

Fisheye Camera's Intrinsic Parameter Estimation Using Trajectories of Feature Points Obtained from Camera Rotation

Akihiko Hishigi, Yuki Tanaka, Gakuto Masuyama, and Kazunori Umeda

Abstract—This paper proposes a simple method of estimating fisheye camera's intrinsic parameters without calibration targets. This method takes advantage of the trajectories of feature points in the scene. The trajectories of feature points are obtained from a rotation movement of the camera around the vertical axis. The feature points are utilized for calibration, thus specific calibration targets are not required. The proposed method estimates intrinsic parameters of a real fisheye camera. The validity of the proposed method is verified by the perspective projection of a distorted image using estimated parameters. In addition, this paper compares the results of perspective projection using the proposed method and the conventional method.

I. INTRODUCTION

Fisheye cameras are widely used as sensors to acquire information from the external world. This camera has a wide angle of view (almost 180 [deg]) and can measure a wide range of information at once. Therefore, it is effective for building sensor systems, such as driver assistance systems and monitoring systems, that require a wide range of measurements at a low cost. However, images obtained from a fisheye camera have intractable distortions. Thus, intrinsic parameters must be estimated accurately to remove the distortion by perspective projection. Intrinsic parameters represent the individual differences of each lens. The accuracy of the estimation of this parameter affects the image-processing result.

Many studies have estimated the intrinsic parameters of the fisheye camera [1]-[6]. However, these methods are often cumbersome, due to the requirement of a special target. In Scaramuzza's method [7], a checkerboard pattern is presented repeatedly to calibrate the fisheye camera. Kannala [8] estimated the intrinsic parameters of the fisheye camera to shoot only once for each piece of the calibration board. However, that is also cumbersome due to the size of calibration board (2×3 [m²]). The quality of the calibration depends on the way or the number of times the target is shown.

Therefore, this paper develops a marker-free calibration method. As seen in Figure 1, this method estimates the intrinsic parameters of a fisheye camera using trajectories of feature points obtained by rotating the camera; the camera is

Akihiko Hishigi and Yuki Tanaka are with the Course of Precision Engineering, School of Science and Engineering, Chuo University, 1-13-27, Kasuga, Bunkyo-ku, Tokyo, Japan, (e-mail: hishigi@sensor.mech.chuo-u.ac.jp).

Gakuto Masuyama and Kazunori Umeda are with the Department of Precision Mechanics, Faculty of Science and Engineering, Chuo University, 1-13-27, Kasuga, Bunkyo-ku, Tokyo, Japan, (e-mail: {masuyama, umeda}@mech.chuo-u.ac.jp).

rotated around the vertical axis through the optical center of the camera. Note that the trajectory used in this paper is defined as the coordinates of a feature point, excluding temporal information from the trajectory in the meaning that is usually used. A brief overview of the proposed method is given in [9]. In this paper, we present details of the method with two minor modifications: camera model and evaluation function. Furthermore, we performed experiments using actual equipment.

The verification of the proposed method is shown by simulation, and an experiment was conducted using the actual equipment. In addition, this paper compares the estimation results with an existing approach.

II. INTRINSIC PARAMETERS OF THE FISHEYE CAMERA

A. Model of the fisheye

This paper uses the generic omnidirectional camera model that was proposed by Scaramuzza [7]. Figure 2 shows an outline of the camera model.

A 3D position, $\mathbf{P} = [X \ Y \ Z]^T$, is represented by the 2D position $\mathbf{p} = [u \ v]^T$ and the image center $\mathbf{p}_0 = [u_0 \ v_0]^T$:

$$\mathbf{P} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \approx \begin{bmatrix} u - u_0 \\ v - v_0 \\ -f(\rho) \end{bmatrix}, \quad (1)$$

where the origin of the image coordinate system is in the upper left corner. The \approx symbol represents equivalence up to scale. ρ represents the distance from the image center, \mathbf{p}_0 , to a 2D position $\mathbf{p} = [u \ v]^T$:

$$\rho = \sqrt{(u - u_0)^2 + (v - v_0)^2}. \quad (2)$$

$f(\rho)$, which is a polynomial of ρ , is expressed as follows: $f(\rho) = a_0 + a_1\rho + a_2\rho^2 + a_3\rho^3 + a_4\rho^4 + \dots$. (3)

