Estimation of extrinsic parameters of multiple fish-eye cameras using calibration markers

Hiroya Okamoto[a], Yuki Tanka[a], Akihiko Hishigi[a], Abdelaziz Khiat[b], Noriko Shimomura[b], Gakuto Masuyama[a] and Kazunori Umeda[a]

[a]Chuo University, 1-13-27 Kasuga, Bunkyo-ku, Tokyo, Japan
[b]Nissan Motor Co., Ltd., 1-1 Morinosato-aoyama, Atsugi, Kanagawa, Japan

ABSTRACT

This paper introduces a simultaneous estimation method of the extrinsic parameters of multiple fish-eye cameras using simple calibration markers. Precise extrinsic parameters of cameras mounted on a car are important, for example, to provide a seamless overhead view image to the driver. Calibration markers are set in the area that are observable from adjacent two cameras. Extrinsic parameters of each camera are estimated individually and then combined and refined using a geometric constraint. Cube markers are chosen as the calibration markers. The method is evaluated by simulation and experiments using a real car. It is shown that extrinsic parameters are obtained by the proposed method and suggested that the errors of intrinsic parameters affect the estimation of extrinsic parameters.

Keywords: Fish-eye camera, in-vehicle camera, extrinsic parameter, camera calibration, PnP problem, calibration

1. INTRODUCTION

Recently, camera systems are widely used in cars to assist drivers. As an example, a system to generate an overhead view image was proposed[1] and is implemented to commercial cars. The system uses multiple fish-eye cameras and generates an overhead image by combining images of the cameras. Therefore, it is important to calibrate multiple fish-eye cameras correctly to generate a seamless overhead image with small errors. There are several parameters to calibrate, such as intrinsic, extrinsic, and optical parameters. Several studies have discussed calibration of intrinsic parameters.[2,3] In this paper, we focus on extrinsic parameters, i.e., camera position and orientation.

The most common method to calibrate a camera is the one proposed by Zhang.[4] The method shows a planar calibration pattern several times to a camera from different directions and estimates both intrinsic and extrinsic parameters. Zhang's method is widely used for calibration of normal cameras with perspective projection. However, the distortion of a fish-eye camera is strong compared to a normal camera and Zhang's method cannot be applied to a fish-eye camera. Scaramuzza proposed a similar method that can be applied to a fish-eye camera.[5] However, their method is for a single camera and thus may cause large gaps at the overlapped regions of multiple images in the overhead image. Okutsu et al. proposed a method to estimate intrinsic and extrinsic camera parameters of multiple fish-eye cameras.[6] However, their method calibrates each camera individually. It is required to calibrate multiple (fish-eye) cameras simultaneously.

In this paper, we introduce a method to obtain extrinsic parameters of multiple fish-eye cameras simultaneously using simple calibration markers. We evaluate the method by simulation and experiments using a real car.

2. PROJECTION OF A FISH-EYE CAMERA

The projection model of a fish-eye camera is usually represented by using a radial relation. Suppose θ [rad] is the angle between the direction to a target point and the optical axis of the camera lens, and ρ [pixel] is the length from the principal point (the point corresponding to the optical axis) and the target point in the image. The typical projection models are equidistance projection and orthogonal projection, that are represented respectively as follows.

$$\rho = \delta\theta \tag{1}$$

$$\rho = \delta \sin\theta \tag{2}$$

Further author information: (Send correspondence to Kazunori Umeda.)
Akihiko Hishigi.: E-mail: hishigi@sensor.mech.chuo-u.ac.jp
Kazunori Umeda.: E-mail: umeda@mech.chuo-u.ac.jp

$$\delta = f/w$$

where f is the focal length [mm] and w is the pixel size of a image sensor [mm]. A real fish-eye camera does not follow the ideal models such as (1) and (2). One of the models to represent a real fish-eye camera is a polynomial model[6] that represents the radial model as

$$\rho = k_1\theta + k_3\theta^3 + k_5\theta^5 + ... \tag{3}$$

where k_1, k_3, k_5 are coefficients. And another typical model is the one presented by Scaramuzza et al.[5] The relation between a three-dimensional (3D) point $\boldsymbol{P} = [X\ Y\ Z]^T$ and image coordinate $\boldsymbol{p} = [u\ v]^T$ and principal point (image center) $\boldsymbol{p}_0 = [u_0\ v_0]^T$ is given as follows.

