

Accurate Calibration of Intrinsic Camera Parameters by Observing Parallel Light Pairs

Ryusuke Sagawa and Yasushi Yagi

Abstract—This study describes a method of estimating the intrinsic parameters of a perspective camera. In previous calibration methods for perspective cameras, the intrinsic and extrinsic parameters are estimated simultaneously during calibration. Thus, the intrinsic parameters depend on the estimation of the extrinsic parameters, which is inconsistent with the fact that intrinsic parameters are independent of extrinsic ones. Moreover, in a situation where the extrinsic parameters are not used, only the intrinsic parameters need to be estimated. In this case, an intrinsic parameter, such as focal length, is not sufficiently robust to combat the image processing noise, that is absorbed by both parameter types, during calibration. We therefore propose a new method that allows the estimation of intrinsic parameters without estimating the extrinsic parameters. In order to calibrate the intrinsic parameters, the proposed method observes parallel light pairs that are projected on different points. This is accomplished by applying the constraint that the relative angle of two parallel rays is constant irrespective of where the rays are projected. This method focuses only on intrinsic parameters and the calibrations are sufficiently robust as demonstrated in this study. Moreover, our method can visualize the error of the calibrated result and the degeneracy of the input data.

I. INTRODUCTION

The intrinsic parameters (such as focal length and principal point) of a perspective camera are critical for geometrical analysis. Various methods have been proposed to calibrate the intrinsic and extrinsic parameters. However, in some situations, only the intrinsic parameters need to be estimated as the extrinsic parameters are not used. For example, if a moving camera, such as that mounted on a robot, is used, the intrinsic parameters can be calibrated *a priori*. In contrast, the extrinsic parameters change during motion and cannot be estimated in advance.

Previous calibration methods calculate both the intrinsic and extrinsic parameters that minimize the error between the observed image points and the points computed by projecting the scene points onto the image. Since both parameters are necessary to compute the projection of a scene point onto an image point, the extrinsic parameters must be estimated simultaneously even if they are not used. The intrinsic parameters therefore depend on the estimation of the extrinsic parameters. This is inconsistent with the fact that intrinsic parameters are independent of extrinsic parameters.

Clearly a minimization of the error of the estimated intrinsic parameters is desirable. However, the results depend heavily on the images which include the calibration objects.

In our preliminary experiment, the calibration results were worse than expected even when the images were carefully processed. This can be attributed to the fact that previous methods require both parameters to be estimated in order to minimize the error of the intrinsic parameters. In this paper, we show that the calibration result of intrinsic parameters depends on the extrinsic parameters.

To calibrate camera parameters, a marker object is often used for easy detection of feature points. In [1], a planar object is used as a marker with a known translation. This is equivalent to using a 3D marker object. The camera parameters are computed by minimizing the reprojection error; for details, refer to [2], [3]. In [4], [5], planar objects are also used, but the relationship, i.e., rotation and translation, between them is unknown. These methods first compute a homography between a planar object and the image. The intrinsic parameters are estimated based on the constraints between the intrinsic parameters and the homography. A nonlinear minimization of the reprojection error is applied to improve the solution obtained by minimizing an algebraic distance in [5]. If a marker object is used at a finite distance to estimate the camera parameters, the solution is optimized within the range of the marker. Hence, the error of projection is large if an object is far from the range. The calibration results of the intrinsic parameters are thus affected by the setting of markers.

In [6], [7], the camera parameters are estimated using pure rotation. In this case, the translation from a camera to a marker object is omitted from the cost function. Thus, the number of extrinsic parameters is reduced and the calibration result can be independent of the marker range. However, the rotation still needs to be included in the cost function of the estimation.

The method proposed in this paper uses pairs of parallel lights. A parallel light is a bunch of parallel rays, such as sunlight, and can be obtained by observing a distant object or a collimator. Some previous approaches used multiple collimators to calibrate intrinsic parameters [8], [9], [10], [11], [12]. Only two parallel lights in each image are, however, necessary in the proposed method, thus allowing it to be applied more easily than these previous methods. Other methods used stars as distant objects [13], [14] and estimated both the intrinsic and extrinsic parameters using parallel lights. Instead of using a distant object, yet other methods used vanishing points for calibration. If the relative angle of vanishing points is 90 degrees, the intrinsic parameters are obtained from a linear solution [15], [16], which is a special case of our proposed method. In [17], [18], the intrinsic and

The authors are with the Institute of Scientific and Industrial Research, Osaka University, 8-1 Mihogaoka, Ibaraki-shi, Osaka, 567-0047, Japan {sagawa, yagi}@am.sanken.osaka-u.ac.jp

extrinsic parameters are computed using vanishing points in a similar manner to [6].

We have already proposed methods for calibrating catadioptric systems using parallel lights [19], [20]. In this paper, we apply the technique to the calibration of a perspective camera. The proposed method estimates the intrinsic parameters without computing any extrinsic parameters to improve the accuracy of calibration. Related methods are described briefly in Section II and the new method of camera calibration is proposed in Section III. Based on the cost function of this new method, a visualization of the calibration error is provided. In Section IV, the sensitivity of the estimated intrinsic parameters is analyzed according to the cost function to be minimized. Finally, we summarize our work in Section V.

II. RELATED METHODS OF CALIBRATING INTRINSIC CAMERA PARAMETERS

In this section, related methods used to calibrate camera parameters are briefly explained. First, we define the intrinsic parameter K .

$$K = \begin{pmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix}, \quad (1)$$

where f_x and f_y are the focal lengths. If the aspect ratio is not 1 then f_x and f_y are different. The principal point of the projection is (c_x, c_y) . In the discussion that follows, we assume that the skewness s is 0, because it can be ignored for cameras based on recent technology.

