

Ko Nishino, Kenji Hara, Robby T. Tan, Daisuke Miyazaki, Katsushi Ikeuchi,
"Photometric Aspects on the Preservation of Cultural Assets,"
in Proceedings of The Eighth International Conference on Virtual Systems and MultiMedia,
pp.926-933, Gyeongju, Korea, 2002.09

Photometric Aspects on the Preservation of Cultural Assets

Ko Nishino Kenji Hara Robby T. Tan Daisuke Miyazaki Katsushi Ikeuchi
Institute of Industrial Science, The University of Tokyo

Abstract

We recently started a project to construct a digital archive of cultural heritage objects. Besides geometrical preservation, preserving the current appearance of historical assets is an important and urgent component of our project, due to the daily decay and stain of these ancestral statues and buildings. For this purpose, we have been developing frameworks and algorithms to capture the appearance of real world objects accurately and efficiently. In this paper, we will overview several of these algorithms, namely, appearance modeling from a densely/sparsely sampled image sequence, and image-based shape modeling of transparent objects. These methods can synthesize appearance under novel viewing and lighting conditions, which becomes important when building multimedia contents of historical assets, as well as precisely preserving the appearance itself.

1 Introduction

As an important component of our project on constructing a digital archive of cultural assets, we are making efforts to precisely model the appearance of existing historical statues and buildings as well as capturing their geometry. Since these ancestral assets are suffering from daily weathering and stain, it is an urgent demand to preserve their appearance as they stand today. Furthermore, once we capture the precise geometric and photometric properties of these objects, we are able to display these social treasures as multimedia contents, for example in a virtual museum system, which enables people to look and feel the great historical heritages without spending time and money to visit their actual locations.

In this paper, we will overview several algorithms we have developed in the photometric framework of our digital archive project. In Section 2, we will overview two methods to accurately preserve and synthesize appearance of real world objects. Both methods assume we know the precise geometry of the target object. The first method takes a large set of images of the object under varying viewing and illumination conditions and can be applied to any kind of object material. On the other hand, the second method takes a very sparse set of images, which reduces the burden of capturing hundreds of images. The downside is that it can only be applied to shiny objects, since we use the specular reflection as clues to recover the illumination distribution. In Section 3, we present two algorithms to recover the shape of transparent objects. While many methods exist to reconstruct the shape of opaque objects, for example active or passive range finders including laser range finders and stereo systems, shape-from-X algorithms, etc., these methods cannot be applied to transparent objects. Since many cultural objects containing transparent components exist, such as glasswork and stained glass, we need a way to capture the precise geometry of these objects. We take an image-based approach utilizing polarization to accomplish this task. The problem of shape recovery observing the polarization degree of surface reflection is that it has ambiguity, i.e. the polarization degree and surface normals are a one-to-two mapping rather than a one-to-one mapping. To disambiguate this problem to obtain unique surface normals at each surface point, we first propose a method to additionally use polarization in thermal radiation. Next, we present a method to use multiple views of the target object, observing the polarization degrees from at least two views.

From the next section, we present an overview of each method with several results of applying the method to real world objects.

2 Appearance Modeling

An extensive amount of research has been conducted in both the computer vision and the computer graphics communities with the goal of efficiently representing a real object on the computer to synthesize photorealistic images from arbitrary viewpoints and novel lighting environments (relighting). This research constitutes three schools: image-based rendering,

inverse rendering and 3D photography. Most of these methods rely on a dense sampling of the appearance variation, requiring a large number of images of the target object.