$$\boldsymbol{P} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \approx \begin{bmatrix} u{-}u_0 \\ v{-}v_0 \\ f(\rho) \end{bmatrix} \tag{4}$$

$\approx$ means that they are equal as homogeneous coordinates. $\rho = \sqrt{(u - u_0)^2 + (v - v_0)^2}$ is the distance from the principal point $\boldsymbol{p}_0 = [u_0\ v_0]^T$ to the image coordinate $\boldsymbol{p} = [u\ v]^T$. $f(\rho)$ is a polynomial of ρ and represented as follows.

$$f(\rho) = a_0 + a_1\rho + a_2\rho^2 + a_3\rho^3 + a_4\rho^4 + ... \tag{5}$$

In this paper, we adopt up to the 4th term and set the coefficients a_0, a_1, a_2, a_3, a_4 as camera intrinsic parameters. With the principal point $\boldsymbol{p}_0 = [u_0\ v_0]^T$, the camera intrinsic parameters of each camera $\boldsymbol{I}_i$ becomes

$$\boldsymbol{I}_i = [a_{0i}\ a_{1i}\ a_{2i}\ a_{3i}\ a_{4i}\ u_{0i}\ v_{0i}]^T \tag{6}$$

where i is the camera number.

3. ESTIMATOIN OF EXTRINSIC PARAMETERS OF MULTIPLE FISH-EYE CAMERAS

In this section, we introduce the method to estimate extrinsic parameters of multiple fish-eye cameras. We use a calibration marker of a cubic shape. We assume that four fish-eye cameras with known intrinsic parameters are set on a car as shown in Fig.1. Note that numbers other than four are possible and cameras are not necessarily on a car.

Markers with given dimensions are set as shown in Fig.1. Each marker is set in the area that are observable from adjacent two cameras. That is, marker A is observed by cameras 3, 1, marker B is observed by cameras 1, 2, marker C is observed by cameras 2, 4 and marker A is observed by cameras 4, 3. For this calibration setting, we solve the Perspective n-Point (PnP) problem and estimate the camera extrinsic parameters, that represent each camera's position and orientation in the world coordinate system. For each camera, we set the optical axis as the Y-axis as shown in Fig.2, and set the pitch, roll, yaw angles in the world coordinate system as α_{cami}, β_{cami}, γ_{cami} respectively, and camera position as $[X_{cami}\ Y_{cami}\ Z_{cami}]^T$. Then the extrinsic parameters $\boldsymbol{E}_i$ are defined as follows.

$$\boldsymbol{E}_i = [X_{cami}\ Y_{cami}\ Z_{cami}\ \alpha_{cami}\ \beta_{cami}\ \gamma_{cami}]^T \tag{7}$$

3.1 Outline of parameter estimation

We estimate the extrinsic parameters by the following procedures.

1. Estimate two cameras' position and orientation in the local coordinate system of each marker, so that the square sum of errors between input points and reprojected points becomes minimum.

2. Set an arbitrary local coordinate system to a world coordinate system. Then transform local coordinates of markers to world coordinates by using the estimated camera's position and orientation for each marker.

3. Estimate each camera's position and orientation in the world coordinate system, i.e., extrinsic parameters $\boldsymbol{E}_i$ by applying the same procedure as 1., using the world coordinates of each marker.

The procedures 1(3) and 2 are described in 3.2 and 3.3 respectively.

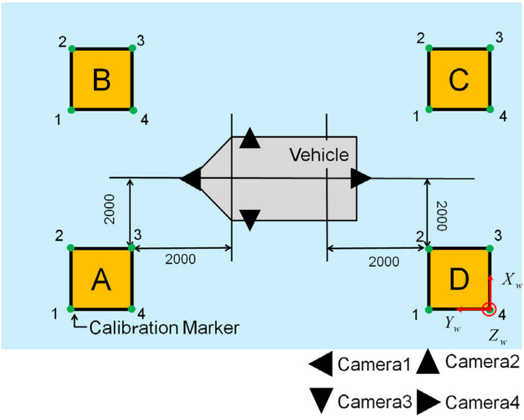


Figure 1. Calibration environment

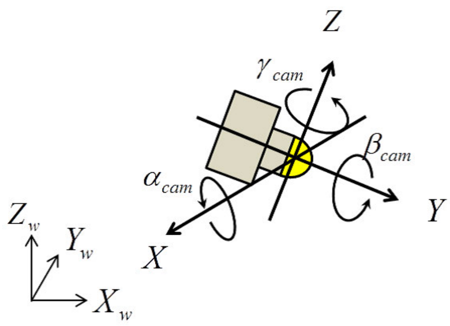


Figure 2. Camera coordinate system

3.2 Estimation of extrinsic parameters

We estimate extrinsic parameters of each camera by the following procedures. j represents the number of vertices that are projected in a fish-eye image. In the procedure 3. of 3.1, the number is twice of the number of vertices of each marker, because the estimation is carried out using two markers simultaneously.

