Shape Reconstruction Based on Similarity in
Radiance Changes under Varying Illumination

Imari Sato
National Institute of Informatics
imarik@nii.ac.jp

Takahiro Okabe   Qiong Yu   Yoichi Sato
The University of Tokyo
{takahiro,qiong,ysato}@iis.u-tokyo.ac.jp

Abstract

This paper presents a technique for determining an object's shape based on the similarity of radiance changes observed at points on its surface under varying illumination. To examine the similarity, we use an observation vector that represents a sequence of pixel intensities of a point on the surface under different lighting conditions. Assuming convex objects under distant illumination and orthographic projection, we show that the similarity between two observation vectors is closely related to the similarity between the surface normals of the corresponding points. This enables us to estimate the object's surface normals solely from the similarity of radiance changes under unknown distant lighting by using dimensionality reduction. Unlike most previous shape reconstruction methods, our technique neither assumes particular reflection models nor requires reference materials. This makes our method applicable to a wide variety of objects made of different materials.

1. Introduction

The appearance of an object is determined by several factors including illumination, viewing geometry, and the shape and surface material of the object. Changing any one of these factors should change the object's appearance. These factors are, in general, nonlinearly related to the object's appearance, and consequently an inverse problem of estimating them becomes difficult unless there is some a priori knowledge of the scene. Most of the previous techniques for solving this inverse problem thus try to estimate some of these factors from images of a scene by assuming that the others are known.

Previous studies of shape reconstruction have demonstrated that the shape of an object can be recovered from multiple images of the object taken under various lighting conditions assuming that there is some knowledge of the scene, e.g., illumination and/or surface materials [19, 20, 7, 12].[1]

---

[1]Most of the shape-from-shading approaches estimate an object's shape from a single image assuming distant illumination and uniform Lambertian reflectance [21].
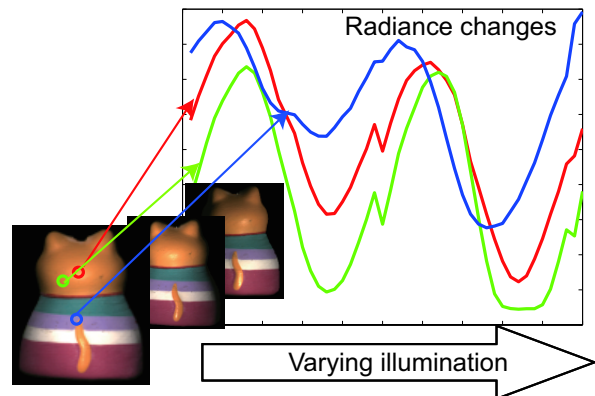

Figure 1. Determining an object's shape based on similarity of radiance changes observed at its surface points.

The classical photometric stereo approach presented in [20, 7] recovers the shape of a Lambertian object from multiple images of the object taken under known light sources. Later, this approach was extended to several directions including photometric stereo methods for non-Lambertian surfaces [8, 3, 12, 4] (see [19] for a good survey) and uncalibrated photometric stereo methods [5, 1, 2].

However, real-world materials sometimes have complex appearances that prevent us from extracting their shapes using analytic reflectance models. To cope with this problem, Hertzmann and Seitz proposed the use of a calibration object [6], which had also been suggested in early work on photometric stereo [7]. Instead of computing a reflectance map based on an analytic reflection model, images of a calibration object of a known shape made of the same material are captured under the same lighting conditions as a target object and used as an empirical reflectance map. Later, Hertzmann and Seitz extended their method so that instead of using a single calibration object, it uses a small number of them to estimate the shape of objects with non-uniform surface materials. However, the issue of what kinds of reference materials are suitable for representing the complex appearances of a wide variety of real-world materials still remains to be addressed.

Recently, Koppal and Narasimhan presented a clustering-based technique to find iso-surface normal clusters of scene points without requiring knowledge about materials and lighting [9]. Their approach shows how effective it is to analyze the temporal variation in the appearance of a scene for obtaining its meaningful geometric structure; e.g., each face of a planar object is found regardless of its texture. However, since the method finds iso-surface normal clusters as the geometric structure of the scene, the 3D shape of the scene cannot be obtained unless another shape reconstruction technique such as uncalibrated photometric stereo is used, as suggested in their work.

In this work, we present a method for recovering an object's shape directly from input images of the object based on *the similarity of the radiance changes* observed at points on its surface under varying illumination (Figure 1). To examine the similarity, we consider an observation vector that represents a sequence of pixel intensities of a surface point under different lighting conditions. Assuming convex objects under distant illumination and orthographic projection, we show that the similarity between two observation vectors is closely related to the similarity between surface normals of the corresponding points.

This gives rise to the interesting theory that within the space containing a set of observation vectors, all of the observation vectors should lie on a manifold that represents the surface normals of the corresponding surface points. Accordingly, we can find the surface normals of an object by using a dimensionality reduction technique that best preserves the intrinsic structure in the input space of observation vectors as illustrated in Figure 2. Our method is capable of estimating the dense distribution of an object's surface normals by analyzing the similarities of radiance changes between scene points.

As far as we know, this is the first attempt to discover an object's shape solely from a set of images of the object based on manifold learning techniques. Unlike most previous shape reconstruction methods, our technique neither uses particular reflection models, nor requires empirical reflection maps. This makes our method applicable to a wide variety of objects with different types of material.

