Shape Reconstruction by Combination of Structured-Light Projection and Photometric Stereo Using a Projector-Camera System

High Quality Reproduction of a Virtual Reflectance Property on a Real Object Surface

Tomoya Okazaki, Takayuki Okatani, and Koichiro Deguchi

Graduate School of Information Sciences, Tohoku University
6-6-01 Aza Aoba, Aramaki, Aoba-ku, Sendai, Japan
{okazaki,okatani,kodeg}@fractal.is.tohoku.ac.jp
http://www.fractal.is.tohoku.ac.jp

Abstract. In this paper, we present a method for synthesizing virtual appearance of an object by projecting images onto the surface of the object using projectors. The object surface is assumed to have a known diffuse reflectance property; its shape is allowed to have an arbitrary shape. Using a system consisting of multiple projectors and a camera, the method first estimates their internal and external parameters as well as the object surface based on the projection of structured patterns, and then measures surface normals by the method of photometric stereo that uses the same projectors as point sources of illumination. By enabling highly accurate calibration of the projectors as well as reconstruction of the object shape and also by reducing the random errors in surface normals that significantly affect final appearance, it is made possible to synthesize high-quality appearance associated with an arbitrary virtual reflectance property.

Keywords: Projector-camera, autocalibration, photometric stereo, augmented reality.

1 Introduction

There are several studies of the methods for realizing various visual effects by projecting images onto the surfaces of real-world objects using image projectors[1,2]. Their potential applications include virtual museums, industrial designs, and 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

T. Wada, F. Huang, and S. Lin (Eds.): PSIVT 2009, LNCS 5414, pp. 410–422, 2009.

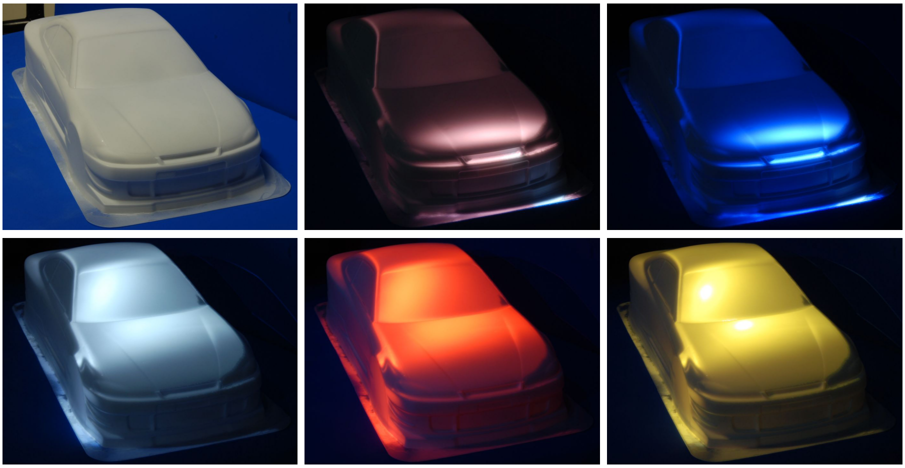


Fig. 1. A projector-camera system virtually reproduces arbitrary reflectance properties on a real-world object

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 in front of them; for example, it can immediately reflects the designer's choice of surface reflectance.

As mentioned earlier, there are similar methods that alter the appearance of real objects by using projectors. Raskar et al. present a method for visually reproducing apparent motion of a target object by changing the projected pattern[3]. Grossberg et al. and Fujii et al. present methods for realizing a desired appearance for a textured object by photometric compensation[4,5]. Yamamoto et al. present a method for reproducing an appearance of a real object that is the same as a reference object by image projection[6]. The purpose of this study is to highly accurately reproduce any virtual reflectance property on a real object surface that have an arbitrary shape by image projection.

In this paper, we consider a system consisting of multiple projectors and one and more cameras. We then assume dedicated objects for appearance synthesis, such that their surface has a simple reflectance property dominated by diffuse reflectance. As its realistic application, we consider a system for assisting a purchaser of an automobile in selecting a 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(Fig.1).

In order to synthesize a desired appearance by our system, besides the virtual reflectance property to be realized (which is given by a BRDF), it is necessary 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 especially important to acquire precise information as to (1) and (2). For this purpose, as was done in some of the previous studies, it makes sense to use the same projectors used for the appearance synthesis also for measuring the object shape; the projectors project structured light (e.g. stripe patterns) 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 on (1) and (2); its purpose is to maximize the visual quality of the synthesized appearance. Considering the above applications, the visual quality must be the main concern. (In our opinion, it seems to have been loosely considered in previous studies.) Toward this end, 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], along with the stereo-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. The existence of multiple projectors is inevitable to eliminate or minimize the shadowed areas on the object surface in the appearance synthesis.