Next, the extrinsic parameters T are represented as follows:

$$T = [R \mid \mathbf{t}], \quad (2)$$

where R is a 3×3 rotation matrix and \mathbf{t} is a 3×1 translation vector. Thus, the projection from a 3D point $\mathbf{M} = (X, Y, Z)^T$ to an image point $\mathbf{m} = (x, y)^T$ is computed as

$$s\hat{\mathbf{m}} = KT\hat{\mathbf{M}}, \quad (3)$$

where $\hat{\mathbf{m}} = (x, y, 1)^T$ and $\hat{\mathbf{M}} = (X, Y, Z, 1)^T$ are the vectors of homogeneous coordinates of \mathbf{m} and \mathbf{M} , respectively. s is a scaling factor.

A. Calibration by Observing a 3D Object

If a 3D object is used as a marker for calibration, the parameters are obtained by minimizing the cost function

$$E_{3D}(P) = \sum_i \left\| \frac{1}{s_i} P \hat{\mathbf{M}}_i - \hat{\mathbf{m}}_i \right\|^2, \quad (4)$$

where $P = KT$ and $\hat{\mathbf{M}}_i$ and $\hat{\mathbf{m}}_i$ are the corresponding points given by image processing. The intrinsic parameters K and extrinsic parameters T are computed by decomposing P .

Although an optimal P is obtained by minimizing $E_{3D}(P)$, the intrinsic parameter K may not be optimal after decomposition because the cost function does not minimize the error of K . Moreover, since the parameters are optimized

for the marker, the error of projection will be large when an object is at a different position to the marker.

B. Calibration by Observing a Plane using Translation and Rotation

This next method calibrates the parameters by multiple observations of a plane [4], [5] that translates and rotates relative to a camera. The method first computes homographies that represent the correspondence between points on the plane and image points. If the plane is on $Z = 0$, the homography H is represented by

$$H = [\mathbf{h}_1 \ \mathbf{h}_2 \ \mathbf{h}_3] = K[\mathbf{r}_1 \ \mathbf{r}_2 \ \mathbf{t}], \quad (5)$$

where \mathbf{h}_k and \mathbf{r}_k are the k th columns of H and R , respectively. Since \mathbf{r}_1 and \mathbf{r}_2 are orthonormal, the following constraints are applicable to each plane:

$$\mathbf{h}_1^T K^{-T} K^{-1} \mathbf{h}_2 = 0, \quad \mathbf{h}_1^T K^{-T} K^{-1} \mathbf{h}_1 = \mathbf{h}_2^T K^{-T} K^{-1} \mathbf{h}_2, \quad (6)$$

where $K^{-T} = (K^T)^{-1}$. $K^{-T} K^{-1}$ is estimated by a closed-form solution and K is obtained by decomposing $K^{-T} K^{-1}$. K is refined in [5] by a nonlinear minimization of

$$E_{\text{plane}}(K, T_1, \dots, T_n) = \sum_{i=1}^N \sum_{j=1}^n \left\| \frac{1}{s_{ij}} K T_j \hat{\mathbf{M}}_i - \hat{\mathbf{m}}_{ij} \right\|^2, \quad (7)$$

where n is the number of planes and N is the number of points on an image.

As in the case of the 3D marker, the intrinsic parameter K , obtained by minimizing Eq. (7), may not be the best one possible and the result is optimized for the position of the planes.

C. Calibration by Observing an Object using only Rotation

The third method utilizes several images of an object obtained by only rotating the camera. The translation from the camera to a marker is therefore omitted from the projection. The rotation between the j th and k th images is computed by $R_{jk} = K^{-1} H_{jk} K$, where H_{jk} is the homography that represents the correspondence. Since $R_{jk} = R_{jk}^{-T}$, the constraint given below is applied to estimate the intrinsic parameter K in [6].

$$(K K^T) H_{jk}^{-T} = H_{jk} (K K^T). \quad (8)$$

K is obtained by decomposing $K K^T$ and refined by a nonlinear minimization of

$$E_{\text{rot}}(K, R_1, \dots, R_n) = \sum_{i=1}^N \sum_{j=1}^n \left\| \frac{1}{s_{ij}} K R_j \hat{\mathbf{M}}_i - \hat{\mathbf{m}}_{ij} \right\|^2. \quad (9)$$

In [7], R_j is assumed known as a turntable is used. To obtain a pure rotation, a distant object is used in [6] and the rotation axis is precisely aligned to the camera in [7].

Translation is removed from the parameterization in this approach. Thus, the result is independent of the position of the marker, but the rotation must be considered in all circumstances.

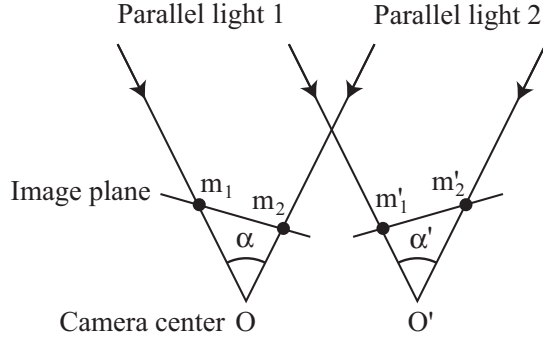


Fig. 1. Two parallel lights are observed from two positions, O and O' . Although the cameras are relatively translated and rotated, the relative angles of the two parallel lights are the same, i.e., $\alpha = \alpha'$.

