Surface Extraction from Multiple Views with Different Zoom Ratio

Yoshinari NAKANISHI [†], Xiaohua ZHANG [†], Kiichi KOBAYASHI [†],

Hideki MITSUMINE [††], and Suguru SAITO [†††]

[†]:NHK Engineering Services Inc.,   [††]: NHK Science and Technical Research Laboratories

[†††]:Tokyo Institute of Technology

Ａ ｂ ｓ ｔ ｒ ａ ｃ ｔ

We propose a high-resolution digitizing approach to the extraction of shape and surface reflectance of a real 3D object. The object is set on a rotary table and taken by a HDTV camera. The camera and the rotation axis of rotary table are well calibrated. The 3D object's shape is recovered through stereoscopic measurement using multiple images by rotating the rotary table horizontally and the camera vertically. To measure a region with a narrow section, we employed the images taken with a long focal length to estimate its shape. The entire 3D shape is obtained by merging all point data in 3D space through the estimated zoom ratio. A surface triangle mesh is generated to represent the object shape. For each 3D object, the intensity variation data from all views is computed for estimating surface reflectance. Experimental results are presented to demonstrate the effectiveness of our approach.

1.   INTRODUCTION

In recent years, a rapid increase in multi-channel systems and HDTV broadcasts has stimulated an ever-increasing demand for digital content. However, at present, many photo-realistic 3D objects are created manually. This is not only time consuming but requires advanced skills such as an ability to use specialized application softwares. 3D objects created in this way are used in combination with actual images. Achieving a sense of reality is therefore important when forming composite pictures of this kind. To this end, it is essential to make high-resolution shape measurements, to obtain texture that excludes the lighting conditions present when taking the image, and to simulate reflection under any lighting.

Under these circumstances, we have constructed a system for extracting the surface shape and reflectance of an object.

For shape measurement, various techniques have been researched including the range-finder, pattern irradiation by a projector, and "shape from X" . However, We have selected the stereo method, considering our need for a system that can be applied to any object and that can easily relate the depth of unit pixels to texture. This method is capable of performing high-resolution shape measurements, and based on this method, we devised a multiple-viewpoint image matching method that statistically processes multiple viewpoint images to measure shape.

Moreover, considering that the stereo method is also capable of measuring narrow sections of an object and areas for which change in surface shape is minute, we have also devised a technique for precisely estimating the zoom ratio between two images taken with different zoom values. With this technique, we can capture each section of an object using appropriate zoom ratios and then merge the results to obtain the shape of the entire object.

Likewise, various methods have been researched for extracting reflectance of an object's surface. For example, there is a method that analyzes a color histogram and a method that describes the brightness of a monochromatic object by a dichromatic reflection model [1], [2]. These methods, however, are not applicable to objects having complex textures. For this purpose, there is a method that separates reflection components using a polarizing filter and a method that determines reflectance from intensity-variation data of sequential images. The former method, however, is not adequate due to the complexity of component separation, and the latter method requires calibration of light-source color beforehand and accurate equipment setup that can be difficult to maintain.

Against the above background, we have devised a new method that directly determines reflectance from sequential images on a pixel-by-pixel basis. Not only does this method not require prior calibration of light-source color, it also employs two point light sources thereby solving the problem of texture vanishment by shading. Furthermore, the fact that reflectance can be extracted using the same images as those used for shape extraction is a great advantage.

