複雑照明下における運動物体の反射特性の推定

杜　　菲† 岡部　孝弘† 佐藤　洋一† 杉本　晃宏††

† 東京大学生産技術研究所　〒 153–8505 東京都目黒区駒場 4–6–1
†† 国立情報学研究所　〒 101–8430 東京都千代田区一ツ橋 2–1–2
E-mail: †{dufei,takahiro,ysato}@iis.u-tokyo.ac.jp, ††sugimoto@nii.ac.jp

Reflectance Estimation of a Moving Object under Complex Illumination

Fei DU†, Takahiro OKABE†, Yoichi SATO†, and Akihiro SUGIMOTO††

† Institute of Industrial Science, University of Tokyo　4–6–1, Komaba, Meguro-ku, Tokyo, 153–8505 Japan
†† National Institute of Informatics　2–1–2, Hitotsubashi, Chiyoda-ku, Tokyo, 101–8430 Japan
E-mail: †{dufei,takahiro,ysato}@iis.u-tokyo.ac.jp, ††sugimoto@nii.ac.jp

Abstract In this paper, we propose a method for recovering the reflectance properties of a moving Lambertian object from an image sequence of the object taken by a fixed camera under unknown, complex illumination. Our proposed method is based on the spherical-harmonic representation of Lambertian reflectance under arbitrary illumination. Then, by combining the geometry reconstructed by Shape-From-Motion (SFM), we recover the albedo of the object and the illumination distribution only from the image sequence of the moving object. The proposed method enables us to synthesize realistic images of the object in arbitrary poses under arbitrary lighting conditions by using reconstructed shape and albedo. A number of experiments by using both synthetic and real images demonstrate the effectiveness of our proposed method.

Key words shape-from-motion(SFM), reflectance properties, Lambertian, complex illumination, spherical-harmonic analysis

1. Introduction

To render a realistic image of an object, we require a detailed model of the object, including both geometric and photometric properties. Because the traditional manual method of constructing models is too labor-intensive to realize, the algorithms which work backward from photographs to model of the scene has increased great interest in recent years.

In this work, we focus on the estimation of photometric properties from real object's images. The problem of recovering photometric properties from real images has been discussed mainly under three different assumptions: variable illuminations [1], [14], [15], different viewing directions [8] and variable poses [7], [11], [16]. Among them, the methods which use the images of the object with variable poses to recover photometric properties is applicable to a moving object, and in this work we focus on this branch.

For a moving object, the algorithm of Shape-From-Motion (SFM) [13] can be used for recovering the 3D shape of an object from the input images of the object taken with varying poses. However, SFM in its original formulation does not take into account intensity variation caused by pose changes, and thus it cannot be used for recovering photometric properties of the object. So a texture-mapping technique is usually used to render new images. However, there are some apparent weaknesses in these synthetic images, such as inconsistency with respect to illumination and seams between mapped textures.

Different from traditional texture-mapping, Debevec *et al.* [5] presented view-dependent texture-mapping that interpolates between given photographs of the scene depending on the user's point of view. This algorithm results in more lifelike animations, but still does not recover the reflectance properties of the objects and can not synthesize images under new lighting conditions.

To overcome the limitations of texture-mapping, some methods recovering reflectance properties of a moving object have been proposed. However, such methods still have

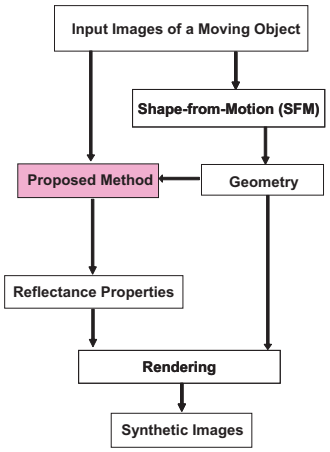

Figure 1 Proposed method.

limitation in its applications, because the methods assume some specific illumination conditions: such as several point light sources [7] or a point source and an ambient light [16].

In contrast to these methods, we present a novel method for recovering the reflectance properties of a convex object in motion under *unknown, complex* illumination conditions. The object is assumed to be Lambertian, so the reflectance properties here are albedo. The illumination distribution, which is relatively distant, can include an arbitrary combination of point source, extended source and diffuse light. The convex shape of the object ensures that there are no cast shadows and interreflection.