III. CALIBRATING INTRINSIC CAMERA PARAMETERS BY OBSERVING PARALLEL LIGHT PAIRS

The aim of this paper is to compute an optimal K by removing the extrinsic parameters from the cost function. To obtain a cost function whose sole parameter is K , we use a constraint similar to Eqs. (6) and (8). Although homographies, that include the extrinsic parameters, must be computed to obtain the constraints, this is not necessary in our method as a constraint is obtained directly by observing the parallel light pairs.

A. A Cost Function Obtained from Parallel Light Pairs

Eqs. (6) and (8) are based on the constraint that both distance and angle are invariant to rotation. Observation of a pair of parallel lights allows us to use this constraint directly. Figure 1 shows a situation in which two parallel lights are observed from two different positions. Although the cameras are relatively translated and rotated, the relative angles of the two parallel lights are the same, i.e., $\alpha = \alpha'$. If \mathbf{m} is the projected point of a parallel light, the ray vector is computed by $K^{-1}\mathbf{m}$. The relative angle α is computed as follows [3]:

$$\cos \alpha = \frac{\mathbf{m}_1^T K^{-T} K^{-1} \mathbf{m}_2}{\|K^{-1} \mathbf{m}_1\| \|K^{-1} \mathbf{m}_2\|}. \quad (10)$$

Therefore, if α is known, the cost function becomes

$$E_{\text{para1}}(K) = \sum_{i=1}^N d_{1i}^2$$

$$d_{1i} = \frac{\mathbf{m}_{i1}^T K^{-T} K^{-1} \mathbf{m}_{i2}}{\|K^{-1} \mathbf{m}_{i1}\| \|K^{-1} \mathbf{m}_{i2}\|} - \cos \alpha \quad (11)$$

where N is the number of light pairs. Since parallel lights are used, the translation is omitted from the function. The rotation is also omitted as the rotation constraint is used directly. In addition, in the case that α is unknown, the cost

function becomes

$$E_{\text{para2}}(K) = \sum_{i=1}^N d_{2i}^2$$

$$d_{2i} = \frac{\mathbf{m}_{i1}^T K^{-T} K^{-1} \mathbf{m}_{i2}}{\|K^{-1} \mathbf{m}_{i1}\| \|K^{-1} \mathbf{m}_{i2}\|} - \frac{\mathbf{m}'_{i1}^T K^{-T} K^{-1} \mathbf{m}'_{i2}}{\|K^{-1} \mathbf{m}'_{i1}\| \|K^{-1} \mathbf{m}'_{i2}\|} \quad (12)$$

K is estimated by nonlinear minimization of these cost functions.

B. Initial Estimation of Parameters

Since a nonlinear minimization is necessary to estimate K from E_{para1} or E_{para2} , an initial estimation is required. If we have four or more points on each image, an initial estimation of K can be computed by Eq. (8). However, we propose a method that assumes simplified intrinsic parameters.

For the initial estimation, the aspect ratio is assumed to be 1 and the principal point is the center of the image, i.e., $f_x = f_y$ and $(c_x, c_y) = (w/2, h/2)$, where w and h are the width and height of the image, respectively. Thus, the only remaining parameter is f_x and $E_{\text{para1}} = 0$ is simplified as follows:

$$\sum_{i=1}^N A_i f_x^4 + \sum_{i=1}^N B_i f_x^2 + \sum_{i=1}^N C_i = 0, \quad (13)$$

where A_i , B_i , and C_i are computed from \mathbf{m}_{i1} , \mathbf{m}_{i2} , and α . By solving Eq. (13), an initial f_x is obtained. In the case of E_{para2} , f_x can be computed by solving a similar equation.

C. Visualization of Error and Degeneracy

The parameters of E_{para1} and E_{para2} are four variables, f_x , f_y , c_x and c_y . Hence, $d_{1i} = 0$ and $d_{2i} = 0$ are surfaces in a four-dimensional space. Since the surface is computed for each pair of input points, the surfaces intersect at the solution given by minimizing E_{para1} or E_{para2} .

To visualize the surfaces, we show 2D graphs that are slices of the 4D surfaces at the solution. Each surface becomes a curve in a 2D slice and, if the curves do not intersect at a single point, this indicates that the input pairs contain noise. The error of the estimated solution is visualized as the distance of the solution and the curves. If the curves are nearly parallel around the intersecting point, the intersecting point becomes sensitive to the noise of the input points. Therefore, the input points are almost a degenerated configuration.

For example, consider the situation in which $f_x = f_y = 900$, $c_x = c_y = 255$ and the image size is 512×512 pixels. The sample input points are shown in Figure 2(a). There are six pairs and each pair is represented by the same color. Figure 2(b) shows the curves that are 2D slices of the 4D surfaces. The colors of the curves correspond to the input pairs in Figure 2(a). All the curves intersect at the solution, i.e., $f_x = 900$ and $c_x = 255$ in Figure 2(b). However, some curves are parallel at the solution. For example, the magenta and black lines are almost the same. Thus, their input pairs are almost degenerate.

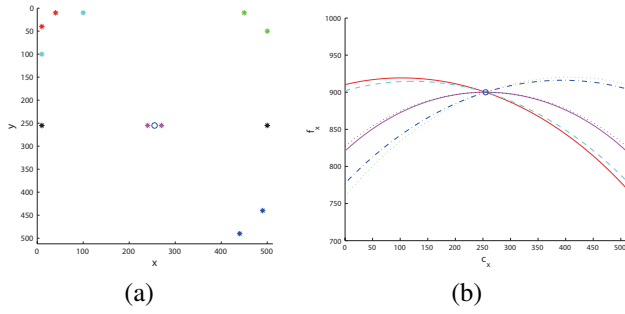


Fig. 2. (a) The sample input points are indicated by '*'. The principal point is depicted by an 'o'. (b) Each curve indicates a slice of the 4D surface at $f_x = 900$ and $c_x = 255$.

