Abstract
In this paper, we propose a compact and low-cost system to measure an accurate 3D shape of the object. The proposed system consists of a camera, a projector and eight LED light sources. We use a structured light technique to recover the shape of the object by using the projector. The surface normal of the object is estimated based on a photometric stereo method applied to the images taken with eight LED light sources. Conventionally, the light sources are assumed as parallel light sources to estimate the surface normal. Therefore, the measurement system becomes large to keep the long distance between the object and light source. In this paper, we attach the light sources near the object and assume that the light sources are point light sources. Based on the position of the point light source and the shape obtained by structured light method, we can calculate the incident direction and the intensity at each point of the object, and estimate the surface normal accurately in the compact system. The recovered shape and estimated surface normal are combined by the technique proposed by Diego Nehab et al. [SIGGRAPH, pp. 536-543, 2007]. We evaluated the proposed system by measuring the sample object. The sample object is also measured by a commercial 3D scanner, and we evaluate our system in comparing both shapes. From the results of the evaluations, we confirmed the effectiveness of our proposed system.

Introduction
With the development of a digital imaging system, a digital archiving system has been used in museums to preserve art works and exhibit them on displays. However, most of objects in such museums are preserved and exhibited as 2D images. Recently, for the accurate preservation of the objects, many researches have been performed to record the objects as 3D data. For example, J. Stumpfel et al. recoded the shapes of the sculptures of the Parthenon with the resolution of 1mm [1], and F. Bernardini et al. measured the 3D shape and the surface normal of Michelangelo’s Florentine Pieta [2], where the surface normal shows the roughness of the surface. In these researches, the 3D measurement system requires high accuracy and resolution to measure the shape and normal accurately. Furthermore, the system has to be compact and inexpensive for practical use.

Recently, there are many commercial 3D scanning devices. These devices can measure the 3D shape and surface normal of an object accurately and the sizes of the devices are enough small for practical use. However, the cost of these systems is so high that these devices are not used practically.

For reducing the cost, 3D scanning methods using digital cameras and projectors based on triangulation have been proposed (e.g. stereo [3], structured light technique [4]). These methods can measure the shape of the object as accurately as the commercial 3D scanners. Although the normal can be estimated by calculating the gradient from the obtained shape, a slight error of the shape greatly affects the accuracy of the normal vector estimation.

On the other hand, photometric stereo method [5] can estimate the surface normal directly from the images of the object under illuminations with different directions. The shape of the object can be recovered by integrating the normal. However, the accuracy of the recovered shape is not high because of the propagation of the errors of the estimated normal. Furthermore, since parallel light sources are assumed in photometric stereo method, the measurement system becomes large to keep the long distance between the object and the light sources.

D. Nehab et al. [6] has presented a hybrid algorithm that combines the shape recovered by the
triangulation method and the normal estimated by photometric stereo method to produce a new shape and normal that approximates both. Although their proposed algorithm improved the accuracy of the shape and normal dramatically, the size reduction of the measurement system is not considered in their paper.

In this paper, we improve the Nehab’s algorithm and propose a compact and low-cost system to measure an accurate 3D shape and normal of the object.

The key to reduce the size of the measurement system is that we don’t assume parallel light sources but assume points light sources in photometric stereo method. We first confirm the effectiveness of the proposed method of photometric stereo with point light sources. Then, the shape and the normal of one sample object are measured using the proposed measurement system. This object is also measured by a commercial 3D scanner, and the accuracy of the shape and normal with the proposed system is evaluated by comparing both shapes and normals. From both evaluations, we confirm the effectiveness of the proposed method and system.

**Measurement System**

The geometry of the proposed measurement system is shown in Fig. 1 (a). The illustrations of the side view of the system and the view from the object are shown in Figs. 1 (b) and (c). The system is similar to the one made by D. Nehab et al. [6]. One digital camera (Nikon D1x, 4016×2624) views an object from the distance of 0.8 meter. A DLP projector (Toshiba TDP-FF1A, 800×600) casts a series of graycode patterns onto the object from the distance of 0.8 meter while the camera captures the images. Eight LED light sources with polarized filters illuminate the objects one by one from the distance of approximately 0.6 meter. The polarized filter is also set in front of the camera for removing the surface reflection, and the diffuse reflection images are taken by the camera. The width, height and depth of the system are approximately 0.6, 0.8 and 0.8 meters. These sizes are as small as the ones of commercial 3D scanning systems. In the following section, we describe the techniques used in the proposed system.

