Attached Shadow Coding: Estimating Surface Normals from Shadows under Unknown Reflectance and Lighting Conditions

Takahiro OKABE
The University of Tokyo
takahiro@u-tokyo.ac.jp

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

Yoichi SATO
The University of Tokyo
ysato@u-tokyo.ac.jp

Abstract

We present a novel technique, termed attached shadow coding, for estimating surface normals from shadows when the reflectance and lighting conditions are unknown. Our key idea is encoding surface points via attached shadows observed under different light source directions and then estimating surface normals on the basis of the similarity of the attached shadow codes. Because shadows do not rely on reflectance properties, our method is applicable to surfaces with various complex reflectances such as anisotropic and composite materials. Moreover, our method is robust against noise because it takes advantage of the combination of weak constraints imposed by a number of light sources. We theoretically show that the distance between the codes at two surface points is equal to the angle between the corresponding surface normals under the assumption of uniform lighting and a convex object. Our method embeds high-dimensional codes into a 3D surface normal space so that the inter-code distances are preserved. Furthermore, we extend the method in order to alleviate the effects of nonuniform lighting and cast shadows. Experimental results demonstrate the effectiveness of our method.

1. Introduction

Recovering the geometric structure of a scene is one of the fundamental problems in computer vision. Unlike techniques for recovering surface depths such as stereo, photometric stereo is a technique for estimating surface orientations from a set of images of a stationary object taken from a fixed viewpoint and under different light sources. The tasks of recovering surface depths and recovering surface orientations complement each other since the derivative of the estimated surface depths often yield inaccurate surface orientations and vice versa. Directly acquiring surface normals is of critical importance for applications such as image synthesis and reflectance recovery because the appearance of an object is frequently sensitive to surface normals.

Since Woodham [27] first addressed photometric stereo under the assumption of the Lambert model and calibrated light sources, its generalization has extensively been studied: one direction of study is generalization to the case of *unknown reflectance properties*, and the other is generalization to the case of *unknown light sources*. The former direction of study makes use of the generic properties of Bidirectional Reflectance Distribution Functions (BRDFs) such as isotropy [16], bilateral symmetry [1], reciprocity [29], halfway-vector symmetry [3], reflective symmetry of the halfway vector [11], and specularity [6], and therefore does not assume specific reflectance models. The latter direction resolves the Generalized Bas-Relief (GBR) ambiguity [4] of uncalibrated photometric stereo on the basis of specularities [7, 8], interreflections [5], nearby light sources [18], isotropic and reciprocal structure [24], and priors on the albedo distribution [2].

Accordingly, the next challenge would be estimating surface orientations of an object with *unknown reflectance* from images taken under *unknown lighting* conditions. In the context of surface normal estimation under unknown reflectance conditions, we shall reconsider the concept of Shape From Shadows (SFS) [22, 9, 19]. Because shadows are independent of BRDFs, SFS could be applicable to surfaces with a variety of reflectance properties. Generally, shadows are classified into *attached shadows* and *cast shadows*. Attached shadows arise when the angle between a surface normal and a light source direction is obtuse, and they yield local constraints on surface normals. On the other hand, cast shadows arise when a light source is obstructed by a part of the same or another object, and they yield global constraints on surface depths. Although previous studies [21, 28] have shown that the use of shadows is an effective way of coping with unknown reflectance properties, these methods do require the light source directions or locations.

Kriegman and Belhumeur [14] addressed the problem of determining shapes from attached shadows in the case of unknown light source directions. This study introduced the interesting fact that the GBR transformation preserves both

2009 IEEE 12th International Conference on Computer Vision (ICCV)
978-1-4244-4419-9/09/$25.00 ©2009 IEEE

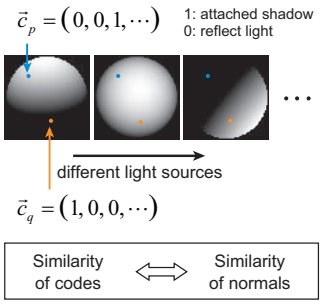

Figure 1. Encode surface points on the basis of attached shadows observed under different light sources.

attached and cast shadows, and theoretically showed that it is possible to reconstruct surface normals and light source directions up to the GBR ambiguity from shadows alone. Specifically, they proposed an algorithm for surface normal recovery that combines constraints imposed by an occluding contour of an object with those obtained at surface points where three attached shadow boundaries intersect all together. However, the applicability of their algorithm in its present form is restricted to some extent as described in their work because the detection of attached shadow boundaries is particularly prone to error.