By observing the curves, the corner pairs, e.g. the red, green and blue pairs, were found to be effective in computing an accurate solution. Since an equation is obtained for each pair of points, four equations are sufficient to determine a solution. Therefore, it is reasonable to acquire the input pairs at the four corners. While it is easy to acquire point pairs at the corners of an image, it is difficult in practice to observe a 3D or plane marker at the edge of an image.

D. Acquisition of a Parallel Light

Our proposed method requires observation of a parallel light. A parallel light can be viewed by adopting one of the following two approaches:

- Using the feature point of a distant marker
- Generating a collimated light

In the former approach, small translation of the camera motion can be ignored because it is much smaller than the distance to the marker. Thus, the ray vector from the feature point is invariant even if the camera moves. The main issue of this approach is a lens focus problem. When the focus setting of the camera is not at infinite focus, the image is acquired with a minimum aperture and long shutter time to avoid blurring.

In the latter approach, a parallel light is generated by a collimator. A simple method is to use a concave parabolic mirror and a point-light source. This is illustrated in Figure 3. By placing pinholes in front of the light sources, they become point-light sources. Since pinholes are placed at the focus of the parabolic mirrors, the reflected rays are parallel. The advantage of this approach is that a small system can be constructed, although optical apparatus is necessary.

E. Measurement of Relative Angle

It is necessary to measure the relative angle of the two parallel lights for the cost function E_{para1} . If distant markers are used as the source of parallel lights, we can measure the relative angle by using 1) a map, 2) a camera on a turntable, or 3) a theodolite. The first method, measuring the relative angle on a map, is the logical choice if corners of a building are used as features for calibration. The error is approximately 0.2 degrees. In the second method, a camera is mounted on a turntable. By rotating the turntable, we can

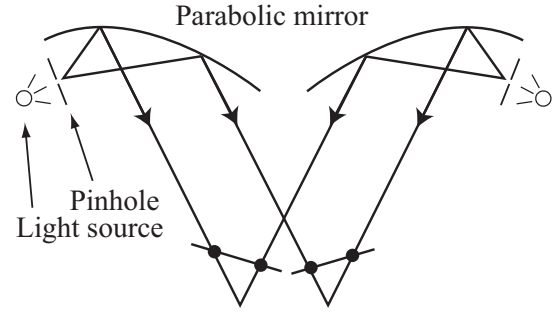


Fig. 3. Collimated lights are generated by two collimators consisting of concave parabolic mirrors, pinholes and light sources.

discover the relative angle at which two feature points are projected on the same point. The error is determined by the resolution of the turntable which is 0.01 degrees in our experiment. In the third method, a device that measures angles, a theodolite, is used. The error when using a theodolite is less than 0.001 degrees. To measure the relative angle of two collimated lights, we used a camera on a turntable in the manner described above.

IV. EXPERIMENTS

In this section, we first compare the robustness of our proposed method with related methods. Simulations are used to evaluate the methods. Then, we calibrate a camera using real images.

A. Evaluating the Robustness of the Proposed Method through Simulation

In the simulation, we compare the accuracy of the intrinsic parameters estimated using E_{3D} , E_{plane} , E_{rot} , E_{para1} , and E_{para2} . The ground truth of the parameters is $f_x = f_y = 900$ and $c_x = c_y = 255$, and the image size is 512×512 pixels.

Input data provided by Zhang [5] is used for the comparison. Since the intrinsic parameters differ from the data provided, we recompute the projected points using the ground truth of the intrinsic parameters and the provided extrinsic parameters. As the data is obtained from three plane markers, we use the provided extrinsic parameter E_{plane} and modify the other extrinsic parameters as follows:

- E_{3D} : the provided three planes are regarded as a 3D object. It is assumed that the relative positions of the three planes are known.
- E_{rot} , E_{para1} , and E_{para2} : only the rotation parameter is used. The points of the latter two images are recomputed by applying the rotation parameter for each image on the points of the first plane.

First, we analyze the error of each cost function when one of the parameters changes from the ground truth. Since it is difficult to compute the derivatives, such as $\partial E / \partial f_x$, analytically, experiments are used for evaluation. Figure 4 and Figure 5 show the error after minimizing each cost function when f_x and c_x are changed, respectively. The other intrinsic

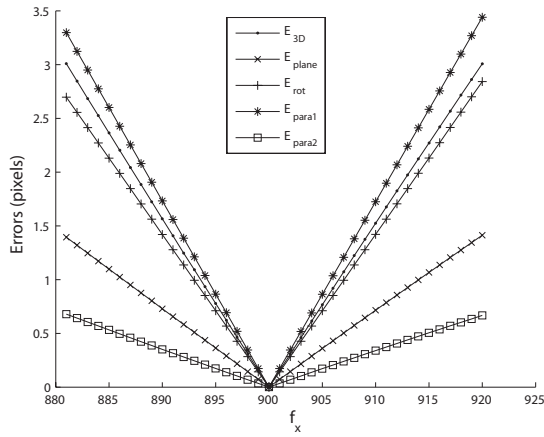


Fig. 4. The errors increase when f_x changes from the ground truth.

