IPSJ Transactions on Computer Vision and Applications Vol. 2 71–83 (Mar. 2010)

Research Paper

A Projector-Camera System for High-Quality Synthesis of Virtual Reflectance on Real Object Surfaces

Tomoya Okazaki,[†1] Takayuki Okatani[†1]
and Koichiro Deguchi[†1]

In this paper, we present a projector-camera system for virtually altering the surface reflectance of a real object by projecting images onto it using projectors. The surface of the object is assumed to have an arbitrary shape and have a diffuse reflectance whose quantitative information is unknown. The system consists of multiple projectors and a camera. The proposed method first estimates the object surface along with the internal and external parameters of the projectors and the camera, based on the projection of structured patterns. It then improves the accuracy of surface normals by using the method of photometric stereo, where the same projectors are used as point sources of illumination. Owing to the combination of triangulation based on structured light projection and the method of photometric stereo, the surface normals of the object along with its surface shape can be accurately measured, which enables high-quality synthesis of virtual appearance. Our experimental system succeeded in giving a number of viewers a visual experience in which several plaster objects appeared as if their surfaces were made of different materials such as metals.

1. Introduction

There are several studies of the system for realizing various visual effects by projecting images onto the surfaces of real objects using image projectors [1),2)]. Their potential applications include virtual museums, industrial designs, and other entertainment uses. Suppose as an example their applications to the design of products such as automobiles and mobile phones; these products have curved surfaces with complicated surface reflectance (e.g., metallic color coating of automobiles). Currently, the designing process of these products usually requires trial productions; it is inevitable to examine how the designed product actually looks in the real world. Even if state-of-the-art CG rendering algorithms are used, the trial productions are necessary because of the limitation of the quality of CG images and/or two-dimensional image displays. Therefore, it could drastically reduce the labor and time required, if a projector-based system enables the designers to see the precise appearance of the product being designed; for example, it can immediately reflect the designer's choice of surface reflectance.

As mentioned above, there are several studies of such a system that uses projectors to alter the appearance of real objects. Raskar, et al. [3)] present a method for visually reproducing apparent motion of a target object by changing the projected pattern. Grossberg, et al. [4)] and Fujii, et al. [5)] present methods for realizing a desired appearance for a textured object by photometric compensation. Yamamoto, et al. [6)] present a method for reproducing an appearance of a real object that is the same as a reference object by image projection. This is not an exhaustive list of related studies.

This paper considers the problem of synthesizing arbitrary virtual reflectance on a real object surface by image projection. The goal is to make an object appear to a viewer as if it were made of different surface material from the real one. Based on the above existing studies, our main concern here is to *maximize the visual quality of the synthesized appearance.* Toward this goal, there are several problems to be solved, to which we will present solutions.

In this paper, we consider a system consisting of multiple projectors and a camera. An object is placed in front of them, to which images are projected to synthesize virtual appearances. Throughout the paper, we assume that the object has an ideal surface material such that it is non-colored (i.e., white/grey) and that its reflectance property is dominantly diffuse so that there will be no interreflections. As a realistic application of this system, we consider assisting a buyer of an automobile by helping to select a body color out of a number of candidates. By using a scaled model of the automobile, the system synthesizes its appearance in a more realistic manner than the conventional presentation methods (e.g., photographs in brochures) so that the viewer can experience the real appearance; see **Fig. 1**.

In order to synthesize a desired appearance in the above way, besides the virtual reflectance property to be realized (which is given by a BRDF), it is necessary

†1 Graduate School of Information Sciences, Tohoku University

Fig. 1 The proposed projector-camera system alters the appearance of a real object in such a way that it appears as if it had an arbitrary surface reflectance.

to acquire 1) the three dimensional shape of the object, 2) the relative poses of the projectors, and 3) the real reflectance property of the object surface. It is important to acquire precise information about all of the three. For this purpose, as was done in some of the previous studies, we employ the configuration of the system in which the same projectors are used for both the appearance synthesis and the measurement of the object shape; the projectors project structured light onto the object surface and then its shape is reconstructed based on stereo in combination with a camera.

In this paper, using the same system configuration, we propose a systematic method of acquiring precise information about (1)–(3) to achieve the above goal of maximizing the visual quality. Specifically, we propose a) to perform the autocalibration of the projector-camera system that can accurately reconstruct the object shape as well as the poses of the projectors and cameras, and b) to use the method of photometric stereo [7], in combination with the triangulation-based shape reconstruction by structured light projection, to obtain accurate normals of the object surface. The method makes full use of the fact that the system has multiple projectors. Multiple projectors are also necessary for eliminating or minimizing the shadowed areas on the object surface in the appearance synthesis.

Since the object shape is already reconstructed by the triangulation-based method that uses structured light projection, it might not seem to make sense to further use photometric stereo to compute the surface normals. The necessity for photometric stereo stems from the fact that it is quite difficult to synthesize high-quality appearance using the surface normals computed from the reconstructed shape.

The reason for the difficulty is two-folded. First, since the surface normal is the derivative of the surface shape, only a slight error of the reconstructed shape results in a relatively large error in the surface normal. Second, the errors in the surface normals affect the synthesized appearance to a great extent. This is significant especially when the desired virtual reflectance includes specular components. More specifically, the highlights of the synthesized appearance will have wrongly distorted shapes in the presence of small random errors of the surface normals. On the other hand, photometric stereo directly computes the surface normals from image brightness; differentiation is not necessary. We think that because of this direct relation of the surface normals obtained by photometric stereo to image brightness, it is more promising to use them to synthesize virtual appearances. The idea of combining the shape reconstructed by triangulation and normals computed by photometric stereo is very similar to the method of Nehab, et al. [8] for integrating the two different sources of data to obtain a more accurate shape. In this paper, focusing on maximizing the visual quality of synthesized appearance, we present another method that is more appropriate for our goal.