In this paper, $f(\rho)$ is a fourth-order polynomial. Including the center of the image, intrinsic parameters to estimate are defined as follows:

$$\mathbf{I} = [a_0 \ a_1 \ a_2 \ a_3 \ a_4 \ u_0 \ v_0]^T. \quad (4)$$

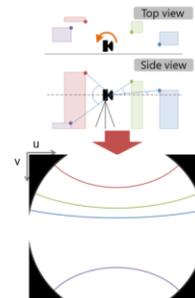


Fig. 1. Estimation of intrinsic parameters [9]

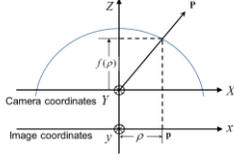


Fig. 2. Camera model

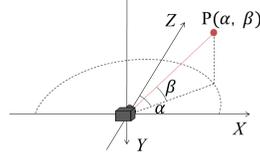


Fig. 3. 3D coordinates

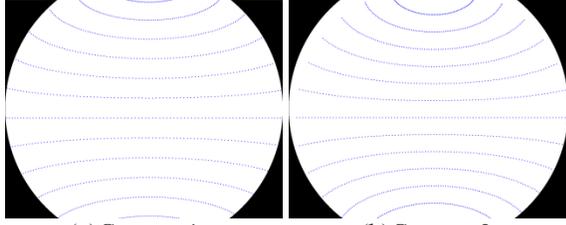


Fig. 4. Shape differences between the two cameras

B. Projection of the 3D point

In this paper, the 3D point, \mathbf{P} , is represented by azimuth α and elevation β utilizing the rotation of the camera, as shown in Figure 3.

\mathbf{p} is projected from eq. (1) on the image as follows:

$$\mathbf{p} = \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} -\tan\alpha \cdot f(\rho) + u_0 \\ \sqrt{\tan^2\alpha + 1} \cdot \tan\beta \cdot f(\rho) + v_0 \end{bmatrix}, \quad (5)$$

where ρ is the real solution of eq. (3). $f(\rho)$ is given from eq. (1) and Figure 3 as follows:

$$\begin{aligned} f(\rho) &= -\frac{Z}{\sqrt{X^2 + Y^2}} \rho \\ &= -\frac{\rho}{\sqrt{\tan^2\alpha + (\tan^2\alpha + 1) \cdot \tan^2\beta}}. \end{aligned} \quad (6)$$

III. THE ESTIMATION METHOD

This method assumes that the fisheye camera is installed so as to satisfy the following conditions.

- The rotation axis of the camera passes through the optical center.
- The optical axis and the rotation axis cross one another perpendicularly.

Trajectories are captured by rotating the camera. These trajectories include information regarding the azimuth α , elevation β , and intrinsic parameters of the camera. Figure 4 shows the shapes of trajectories based on different intrinsic parameters in the simulation. Our idea is to estimate the intrinsic parameters based on the geometric shapes of the trajectories, which reflect the differences in the intrinsic parameters.

In this method, the intrinsic parameters of the fisheye camera are estimated by minimizing the evaluation function. Arguments of the evaluation function are intrinsic parameters and the 3D coordinates of observation points. The flow of the estimation is shown as follows.

A. Definition of the evaluation function

First, the following constraint is obtained by erasing α from eq. (5):

$$v_{ri} - v_0 - \sqrt{(u_{fi} - u_0)^2 + f^2(\rho)} \cdot \tan\beta = 0, \quad (7)$$

where u_{fi} is obtained from observation point $\mathbf{p}_{fi} = [u_{fi} \ v_{fi}]^T$. Then we can calculate the re-projected point $\mathbf{p}_{ri} = [u_{ri} \ v_{ri}]^T$, where $u_{ri} = u_{fi}$ and v_{ri} is the solution of eq.(7).

An evaluation function, E , is defined as the sum of the squares of the difference between observation point \mathbf{p}_{fi} and re-projection point \mathbf{p}_{ri} as follows:

$$E = \sum_{i=1}^N (v_{fi} - v_{ri})^2, \quad (8)$$

where N is the number of observation points. Evaluation function E is calculated by β of the observation point, \mathbf{I} and \mathbf{p}_{fi} .

Intrinsic parameters, \mathbf{I} , are estimated by the optimization technique. The elevation angle, β , of the feature point, as shown in the next section, is estimated by iterative calculation. In this paper, a modified Powell method [10] is used to optimize an evaluation function.