1. Capture a fish-eye image that contains two marker images.
2. Obtain the coordinates of vertices of a marker $\boldsymbol{m}_{fj} = [u_j \ v_j]^T$ in the image coordinate system.
3. Reproject the 3D coordinates of vertices of a marker in the world or marker's coordinate system to the image coordinate system, and obtain the image coordinate $\boldsymbol{m}_{wj} = [u_{wj} \ v_{wj}]^T$.
4. Estimate the extrinsic parameters $\boldsymbol{E}_i$ by minimizing the square sum of reprojection errors D between $\boldsymbol{m}_{fj}$ and $\boldsymbol{m}_{wj}$.

The procedures are applied to each camera i $(i = 1, 2, 3, 4)$ independently.

The reprojection error function D is obtained as follows. The reprojected point $\boldsymbol{m}_{wj}$ is represented using the camera intrinsic and extrinsic parameters $\boldsymbol{I}_i$ and $\boldsymbol{E}_i$ as follows (see Fig.3).

$$\boldsymbol{m}_{wj} = \begin{bmatrix} u_{wj} \\ v_{wj} \end{bmatrix} = \frac{\rho}{\sqrt{X_{cj}^2 + Y_{cj}^2}} \begin{bmatrix} X_{cj} \\ Y_{cj} \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} \tag{8}$$

From (5), ρ is calculated by solving the 4th-degree equation

$$f(\rho) = a_0 + a_1\rho + a_2\rho^2 + a_3\rho^3 + a_4\rho^4. \tag{9}$$

$f(\rho)$ is given by $f(\rho) = \dfrac{Z_{cj}}{\sqrt{X_{cj}^2 + Y_{cj}^2}}\rho$ from (4). $\boldsymbol{X}_{cj} = [X_{cj} \ Y_{cj} \ Z_{cj}]^T$ is the coordinates that the world coordinates $\boldsymbol{X}_{wj}$ of measured point $\boldsymbol{P}_j$ are transformed in the camera coordinate system, and represented as

$$\begin{aligned} \tilde{\boldsymbol{X}}_{cj} &= M\tilde{\boldsymbol{X}}_{wj} \\ M &= \left[\begin{array}{c|c} R & \boldsymbol{t} \\ \hline \boldsymbol{0}^T & 1 \end{array}\right] \end{aligned} \tag{10}$$

$\tilde{\boldsymbol{X}}$ represents the homogeneous coordinates of $\boldsymbol{X}$, i.e., $\tilde{\boldsymbol{X}} = [\boldsymbol{X}^T \ 1]^T$. R is the 3×3 rotation matrix with roll β_{cami}, pitch α_{cami}, yaw γ_{cami} in this order, and $\boldsymbol{t}$ is a 3D translation vector. R and $\boldsymbol{t}$ are calculated from camera extrinsic parameters

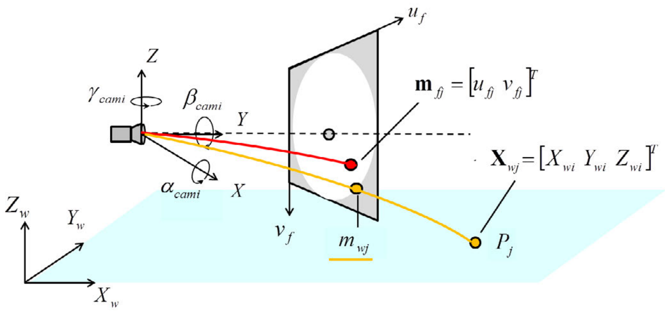


Figure 3. Reprojection on fish-eye image

$\boldsymbol{E}_i$. The point reprojected in the fish-eye camera coordinate system $\boldsymbol{m}_{wj}$ can be expressed as the function of $\boldsymbol{E}_i$ and $\boldsymbol{X}_{wj}$. We estimate the extrinsic parameters $\boldsymbol{E}_i$ so that the square sum of the distance between the reprojected point $\boldsymbol{m}_{wj}$ and the projected point of the original measured point $\boldsymbol{m}_{fj}$ becomes minimum. The sum D is given as

$$D = \sum_{j=1}^{N} \|\boldsymbol{m}_{fj} - \boldsymbol{m}_{wj}(\boldsymbol{E}_i,\ \boldsymbol{X}_{wj})\|^2 \tag{11}$$

where N is the number of vertices of a marker projected in a fish-eye image. We use modified Powell's method[7] for the minimization.