The rest of this paper is organized as follows. Section 2 considers the similarity of radiance changes observed at points on an object's surface under varying illumination, and Section 3 outlines our idea of estimating an object's surface normals from the similarity of radiance changes. Section 4 discusses the issue of how the similarity of two observation vectors is related to the similarity of their surface normals. Section 5 describes our algorithm for recovering the object's shape by using a dimensionality reduction technique. Section 6 shows experimental results for synthetic and real data and Section 7 presents concluding remarks.
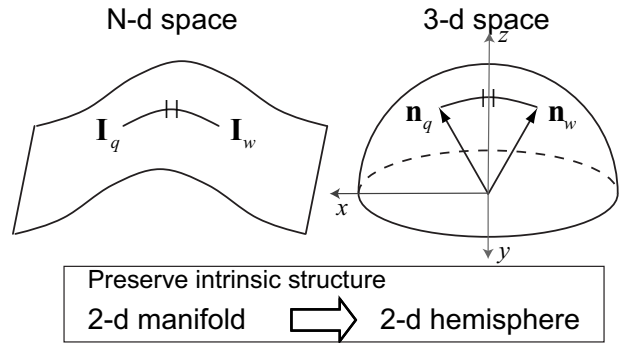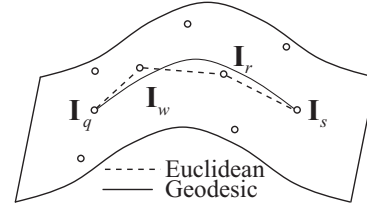

Figure 2. 3-d embedding of observation vectors.

Figure 3. Geodesic distance measured along a manifold.

2. Similarity of Radiance Changes

Let us first analyze the similarity of radiance changes observed at points on an object's surface under n different light sources distributed around the object to illuminate it from various directions. Suppose I_p^k is the brightness of each pixel or each corresponding surface point p seen under the kth illumination with unit radiance.

A vector containing the set of I_p^k $(k = 1, 2, \cdots, n)$ in its elements is normalized

$$\mathbf{I}_p = \frac{\left(I_p^1, \cdots, I_p^n\right)^T}{\sqrt{\sum_{k=1}^{n}(I_p^k)^2}} \quad (1)$$

and referred to as an *observation vector*. From input images with m surface points $(p = 1, \cdots, m)$, we obtain m observation vectors. In the following, a n-dimensional space containing a set of m observation vectors is referred to as an *observation space*.

Assuming distant light, orthographic projection, isotropic reflection, and no local effects such as cast shadows and interreflections, we consider the brightness of p under the kth illumination described in the form:

$$I_p^k = \rho_p g(\theta_k', \theta_v', \phi_k' - \phi_v'), \quad (2)$$

where (θ_k', ϕ_k') and (θ_v', ϕ_v') are the incident and reflection directions defined in the local coordinates with respect to the point p. Function g represents how much of the incident light from the direction (θ_k', ϕ_k') is reflected on the point toward (θ_v', ϕ_v'). ρ_p is a reflection coefficient that varies over the surface.

Let $\mathbf{n}_p$ be the surface normal of p and $\mathbf{l}_k$ and $\mathbf{v}$ denote the incident direction of the kth illumination and the direction to the viewer represented by the global coordinate, respectively. Note that $\mathbf{l}_k$ and $\mathbf{v}$ are constant over all surface points of the object under the assumptions of distant illumination and orthographic projection.

The angles θ'_k, θ'_v, and $\phi'_k - \phi'_v$ can be computed from $\mathbf{n}_p$, $\mathbf{v}$, and $\mathbf{l}_k$,[2] and (2) is thus rewritten as

$$I_p^k = \rho_p g'(\mathbf{n}_p, \mathbf{l}_k, \mathbf{v}), \qquad (3)$$

where the function g' represents how much of the incident light $\mathbf{l}_k$ is reflected on the point of its normal $\mathbf{n}_p$ toward $\mathbf{v}$.

From the normalization process in (1), reflectance coefficient ρ_p comes to have no influence on the observation vector $\mathbf{I}_p$. From this, we see that differences in the observation vectors among all surface points are caused by differences in their surface normals. Namely, $\mathbf{I}_q = \mathbf{I}_w$ is equivalent to $\mathbf{n}_q = \mathbf{n}_w$ for surface points q and w, which was used as the orientation-consistency cue in [6].

As we will see later in Section 4, the similarity between two observation vectors is closely related to the similarity between the surface normals of the corresponding surface points. This enables us to estimate the distribution of an object's surface normals by analyzing the similarities of its observation vectors.

Note that the form in (2) can represent a wide variety of BRDFs including commonly used reflection models such as Lambertian, Oren-Nayar [13], Phong, Torrance-Sparrow[16], and isotropic Gauss[17] models. On the other hand, some BRDFs do not follow this form exactly: e.g., a dichromatic reflection model whose relative strength between its diffuse and specular components varies over the object surface. However, as shown later in our experimental results, our method is still able to recover object shapes reasonably well even for BRDFs that do not follow (2) exactly.

3. Shape Recovery by Embedding of Observation Vectors

The set of m observation vectors seen under n different light sources can be thought of as m points in n-dimensional space. As we see in the previous section, differences in the observation vectors among all surface points are caused by differences in their surface normals in the case where their n elements are described in the form (2).

It can be further said from this that within the observation space, all of the m points lie on an intrinsically two-dimensional manifold that represents the surface normals of the corresponding m surface points because surface normals have two degrees of freedom.[3]

The key idea of our work is that since surface normals are defined on a unit sphere in a three-dimensional space, if we find three-dimensional embedding of the observation vectors using the dimensionality reduction that best preserves its intrinsic structure, the resulting three-dimensional space should reveal the distribution of the object's surface normals as illustrated in Figure 2.

It is important to note that what we count on here is the two-dimensional manifold hidden in the observation space and the similarity of two observation vectors needs to be determined by their shortest distances measured along a manifold. Such a distance is called *geodesic distances*.

In our method, a nonlinear embedding method, called isometric feature mapping (Isomap) is used for dimensionality reduction [15].[4] Isomap assumes that, for neighboring data points, distances defined in an input space such as a Euclidean distance provides good approximations of the geodesic distances. Accordingly geodesic distances for faraway points are computed by adding up a sequence of distances between neighboring points.

