2007－CVIM－158 (32)
2007／3／20

Reconstructing an object's shape from its appearance manifold under moving light

YuQiong† ImariSato† TakahiroOkabe† Yoichi Sato†

†Institute of Industrial Science, The University of Tokyo
4-6-1 Komaba, Meguro-ku,
Tokyo 153-8505, JAPAN
E-mail : qiong@iis.u-tokyo.ac.jp; imarik@nii.ac.jp
takahiro@iis.u-tokyo.ac.jp; ysato@iis.u-tokyo.ac.jp

Abstract
This paper presents a technique for recovering the shape of an object from its appearance manifold composed of a set of images that can be taken from a fixed viewpoint camera under a moving light source. We assume the distant illumination, a convex object shape and the variations of the object's appearance under a moving light are caused by the difference in the surface normals. These input images give an appearance manifold and a dimensionality reduction technique called 'Isomap' can recover the embedding three-dimensional surface normals of the object from this appearance manifold. Furthermore, the boundary points are used as the reference points to transform the result three-dimensional vectors into the true distribution of the surface normals. The proposed method is available for a wide range of reflectance materials and is easy to implement. The only requirement for our method is to take the different images of an object under different lighting directions.
Keywords appearance manifold, appearance profile, Isomap, dimensionality reduction.

1 Introduction

The appearance of an object is determined by several factors such as illumination, viewing position, the surface shape and the reflectance property of the object. Changing any one of these factors should lead to the object having a different appearance. Based on the relation among these factors, in the computer vision field, some important inverse problems have been addressed. In the real world, for most of objects' appearance, these factors are related nonlinearly if the object has complex reflectance properties. Most of the previous inverse approaches thus estimate some of these factors from images of a scene assuming one or two of the factors are given. How to construct the shape of the complex reflectance property object makes the topic of shape reconstruction still be a hot research in the computer vision field.

The previous studies have demonstrated that the shape of an object can be recovered from a single image or multiple images of the object. Always these studies based on the images need assume that some knowledge about the scene is given: the illumination is known or the surface materials are known. Furthermore, some works can only success in constructing the shape of a lambertian object. For example most of the shape-from-shading approaches estimate an object's shape from a single image assuming distant illumination and uniform lambertian reflectance [ZTCS99].

The classical photometric stereo approach presented in [Hor86, Woo81] recovers the shape of a lambertian object from multiple images of the object taken under known light sources. Photometric stereo has been intensively studied as a fundamental computer vision problem. For instance, uncalibrated photometric stereo approaches estimate the shape of a lambertian object up to a linear ambiguity under unknown lighting [BKY99, BJ03]. Some researches applied other analytic reflectance models to photometric stereo to deal with non-lambertian surfaces [NIK90] and the previous approaches have shown promising results for objects with

various surface materials. Furthermore, some approaches use the combination of the photometric stereo and the geometric method to get a good shape reconstruction result [VH06].

However, as noted in [HS05], real world materials sometimes have complex appearances that prevent us from extracting their shapes by using the analytic reflectance models. To cope with this problem, the use of a calibration object was proposed in the early works on photometric stereo [Sil80, KH86]. A more sophisticated example-based photometric stereo approach [HS05]. can handle the objects with non-uniform surface materials and do not need a particular calibration object for each target object.

Now the question we ask next is, given only images of an object captured under various lighting conditions, is it possible to achieve the shape of the object without any calibration object? Is there any information left in the input images that we can use as reference? Recently, Koppal and Narasimhan presented a novel approach for clustering surface normals of a scene of unknown geometry and surface materials under unknown illuminations [KN06]. Their approach shows how effective it is to analyze the temporal variation in the appearance of a scene for clustering surface normals for representing its meaningful geometric structure.

Our proposed method can directly recover an object's shape from its appearance changes under a freely moving unknown light. In the procedure of shape reconstruction, we need not to make any analytic reflectance model or to cluster the surface normals into the groups for estimating the surface normals. Assuming distant illumination and a convex object shape, the temporal variations in the appearance of the object surface under a moving light source reflect the difference in the surface normals. Through analyzing the difference between the different positions on the object's surface in the different appearances, we can make it possible to discover the shape of the object from this high-dimensional input appearance manifold. The merit of this proposed method is easy to implement, do not need complex equipment and the only requirement is the images of an object taken from a fixed view position under different lighting directions. Also no calibration is required for lights and camera. and the order of the input images does not affect the estimation results.