Our work is based on the analysis of the Lambertian reflectance in the frequency domain by Ramamoorthi – Hanrahan [10] and Basri – Jacobs [2]. They showed that a set of images of a convex Lambertian object in a fixed pose under arbitrary illumination is represented approximately by using low-order spherical harmonics. Combining this theoretical analysis with the geometry reconstructed by SFM, we obtain a set of constraints on the harmonic coefficients of illumination distribution and the albedo of the object's surface. Therefore, we can recover the albedo of an object even from an image sequence of the object in motion under complex illumination conditions. After both geometry and photometry have been successfully obtained, we are able to synthesize realistic images of the object with arbitrary poses under arbitrary illuminations. We show a flowchart of our proposed method in Fig.1.

2. Previous Works

The problem of recovering photometric properties from real images has been discussed mainly under three different assumptions: variable illuminations [1], [14], [15], different viewing directions [8], [9] and variable poses [7], [11], [16]. In the following parts, we will review some methods of each field and summarize assumptions about illumination conditions for recovering reflectance properties.

Methods [1], [6], [14], [15] recover object's shape and reflectance properties by using multiple images of a fixed object scene taken under variable illumination conditions. Among them, methods [6], [14], [15] are applicable under limited illumination conditions: such as a variable point source [6], [14] or a variable point light source and an ambient light [15]. Contrast to these methods, Basri and Jacobs [1] recently proposed a method for recovering the albedo and shape of a Lambertian object under variable, general illumination conditions. This work is also based on the low-order representation of Lambertian reflectance [2], [10] and recovers the shape and albedo by performing a optimization method.

Table 1 Comparison with previous work.

Method	View point	Pose	Illumination
Basri and Jacobs [1]	fixed	fixed	variable
Nishino et al. [8]	variable	fixed	fixed
Proposed	fixed	variable	fixed

Methods [8], [9], assuming known geometry, recover reflectance properties of the object under general illumination conditions from several images with different viewing directions. The method proposed by Nishino *et al.* [8] first used the specular components to recover illumination distribution, and then estimated the reflectance parameters of the object. Ramamoorthi–Hanrahan [9] proposed a signal-processing framework for inverse rendering which describes convolution form of reflection as a product of spherical harmonic coefficients of reflectance and the lighting. Based on this theoretical analysis, they factored the lighting and reflectance also from the images of the object taken from sparse viewpoints under unknown complex lightings.

In the context of the assumption of variable poses, several methods [7], [16] recover shape and reflectance properties from the images of the object with different poses. Maki *et al.* [7] proposed a method for recovering the shape and albedo of a moving, Lambertian object from multiple images illuminated by several point sources. Zhang *et al.* [16] proposed a method for recovering the shape and albedo of a moving, Lambertian object under a single point light source with an ambient light. Although the reflectance properties of a moving object have been precisely recovered, the applicability of these methods [7], [16] are limited because they assume simple illumination conditions.

Our work is more practical than the previous works [6], [7], [11], [14]〜[16] since our proposed method is suitable for arbitrary general illumination conditions.

Then we compare the assumptions of our work with the other two approaches [1], [8], which also recover reflectance properties under general illumination distributions, as Table 1 shows. We can see that, only the assumption of our proposed method is applicable to a moving object and overcomes the limitation of texture-mapping technique. Additionally, the method proposed by Basri and Jacobs [1] encounters the GBR ambiguity [3] and can not recover the shape and albedo

without additional constraints. The method proposed by Nishino *et al.*[8] and Ramamoorthi–Hanrahan[9] suppose that the geometry is known. While our proposed method, by combining with SFM, allows us to recover reflectance properties as well as shape only from an image sequence of a *moving* object under *complex, unknown* illumination conditions.

Finally, we mention a recent paper related to ours. Simakov *et al.*[12] proposed a method for reconstructing the shape of an object from motion under complex illumination condition. The method is also based on the theoretical analysis by using spherical harmonics[2], [10]. However, the method focuses on dense shape reconstruction and does not ensure the consistency of the illumination distribution. On the other hand, our work focuses on recovering the *photometric* properties of an object in motion.

3. Proposed Method

In this section, we begin to explain how to recover the albedo from multiple images of a moving object with known geometry. These images can be obtained by tracking a number of feature points on the object according to SFM. We used this technique to perform experiments with real images.