In this paper, we propose a novel technique termed *attached shadow coding* for estimating surface normals from shadows under unknown reflectance and unknown lighting conditions. Our key idea is encoding surface points via attached shadows observed under different light source directions and then estimating surface normals on the basis of the similarity of attached shadow codes (see Figure 1). We assume that BRDFs are positive, *i.e.* a surface point reflects some light if the point is visible from a light source. Then, we prove that the distance (dissimilarity) between the codes at two surface points is equal to the angle between the corresponding surface normals under the assumption of uniformly distributed light sources and a convex object. This enables us to embed the high-dimensional codes into a 3D surface normal space by using MultiDimensional Scaling (MDS) so that the distances between the codes are preserved. Furthermore, we discuss how surface normal estimation is affected by the deviation from the above assumptions and extend the proposed method in order to alleviate the effects of nonuniform lighting and cast shadows.

Our main contributions are summarized as follows. (i) Our method based on attached shadows is applicable to various reflectance properties such as anisotropic BRDFs and composite materials because shadows do not rely on reflectance properties if BRDFs are positive. (ii) Our method is robust against noise because it takes advantage of the combination of weak constraints imposed by a number of light sources; whether a surface point is visible from a light source direction or not. (iii) Our method does not require calibrated light sources, and as a result, makes the capture process easy.

1.1. Related work

The problem of estimating surface normals when both the reflectance and lighting conditions are unknown has been studied by some researchers.

Hertzmann and Seitz [10] proposed an example-based method for estimating surface normals by using a reference object with a known shape and made of the same material as the target object. Their method simultaneously captures images of the target object and the reference object under the same light sources, and therefore the surface normal estimation results in appearance matching. They extended their method by using a small number of reference objects so that the variation in surface reflectance is represented by a linear combination of the reference objects' BRDFs. However, the question of how to choose a suitable set of reference objects to best fit a given target material still remains to be addressed. They also demonstrated that the shape of an anisotropic material can be recovered if the target has a two-dimensional shape such as a cylindrical surface.

Our method is similar in spirit to the following two studies which estimate surface normals under the condition of unknown reflectance and unknown lighting without reference objects. Koppal and Narasimhan [12] proposed a method for discovering iso-normal clusters on the basis of an *appearance profile* which is a vector of intensities measured at a surface point under varying illumination. They demonstrated that the temporal variations of appearance profiles observed under a smoothly moving light source are useful for acquiring the geometric structure of a scene. Sato *et al*. [20] proposed a method for estimating surface normals on the basis of the similarity of appearance profiles. They demonstrated that the geodesic distance between appearance profiles themselves is useful for estimating surface normals by dimensionality reduction via isomap [25]. Isomap is also utilized for estimating light source directions from images alone [26].

However, these previous methods share a common limitation. That is, they assume isotropic BRDFs so that appearance profiles depend only on surface normals and are independent from the local tangent frame, *i.e.* the rotation of the local coordinates with respect to a surface normal. More specifically, Koppal's method assumes that the reflectance is described by a linear combination of isotropic BRDFs, and Sato's method assumes that the reflectance is proportional to a single isotropic BRDF. In addition, Koppal's method of iso-normal clustering requires another existing technique for acquiring the normals of curved surfaces, whereas Sato's method implicitly assumes that surface normals are densely distributed over a hemisphere so that the geodesic distances can be correctly calculated.

The advantages of our method over these related works are that our method is applicable even to anisotropic BRDFs and composite materials because attached shadows are independent of BRDFs, and our method can deal with surfaces with sparsely or densely distributed normals because our distance measure is global and does not require any calculation of geodesic distances.

2. Encoding via Attached Shadows

Our method estimates surface normals on the basis of the similarity of attached shadow codes. In this section, we define the attached shadow code and its distance measure.

2.1. Attached shadow codes

Let us denote the observed brightness at the p-th surface point $(p = 1, 2, 3, ..., P)$ under the n-th light source direction $(n = 1, 2, 3, ..., N)$ by I_{pn}. We assume that $I_{pn} > 0$ iff $\boldsymbol{n}_p^T \boldsymbol{l}_n > 0$ where $\boldsymbol{n}_p$ is the surface normal at the p-th point and $\boldsymbol{l}_n$ is the direction of the n-th light source, *i.e.* BRDFs are not non-negative but positive. We assume that BRDFs have non-negligible DC or low-frequency components, and so highly specular BRDFs such as mirror-like reflectance are outside the scope of our study.

Thus, we consider that the p-th point is in attached shadows with respect to the n-th light source direction if $I_{pn} = 0$[1] and then binarize its brightness as

$$c_{pn} = \begin{cases} 1 & I_{pn} = 0 \\ 0 & I_{pn} > 0. \end{cases} \tag{1}$$

As shown in Figure 1, we encode surface points based on attached shadows observed under different light sources and call the vector $\boldsymbol{c}_p = (c_{p1}, c_{p2}, c_{p3}, \cdots, c_{pN})$ an *attached shadow code*.

2.2. Distance measure

We define the similarity s_{pq} between the attached shadow codes $\boldsymbol{c}_p$ and $\boldsymbol{c}_q$ by

$$s_{pq} = \frac{1}{N} \sum_{n=1}^{N} c_{pn} c_{qn}, \tag{2}$$

i.e. the ratio between the number of light sources under which both the points p and q are in attached shadows and the total number of light sources.