Image-based rendering is a rendering method that basically does not assume any geometry. Taking only 2D images as the input, image-based methods set their basis on the fact that light rays can be parameterized as a 7D function called the *plenoptic function* [1]. Considering each pixel in real images as a sample of this plenoptic function, image-based methods synthesize virtual images by selecting the most appropriate sample of rays, or by interpolating between the sampled rays [15]. Levoy and Hanrahan [10] represent the light rays in free space (space free of occluders) in 4D by using two 2D planes, thereby enabling the plenoptic function to be described in a compact manner. Gortler et al. [6] adopt a similar two 2D plane representation, but additionally they use rough geometric information, in a volumetric representation, derived from images, to correct the basis function of the rays. Shum et al. [23] proposed a 3D representation of the plenoptic function. The method is especially suited to rendering scenes with the viewing direction inside-out, which is the case when modeling real world scenes rather than objects. Since these plenoptic function-based methods require only real images as the input, they provide high generality, i.e., they can be applied to a wide variety of objects and scenes. Because of the principle of interpolation, these approaches tend to require a large amount of input images: a very dense sampling of the appearance variation of objects under a static illumination environment. Even though compression techniques such as vector quantization- or discrete cosine transformation (DCT)-based approaches can significantly reduce the total amount of information that has to be stored on the computer, they still require the user to take hundreds of images as the input.

“Inverse rendering” is another major approach in this area. Taking a dense set of images and a 3D model of the target object model reconstructed *a priori*, inverse rendering methods estimate an approximation of the full bidirectional reflectance distribution function (BRDF) [17] of the object surface by fitting a particular reflection model to the pixel values observed in input images [22] or by solving the inverse radiosity problem [28]. As these methods estimate the BRDF parameters of the object surface, view-dependent rendering and relighting can be accomplished with a very compact representation of the object. In these methods the radiance and positions of the light sources need to be known to compute the BRDF, and direct information of lighting environment has to be provided in some way, e.g., with high dynamic range images [5, 13] of the light sources, making the method less attractive for practical use.

Recent research in the so-called “3D photography” domain have proposed methods that lie somewhere between these two major research streams of image-based rendering and inverse rendering. By taking advantage of the latest advances in 3D sensing equipments, such as laser range scanners [4, 3, 12] and structured light scanners [21], these approaches try to make full use of the 3D geometry as well as the images. The method presented in Section 2.1, Wood et al. [27] and W-C. Chen et al. [2] can be considered as setting one of the 2D planes in the light field approaches attached to the object surface, in concrete, on the coarse triangular patches or dense surface points, respectively. By deriving information from the geometry in this way, these approaches succeed in achieving higher compression ratio without losing smooth view-dependent variation such as the movement of highlights.

In the next two sections, we overview two methods we have developed to model the appearance of real objects from observation. Both methods assume we know the geometry of the target object. The geometric modeling procedure is described in our accompanying paper.

2.1 Densely Sampled Surface Light Field

Given a 3D model of the target object and an image sequence of the object with varying viewing directions and light source directions, how can we efficiently represent the object so that it can be rendered from arbitrary viewpoints? Furthermore, how can we change the lighting conditions based on that representation?

Unlike pure image-based rendering methods, ours assumes that we know the geometry of the scene, i.e., the shape of the object and the viewpoint. Knowledge of the geometry of the object contributes much information toward making the representation compact. To take full advantage of the geometry of the object, we handle the appearance variation, caused by view direction and light source direction changes on the surface of the object. Note that, although it is the same 2D as the images in image-based rendering methods, it significantly reduces the redundancy to represent the appearance. In an image-based rendering point of view; this can be considered as pasting one of the two 2D slices to parameterize light rays on the object surface, more concretely, on the triangular patches of the mesh model of the object. The term “Surface Lightfields” introduced by Miller et al. [11] captures this idea well, although they use it in a different context.