2.   SYSTEM DESCRIPTION

Fig. 1 shows our experimental setup. This system consists of a HDTV camera, curved boom equipment, a rotary table and light sources. The object to be measured is placed on the rotary table, and images of the object are taken from multiple viewpoints by rotating the table

(while keeping the camera fixed). The position of the camera can also be changed using the boom equipment. This setup uses two light sources that can be treated as point light sources as long as they are placed sufficiently far from the object.
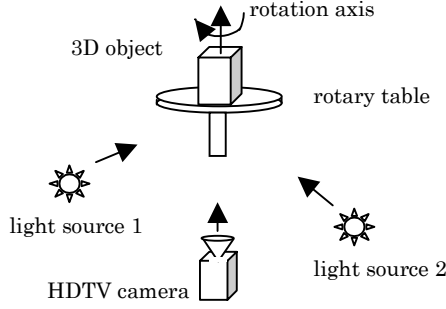

Fig. 1: Experimental setup

The camera and rotation axis of the rotary table are calibrated beforehand. For camera movement using the boom equipment, the relative position of the camera is calibrated. Camera calibration after changing the zoom ratio is described in Section 3. The positions and directions of the two point light sources are calibrated beforehand by trigonometry. Camera white balance is also adjusted in accordance with light-source color. The above calibrations and adjustments enable measurement of occlusion areas and narrow areas that move little due to rotation.

3. CALIBRATION

The camera and rotation axis of the rotary table are calibrated using the images of a checkered panel taken from various angles. When shooting a particular part of an object with different zoom values, the zoom ratio will be calibrated based on phase correlation method in the frequency space of the image. In this method, two images having overlap sections are taken with different zoom values, and the overlap sections are extracted from these images as sections I_1 and I_2. Now, given that these images are spatially offset by (x_0, y_0), the following equation holds.

$$I_2(x, y) = I_1(x - x_0, y - y_0) \qquad (1)$$

Next, the above two images are subjected to a discrete Fourier transform, and the resulting cross-power spectrum can be given as follows.

$$\frac{F_2(\xi,\eta)F_1^*(\xi,\eta)}{\left|F_2(\xi,\eta)F_1^*(\xi,\eta)\right|} = e^{-j2\pi(\xi x_0 + \eta y_0)} \qquad (2)$$

Here, F* denotes the conjugate complex number of Fourier transform F. If an inverse Fourier transform is now performed on the above, we obtain the phase-correlation curved surface whose peak positions correspond to offset (x_0, y_0) of sections I_1 and I_2.

We next determine the rotational angle and zoom ratio for I_1 and I_2. Denoting rotational angle as θ_0 and zoom ratio as α, the Fourier transform becomes as follows.

$$F_2(\xi, \eta) = \frac{1}{|\alpha \cdot \alpha|} F_1(\xi/\alpha, \eta/\alpha) \qquad (3)$$

Now, denoting the respective magnitude images as M_1 and M_2, their relationship can be expressed as follows in polar coordinates.

$$M_2(\rho, \rho, \theta) = M_1(\rho, \rho/\alpha, \theta_0) \qquad (4)$$

Converting this to logarithmic space, we get the following equation.

$$M_2(\log\rho, \theta) = M_1(\log\rho - \log\alpha, \theta - \theta_0) \qquad (5)$$

From this, $\log\alpha$ and θ_0 can be determined as in the above method for (x_0, y_0).

Assuming that other internal camera parameters are not affected by zoom change, only external camera parameters need be calibrated after zoom change.

4. EXTRACTION OF SURFACE SHAPE

4.1 Two Step Block Matching Method Using Multiple Viewpoint Images

In the stereo matching method, the correlation between two images increases as the distance between two viewpoints decreases, which facilitates matching detection. On the other hand, error in depth measure-ments increases as the distance between the viewpoints (length of baseline) decreases. We therefore propose a two-step block matching method using multiple viewpoint images as a high-resolution measuring method that is compatible with these two contradictory properties.

4.1.1. Multiple Viewpoint Images

Our proposed method makes use of images from multiple viewpoints obtained by fixing the camera and rotating the rotary table. This method selects an image for which measurement points on the object's surface have been captured under favorable conditions, and uses images taken within the range $\pm\theta°$ as multiple viewpoint images. For example, when measuring points captured near the center of an image having a rotation angle of 0°, images may be taken at 2° increments within ±10° about that image at 0° for a total of 11 images.

