MVA2009 IAPR Conference on Machine Vision Applications, May 20-22, 2009, Yokohama, JAPAN

Generation of Overhead View Images by Estimating Intrinsic and Extrinsic Camera Parameters of Multiple Fish-Eye Cameras

Ryota Okutsu[1], Kenji Terabayashi[1], Yohei Aragaki[2], Noriko Shimomura[2], and Kazunori Umeda[1]

[1]Chuo University
Tokyo, Japan
okutsu@sensor.mech.chuo-u.ac.jp

[2]Nissan Motor Co., Ltd.
Kanagawa, Japan
aragaki@mail.nissan.co.jp

Abstract

In recent years, systems to support driving a car using cameras are increasing. As an example, a system to generate overhead view images is proposed and used to assist drivers. In this paper, we propose a method for generating overhead view images with small errors by estimating intrinsic and extrinsic camera parameters of multiple fish-eye cameras. The accuracy of generated images is evaluated quantitatively by experiments and effectiveness of the proposed method is shown.

1 Introduction

In recent years, systems to support driving a car using cameras are increasing. It is difficult for drivers to perceive the distance with a usual image. As an example, a system to generate overhead view images is proposed[1] and used to assist drivers. This system generates a view of all surrounding area of a car[2][3]. It is important for generating overhead view images with small errors using multiple fish-eye cameras to correctly calculate extrinsic camera parameters.

A fish-eye camera has a wide angle, which is effective for reducing number of cameras. Studies on estimating intrinsic camera parameters of fish-eye cameras[4][5][6] are less frequent than ones of usual pinhole cameras[7], and it is important to estimate intrinsic camera parameters of fish-eye cameras.

In this paper, we propose a method for generating overhead view images with small errors by estimating intrinsic and extrinsic camera parameters of multiple fish-eye cameras.

2 Description of A Fish-Eye Camera

2.1 Projection model

There are several kinds of projection models of a fish-eye lens. Typical projection models are the following two.

$$r_f = \delta\theta \ (EquidistanceProjection) \quad (1)$$
$$r_f = \delta \sin\theta \ (OrthogonalProjection) \quad (2)$$

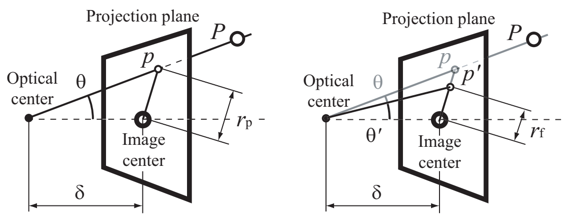
where θ is an angle which a projection line and axis of lens make and r_f is a distance from a projection point to axis of lens. And δ is given by the ratio of focal length f[mm] and pixel size w[mm] as

$$\delta = f/w. \quad (3)$$

In contrast, a perspective projection model of a usual pinhole camera is given by

$$r_p = \delta \tan\theta \ (PerspectiveProjection). \quad (4)$$

Difference between a projection of a pinhole camera and one of a fish-eye lens is shown in Fig.1. In the former case, a point in a scene P is projected on the point p. In the latter case, P is projected on the point p'.



(a) Pinhole (b) Fish-eye

Figure 1: Difference of projection

2.2 Fish-eye camera model

In general, a projection of fish-eye lens does not obey ideal models as eq.(1) and (2). In this section, we describe a fish-eye camera model by referring Nakano et al.[6].

2.2.1 Radial distortion

Typical projection models as eq.(1) and (2) are linear against θ or are sinusoidal function of θ. In this research, we define the fish-eye camera model as the 5th polynomial as follows.

$$r_f = k_1\theta + k_2\theta^2 + k_3\theta^3 + k_4\theta^4 + k_5\theta^5 \quad (5)$$

These parameters k_1, k_2, k_3, k_4, k_5 express radial distortions.