B. Flow of the estimation

Figure 5 shows the flow of the estimation of this paper. First, the intrinsic parameters are initialized. Appropriate initial values are given to $a_0 \sim a_4$. u_0 and v_0 are calculated from the symmetry of each trajectory. Details about the initialization process are described in section III.C.

First, the elevation angle of each point in one trajectory is solved from eq. (7) while fixing intrinsic parameters. An average of the elevation angles is then defined as β of the trajectory. The intrinsic parameters are estimated so as to minimize the evaluation function using obtained β . Iterating successive flow, each intrinsic parameter and the elevation are updated. Note that the estimation of $a_0 \sim a_4$ is repeated while increasing the number of intrinsic parameters. We found that this heuristic approach improves the stability of the estimation process.

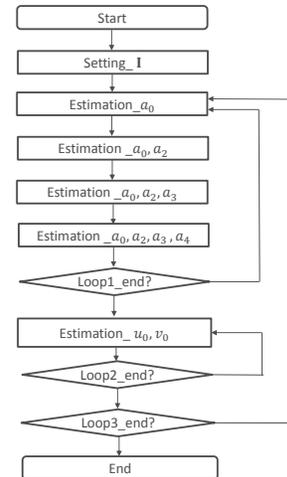


Fig. 5. Flow of estimation

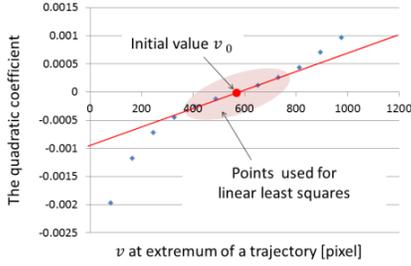


Fig. 6. Calculation of v_0

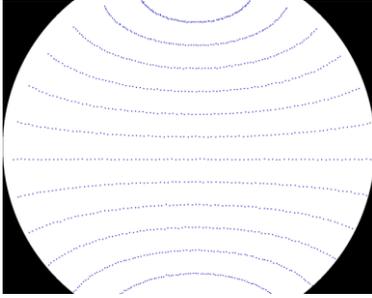


Fig. 7. An example of an artificial feature points

C. Initializing values

Initial values affect the accuracy of estimated results. $a_0 \sim a_4$ are almost the same among fisheye cameras. Thus, we set their initial values empirically. The manual initial guess on intrinsic parameters can be very rough and thus is not hard.

An initial value of u_0 is determined by using a trajectory approximated by a quadratic curve. u_0 is set as the average of u at every extremum of the quadratic curves. However, u at an extremum tends to deviate when the trajectory is nearly linear. Therefore, we may exclude trajectories from the initialization comparing v at the endpoints of each trajectory.

v_0 is calculated from correspondence between the curvature and the vertex coordinate. The quadratic coefficient and v at the extremum of each trajectory are plotted on a 2D plane. Figure 6 shows an example of the plot. The horizontal axis represents v at extremum, and the vertical axis represents a quadratic coefficient. We use points positioned near the center of the fisheye image to apply linear approximation. Then v_0 is determined from the approximated line by substituting 0 for the quadratic coefficient.

IV. SIMULATION

When the trajectories of feature points are captured using the actual camera, ideal trajectories cannot be obtained due to various factors, including errors in installing the camera or errors in feature point detection and tracking. In order to verify the accuracy of the proposed method and the influence of errors, the simulation is conducted using an artificial image.

A. Setting of the environment

An artificial image was produced with the assumption that the optical axis of the camera is rotated horizontally around the optical center in parallel with the horizontal plane. To verify the influence of the measurement error of the feature point, normally distributed artificial noise was added to u_{fi} and v_{fi} , independently. The mean and standard deviations were 0.0 and $\sigma = 0.0, 0.5, 1.0, 2.0$, respectively. We performed 10 experiments in each setting. Figure 7 shows the example of an artificial image used in the experiment. There were 820 observation points in each image. Initial values of $a_0 \sim a_4$ are shown in Table I. To estimate the initial value v_0 stably, four trajectories near the image center were used for estimation.