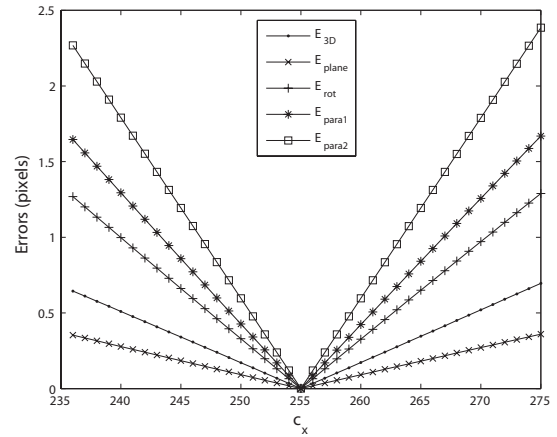


Fig. 5. The errors increase when c_x changes from the ground truth.

parameters are fixed to the ground truth. After minimization of the cost functions, the only remaining parameters are the extrinsic parameters. The errors are calculated by computing $\sqrt{E/N}$, which is the root-mean-square error (RMS). While E_{3D} , E_{plane} , and E_{rot} are computed in pixels, E_{para1} and E_{para2} are computed from the errors of cosine values. Thus, for the sake of comparison, E_{para1} has been modified to be computed in pixels as follows:

$$d_1 = \min_{\mathbf{p}} \|\mathbf{m}_2 - \mathbf{p}\| \text{ subject to } \cos \alpha = \frac{\mathbf{m}_1^T \mathbf{K}^{-T} \mathbf{K}^{-1} \mathbf{p}}{\|\mathbf{K}^{-1} \mathbf{m}_1\| \|\mathbf{K}^{-1} \mathbf{p}\|} \quad (14)$$

where \mathbf{p} is a point on the image. E_{para2} is modified in the same way.

The results of Figure 4 and Figure 5 indicate the sensitivity of the cost function to changes in the parameters. If the error increases quickly when a parameter changes from the ground truth, the local minimum is expected to be determined robustly. In other words, this indicates robustness towards the noise of image processing. For example, if E_{plane} is used and the error of the feature position is 1 pixel, f_x may change 15 pixels, although f_x may change only 5 pixels if E_{para1} is used. This indicates that E_{para1} is a robust criterion with respect to both f_x and c_x . In contrast, E_{para2} is not as robust with respect to f_x . The reason is that the relative angle difference alters only slightly because both relative angles, α and α' , change similarly when f_x is modified. As for other criteria, the errors of E_{3D} , E_{plane} , and E_{rot} are absorbed by the extrinsic parameters and tend to be small.

Next, we compute the intrinsic parameters by adding noise to the position of the input points. The added Gaussian noise has standard deviations of 0, 0.1, 0.5, and 1.0 pixels. As for E_{para1} , since the relative angle α between the two points has to be given, we add noise to α with standard deviations of 0, 0.1, and 0.5 degrees. Figure 6 and Figure 7 show the RMS errors of the estimated f_x and c_x , respectively. The RMS errors of E_{para1} are much smaller than the other criteria even when the noise of α is large. It can be concluded that,

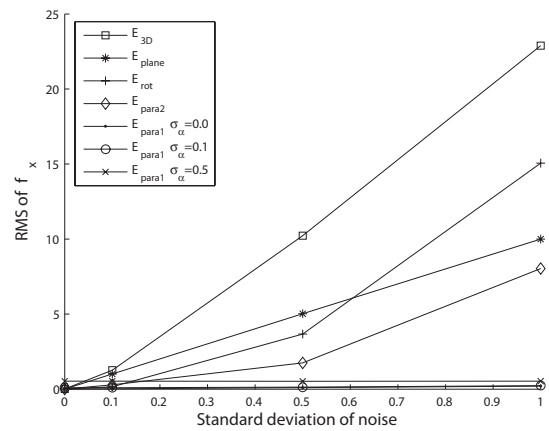


Fig. 6. The RMS error of the estimated f_x : the standard deviations of the Gaussian noise added to the position of the input points are 0, 0.1, 0.5, and 1.0 pixels. For E_{para1} , noise is added to the relative angle α , with standard deviations of 0, 0.1, and 0.5 degrees.

if the relative angle of two parallel lights can be measured to an accuracy of 0.1 degrees, our proposed method is a better solution than previous methods. This result is true in the case where the noise of the input points is small. However, the method continues to provide adequate solutions even if the accuracy is slightly less, e.g. 0.5 degrees. These results illustrate that using a map to measure angles is a viable option.

B. Experiments using Real Images

Next, we calibrate a SONY DCR-VX2000 video camera using four different methods. We show that the proposed method gives good intrinsic parameters that are independent of extrinsic parameters by comparing the results. In this experiment, since the lens distortion has previously been calibrated using [21], no distortion exists in the images.

The first method is the proposed method that uses pairs of parallel lights based on E_{para1} . In this experiment, we use feature points of a distant building as shown in Figure 8. The distance from the camera to the building is about 65m. Thus, the small translation of camera motion can be ignored. The

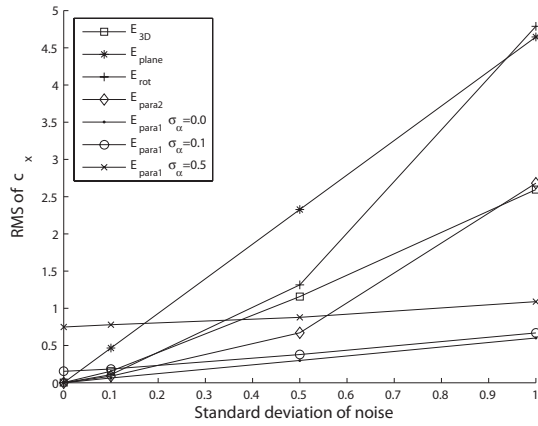


Fig. 7. The RMS error of the estimated c_x : noise, similar to that in Figure 6, is added to the position of input points and the relative angle α .