**Proposed Method**

The flow of the process in the proposed method is shown in Fig. 2. First, the 3D shape and the normal of the object are recovered from the images projected with graycode patterns using structured light technique. Using the recovered shape and photometric stereo method, the surface normal is estimated from the diffuse reflection images illuminated by the eight LEDs. Finally, by combining the recovered shape and the estimated normal using the Nehab’s algorithm, the improved shape and normal of the object are obtained.

**Structured light technique**

We measure the shape of the object by structured light technique, respectively. Before the measurement, it is necessary to calibrate the camera and the projector. The camera and projector calibrations are performed by the method of R. Sukthankar et al. [7], however we ignore the lens distortion due to the slight effects to the results.

First, we describe the camera calibration. Using the $3\times4$ camera projection matrix $P$, the camera projection model is given by
Figure 2. Flow of the process in the proposed measurement system.

Figure 3. (a) The calibration board and the world coordinate system and (b) the captured image of the calibration board by the camera.

Figure 4. Description of the stripe value. When, the position \((x, y, z)\) is projected by the \(w\)th stripe in the projected image, the stripe value is \(w\).

\[
\begin{bmatrix}
y\
x\
z
\end{bmatrix} = \begin{bmatrix}
P_{c11} & P_{c12} & P_{c13} & P_{c14} \\
P_{c21} & P_{c22} & P_{c23} & P_{c24} \\
P_{c31} & P_{c32} & P_{c33} & P_{c34}
\end{bmatrix} \begin{bmatrix}
y\
x\
z
\end{bmatrix} + \begin{bmatrix}
a_{c1} \\
a_{c2} \\
a_{c3} \\
a_{c4}
\end{bmatrix},
\]

where \((x, y, z)\) is 3D positions in the world coordinates system, \((u, v)\) is 2D positions in the digital image coordinates system and \(a_c\) is a scale factor.

The camera calibration is performed by determining the 12 unknown camera parameters of \(P_c\). The camera parameters can be determined by extracting several sets of 3D positions and 2D positions using the image of the calibration object. Figure 3 (a) shows the calibration board and the world coordinates system, and Fig. 3 (b) shows the captured image of the calibration board.

Suppose that \(m \geq 8\) sets of 3D positions and 2D position are given, Equation (1) is rewritten as follows,

\[
\begin{bmatrix}
y_i \\
x_i \\
z_i
\end{bmatrix} = \begin{bmatrix}
P_{c11} & P_{c12} & P_{c13} & P_{c14} \\
P_{c21} & P_{c22} & P_{c23} & P_{c24} \\
P_{c31} & P_{c32} & P_{c33} & P_{c34}
\end{bmatrix} \begin{bmatrix}
y_i \\
x_i \\
z_i
\end{bmatrix} + \begin{bmatrix}
a_{c1}y_i \\
a_{c2}y_i \\
a_{c3} \\
a_{c4}
\end{bmatrix},
\]

where \(B_c\) is a following \(12 \times 2m\) matrix

\[
B_c = \begin{bmatrix}
x_i & y_i & z_i & 1 & 0 & 0 & 0 & 0 & -u_i & -v_i & -u_i & -v_i \\
0 & 0 & 0 & 1 & x_i & y_i & z_i & 1 & -u_i & -v_i & -u_i & -v_i \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
x_m & y_m & z_m & 1 & 0 & 0 & 0 & 0 & -u_m & -v_m & -u_m & -v_m \\
0 & 0 & 0 & 1 & x_m & y_m & z_m & 1 & -u_m & -v_m & -u_m & -v_m
\end{bmatrix},
\]

and \(p_c\) is a following unknown camera parameter vector

\[
p_c = \begin{bmatrix}
P_{c11} \\
P_{c12} \\
P_{c13} \\
P_{c14} \\
P_{c21} \\
P_{c22} \\
P_{c23} \\
P_{c24} \\
P_{c31} \\
P_{c32} \\
P_{c33} \\
P_{c34}
\end{bmatrix}.
\]
constrained \( | \mathbf{p} | = 1 \). Equation (3) can be solved to find the unit vector \( \mathbf{p} \) that minimize \( | \mathbf{B} \mathbf{p} | \) as a least square solution and the solution is given by the eigenvector corresponding to the smallest eigenvalue of \( \mathbf{B}^T \mathbf{B} \).