B. Experiment results

The estimated value of each intrinsic parameter and the final evaluation function are shown in Table II. Figure 8 shows the observation points and re-projected points at the end of the experiment. Only the upper right corner of the image is shown. Each blue dot represents an observation point. Each red rectangle represents a re-projection point. As in Table II, the differences of the image center between the average and the true value were within 1[pixel] in any experiment. Thus, convergence can be correctly confirmed stably even if the measurement error was large. However, the evaluation function was not 0 even if the standard deviation was 0.0[pixel]. The reason is considered to be that the evaluation function converges to a local optimum solution due to the calculation of the elevation angle.

It is difficult to verify the accuracy of $a_0 \sim a_4$. Instead, we compared the trajectories of the re-projected points obtained with the following conditions:

- I. $a_0 \sim a_4$: estimated value
 u_0, v_0 : true value.
- II. $a_0 \sim a_4$: true value
 u_0, v_0 : true value.

The trajectories were assessed by the average of the distance of each re-projected point pair.

The results are shown in Figure 9. The horizontal axis represents σ of the noise. The vertical axis represents the average of the distance between re-projected points. Error bars represent the standard deviation. When σ was 1.0[pixel] and 2.0[pixel], a large re-projection error was observed once for 10 trials. These were excluded from the estimated results as the estimation failure. As in Figure 9, if the standard deviation was 0.0[pixel], the average of the re-projection error was 0.17[pixel] per point. As in Table II, the re-projection error was large despite the small value of the evaluation function. The reason is also considered to be that the evaluation function converges to a local optimum solution. If the standard deviation was 2.0[pixel], the average re-projection error was 1.38[pixel].

Further improvements of accuracy and stability would be required for practical use. The problem, which converges to a local optimum solution, must be solved. We are considering giving another constraint on the optimization process to avoid premature convergence.

TABLE I INITIAL INTRINSIC PARAMETERS

a_0	a_1	a_2	a_3	a_4
-100	0	0	0	0

TABLE II AVERAGE VALUE AND STANDARD DEVIATION OF 10 EXPERIMENTS

	True value	$\sigma = 0.0$		$\sigma = 0.5$		$\sigma = 1.0$		$\sigma = 2.0$	
		Ave.	S.D	Ave.	S.D	Ave.	S.D	Ave.	S.D
a_0	-391.58	-391.40	0.00	-386.58	3.41	-385.47	4.70	-385.83	5.65
a_1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$a_2(\times 10^{-4})$	9.57	9.63	0.00	7.63	1.52	6.31	3.57	6.36	1.86
$a_3(\times 10^{-7})$	-6.12	-6.60	0.00	0.09	5.31	5.38	13.7	5.58	5.56
$a_4(\times 10^{-9})$	1.14	1.21	0.00	0.65	0.49	0.05	1.46	0.40	0.52
u_0 [pixel]	689.61	689.60	0.00	689.05	0.46	689.38	0.43	689.23	1.29
v_0 [pixel]	569.30	570.01	0.00	569.95	0.35	569.73	0.55	570.00	1.21
E [pixel ²]	---	32.70	0.00	794.11	334.44	1508.91	444.50	8120.03	9629.94