Let us first assume that distant light sources are uniformly distributed around a scene and that a target object is convex (we relax these assumptions in Section 4). In this case, as shown in Appendix, the similarity s_{pq} between two attached shadow codes at the points p and q is related to the angle θ_{pq} between the corresponding surface normals $\boldsymbol{n}_p$ and $\boldsymbol{n}_q$ as

$$s_{pq} = \frac{\pi - \theta_{pq}}{2\pi}. \tag{3}$$

Hence, the Euclidean distance between the surface normals $d_{pq} = |\boldsymbol{n}_p - \boldsymbol{n}_q| = 2\sin(\theta_{pq}/2)$ is represented by

$$d_{pq} = 2\sin\left(\frac{\pi - 2\pi s_{pq}}{2}\right). \tag{4}$$

Note that our distance measure is global, *i.e.* eq. (4) holds for arbitrary pairs of surface points. This is why our method is applicable to surfaces with sparsely distributed surface normals unlike the existing method [20] which uses local distance measure and as a result requires geodesic distances.

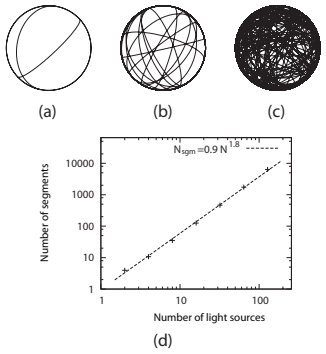

Figure 2. Number of distinct codes versus number of light sources.

2.3. Number of distinct codes

Before getting into the details of our method, we shall investigate the relationship between the number of distinct codes and the number of light sources (or equivalently the number of images). Because surface normals visible from a fixed viewpoint can be represented by points on a hemisphere and a distant point light source binarizes surface normals according to the angle between the surface normal and the light source direction by a great circle, the number of distinct codes is equal to the number of regions on the hemisphere segmented by the great circles defined by the light sources. In Figure 2, we show hemispheres divided up by (a) 2, (b) 16, and (c) 128 great circles, and (d) the number of segments N_{sgm} versus the number of light sources N. Our numerical analysis shows that the number of segments approximately grows as $N_{\mathrm{sgm}} = 0.9 \times N^{1.8}$. This is comparable to the fact that the number of intersecting points of the great circles grows as $O(N^2)$.

Based on this observation, we can roughly calculate the relationship between the accuracy of estimated surface normals $\delta\theta$ and the number of images N. Assuming that N_{sgm} equal-sized disks with $\pi(\delta\theta)^2$ area cover the entire solid

[1] We assume that a scene is illuminated only by a distant point light source or that the image taken under ambient illumination is subtracted in advance. Actually, taking dark-current noise into consideration, we use a certain threshold to detect shadows.

angle of a hemisphere (2π), we obtain an optimistic estimate of the accuracy $\delta\theta \simeq \sqrt{2/N_{\mathrm{sgm}}}$. This means that our method requires a large number of images, for example $N = 128$ ($\delta\theta \simeq 1^\circ$). However, we can simply capture images by using a video camera and by waving a distant point light source around an object because our method does not require light source directions.

3. Estimating Normals by Embedding Codes

Based on the relationship between the similarity of attached shadow codes and that of surface normals, our method embeds the high-dimensional codes into a 3D surface normal space by using MDS. The ambiguity caused by the embedding is resolved by using known surface normals along an occluding contour.

3.1. Embedding into a 3D space

As shown in the previous section, the Euclidean distance d_{pq} between two surface normals $\boldsymbol{n}_p$ and $\boldsymbol{n}_q$ is calculated from their attached shadow codes $\boldsymbol{c}_p$ and $\boldsymbol{c}_q$. This means that we can capture the relative configuration of surface normals solely from the codes.

We embed high-dimensional codes into a 3D surface normal space so that the inter-code distances are preserved and acquire the coordinates $\boldsymbol{m}_p$ of each code in the 3D space. Here, we use MDS, which is a well-known technique for distance-preserving dimensionality reduction and minimizes

$$\sum_{pq} \left[(\boldsymbol{m}_p - \boldsymbol{m}_q)^2 - d_{pq}^2 \right]^2, \tag{5}$$

i.e. the difference between the squared original distance d_{pq}^2 of the high-dimensional codes and the squared distance $(\boldsymbol{m}_p - \boldsymbol{m}_q)^2$ in the 3D space. The optimization results in an eigenvalue problem and its implementation is straightforward [15, 17][2].

3.2. Projection onto a hemisphere

The embedding via MDS causes an ambiguity because inter-code distances are invariant when the embedded points are transformed by translations, rotations, and mirror reflections. Moreover, the embedded 3D vector $\boldsymbol{m}_p$ does not necessarily have unit length. Therefore, in a similar manner to the previous studies [26, 20], we project the vector $\boldsymbol{m}_p$ onto a (hemi)sphere by gnomonic projection.