3.1 Spherical Harmonic Representation of Lambertian Reflectance

For Lambertian surface, observed brightness can be represented as

$$I(x,y,z) = \rho(x,y,z)\int_{\Omega} L(\theta_i,\phi_i)A(\theta_i')\sin\theta_i \mathrm{d}\theta_i \mathrm{d}\phi_i, \quad (1)$$

where $A(\theta_i') = \max(\cos\theta_i', 0)$. Here, $I(x,y,z)$ is observed brightness of one point on the object whose normal direction is (x,y,z). Function $L(\theta_i,\phi_i)$ represents illumination distribution, which depends on global incident angle (θ_i,ϕ_i). $A(\theta_i')$ represents the reflectance properties of Lambertian surface with respect to local incident angle θ_i'. $\rho(x,y,z)$ between 0 and 1 is the albedo of this point.

Then, we can expand the functions $L(\theta_i,\phi_i)$ and $A(\theta_i')$ in terms of spherical harmonics $Y_{l,m}(\theta_i,\phi_i)$ as follows:

$$L(\theta_i,\phi_i) = \sum_{l=0}^{\infty}\sum_{m=-l}^{+l} L_{l,m}Y_{l,m}(\theta_i,\phi_i), \quad (2)$$

$$A(\theta_i') = \sum_{l=0}^{\infty} A_l Y_{l,0}(\theta_i'). \quad (3)$$

The coefficients $L_{l,m}$ and A_l can be computed in the standard way by integrating $L(\theta_i,\phi_i)$ and $A(\theta_i')$ against the spherical harmonics $Y_{l,m}$

$$L_{l,m} = \int_{\theta_i=0}^{\pi}\int_{\phi_i=0}^{2\pi} L(\theta_i,\phi_i)Y_{l,m}(\theta_i,\phi_i)\sin\theta_i \mathrm{d}\theta_i \mathrm{d}\phi_i, \quad (4)$$

$$A_l = 2\pi\int_{\theta_i'=0}^{\pi} A(\theta_i')Y_{l0}(\theta_i')\sin\theta_i' \mathrm{d}\theta_i'. \quad (5)$$

By substituting equations (2) and (3) into equation(1) and according to orthonormality of the spherical harmonics, Ramamoorthi – Hanrahan[10] and Basri – Jacobs[2] represented the convolution form of reflection in terms of spherical harmonic function as follows

$$I(x,y,z) = \rho(x,y,z)\sum_{l=0}^{\infty}\sum_{m=-l}^{+l}\sqrt{\frac{4\pi}{2l+1}}A_l L_{l,m}Y_{l,m}(x,y,z). \quad (6)$$

According to equation(5), we will find that the asymptotic behavior of A_l for large even l is $A_l \sim l^{-2}$ while $A_l = 0$ for odd $l > 1$. Accordingly more than 99% of the energy is captured by $l \leqq 2$. Thus we can omit $A_l (l > 2)$ and approximate reflectance of the Lambertian surface by using only the first nine terms of equation (6) [2], [10].

$$I(x,y,z) \approx \rho(x,y,z)\sum_{l=0}^{2}\sum_{m=-l}^{+l}\sqrt{\frac{4\pi}{2l+1}}A_l L_{l,m}Y_{l,m}(x,y,z). \quad (7)$$

3.2 Recovering Illumination Distribution

The equation (7) is the basis of our proposed method. We use a total of N visible points and P variable poses of the object. We substitute the specific expression of $Y_{l,m}(x,y,z)$ in terms of normal direction (x,y,z) and suppose that $\bar{\rho}(x,y,z) = \rho(x,y,z)L_{0,0}$, $\bar{L}_{l,m} = L_{l,m}/L_{0,0}$; then equation (7) is rewritten by

$$I_n^{(p)} \approx \bar{\rho}_n\left(\frac{\sqrt{\pi}}{2} + \sqrt{\frac{\pi}{3}}\bar{L}_{1,0}z_n^{(p)} + \sqrt{\frac{\pi}{3}}\bar{L}_{1,1}y_n^{(p)} + \cdots\right)$$
$$(n = 1,2,...,N; p = 1,2,...,P).$$

Here n and p denote points and poses of the object.