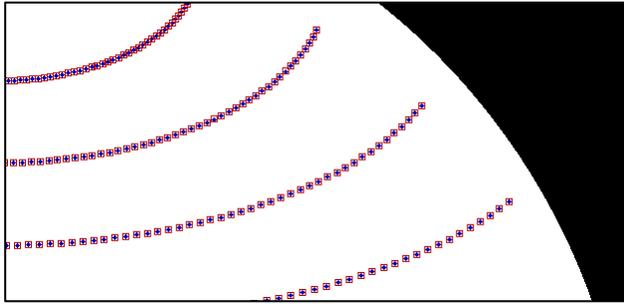
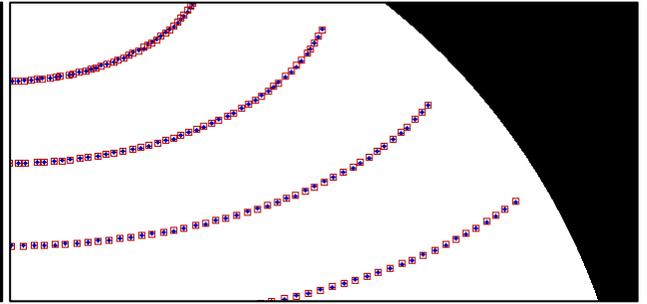
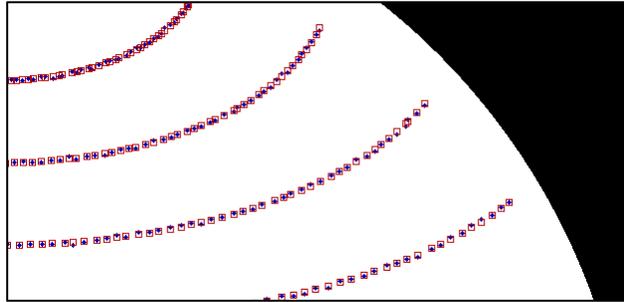
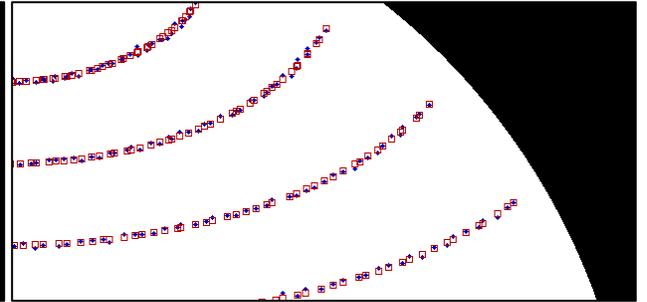
a) $\sigma = 0.0$ [pixel]($E = 32.70$ [pixel²])b) $\sigma = 0.5$ [pixel] ($E = 310.26$ [pixel²])c) $\sigma = 1.0$ [pixel]($E = 1283.11$ [pixel²])d) $\sigma = 2.0$ [pixel] ($E = 5463.65$ [pixel²])

Fig. 8. Examples of the observation points and re-projection points of each experiment

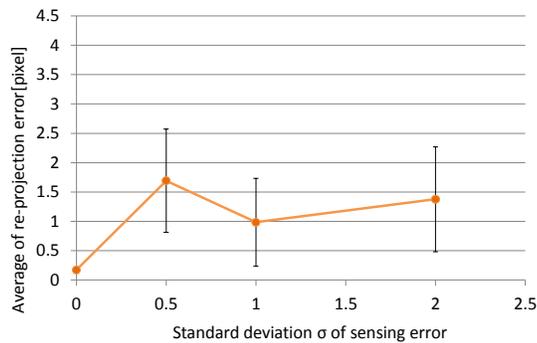
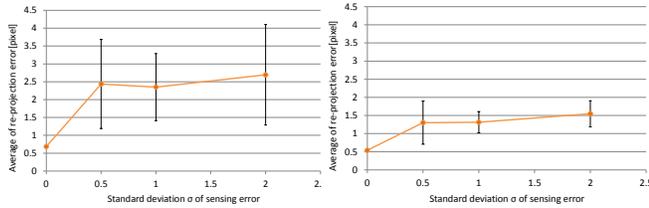


Fig. 9. Re-projection error



(a) Current camera model (b) Previous camera model
Fig. 10. Re-projection error for previous evaluation function

C. Comparison with the previous camera model and evaluation function

To validate the differences of camera models, we applied the previous evaluation function [9] to the current camera model and the previous camera model, respectively. The results are shown in Figure 10. As in Figure 10, the previous camera model is more accurate than the current camera model. However, the previous camera model was unstable and the evaluation function did not converge in several times. In 10 trials for each standard deviation, failures of 3, 2 and 3 times occurred for 0.5[pixel], 1.0[pixel] and 2.0[pixel] respectively. On the other hand, in the current camera model, there was no case that estimation diverged. To summarize, the current model's estimation is easier to estimate without requiring the azimuth angle and more stable compared with the previous model, but is less accurate.

By comparing Figure 9 and Figure 10(a), we can validate the differences of evaluation functions. We can say that the current evaluation function has higher accuracy. The stability is slightly inferior to the estimation using the previous evaluation function, since a few failures occurred for the current evaluation function as mentioned in section IV.B.