Figure 1 displays an overview of the proposed method, which we will refer to as *Eigen-Texture Rendering* method. First, we create a 3D model of the target object from a sequence of range images. We use a light stripe range finder [21] to acquire the range images, and then the registration and integration of range images are done with the methods presented in our accompanying paper. Next, we take a number of color images and align and paste them onto the surface of the

object model. Each color image is divided into small areas that correspond to triangular patches on the 3D model. Then the color images corresponding to triangular patches are warped to have a predetermined normalized triangle shape by bi-linearly interpolating the pixel values. We refer to this normalized triangular shape corresponding to each triangular patch as a *cell* and to its color image as a *cell image*. A sequence of cell images from the same cell is collected as shown in Figure 1. Here, this sequence depicts appearance variations on the same physical patch of the object under various viewing conditions. Principle component analysis (PCA) is applied for each cell image sequence to compress each cell image set separately. Note that the compression is done in a sequence of cell images. The appearance changes in the cell image sequence are mainly due to the change of brightness, when the object surface is well approximated with the triangular patch based 3D model. Thus, high compression ratio can be expected with PCA. Furthermore, it is possible to interpolate appearances in the eigenspace. View-dependent image synthesis can be accomplished by taking a linear combination of basis cell images per triangular patch (see Figure 2). Virtual images under a complicated illumination condition can be generated by a summation of component virtual images sampled under single illuminations thanks to the linearity of image brightness.

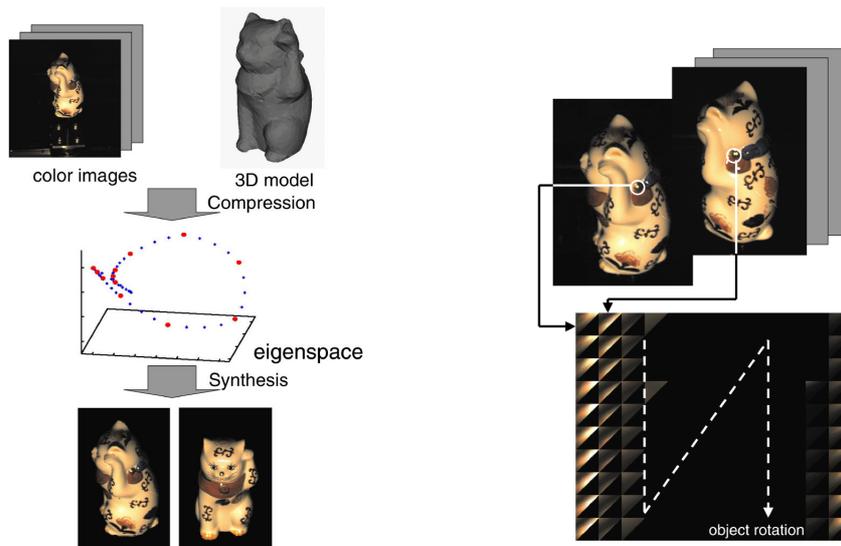


Figure 1: *Left: Outline of the Eigen-Texture Rendering method. Right: A sequence of cell images.*

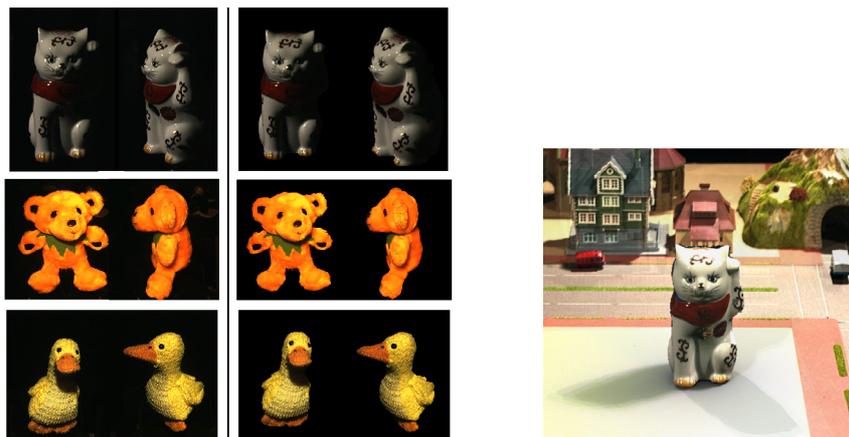


Figure 2: *Left image left: Input color images, Left image right: Synthesized images (by using cell-adaptive dimensional eigenspaces). Right image: Integration results with real background.*