2 Shape from Appearance Manifold

Consider a set of images of an object captured under n different illuminations seen from a fixed view point. Let I_i^p be the intensity of a each pixel or each corresponding surface point p seen under the ith illumination; then the observation vector, also known as the *appearance profile*, for this point is

$$\vec{o_p} = [I_1^p, \cdots, I_n^p]^T . \quad (1)$$

From input images with m surface points ($p = 1, \cdots, m$), we obtain m observation vectors. Note that observation vectors can be also thought of as data points in an n-dimensional vector space, $o_p = (I_1^p, \cdots, I_n^p)$. To examine the variations in the appearance of surface points under a moving light source, the observation vectors are normalized and used as inputs for shape recovery. Since both illumination and viewing directions are consistent over the object surface, it seems reasonable to think that the variations of the observation vectors reflect the distribution of the object's surface normals and data points o_p lie on a manifold whose intrinsic reveals the distribution of the object's surface normals.

We refer to this manifold as an *appearance manifold* and use an effective embedding method to find a three-dimensional embedding of this appearance manifold. We later investigate the validity of the condition of the appearance manifold by using analytic reflection models.

2.1 Embedding of Appearance Manifold

Given high-dimensional input data, the classical techniques for dimensional reduction such as principal component analysis (PCA) and multidimensional scaling (MDS) are able to

find the intrinsic structure of the data lying on or near its linear subspace.

In the case of an appearance manifold, however, the set of n-dimensional input data generally has nonlinear structures, and thus it is difficult to discover its low-dimensional embedding by using the above techniques. To reveal the intrinsic structures underlying an appearance manifold, we employ a nonlinear embedding method, called isometric feature mapping (Isomap), proposed by Tenenbaum et al. in [TSL00].

Isomap has been used to find perceptually meaningful low dimensional manifolds of natural images, such images of a face with different poses and lighting directions. Isomap learns a manifold of input data as a graph by connecting k-nearest neighbors among all data points. A low-dimensional embedding of this manifold is estimated such that the geodesic distances between all pairs of points are preserved even after dimensionality reduction.

In the case of an appearance manifold, differences between all pairs of n-dimensional data points o_p are due to differences in their surface normals. Therefore, if a three-dimensional embedding of this appearance manifold is discovered by Isomap, it should reveal the distribution of surface normals of the object. This is the key idea of our algorithm for shape recovery through the appearance manifold.

3 Algorithm

As defined in section 2, for one point p, its corresponding *appearance profile* is remembered as a vector o_p: $o_p = (I_1^p, \cdots, I_n^p)$. If the input images captured under n different illuminations and there are m surface points in one image, the input appearance manifold can be seen as m n-dimensional vectors. The surface normals are estimated from their appearance manifold as described in the following steps shown in the figure 1:

1. **Discover low-dimensional representation.** The nonlinear embedding method, Isomap, is employed for discovering three-dimensional structure underlying the appearance manifold (Section 3.1).

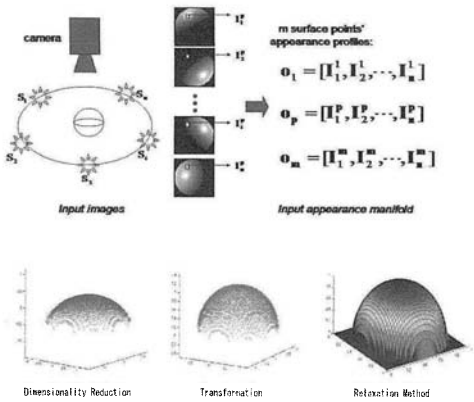

Figure 1: The main steps of our method

2. **Transform outputs into a surface normal distribution.** Since the converged solution from Isomap does not necessarily correspond to the true distribution of object's surface normals. We use the occluding boundary as reference points to transform the output from Isomap (Section 3.2).

3. **Estimate shape from surface normals.** The height field of the object surface is recovered from the transformed data points by using the relaxation method.

3.1 Dimensionality Reduction

The dimensionalirty reduction method, Isomap, can recover the nonlinear embedding structures from the high-dimensional data and the results best preserve the intrisic difference between the vectors in high dimensional space. A matrix D of distances between all pairs from o_p for $(p = 1, \cdots, m)$ as input. Element (i, j) of the distance matrix D, denoted as $d_M(i, j)$, is the distance between o_i and o_j. In our work, $d_M(i, j)$ is measured using the standard Euclidean distance. Given a distance matrix, the Isomap algorithm works as follows:

1. Define a graph G over all data points o_p based on the distance matrix. In the

graph G, points i and j are connected if i is one of the k nearest neighbors of j.