Since the geometry of the object is recovered according to SFM, the values of $I(x,y,z)$ of all the visible points can be obtained from one image of the object. So in these equations, only $\bar{\rho}(x,y,z)$ and $L_{l,m}(l = 0,1,2; -l \leqq m \leqq l)$ are unknowns.

Because the unknowns $\bar{\rho}$ and $L_{l,m}$ are in multiplication form, we should first remove the unknowns $\bar{\rho}$ to obtain equations only according to $L_{l,m}$. Thus we need to obtain the ratios of brightness between the same point observed with different poses. Supposing that $I_n^{(p+1)}/I_n^{(p)} = k_n^{(p)}$, we can obtain $N(P-1)$ equations about $\bar{L}_{l,m}$:

$$-\frac{\sqrt{\pi}}{2}(k_n^{(p)} - 1) \approx \sqrt{\frac{\pi}{3}}\bar{L}_{1,0}(k_n^{(p)}z_n^{(p)} - z_n^{(p+1)}) + \sqrt{\frac{\pi}{3}}\bar{L}_{1,1}(k_n^{(p)}y_n^{(p)} - y_n^{(p+1)}) + \cdots$$
$$(n = 1,2,...,N; p = 1,2,...,P-1). \quad (8)$$

By now we have removed the unknown $\bar{\rho}_n (n = 1,...,N)$.

As all the equations are approximate equations, we obtain the optimal solutions of $\bar{L}_{l,m}$ in the sense of the least square error.

Table 2 Recovered albedo of football.

——	1	2	3	4	5	6	7	8	9	10
Recovered ρ	0.100	0.910	0.102	0.988	0.976	0.980	1.027	0.102	0.974	0.102
Ground truths	——	1.000	0.100	1.000	1.000	1.000	1.000	0.100	1.000	0.100

Table 3 Recovered $L_{l,m}$.

——	$L_{0,0}$	$L_{1,0}$	$L_{1,1}$	$L_{1,-1}$	$L_{2,0}$	$L_{2,-1}$	$L_{2,1}$	$L_{2,-2}$	$L_{2,2}$
Recovered	1.000	-0.117	-0.295	-0.372	-0.344	0.110	-0.180	-0.166	-0.191
Ground truths	1.000	-0.253	0.463	-0.660	-0.047	0.238	-0.124	-0.0003	-0.0001

3.3 Recovering Albedo of Visible Points

Then we substitute the solution of $\bar{L}_{l,m}$ into the following equation to recover the $\bar{\rho}_n$ of every visible point.

$$\bar{\rho}_n = \frac{I_n}{\frac{\sqrt{\pi}}{2} + \sqrt{\frac{\pi}{3}}\bar{L}_{1,0}z_n + \sqrt{\frac{\pi}{3}}\bar{L}_{1,1}y_n + ...} \quad (n = 1...N). \tag{9}$$

Because the albedo is a relative value, we should set a standard albedo and obtain the relative value of albedo $\tilde{\rho}_n$. Supposing that $\tilde{\rho}_1 = \rho_{ref}$, we can obtain the relative albedo of other points by

$$\tilde{\rho}_n = \rho_n/\rho_1 \times \tilde{\rho}_1 = \bar{\rho}_n/\bar{\rho}_1 \times \rho_{ref} \quad (2 \leqq n \leqq N). \tag{10}$$

We show the specific steps of our proposed method in Fig.2.

4. Experimental Results

We now present experimental results obtained with our algorithm by using synthetic and real images. We compare our recovered albedo/texture with the ground truths to verify the effectiveness of our method.

4.1 Experiments with Synthetic Images

We used two objects, a football and a textured cube, to verify the validation of our proposed method[1)]. Fig.3 shows (a) the incident illumination distributions and (b) the rendered images of football and cube.

We chose a set of 10 arbitrary points on the football and observed their brightness with 7 different poses of the football. Then according to equation (8)(9)(10), we solved the optimal solutions of $\bar{L}_{l,m}$ and albedo $\tilde{\rho}_n$ by using the observed brightness and normal directions of these points. Here, we supposed a standard albedo of the first point $\tilde{\rho}_1 = 0.100$. The first row of Table 2 illustrates the recovered relative albedo of nine arbitrary points on the football compared with point 1. The second row shows the ground truths. We see that the recovered albedo is approximately consistent with the ground truths.