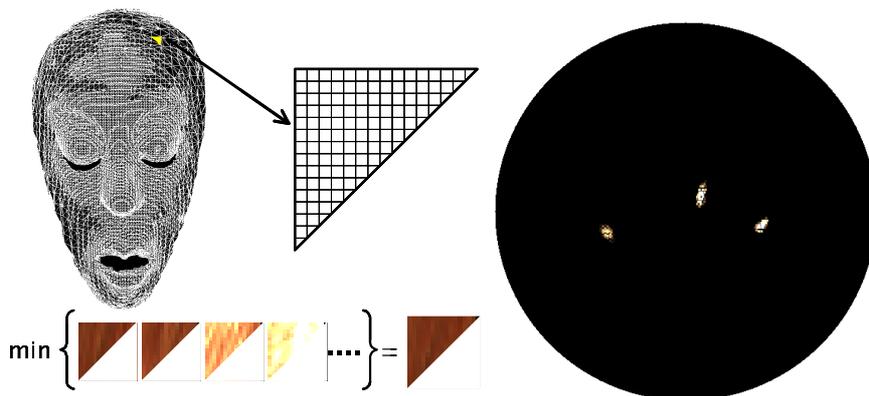


Figure 3: *Left: Diffuse-texture-map. Right: Estimated illumination distribution (looking down the illumination hemisphere from the top).*

2.2 Sparsely Sampled Surface Light Field

Following an inverse rendering approach, we present in this section a framework to accomplish photorealistic rendering, including view-dependent rendering and relighting, from a sparse set of images. These images sparsely sample the surface light field of a real world object under a static illumination environment. With the aid of the geometric information of the target object, and assuming that the specular reflectance property is homogeneous over the object surface, we estimate three important elements necessary for photorealistic forward rendering: diffuse texture, approximate Bidirectional Reflectance Distribution Function (BRDF) and the illumination distribution.

We assume that the surface reflection can be approximated by a dichromatic reflection model with a Lambertian diffuse reflection and simplified Torrance-Sparrow specular reflection. Given a relatively small number of images (on the order of one) of an object, the geometry of the object (as a 3D mesh model) and the positions of the cameras corresponding to each image, we first separate the diffuse and specular reflection components on each object surface point. By constructing a *diffuse texture map* that represents the diffuse reflection and subtracting the diffuse-texture-mapped images from the input images one by one, we obtain a sequence of specular reflection images. Then, we derive an initial estimation of the illumination distribution by shooting back each pixel value to the perfect mirror direction and mapping those values to a hemisphere, which we call the *illumination hemisphere*. We then formulate the specular reflection mechanism as a 2D convolution on the surface of the illumination hemisphere and estimate the reflection parameters and the true illumination distribution by solving a blind 2D deconvolution problem. The algorithm for this blind deconvolution problem is based on the Alternating Minimization (AM) algorithm with a Total Variation (TV) regularization term imposed on the illumination distribution. A robust technique based on M-estimators is incorporated to combat noise inherited from the diffuse texture extraction procedure.

From six images of a mask made in Bali, we derived the diffuse-texture-map by virtually considering a 3D grid in each triangular patch and mapping the pixel value whose intensity was minimum through out the image sequence (see Figure 3). After constructing an initial illumination hemisphere and accomplishing deconvolution of the surface reflectance property and the lighting distribution, we obtain an estimate of the illumination environment as depicted in Figure 3 (Note the three small clusters corresponding to three point light sources that were placed when capturing the input image sequence). Using these three components, namely, the diffuse-texture-map, recovered illumination hemisphere and the specular reflection parameters, we are able to render synthetic image of the object from arbitrary viewpoints and novel illumination conditions as shown in Figure 4.

3 Shape Acquisition of Transparent Objects

Many techniques to measure object shape have been developed especially in the field of optical engineering. These techniques can be classified into two categories: point and surface type. A point type method, such as a laser range sensor, measures object shape by projecting a spotlight, often a laser beam, over the object surface, and by measuring the returned timing or the returned direction. A surface method, such as Moire topography, determines the shape of an object by projecting a planar light and measuring the interference of the light with the surface.



