Calibration of a Structured Light System by Observing Planar Object from Unknown Viewpoints

Koichiro Yamauchi, Hideo Saito and Yukio Sato
Keio University, Yokohama 223-8522, Japan
{yamauchi, saito}@ozawa.ics.keio.ac.jp, satoy@ics.keio.ac.jp

Abstract—A calibration method for a structured light system by observing a planar object from unknown viewpoints is proposed. A structured light system captures a 3D shape by a camera that observes a light stripe on an object illuminated by a projector. The 3D shape, obtained from the system defined by a pinhole model for the projection of a light stripe, is solved using the equation of a plane model for the projector. The coefficients of each light stripe's equation are estimated using the 4×3 image-to-camera transformation matrix that is expressed by camera parameters. Experimental results demonstrate a high degree of accuracy when following the proposed approach.

I. INTRODUCTION

Studies in computer vision on 3D recovery using a light stripe have been conducted [1]. Study results from this area may advance with the development of object modeling and recognition. For example, creation of custom-made clothing using the 3D shape of a human body has already been realized in the apparel business [2].

A structured light system based on a camera-projector pair allows the 3D reconstruction of an object. We can obtain highly accurate results by adding an appropriate geometric model to the system; however, in many cases, these models impose constraints on the projector of the structured light system.

Typically, we can consider the 3D shape from a structured light system defined by the pinhole model to approximate both the camera and the projector. The camera is modeled by the 3×4 projection matrix, and the projector by the 2×4 projection matrix [3]. It is assumed that all the light stripes are emitted from the optical center of the projector.

Second, the baseline model, which is defined by the distance between the camera and the projector, is proposed [4]. If the baseline is not defined as a variable parameter, the projector must assume the pinhole model. In addition, the light stripe must be vertical to the baseline.

Finally, the equation of a plane model is proposed [5]. In this model, some improvements are made in the projector, which can be defined by the equation of a plane. Even if light stripes are emitted in different directions, this model can represent them more accurately than the above models.

The coefficients of the equation of a plane are estimated using calibration rigs such as cubes, turntables, and slide stages. Although these rigs can be placed in the correct position, they are cubic or large-sized objects [6]. Therefore, simple calibration rigs are required to simplify a user’s tasks.

In this paper, we propose a calibration method for a structured light system by observing a planar object from unknown viewpoints. Fig. 1 shows a calibration scene of the structured light system. The coefficients of each light stripe’s equation are estimated using the 4×3 image-to-camera transformation matrix that is expressed by camera parameters. Our method provides a high degree of accuracy when compared to other conventional methods.

II. GEOMETRIC MODEL

A structured light system is composed of a camera and a projector. This system allows 3D reconstruction when the camera observes a light stripe on a target object illuminated by the projector. Fig. 2 is the geometric model of the structured light system. Both the camera and the projector are represented in the camera coordinate system. The camera model is based on the pinhole model of perspective projection. The projector model is based on the equation of the plane model.

A. Camera Model

The pinhole model of perspective projection is defined by intrinsic and extrinsic parameters. The projection from a 3D point $M_w = [x_w, y_w, z_w]$ in the world coordinate system $(O_w-X_w-Y_w-Z_w)$ to a 2D image point $m = [u, v]$ in the image plane is given by the following equation:

$$\tilde{m} \simeq A \begin{bmatrix} R \ t \end{bmatrix} M_w$$

where

$$A = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

In this formulation, the tilde in $\tilde{m}$ and $M_w$ indicates homogeneous coordinates. The matrix $A$ is the camera calibration matrix.
matrix, which includes the focal length, the image center
c\in [u_0, v_0], the skew, and the aspect ratio. The rotation
matrix R and the translation vector t, which translate to a 3D
point \( \mathbf{M}_c = [x_c, y_c, z_c] \) in the camera coordinate system \((O_x, X_c, Y_c, Z_c)\), encapsulate the camera orientation and position
[7].