V. INTRINSIC PARAMETER ESTIMATION EXPERIMENT WITH ACTUAL EQUIPMENT

A. Generation of feature point trajectories

Figure 11 shows fisheye camera mounted on a tripod. CCD camera is Point Grey Research Dragonfly2. Its number of effective pixels is 1024×768. The fisheye lens is SPACE TV 1634 M. Its focal length is 1.6[mm] and angle of view is 180.0 [deg] × 114.1 [deg]. The experiment with actual equipment used images captured by rotating the mounted camera. The trajectory of each feature point was obtained by sequentially matching feature points of two successive images. In this paper, the AKAZE feature was used for detection and matching using feature points [11]. The AKAZE gives features with sub-pixel accuracy. Euclidian distances between feature points matched with two successive images, are calculated to judge whether matching is successful or not. If the distance between the feature points was larger than the threshold value in two successive images, these were regarded as incorrectly matched.

The camera was rotated manually. Images were captured continuously at a regular interval. Figure 12 depicts transitions of the upper left corner of the whiteboard. It is desired to obtain dense feature points by tracking between

successive images. To estimate intrinsic parameters stably and accurately, we chose full-trajectory manually. Figure 13 shows the environment of the experiment. Trajectories, which were generated by tracking, are shown in Figure 14.



Fig. 11. Fisheye camera mounted on a tripod

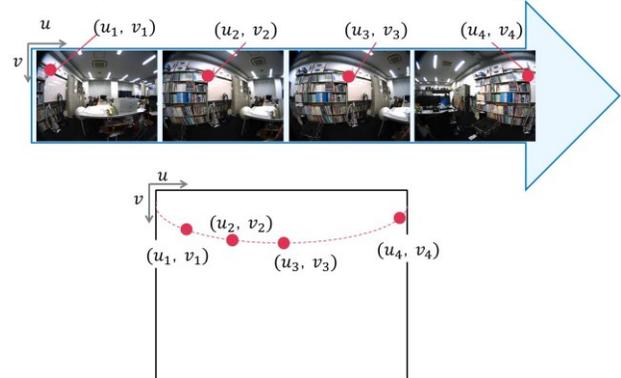


Fig. 12. Capturing feature point trajectories



Fig. 13. Environment of the experiment

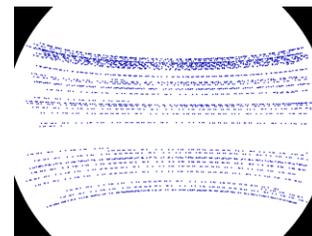


Fig. 14. Feature point trajectories

TABLE III INTRINSIC PARAMETERS

	Proposed method	Existing method
a_0	-415.39	-403.67
a_1	0.00	0.00
$a_2(\times 10^{-4})$	12.70	9.39
$a_3(\times 10^{-7})$	-20.60	-4.42
$a_4(\times 10^{-9})$	3.51	1.29
u_0 [pixel]	548.95	546.94
v_0 [pixel]	386.86	386.58

B. Conditions of intrinsic parameter estimation

Intrinsic parameters $a_0 \sim a_4$ to be initialized are the same values as those in Table I. Image center were calculated for the above method which was explained in section III.C.

C. Results of the experiment

Figure 15 shows the re-projected points obtained by estimated intrinsic parameters. As in Figure 15, the minimization of the evaluation function was generally performed correctly. Also, Table III shows the results of estimated parameters and the existing method.

It is impossible to know the true value of the intrinsic parameters. The evaluation was performed by the perspective projection of the fisheye image using the intrinsic parameters to be estimated. Also, the evaluation was compared with the existing methods. Using Scaramuzza's method [6], the intrinsic parameters were calculated with the same fisheye camera. Twenty images that described the checkerboard pattern were used for estimation. An input image to be converted is shown in Figure 16. The perspective projection, using the results of each technique, is shown in Figure 17. As in Figure 17(a), we can see straight lines in the checkerboard pattern. As compared with Figure 17(b), the distortion to the same extent has been removed as well. Therefore, the intrinsic parameters that were estimated can be considered to be approximately correct. We conducted quantitative evaluation by fitting a line to lattice points of a checkerboard pattern after transforming a fisheye image to a perspective image using the estimated intrinsic parameters. We extracted 10 vertical lines and 7 horizontal lines from Figure 17. Deviation of lattice points are evaluated from the obtained line. The average of standard deviation is 0.35[pixel] and 0.16[pixel] for the proposed method and the conventional method respectively. Although the deviation is sufficiently small even in the proposed method, the estimation accuracy is inferior to the traditional method. That is considered that there is a room for further improvement of accuracy.