Figure 4: *Left: Novel viewpoint, Right: Novel illumination condition*

In addition, the computer vision community has extensively developed such techniques. Shape-from-shading, for example, analyzes shading information in an image with a reflectance map in order to relate image brightness to surface orientations. Photometric stereo obtains information from three images, taken from the same position, under three different illumination conditions. Binocular stereo and motion analysis use image differences in a series of images taken from different positions.

Most of these methods are, however, designed to obtain the shape of opaque surfaces. Namely, these techniques are based on analysis of the body reflection component of an object surface. Models of transparent objects, which have only surface reflection, cannot be created using these techniques. Polarization is a useful cue to analyze surface reflection. Kosikawa [9] proposed to use the degree of polarization, employing polarized light sources to determine the shape of metal surfaces. Wolff [24] proposed to analyze the degree of polarization in visible light. Wolff et. al. [25] indicated that the surface normal of the object surface is constrained by analyzing the polarization of the object; Rahmann [18] addressed the potential of recovering shape from polarization; and Jordan et. al. [7, 8] and Wolff et. al. [26] analyzed the degree of polarization in the infrared wavelength. Saito et. al. [20, 19] employed the analysis of the degree of polarization and developed a method with which to measure the surface of a transparent object. Employing an extended light source originally developed by Nayar et. al. [16], they illuminated a transparent object and were able to obtain surface reflection components over the entire visible surface. Then, by measuring the degree of polarization, they determined surface orientations. Unfortunately, however, the degree provides two solutions corresponding to one polarization degree. Thus, the method can be applied to measuring a limited class of objects or to surface inspection where rough surface orientation is predetermined; it cannot be applied to a general class of objects.

In the next two sections, we overview our latest work on shape recovery of transparent object utilizing polarization. To overcome the problem of ambiguity in polarization based surface reflection analysis, we propose to use additional information, namely, the degree of polarization of thermal radiation and multiple observations of polarization curves.

3.1 Thermal Radiation

Generally speaking, natural light is unpolarized; it oscillates in all directions on the plane of oscillation, which is perpendicular to the path of the light. Natural light, however, becomes polarized once it goes through a polarization material or is reflected on a surface. The intensity varies depending on the direction on the plane of oscillation, and therefore a difference can be observed when the polarization filter is rotated in front of a CCD camera. The variance is described as a sinusoidal function of rotation angles. We will denote the maximum and minimum brightness in the observed intensities as I_{\max} and I_{\min} . Given that the sum of the maximum and minimum brightness is the total brightness of the reflected light, I_{specular} ,

$$I_{\max} = \frac{F_s}{F_p + F_s} I_{\text{specular}}, \quad I_{\min} = \frac{F_p}{F_p + F_s} I_{\text{specular}}. \quad (1)$$

By this equation, the direction parallel to the plane of incidence provides the minimum brightness I_{\min} . Namely, by measuring the angle to give the minimum brightness, we can determine the direction of the plane of incidence, θ . Figure 5 shows the geometric relationship between a light source, camera and the incident plane on the object surface. Furthermore,

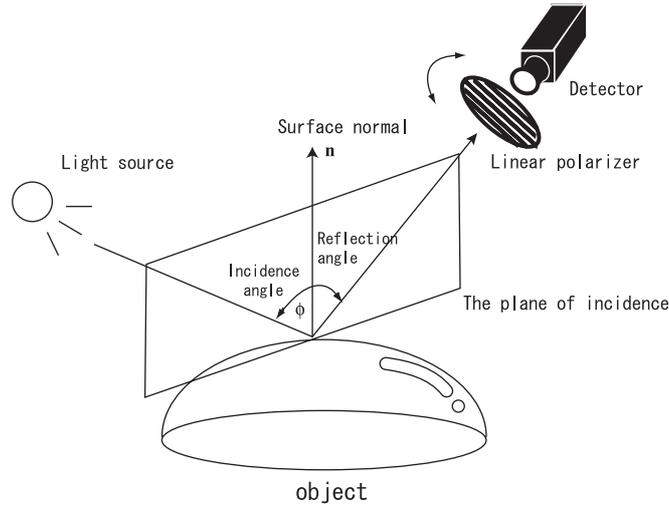


Figure 5: *The geometric relationship between the light source, camera and the incident plane.*