2. Initialize geodesic distances $d_G(i,j)$ for all i,j pairs from m data points: $d_G(i,j) = d_M(i,j)$ if i and j are connected as neighbors, and $d_G(i,j) = \infty$ otherwise. Then compute the shortest path distances between all pairs of points in G by using Floyd's algorithm and update the geodesic distance $d_G(i,j)$.

3. Apply MDS to the matrix that contains graph distances D_G: The (i,j) elements of this matrix are $d_G(i,j)$. The MDS algorithm finds a three-dimensional embedding such that the intrinsic geometry of the appearance manifold is best preserved even after the dimensionality reduction [KJ79].

3.2 Transform Solution from Isomap into Surface Normal Distribution

There is an important issue when the output from isomap is used for recovering the object shape. Although isomap computes a globally optimal solution, the converged solution does not necessarily correspond to the true distribution of the object's surface normals; what isomap computes is the relative relationships of its surface normals. The occluding boundary of the object is used as reference to transform the output from isomap to the distribution of surface normals of the object.

The occluding boundary is the curve on the object surface that is projected as a silhouette in the input image. Occluding contours of an object can be found by applying a gradient-based edge detector to the image of the object. The surface normal on the occluding boundary lies in a plane parallel to the image plane.

Let $C_b = (c_x^b, c_y^b, c_z^b)$ be surface normals of those boundary points b. Assuming that C_b are on a plane through the origin $(0,0,0)$ with its surface normal $(0,0,1)$, c_x^b and c_y^b of those points can be computed as their gradient directions in the 2D image coordinate system defined in the input image. This transforms the viewing direction $\vec{w_o}$ into the direction $(0,0,1)$ as well from the definition of the occluding boundary.

Let $E_b = (e_x^b, e_y^b, e_z^b)$ be the output three-dimensional coordinates of the corresponding boundary points from Isomap. A transformation $M : R^3 \rightarrow R^3$ is estimated such that the transformed points $M(E_b)$ minimize

$$\sum_{\text{all } b} \|M(E_b) - C_b\|^2, \tag{2}$$

where the correspondences between E_b and C_b are given based on the initial pixel location of o_b, which is maintained through the dimensionality reduction process. Then we achieve the transformation relation matrix T and R from a set of mapping steps between C_b and E_b. Finally, we modify the isomap three-dimensional results I by the equation: $R\dot{I} + T$ to get the results. In the process of transformation, the rotation and transformation matrix in two-dimensional and three-dimensional space are used, some optimization problems are solved by using the function in the MATLAB.

4 Applicable Materials

Since illumination and camera position are assumed to be sufficiently distant from an object, illumination and viewing directions are consistent over the object surface; then the reflection equation for a point p when it is illuminated by the ith illumination with unit radiance is

$$I_i^p = f_p(\vec{\omega_i}, \vec{\omega_o})(\vec{n_p} \cdot \vec{\omega_i}), \tag{3}$$

where $\vec{n_p}$ is the surface normal of p, $\vec{\omega_i}$ and $\vec{\omega_o}$ are incident and reflection directions that are consistent over all surface points of the object, and $f_p(\vec{\omega_i}, \vec{\omega_o})$ represents a bidirectional reflectance distribution function (BRDF) of the point p that represents how much of the incident light $\vec{\omega_i}$ is reflected on the object surface toward $\vec{\omega_o}$.

4.1 Lambertian Surfaces

Let us start with the case where objects have uniform Lambertian reflectance. The BRDF for a Lambertian surface is known to be a constant. From (3), the equation for a Lambertian

surface is given as

$$I_i^p = \mathrm{kd}(\vec{n_p} \cdot \vec{\omega_i}), \tag{4}$$

where kd is a constant albedo over the object surface.

Suppose that q and w are two points on an object surface; then their intensities are computed from (4) as: $I_i^q = \mathrm{kd}(\vec{n_q} \cdot \vec{\omega_i})$ and $I_i^w = \mathrm{kd}(\vec{n_w} \cdot \vec{\omega_i})$. Comparing these equations, it can be clearly seen that under the same incident direction $\vec{w_i}$, their intensity difference is caused by the difference between their surface normals $\vec{n_q}$ and $\vec{n_w}$.