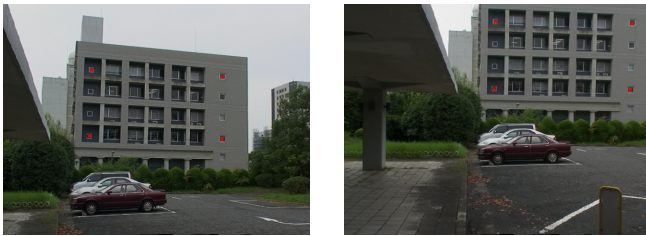


Fig. 8. Examples of the input images: the detected feature points are indicated with 'x'.

focus is fixed to the near range and the images are captured with a minimum aperture and long shutter time. The image size is 640×480 pixels. The feature points are detected as the centers of windows in the building. The relative angles are 8.92 - 21.53 degrees as measured by a theodolite.

We captured 5 images each containing 4 feature points. Therefore, 30 pairs of input points were used as shown in Figure 9. Each pair is represented by the same color and connected by a line. The principal point is depicted by an 'o'. The intrinsic parameters were calibrated using the cost function E_{para1} . Figure 10 and Figure 11 show the curves that correspond to the pairs of parallel lights. The curve computed from each pair is depicted by the corresponding color and line style. The estimated parameters are indicated by an 'o' at the intersection of the curves.

The second method is a calibration with planar objects for which we used an implementation by Bouguet [22]. Four checker patterns are captured as shown in Figure 12 and the corner points are used as input points.

The third and fourth methods are calibrations using a 3D object. An indoor scene is used for the third method and an outdoor scene for the fourth one. The images are shown in Figure 13. In the indoor scene, the input points are detected as the gravity of white circles. The corner detector of OpenCV [23] is used for the outdoor scene. The 3D positions of input points are measured using a theodolite, the error of which is about 2mm. The scale of the scenes is about 3m and 100m for the indoor and outdoor scenes,

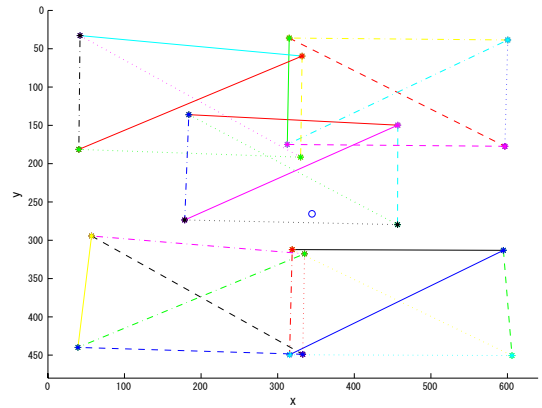


Fig. 9. The input points are indicated by '**'. (Note: The caption text in the image is '**', but the visual marker is 'o').

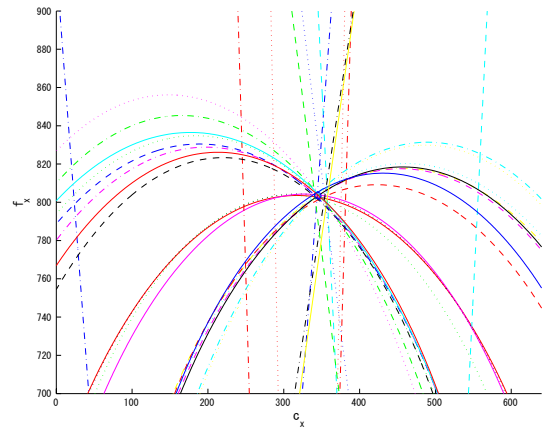


Fig. 10. Each curve is a slice of a 4D surface in the c_x - f_x plane.

respectively.

To evaluate the accuracy of the intrinsic parameters estimated by the four different methods, we computed the reprojection error of the points not used for calibration in the indoor and outdoor scenes. Hereafter, we refer to these as the 'Indoor Test' and 'Outdoor Test', respectively.

In the Indoor Test, we use three images captured from different positions. The camera positions and 3D positions of input points are estimated by a bundle adjustment. The marker on a tripod is not used for calibration or the bundle adjustment in Figure 13(a). The 3D position of the marker is computed as the intersection of three viewing rays using the mid-point method [24]. The reprojection error is computed as the distance between the input point and the point projected from the mid-point onto each image. Figure 14 shows a top view of the situation. Points marked with '+' and '*' denote the estimated 3D positions of the input points and camera positions, respectively. The lines are the viewing rays and the intersection is indicated by an 'o'.

In the Outdoor Test, we first estimate the camera position using the 7 points on the near building shown in Figure 13(b). The 3 points on the far building, which is on the right side in the image, are used for testing. The 3D positions of the

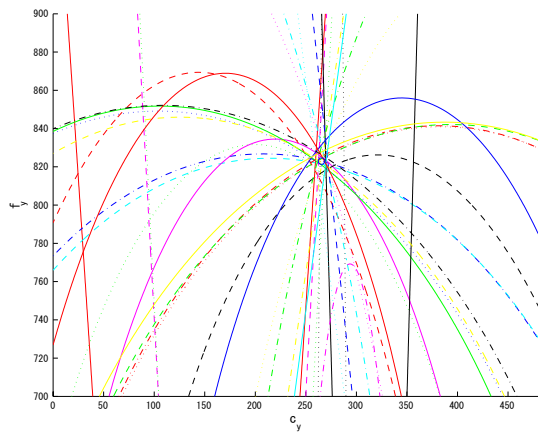


Fig. 11. Each curve is a slice of a 4D surface in the c_y - f_y plane.