Assuming that a surface is smooth and its occluding contour is visible, we use surface normals $\boldsymbol{n}_p$ obtained along the occluding contour (or its convex portion) OC as references. More specifically, we determine the center of a unit sphere and the origins of the elevation and azimuth angles by nonlinear minimization of the cost function $\sum_{p \in OC} \left[\boldsymbol{n}_p - \mathrm{GP}(\boldsymbol{m}_p) \right]^2$, where $\mathrm{GP}(\boldsymbol{m}_p)$ is a gnomonic projection of $\boldsymbol{m}_p$. The constraints imposed by the fact that the embedded vectors $\boldsymbol{m}_p$ ($p \in OC$) of points on the occluding contour exist on a plane in the 3D surface normal space make the optimization tractable.

[2]For images with 68×68 pixels, it takes about 7 seconds by using MATLAB on a typical Core2 PC.

4. Beyond Uniform Lighting and Convex Surfaces Assumptions

In this section, we discuss how surface normal estimation is affected by the deviation from the assumptions of uniform lighting and a convex object, and then we relax these assumptions.

4.1. Nonuniform lighting

Our method can *uniquely* determine surface normals if distant light sources uniformly distribute around a convex object because the distance measure in eq. (4) is exactly correct. From a practical point of view, it is worth discussing how surface normal estimation is affected by nonuniformly distributed light sources.

First, we consider the case in which uniform lighting is distorted by the GBR transformation and reveal the relationship with the previous study by Kriegman and Belhumeur [14]. The GBR transformation, which is represented by a 3×3 matrix G, preserves attached shadows because the inner product of a light source direction and a surface normal is invariant; $\boldsymbol{l}_n^T \boldsymbol{n}_p = \boldsymbol{l}_n^T G^{-1} G \boldsymbol{n}_p = (G^{-T} \boldsymbol{l}_n)^T (G \boldsymbol{n}_p)$.

Assuming that a set of light sources $\boldsymbol{l}_n$ are uniformly distributed around an object, let us consider a set of light sources $\boldsymbol{l}'_n$ distorted by the GBR transformation:

$$\boldsymbol{l}'_n = G^{-T} \boldsymbol{l}_n = \frac{1}{\lambda} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \mu & \nu & \lambda \end{pmatrix} \boldsymbol{l}_n. \tag{6}$$

Because the inner product of the transformed light source and a surface normal satisfies

$$\boldsymbol{l'}_n^T \boldsymbol{n}_p = (G^{-T} \boldsymbol{l}_n)^T \boldsymbol{n}_p = \boldsymbol{l}_n^T (G^{-1} \boldsymbol{n}_p) = \boldsymbol{l}_n^T \boldsymbol{n'}_p, \tag{7}$$

the attached shadow codes of an object with surface normals $\boldsymbol{n}_p$ observed under transformed light sources $\boldsymbol{l}'_n$ are identical to those of an object with transformed surface normals $\boldsymbol{n}'_p$ observed under uniformly distributed light sources $\boldsymbol{l}_n$. Unfortunately, in this case, our method interprets as if the distribution of light sources were uniform and recovers the distorted surface normals $\boldsymbol{n}'_p$ because it assumes uniform lighting. Consequently, our method cannot determine surface normals uniquely when the distribution of light sources is distorted by the GBR transformation by chance[3]. Note

[3]Although eq. (7) holds for an arbitrary invertible matrix, the integrability constraint reduces the ambiguity of surface normals up to the GBR transformation.

that known surface normals along an occluding contour do not resolve the ambiguity because they are preserved under the GBR transformation [14].

Second, we consider general nonuniform distributions that include global distortions other than the GBR transformation and local inhomogeneities. In this case, the surface normals estimated by using our method as described in Section 3 are not accurate because the relationship between the similarity and the angle deviates from eq. (3).

To alleviate the effect of nonuniform lighting, we assign a weight w_n^{light} to each light source and extend the definition of similarity in eq. (2) to

$$s_{pq} = \frac{1}{N} \sum_{n=1}^{N} w_n^{\text{light}} c_{pn} c_{qn}. \tag{8}$$

The weights should be determined according to the density of light sources so that light sources in densely distributed directions have smaller weights and those in sparsely distributed directions have larger weights. More specifically, we determine the weights solely from attached shadow codes by minimizing

$$\sum_{p=1}^{P} \left(\sum_{n=1}^{N} w_n^{\text{light}} c_{pn} - \frac{N}{2} \right)^2 \tag{9}$$

subject to $\sum_{n=1}^{N} w_n^{\text{light}} = N$ and $w_n^{\text{light}} \geq 0$. This cost function requires that each surface point is illuminated by just half of all light sources, as is the case for uniform lighting.