Since this is true for all n illumination directions $(i = 1, \cdots, n)$, it can be concluded that the differences in observation vectors $\vec{o_p} = [I_1^p, \cdots, I_n^p]^T$ among m surface points result from their surface normal differences, and thus points o_p construct an appearance manifold.

4.2 Textured Lambertian Surfaces

Suppose that two surface points q and w have the same surface normal $\vec{n}$, but different diffuse parameters $\mathrm{kd}_q \neq \mathrm{kd}_w$. From (4), their intensities under the ith illumination are $I_i^q = \mathrm{kd}_q(\vec{n} \cdot \vec{\omega_i})$ and $I_i^w = \mathrm{kd}_w(\vec{n} \cdot \vec{\omega_i})$, and their observation vectors become

$$\begin{aligned} \vec{o_q} &= \mathrm{kd}_q\,[(\vec{n} \cdot \vec{\omega_1}), \cdots, (\vec{n} \cdot \vec{\omega_n})]^T, \\ \vec{o_w} &= \mathrm{kd}_w\,[(\vec{n} \cdot \vec{\omega_1}), \cdots, (\vec{n} \cdot \vec{\omega_n})]^T. \end{aligned} \tag{5}$$

We can see that kd_q and kd_w are just scalar values of the same vector $[(\vec{n} \cdot \vec{\omega_1}), \cdots, (\vec{n} \cdot \vec{\omega_n})]^T$ in these equations.

As a result, by normalizing each observation vector $\vec{o_p}$ by its length $\|\vec{o_p}\|$, the effect of diffuse parameters kd_p or kd_w can be canceled. Accordingly, the variations in these vectors are due to their surface normal differences, and this leads to an appearance manifold for the textured Lambertian surface.

4.3 Specular Surfaces

The proposed approach can be applied to non-Lambertian surfaces as well. Let us take several reflection models as examples. Supposing uniform reflectance properties over the object surface, the intensity of a surface point is computed as

Blinn-Phong model:

$$I_i^p = \mathrm{ks}\frac{n+2}{2p}\cos^n \angle(\vec{h_i}, \vec{n_p}) \tag{6}$$

Torrance-Sparrow model [TS67]:

$$I_i^p = \mathrm{ks}\frac{1}{(\vec{v} \cdot \vec{n_p})}\exp(\frac{\angle(\vec{h_i}, \vec{n_p})^2}{2\sigma^2}). \tag{7}$$

Ward isotropic reflection model [War92]:

$$I_i^p = \mathrm{ks}\frac{1}{\sqrt{(\vec{n} \cdot \vec{\omega_i})(\vec{n} \cdot \vec{\omega_o})}}\frac{\exp(-\tan^2 \angle(\vec{h_i}, \vec{n_p})/\sigma^2)}{4\pi\sigma^2}, \tag{8}$$

where ks is a constant for the specular reflection component, and σ is the standard deviation of a surface slope. $\vec{h_i}$ is the bisector of the light source direction $\vec{\omega_i}$ and the viewing direction $\vec{v}$, and the function $\angle(\vec{h}, \vec{n})$ computes the angle between two vectors.

Comparing intensities I_i^p for all surface points $(p = 1, \cdots, m)$ illuminated by the ith illumination from the direction $\vec{w_i}$, we can see that only the surface normals $\vec{n_p}$ differ in their reflection equations in the case where reflectance properties ks, sigma or n are uniform over the object surface.

This is true for all illumination directions $(i = 1, \cdots, n)$. One can say that the temporal variations in their appearance also reflect the differences in their surface normals, and the appearance changes of the specular surface should thus construct an appearance manifold.

4.4 Surfaces with both components

If the object surfaces consist of both uniform specular and diffuse reflectance components, the intensities of their surface points are computed as the addition of its diffuse component and specular component from the Dicromatic reflection model. Based on our analyses of Lambertian surfaces (Section 4.1) and specular surfaces (Section 4.3), it can be said the differences among the observation vectors are still due to the differences in the surface normals $\vec{n_p}$ among p. Our algorithm is therefore able to extract the surface normals of objects with both diffuse reflectance and specular reflectance properties from their appearance manifold.

The most complex reflectance model is a textured specular and diffuse reflectance model. This type reflectance property object is still a challenge for our approach.

5 Experimental Results

We have tested our proposed approach for shape recovery by using both synthetic and real data.