2. Method for Acquiring Necessary Information for Appearance Synthesis

2.1 Problem Formulation

The system consists of three or more projectors and a camera. A target object is placed within the intersection of the viewing area of the camera and the projection areas of the projectors. The surface of the target object is assumed to be made from a uniform material that is non-colored (i.e., white or grey). It is also required to have a reflectance property that is dominantly diffuse so that there will be only negligible specular reflections. We assume that the external parameters of the projectors and the camera are unknown, while their internal parameters

are known except for their focal lengths. As in the case of multi-camera systems, this setting is convenient, since the focal lengths will vary whenever reconfiguring the system. The other internal parameters are constant, and it is sufficient to calibrate them once.

2.2 Establishing Point Correspondences by the Phase Shifting Method

The method starts with establishing the point correspondences between each projector image and the associated camera image using the phase shifting method [9]. While a projector projects a sinusoidal brightness pattern onto the object surface, the camera captures its image, from which the phase of the initial sinusoidal pattern is calculated. In order to stably perform the phase unwrapping, the object shape is roughly estimated by projecting binary patterns onto the surface. This process is performed in turn for each projector.

2.3 Autocalibration-based Shape Reconstruction

If the internal and external parameters of the projectors and the camera are *all* unknown, it is only possible to obtain the projective reconstruction from the point correspondences. However, since the focal lengths are the only unknown internal parameters here, the projective ambiguity can be removed, as is well known for multi-camera systems [10].

To be specific, we first estimate the fundamental matrix between each projector and the camera from a decimated set of the point correspondences obtained above. Then, applying the method of Kanatani [11], the fundamental matrix is decomposed, and the focal lengths of the projectors and the camera are calculated. From this, the external parameters of the projector relative to the camera are determined.

Using these estimates as initial values, we then carry out bundle adjustment to improve accuracy of the estimation. We minimize the sum of the squared distances between the measured coordinates and their estimates in the camera images of the decimated corresponding points. The parameters to be determined in the optimization are the focal lengths and the external parameters of the projectors and the camera, and the depths of the points. The overall scaling ambiguity of the system is constrained by setting the distance from the first projector to the camera to be 1.

Finally, using the estimated poses of the projectors and the camera, the object shape is reconstructed in a dense manner from all the point correspondences. The estimated object shape and projector poses are represented in a single common Euclidean coordinate system.

2.4 Recovering the Surface Normals by Photometric Stereo

Next we apply the method of photometric stereo using the same projectors as simple illumination sources. Photometric stereo [7] estimates the normal of an object surface from its multiple images taken under different illuminations. The illuminant directions are required to be known, and in our case, we use the projector poses that have been obtained in the above geometric calibration stage for calculating them. To capture the images for photometric stereo, we project a uni-colored image from each projector to the object. We assume the illumination to be a point source of illumination. Although the projectors have projection optics, this is a good approximation when the projectors are distant from the object surface. The optical centers of the projector lenses are used for the positions of the point sources.

Three or more images are captured by projecting a uni-colored pattern in turn from each projector. Let b_{pi} be the image brightness of image point i under the illuminant direction p. Assuming the surface reflectance of the object to be Lambertian, we have $b_{pi} = \rho_i \mathbf{n}_i^\top \mathbf{l}_{pi}$, where ρ_i is the albedo, $\mathbf{n}_i$ is the surface normal, and $\mathbf{l}_{pi}$ represents the illumination orientation and strength of projector p for the surface point corresponding to image point i. The brightness $[b_{1i}, \ldots, b_{mi}]^\top$ for m different illuminant directions ($p = 1, \ldots, m$) are then represented as

$$\begin{bmatrix} b_{1i} \\ \vdots \\ b_{mi} \end{bmatrix} = \begin{bmatrix} \mathbf{l}_{1i}^\top \\ \vdots \\ \mathbf{l}_{mi}^\top \end{bmatrix} (\rho_i \mathbf{n}_i). \tag{1}$$

We perform in advance photometric calibration for the projectors by the method described later. Using the result, we can determine $\mathbf{l}_{pi}$ ($p = 1, \ldots, m$) up to overall scale (i.e., the matrix $[\mathbf{l}_{1i}, \ldots, \mathbf{l}_{mi}]^\top$ is determined up to scale). The solution of the above equation yields the normal $\mathbf{n}_i$. Although the albedo ρ_i cannot be determined due to the scaling ambiguity of the illumination matrix, it is not necessary to estimate here, since we have assumed that the reflectance

property is uniform across the object surface. To solve the above equation and determine $\mathbf{n}_i$, three or more illuminant directions (i.e., projectors) are necessary; $m \geq 3$. If $m > 3$, the solution is determined in a least squares sense.

We refer to the normals thus obtained as PS (Photometric-Stereo) normals. We refer to the normals computed from the surface shape by difference approximation as DA (Difference Approximation) normals. As will be observed later, as compared with DA normals, PS normals tend to have large systematic errors or biases, because of the difference between the real reflectance property of the object and the assumed ideal Lambertian model. Note that although we consider only diffuse reflectance here, there are still various types of them [12]. If a precise model is known for the reflectance property of a target object, its surface normals will be able to be more precisely recovered by performing photometric stereo based on the model. In that case, there are theoretically no biases in the recovered PS normals. (Although they will still have some errors due to interreflections, we assume here such errors to be small and negligible.)

We do not employ this method here. Instead of requiring a precise knowledge of the reflectance property of the object, we choose to estimate it from the measured 3D shape[⋆1]. The details are given in the next subsection. In addition to the obvious advantage that prior knowledge about the reflectance property is not necessary, the method presented below has an additional advantage that photometric calibration of the projectors does not need to be highly accurate. These contribute to increasing the usability of the proposed system. By photometric calibration, we mean the calibration of the illumination strengths of the projectors, which enables the determination of $[\mathbf{l}_{1i}, \ldots, \mathbf{l}_{mi}]^\top$ of Eq. (1). Even if this illumination matrix has some errors, they will be compensated for in the method described below.