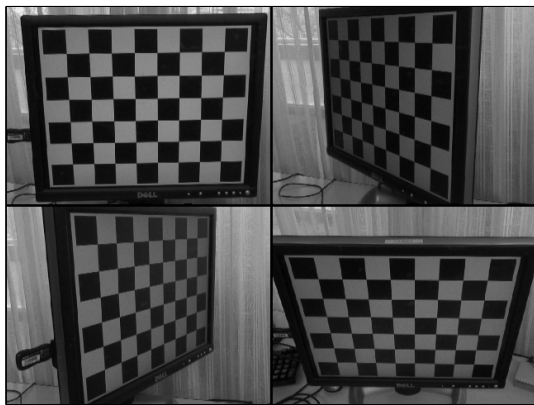


Fig. 12. Four checker patterns captured on a planar display.

input points are given in this test. The reprojection error is computed as the difference between the input points and the projected points on the far building. Three images not used for the calibration, are used for testing.

Table I shows the results of the calibration and tests. The 'Parallel' method is the proposed method and the 'Planes' method is a calibration using planar objects. The 'Indoor' and 'Outdoor' methods refer to calibration with a 3D object. The 'Indoor' and 'Outdoor' methods use the markers in Figure 13(a) and (b), respectively. The calibration errors are the RMS errors computed from Eq. (14) for the proposed method and the reprojection error of the input points for calibration for the other three methods. These errors denote the accuracy of the input points when fitted to the cost models. Since the errors are less than one pixel, the input points are sufficiently accurate. The fitting errors are the RMS reprojection errors of the input points used for estimating the extrinsic parameters. These errors denote the accuracy of the input points when fitted to the projection model. Since the fitting errors are less than one pixel in both the Indoor and Outdoor Tests, the extrinsic parameters have been estimated accurately.

The reprojection errors are the RMS errors of the points that are not used for estimating the extrinsic parameters.

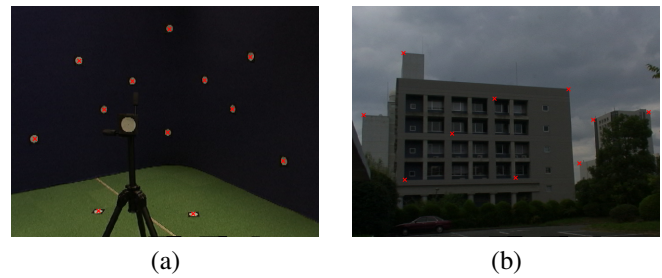


Fig. 13. (a) The indoor scene with 12 input points. (b) The outdoor scene with 10 input points.

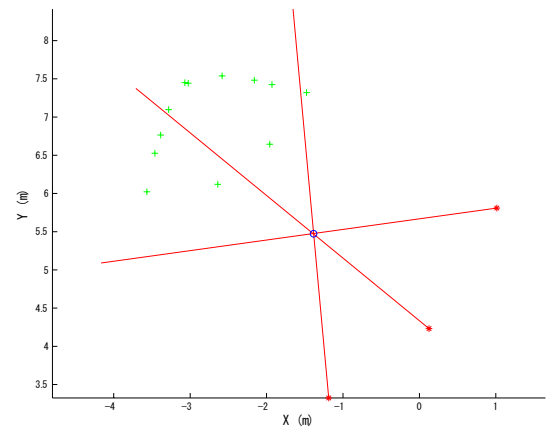


Fig. 14. The 3D position 'o' of the marker on a tripod is computed as the intersection of three viewing rays in the Indoor Test.

If the extrinsic parameters are accurately estimated, the reprojection errors depend on the accuracy of the intrinsic parameters. The errors of the proposed method are the smallest in both tests and between 43%–92% of the other three methods. Furthermore, the errors of the other methods become larger if the test and calibrated environments differ. This indicates that the estimated intrinsic parameters depend on the calibrated environments. The reason the proposed method produces good results in both tests is that it does not need to estimate the extrinsic parameters.

The MATLAB program and data for this experiment are available online [25].

V. CONCLUSION

In this paper, we have proposed a method for calibrating the intrinsic parameters of a perspective camera. Since previous methods minimize the reprojection error computed by projecting a 3D point onto an image, estimating intrinsic parameters cannot be done without estimating the extrinsic parameters, which is inconsistent with the fact that intrinsic parameters are independent of extrinsic parameters. As shown in the experiments, the error becomes larger if the environment differs from the calibrated environment.

The proposed method omits the extrinsic parameters from the cost function by observing pairs of parallel lights. The results show that the accuracy of the estimated parameters improves. Moreover, because the constraint given by a pair

TABLE I

THE RMS ERRORS ARE COMPUTED USING THE CALIBRATION RESULTS OF THE FOUR METHODS. THE ERRORS OF THE PROPOSED METHOD ARE THE SMALLEST OF THE FOUR METHODS FOR BOTH TESTS. FURTHERMORE, THE ERRORS OF THE OTHER METHODS BECOME LARGER IF THE TEST ENVIRONMENTS DIFFER FROM THE CALIBRATED ENVIRONMENTS.