The projector calibration can be performed similar to the camera calibration with a one dimensional case. Using the 2×4 projector projection matrix \( \mathbf{P}_p \), the projector projection model is given by

\[
\mathbf{a}_p \begin{pmatrix} w \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} p_{p,11} & p_{p,12} & p_{p,13} & p_{p,14} \\ p_{p,21} & p_{p,22} & p_{p,23} & p_{p,24} \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{6}
\]

where \( a_p \) is a scale factor and \( w \) is a stripe value at \((u,v)\).

The stripe value \( w \) means that the 3D position \((x, y, z)\) is projected by the \( w \)th stripe of the projected image. Since we use the horizontal stripe patterns, the range of the stripe value is equal to the vertical resolution of the projector. Figure 4 also shows the description of the stripe value.

The stripe value can be obtained by combining set of images of the calibration board projected graycode patterns as shown in Fig. 5 (a) [8]. We also estimate the substripe value using the method proposed by R. J. Valkenburg, and A. M. McIvor [9]. The stripe value image of the calibration board is shown in Fig. 5 (b).

Substituting the given \( m \) set of 3D positions and 1D strip values into Equation (6), we obtain

\[
\mathbf{B}_p \mathbf{s} = \mathbf{0}, \tag{7}
\]

where \( \mathbf{B}_p \) is a following 8×\( m \) matrix

\[
\mathbf{B}_p = \begin{pmatrix} x_1 & y_1 & z_1 & -w_x x_1 & -w_y y_1 & -w_z z_1 & -w y_1 & -w z_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_m & y_m & z_m & -w_x x_m & -w_y y_m & -w_z z_m & -w y_m & -w z_m \end{pmatrix} \tag{8}
\]

and \( \mathbf{P}_p \) is a following unknown projector parameter vector

\[
\mathbf{P}_p = \begin{pmatrix} p_{p,11} & p_{p,12} & \cdots & p_{p,14} \\ p_{p,21} & p_{p,22} & \cdots & p_{p,24} \end{pmatrix} \tag{9}
\]

constrained \( | \mathbf{p} | = 1 \). The unit vector \( \mathbf{p} \) is given by the eigenvector corresponding to the smallest eigenvalue of \( \mathbf{B}_p^T \mathbf{B}_p \) same as the camera calibration.

From the Equations (1) and (6), the following equation can be obtained,

\[
\begin{pmatrix} p_{11} \mathbf{s} & p_{12} \mathbf{s} & \cdots & p_{14} \mathbf{s} \\ p_{21} \mathbf{s} & p_{22} \mathbf{s} & \cdots & p_{24} \mathbf{s} \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} p_{11} \mathbf{s} & p_{12} \mathbf{s} & \cdots & p_{14} \mathbf{s} \\ p_{21} \mathbf{s} & p_{22} \mathbf{s} & \cdots & p_{24} \mathbf{s} \end{pmatrix} = \begin{pmatrix} p_{11} \mathbf{s} \mathbf{s} & p_{12} \mathbf{s} \mathbf{s} & \cdots & p_{14} \mathbf{s} \mathbf{s} \\ p_{21} \mathbf{s} \mathbf{s} & p_{22} \mathbf{s} \mathbf{s} & \cdots & p_{24} \mathbf{s} \mathbf{s} \end{pmatrix} \tag{10}
\]

\[
\mathbf{Q} \mathbf{s} = \mathbf{f},
\]

where \( \mathbf{s} \) denotes the 3D positions vector \((x, y, z)^T\) at \((u,v)\).

After the camera and projector calibrations, we take a set of graycode patterns images of the object, and the stripe value image of the object is calculated by the similar way in the projector calibration. The stripe value image allows to obtain the position map \( \mathbf{s}(u,v) \) as the shape of the object by solving the Equation (10) for \( \mathbf{s} \) at all \((u,v)\).

Using the recovered shape, the normal can be estimated by calculating gradients to the neighbor pixels. The gradients \( g_x(u,v) \) and \( g_y(u,v) \) in the direction of \( x \) and \( y \) at \((u,v)\) can be calculated by the following equations,

\[
g_x(u,v) = \frac{z(u+1,v) - z(u-1,v)}{x(u+1,v) - x(u-1,v)} \tag{11}
\]

\[
g_y(u,v) = \frac{z(u,v+1) - z(u,v-1)}{y(u,v+1) - y(u,v-1)} \tag{12}
\]
where \(x(u,v), y(u,v)\) and \(z(u,v)\) are the 3D position \((x, y, z)\) at \((u,v)\). Using the gradients, the normal vector \(n_s\) can be obtained as the following normalized tangent vector,

\[
\mathbf{n}_s = \frac{(-g_y, -g_z, 1)^T}{\|(-g_y, -g_z, 1)^T\|}.
\]

Figures 6 (a) and (b) shows the recovered shape and the estimated normal of the object using structured light technique. The color R, G and B in the normal map shows the absolute value of \(X, Y\) and \(Z\) components of the normal vector. illumination direction and the distance for the light source are varied dependent on the 3D position.