For instance, in Figure 3, the geodesic distance between two observation vectors $\mathbf{I}_q$ and $\mathbf{I}_s$ denoted to as $d_G(q, s)$ is approximated by adding up a sequence of Euclidean distances between neighboring points as

$$d_G(q, s) \simeq |\mathbf{I}_q - \mathbf{I}_w| + |\mathbf{I}_w - \mathbf{I}_r| + |\mathbf{I}_r - \mathbf{I}_s|.$$

Tenenbaum et al. demonstrated that Isomap could find low-dimensional representations of natural images, e.g., images with different poses and lighting directions [15]. Previous studies have also shown that scene conditions under which input images are captured could be recovered through dimensionality reduction: Pless and Simon estimated viewing directions or object poses in [14], and Winnemoller et al. recovered the light source directions in [18]. However, no theoretical insight into why dimensionality reduction is effective in extracting such information from the input image space has been provided.

In the next section, we will show that the Euclidean distance between two neighboring observation vectors is indeed a good approximation of the difference between their surface normals. This means that Isomap is suitable for revealing the intrinsic structure of the observation space.

It should be also noted that previously proposed methods of analyzing images using dimensionality reduction aimed at recovering scene conditions under which images of unknown subjects are captured, e.g., viewing direction or light source directions. On the other hand, our goal is to extract the information about a photographed subject from images captured under unknown conditions.

---

[2]Since $\cos\theta'_k = \mathbf{n}_p^T\mathbf{l}_k$, $\cos\theta'_v = \mathbf{n}_p^T\mathbf{v}$, $\cos(\phi'_k - \phi'_v) = [\mathbf{l}_k^T\mathbf{v} - (\mathbf{n}_p^T\mathbf{l}_k)(\mathbf{n}_p^T\mathbf{v})]/[\sin\cos^{-1}(\mathbf{n}_p^T\mathbf{l}_k)\sin\cos^{-1}(\mathbf{n}_p^T\mathbf{v})]$.

[3]Note that this manifold has two dimensions because surface normals (n_x, n_y, n_z) are described in elevation and azimuthal angles (θ, ϕ).

[4]The set of observation vectors generally has intrinsically nonlinear structures, and classical techniques for dimensional reduction, such as principal component analysis (PCA) and multidimensional scaling (MDS) are not effective in revealing such structures.

4. Theoretical Insight into Similarity of Observation Vectors

This section discusses the important issue of how the similarity of two observation vectors is related to the similarity of their surface normals. To get some insight into good similarity measures for finding a manifold representing surface normals, we derive a distance measure for two commonly used reflectance models.

We show that the obtained measures are approximately equal to or closely related to the Euclidean distance $|\mathbf{I}_i - \mathbf{I}_j|$ when the distance between two observation vectors of two surface points i and j is small ($\mathbf{I}_i \simeq \mathbf{I}_j$). Note that, even though the two similarity measures are derived from parametric reflection models, our method is by no means restricted to particular reflection models.

Recall that Isomap assumes that, for neighboring data points, a distance defined in an input space provides good approximations of the geodesic distances. If two observation vectors are detected as neighbors with a small distance, this should also indicate that the distances between their surface normals are small.

4.1. Discussion of Body Reflection

First, we consider radiance changes of body reflection as an example. As shown in Appendix A, assuming that a convex object with a Lambertian surface is illuminated by a set of directional light sources distributed around the object with uniform density, the inner product of observation vectors $\mathbf{I}_i$ and $\mathbf{I}_j$ is given by

$$\mathbf{I}_i^T \mathbf{I}_j \simeq \mathbf{n}_i^T \mathbf{n}_j, \tag{4}$$

when the corresponding surface normals satisfy $\mathbf{n}_i \simeq \mathbf{n}_j$.

Because the observation vectors are normalized so that they have unit length, it is clear that

$$\mathbf{I}_i^T \mathbf{I}_j = 1 - (|\mathbf{I}_i - \mathbf{I}_j|^2/2). \tag{5}$$

In addition, when $\mathbf{I}_i \simeq \mathbf{I}_j$, the angle between $\mathbf{n}_i$ and $\mathbf{n}_j$ is small. Then, the following approximation

$$\mathbf{n}_i^T \mathbf{n}_j = \cos[\arccos(\mathbf{n}_i^T \mathbf{n}_j)] \simeq 1 - \frac{\arccos^2(\mathbf{n}_i^T \mathbf{n}_j)}{2} \tag{6}$$

holds. From (4), (5), and (6), the Euclidean distance between observation vectors $\mathbf{I}_i$ and $\mathbf{I}_j$ is given by

$$|\mathbf{I}_i - \mathbf{I}_j| \simeq \arccos(\mathbf{n}_i^T \mathbf{n}_j), \tag{7}$$

when $\mathbf{n}_i \simeq \mathbf{n}_j$. This means that $|\mathbf{I}_i - \mathbf{I}_j|$ is approximately equal to the ideal distance that is the angle between $\mathbf{n}_i$ and $\mathbf{n}_j$ that represents their relationship.

4.2. Discussion of Surface Reflection

Second, we investigate radiance changes of surface reflection as another example. As shown in Appendix B, the inner product of observation vectors $\mathbf{I}_i$ and $\mathbf{I}_j$ is given by

$$\mathbf{I}_i^T \mathbf{I}_j \simeq \exp\left[-\frac{\arccos^2(\mathbf{n}_i^T \mathbf{n}_j)}{2\sigma^2}\right] \tag{8}$$

for the simplified Torrance-Sparrow model, when $\mathbf{n}_i \simeq \mathbf{n}_j$ and surface roughness $\sigma \ll \pi$.