In the implementation of our method, we find the optimal weights w_n^{light} from a certain interval $[w_l, w_u]$ in order to prevent overfitting. We can evaluate whether the predetermined interval is appropriate or not in the post-processing as follows. We estimate light source directions based on the consistency between the estimated surface normals and observed attached shadow codes. Then, we calculate the 2nd moment of the light source distribution $M = \sum_{n=1}^{N} w_n^{\text{light}} \boldsymbol{l}_n \boldsymbol{l}_n^T / N$, which is equal to $M^{\text{uniform}} = \text{diag}(1/3, 1/3, 1/3)$ for uniform lighting. We evaluate the deviation $\epsilon = |M - M^{\text{uniform}}|$ in terms of the Frobenius norm.

4.2. Nonconvex surfaces

Nonconvex surfaces have not only attached shadows but also cast shadows. Since we cannot distinguish these shadows solely by the intensities of surface points, cast shadows cause two problems. (i) The attached shadow codes of points in concave areas are contaminated by cast shadows. (ii) The effects of the contaminated codes propagate to other surface points in convex areas via inter-code distances in the embedding by MDS, and they degrade the overall estimation accuracy of surface normals. Here, we will mainly deal with the latter problem. It should be noted that the accuracy of estimated surface normals in concave areas can be improved (see Section 5.1).

To cope with the latter problem, we extend the cost function in eq. (5) as

$$\sum_{pq} w_p^{\text{pixel}} w_q^{\text{pixel}} \left[(\boldsymbol{m}_p - \boldsymbol{m}_q)^2 - d_{pq}^2 \right]^2, \tag{10}$$

where w_p^{pixel} is a weight for each pixel. We consider surface points in concave areas to be outliers and assign them smaller weights, whereas we consider surface points in convex areas to be inliers and assign them larger weights.

We simultaneously estimate both the embedding and the weights by using an EM-like algorithm in a similar manner to manifold clustering [23]. In the *M-step*, we embed attached shadow codes into a 3D surface normal space by minimizing the cost function in eq. (10) for given weights. In the *E-step*, we update the weights by using the result of the embedding. More specifically, we evaluate the error of embedding e_p at a surface point p by

$$e_p = \frac{\sum_q w_q^{\text{pixel}} ||\boldsymbol{m}_p - \boldsymbol{m}_q| - d_{pq}|}{\sum_q w_q^{\text{pixel}}}, \tag{11}$$

and then update the weight as

$$w_p^{\text{pixel}} = \frac{1}{1 + \exp\left[(e_p - a)/b \right]}. \tag{12}$$

For points in concave areas, the embedding causes larger errors e_p, and as a result, the weights w_n^{light} get smaller. We use a sigmoid function because our purpose is not multi-class clustering but outlier removal. For initializing the weights, we consider the surface points on an occluding contour (or its convex portion) to be inliers.

5. Experiments

5.1. Synthetic images

We evaluated the performance of our method against various factors such as number of images, noise, nonuniform lighting, and cast shadows.

First, we investigated the relationship between the accuracy of our method and the number of images. As shown in Figure 3, we estimated the surface normals of (a) a sphere from (b) its images rendered under light sources uniformly distributed around it. We show (c) the ground truth of surface normals and the estimated surface normals from (d) 128, (e) 256, and (f) 512 images. Here, the absolute values of the x, y, and z components of the surface normals are linearly coded in R, G, and B channels. One can see that (g) the error decreases as the number of images increases, and our method achieves an accuracy of about 3° for 512

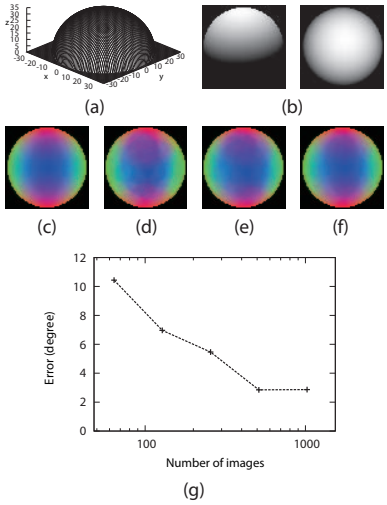

Figure 3. Relationship between the accuracy of our method and the number of images.

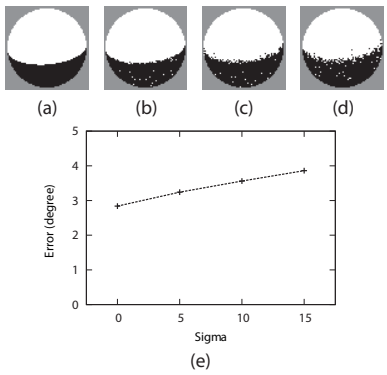

Figure 4. Error of the estimated surface normals versus noise level.

and 1024 images. The accuracy is worse than the rough estimate in Section 2.1 mainly because the reference surface normals along an occluding contour calculated by using the Sobel operator are not accurate enough. In the following experiments using synthetic data, we used 512 images.