The reason distortion remains slightly is thought to be that a position of the installed camera violates presupposition. As for the discrepancy between the optical center and the rotation axis, it is possible to reduce the influence of the discrepancy by measuring distant feature points. Another cause of slight distortion is considered to be the uneven presence of trajectories. Trajectories tracked completely successfully were only used for estimation in this experiment. Accuracy will be further improved by using non-full trajectories; tracking may fails in the middle and then succeed again.

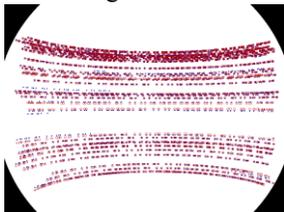
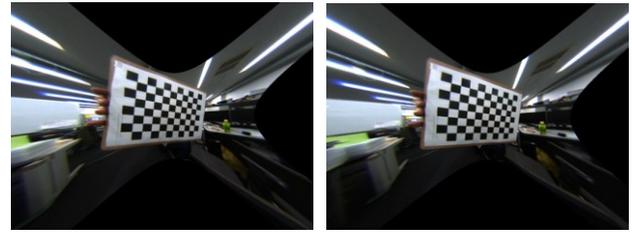


Fig. 15. Re-projected points



Fig. 16. Input image



(a) Proposed Method (b) Conventional Method

Fig. 17. Results of perspective projection

VI. CONCLUSION

This paper proposed a method of estimating fisheye camera's intrinsic parameters without calibration targets. The experiment was conducted using the actual equipment. As a result, the minimization of the evaluation function was performed correctly, and it was possible to obtain an image that removed the distortion in the perspective projection with the estimated parameters. In future works, further accuracy and stability improvements will be achieved by solving the problems: the discrepancy between the optical center and the rotation axis, and the uneven presence of trajectories.

REFERENCES

- [1] Z. Zhang, "A Flexible New Technique for Camera Calibration," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 22, No. 11, pp.1330-1334, 2000.
- [2] P. Carr, Y. Sheikh, and I. Matthews, "Point-less Calibration: Camera Parameters from Gradient-based Alignment to Edge Images," Proceedings of 2012 IEEE Workshop on Applications of Computer Vision (WACV'12), pp.377-384, 2012.
- [3] C. Hughes, P. Denny, M. Glavin, and E. Jones, "Equidistant Fisheye Calibration and Rectification by Vanishing Extraction," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 32, No.12, pp.2289-2296, 2010.
- [4] H. Bakstein and T. Pajdle, "Panoramic Mosaicing with a 180° Field of View Lens," Proceedings of the 3rd Workshop on Omnidirectional Vision (OMNIVIS'02), pp.60-67, 2002.
- [5] L. Heng, B. Li, and M. Pollefeys, "CamOdoCal: Automatic Intrinsic and Extrinsic Calibration of a Rig with Multiple Generic Cameras and Odometry," Proceedings of the 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2013), pp.1793-1800, 2013.
- [6] M. Rufli, D. Scaramuzza, and R. Siegwart, "Automatic Detection of Checkerboards on Blurred and Distorted Images," Proceedings of the 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2008), pp.3121-3126, 2008.
- [7] D. Scaramuzza, A. Martinelli, and R. Siegwart, "A Toolbox for Easily Calibrating Omnidirectional Cameras," Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2006), pp.5695-5701, 2006.
- [8] J. Kannala and S.S. Brandt, "A Generic Camera Model and Calibration Method for Conventional, Wide-angle, and Fisheye Lenses," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.28, No.8, pp.1335-1340, 2006.
- [9] Y. Tanaka, G. Masuyama, and K. Umeda, "Intrinsic Parameter Estimation of a Fisheye Camera Using Trajectories of Feature Points when Rotating the Camera," Proceedings of the 6th International Conference on Advanced Mechatronics (ICAM2015), pp.133-134, 2015.
- [10] M. J. D. Powell, "An Efficient Method for Finding the Minimum of a Function of Several Variables without Calculating Derivatives," Computer Journal, Vol.7, No.2, pp.155-162, 1964.
- [11] P.F. Alcantarilla, J. Nuevo, and A. Bartoli, "Fast Explicit Diffusion for Accelerated Features in Nonlinear Scale Space," Proceedings of British Machine Vision Conference (BMVC), pp.1-11, 2013.