Because $\mathbf{n}_i \simeq \mathbf{n}_j$ is equivalent to $\mathbf{I}_i \simeq \mathbf{I}_j$, the following approximation

$$\ln(\mathbf{I}_i^T \mathbf{I}_j) = \ln\left(1 - \frac{|\mathbf{I}_i - \mathbf{I}_j|^2}{2}\right) \simeq -\frac{|\mathbf{I}_i - \mathbf{I}_j|^2}{2} \tag{9}$$

holds. Hence, the Euclidean distance between observation vectors $\mathbf{I}_i$ and $\mathbf{I}_j$ is given by

$$|\mathbf{I}_i - \mathbf{I}_j| \simeq \frac{\arccos(\mathbf{n}_i^T \mathbf{n}_j)}{\sigma}. \tag{10}$$

In other words, the distance $|\mathbf{I}_i - \mathbf{I}_j|$ is proportional to the angle between $\mathbf{n}_i$ and $\mathbf{n}_j$. This means that the geodesic distance is approximately equal to the ideal distance up to some scale factor. Note that this scale factor can be removed by normalizing geodesic distances between all pairs of observation vectors so that $0 \leq d_G(i, j) \leq \pi$.

5. Algorithm for Shape Recovery

Given m sets of n-dimensional observation vectors, the surface normals and heights of the corresponding m surface points are estimated as described in the following steps:

1. Provide a geodesic distance matrix that contains geodesic distances between all pairs of observation vectors (Section 5.1).
2. Find the three-dimensional embedding of the observation vectors based on the geodesic distance matrix (Section 5.2).
3. Recover surface normals and object heights. The occluding boundary is used as a reference to transform the output from Isomap to the true distribution of the object's surface normals (Section 5.3).

5.1. Geodesic Distance Matrix

The geodesic distances between all pairs of observation vectors are approximated by adding up a sequence of Euclidean distances between neighboring points as follows:

1. Compute Euclidean distances $d(i, j)$ between all pairs of observation vectors $\mathbf{I}_i$ and $\mathbf{I}_j$ as

$$d(i, j) = |\mathbf{I}_i - \mathbf{I}_j|. \tag{11}$$

2. Initialize a graph G where points i and j are connected if i is one of the k nearest neighbors of j based on $d(i, j)$. Geodesic distances $d_G(i, j)$ for all i, j pairs of m observation vectors are initialized such that $d_G(i, j) = d(i, j)$ if i and j are connected as neighbors and $d_G(i, j) = \infty$ otherwise.

3. Compute the shortest path distances between all pairs of observation vectors in G by using Floyd's algorithm and update the geodesic distance $d_G(i, j)$. This provides a matrix of the geodesic distances between all pairs of observation vectors $D_g = \{d_G(i, j)\}$.

5.2. Dimensionality Reduction Based on Geodesic Distance

After the matrix of geodesic distances D_g is obtained, classical MDS is applied to this matrix to find a three-dimensional embedding of the observation vectors such that the estimated intrinsic geometry is best preserved through dimensionality reduction.

Let $\mathbf{e}_p$ be a three-dimensional vector corresponding to observation vector $\mathbf{I}_p$ after dimensionality reduction. Then $\mathbf{e}_p$ are determined by minimizing the cost function

$$E = |\tau(D_g) - \tau(D_3)|, \tag{12}$$

where D_3 is the matrix of Euclidean distances in the three-dimensional space $d_3(i, j) = |\mathbf{e}_i - \mathbf{e}_j|$ and $|A|$ is the Frobenius norm $\sqrt{\sum_{ij} A_{ij}^2}$ of matrix A. The τ operator converts distances to inner products to support efficient optimization.[5] By setting the coordinates of $\mathbf{e}_p$ to the top three eigenvectors of the matrix $\tau(D_g)$, we are able to find the global minimum of the cost function (12) [11].

5.3. Solving Ambiguity of Object's Surface Normals

There are distance-preserving transformations such as translations, reflections, and rotations in a three-dimensional space, and the estimated three-dimensional vector $\mathbf{e}_p$ does not necessarily correspond to the true distribution of the object's surface normals. In other words, what Isomap computes is the relative relationships of these surface normals. In addition, $\mathbf{e}_p$ does not necessarily has a unit length. For this reason, we need a further step to solve this ambiguity and obtain the true distribution of the object's surface normals. We project $\mathbf{e}_p$ onto a unit sphere to ensure $|\mathbf{e}_b| = 1$ and find a valid transformation that projects $\mathbf{e}_p$ onto the unit sphere based on the known surface normals along the occluding boundary .

Under the orthographic projection viewed from the direction $(0, 0, 1)$, the true surface normals of the occluding boundary can be obtained as their gradient directions in the 2D input image coordinate systems. Let $\mathbf{c}_b$ be the obtained true surface normals of those boundary points and $\mathbf{e}_b$ be their corresponding embedded vectors. Then a transformation $\varsigma()$ that projects the embedded vectors $\mathbf{e}_b$ onto a unit sphere is found by minimizing the cost function $\sum_{\text{all } b} |\mathbf{c}_b - \varsigma(\mathbf{e}_b)|$. Here the center and orientation ($\theta = 0$ and $\phi = 0$ directions) of the unit sphere defines this transformation, and $\varsigma()$ are thus estimated with respect to those parameters.

[5]$\tau(D) = -HSH/2$, where S is the matrix of squared distances $S_{ij} = D_{ij}^2$ and H is the centering matrix $H_{ij} = \delta_{ij} - 1/m$. See Tenenbaum [15] and Mardia [11] for details.

After projecting all embedded vectors $\mathbf{e}_p$ onto the unit sphere based on the estimated $\varsigma()$, the recovered surface normals are stored in their original pixel locations to provide a normal field, as shown in Figure 6(b). Then the height field of the object is estimated from the recovered normal field by using the shapelet technique [10].