2.5 Correction of Surface Normals Based on Recovered 3D Shape

As mentioned above, we apply the method of photometric stereo by assuming the ideal Lambertian reflectance and then estimate the reflectance property using the 3D shape that has been measured by triangulation. Our goal is to obtain precise surface normals, and thus we search for a mapping that corrects the PS normals to more accurate normals instead of estimating the reflectance property itself. Let $\mathbf{n} = [n_x, n_y, n_z]^\top$ be the PS normal of a surface point and $\mathbf{n}^*$ be its correct normal. Assuming that there exists a nonlinear function $\mathbf{n}^* = \mathbf{F}(\mathbf{n})$, we wish to estimate $\mathbf{F}(\cdot)$. This function implicitly represents the real reflectance property.

We model $\mathbf{F}$ by t-th order polynomials as

$$n_x^* = F_x(\mathbf{n}) = \sum_{\alpha=0}^{t}\sum_{\beta=0}^{t}\sum_{\gamma=0}^{t} a_{\alpha\beta\gamma} n_x^\alpha n_y^\beta n_z^\gamma, \tag{2a}$$

$$n_y^* = F_y(\mathbf{n}) = \sum_{\alpha=0}^{t}\sum_{\beta=0}^{t}\sum_{\gamma=0}^{t} b_{\alpha\beta\gamma} n_x^\alpha n_y^\beta n_z^\gamma, \tag{2b}$$

$$n_z^* = F_z(\mathbf{n}) = \sum_{\alpha=0}^{t}\sum_{\beta=0}^{t}\sum_{\gamma=0}^{t} c_{\alpha\beta\gamma} n_x^\alpha n_y^\beta n_z^\gamma, \tag{2c}$$

where n_x^*, n_y^*, and n_z^* are the components of $\mathbf{n}^* = [n_x^*, n_y^*, n_z^*]^\top$. We estimate the polynomial coefficients $a_{\alpha\beta\gamma}$, $b_{\alpha\beta\gamma}$, and $c_{\alpha\beta\gamma}$ $(\alpha, \beta, \gamma = 0, \cdots, t)$. The order t of the polynomials needs to be appropriately chosen, since a too small value will result in insufficient accuracy while a too large value will result in overfitting. In our experiments, we found that $t = 3$ yielded the best results.

These coefficients can be simply determined if the correct surface normals $\mathbf{n}^*$ are available. Since they are not available, we propose to use the DA normals that are obtained by differentiating the measured 3D shape. We first explain the method and then give a rationale for it.

Let $\hat{\mathbf{n}}^*$ be the DA normal. Then, we estimate the polynomial coefficients by minimizing the sum of the squared differences between the DA normals and the corrected normals over the entire object surface:

$$\min \sum_i h_i \|\hat{\mathbf{n}}_i^* - \mathbf{F}(\mathbf{n}_i)\|^2 \tag{3}$$

where h_i is an indicator variable of inlier/outlier (point i is an inlier if $h_i = 1$ and an outlier otherwise). Since the captured images used for the method of photometric stereo can have shadowed regions or regions of strong interreflections that are inappropriate for estimating the parameters, a mechanism of automat-

⋆1 What we estimate here is more like a set of reflectance maps that encode the illuminations along with the reflectance property.

ically excluding such image regions is necessary. The variable h_i for each point is estimated along with the polynomial coefficients using a RANSAC algorithm. Starting from $h_i = 1$ for all points, we iterate the above least squares minimization followed by the update of h_i for each point based on the residual error. Specifically, at each iteration, we set $h_i = 1$ if $\|\hat{\mathbf{n}}_i^* - \mathbf{F}(\mathbf{n}_i)\|$ is smaller than a pre-defined threshold for point i and set $h_i = 0$ otherwise. The threshold can be determined in a reasonable manner by using a statistical estimate of the angular errors of the DA normals. In our experiments, the RANSAC loop converges after 10–20 iterations.

As mentioned earlier, the DA normals have large random errors, whereas they have only small systematic errors. The reason why we use the DA normals, despite the fact that they can have large errors, is based on the following observation. We consider only objects having a uniform surface material, and thus we may assume the polynomial coefficients to be constant across the object surface. Therefore, we are to determine a small number of parameters from a large number of data. We then expect that the averaging effect will statistically eliminate the errors of the DA normals that are considered to be only random, and therefore the parameters can be precisely estimated.

A method for improving the accuracy of the normals measured by photometric stereo using the shape measured by triangulation is proposed by Nehab, et al. [8]. We highlight here differences between our study and their study. In their study [8], they observe that the biases present in the PS normals are low frequency and the errors in the DA normals are high frequency. They then argue that more accurate normals can be estimated by combining appropriate frequency bands, and present a method for combining them. The PS normals and the DA normals are first smoothed in the image domain to the same degree. The smoothing is performed in such a way that the coordinate functions of the normals are convolved with a Gaussian kernel and then they are renormalized to form a unit vector. Next, rotation that transforms the smoothed PS normal to the original PS normal is calculated at each point. Finally, for each point, the calculated rotation is applied to the smoothed DA normal; the rotation transforms the smoothed DA normal to a new normal, which is the corrected surface normal.

Although their observation with respect to the statistical natures of the PS normals and the DA normals is very similar to ours, their method of correcting the normals is considerably different from our method. An advantage of their method is that the correction of the normals is performed locally at each point and thus the method is expected to be able to deal with the case where the reflectance property varies across the object surface. (How local the correction is performed depends on the scale of the smoothing.) However, if it is known that the object surface has a uniform reflectance property, their method cannot make full use of the knowledge. Our method globally estimates the unknown reflectance property by assuming it to be uniform, which is considered to contribute to improve the overall accuracy of the correction of the surface normal. Our method is more appropriate for the purpose of this study.

2.6 Issues on Calibration of the System

2.6.1 Calibration of the Principal Points of Projectors

We assume for the camera and the projectors that the internal parameters but the focal length are all known. Thus, it is necessary to estimate them in advance. For the camera, it is possible to use existing calibration tools such as Camera Calibration Toolkit for Matlab. For the projectors, we employ the following calibration procedures.