**Conventional photometric stereo method**

After the shape and normal are recovered, we also estimate the surface normal from diffuse reflection images by photometric stereo method. Before introducing our proposed photometric method with point light sources, we describe conventional photometric stereo method with parallel light sources [5]. Suppose that the surfaces of the object is Lambertian surfaces, a pixel value \(i\) in the a diffuse reflection image is given by

\[
i = k\rho (\mathbf{l} \cdot \mathbf{n}_p),
\]

where \(k\) is the intensity of the light source, \(\rho\) is diffuse reflectance, \(\mathbf{l}\) is the unit vector of illumination direction and \(\mathbf{n}_p\) is the surface normal vector at \((x,v)\) (see Fig. 7). Given three pixel values \(i_1, i_2\) and \(i_3\) for illumination direction vectors \(\mathbf{l}_1, \mathbf{l}_2\) and \(\mathbf{l}_3\), Equation (14) can be rewritten in matrix form as follows,

\[
\begin{bmatrix}
i_1 \\
i_2 \\
i_3
\end{bmatrix} = k\rho \begin{bmatrix}
\mathbf{l}_1^T \\
\mathbf{l}_2^T \\
\mathbf{l}_3^T
\end{bmatrix} \mathbf{n}_p.
\]

If all three illumination directions \(\mathbf{l}_1, \mathbf{l}_2\) and \(\mathbf{l}_3\) do not lie in the same plane, the matrix \(L\) is non-singular and can be inverted as follows,

\[
k\rho \mathbf{n}_p = L^{-1} i.
\]

From the Equation (16), the normal vector \(\mathbf{n}_p\) can be calculated with the constraint of \(\|\mathbf{n}_p\| = 1\). If pixel values with more than three illumination directions are given, normal vector can be obtained as a least squares solution by calculating the pseudo-inverse matrix of \(L\) and solving Equation (16).

**Proposed photometric stereo method**

In the conventional photometric stereo method, the light sources are assumed as parallel light sources. However, in the proposed compact measurement system, the distance between the object and LED light sources is not enough large to assume the light sources as parallel light sources. Therefore, we assume that the LEDs are point light sources, and propose new photometric stereo method with point light sources.

Under the assumption of point light sources, the incident light directions are varied dependent on the 3D position. The intensity of the illuminations is also decreased in proportional to the square of the distance between the light source and the position. Therefore, Equation (14) is rewritten as follows,
where $d$ is the distance between the 3D position at $(u,v)$ and the light source, and $l'$ is the new unit vector of illumination direction with considering the variation dependent on the 3D position. Using the recovered shape and the known light position, $d$ and $l'$ can be calculated as shown in Fig. 8.

By giving three pixel values $i_1$, $i_2$, and $i_3$ with illumination direction $l'_1$, $l'_2$, and $l'_3$ and distance $d_1$, $d_2$, and $d_3$, Equation (14) can be rewritten in matrix form as follows,

$$
\begin{bmatrix}
i_1 \\
i_2 \\
i_3
\end{bmatrix} = k \rho
\begin{bmatrix}
1/d_1^2 & 0 & 0 \\
0 & 1/d_2^2 & 0 \\
0 & 0 & 1/d_3^2
\end{bmatrix}
\begin{bmatrix}
l'_1 \\
l'_2 \\
l'_3
\end{bmatrix}
\begin{bmatrix}
n_p
\end{bmatrix}
$$

$$
i = k \rho D L' n_p,
$$

Similar to the conventional photometric stereo method, the matrices $L'$ and $D$ are inverted as follows

$$
k \rho n_p = D^{-1} L'^{-1} i.
$$

By solving Equation (19) for $n_p$ at all $(x,y)$, we can estimate the surface normals under the assumption of point light sources. If pixel values with more than three illumination direction are given, we can obtain the normal by calculating the pseudo-inverse matrix of $L'$ as the same way in the conventional photometric stereo.

The shape can be recovered by integrating the estimated normal [10]. Figures 9 (a) and (b) show the recovered shape and the estimated normal map by the proposed photometric stereo method.