Since the object shape is already reconstructed by the stereo-based method that uses structured light projection, it might appear awkward to further use photometric stereo to compute the surface normals. The necessity for photometric stereo stems from the fact that when using the surface normals computed from the reconstructed shape, it is quite difficult to synthesize high-quality appearance. There are two reasons for the difficulty. One is that 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. The other is that 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 (this is almost always the case). More specifically, the highlights of the synthesized appearance will have distorted erroneous 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. Since the estimated surface normals have direct relation to the appearance of the object, it is more appropriate to synthesize virtual appearance. Combining depths and normals for better shape reconstruction has been proposed by Nehab et al.[8]. In their method, there is a risk that small shapes are lost since it smooths the reconstructed shape to remove a slight error. In contrast, the shape is not smoothed in our method, since the surface normals are connected based on the nature of the reconstructed shapes that their errors are random. Therefore, small shapes can be maintained. Details are described in Sec.2.5.

2 Method for Acquiring Necessary Information for Appearance Synthesis

2.1 Problem Formulation

The system consists of three and more projectors and one camera. We assume the external parameters of the projectors and the camera to be unknown and their internal parameters to be known except for their focal lengths. As in the case of multi-camera systems, this setting is based on the fact that the focal lengths will vary whenever reconfiguring the system, whereas the other internal parameters are assumed to be constant (therefore it is sufficient to calibrate 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 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 of the projectors.

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 only unknown internal parameters, 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 [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, the method of bundle adjustment is performed. Since the corresponding points in the projector image are not measurements but true values, we minimize the sum of the squared distance between the measurements and their estimates with respect to the corresponding points in the camera images. 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 the all 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

Then, using the same projectors as simple illumination sources, the method of photometric stereo is applied. Photometric stereo [7] estimates the normal of an object surface from its multiple images taken under different illuminations, and it assumes the illuminant directions to be known. In our case, we have obtained the projector poses relative to the object surface as above and use them here. When projecting an uni-colored pattern onto the object surface from a projector, we assume it to be a point source of illumination; the position of the point source coincides with the projector position. Although the projectors have projection optics, this is a good approximation when the projectors are distant from the object surface.

The same camera captures three and more shaded images for different illuminant directions by projecting an 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 $b_{pi} = \rho_i \mathbf{n}_i^\top \mathbf{l}_p$, where ρ_i is the albedo, $\mathbf{n}_i$ is the surface normal, and $\mathbf{l}_p$ is the orientation of the projector p from the surface point of interest multiplied by the strength of the illumination. The brightness $b_{1i}, \ldots, b_{mi}^\top$ for m different illuminant directions ($p = 1, \ldots, m$) are given by

$$\begin{bmatrix} b_{1i} \\ \vdots \\ b_{mi} \end{bmatrix} = \begin{bmatrix} \mathbf{l}_1^\top \\ \vdots \\ \mathbf{l}_m^\top \end{bmatrix} (\rho_i \mathbf{n}_i). \tag{1}$$

The relative strengths of the projector light sources are calibrated in advance. Then, $[\mathbf{l}_1, \cdots, \mathbf{l}_m]^\top$ can be calculated using also the estimated projector poses relative to the object surface. Solving Eq.(1) in a least squares sense with respect to the unknown $\rho_i \mathbf{n}_i$, the surface normal $\mathbf{n}_i$ and the albedo ρ_i are determined. When the surface reflectance is not Lambertian but it is known, these can be determined by performing nonlinear optimization.

As will be observed later, as compared with those computed from the surface shape by difference approximation, the surface normals thus obtained tend to have much smaller random errors between neighboring surface points. This local accuracy of the estimation is the main reason that we employ here the method of photometric stereo. On the other hand, it is unavoidable that the estimated surface normals have systematic biases due to the modeling errors of the real reflectance property of the object and interreflections, which is a well known limitation of photometric stereo. In the next subsection, we present a method for correcting such systematic biases by estimating the real reflectance property.

2.5 Correction of Surface Normals Based on Recovered 3D Shape

As described above, the surface normals obtained from photometric stereo will have small random errors but large systematic errors. The main source of the systematic errors is the modeling error of the real reflectance property of the object. Thus, the systematic errors can be mitigated by estimating the real reflectance property from the measured data.