First, the aspect ratio of an image pixel is determined from the factory sheet of the imaging engine of the projector. We then assume the skew to be 0. With respect to the principal point, it has usually a vertical offset for ordinary projectors such as PC projectors. Thus, special care is necessary to determine it. In our experience, it is easy and accurate to use the focus-of-expansion (FOE) of the projected image when manually varying the zoom value of the projector. The FOE of the projected image gives the projection of the image point that is on the optical axis of the projector lens. Since this image point coincides with the principal point, by identifying the projection of this image point and transferring this to the projector image, we obtain the principal point. The transfer (i.e., back-projection) is given by a 2D projective transformation (or a planar homography), which can be estimated by using a stationary camera. The identification of the FOE is also possible by using the same camera.

2.6.2 Photometric Calibration of Projectors

As mentioned above, we need only to determine the relative illumination

strengths of the projectors $1, \ldots, m$, since we do not determine the albedo ρ_i of each point. Although there can be a variety of methods that determine the relative illumination strengths, we choose a simple method that uses a planar diffuse surface. The surface is placed in a fronto-parallel manner toward the camera so that the surface and the image plane of the camera will be parallel. Since we are using the camera-centered coordinate system, the surface will have the normal $\mathbf{n} = [0, 0, -1]$. We acquire m images by turning on each projector in turn. Let $[b_1, \ldots, b_m]^\top$ be the brightnesses of the central image point for the m images. Let $\mathbf{l}_p$ $(p = 1, \ldots, m)$ be the multiplication of the illuminant strength of each projector and the vector connecting the surface point corresponding to the image point to each projector. There should be a relation $[b_1, \ldots, b_m]^\top \propto [\mathbf{l}_1, \ldots, \mathbf{l}_m]^\top [0, 0, -1]^\top$, from which we can determine the relative illuminant strengths of the projectors.

3. Computation of Projector Images

Using the series of methods described so far, we can obtain all the necessary information to synthesize a desired appearance by projecting images onto the surface of a real object. It is then necessary to be able to compute the projector images to realize the desired appearance. A method for the computation is shown in a literature [1], which assumes the projector to be an ideal point source. Aiming at more precise appearance synthesis, we introduce more realistic assumptions on the projector optics and reconsider how the projector images should be determined.

For the moment, we assume that the system has only a single projector, for the sake of simplicity. **Figure 2** shows the geometry of a projector and the object surface. Suppose a particular image pixel of the projector and its corresponding point (i.e., its projection) on the object surface. Let L_p be the radiance of this projector pixel and E_o be the irradiance of the surface point. (We assume here there is no (real) ambient illumination in space.) Approximating the projector optics by a single thin lens with diameter d, the relation between L_p and E_o is given by Ref. 13):

$$E_o = L_p \left(\frac{\pi d^2}{4 f^2} \right) \frac{\cos^3 \theta \cos \alpha}{r^2}, \tag{4}$$

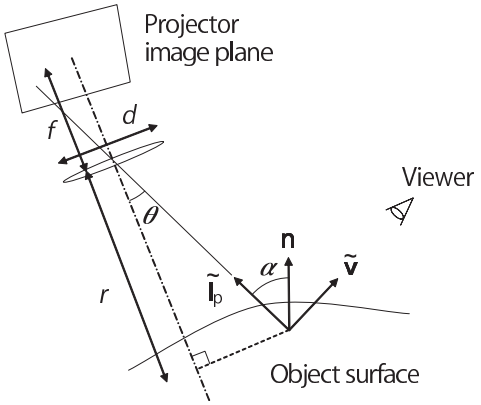


Fig. 2 Geometry of a projector and the object surface.

where f is the focal length of the projector lens, θ is the angle made by the line connecting the surface point and the optical center of the projector lens with its optical axis, α is the angle made by the same line with the surface normal at the surface point, and r is the distance from the projector lens to the surface point measured along the optical axis (Fig. 2).

Let I_p be the image brightness of this projector pixel. We assume here that the optics of the projectors is designed so that when an image (I_p) is projected onto a fronto-parallel screen, the irradiance (E'_o) of the screen will be proportional to the image at each screen point, i.e., $I_p \propto E'_o$. Applying the single thin lens model to this optical system [13], we have

$$I_p \propto E'_o = L_p \left(\frac{\pi d^2}{4 f^2 r_0^2} \right) \cos^4 \theta \propto L_p \cos^4 \theta, \tag{5}$$

where r_0 is the perpendicular distance to the screen from the projector. This equation means that the image irradiance L_p of a projector pixel is proportional to its input image brightness I_p with a correction to compensate for the cosine fourth law. By substituting this result into Eq. (4) to eliminate L_p, we can derive the relation between the image brightness I_p of a projector pixel and the surface irradiance E_o of the corresponding surface point as

$$E_o \propto I_p \frac{\cos \alpha}{r^2 \cos \theta}. \tag{6}$$

Let $\mathbf{X}_v$ be the spatial position of the viewer eye that we want to present the virtual appearance. We denote the true BRDF of the object surface by $f_o(\tilde{\mathbf{v}}; \tilde{\mathbf{l}}_p)$, where $\tilde{\mathbf{v}}$ and $\tilde{\mathbf{l}}_p$ are the direction of the viewer $\mathbf{X}_v$ and the projector lens in the local surface coordinates of the surface point of interest, respectively. (The tilde $\tilde{\cdot}$ indicates that the vectors are defined in the local coordinate system of the surface.) When the incident light from the projector to the surface point reflects in the direction of $\tilde{\mathbf{v}}$, the corresponding radiance L_o can be written as $L_o = f_o(\tilde{\mathbf{v}}; \tilde{\mathbf{l}}_p) E_o$. Therefore, by combining this with Eq. (6), the brightness of the projector pixel to realize the desired radiance $\hat{L}_o$ in the direction of the viewer $\mathbf{X}_v$ is given by

$$I_p \propto \frac{\hat{L}_o r^2 \cos\theta}{f_o(\tilde{\mathbf{v}}; \tilde{\mathbf{l}}_p) \cos\alpha}. \tag{7}$$

We have considered the case of a single projector so far. The system has in reality multiple (at least three) projectors. We generate the images of individual projectors by dividing the desired radiance $\hat{L}_o$ of Eq. (7) into multiple values $\hat{L}_{o1}, \ldots, \hat{L}_{om}$ so that $\hat{L}_o = \sum_{p=1}^{m} \hat{L}_{op}$ and assigning them to each of the projectors. For projector p, the assigned radiance $\hat{L}_{op}$ is calculated based on the cosine of the angle between the surface normal and the direction of the projector as

$$\hat{L}_{op} = \hat{L}_o \frac{\max(\cos\alpha_p, 0)}{\sum_{p=1}^{m} \max(\cos\alpha_p, 0)}, \tag{8}$$

where $\cos\alpha_p = \mathbf{n}^\top \mathbf{l}_p / \|\mathbf{l}_p\|$. The image brightness I_p of projector p is calculated by replacing $\hat{L}_o$ with $\hat{L}_{op}$ in Eq. (7).