In addition, we consider the discrepancy between the real
image coordinates \( \mathbf{m} = [\bar{u}, \bar{v}] \) and the corresponding ideal
good image coordinates \([u, v]\) of perspective projection.

\[
\begin{align*}
\bar{u} &= u + (u - u_0)\left[k_1(x^2 + y^2) + k_2(x^2 + y^2)^2\right] \\
\bar{v} &= v + (v - v_0)\left[k_1(x^2 + y^2) + k_2(x^2 + y^2)^2\right]
\end{align*}
\]

where \( k_1, k_2 \) are the radial distortion coefficients, and \([x, y]\)
are the normalized image coordinates. The center of radial
distortion is the same as the principal point [8].

B. Projector Model

Let us consider the case in which the light stripes are
emitted in different directions. It is difficult to assume that the
projector model is based on the pinhole model of perspective
projection, because the light stripe is not directly illuminated
from the optical center of the projector. Therefore, we use the
equation of a plane model to accurately represent the
projector instead of considering the projection of the light
stripe from the optical center of the projector. Therefore, we use
the projector model is based on the pinhole model of perspective
projection. From (1) and (4), we derive the linear equation
\([x_c/z_c, y_c/z_c, 1/z_c]\)
as follows:

\[
\begin{bmatrix}
\alpha & \gamma & 0 \\
0 & \beta & 0 \\
\alpha_i & b_i & d_i
\end{bmatrix}
\begin{bmatrix}
x_c/z_c \\
y_c/z_c \\
1/z_c
\end{bmatrix}
= \begin{bmatrix}
u - u_0 \\
v - v_0 \\
-c_i
\end{bmatrix}
\]

Therefore, the camera coordinates \([x_c, y_c, z_c]\) are expressed as

\[
\begin{align*}
x_c &= \frac{(u - u_0) - \frac{\gamma}{\alpha}(v - v_0)}{z_c} \\
y_c &= \frac{v - v_0}{\beta} \\
z_c &= \frac{-c_i}{\alpha_i} - \frac{(u - u_0) - \frac{\gamma}{\alpha}(v - v_0) - \frac{b_i(v - v_0)}{\alpha}}{\frac{c_i}{\beta}}
\end{align*}
\]

The coordinate \( z_c \) is computed by the triangulation principle
using one side and two angles of a triangle. Then, the
cordinate \( x_c \) and \( y_c \) are calculated based on the scaling
relation of the camera.

III. CALIBRATION METHOD

We present a calibration method for the structured light
system by observing the reference plane from unknown
viewpoints. Fig. 1 is a calibration scene of a structured light
system. The reference plane contains a checkered pattern so
that the calibration points can be detected as the intersection
of straight lines. Our approach to calibrating a structured
light system incorporates two separate stages: camera cali-
bation and projector calibration.

A. Camera Calibration

In the camera calibration stage, camera parameters are
obtained by Zhang’s method [9]. First, the camera calibration
matrix A is estimated from perspective projections and
homographies between the image and the world coordinates.
Then, the rotation matrix R, translation vector t, and radial
distortion coefficients \( k_1, k_2 \) are computed. Finally, the
camera parameters are optimized with a nonlinear refinement
based on the maximum likelihood criterion.

B. Projector Calibration

In the projector calibration stage, we estimate the equation
of a plane for the light stripe using the \( 4 \times 3 \) image-to-camera
transformation matrix. Although, in the work of Huynh et
al. [10] this matrix is solved by world point to image point.
correspondences, it is represented by the camera parameters as follows:

\[
\tilde{M}_c = \begin{bmatrix} M_c & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} Q & \mathbf{k}^T \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ 1 \end{bmatrix} \\
\begin{bmatrix} Q & \mathbf{k}^T \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} \mathbf{I} \\ \mathbf{r}_i^T \mathbf{t}^{-1} \mathbf{r}_i^T \end{bmatrix} \mathbf{A}^{-1} \tilde{m} \approx \begin{bmatrix} Q & \mathbf{k}^T \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} \mathbf{I} \\ \mathbf{r}_i^T \mathbf{t}^{-1} \mathbf{r}_i^T \end{bmatrix} \mathbf{A}^{-1} \tilde{m} \tag{12}
\]

where \( \mathbf{Q} = \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix} \mathbf{k} = [0, 0, 1] \)

Here, \( \mathbf{r}_i \) is the \( i \)th column of the rotation matrix \( \mathbf{R} \). Once the camera parameters are obtained by Zhang’s method, the orientation and position of the reference plane are determined. Unlike other methods that necessitate recalculation, here, we use the \( 4 \times 3 \) image-to-camera transformation matrix that can be computed directly. This matrix has eight degrees of freedom, which are the same as homography in 2D space.