6. Experimental Results

6.1. Synthetic Data

We evaluated the accuracy of our technique for shape reconstruction using synthetic data. The images of objects with different surface materials were synthesized under 450 distant light sources that were randomly distributed around the objects with (1) diffuse (textured Lambertian), (2) specular (Torrance-Sparrow), (3) diffuse (uniform Lambertian) + specular (Torrance-Sparrow), and (4) diffuse (textured Lambertian) + specular (Torrance-Sparrow) surfaces.

The residual variance after the dimensionality reduction is shown in Figures 4. We can clearly see that residual variances decrease as the dimensionality increases and the curve levels off after three dimensions. This indicates that three-dimensional embedding of the observation vectors were successfully found. It should be noted that in the process of dimensionality reduction, there is only one parameter k representing the number of neighbors, which was adjusted to minimize the cost function (12).

The estimated surface normals and shapes of the objects and some of the input images used for this experiments are shown in Figure 5. In this figure, the estimated surface normals are stored directly in the RGB values of the surface normal map images where absolute values of x, y, and z coordinates are shown in red, green, and blue, respectively: x and y axes correspond to the horizontal and vertical axis of the image plane, and z axis is defined to be perpendicular to the image plane.

Compared with the ground truth provided in Figure 5 (a), we clearly see that our technique was able to estimate the surface normals and shapes of the object quite well for all four cases in spite of the fact that they were recovered solely from input images without requiring light source directions or reflection models.

Table 1 provides a quantitative evaluation of the technique that summarizes the root-mean-square errors in the estimated

Case	Diffuse	Texture	Specular	Distance	Normal	Height
(1)	⋄	⋄		d	5.7	0.066
				d_d	5.7	0.066
(2)			⋄	d	8.5	0.093
				d_s	6.9	0.052
(3)	⋄		⋄	d	5.2	0.059
(4)	⋄	⋄	⋄	d	6.0	0.072

Table 1. Root-mean-square errors in the estimated surface normals and object's heights.

(a) synthetic data   (b) real data
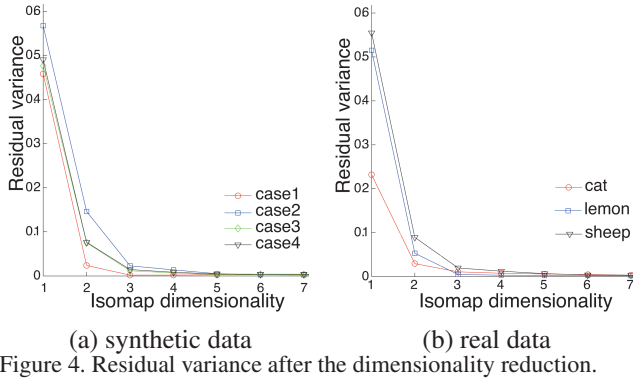
Figure 4. Residual variance after the dimensionality reduction.

(a) ground truth   (b) some input images

(c) case 1   (d) case 2

(e) case 3   (f) case 4

Figure 5. Synthetic data: estimated surface normals and object heights.
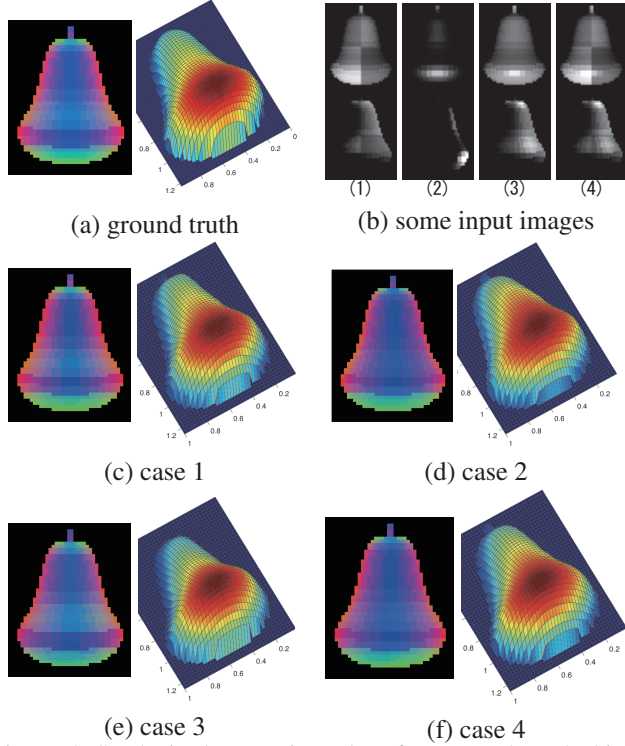

Figure 6. Results for real images: (a) input images, (b) surface normals, and (c) object's heights.

surface normal directions and heights for all cases. Here, the error in surface normal directions are shown in degrees, and the heights are normalized between 0 and 1. For cases (1) and (2), we also tested the ideal distances analytically derived in the previous section. The distance between two neighboring observation vectors $d(i,j) = |\mathbf{I}_i - \mathbf{I}_j|$ defined in (11) was replaced with the ideal distances derived from (4) and (8): $d_d(i,j) = \arccos(\mathbf{I}_i^T \mathbf{I}_j)$ for Lambertian, and $d_s(i,j) = \sqrt{-\ln(\mathbf{I}_i^T \mathbf{I}_j)}$ for the simplified Torrance-Sparrow model. The column Distance in this table indicates which distance was used for the estimation.