4.1.2. Matching Algorithm

The multiple-viewpoint image matching method uses a group of images obtained by fixing the camera

and rotating the target object on a rotary table. As described above, this method selects an image for which measurement points on the object's surface have been captured under favorable conditions. Next, it determines a line-of-sight equation for the position of a measure-ment point using camera calibration data. Then, after assuming the distance between the measurement point and the rotation axis of the rotary table (rotation radius), the method determines the point of intersection between the above line-of-sight and the cylinder defined by the rotation axis and the rotation radius.

The method now rotates the point of intersection by the angle corresponding to each image, projects the point onto each image, and extracts, for each image, a block image with this point at center. Finally, by evaluating the correlation between block images extracted in this way using an evaluation function, the rotation radius exhibiting maximum correlation here can be detected (matched) and the three-dimensional coordinates of the measurement point determined.

In our two step block matching method, block images are statistically evaluated as described. Given N multiple viewpoint images, there are $N(N-1)/2$ combinations of two block images from the N-extracted block images. The evaluation function first determines the correlation coefficient of each combination and then computes the total sum of all coefficients. At this time, a weight is introduced to each of these correlation coefficients in proportion to the distance between viewpoints. Considering that measurement error is small for long base lines, this weighting is to give a higher evaluation to a correlation coefficient of high-reliability combination.

To achieve measurements at even higher resolu-tion, matching detection is performed hierarchi-cally in two steps called primary and secondary matching. Primary matching performs matching detection by the algorithm described above. Secondary matching, on the other hand, first magnifies the multiple viewpoint images by four times in the vertical and horizontal directions and then performs matching detection by limiting the matching search to a range near the results of primary matching. In addition, the graph of evaluation-function values versus rotation radius in secondary matching is used to perform a linear approximation of the first-order derivatives of five peak points, and the zero cross point of the resulting line is taken to be the final matching result.

5. EXTRACTION OF SURFACE REFLEC-TANCE

5.1 Reflection Model under two point light sources

We describe the reflection model referring to the geometry of the experimental setup shown in Fig. 1. For the sake of understanding, we divide this geometry into (a) and (b) sections. Fig. 2 (a) omits the normal vector while Fig. 2 (b) omits the direction vectors of the two light sources. All vectors in (a) and (b), however, exist in the same coordinate system. To keep things simple, the optical axis of the camera coincides with the Z-axis and the rotation axis of the rotary table with the Y-axis in this discussion. In the figure, N is the normal vector of a measurement point and L_1 and L_2 are the unit direction vectors of the two light sources.
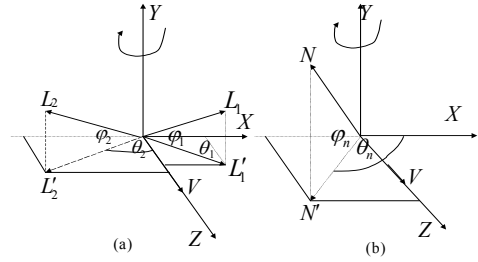

Fig. 2: Geometry of experimental setup

The projections of N, L_1, and L_2 onto the XOZ plane and the angles thereby formed with that plane are N', L_1', and L_2' and φ_n, φ_1, and φ_2, respectively. In addition, the angle formed between N' and the X-axis is θ_n, that between L_1' and the X-axis is θ_1, and that between L_2' and the Z-axis is θ_2. Angles φ_1, φ_2, θ_1, and θ_2 are measured beforehand at the calibration step.