The overall steps for generating the projector images are as follows. First, we specify a virtual BRDF and illumination for which we want to synthesize virtual appearance. The direction of the virtual illumination and that of the viewer position at each surface point are computed using the recovered 3D shape of the object. In order to convert these directions to the surface local coordinate system, the surface normals estimated by the method of Section 2.5 are used. Then, the desired radiance $\hat{L}_o$ that we wish to present to the viewer is calculated. This is substituted into Eq. (8) and then Eq. (7), where r and θ are computed also from the 3D shape, and α is computed from the surface normal. In the experiments shown later, we approximate the real reflectance $f_o(\tilde{\mathbf{v}}; \tilde{\mathbf{l}}_p)$ of the object surface with an ideal Lambertian reflectance and set $f_o(\tilde{\mathbf{v}}; \tilde{\mathbf{l}}_p) = 1$. It may be possible to determine it from the mapping $\mathbf{F}$ obtained in the step of the surface correction.

4. Experimental Results

4.1 Measurement of Surface Normals

We conducted an experiment to confirm the effectiveness of the method described in Section 2 that measures the surface normals of an object. **Figure 3** shows the experimental system. Three projectors (1,024 × 768 pixels) and a camera (1,280 × 960 pixels, Point Grey Research Grasshopper) are used.

4.1.1 Correction of Surface Normals

We first show how the surface normals are corrected by the proposed method. **Figure 5** shows an example. A plaster bust shown in **Fig. 4** is used as an object. Similarly as above, we use the terms, PS, DA, and "corrected" normals to distinguish the three kinds of normals. Figure 5 (a) shows the differences between the PS normals and the DA normals, more specifically, the sines of the angular differences of the two normals. Figure 5 (b) shows those between the corrected normals and the DA normals. The order of the polynomial used for the correction is $t = 3$. Figure 5 (c) shows the inlier/outlier distribution; the image points that are found to be the inliers are shown in white and outliers in black. It is observed in Fig. 5 (a) that there are systematic differences or biases in the PS normals over the entire image; the tendency is more significant around

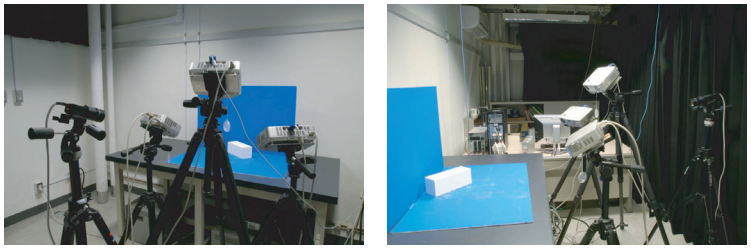
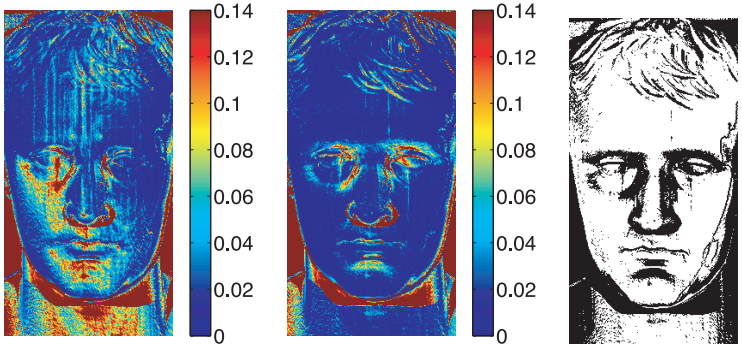


Fig. 3 Experimental configuration for examining the accuracy of surface normals. The system consists of three projectors and a camera.

(a) (b)

Fig. 4 (a) One of the three images captured for photometric stereo. (b) Measured shape by triangulation.



(a) (b) (c)

Fig. 5 (a) Angular differences between the PS normals before the correction and the DA normals. The sine of angular difference is plotted. (b) Angular differences between the corrected normals and the DA normals. (c) Inlier/outlier map. Inliers are shown in white and outliers in black. See text.

the left cheek and the neck of the bust. It is then observed in Fig. 5 (b) that these biases are well corrected by the proposed method except for a few regions, such as eyebrows, the sides of the nose. The erroneous regions correspond to shadowed areas or interreflections; they cannot be appropriately dealt with in our method. In fact, they are identified as outliers as shown in Fig. 5 (c). For the inliers, the corrected normals have only small differences from the DA normals in an average sense, which demonstrates that the proposed method works correctly.

Table 1 Angular differences between the averaged normals and the true normal for a parallelepiped.

	Angular difference (in degrees)
DA normal	3.0
PS normal	50.0
Corrected	5.4

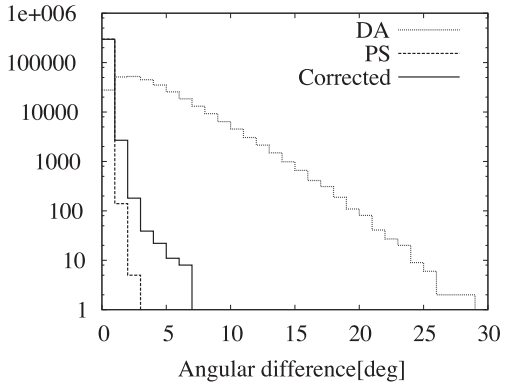


Fig. 6 Histograms of differences of neighboring surface normals at neighboring points of a planar surface. The horizontal axis indicates the difference in angles. The vertical axis is in log scale.