Since we wish to obtain accurate surface normal after all, instead of estimating the reflectance property itself, we estimate here nonlinear transform from the normal $\mathbf{n}$ that has been obtained by assuming Lambertian reflectance to the normal $\tilde{\mathbf{n}}$ that would be obtained by assuming true reflectance. Namely, $\tilde{\mathbf{n}} = F(\mathbf{n})$. We then represent F by k-th order polynomial as

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

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

$$\tilde{n}_z = F_z(\mathbf{n}) = \sum_{\alpha=0}^{k}\sum_{\beta=0}^{k}\sum_{\gamma=0}^{k} c_{\alpha\beta\gamma} n_x^\alpha n_y^\beta n_z^\gamma. \tag{2c}$$

Then, we estimate the coefficients $a_{\alpha\beta\gamma}$, $b_{\alpha\beta\gamma}$, and $c_{\alpha\beta\gamma}$ for.

In order to estimate these parameters, the correct surface normal $\tilde{\mathbf{n}} = [\tilde{x}, \tilde{y}, \tilde{z}]^\top$ is necessary. We use the surface normals obtained by differentiating the recovered 3D shape. As mentioned earlier, these normals have large random errors, whereas they have only small systematic errors. Then, we estimate the parameter by minimizing the sum of the Euclidean distance between $\mathbf{n}$ and $\tilde{\mathbf{n}}$ (i.e. the normals obtained from the 3D shape) over the entire object surface:

$$(\tilde{n}_x - n_x)^2 + (\tilde{n}_y - n_y)^2 + (\tilde{n}_z - n_z)^2. \tag{3}$$

Since the number of data participating in this minimization is much larger than the degrees of freedom of the parameters to be estimated, this minimization is expected to yield good parameter estimates. The parameter estimates are obtained by a linear least squares method.

2.6 On the Calibration of Projector Image Center

We assume for both 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 the 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, a 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.

3 Computation of Projector Images

A virtual appearance is reproduced by projecting images from the projectors onto the surface of the real object. The images that are input to the projectors are computed in the following manner.

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 in the direction of the projector lens and E_o be the irradiance of the surface point. (We assume here there is no (real) ambient illumination in space.)

Further let I_p be the image brightness of this projector pixel. The optical system of a projector is usually designed so that when an image is projected onto a fronto-parallel screen, the irradiance (E_o') of the screen is proportional to I_p. By assuming the distance to the screen to be r_0, we have

$$I_p(\propto E_o') \propto L_p \left(\frac{\pi d^2}{4r_0^2}\right) \cos^4 \theta. \tag{4}$$

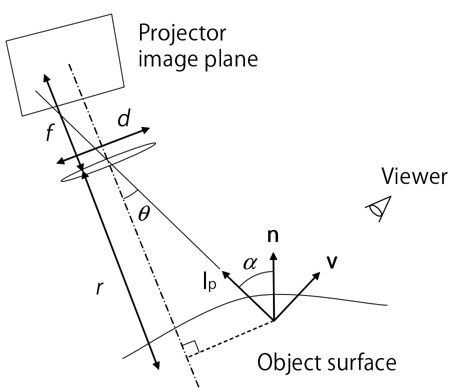


Fig. 2. Geometry of a projector and the object surface

The cosine fourth law may, in reality, not be accurate considering the optical construction of projectors. In that case, we will need to photometrically calibrate the projectors.

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 $\tilde{\mathbf{l}}_p$ indicate the direction of the viewer $\mathbf{X}_v$ and the projector lens, respectively, in the local surface coordinates of the surface point of interest. 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.(4), 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

$$I_p \propto \frac{\hat{L}_o r^2 \cos\theta}{f_o(\tilde{\mathbf{v}}; \tilde{\mathbf{l}}_p)\cos\alpha}. \tag{5}$$

The projector image that realizes a desired appearance of the object is computed in the following steps. First, virtual illumination and virtual BRDF of the object are selected. The directions of the virtual illumination and viewer position at each surface point are computed using the recovered 3D shape of the object. In order to convert these directions to the representation in the surface local coordinates, the surface normals estimated by the proposed method are used. Then, the desired radiance $\hat{L}_o$ that we wish to present to the viewer is determined. This is substituted into Eq.(5), where r and θ are computed from the 3D shape, and α is computed from the surface normal. With respect to the real reflectance $f_o(\tilde{\mathbf{v}}; \tilde{\mathbf{l}}_p) = 1$, we use a mathematical model in the experiments shown in the next section, which approximates the reflectance of the object surface material. It could be possible to derive the reflectance from the estimated nonlinear transform F in the method presented in 2.5 and to use it as $f_o(\tilde{\mathbf{v}}; \tilde{\mathbf{l}}_p) = 1$ here.