3.3 World coordinates of markers

We detail the procedure to obtain the marker's 3D coordinates in the world coordinate system described in 3. of 3.1. Assume k represents the marker's label A, B, C, D. We set marker's 3D coordinates in the local and world coordinate systems as $\boldsymbol{X}_k = [X_k\ Y_k\ Z_k]^T$ and $\boldsymbol{X}_{wk} = [X_{wk}\ Y_{wk}\ Z_{wk}]^T$ respectively. From Fig.4, transform from local coordinate system to world coordinate system is represented by using the camera's position and orientation $Rt_{(i,k)}$ as

$$\begin{aligned}
\tilde{\boldsymbol{X}}_{wA} &= M_A \tilde{\boldsymbol{X}}_A, M_A = Rt^{-1}_{(3,\ D)} Rt_{(3,\ A)} \\
\tilde{\boldsymbol{X}}_{wB} &= M_B \tilde{\boldsymbol{X}}_B, M_B = M_C Rt^{-1}_{(2,\ C)} Rt_{(2,\ B)} \\
\tilde{\boldsymbol{X}}_{wC} &= M_C \tilde{\boldsymbol{X}}_C, M_C = Rt^{-1}_{(4,\ D)} Rt_{(4,\ C)} \\
\tilde{\boldsymbol{X}}_{wD} &= \tilde{\boldsymbol{X}}_D
\end{aligned} \tag{12}$$

$Rt_{(i,k)}$ is given similarly to the homogeneous transform matrix M in (10). We set the origin of the world coordinate system to the vertex 4 of marker 4 in Fig.4.

4. SIMULATION

We conducted some simulation to evaluate the accuracy of the proposed extrinsic parameter estimation. We used a cubic object as a calibration marker.

4.1 Simulation condition

We explain conditions of simulation. Table 1 shows the extrinsic parameters of each camera $\boldsymbol{E}_i$. The parameters are the same as the ones in the experiments of the next section. As the projection model, we used (5). We estimated the intrinsic parameters of Camera 1 for the experiments by the method by Okutsu.[8] Table 2 shows the estimated intrinsic parameters. In this simulation, we assumed that the intrinsic parameters have no errors because we focus on extrinsic parameter estimation. The calibration marker is a cube with 1200mm length, that is almost the same as the one in the experiments, and we used eight vertices as the feature points.

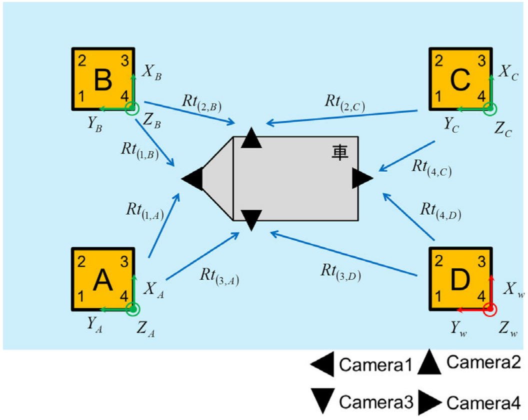


Figure 4. Conditions of extrinsic parameter estimation

Table 1. Camera extrinsic parameters $\boldsymbol{E}_i$

Camera	X_{cam} [mm]	Y_{cam} [mm]	Z_{cam} [mm]	α_{cam} [deg]	β_{cam} [deg]	γ_{cam} [deg]
1	3500	7250	650	-20	0	0
2	2500	5800	800	-20	0	-90
3	4500	5800	800	-20	0	90
4	3500	2500	670	-20	0	180

We assumed that the markers are set as in Fig.1 and produced simulation data. Fig.5 shows the produced fish-eye images. Size of the simulation image is 1328×1048 pixel.