In this paper, we use the Torrance-Sparrow model [4] as a reflection model. Here, the intensity of reflected light is expressed as the sum of the diffuse reflection component (Id) and specular reflection component (Is) as given in Eq. (6).

$$I = I_d + I_s$$
$$= K_d \sum_{i=1}^{2} (L \cdot N) + K_s \sum_{i=1}^{2} \exp(-\frac{\alpha_i^2}{2\sigma^2}) \qquad (6)$$

Here, Kd and Ks are the coefficients of diffuse reflection and specular reflection, respectively. Symbols α_1 and α_2 denote the angle formed between the viewpoint-direction vector and the light-source direction vector and the angle formed between the bisector and normal vector. The symbol σ denotes the standard deviation of the facet slope indicating the roughness of the surface.

In polar coordinates, L_1 and L_2 become (φ_1, θ_1) and (φ_2, θ_2), respectively, and the normal vector after rotating the measurement point by θ becomes $(\varphi_n, \theta_n + \theta)$. The first term of Eq. (6), that is,

the diffuse component, can now be expressed as shown in Eq. (7) below.

$$I_d = A\sin\theta + B\cos\theta + C,\qquad(7)$$

Here, A, B, and C are as follows.

$$\begin{cases} A = K_d t_1 \cos\varphi_n \\ B = K_d t_2 \cos\varphi_n \\ C = K_d t_c \sin\varphi_n \end{cases}\qquad(8)$$

$$\begin{cases} t_1 = t_a \cos\theta_n + t_b \sin\theta_n \\ t_2 = t_b \cos\theta_n - t_a \sin\theta_n \end{cases}\qquad(9)$$

$$\begin{cases} t_a = \cos\varphi_1 \cos\theta_1 - \cos\varphi_2 \sin\theta_2 \\ t_b = \cos\varphi_1 \sin\theta_1 + \cos\varphi_2 \cos\theta_2 \\ t_c = \sin\varphi_1 + \sin\varphi_2 \end{cases}\qquad(10)$$

As shown by Eqs. (8) − (10), coefficients A, B, and C are independent of rotation angle θ of the rotary table but dependent on light source and viewpoint directions.

The specular reflection component can be ex-pressed as follows for the case of two point light sources if we approximate the change in intensity of specular reflection by a Gauss function the same as in [5].

$$I_s = D_1 \exp\left(-\left(\frac{E_1 - \theta}{F}\right)^2\right) + D_2 \exp\left(-\left(\frac{E_2 - \theta}{F}\right)^2\right)$$

(11)

Here, variables E_1 and E_2 indicate the position of the peak values in the intensity variation curve for measure-ment points and F is $\sqrt{2}\sigma$. In addition, D_1 and D_2 are coefficients corresponding to K_s of Eq. (6).

As described above, intensity variation of measurement points can be expressed in terms of eight variables: A, B, C, D_1, D_2, E_1, E_2, and F. In the following section, we determine these eight variables by parameter fitting using iterative calculations and then compute reflection coefficients K_d and K_s and surface roughness coefficient σ.

5.2 Parameter Fitting

Although intensity variation of measurement points can be expressed in terms of eight variables, intensity variation as actually observed can be obtained by using accurate calibration results and projecting the results of the measurements onto the multiple viewpoint images. Fig. 3 shows an example of intensity variation obtained in this way.

To extract object reflectance (K_d, K_s, σ), we first determine the eight variables described above. As shown by Eqs. (7) and (11), the diffuse component is a linear function in terms of variables A, B, C but the specular component in a non-liner function in D_1, D_2, E_1, E_2, and F. We therefore determine

these variables by minimizing the error function shown in Eq. (12) using the Levenberg-Marquardt method.

$$E = \sum_k W_{\theta_k}\left(I(\theta_k; A, B, C, D_1, E_1, D_2, E_2, F) - I_k\right)\qquad(12)$$
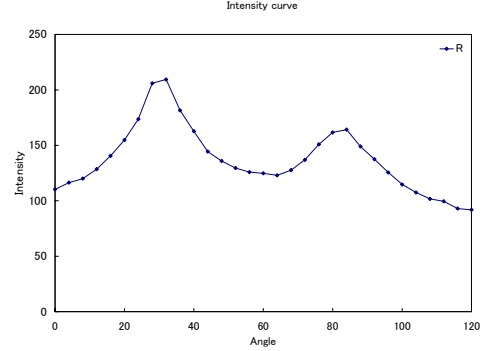
Fig. 3: An example of intensity variation curve