**Figure 9. (a) Recovered shape and (b) estimated normal by the proposed photometric stereo method.**

**Figure 10. (a) Improved shape and (b) improved normal by the hybrid algorithm.**

**Hybrid Algorithm**

Finally, we improve the accuracy of the shape and the normal by the Nehab's algorithm [6]. The improved normal is produced by combining two kinds of normal maps, the normal map $n_s$ estimated using structured light technique and the normal map $n_p$ estimated by photometric stereo method. First in the algorithm, it is necessary to produce the smoothed normal maps $\tilde{n}_s$ and $\tilde{n}_p$ by blurring the original normal maps $n_s$ and $n_p$. From the normal maps $n_s$ and $\tilde{n}_p$, the rotation matrix $R$ from $\tilde{n}_p$ to $n_p$ can be calculated. Given the rotation matrix $R$, the transformation $\tilde{n}_p$ to $n_p$ is shown by the following equation

$$
n_p = R \tilde{n}_p.
$$

The improved normal $n$ can be obtained by applying the rotate matrix $R$ to the blurred normal $\tilde{n}_s$ as follows,

$$
n = R \tilde{n}_s.
$$

After the normal is improved, the improved shape can be obtained by deforming the shape recovered using structured light technique. The deformation is performed to obtain the improved normal map from the improved shape by the non-linear least square optimization.

Figures 10 (a) and (b) show the improved shape and the normal map of the object. Thus, we obtain the accurate 3D shape and surface normal of the object.

**Evaluation**

We evaluated the proposed method and system by two kinds of measurements. In the measurements, we used a sample object shown in Fig. 11. For the evaluation of the proposed photometric method, we measured the
surface normal in the following three kinds of conditions.

(i) Taking images of the object illuminated with eight directions from the distance of more than 1.5 meter (can assume parallel light sources), and estimating the normal map by conventional photometric stereo method to obtain ground truth.

(ii) Taking images with the proposed measurement system, and estimating the normal map by the conventional photometric stereo method.

(iii) Taking images with the proposed measurement system, and estimating the normal map by the proposed photometric stereo method using the recovered shape.

Figures 12 (a), (b) and (c) show the normal maps estimated with the above condition of (i), (ii) and (iii). Using the estimated normal maps, the angular errors of each estimated normals to the ground truth were calculated as shown in Figs. 13 (a) and (b).

From the results, the accuracy of conventional photometric stereo method became lower with leaving the center of the object. This was caused by not considering the variation of the incident light directions and the intensity of the light sources. On the other hand, the proposed photometric stereo method could estimate the surface normal accurately.

Next, we evaluated the accuracy of the recovered shape and the estimated normal using the proposed measurement system by comparing the shape and the normal obtained by a commercial 3D scanner (Konica Minolta VIVID 910). The shape and the normal of the object were measured by the proposed measurement system and the commercial 3D scanner. The shape and normal obtained by the proposed measurement system are shown in Figs. 14 (a) and (c), and the shape and normal obtained by the commercial scanner are shown in Fig. 14 (b) and (d). Using these results, we calculated the error maps with the shape and the normal as shown in Figs. 15 (a) and (b).

From the Figs. 15, the proposed system could recover the shape of the object as accurately as the commercial 3D scanner. However, the accuracy of the estimated normal was low. We attribute this to the noise of the normal the commercial scanner.

However, the accuracy of the estimated normal was low. We attribute this to the noise of the normal estimated by the commercial 3D scanner because the normal was estimated from the shape recovered by the commercial scanner.

From the both evaluations, we confirmed the effectiveness of the proposed photometric stereo method and the proposed measurement system.
Conclusion

We introduced the compact, low-cost and accurate 3D measurement system. To reduce the sizes of the measurement system, we proposed new photometric stereo method with point light sources using a recovered shape. The effectiveness of the proposed photometric stereo method was evaluated by measuring surface normal of an object. The accuracy of the shape and the normal with the proposed measurement system is also confirmed by comparing to the ones obtained by a commercial 3D scanner.

The cost of a commercial scanner is generally more than ten thousand dollars. On the other hand, since the proposed system requires a digital camera, a projector and LED light sources, the total cost can be reduced to approximately three thousand dollars.

For more accurate measurement, it is required to measure the luminous intensity distribution of LED light sources [11]. We also have to consider the lens distortion. Finally, we would like to apply our 3D scanning system to digital archiving systems.

Acknowledgements

This research is partly supported by JSPS Grants-in-Aid for Scientific Research (16760031).

References