(The small differences that are observed for the inliers are due to the random errors inherent to the DA normals.)

4.1.2 Evaluation of Accuracy of Surface Normals

We next qualitatively examine the accuracy of the obtained surface normals. We first show the results obtained for an object having a simple shape, a rectangular parallelepiped made of plaster. Choosing a particular face of the parallelepiped, we examine the accuracy of the normals as follows. We first calculate the means of the surface normals belonging to the parallelepiped face. Since the face is ideally planar, the averaged normals should be close to the true normal of the face. **Table 1** shows the differences between the averaged normals and the true surface normal for the DA, PS, and corrected normals. We next evaluate how much the normals at neighboring two points differ. The face is planar and thus the differences should be small, or ideally, zero. **Figure 6** shows the

histograms of the angular differences for the three kinds of normals.

The following is observed from Table 1 and Fig. 6. In the case of the DA normals, the averaged normal has only a small error whereas the neighboring surface normals tend to have large differences. In contrast to this, in the case of the PS normals, the averaged normal has a very large error whereas the neighboring normals has only small differences. Thus, the DA normals are accurate in an average sense but are not smooth, while the PS normals are very smooth but are not accurate in an average sense. These agree with our observations described earlier. It is also seen from Table 1 and Fig. 6 that the corrected surface normals are smooth as well as accurate in an average sense, which confirms the effectiveness of the proposed method.

We next show results for an object having a more complicated shape. Since objects having a complicated but precisely known shape are difficult to prepare, we use here the fact that quality control is possible for the measurement of the object shape[⋆1]. For a target object, we measured its shape twice with different measuring parameters; one results in accurate shape having maximum quality, and the other is less accurate shape that is obtained by deliberately lowering accuracy[⋆2]. We then regard the DA normals obtained from the accurate measurement as the ground truth, and apply the proposed method to the less accurate measurement. The PS normals obtained independently of the shape measurements are corrected by using the less accurate DA normals. Finally, we compare the corrected normals thus obtained with the ground truth (i.e., the accurate DA normals). **Figure 7** shows the results; Fig. 7 (a) shows the angular differences between the accurate DA normals and the less accurate DA normals, and (b) shows those between the accurate DA normals and the corrected normals. Figure 7 (c) shows the inlier/outlier distribution when the correction of the normals is performed.

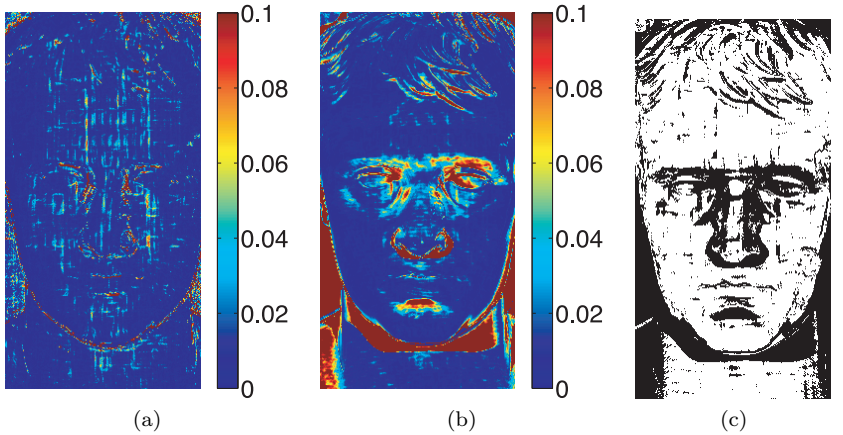


Fig. 7 (a) Angular differences between the accurate DA normals and the less accurate DA normals. (b) Angular differences between the accurate DA normals and the corrected normals obtained from the PS normals and the less accurate DA normals. (c) Inlier/outlier distribution when the correction is performed.

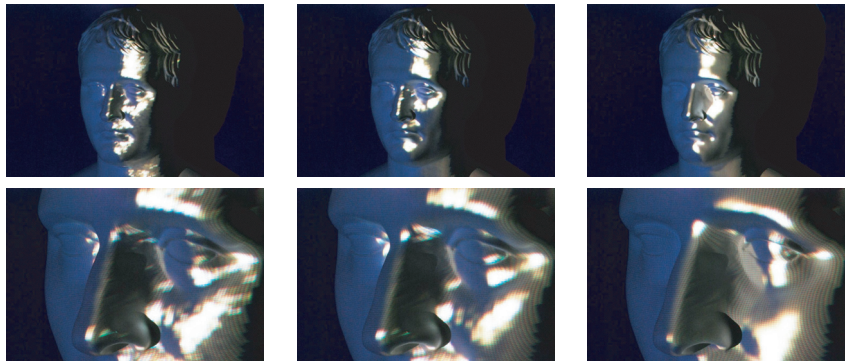
It is observed in Fig. 7 (a) that the less accurate DA normals considerably differ from the accurate DA normals, which appears to be noisy artifacts. On the other hand, it is observed in Fig. 7 (b) that the corrected normals have only small differences from the accurate DA normals for smooth regions of the object surface. Although there are large errors in some regions such as the eyebrows and others, they correspond to the outliers shown in Fig. 7 (c) and these errors are unavoidable in the proposed method. It is noteworthy that even the accurate DA normals that are regarded here as the ground truth do not achieve the accuracy demanded for our purpose. In fact, it is observed in Fig. 7 (b) that there are noisy textures in the smooth surface regions, although they are much smaller than Fig. 7 (a).

4.1.3 Qualitative Evaluation of Surface Normals

Using the surface shape along with precise normals thus obtained, arbitrary surface reflectance can be virtually synthesized on the object surface; the images to be projected are computed according to the method described in Section 3.