For the textured cube, we computed the albedo of one representative point on one surface. Since all the points of one surface have the same normal direction, only albedo affects their brightness. Thus we recovered the texture of one surface based on ratios of brightness between the representative point and other points on the surface. Fig.4 (a) is the original texture of the cube and (b) shows the recovered texture of two surfaces of the cube.

Table 3 shows the recovered $L_{l,m}$ from images of the football and their ground truths. We see that each $L_{l,m}$ is not accurately recovered. The reason is that, in our method, we observed the moving object from a fixed viewpoint, and the illumination distribution over the hemisphere centered at the viewing direction has a great effect on the brightness of observed parts of the object, while that over the hemisphere on the other side has only slight contribution to the observed brightness. So, the recovered optimal solutions of $L_{l,m}$ in the sense of the least square error from the observed brightness can only approximate the illumination distribution over the hemisphere centered at the viewing direction, which affects

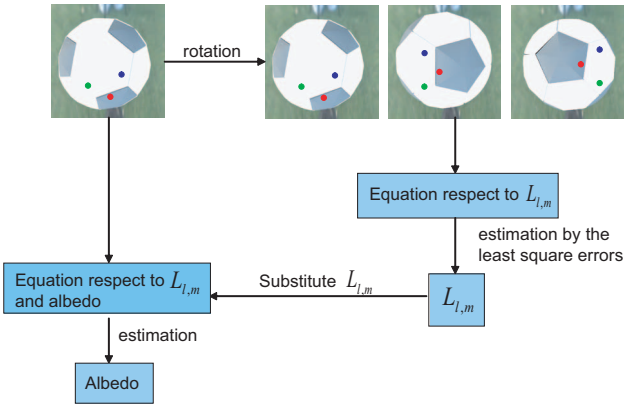

Figure 2 Specific steps of the proposed method.

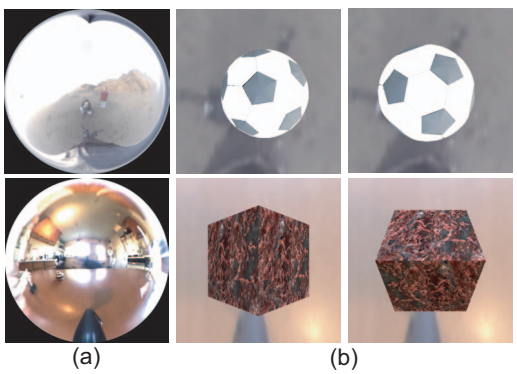

Figure 3 Synthetic images under general illumination: (a) Light probe images. (b) Synthetic images under general illumination conditions.

(1): We used software Radiance and Light Probe Image [4] to render the synthetic images under general illumination.

observed brightness greatly. Fig.5 verifies the above discussion. Fig.5 (a) is the original probe image of the beach, from which we obtained the lighting distribution $L(\theta_i, \phi_i)$; then, according to the equation (4), we computed the real values of $L_{l,m}$ as ground truths. Fig.5 (b) shows the low-order approximation of illumination distribution represented with real $L_{l,m}$. Fig.5 (c) shows the illumination distribution estimated by using our proposed method. Comparing (b) and (c) allows us to observe that the illumination of the upper hemispheres indicated with blue boundaries, which is centered at the viewing direction, are similar to each other. Then we applied these two low-order approximations of illumination to synthesizing images of a football observed from the viewing direction. Their results are shown in Fig.5 (d) and (e). Comparing (d) and (e), we see that the two illumination distributions of (b) and (c) lead very similar observed brightness, which ensure that the accuracy of recovered albedo is not affected by the errors of recovered $L_{l,m}$ according to equation (9).

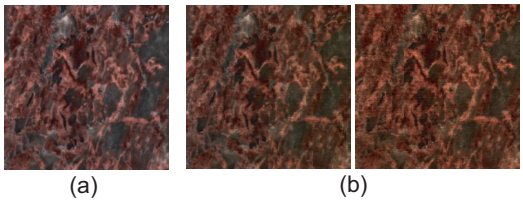

Figure 4 Recovered texture of the cube: (a) Original texture. (b) Recovered textures of two faces of the cube.