Second, we demonstrated the robustness of our method against noise. We added zero-mean Gaussian noise with standard deviations of σ =5, 10, and 15 to 8 bit images and detected shadows by using a threshold $t = 2\sigma$. Figure 4 shows that the detection of shadows becomes inaccurate as the noise level increases from (a) 0 to (d) 15, but (e) the error of the estimated surface normals grows slowly. This result shows the robustness of our method that combines weak constraints imposed by a large number of light sources.

Third, we investigated how light sources distorted by the GBR transformation affect surface normal estimation. As explained in Section 4.1, Figure 5 shows that (a) the estimated surface normals are different from (b) the ground truth, but quite similar to (c) the ground truth of surface normals distorted by eq. (7).

Figure 5. Effect of light sources distorted by the GBR transformation.

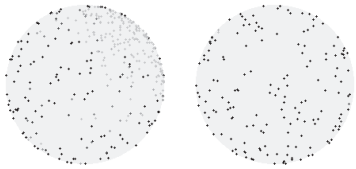

Figure 6. Estimated weights of nonuniformly distributed light sources.

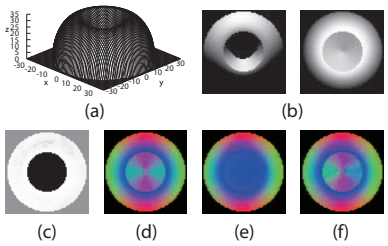

Figure 7. Estimated weights and surface normals of a nonconvex object.

Fourth, we demonstrated how well the weights of light sources work for alleviating the effect of nonuniform lighting. Figure 6 shows the orthogonal projection of the light source distribution under which synthetic images were rendered; the left disk represents the upper hemisphere and the right disk represents the lower hemisphere. In the upper right quadrant of the upper hemisphere, the light sources are five times as dense as in the other directions. The weights are determined by eq. (9) from the interval $w_n^{\text{light}} = [0.3, 1.5]$. The darker the cross is, the larger the weight is. One can see that the weights in densely distributed directions get smaller. By incorporating the weights of light sources, the error of the estimated surface normals decreases from 10.6° ($\epsilon = 0.075$) to 5.1° ($\epsilon = 0.016$). On the other hand, we obtained a worse result of 15.1° for an inappropriate interval $w_n^{\text{light}} = [0, 10]$ ($\epsilon = 0.043$), due to overfitting.

Finally, we investigated the effect of our outlier removal framework proposed in Section 4.2. The shape of the target object with a concavity and its images are shown in Figure 7 (a) and (b). One can see that (c) the estimated weight map successfully distinguishes inliers (white) from outliers (black), and that (e) the estimated surface normals of inliers are similar to (d) the ground truth. Here, we set the parameters in eq. (12) to be $a = \text{mean}(e_p)$ and $b = a/5$.

Furthermore, we revised surface normals of the outliers

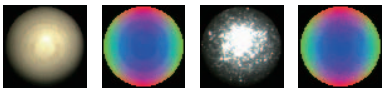

Figure 8. Estimated surface normals of wood and shiny plastic.

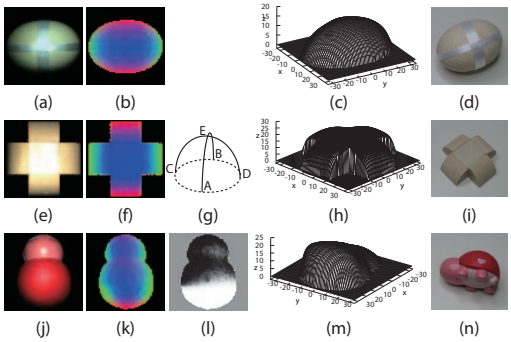

Figure 9. Estimated surface normals of an object with composite materials, an object with sparse surface normals, and a nonconvex object.

and obtained the result shown in (f). As described in Section 4.1, we calculated the light source directions from the estimated surface normals of the inliers. Then, we estimated surface normals in a concave area based on the attached shadow boundaries by using RANSAC, although other existing techniques can be used. The overall error of surface normals decreases from (e) 12.8° to (f) 5.8°.

5.2. Real images

First, as shown in Figure 8, we evaluated the accuracy of our method by using targets with known shapes. We tested spheres made of two materials: (a) wood, which has both diffuse and specular reflection components, and (c) shiny plastic, which has strong specularities and slight bumps. We used our existing equipment designed for capturing images under different light source directions; we sampled 9 directions in elevation angle and 18 directions in azimuth angle, and thus acquired 162 images in total. The threshold t for detecting shadows was determined based on the first quartile of measured intensities ($\simeq$ the median of shadowed pixel intensities). We confirmed that our method is not sensitive to the choice of the threshold.

The errors of the estimated surface normals are (b) 3.5° and (d) 3.5°. Since we used roughly equally-spaced light sources, the accuracy is comparable to or better than the results in Figure 3, but some artifacts are visible. Note that our method does not require saturations or the dynamic range of reflected radiance to be accounted for because it makes use of only the darkness of surface points and does not require their brightness. If not otherwise specified in the following, we used 162 captured images.

