the illumination of a scene even in cases where one does not have physical access to the scene.

an already captured image. These results suggest that the eye is a powerful light probe which may have implications for graphics that go beyond relighting.

A significant research problem we have not addressed in this paper is the separation of iris and corneal reflections from an eye image. We have assumed the computed environment maps to directly correspond to illumination distributions. To obtain high quality illumination distributions, however, we need to remove the colors and brightnesses introduced by the iris texture. Given that iris colors and textures vary quite a bit across people, this is a challenging task. However, it is worth noting that iris textures have a strong radial structure. As a result, the appearance of the iris tends to be strongly uncorrelated with the appearance of the corneal reflection. Therefore, it may be possible to use statistical models of iris appearances to at least partially remove iris textures from eye images. We are currently exploring such an approach.

Acknowledgements

This research was conducted at the Columbia Vision and Graphics Center in the Department of Computer Science at Columbia University. It was funded by an NSF ITR Grant (IIS-00-85864). We thank Todd Zickler for his implementation of [Frankot and Chellappa 1988], Harish Peri, Maryam Kamvar, Emiko Nishino and many others for posing for the experiments. We thank Srinivasa Narasimhan, Kshitiz Garg and other lab members for detailed and helpful comments on early drafts. We also thank the anonymous reviewers for constructive suggestions on the paper. **Photo Credits:** Figure 10(a1) courtesy of Miramax Film Corp; Figure 10(b1) courtesy of Paramount Pictures.

References

AGARWAL, S., RAMAMOORTHI, R., BELONGIE, S., AND JENSEN, H. 2003. Structured Importance Sampling of Environment Maps. In *Proc. of ACM SIGGRAPH 2003*, 605–612.

BAKER, S., AND NAYAR, S. 1999. A Theory of Single-Viewpoint Catadioptric Image Formation. *IJCV 35*, 2 (Nov.), 1–22.

BAKER, T. 1943. Ray tracing through non-spherical surfaces. *Proc. of The Royal Society of London 55*, 361–364.

BARSKY, B., BARGTEIL, A., GARCIA, D., AND KLEIN, S. 2002. Introducing Vision-Realistic Rendering. In *Proc. of EGWR 2002*, 1–7.

BARSKY, B. 2003. Geometry for Analysis of Corneal Shape. In *Computational Geometry: Lectures at Morningside Center of Mathematics*. Amer Mathematical Society, 33–56.

BLANZ, V., AND VETTER, T. 1999. A Morphable Model for the Synthesis of 3D Faces. In *Proc. of ACM SIGGRAPH 99*.

BOIVIN, S., AND GAGALOWICZ, A. 2001. Image-based rendering of diffuse, specular and glossy surfaces from a single image. In *Proc. of ACM SIGGRAPH 2001*, 197–116.

BOLT, R. 1982. Eyes at the Interface. In *Proc. of ACM CHI 82*, 360–362.

BURKHARD, D., AND SHEALY, D. 1973. Flux Density for Ray Propoagation in Gemoetrical Optics. *JOSA 63*, 3 (Mar.), 299–304.

CORNBLEET, S. 1984. *Microwave and Optical Ray Geometry*. John Wiley and Sons.

DAUGMAN, J. 1993. High Confidence Visual Recognition of Persons by a Test of Statistical Independence. *IEEE TPAMI 15*, 11 (Nov.), 1148–1161.

DEBEVEC, P., AND MALIK, J. 1997. Recovering High Dynamic Range Radiance Maps from Photographs. In *Proc. of ACM SIGGRAPH 97*, 369–378.

DEBEVEC, P., HAWKINS, T., TCHOU, C., DUIKER, H.-P., AND SAROKIN, W. 2000. Acquiring the Reflectance Field of a Human Face. In *Proc. of ACM SIGGRAPH 2000*, 145–156.

DEBEVEC, P. 1998. Rendering Synthetic Objects into Real Scenes: Bridging Traditional and Image-based Graphics with Global Illumination and High Dynamic Range Photography. In *Proc. of ACM SIGGRAPH 98*, 189–198.

FLOM, L., AND SAFIR, A., 1987. Iris Recognition System. US patent 4,641,349.

FRANKOT, R., AND CHELLAPPA, R. 1988. A Method for Enforcing Integrability in Shape from Shading Algorithms. *IEEE TPAMI 10*, 4, 439–451.

GEORGHIADES, A. 2003. Recovering 3-D Shape and Reflectance From a Small Number of Photographs. In *Proc. of EGSR 2003*, 230–240.

HALSTEAD, M., BARSKY, B., KLEIN, S., AND MANDELL, R. 1996. Reconstructing Curved Surfaces From Specular Reflection Patterns Using Spline Surface Fitting of Normals. In *Proc. of ACM SIGGRAPH 96*, 335–342.

HUTCHINSON, T., WHITE, K., REICHERT, K., AND FREY, L. 1989. Human-computer Interaction using Eye-gaze Input. *IEEE TSMC 19* (Nov./Dec.), 1527–1533.

JACOB, R. 1990. What You Look At is What You Get: Eye Movement-Based Interaction Techniques. In *Proc. of ACM CHI 90*, 11–18.