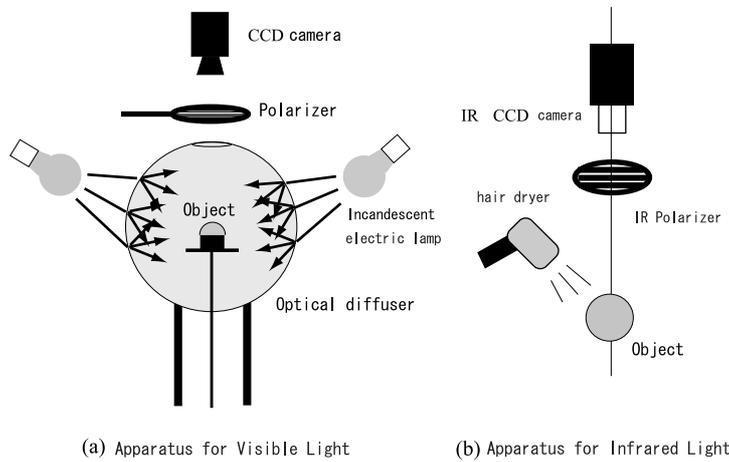


Figure 6: *Experimental setup. Left: for visible light, right: for thermal radiation.*

by acquiring the degree of polarization, we can estimate the surface normal at each surface point. However the relationship between the degree of polarization and the incident angle is not one-to-one but a quadratic convex curve which has an extreme at the Brewster angle. From this function, an observed degree of polarization provides two possible incident angles, except at the Brewster angle. Hence, we are able to determine the surface normal only up to two possibilities. It is necessary to have a method to resolve this ambiguity.

We propose to use thermal radiation as an additional information to disambiguate the surface normals for each surface point. The relation between the polarization ratio and the emitting angle is a one-valued function; there is a one-to-one correspondence between the ratio and the emitting angle. Thus, once we measure the ratio in an infrared light, we can uniquely determine the emitting angle. While the polarization ratio of thermal radiation itself is small and unreliable to be used for direct estimation of light emittance angle and consequently the surface normal, it can provide enough information to disambiguate the one-to-two mapping in the visible light case.

Figure 6 shows the apparatus to measure the degree of polarization of both visible light and thermal radiation. Figure 7 depicts the shape reconstruction result of an acrylic shellfish object.

3.2 Stereopsis in Polarization

The defect of the method proposed in the former section is that one should set up measurement equipments of infrared light in addition to that of visible light. For the sake of convenience in the experimental setup, we propose another method to solve the problem of recovering surface normals through polarization in visible light, which is degenerate because