⋆1 The accuracy of the shape measurement depends on how accurately we can establish the geometric correspondences between the projector image and the camera image in the phase-shifting method. It is determined by several factors such as the wavelength of the sinusoidal wave and the accuracy of image brightness. Note that the accuracy of image brightness is increased to some degree by accumulating multiple images of the same scene and simply averaging them.

⋆2 For the accurate shape, we set the sinusoidal wave wavelength to be 20 pixels and the image accumulation count to be 30. For the less accurate shape, we set the wavelength to be 40 pixels and the image accumulation count to be 10. As a result, the former achieves more accurate measurement than the latter, with the sacrifice of the measurement time.



(a) (b) (c)

Fig. 8 Comparison of the synthesized appearances. (a) 3×3 filter. (b) 5×5 filter. (c) Proposed method.

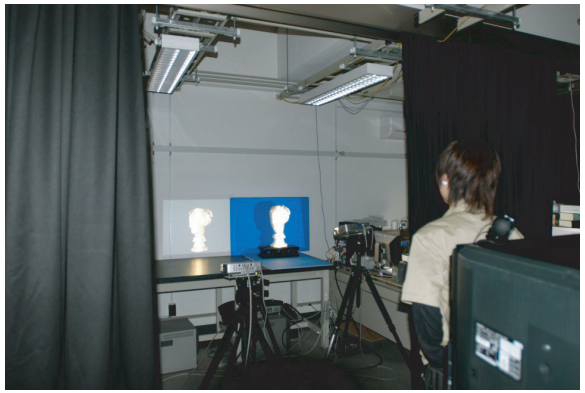


Fig. 9 Experimental setup for evaluating quality of the synthesized appearance. A planar screen and a target object are placed side by side. An image is projected onto the planar screen such that the image appears the same as the synthesized appearance when viewed from the same viewpoint.

Figure 8 demonstrates how large the impact of the accuracy of the surface normals is on the visual quality of the final appearances. The figure shows those for the case where the surface normals are derived from the reconstructed shape and for the case where they are estimated by the proposed method. For the virtual reflectance to be synthesized, the dichromatic reflectance model is assumed, where the diffuse component is given by the Lambertian model and the specular component is given by the Torrance-Sparrow model. The images in column (a) and (b) show the results when the surface normals are computed from the reconstructed shape and then smoothed by 3×3 and 5×5 pixel filters, respectively. Those in the column (c) show the results of the proposed method. It is observed that the highlights are randomly distorted in (a) and (b), whereas they are smooth and natural in (c). Note that the object surface is in reality smooth and does not have the undulations that yield those highlight distortions seen in (a) and (b).

4.2 Subjective Evaluation of Visual Quality

It is not easy to precisely evaluate quality of the synthesized appearance, i.e., how it is perceived by a viewer. One possible method is to regard quality as physical errors in brightness; brightness distribution on the surface of the target object is measured by using a camera and its difference from the ideal brightness distribution is calculated. However, this is not the same as directly measuring to what extent our original goal, "to make a real object appear as if its surface had a specified reflectance model," is achieved.

Therefore, we conducted experiments using a number of subjects to obtain subjective evaluation of quality. In order to examine if our system has a clear advantage to the case where an image is projected onto a planar screen by a projector, a planar screen and a target object were placed side by side. **Figure 9** shows the experimental configuration. The planar screen and the target objects are both made of plaster. An image was synthesized and projected onto the planar screen so that the image on the screen and the target object with synthesized appearance appeared to be geometrically the same when viewed from the same viewpoint; the viewpoint was determined in advance. We used projectors of the same product for the planar screen and the target object.

Then, the image on the screen and the target object with synthesized appearance were shown to the subjects, and they were asked which appears more realistic. In the experiments, we used several reflectance properties based on the dichromatic reflectance model, where the diffuse component is given by the Lambertian model and the specular component is given by the Phong or Torrance-

Fig. 10 Examples of synthesized appearances for two objects when assuming several virtual reflectance models.

Sparrow models; their model parameters were changed within certain ranges.

The subjects were asked to answer how real the synthesized appearance is. The answers were recorded using a continuous scale (0–100). They were told to use an identical criterion for the two cases under comparison: the image projection onto the planar screen and the proposed system. Since each subject will have different criteria for reality, it is meaningless to compare the evaluation scores among different subjects. However, it makes sense to compare within each subject in terms of which is perceived as more realistic.

We conducted the experiments using 18 subjects (ages from 21 to 40) and they all answered that the proposed system offered more realistic appearances. We

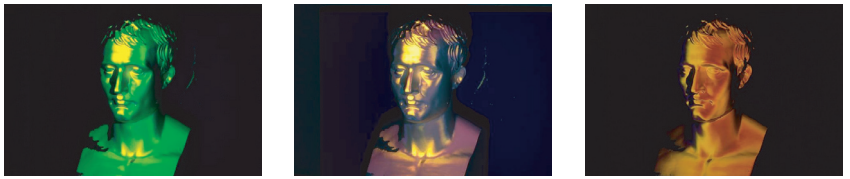


Fig. 11 Some results for wavelength-dependent reflectance models.

think that this result shows the effectiveness of the proposed system. In addition, it is noteworthy that when we changed the reflectance properties and showed the synthesized appearance to the subjects, a person commented "it is interesting to see that the same object appears to be made of different materials," and another person commented "I wonder how the authors prepare multiple objects made of different material that have the same shape" (We prepared only one object, of course.) We think that these comments verify the high quality of the synthesized appearances.

4.3 Demonstration of Reproduction of Various Reflectance Properties

By using the proposed method, it is straightforward to go from the capture of the object shape to the synthesis of appearance simulating any reflectance model, as shown in **Fig. 10**. Wavelength-dependent reflectance models [14] can also be dealt with, as shown in **Fig. 11**. It is observed for all these results that quality of the synthesized appearances is high; glosses of metallic surfaces are well reproduced that precisely reflect the delicate undulations of the object surface shape.