Following these steps, the brightness I_p of the image that is input to the projector is determined. Note that when the projector and the object surface is sufficiently distant, it is a good approximation to assume the projector to be a point light source. We assumed so in the experiments in what follows.

4 Experimental Results

We conducted experiments to test the efficacy of the proposed method. Fig.3 shows the overview of the experimental setup. The system consists of four NEC VT595 projectors and one DELL 3400 MP projector (all have 1024×768 pixels) and a Point Grey Research Frea2 camera (1024×768 pixels) with a Fujinon HF12.5SA-1 lens. We used several objects that have diffuse reflectance properties; they are assumed to be Lambertian.

The proposed method starts with the pattern projection shown in Fig.3 to obtain the point correspondences. Then, the focal lengths as well as the external parameters of the projectors and the camera is estimated; the result is shown in Fig.4. Using these parameters, the shape of the object is reconstructed. Next,

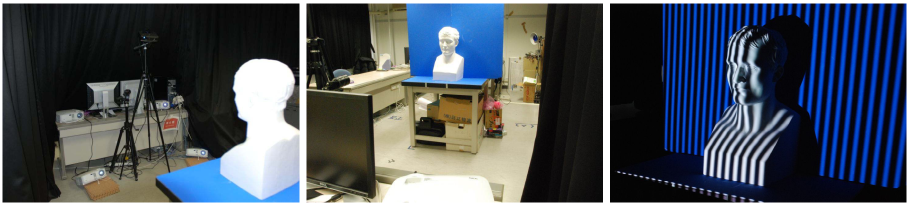


Fig. 3. The experimental setup (left and middle) and the structured pattern projection (right)

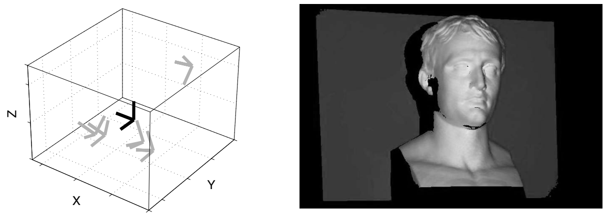


Fig. 4. A result of the autocalibration-based shape reconstruction. Left: Estimated poses of the projectors and the camera. Right: Reconstructed shape.

the surface normals are computed by the photometric stereo method. While the five projectors project a uni-colored pattern in turn, the camera captures the images, as shown in Fig.5.

In order to demonstrate the efficacy of using photometric stereo in combination with the (geometric) stereo-based shape reconstruction, we show in Fig.6 the surface normals computed from the reconstructed shape by difference approximation and those obtained by the proposed method. In the images, the x components of the surface normals are represented as brightness. They are mostly identical when globally comparing the two images. However, when comparing in a finer scale, it is observed that they are considerably different. There exist several artifacts in the surface normals obtained from the reconstructed shape The possible causes for the artifacts are the errors of the phase estimation in the phase shifting method, aliasing due to the quantization of the projector and the camera images, etc. They are not easy to eliminate by, for example, spatially smoothing the surface normals, as will be demonstrated below.

Using the surface shape along with precise normals thus obtained, arbitrary surface reflectance can be virtually reproduced on the object surface; the images to be projected are synthesized according to the physics-based model [2] between a projector image and the projected image on the object. In the experiments, we reproduced 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 Phong or Torrance-Sparrow models; their model parameters were changed within certain ranges. When synthesizing the appearance, arbitrary virtual illumination can be used.