In Table 1, for all cases, our method could estimate surface normals and object's heights with reasonably high accuracy. This shows that our algorithm successfully found the intrin-

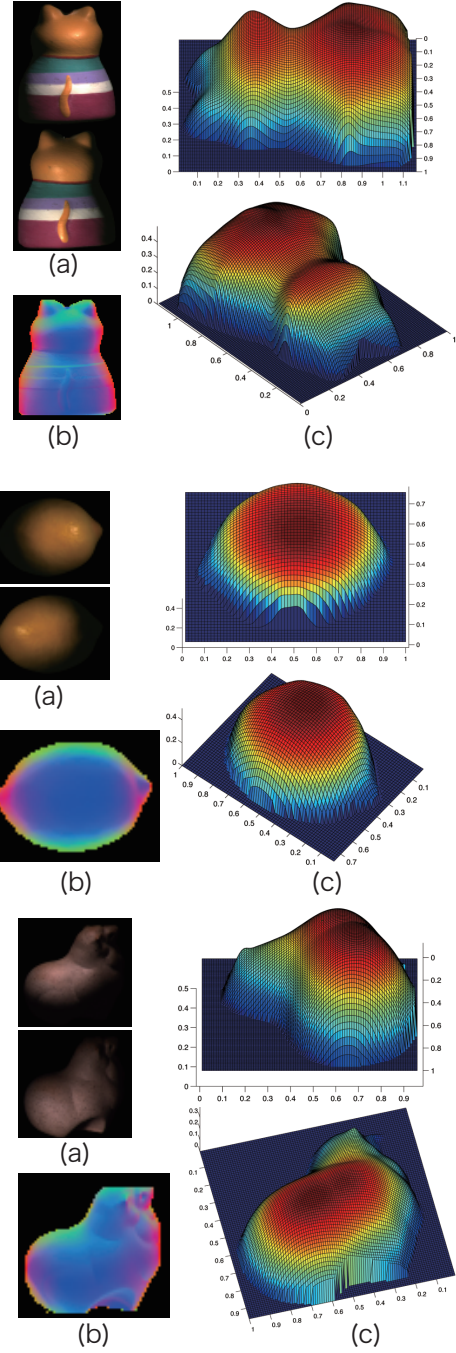sic structure of the observation space to be used for shape reconstruction. Especially, in the cases of (1) and (2), about the same accuracy was achieved for the surface normals using the distance d and the ideal distances d_d and d_s to measure the similarity between two neighboring observation vectors. This indicates that the distance d well approximates the geodesic distance between these vectors.

On the other hand, strictly speaking, case (4) does not satisfy the necessary conditions of our approach since its radiance changes cannot be described in the form (2). Nevertheless, we could achieve about the same accuracy as in case (3). Body reflection is often observed at most points on an object's surface under varying illumination, while specular reflection are less likely to be observed. We consider that the diffuse reflection components contributed more to the shape reconstruction than did the specular reflection components in this case.

6.2. Real Objects

We also tested our technique using real images of several objects with different surface materials. Figure 6(a) shows the objects used in this experiment: cat (painted bisque), lemon (plastic), and sheep (ceramic) ornaments. Figure 6 (b) and (c) show the estimated surface normals and the objects' heights. For each object, $130 \sim 150$ images were captured by moving a point light source around the object, which was roughly 1 m away from the object of $5 \sim 10$ cm in diameter.

In this figure, we see that our method was able to estimate their shapes regardless of the complexity of their surface materials and shapes. In the case of the cat ornament, this object has both convex and concave shapes and thus some cast shadows were observed in the images. Strictly speaking, this does not follow our assumption of a convex object. However, it was found through the experiment that cast shadows did not cause that much trouble in estimating an object's surface normals as long as its unshadowed appearances were observed enough in other input images under different lighting conditions. It seems that our method is less sensitive to this kind of obstruction in cases where observation vectors contain correct values for most of their elements.

7. Conclusion

We presented a novel technique for recovering an object's shape directly from input images of the object that is based on the similarity of the radiance changes observed at points on its surface under varying illumination. The main contribution of our work is that we have shown that the similarity of the radiance changes is closely related to the similarity between surface normals of the corresponding points and introduced a new theory that within the observation space, all of the observation vectors of these points lie on a manifold that represent their surface normals. This enables us to find surface normals of an object by using a dimensionality reduction technique that best preserves the intrinsic structure in the observation space as captured in geodesic distances on a manifold. A future research direction of this work would be to analyze the manifold in more detail by using objects of known shapes with various kinds of surface materials and to investigate how many images are generally sufficient in order to obtain an object's shape correctly.

Appendix A

Assuming the Lambertian model, let us consider radiance changes observed at a point i on an object's surface under a set of directional light sources $\{l_1, l_2, \cdots, l_F\}$. Here, the light source vector l_k is a product of its brightness and a unit vector representing the direction of the light source, so the unnormalized observation vector $\mathbf{I}'_i$ ($\mathbf{I}_i = \mathbf{I}'_i / \sqrt{|\mathbf{I}'_i|^2}$) is given by

$$\mathbf{I}'_i = (\max(\rho_i \mathbf{n}_i^T l_1, 0), \cdots, \max(\rho_i \mathbf{n}_i^T l_F, 0))^T, \qquad (13)$$

where ρ_i and $\mathbf{n}_i$ are the albedo and normal, respectively. Note that we represent attached shadows using operator $\max(*, 0)$ but do not take account of cast shadows.

The inner product of vectors $\mathbf{I}'_i$ and $\mathbf{I}'_j$ is derived as

$$\mathbf{I}'_i{}^T \mathbf{I}'_j = \sum_{k=1}^{F} \max(\rho_i \mathbf{n}_i^T l_k, 0) \max(\rho_j \mathbf{n}_j^T l_k, 0) \qquad (14)$$

$$= \rho_i \rho_j \mathbf{n}_i^T \left(\sum_{\{k | \mathbf{n}_i^T l_k > 0, \mathbf{n}_j^T l_k > 0\}} l_k l_k^T \right) \mathbf{n}_j, \qquad (15)$$

where $\sum_{\{k | \mathbf{n}_i^T l_k > 0, \mathbf{n}_j^T l_k > 0\}}$ means the summation over light source vectors satisfying both $\mathbf{n}_i^T l_k > 0$ and $\mathbf{n}_j^T l_k > 0$. That is, the summation is taken when neither i nor j are in attached shadows.