Simulation was carried out as follows. The vertices of a marker in the produced fish-eye image are used as the feature points, and Gaussian noises with $\sigma = 0, 0.2, 0.6, 0.8, 1.0, 2.0, 3.0$ pixel were added. Then extrinsic parameters of each camera were estimated by the proposed method. The number of simulation was 100 for each σ, and the averages of absolute errors of estimated parameters and their standard deviation were calculated.

4.2 Simulation results

Fig.6 shows the simulation results. Horizontal axis represents the σ [pixel] of the Gaussian noise added to the feature points, and vertical axis represents the average absolute error of position [mm] and orientation [deg]. Error bars represent the standard deviation.

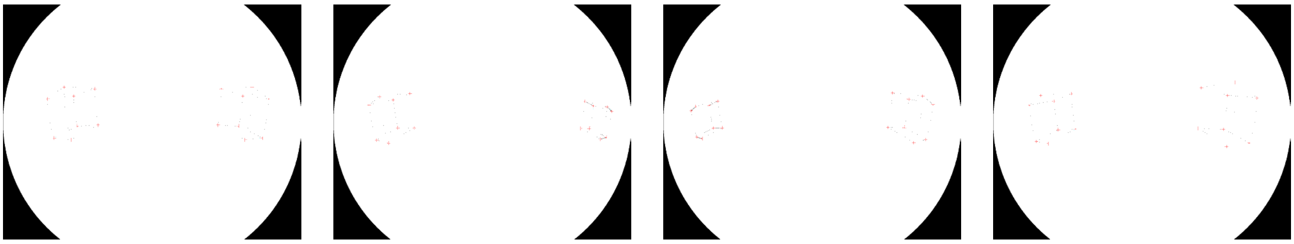
The results show that the extrinsic parameters are estimated without errors if the feature points are given without errors, and that both the position errors and orientation errors increase roughly proportional to the given error size. When σ is 1 pixel or less, position errors are less than 50 mm and orientation errors are less than 1 deg except Camera1's roll angle β_{cam1}. Although the required accuracy depends on real applications, the errors for $\sigma \leq 1$ seem to be reasonably small. And extraction of feature points with the accuracy of 1 pixel is not very difficult in general.

5. EXPERIMENTS WITH A REAL CAR

We conducted experiments to estimate extrinsic parameters of four fish-eye cameras that were attached on a real car using cube markers with the same dimensions in the simulation. In this experiments, we estimated the intrinsic parameters $\boldsymbol{I}_i$ of the fish-eye cameras by using the method of Scaramuzza et al.[5] that is implemented in Matlab. Table 3 shows the estimated

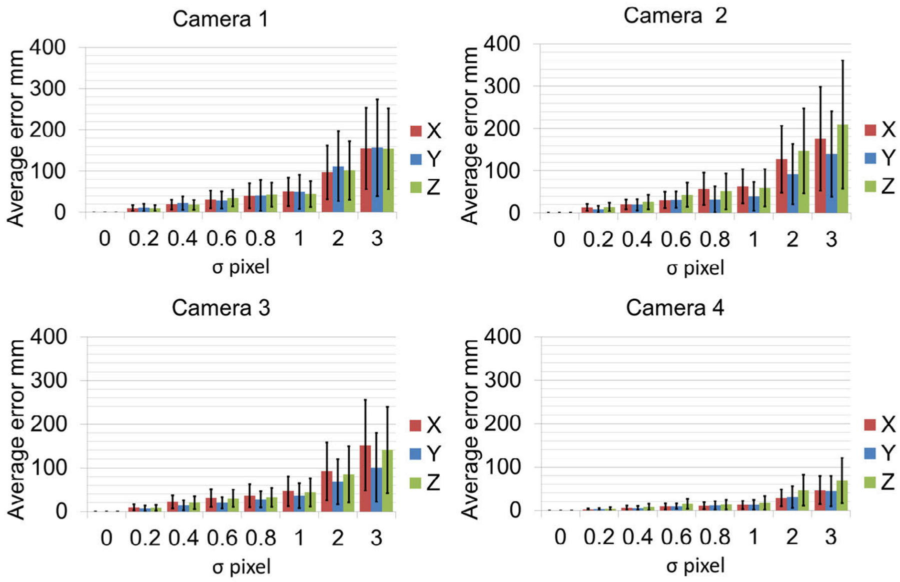
Table 2. Camera intrinsic parameters $\boldsymbol{I}_1$