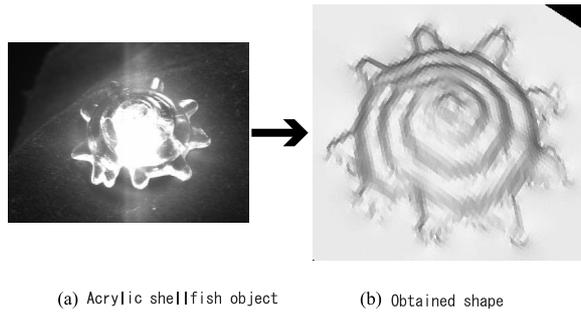


Figure 7: *Shape reconstruction result.*

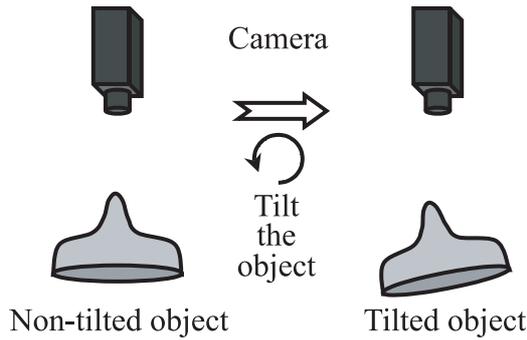


Figure 8: *Two views of the target object with object rotation.*

of the one-to-two mapping explained in the former section. To disambiguate the correspondence between polarization degree and the surface normal, we developed an algorithm which incorporates an additional view of the object in interest. Instead of taking a single image of the surface reflection with the apparatus depicted in the left hand side image of Figure 6, we capture two images: one image in a normal setup and another image with the object slightly tilted (see Figure 8).

We segment these data into some regions with regards to the Brewster angle. We calculate the difference of the polarization degree between these two sets of data at the corresponding point: the point where surface normal lies along the rotation direction and where polarization degree is minimum in the region bounded by the Brewster angles. From that difference, we determine the correct surface normal. Further details of how to establish the correspondence between the two views can be found in [14]. Figure 9 shows one result of shape recovery through stereopsis in polarization degrees.

4 Conclusion

In this paper, we overviewed several algorithms we have developed in the photometric framework of our project on preservation and restoration of cultural assets. In particular, we have presented appearance modeling from a densely/sparsely sampled image sequence, and image-based shape modeling of transparent objects. These methods can synthesize appearance under novel viewing and lighting conditions, which becomes important when building multimedia contents of historical assets, as well as precisely preserving the appearance itself. We plan to apply these methods to existing historical objects.