5.1 Synthetic Data

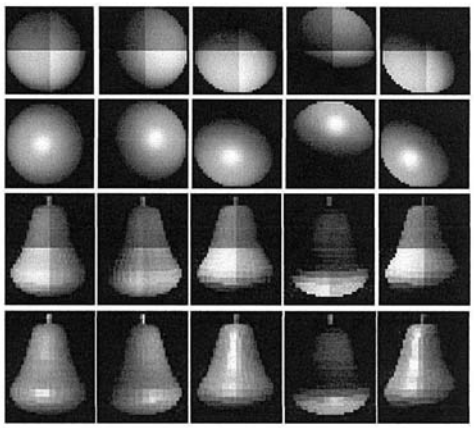

Figure 2: Input images of the CG objects: (1)a diffuse and texture sphere; (2)a diffuse and specular sphere;(3)a diffuse and texture pear;(4)a diffuse and specular pear.

The images of objects with different surface materials are synthetically provided under sparsely distributed 200 light sources based on the Lambertian and the Torrance-Sparrow reflection models: (1) a textured lambertian surface and (2) a surface with both diffuse and specular reflection components (kd = 0.6, ks = 0.4, and $\sigma = 0.15$). Figure 2 shows the shapes and surface normals of the objects and some of the input images used for the experiments. Each image contain about 2500 surface points, and thus 200-dimensional input space consists of 2500 observation vectors. SurNormal maps are images that store normals directly in the RGB values of an image.

Three-dimensional embedding of this input space is estimated by Isomap and then transformed to the distribution of surface normals of the object by using its occluding boundary as reference. In the process of dimensionality reduction, the parameter k that represents the number of neighbors needs to be adjusted by using the results.

The recovered shape and surface normals shown in figure4 and figure5 highly resemble the ground truth shown in figure3 for both all four examples. The quantitative evaluation of the estimated surface normals is provided in Table1. Here the root mean square error μ and its variance σ are computed to evaluate the accuracy. For all examples, μ is within 8 degree with small variances. It can be said from this that our purely image-based algorithm could achieve reasonably high accurate without assuming any analytic reflection models to describe the appearance of the objects.

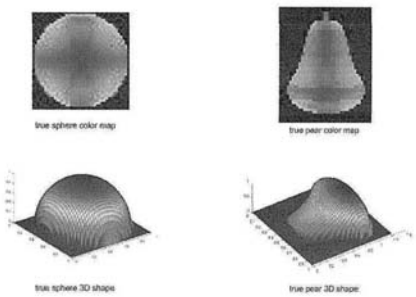

Figure 3: True surface normals of the CG objects.

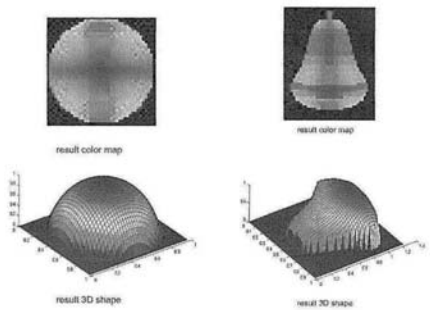

Figure 4: Results of the diffuse and texture objects.

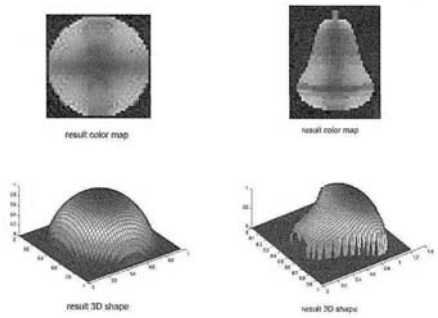

Figure 5: Results of the diffuse and specular objects.

shape	diffuse	specular	texture	μ	σ
sphere	Y	N	Y	0.1474	0.030
sphere	Y	Y	N	0.0606	0.020
pear	Y	N	Y	0.1390	0.0072
pear	Y	Y	N	0.1295	0.0052

Table 1: Errors of the CG sphere and the CG pear.

5.2 Real objects

The process of testing the real objects is to set the camera to take the images at regular intervals as we are waving a light source around the object from the approximate fixed distance. Here we need to pay attention to avoid making the cast shadow on the objects. This process is easy to implement and do not need long time.

The orange (plastic), cat (ceramic)is tested for our method. For each object, 180 $\sim$ 200 images are captured by moving a point light source around the object. The point light source and a camera was roughly 1m away from the object (5 $\sim$ 10cm in diameter). Figure 6 shows some images for a plastic orange and the surface normals results are shown by the color map and 3D shape in the figure 7. Figure 8 shows some images for a ceramic cat and the results are shown in the figure 9. Furthermore, interestingly enough, even the fine details of the shapes, e.g, the tail of the cat, were extracted as we see in the result figures. This shows that how significant it is to use the appearance changes observed on an object surface for discovering the object's shape.