k_1	k_3	k_5
338.518	24.650	-1.364

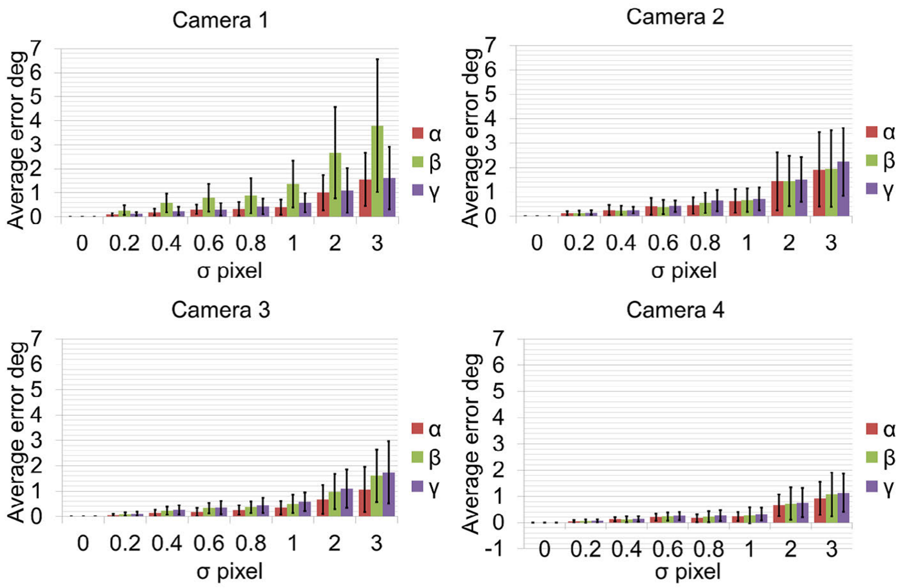


(a) Camera 1 (b) Camera 2 (c) Camera 3 (d) Camera 4

Figure 5. Fish-eye images of cube markers



(a) Average error of position



(b) Average error of pose

Figure 6. Simulation results

Table 3. Camera intrinsic parameters $\boldsymbol{I}_i$

Camera	a_0	a_1	$a_2(\times 10^{-4})$	$a_3(\times 10^{-7})$	$a_4(\times 10^{-9})$	u_0	v_0
1	-391.58	0.00	9.57	-6.12	1.14	689.61	569.30
2	-392.53	0.00	9.59	-6.14	1.14	685.09	556.87
3	-392.47	0.00	9.82	-6.78	1.20	677.74	554.75
4	-392.23	0.00	9.50	-6.92	1.12	686.47	577.17

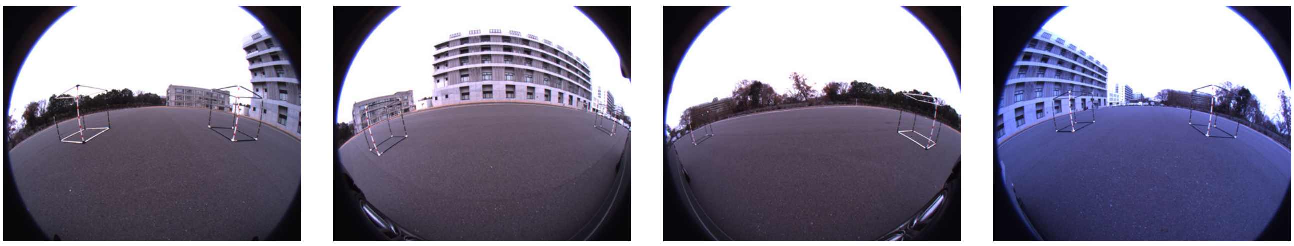
Figure 7. Cube marker

intrinsic parameters. The marker was constructed using pipes and joints. Fig.7 shows the marker in the fish-eye image. Fig.8 shows the examples of images used in the experiments. From such images, we can extract vertices of markers by transforming the marker region to remove local distortion and then extracting lines of pipes by applying Hough transform to the marker image, and finally detecting the intersection points of the lines as the vertices (see Fig.9). However, we detected the vertices manually with subpixel resolution in the experiments. The initial values of the extrinsic parameters were set to zero. We set the origin of the world coordinate system to the Camera 1 position with height z=0. Table 4 shows the ground truth of the extrinsic parameters.