Second, as shown in Figure 9, we tested our method on

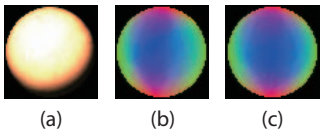

Figure 10. Surface normals estimated from images taken under a waving distant point light source.

(a) an object made of composite materials, (e) an object with sparse surface normals, and (j) a nonconvex object. One can see that the estimated surface normals in (b), (f), and (k) seem to be qualitatively consistent with the shapes. The recovered shapes by using shapelets [13] are shown in (c), (h), and (m). The images of these objects taken from different viewpoints are shown in (d), (i), and (n) for reference.

The first object (a) made of wood and satin is a composite and anisotropic material. The ability to deal with these complex reflectance properties is one of the advantages of our method over the existing techniques [12, 20]. The surface normals of (e) the second object exist on (g) two orthogonal meridians AB and CD in the surface normal space. Because our distance measure in eq. (4) is global, the angle between A and C is correctly calculated as $\pi/2$. On the other hand, the geodesic distance between A and C, which is the sum of distances AE and EC, is π and is equal to that between A and B. Therefore, the previous method [20] using geodesic distances cannot capture the configuration of surface normals. Our method can deal with not only convex objects but also nonconvex objects as shown in (j). The observation that (l) the weights on the lower carapace of a turtle are large means that the lower carapace is convex, which is consistent with the shape of the turtle.

Finally, as shown in Figure 10, we estimated the surface normals of (a) a wood sphere from 366 images taken under a waving distant point light source. By incorporating the weights of light sources, the error decreases from (b) 10.4° ($w_n^{\text{light}} = 1$, $\epsilon = 0.106$) to (c) 6.9° ($w_n^{\text{light}} = [0.6, 1, 4]$, $\epsilon = 0.034$).

6. Conclusions and Future Work

We presented a novel technique, termed attached shadow coding, for estimating surface normals under unknown reflectance and lighting conditions. Our method is applicable to surfaces with various reflectance properties, is robust against noise, and does not require calibrated light sources.

Our method regards surface points in concave areas as outliers, and hence, the estimation of their surface normals is not necessarily accurate. Because we can calculate light source directions from the estimated surface normals of the inliers, combining our method with other existing techniques assuming *known* lighting such as shape from shadows and specularities could overcome this limitation.

References

[1] N. Alldrin and D. Kriegman. Toward reconstructing surfaces with arbitrary isotropic reflectance: a stratified photometric stereo approach. In *Proc. ICCV 2007*, pages 1–8, 2007.

[2] N. Alldrin, S. Mallick, and D. Kriegman. Resolving the generalized bas-relief ambiguity by entropy minimization. In *Proc. CVPR 2007*, pages 1–7, 2007.

[3] N. Alldrin, T. Zickler, and D. Kriegman. Photometric stereo with non-parametric and spatially-varying reflectance. In *Proc. CVPR 2008*, pages 1–8, 2008.

[4] P. Belhumeur, D. Kriegman, and A. Yuille. The bas-relief ambiguity. In *Proc. CVPR '97*, pages 1060–1066, 1997.

[5] M. Chandraker, F. Kahl, and D. Kriegman. Reflections on the generalized bas-relief ambiguity. In *Proc. CVPR 2005*, pages I–788–795, 2005.

[6] T. Chen, M. Goesele, and H.-P. Seidal. Mesostructure from specularity. In *Proc. CVPR 2006*, pages II–1825–1832, 2006.

[7] O. Drbohlav and R. Sara. Specularities reduce ambiguity of uncalibrated photometric stereo. In *Proc. ECCV 2002 (LNCS2351)*, pages 46–60, 2002.

[8] A. Georghiades. Incorporating the Torrance and Sparrow model of reflectance in uncalibrated photometric stereo. In *Proc. ICCV 2003*, pages 816–823, 2003.

[9] M. Hatzitheodorou and J. Kender. An optimal algorithm for the derivation of shape from shadows. In *Proc. CVPR '88*, pages 486–491, 1988.

[10] A. Hertzmann and S. Seitz. Example-based photometric stereo: shape reconstruction with general, varying BRDFs. *IEEE Trans. PAMI*, 27(8):1254–1264, 2005.

[11] M. Holroyd, J. Lawrence, G. Humphreys, and T. Zickler. A photometric approach for estimating surface normals and tangents. *ACM Trans. Graphics (Proc. SIGGRAPH ASIA 2008)*, 27(5), 2008.

[12] S. Koppal and S. Narasimhan. Clustering appearance for scene analysis. In *Proc. CVPR 2006*, pages II–1323–1330, 2006.

[13] P. Kovesi. Shapelets correlated with surface normals produce surfaces. In *Proc. ICCV 2005*, pages II–994–1001, 2005.