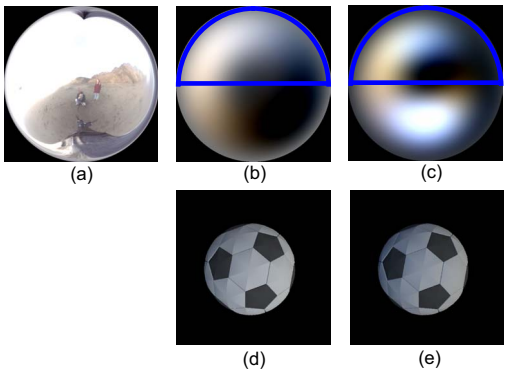

Figure 5 Recovered illumination distribution: (a) Original probe image. (b) Low-order simulation of illumination represented with real $L_{l,m}$. (c) Estimation of illumination obtained by our method. (d) Synthetic images by using illumination in (b). (e) Synthetic images by using illumination in (c).

4.2 Sensitivity Analysis against Geometrical Errors

To precisely recover albedo on the basis of estimated geometry by SFM, we should first analyze the sensitivity of our proposed method against geometric errors.

We chose ten arbitrary points on the football and perturbed their normals by adding artificial errors. For each point, we let the angle between the new normal $\vec{n}'$ and original normal $\vec{n}$ be $(\triangle\theta, \triangle\phi)$, in which $\triangle\theta$ is a random number according to the normal distribution and $\triangle\phi$ is a uniform random number from ranging 0^0 to 360^0 (Fig.6). Then we performed the experiments and observed the relationship between errors of the recovered albedo and the errors added to the geometry.

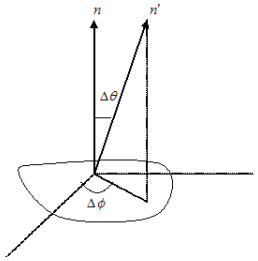

Figure 6 Artificial geometrical errors

Fig.7 shows the result of two representative points. The horizontal axis represents the standard deviation σ of the distribution of $\triangle\theta$; the vertical axis represents the absolute errors of the recovered albedo. For every σ, we performed the experiment 500 times and computed the mean and the standard deviation of the errors of recovered albedo. From Fig.7 we see that, even though some errors in the geometry exist, our method can recover albedo with a high degree of accuracy.

Because our proposed method is robust against the errors of geometry, combining our method with SFM, allows us to recover reflectance properties as well as shape from the image sequence of a moving object.

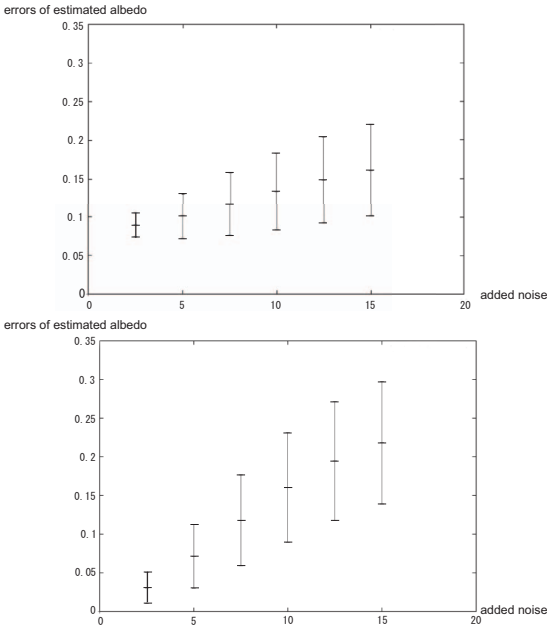

Figure 7 Sensitivity analysis against geometrical errors

4.3 Experiments with Real Images

In experiments with real images, a cube was rotated and its images were captured by using a Sony DFW-VL500 video camera. We posted a piece of matte paper on the surface of the cube. The images were taken on the condition that orthographic projection is sufficiently hold. By tracking several feature points, the geometry of the cube was obtained using SFM [13]. Note that, since geometry estimation is not the focus of our research, we tracked the feature points manually. Fig.8 (a) and (b) show two examples of the cube whose features have been tracked. Fig.8 (c) shows the recovered geometry of the cube indicated with red lines.

We computed the texture of the cube (matte paper) by applying our algorithm with the estimated geometry. Fig.8 (d) is the input image of the cube and (e) is the recovered texture. We see that our algorithm can recover the albedo/texture of a Lambertian surface with high accuracy under complex, unknown illumination.