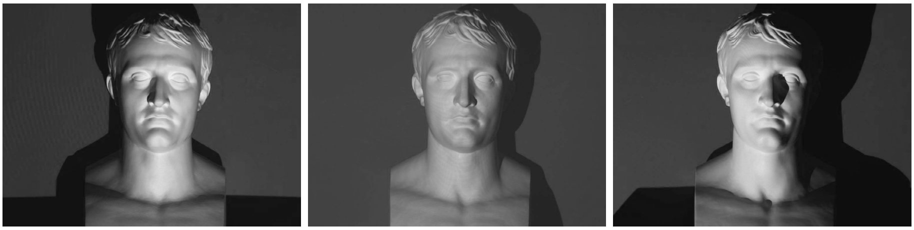


Fig. 5. Selected three images used for photometric stereo

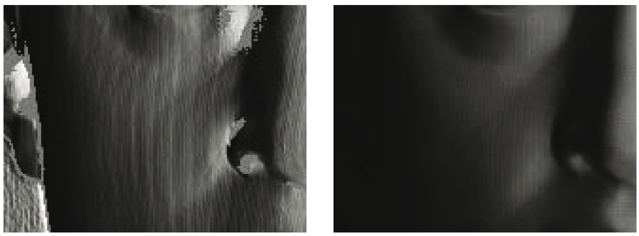
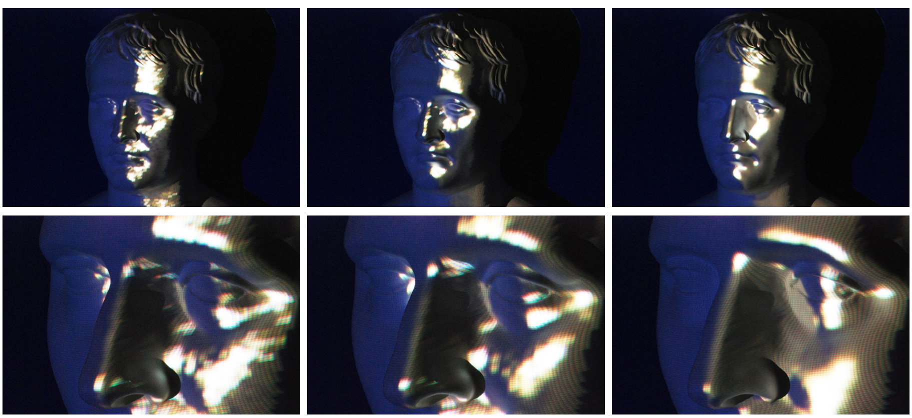


Fig. 6. Visualized surface normals. Left: When computed from the shape. Right: Proposed method.



(a) (b) (c)

Fig. 7. Comparison of the synthesized appearances. (a) 3x3 filter. (b) 5x5 filter. (c) Proposed method.

Again, in order to demonstrate the efficacy of the proposed method, we show in Fig.7 the synthesized appearances between the case where the surface normal is computed from the reconstructed shape and the case where they are estimated by the proposed method. The images in column (a) and (b) show the results when the surface normals are computed from the reconstructed shape and then

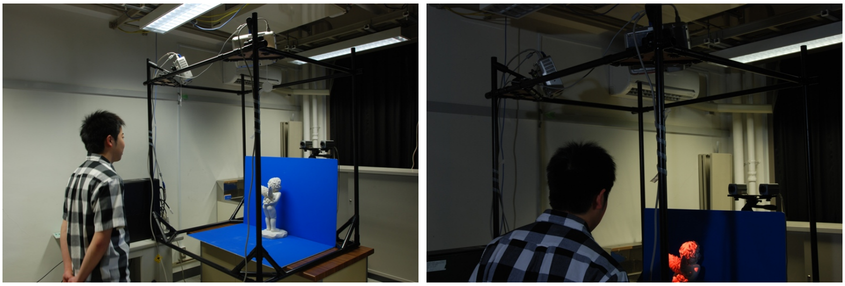


Fig. 8. Left: An overview of a working system of the virtual reflectance reproduction. A camera and two projectors were used in the calibration stage, and they are removed in the stage of virtual appearance display. Right: Image-based head tracking is now incorporated into the system to enable the appearance change in response to viewer head motion.

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 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).

By using the proposed method, it is possible to easily acquire the surface shape information that enables natural appearance simulating any arbitrary reflectance property.

Fig.9 shows several results; also see Fig.1. It is observed that high-quality appearance is realized; glosses of metallic surfaces are reproduced that are completely different from the real reflectance of the objects, and they precisely reflects 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 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 resolve 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.