When $\mathbf{n}_i \simeq \mathbf{n}_j$, the summation is approximated as

$$\mathbf{I}'_i{}^T \mathbf{I}'_j \simeq \rho_i \rho_j \mathbf{n}_i^T \left(\sum_{\{k | (\mathbf{n}_i + \mathbf{n}_j)^T l_k > 0\}} l_k l_k^T \right) \mathbf{n}_j \qquad (16)$$

$$= \rho_i \rho_j \mathbf{n}_i^T \begin{pmatrix} \overline{l_x^2} & \overline{l_x l_y} & \overline{l_x l_z} \\ \overline{l_x l_y} & \overline{l_y^2} & \overline{l_y l_z} \\ \overline{l_x l_z} & \overline{l_y l_z} & \overline{l_z^2} \end{pmatrix} \mathbf{n}_j, \qquad (17)$$

because both i and j are in attached shadows for almost the same light sources. Here, we denote $l_k = (l_{kx}, l_{ky}, l_{kz})^T$ and define $\overline{l_x^2} = \sum_k l_{kx}^2$.

Assuming that directional light sources with brightness l distribute around the object with uniform density, the above summations can be converted to integrals such as

$$\overline{l_x^2} = l^2 \int_0^{\pi/2} \int_0^{2\pi} (\sin\theta \cos\phi)^2 \sin\theta d\theta d\phi = \frac{2\pi}{3} l^2. \qquad (18)$$

By computing the other elements of the matrix in (17) similarly, we obtain

$$\mathbf{I}'_i{}^T \mathbf{I}'_j \simeq \frac{2\pi}{3} l^2 \rho_i \rho_j \mathbf{n}_i^T \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{n}_j \qquad (19)$$

The norm of $\mathbf{I}'_i$ is given by

$$|\mathbf{I}'_i|^2 = \rho_i^2 l^2 \int_0^{\pi/2} \int_0^{2\pi} (\cos\theta)^2 \sin\theta d\theta d\phi = \frac{2\pi}{3} l^2 \rho_i^2. \qquad (20)$$

Therefore, the inner products of the normalized observation vectors are related to the inner products of the surface normals at the corresponding points as

$$\mathbf{I}_i^T \mathbf{I}_j = \frac{\mathbf{I}'_i{}^T \mathbf{I}'_j}{\sqrt{|\mathbf{I}'_i|^2 |\mathbf{I}'_j|^2}} = \mathbf{n}_i^T \mathbf{n}_j, \qquad (21)$$

because $\sqrt{|\mathbf{I}'_i|^2 |\mathbf{I}'_j|^2} = \frac{2\pi}{3} l^2 \rho_i \rho_j$.

Appendix B

We derive the relationship between the similarity of observation vectors and the similarity of surface normals for the simplified Torrance-Sparrow model. Let us consider a polar coordinate system (Figure 7) whose north pole coincides with the viewing direction $\mathbf{v}$, and denote surface normals $\mathbf{n}_i$,
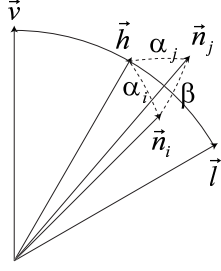
Figure 7. Polar coordinate system.

$\mathbf{n}_j$, and a light source vector $\mathbf{l}$ by (θ_i, ϕ_i), (θ_j, ϕ_j), and (θ, ϕ). Then, half vector $\mathbf{h}$, which bisects $\mathbf{v}$ and $\mathbf{l}$, is given by $(\theta/2, \phi)$. We denote the angles between $\mathbf{h}$ and $\mathbf{n}_i$, $\mathbf{h}$ and $\mathbf{n}_j$, and $\mathbf{n}_i$ and $\mathbf{n}_j$ by α_i, α_j, and β.

For the sake of simplicity, we do not take the Fresnel and geometric attenuation terms into consideration. We also obtain radiance I_i at a point i illuminated by a single directional light source $\mathbf{l}$ with brightness l as

$$I_i = \frac{\rho_i l}{\mathbf{n}_i^T \mathbf{v}} \exp\left(-\frac{\alpha_i^2}{\sigma^2}\right). \tag{22}$$

Here, ρ_i is the reflectance of surface reflection, and σ represents surface roughness. The angle α_i satisfies

$$\cos\alpha_i = \mathbf{h}^T \mathbf{n}_i = \sin\frac{\theta}{2}\sin\theta_i\cos(\phi - \phi_i) + \cos\frac{\theta}{2}\cos\theta_i. \tag{23}$$

Assuming that the directional light sources with brightness l are distributed around the object with uniform density, the inner product of observation vectors $\mathbf{I}'_i$ and $\mathbf{I}'_j$ is represented by an integral with respect to $\mathbf{l}$ as

$$\mathbf{I}_i'^T \mathbf{I}_j' = \frac{\rho_i l}{\mathbf{n}_i^T \mathbf{v}} \frac{\rho_j l}{\mathbf{n}_j^T \mathbf{v}} \int\int_{\{\mathbf{l}|\mathbf{n}_i^T \mathbf{l}>0, \mathbf{n}_j^T \mathbf{l}>0, |\mathbf{l}|=1\}} \exp\left(-\frac{\alpha_i^2 + \alpha_j^2}{\sigma^2}\right) d\mathbf{l} \tag{24}$$

in a similar manner to the Lambertian model.

As is often the case with real materials, we assume that the specular reflection has a sharp peak and then $\sigma \ll \pi$. Thus, only combinations of $\mathbf{n}_i$, $\mathbf{n}_j$, and $\mathbf{l}$ that satisfies $\alpha_i^2 + \alpha_j^2 \ll \pi^2$ can contribute to the integral in (24). Therefore, we assume $\mathbf{n}_i \simeq \mathbf{n}_j$, and consider $\mathbf{l}$ such that $\mathbf{h} \simeq \mathbf{n}_i \simeq \mathbf{n}_j$.