Table 5 shows the estimation errors of extrinsic parameters. The orientation errors are rather small: less than 1.5 deg. However, position errors are not small. By comparing with the simulation results, it is equivalent to the errors of feature points of as large as 3 pixel or so. We think that the large errors are due to the errors of intrinsic parameters. It will be required to improve the projection model of the fish-eye camera and develop a more precise calibration method of intrinsic parameters.

6. CONCLUSIONS

We introduced a method to estimate extrinsic parameters of multiple fish-eye cameras simultaneously. A known calibration marker is set in the area that are observable from adjacent two cameras and individually estimated extrinsic parameters are combined and refined. Simulation and experiments using a real car were conducted using cube markers, and showed that



(a) Camera1 (b) Camera2 (c) Camera3 (d) Camera4

Figure 8. Samples of input images

Table 4. Ground truth of extrinsic parameters

Camera	X_{cam} [mm]	Y_{cam} [mm]	Z_{cam} [mm]	α_{cam} [deg]	β_{cam} [deg]	γ_{cam} [deg]
1	0.0	0.0	650.0	-19.1	1.1	-0.2
2	1060.0	-1355.0	800.0	-21.2	-90.0	-2.8
3	-1055.0	-1342.0	800.0	-19.8	90.2	-2.1
4	-510.0	-5060.0	650.0	-20.8	179.1	0.3

Table 5. Error of extrinsic parameters

Camera	X_{cam} [mm]	Y_{cam} [mm]	Z_{cam} [mm]	α_{cam} [deg]	β_{cam} [deg]	γ_{cam} [deg]
1	0.0	0.0	18.0	-0.1	1.3	-1.4
2	-121.0	-148.2	59.4	0.6	-0.1	1.1
3	82.1	-164.2	-22.1	-0.1	-0.5	-1.2
4	-30.0	-132.7	10.0	0.4	-0.3	0.7

extrinsic parameters are obtained by the proposed method and suggested that the errors of intrinsic parameters affect the estimation of extrinsic parameters.

Improvement of the projection model of the fish-eye camera and construction of a more precise calibration method of intrinsic parameters, and application to real problems will be our future works.

REFERENCES

[1] K. Oizumi, e. a., "Development of "all-around view" system," *SAE Technical Paper* , 2003–01–0017 (2003).

[2] F. Devernay, e. a., "Straight lines have to be straight," *Machine Vision and Applications* **13**, 14–24 (2001).

[3] J. Kannala, e. a., "A generic camera calibration method for conventional, wide-angle, and fish-eye lenses," *IEEE Trans. PAMI* **28**(8), 1335–1340 (2006).

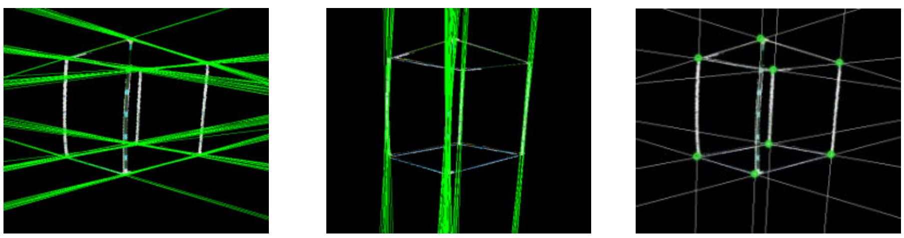
[4] Zhang, Z., "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**(11), 1330–1334 (2000).

[5] D. Scaramuzza, A. M. and Siegwart, R., "A toolbox for easily calibrating omnidirectional cameras," in [*IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2006)*], 5695–5701 (2006).

[6] R. Okutsu, e. a., "Generation of overhead view images by estimating intrinsic and extrinsic camera parameters of multiple fish-eye cameras," in [*Proc. IAPR Conference on Machine Vision Applications (MVA2009)*], 447–450 (2009).

[7] Powell, M., "An efficient method for finding the minimum of a function of several variables without calculating derivatives," *Computer Journal* **7**(2), 155–162 (1964).

[8] R. Okutsu, K. T. and Umeda, K., "Calibration of intrinsic parameters of a fish-eye camera using a sphere," *IEICE Trans. Inf.& Syst.(Japanese Edition)* **J93-D**(12), 2645–2653 (2010). in Japanese.



(a) Line detection ($|\theta| > \pi/4$) (b) Line detection ($|\theta| \leq \pi/4$) (c) Extraction of vertices

Figure 9. Extraction of vertices using Hough transform