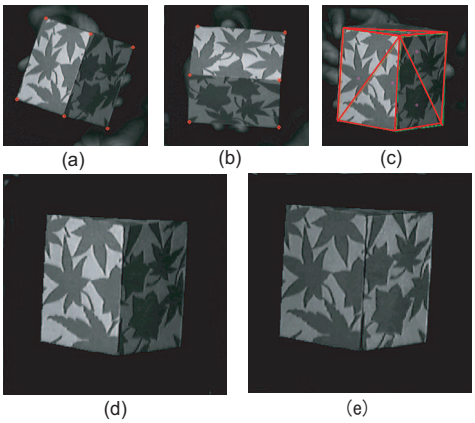

Figure 8 methods results with real images: (a) and (b) Images of tracking features. (c) Recovered geometry. (d) Original images. (e) Recovered texture.

5. Conclusions

In this paper, we presented a method for recovering the albedo of a moving, Lambertian object under a general and distant illumination condition. From multiple images of a moving object, we first estimated the geometry of the object according to SFM, and then recovered albedo based on the low-order spherical harmonics representation of Lambertian reflectance. We have performed a number of experiments with synthetic and real images and confirmed the effectiveness of our proposed method.

In the future work, we plan to extend our proposed method to more complex reflectance properties beyond the Lambertian model and try to recover the illumination distribution more robustly.

Acknowledgements

This research was in part supported by Grant-in-Aid for Scientific Research of the Ministry of Education, Culture, Sports, Science and Technology of Japan (No. 13224051 and No. 14380161).

References

[1] R. Basri and D. Jacobs, "Photometric stereo with general, unknown lighting", In *Proc. IEEE CVPR 2001*, pp.374–381, 2001.

[2] R. Basri and D. Jacobs, "Lambertian reflectance and linear subspaces", *IEEE Trans. PAMI*, 25(2): 218–233, 2003.

[3] P. Belhumeur, D. Kriegman, and A. Yuille, "The bas-relief ambiguity", In *Proc. IEEE CVPR 1997*, pp.1040–1046, 1997.

[4] P. Debevec, "Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography", In *Proc. ACM SIGGRAPH 98*, pp.189–198, 1998.

[5] P. E. Debevec, C. J. Taylor, J. Malik, "Modeling and rendering architecture from photographs: A hybrid geometry and image-based approach", In *Proc. ACM SIGGRAPH 96*, pp. 11–20, 1996.

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

[7] A. Maki, M. Watanabe, and C. Wiles, "Geotensity: Combining motion and lighting for 3D surface reconstruction", *Int'l. J. Computer Vision*, 48(2):75–90, 2002.

[8] K. Nishino, Z. Zhang, and K. Ikeuchi, "Determining reflectance parameters and illumination distribution from sparse set of images for view-dependent image synthesis", In *Proc. IEEE ICCV 2001*, pp.599–606, 2001.

[9] R. Ramamoorthi and P. Hanrahan, "A signal-processing framework for inverse rendering", In *Proc. ACM SIGGRAPH 2001*, pp.117–128, 2001.

[10] R. Ramamoorthi and P. Hanrahan, "On the relationship between radiance and irradiance: Determining the illumination from images of a convex Lambertian object", *J.Optical Soc. Am. A*, 18(10): 2448–2459, 2001.

[11] Y. Sato, M. Wheeler, and K. Ikeuchi, "Object shape and reflectance modeling from observation", In *Proc. ACM SIGGRAPH 97*, pp.379–387, 1997.

[12] D. Simakov, D. Frolova and R. Basri, "Dense shape reconstruction of a moving object under arbitrary, unknown lighting", In *Proc. IEEE ICCV 2003*, pp.1202–1209, 2003.

[13] C. Tomasi and T. Kanade, "Shape and motion from image streams under orthography: a factorization method", *Int'l. J. Computer Vision*, 9(2):137–154, 1992.

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

[15] A. Yuille, D. Snow, R. Epstein, and P.Belhumeur, "Determining generative models of objects under varying illumination: shape and albedo from multiple images using SVD and integrability", *Int'l. J. Computer Vision*, 35(3):203–222, 1999.

[16] L. Zhang, B. Curless, A. Hertzmann, and S. M. Seitz, "Shape and motion under varying illumination: Unifying multiview stereo, and structure from motion", In *Proc. IEEE ICCV 2003*, pp.618–625, 2003.