Here, W is a weighting function given as follows.

$$W_{\theta_k} = \left[a + \exp\left(-\left(\frac{E_1 - \theta_k}{F}\right)^2\right) + \exp\left(-\left(\frac{E_2 - \theta_k}{F}\right)^2\right)\right]^{-\frac{m}{n}}\qquad(13)$$

This weighting function is added as it is assumed that estimations of diffuse components are more reliable than those of specular components. This is due to the difficulty of modeling the specular component and to the higher noise in highlight sections caused by the camera's limited dynamic range. The purpose of constant a in Eq. (13) is to prevent the value of the weighting function from becoming infinite and is empirically set to a value from 10^{-9} to 10^{-3}. Likewise, setting $m \geq 0$ and $n > 0$ controls the shape of the weighting function. A large value of the ratio m/n means that dependency on diffuse components increases in the fitting .

When using the Levenberg-Marquardt method, appropriate initial values must be given to prevent the process from falling into a local solution. First, intensity-variation data (RGB data) of observed measurement points are converted to brightness-signal values. Two peak positions are then found from this converted data and used as initial values E_1 and E_2. Next, peak values corresponding to these E_1 and E_2 positions are determined for each RGB channel and used as initial   values D_1 and D_2 of each channel. Variable F may be given an empirical value from 0.01 to 1.0. The values of D_1 and D_2 differ for each RGB channel and can be treated as vectors within RGB space. Similarly, the RGB values of measurement points can also be treated as vectors in RGB space. The average vector for D_1 and D_2 is therefore computed and the vector of observed RGB values that forms the largest angle with this average vector is taken to be the initial-value vector of variables A, B, and C in

relation to diffuse compo-nents. The above method, while similar to that proposed by Sato et al. [5], does not require prior determination of light-source color and does not fix object color or light-source color.

5.3 Separation of diffuse and specular components

We have so far determined the eight variables in Eqs. (7) and (11) and obtained an intensity variation curve. We now insert these variables in those equations and determine diffuse component I_d' and specular component I_s'. These I_d' and I_s' values are estimates, however, and are affected by calculation error and noise. To therefore achieve consistency with measured intensity values, we determine diffuse and specular components using the following equation.

$$I_d = I * \frac{I_d'}{I_d' + I_s'} \quad I_s = I * \frac{I_s'}{I_d' + I_s'} \quad (14)$$

This method can achieve good results in most cases. An object with strong specular reflection, however, suffers from saturated brightness values and a flat specular peak section that prevents reflection components from being adequately separated. This is because a Gauss-function approximation of the specular component is not appropriate in this case. Estimation of the diffuse component is therefore thought be more reliable than that of the specular component. We therefore compute I_s by subtracting the value of I_d estimated by the above method from the observed brightness value.

5.4 Calculation of reflection parameters

We now calculate reflection parameters K_d, K_s, and σ from the eight variables described above and light-source direction vectors L_1 and L_2. Angle θ_n can be determined by the following equation from Eqs. (8) and (9).

$$\theta_n = \arctan\left(\frac{t_b - \delta t_a}{\delta t_b + t_a}\right) \quad (15)$$

Here, $\delta = A/B$ and t_a and t_b can be determined from Eq. (10). Next, using the value of θ_n found here, t_1 and t_2 is determined from Eq. (9) and φ_n is determined by the following equation from Eq. (8).

$$\varphi_n = \arctan\left(\frac{C}{t_c} \cdot \frac{t_1 + t_2}{A + B}\right) \quad (16)$$

In this way, we get the normal vector of a measurement point, and compute diffuse reflection parameter K_d as follows from normal vector N, direction vectors L_1 and L_2, and I_d.

$$K_d = \frac{I_d}{\max\{L_1 \cdot N, 0\} + \max\{L_2 \cdot N, 0\}} \quad (17)$$

In many cases, the two light-source direction vectors and the line-of-sight direction vectors are not on the same plane. It is for this reason that the exponential part of Eq.