Accordingly, taking Taylor series expansion of (23) and ignoring higher order terms with respect to α_i, $(\theta/2 - \theta_i)$, and $(\phi - \phi_i)$, we obtain

$$\alpha_i^2 \simeq \left(\frac{\theta}{2} - \theta_i\right)^2 + \sin\frac{\theta}{2}\sin\theta_i(\phi - \phi_i)^2. \tag{25}$$

Similarly, we can derive

$$\beta^2 \simeq (\theta_i - \theta_j)^2 + \sin\theta_i\sin\theta_j(\phi_i - \phi_j)^2. \tag{26}$$

Substituting these approximations into (24), we can derive

$$\mathbf{I}_i'^T \mathbf{I}_j' \simeq \frac{\rho_i l}{\mathbf{n}_i^T \mathbf{v}} \frac{\rho_j l}{\mathbf{n}_j^T \mathbf{v}} \exp\left(-\frac{\beta^2}{2\sigma^2}\right) \int\int_{\{\mathbf{l}|(\mathbf{n}_i+\mathbf{n}_j)^T \mathbf{l}>0, |\mathbf{l}|=1\}}$$
$$\exp\left[-\frac{1}{\sigma^2}\left\{\frac{1}{2}(\theta - \theta_i - \theta_j)^2 \right.\right.$$
$$\left.\left. + \sin\frac{\theta}{2}(\sin\theta_i + \sin\theta_j)\left(\phi - \frac{\phi_i + \phi_j}{2}\right)^2\right\}\right] d\mathbf{l}. \tag{27}$$

The norm of the observation vector is calculated as

$$|\mathbf{I}'_i|^2 \simeq \frac{(\rho_i l)^2}{(\mathbf{n}_i^T \mathbf{v})^2} \int\int_{\{\mathbf{l}|\mathbf{n}_i^T \mathbf{l}>0, |\mathbf{l}|=1\}}$$
$$\exp\left[-\frac{1}{\sigma^2}\left\{\frac{1}{2}(\theta - 2\theta_i)^2 + 2\sin\frac{\theta}{2}\sin\theta_i(\phi - \phi_i)^2\right\}\right] d\mathbf{l}. \tag{28}$$

Therefore, the inner products of the normalized observation vectors are related to the angles between surface normals as

$$\mathbf{I}_i^T \mathbf{I}_j = \frac{\mathbf{I}_i'^T \mathbf{I}_j'}{\sqrt{|\mathbf{I}'_i|^2 |\mathbf{I}'_j|^2}} \simeq \exp\left(-\frac{\beta^2}{2\sigma^2}\right). \tag{29}$$

References

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

[2] M. K. Chandraker, F. Kahl, and D. J. Kriegman. Reflections on the generalized bas-relief ambiguity. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages I–788–795, 2005. 1

[3] E. N. Coleman and R. Jain. Obtaining 3-dimensional hape of textured and specular surfaces using four-source photometry. *Computer Vision, Graphics, and Image Processing*, 18(4):309–328, 1982. 1

[4] A. S. Georghiades. Incorporating the torrance and sparrow model of reflectance in uncalibrated photometric stereo. In *Proc. Int'l. Conf. Computer Vision*, pages 816–825, 2003. 1

[5] H. Hayakawa. Photometric stereo under a light-source with arbitrary motion. *J. Optical Soc. Am. A*, 11(11):3079–3089, 1994. 1

[6] A. Hertzmann and S. M. Seitz. Example-based photometric stereo: Shape reconstruction with general, varying brdfs. *IEEE Trans. Pattern Analalysis and Machine Intelligence*, 27(8):1254–1264, 2005. 1, 3

[7] B. Horn. Robot vision. *MIT Press, Cambridge, MA*, 1986. 1

[8] K. Ikeuchi. Determining surface orientations of specular surfaces by using the photometric stereo method. *IEEE Trans. Pattern Analysis and Machince Intelligence*, 3(6):661–669, 1981. 1

[9] S. Koppal and S. G. Narasimhan. Clustering appearance for scene analysis. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages II–1323–1330, 2006. 2

[10] P. Kovesi. Shapelets correlated with surface normals produce surfaces. In *Proc. Int'l. Conf. Computer Vision*, pages 994–1001, 2005. 5

[11] K. V. Mardia, J. T. Kent, and J. M. Bibby. *Multivariate Analysis*. Academic Press, London, 1979. 5

[12] S. K. Nayar, K. Ikeuchi, and T. Kanade. Determining shape and reflectance of hybrid surfaces by photometric sampling. *IEEE Trans. Robotics and Automation*, 6(4):418–443, 1990. 1

[13] M. Oren and S. Nayar. Generalization of the lambertian model and implications for machine vision. *Int'l J. Computer Vision*, 14(3):227–251, 1995. 3

[14] R. Pless and I. Simon. Embedding images in non-flat spaces. In *Proc. Conf. Imaging Science, Systems, and Technology*, pages 182–188, 2002. 3

[15] J. B. Tenenbaum, V. Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290:2319–2323, 2000. 3, 5

[16] K. E. Torrance and E. M. Sparrow. Theory for off-specular reflection from roughened surface. *J. Optical Society of America*, 57:1105–1114, 1967. 3

[17] G. J. Ward. Measuring and modeling anisotropic reflection. *Proc. SIGGRAPH '92*, pages 265–272, 1992. 3

[18] H. Winnemoller, A. Mohan, J. Tumblin, and B. Gooch. Light waving: estimating light positions from photographs alone. In *Proc. Eurographics 2005*, pages 433–488, 2005. 3

[19] L. Wolff, S. Shafer, and G. E. Healey. *Physics-based vision: Principals and Practice, Shape Recovery*. Jones and Bartlett., 1992. 1

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

[21] R. Zhang, P.-S. Tsai, J. E. Cryer, and M. Shah. Shape from shading: A survey. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 21(8):690–706, 1999. 1