[14] D. Kriegman and P. Belhumeur. What shadows reveal about object structure. *JOSA A*, 18(8):1804–1813, 2001.

[15] J. Kruskal and M. Wish. *Multidimensional scaling*. Sage Publications, 1978.

[16] J. Lu and J. Little. Reflectance and shape from images using a collinear light source. *IJCV*, 32(3):213–240, 1999.

[17] K. Mardia, J. Kent, and J. Bibby. *Multivariate analysis*. Academic Press, 1979.

[18] T. Okabe and Y. Sato. Does a nearby point light source resolve the ambiguity of shape recovery in uncalibrated photometric stereo? Technical report, IPSJ-CVIM07157021, 2007 (in Japanese).

[19] D. Raviv, Y.-H. Pao, and K. Loaparo. Reconstruction of three-dimensional surfaces from two-dimensional binary images. *IEEE Trans. RA*, 5(5):701–710, 1989.

[20] I. Sato, T. Okabe, Q. Yu, and Y. Sato. Shape reconstruction based on similarity in radiance changes under varying illumination. In *Proc. ICCV 2007*, pages 1–8, 2007.

[21] S. Savarese, M. Andreetto, H. Rushmeier, F. Bernardini, and P. Perona. 3D reconstruction by shadow carving: theory and practical evaluation. *IJCV*, 71(3):305–336, 2007.

[22] S. Shafer and T. Kanade. Using shadows in finding surface orientations. *CVGIP*, 22(1):145–176, 1983.

[23] R. Souvenir and R. Pless. Manifold clustering. In *Proc. ICCV 2005*, pages I–648–653, 2005.

[24] P. Tan, S. Mallick, L. Quan, D. Kriegman, and T. Zickler. Isotropy, reciprocity, and the generalized bas-relief ambiguity. In *Proc. CVPR 2007*, pages 1–8, 2007.

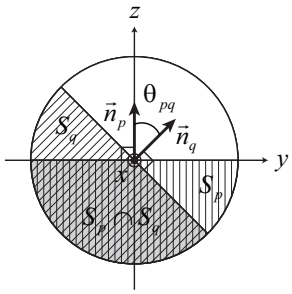

Figure 11. Surface points p and q are in attached shadows with respect to the light sources on $S_p \cap S_q$.

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

[26] H. Winnemöller, A. Mohan, J. Tumblin, and B. Gooch. Light waving: estimating light positions from photographs alone. In *Proc. Eurographics 2005*, pages 433–438, 2005.

[27] R. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):139–144, 1980.

[28] S. Yamazaki, S. Narasimhan, S. Baker, and T. Kanade. Coplanar shadowgrams for acquiring visual hulls of intricate objects. In *Proc. ICCV 2007*, pages 1–8, 2007.

[29] T. Zickler, P. Belhumeur, and D. Kriegman. Helmholtz stereopsis: exploiting reciprocity for surface reconstruction. In *Proc. ECCV 2002 (LNCS2352)*, pages 869–884, 2002.

Appendix

Let us represent the directions of distant point light sources by points on a unit sphere and consider two surface points p and q with surface normals $\boldsymbol{n}_p$ and $\boldsymbol{n}_q$ respectively. Without loss of generality, let us consider $\boldsymbol{n}_p = (0, 0, 1)^T$ and $\boldsymbol{n}_q = (0, \sin\theta_{pq}, \cos\theta_{pq})^T$ where θ_{pq} is the angle between the surface normals. As shown in Figure 11, the surface point p (q) is in attached shadows with respect to the light sources on a hemisphere S_p (S_q) because $\boldsymbol{n}_p^T\boldsymbol{l} < 0$ ($\boldsymbol{n}_q^T\boldsymbol{l} < 0$) for $\boldsymbol{l} \in S_p$ ($\boldsymbol{l} \in S_q$). Therefore, both the points p and q are in attached shadows with respect to the light sources on the intersection of the two hemisphere, i.e. $S_p \cap S_q$.

Because we assume that light sources are uniformly distributed on the sphere, the number of light sources under which both the points p and q are in attached shadows is proportional to the surface area of the shaded region: $A(S_p \cap S_q)$. On the other hand, the total number of light sources is proportional to the surface area of the whole sphere (4π). Hence, the similarity s_{pq} is given by $A(S_p \cap S_q)/(4\pi)$. We can calculate the surface area of the shaded region by integrating an infinitesimal surface element $\sin\psi d\psi d\phi$ over $0 \leq \psi \leq \pi$ and $0 \leq \phi \leq (\pi - \theta_{pq})$:

$$A(S_p \cap S_q) = \int_{\phi=0}^{\pi-\theta_{pq}} \int_{\psi=0}^{\pi} \sin\psi d\psi d\phi = 2(\pi - \theta_{pq}). \quad (13)$$

Here, ψ and ϕ are spherical coordinates such that $x = \cos\psi$, $y = -\sin\psi\cos\phi$, and $z = -\sin\psi\sin\phi$.