(11) is not 1.0. The existence of a specular section on the intensity variation curve is therefore assumed and Ks is determined by the following method. Specifically, D_1 is taken to be greater than D_2 and K_s is determined using only large-peak D_1. In this case, the intensity value of the measurement point for a rotation of E_1 takes on a maximum value. Now, letting L_{b1} denote the bisector to the angle formed by direction vector L_1 of light-source-1 and the line-of-sight vector, K_s can be computed as follows.

$$K_s = D_1 \exp\left(-(\beta / F)^2\right) \quad (18)$$

Here, β is the angle formed between L_{b1} when intensity is maximum and the normal vector to the measurement point. In addition, σ can be found from the equation $\sigma = \sqrt{2}F/2$ from the definition of F.

6. EXPERIMENTAL RESULTS

Fig. 4 shows an example of object images taken with different zoom ratios. The object is a small flower vase with a narrow upper section (neck). Measurements of this section are difficult with the image of Fig. 4(a) but can be done with that of Fig. 4(b) taken after zooming in.
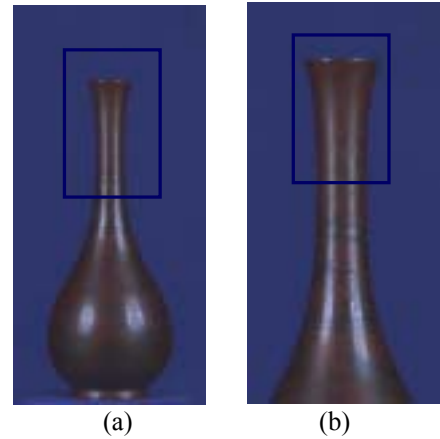

(a)   (b)
Fig. 4: Images used for shape measurements

Fig. 5 shows the results of measuring the surface shape of this object using these two images. The former figure shows the results of measuring one vertical line in the sampled image. The error here between the approximating curve of the measurement results and the measurement results themselves was about 0.29 pixel. This satisfies the target resolution of 0.6 pixel.

Fig. 6 shows the results of merging the point cloud measured from four directions(front/back/left/right) in the images of the object taken with different zoom values and creating polygons by a method that enhances the matching-triangles method.

Fig. 7 shows the results of extracting reflectance. Fig. 7(a) is the original image, while Fig. 7(b) and 7(c) show diffuse and specular reflection components, respectively, that have been separated from the original

image. Fig.7(d), moreover, shows an image of extracted color (albedo map). These figures demonstrate that the separation of diffuse and specular components and estimation of object color can be performed well.
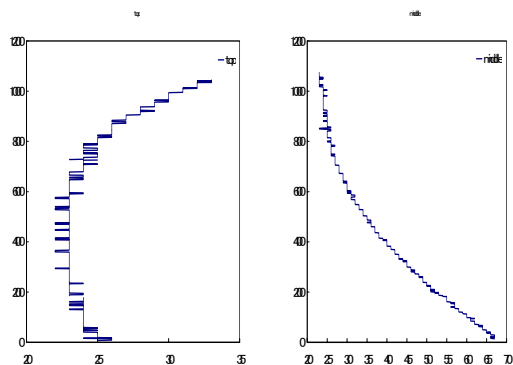
However, while specular components are very few in