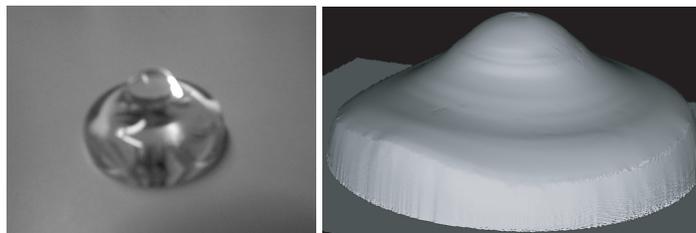


Figure 9: *Shape recovery result. Left: real image, right: recovered 3D model.*

References

- [1] E.H Adelson and J.R. Bergen. *The Plenoptic Function and the Elements of Early Vision*, pages 3–20. Computational Models of Visual Processing, M.Landy and J.A.Movshon(eds),MIT Press, 1991.
- [2] W-C. Chen, R. Grzeszczuk, and J-Y. Bouguet. Light Field Mapping: Hardware-Accelerated Visualization of Surface Light Field. In *ACM SIGGRAPH 01 Course Notes, Course 46 Acquisition and Visualization of Surface Light Fields*, Aug. 2001.
- [3] Cyberware. <http://www.cyberware.com>.
- [4] Cyra. <http://www.cyra.com>.
- [5] P.E. Debevec and J. Malik. Recovering High Dynamic Range Radiance Maps from Photographs. In *Computer Graphics Proceedings, ACM SIGGRAPH 97*, pages 369–378, Aug. 1997.
- [6] S.J. Gortler, R. Grzeszczuk, R. Szeliski, and M.F. Cohen. The Lumigraph. In *Computer Graphics Proceedings, ACM SIGGRAPH 96*, pages 43–54, Aug. 1996.
- [7] D. L. Jordan and G. D. Lewis. Measurements of the effect of surface roughness on the polarization state of thermally emitted radiation. *Opt. lett.*, 19:692–694, 1994.
- [8] D. L. Jordan, G. D. Lewis, and E. Jakeman. Emission polarization of roughened glass and aluminum surfaces. *Appl. Opt.*, 35:3583–3590, 1996.
- [9] K. Koshikawa. A polarimetric approach to shape understanding of glossy objects. In *Proc. of Int. Joint Conf. Artificial Intelligence*, pages 493–495, 1979.
- [10] M. Levoy and P. Hanrahan. Light Field Rendering. In *Computer Graphics Proceedings, ACM SIGGRAPH 96*, pages 31–42, Aug. 1996.
- [11] G.S.P Miller, S. Rubin, and D. Ponceleon. Lazy Decompression of Surface Light Fields for Precomputed Global Illumination. In *Eurographics Rendering Workshop*, pages 281–292, Jun. 1998.
- [12] Minolta. *Vivid 900*, 2001. <http://www.minolta-rio.com/vivid/>.
- [13] T. Mitsunaga and S.K. Nayar. Radiometric Self Calibration. In *Proc. of Computer Vision and Pattern Recognition '99*, pages 374–380, Jun. 1999.
- [14] D. Miyazaki. *Measuring Surface Shape of Transparent Objects Based on the Analysis of Polarization*. PhD thesis, Dept. of Info. Science, Grad. School of Science, The Univ. of Tokyo, 2002.
- [15] T. Naemura, T. Yanagisawa, M. Kaneko, and H. Harashima. Efficient projection of ray space for a ray based description of 3-D real world. In *IEICE IE95–119 (in Japanese)*, Feb. 1996.
- [16] S. K. Nayar, K. Ikeuchi, and T. Kanade. Determining shape and reflectance of hybrid surface by photometric sampling. *IEEE Trans. Rob. Auto.*, 6:418–431, 1990.
- [17] F.E. Nicodemus, J.C. Richmond, J.J. Hsia, I.W. Ginsberg, and T. Limperis. Geometric Considerations and Nomenclature for Reflectance. National Bureau of Standards (US), 1977.
- [18] S. Rahmann. Polarization images: a geometric interpretation of shape analysis. In *Proc. Int. Conf. on Pattern Recognition*, pages 542–546, 2000.
- [19] M. Saito, Y. Sato, K. Ikeuchi, and H. Kashiwagi. Measurement of surface orientations of transparent objects by use of polarization in highlight. *J. Opt Soc. Am. A*, 16(9):2286–2293, 1999.
- [20] M. Saito, Y. Sato, K. Ikeuchi, and H. Kashiwagi. Measurement of surface orientations of transparent objects using polarization in highlight. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages 381–386, 1999.
- [21] K. Sato and S. Inokuchi. Range-imaging system utilizing nematic liquid crystal mask. In *First International Conference on Computer Vision*, pages 657–661, 1987.
- [22] Y. Sato, M.D. Wheeler, and K. Ikeuchi. Object shape and reflectance modeling from observation. In *Computer Graphics Proceedings, ACM SIGGRAPH 97*, pages 379–387, Aug. 1997.
- [23] H-Y. Shum and L-W.He. Rendering with Concentric Mosaics. In *Computer Graphics Proceedings, ACM SIGGRAPH 99*, pages 299–306, Aug. 1999.
- [24] L. B. Wolff. Polarization-based material classification from specular reflection. *IEEE Trans. Patt. Anal. Mach. Intell.*, 12(11):1059–1071.
- [25] L. B. Wolff and T. E. Boult. Constraining object features using a polarization reflectance model. *IEEE Trans. Patt. Anal. Mach. Intell.*, 13(7):635–657, 1991.
- [26] L. B. Wolff, A. Lundberg, and R. Tang. Image understanding from thermal emission polarization. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages 625–631, 1998.
- [27] D.N. Wood, D.I. Azuma, K. Aldinger, B. Curless, T. Duchamp, D.H. Salesin, and W. Stuetzle. Surface Light Fields for 3D Photography. In *Computer Graphics Proceedings, ACM SIGGRAPH 00*, pages 287–296, Jul. 2000.
- [28] Y. Yu, P. Debevec, J. Malik, and T. Hawkins. Inverse Global Illumination: Recovering Reflectance Models of Real Scenes From Photographs. In *Computer Graphics Proceedings, ACM SIGGRAPH 99*, pages 215–224, Aug. 1999.