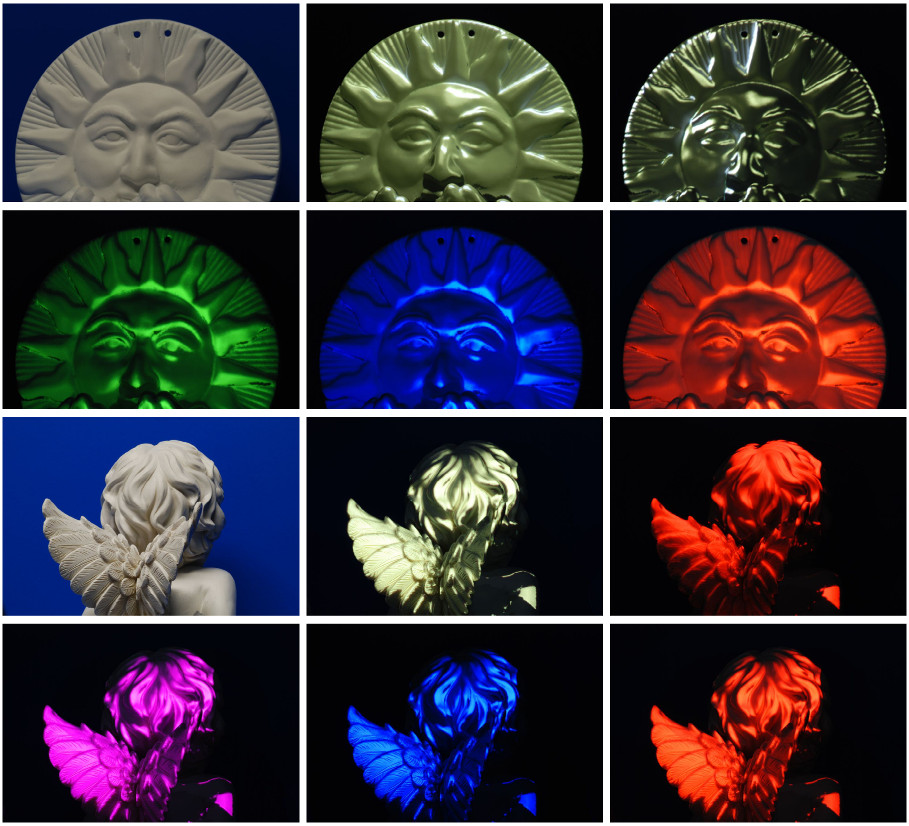


Fig. 9. Examples of the appearance synthesis when assuming various virtual reflectance properties

In this paper, assuming fixed viewer position, we have not considered the case where the viewer moves. However, the appearance of an object will depend on the viewer position, and therefore it is necessary to consider viewer movement. In fact, we have implemented a method that tracks the viewer head motion with a 6D sensor and/or cameras, and uses it to synthesis viewer-dependent appearance of objects(Fig.8).

References

1. Raskar, R., Welch, G., Low, K., Bandyopadhyay, D.: Shader Lamps: Animating Real Objects With Image-Based Illumination. In: Eurographics Rendering Workshop 2001 (2001)
2. Raskar, R., van Baar, J., Beardsley, P., Willwacher, T., Rao, S.: Geometrically aware and self-configuring projectors. In: Proceedings of SIGGRAPH 2003 (2003)
3. Raskar, R., Ziegler, R., Willwacher, T.: Cartoon dioramas in motion. In: Proceedings of International Symposium on Non-Photorealistic Animation and Rendering (2002)

4. Grossberg, M.D., Peri, H., Nayar, S.K., Belhumeur, P.N.: Making one object look like another: Controlling appearance using a projector-camera system. In: Proc. CVPR 2004, pp. 452–459 (2004)
5. Fujii, K., Grossberg, M.D., Nayar, S.K.: A projector-camera system with real-time photometric adaptation for dynamic environments. In: Proc. CVPR 2005, pp. 814–821 (2005)
6. Yamamoto, S., Tsurase, M., Ueda, K., Tsumura, N., Nakaguchi, T., Miyake, Y.: Reproducing an appearance of the objects using high bright projector. In: AIC 2005 Annual Conference, pp. 1043–1046 (2005)
7. Woodham, R.J.: Photometric method for determining surface orientation from multiple images. Optical Engineering 19(1), 139–144 (1980)
8. Nehab, D., Rusinkiewicz, S., Davis, J., Ramamoorthi, R.: Efficiently combining positions and normals for precise 3d geometry. In: SIGGRAPH 2005: 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 35, 51–60 (1996)
10. Pollefeys, M., Koch, R., Gool, L.V.: Self-calibration and metric reconstruction inspite of varying and unknown intrinsic camera parameters. International Journal of Computer Vision 32(1), 7–25 (1999)
11. Kanatani, K.: Gauge-based reliability analysis of 3-d reconstruction from two uncalibrated perspective views. In: Proc. ICPR 2000, vol. I, pp. 76–79 (2000)