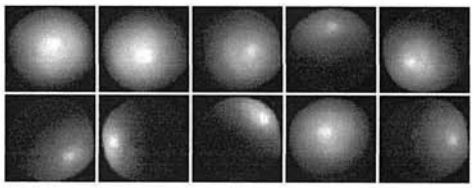

Figure 6: Some input images of a plastic orange

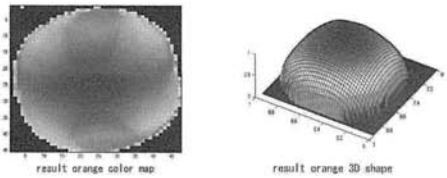

Figure 7: Results of a plastic orange

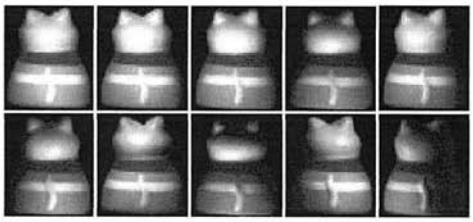

Figure 8: Some input images of a ceramic cat

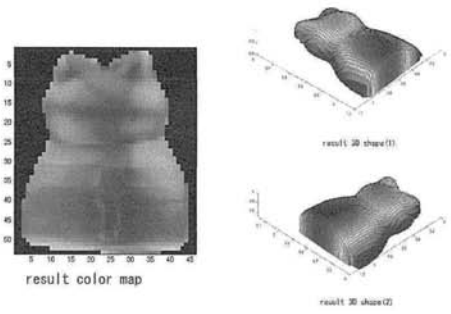

Figure 9: Results of a ceramic cat

6 Future Work

We can estimate the objects' surface normals based on the objects' different images in the different illuminations. But some object such as a texture object with specular and diffuse properties is still a challenge object for our method because these pixels vectors' differences are not only caused by the surface normals. Our method can be suitable by doing some additional modification for the input images such as separating the images into specular part and diffuse part. The theoretical analysis for the number of required images and the applicable material types is also the future work for our research.

References

[BJ03] Ronen Basri and David W. Jacobs. Lambertian reflectance and linear subspaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(2):218–233, 2003.

[BKY99] P. N. Belhumeur, D. J. Kriegman, and A. L. Yuille. The bas-relief ambiguity. *International Journal of Computer Vision*, 35(1):33–44, 1999.

[Hor86] B. Horn. Robot vision. *MIT Press, Cambridge, MA*, 1986.

[HS05] A. Hertzmann and S. M. Seitz. Example-based photometric stereo: Shape reconstruction with general, varying brdfs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1254–1264, 2005.

[KH86] Berthold Klaus and Paul Horn. *Robot Vision.* MIT Press, 1986.

[KJ79] J.T.Kent K.V.Mardia and J.M.Bibby. *Multivariate Analysis.* Academic Press, 1979.

[KN06] S. Koppal and S. G. Narasimhan. Clustering appearance for scene analysis. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1323–1330, 2006.

[NIK90] S. Nayar, Katsushi Ikeuchi, and Takeo Kanade. Determining shape and reflectance of hybrid surfaces by photometric sampling. *IEEE Transactions on Robotics and Automation*, 6(4):418–431, August 1990.

[Sil80] W. M. Silver. Determining shape and reflectance using multiple images. *Master's thesis, MIT, Cambridge,MA*, 1980.

[TS67] K. E. Torrance and E. M. Sparrow. Theory for off-specular reflection from roughened surface. *J. Optical Society of America*, 57:1105–1114, 1967.

[TSL00] J. B. Tenenbaum, V. Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290:2319–2323, 2000.

[VH06] George Vogiatzis and Carlos Hernandez. Reconstruction in the round using photometric normals and silhouettes. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1847–1854, 2006.

[War92] G. J. Ward. Measuring and modeling anisotropic reflection. *Proc. SIGGRAPH '92*, pages 265–272, 1992.

[Woo81] Robert J. Woodham. Analysing images of curved surfaces. *Artif. Intell.*, 17(1-3):117–140, 1981.

[ZTCS99] Ruo Zhang, Ping-Sing Tsai, James Edwin Cryer, and Mubarak Shah. Shape from shading: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence.*, 21(8):690–706, 1999.