KAUFMAN, P., AND ALM, A., Eds. 2003. *Adler's Physiology of the Eye: Clinical Application*, 10th ed. Mosby.

LEFOHN, A., CARUSO, R., REINHARD, E., BUDGE, B., AND SHIRLEY, P. 2003. An Ocularist's Approach to Human Iris Synthesis. *IEEE CGA 23*, 6 (Nov./Dec.), 70–75.

LENSCH, H., KAUTZ, J., GOESELE, M., HEIDRICH, W., AND SEIDEL, H.-P. 2001. Image-Based Reconstruction of Spatially Varying Materials. In *Proc. of EGWR 2001*, 104–115.

LIN, Z., WONG, T.-T., AND SHUM, H.-Y. 2001. Relighting with the Reflected Irradiance Field: Representation, Sampling and Reconstruction. In *Proc. of IEEE CVPR 2001*, vol. 1, 561–567.

MAGDA, S., KRIEGMAN, D., ZICKLER, T., AND BELHUMEUR, P. 2001. Beyond Lambert: Reconstructing Surfaces with Arbitrary BRDFs. In *Proc. of IEEE ICCV 01*, vol. II, 391–398.

MARSCHNER, S., AND GREENBERG, D. 1997. Inverse Lighting for Photography. In *Proc. of IS&T/SID CIC*, 262–265.

MATUSIK, W., PFISTER, H., NGAN, A., BEARDSLEY, P., ZIEGLER, R., AND MCMILLAN, L. 2002. Image-based 3D Photography using Opacity Hulls. In *Proc. of ACM SIGGRAPH 2002*, 427–437.

MITSUNAGA, T., AND NAYAR, S. 1999. Radiometric Self Calibration. In *Proc. of IEEE CVPR 99*, vol. 1, 1374–1380.

MOSTAFAWY, S., KERMANI, O., AND LUBATSCHOWSKI, H. 1997. Virtual Eye: Retinal Image Visualization of the Human Eye. *IEEE CGA 17*, 1 (Jan./Feb.), 8–12.

NAYAR, S., AND MITSUNAGA, T. 2000. High Dynamic Range Imaging: Spatially Varying Pixel Exposures. In *Proc. of IEEE CVPR 00*, 1472–1479.

NISHINO, K., AND NAYAR, S. 2004. Corneal imaging system: Environment from eyes. Tech. rep., Dept. of Computer Science, Columbia University. In preparation.

RAMAMOORTHI, R., AND HANRAHAN, P. 2001. A Signal-Processing Framework for Inverse Rendering. In *Proc. of ACM SIGGRAPH 2001*.

SAGAR, M., BULLIVANT, D., MALLINSON, G., AND HUNTER, P. 1994. A Virtual Environment and Model of the Eye for Surgical Simulation. In *Proc. of ACM SIGGRAPH 94*, 205–212.

SATO, Y., WHEELER, M., AND IKEUCHI, K. 1997. Object shape and reflectance modeling from observation. In *Proc. of ACM SIGGRAPH 97*, 379–387.

SATO, I., SATO, Y., AND IKEUCHI, K. 2003. Illumination from Shadows. *IEEE TPAMI 25*, 3, 290–300.

STEIN, C. 1995. Accurate Internal Camera Calibration using Rotation, with Analysis of Sources of Errors. In *Proc. of ICCV 95*, 230–236.

SZELISKI, R., AND SHUM, H.-Y. 1997. Creating Full View Panoramic Image Mosaics and Environment Maps. In *Proc. of ACM SIGGRAPH 97*, 251–258.

TELLER, S., ANTONE, M., BODNER, Z., BOSSE, M., COORG, S., JETHWA, M., AND MASTER, N. 2003. Calibrated, Registered Images of an Extended Urban Area. *IJCV 53*, 1, 93–107.

TORRANCE, K., AND SPARROW, E. 1967. Theory for off-specular reflection from roughened surfaces. *JOSA*, 57, 1105–1114.

TSUMURA, N., DANG, M., MAKINO, T., AND MIYAKE, Y. 2003. Estimating the Directions to Light Sources Using Images of Eye for Reconstructing 3D Human Face. In *Proc. of IS&T/SID's CIC*, 77–81.

VON HELMHOLTZ, H. 1909. *Physiologic Optics*, third ed., vol. 1 and 2. Voss, Hamburg, Germany.

WOODHAM, R. 1978. Photometric Stereo: A Reflectance Map Technique for Determining Surface Orientation from a Single View. In *Proc. of SPIE*, vol. 155, 136–143.

YAGI, Y., AND YACHIDA, M. 1991. Real-time Generation of Environmental Map and Obstacle Avoidance using Omnidirectional Image Sensor with Conic Mirror. In *Proc. of IEEE CVPR 91*, 160–165.

YU, Y., DEBEVEC, P., MALIK, J., AND HAWKINS, T. 1999. Inverse Globall Illumination: Recovering Reflectance Models of Real Scenes From Photographs. In *Proc. of ACM SIGGRAPH 99*, 215–224.