(a) Upper section   (b) Middle section
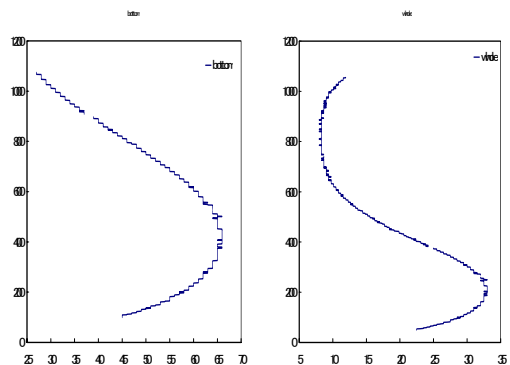
(c) Lower section   (d) After merging

Fig. 5: Results of shape measurements using a zoomed-in image

object texture, some locations can be seen where they remain. This is attributed to the occurrence of error in registration between the images when measuring intensity variation of measurement points. This problem must be addressed in future research. The experi-mental results presented above, while indicating some room for improvement, show that the algorithm proposed here can sufficiently extract shape and reflectance of an object's surface.
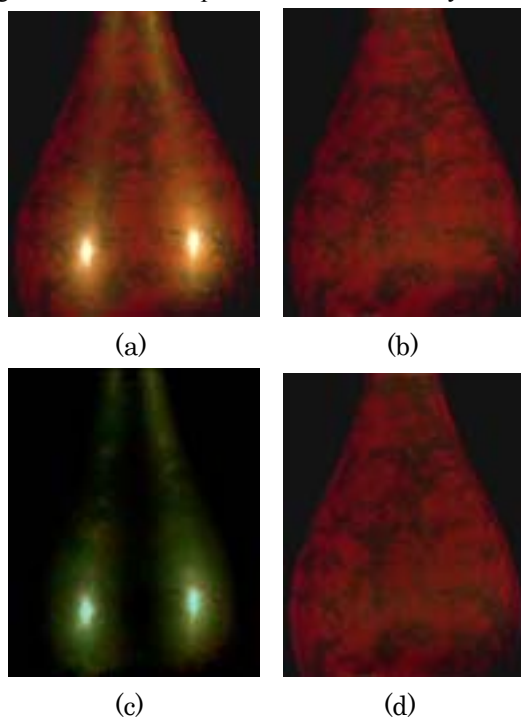

Fig. 6:Extracted shape

7. CONCLUSION

In the research reported here, we have developed a technique for extracting shape and reflectance of an object's surface from multiple images. This technique can 1) accommodate change in zoom ratio in addition to viewpoint position and direction; 2) perform high-resolution shape measurements by statistical

(a)   (b)

(c)   (d)

processing of multiple viewpoint images, and 3) perform robust

Fig. 7: Experimental results with real images

extraction of surface reflectance and object color using two point light sources whose colors are unknown. The results of experiments, moreover, were found to be favorable for both shape measurement and extraction of reflectance thereby demonstrating the effectiveness of the proposed technique. In future research, we plan to apply this technique to various types of objects to determine its range of effectiveness and applicability.

REFERENCES
[1] G. J. Klinker, S. A. Shafer, and T. Kanade, "Using a color reflection model to separate highlights from object color", International Conference on Computer Vision, pp. 145-150, 1987.
[2] G. J. Klinker, S. A. Shafer, and T. Kanade, "The measure-ment of highlight in color images", International Journal of Computer Vision (IJCV), Vol. 2, No. 1, pp. 7-32, 1988.
[3] M. Tadenuma, Y. Nojiri, "Relationship between Fluctu-ations of Disparity Data and Picture Quality in a

Stereoscopic Images", Proceedings of 1998 ITE Annual Convention, 14-6

[4] K. E. Torrance and E. M. Sparrow, "Theory for Off-Specular Reflection from Roughened Surfaces", Journal of the Optical Society of America, (JOSA), No. 57, pp. 1105-1114, 1967.

[5] Y. Sato and K. Ikeuchi, "Temporal-color space analysis of reflection", Journal of Optical Society of America A, Vol. 11, No. 11, pp. 2990-3002, Nov. 1994