The \( i \)th light stripe is illuminated on the reference plane as shown in Fig. 3 so that a segment of the line is projected onto the image plane. Equation (12) allows the \( i \)th light stripe image-to-camera transformation. Therefore, the coefficients of the equation of the \( i \)th plane (\( a_i, b_i, c_i, \) and \( d_i \)) can be computed by the least squares method with at least three image coordinates. This is how all the light stripes are estimated.

IV. RESULTS

A. Calibration

The "Handy 3D Camera Cartesia", which is a structured light system of SPACEVISION Incorporated [11], has been calibrated. This system consists of a camera, the focal length and resolution of which are 8 mm and 640×480 pixels, respectively, and a projector, the number of light stripes of which is 254. Three surface images and light stripe images (a luminance value corresponds to the light stripe number) are captured by observing the reference plane from three viewpoints, as shown in Fig. 4. The light stripe images are obtained when the projector emits a structured light pattern.

Table 1 describes the intrinsic camera parameters estimated by Zhang’s method. Fig. 5 is the calibration result of the projector parameters, which include the baseline, projection angle, and tilt angle, instead of the equation of a plane. When the number of light stripes increases, the baseline gradually reduces, the projection angle increases, and the tilt angle remains almost constant. The camera and projector parameters allow the 3D reconstruction of an object by the triangulation principle.

B. Evaluation

We have evaluated the accuracy of proposed calibration by comparing with other methods as follows:

(i) The pinhole model calibrated with a slide stage: The camera is modeled by the \( 3 \times 4 \) projection matrix, and the projector is modeled by the \( 2 \times 4 \) projection matrix [12]. The parameters are estimated using a slide stage.

(ii) The equation of a plane model calibrated with a slide stage: The camera model is based on the pinhole model. The projector model is the equation of a plane model. The parameters are estimated using a slide stage.

(iii) The equation of a plane model calibrated with the reference plane (the proposed method): The camera model is based on the pinhole model. The projector model is the equation of a plane model. The parameters are estimated using the reference plane.

Evaluations of the above three techniques are performed using five equal spheres of 25 mm radius, which are placed in front of the structured light system. Fig. 6 shows the measurement results of these spheres, according to which, (i) appears to be externally distorted when compared to (ii) and (iii).
The sphericity error is formulated by
\[
E = \frac{1}{S} \sum_{p=1}^{S} (r_p - \hat{r})^2
\]  
(13)
where \( S \) is the number of measurement points, \( r_p \) is the actual measurement value, and \( \hat{r} \) is the theoretical value. \( r_p \) is the distance between a measurement point and a center point that is computed by fitting the ideal sphere to all of the measurement points. Table 2 illustrates the evaluation result of the sphericity error. Thus, the equation of a plane model is more appropriate for the structured light system than the pinhole model. From (ii) and (iii), we infer that there are minor differences between using the slide stage and the reference plane.

Therefore, it has proven that our approach to calibrate the system, defined that the projector model is using the equations of a plane model, achieves high accuracy measurements. The calibration using a planar object obtains similar results to the traditional method using a slide stage.

V. CONCLUSION

In this paper, a calibration method for a structured light system by observing a planar object from unknown viewpoints was presented. A structured light system captures a 3D shape by a camera that observes a light stripe on an object illuminated by a projector. The projector model is based on the equation of a plane model. We proposed an estimation approach for the coefficients of the equation of a plane is based on the 4×3 image-to-camera transformation matrix which can be computed directly from the camera parameters. Furthermore, we verified our method, which uses a simple planar object, provides a high degree of accuracy in the experiment.

REFERENCES