Calibration Method	Calibration Error (pixels)	Indoor Test		Outdoor Test	
		Fitting Error (pixels)	Reprojection Error (pixels)	Fitting Error (pixels)	Reprojection Error (pixels)
Parallel	0.366	0.123	0.908	0.401	1.473
Planes	0.394	0.122	1.120	0.390	3.443
Indoor	0.196	0.126	1.068	0.420	1.971
Outdoor	0.400	0.130	1.477	0.715	1.595

of parallel lights is represented as a 4D surface, we can easily visualize the error of the calibrated result and the degeneracy of the input data. For the parallel light source, either a distant object or a collimated light can be used. If a distant object is used, a map for measuring the angle is the simplest calibration method and no apparatus is necessary. In contrast, if a collimated light is used, a precise and small system for calibration can be constructed. It is envisaged that the generality of the camera calibration method developed will be explored and expanded in future studies.

REFERENCES

- [1] R. Tsai, "A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses," *IEEE Journal of Robotics and Automation*, vol. 3, no. 4, pp. 323–344, 1987.
- [2] O. Faugeras, *Three-dimensional computer vision: a geometric viewpoint*. Cambridge, MA, USA: MIT Press, 1993.
- [3] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, ISBN: 0521540518, 2004.
- [4] P. Sturm and S. Maybank, "On plane-based camera calibration: A general algorithm, singularities, applications," in *Proc. the IEEE Conference on Computer Vision and Pattern Recognition*, Fort Collins, USA, June 1999, pp. 432–437.
- [5] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000, <http://research.microsoft.com/~zhang/Calib/>.
- [6] R. Hartley, "Self-calibration from multiple views with a rotating camera," in *Proc. the 3rd European Conference on Computer Vision*, vol. 1, Stockholm, Sweden, May 1994, pp. 471–478.
- [7] G. Stein, "Accurate internal camera calibration using rotation, with analysis of sources of error," in *Proc. Fifth International Conference on Computer Vision*, Cambridge, Massachusetts, June 1995, pp. 230–236.
- [8] B. Hallert, "The method of least squares applied to multicollimator camera calibration," *Photogrammetric Engineering*, vol. 29, no. 5, pp. 836–840, 1963.
- [9] —, "Notes on calibration of cameras and photographs in photogrammetry," *Photogrammetria*, no. 23, pp. 163–178, 1968.
- [10] R. Karren, "Camera calibration by the multicollimator method," *Photogrammetric Engineering*, vol. 34, no. 7, pp. 706–719, 1968.
- [11] P. Carman, "Camera calibration laboratory at N.R.C." *Photogrammetric Engineering*, vol. 35, no. 4, pp. 372–376, 1969.
- [12] P. Carman and H. Brown, "The NRC camera calibrator," *Photogrammetria*, vol. 34, no. 4, pp. 147–165, 1978.
- [13] H. Schmid, "Stellar calibration of the orbicon lens," *Photogrammetric Engineering*, vol. 40, no. 1, pp. 101–111, 1974.
- [14] A. Klaus, J. Bauer, K. Karner, P. Elbischger, R. Perko, and H. Bischof, "Camera calibration from a single night sky image," in *Proc. the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, Washington, DC, USA, 2004, pp. 151–157.
- [15] B. Caprile and V. Torre, "Using vanishing points for camera calibration," *International Journal of Computer Vision*, vol. 4, no. 2, pp. 127–140, March 1990.
- [16] R. Cipolla, T. Drummond, and D. P. Robertson, "Camera calibration from vanishing points in images of architectural scenes," in *Proc. British Machine Vision Conference*, vol. 2, Nottingham, UK, September 1999, pp. 382–391.
- [17] K. Daniilidis and J. Ernst, "Active intrinsic calibration using vanishing points," *Pattern Recognition Letters*, vol. 17, no. 11, pp. 1179–1189, September 1996.
- [18] J. Guillemaut, A. Aguado, and J. Illingworth, "Using points at infinity for parameter decoupling in camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 2, pp. 265–270, February 2005.
- [19] R. Sagawa, N. Aoki, Y. Mukaigawa, T. Echigo, and Y. Yagi, "Mirror localization for a catadioptric imaging system by projecting parallel lights," in *Proc. IEEE International Conference on Robotics and Automation*, Rome, Italy, April 10–14 2007, pp. 3957–3962.
- [20] R. Sagawa, N. Aoki, and Y. Yagi, "Mirror localization for catadioptric imaging system by observing parallel light pairs," in *Proc. 8th Asian Conference on Computer Vision*, Tokyo, Japan, Nov. 2007, to appear.
- [21] R. Sagawa, M. Takatsuji, T. Echigo, and Y. Yagi, "Calibration of lens distortion by structured-light scanning," in *Proc. 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Edmonton, Canada, August 2005, pp. 1349–1354.
- [22] J.-Y. Bouguet, "Camera Calibration Toolbox for Matlab," http://www.vision.caltech.edu/bouguetj/calib_doc/, 2007.
- [23] Intel Corporation, "Open Source Computer Vision Library," <http://www.intel.com/technology/computing/opencv/index.htm>, 2007.
- [24] P. A. Beardsley, A. Zisserman, and D. W. Murray, "Sequential updating of projective and affine structure from motion," *International Journal of Computer Vision*, vol. 23, no. 3, pp. 235–259, 1997.
- [25] http://www.am.sanken.osaka-u.ac.jp/~sagawa/data/ICRA08_calib_data.zip, MATLAB code and data for Section IV-B.