5. Summary

In this paper, we present a method for synthesizing a high-quality virtual appearance of an object when assuming an arbitrary reflectance property on the object surface. Using a system of multiple projectors and a camera, the method first estimates the surface shape as well as the internal and external parameters of the projectors and the camera based on structured light projection. Based on the estimation, it then performs photometric stereo using the same projectors as simple illuminations to measure the normals of the object surface. As is shown

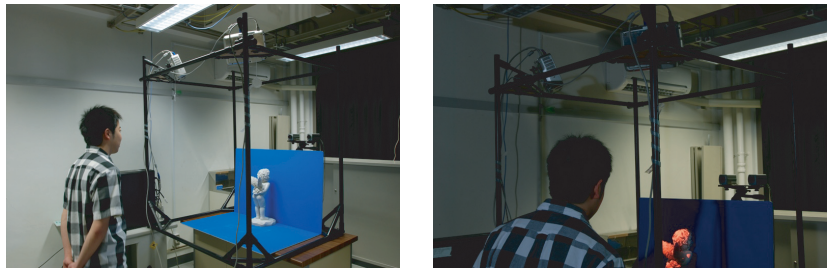


Fig. 12 Overview of an experimental system in which stereo cameras are used to track the viewer's head so that the change in the viewpoint is taken into account when synthesizing virtual appearance.

in the experimental results, when the surface normals are computed from the reconstructed surface by difference approximation, they will have random errors between neighboring surface points, which considerably deteriorates the visual quality of the synthesized appearance. The proposed method resolves this problem. Along with the accurate (auto) calibration of the projector-camera system as well as the accurate shape reconstruction, it enables the reproduction of any virtual reflectance property with high visual quality.

In this paper, assuming the viewer position to be fixed, we do not consider the case where the viewer moves. The appearance of an object will depend, of course, on the viewer position, and therefore it is necessary to consider the motion of the viewer. For this purpose, we have implemented a head-tracking function in our system, where that the viewer head is continuously tracked using stereo cameras, and uses the measured head position to synthesize viewer-dependent appearance of objects (**Fig. 12**).

References

1) Raskar, R., Welch, G., Low, K. and Bandyopadhyay, D.: Shader Lamps: Animating Real Objects With Image-Based Illumination, *Eurographics Rendering Workshop 2001* (2001).
2) Raskar, R., van Baar, J., Beardsley, P., Willwacher, T. and Rao, S.: Geometrically Aware and Self-Configuring Projectors, *Proc. SIGGRAPH 2003* (2003).
3) Raskar, R., Ziegler, R. and Willwacher, T.: Cartoon dioramas in motion, *Proc. International Symposium on Non-Photorealistic Animation and Rendering* (2002).
4) Grossberg, M.D., Peri, H., Nayar, S.K. and Belhumeur, P.N.: Making One Object Look Like Another: Controlling Appearance Using a Projector-Camera System, *Proc. CVPR2004*, pp.452–459 (2004).
5) Fujii, K., Grossberg, M.D. and Nayar, S.K.: A Projector-Camera System with Real-Time Photometric Adaptation for Dynamic Environments, *Proc. CVPR2005*, pp.814–821 (2005).
6) Yamamoto, S., Tsurase, M., Ueda, K., Tsumura, N., Nakaguchi, T. and Miyake, Y.: Reproducing an appearance of the objects using high bright projector, *AIC05 Annual Conference*, pp.1043–1046 (2005).
7) Woodham, R.J.: Photometric method for determining surface orientation from multiple images, *Optical Engineering*, Vol.19, No.1, pp.139–144 (1980).
8) Nehab, D., Rusinkiewicz, S., Davis, J. and Ramamoorthi, R.: Efficiently combining positions and normals for precise 3D geometry, *SIGGRAPH '05: ACM SIGGRAPH 2005 Papers*, pp.536–543 (2005).
9) Surrel, Y.: Design of algorithms for phase measurements by the use of phase-stepping, *Applied Optics*, Vol.35, pp.51–60 (1996).
10) Pollefeys, M., Koch, R. and Gool, L.V.: Self-calibration and metric reconstruction inspite of varying and unknown intrinsic camera parameters, *International Journal of Computer Vision*, Vol.32, No.1, pp.7–25 (1999).
11) Kanatani, K.: Gauge-based Reliability Analysis of 3-d Reconstruction from Two Uncalibrated Perspective Views, *Proc. ICPR 2000*, pp.1076–1079 (2000).
12) Oren, M. and Nayor, S.K.: Generalization of the Lambertian model and implications for machine vision, *International Journal of Computer Vision*, Vol.14, No.3, pp.227–251 (1995).
13) Horn, B.K.P.: *Robot Vision*, MIT Press, Cambridge, MA (1986).
14) Sun, Y.: Rendering biological iridescences with RGB-based renderers, *ACM Trans. Graph.*, Vol.25, No.1, pp.100–129 (2006).

(Received March 31, 2009)
(Accepted September 24, 2009)
(Released March 11, 2010)

(Communicated by *Stephen Lin*)

Tomoya Okazaki received his B.E and M.S. degrees from Tohoku University in 2007 and 2009, respectively. He has been engaged in research on physics-based vision and projector-camera systems. He is currently a researcher at Multimedia Laboratory, Corporate Research & Development Center, Toshiba Corporation.

Takayuki Okatani received his B.Eng., M.Eng., and Dr. Eng. degrees in mathematiacl engineering and information physics from the University of Tokyo in 1994, 1996, and 1999, respectively. From October 2003 to November 2004, he was a visiting scientist of Rutgers University, NJ, USA. He is currently an associate professor of the Graduate School of Information Sciences, Tohoku University, Japan.

Koichiro Deguchi received his B.Eng., M.Eng., and Dr. Eng. degrees from the University of Tokyo in 1974, 1976 and 1980, respectively. From 1984 to 1987, he was an associate professor of Department of Information Engineering, Yamagata University, from 1988 to 1998, he was an associate professor of Department of Mathematical Engineering and Information Physics, University of Tokyo, and in 1998, he joined Tohoku University as a professor of the Graduate School of Information Sciences. During 1991 and 1992, he was also a visiting associate professor of Washington University at Seattle, USA. His recent research interests include computer vision, robotics, and image measurement.