2.2.2 Shift of optical center

Optical axis of a lens does not come to the image center (called shift of optical center). $\boldsymbol{m_f} = [u_f\ v_f]^T$ is a point on a fish-eye image which is captured through a fish-eye camera. $\boldsymbol{c} = [c_u\ c_v]^T$ is the position of optical center on a fish-eye image. Relation between $\boldsymbol{m_f}'$ and $\boldsymbol{m_f}$ on a coordinate system whose origin is $\boldsymbol{c}$ is represented as

$$\boldsymbol{m_f}' = \begin{bmatrix} u_f' \\ v_f' \end{bmatrix} = \begin{bmatrix} \gamma & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_f - c_u \\ v_f - c_v \end{bmatrix} \quad (6)$$

where γ is a ratio of a pixel size of u_f and v_f. In this paper, we assume $\gamma = 1$ because CCD elements are usually square. Parameters of (6) is represented as follows in polar coordinates.

$$r_f = \sqrt{{u_f'}^2 + {v_f'}^2},\ \phi = \arctan\left(\frac{v_f'}{u_f'}\right) \quad (7)$$

2.2.3 Intrinsic parameters of a fish-eye camera

A fish-eye camera is known to have some tangential distortions[8]. In this paper, we do not consider them because they are usually tiny. As a result, the fish-eye camera model is eq.(5) and the intrinsic parameters are represented as

$$\boldsymbol{I} = [k_1\ k_2\ k_3\ k_4\ k_5\ c_u\ c_v]^T. \quad (8)$$

3 Generation of An Integrated Overhead View Image

In this section, a method for generating an integrated overhead view image from fish-eye images is described. First, we convert a fish-eye image into a pinhole image whose model is a perspective projection model. Second, we convert this image into an overhead view image. Finally, we integrate multiple overhead view images and generate an integrated overhead view.

3.1 Convertion to a pinhole image

Converting a fish-eye image into a pinhole image is described by a relation between $\boldsymbol{m_p} = [u_p\ v_p]^T$ and $\boldsymbol{m_f} = [u_f\ v_f]^T$. The former is a point on a pinhole image coordinate system and the latter is a point on a fish-eye image one. A perspective projection model is represented as

$$r_p = \delta_p \tan\theta. \quad (9)$$

And $\boldsymbol{r_p}$ is represented as

$$r_p = \sqrt{u_p^2 + v_p^2}. \quad (10)$$

From eq.(9) and (10), the angle θ is given as follows.

$$\theta = \arctan\left(\frac{\sqrt{u_p^2 + v_p^2}}{\delta_p}\right) \equiv \theta(\boldsymbol{m_p}) \quad (11)$$

Therefore, θ is a function of $\boldsymbol{m_p}$. A direction angle ϕ in a fish-eye image coordinate system which has a compensated shift of optical center is equal to one in a pinhole image coordinate system. Therefore,

$$u_f = \frac{r_f}{r_p} u_p + c_u,\ v_f = \frac{r_f}{r_p} v_p + c_v. \quad (12)$$

By substituting eq.(5),(10) and (11) for eq.(12), the following equation is derived.

$$\begin{aligned} \boldsymbol{m_f} &= \begin{bmatrix} u_f \\ v_f \end{bmatrix} \\ &= \begin{bmatrix} \dfrac{r_f(\boldsymbol{m_p})}{\sqrt{u_p^2 + v_p^2}} u_p + c_u \\ \dfrac{r_f(\boldsymbol{m_p})}{\sqrt{u_p^2 + v_p^2}} v_p + c_v \end{bmatrix} \\ &\equiv \begin{bmatrix} F_u(\boldsymbol{m_p},\ \boldsymbol{I}) \\ F_v(\boldsymbol{m_p},\ \boldsymbol{I}) \end{bmatrix} \end{aligned} \quad (13)$$

This equation shows that $\boldsymbol{m_f}$ is a function of $\boldsymbol{m_p}$.

3.2 Convertion to an overhead view image

Converting a pinhole image into an overhead view image is described by a relation between $\boldsymbol{m_i} = [u_i\ v_i]^T$ and $\boldsymbol{m_p} = [u_p\ v_p]^T$. The former is a point in a overhead view image coordinate system and the latter is one in a pinhole image coordinate system. The axis Z_w is vertical direction on the ground. Positions in a pinhole image coordinate system and in a world coordinate system are given by homogeneous coordinates as follows.

$$\begin{aligned} \tilde{\boldsymbol{m}}_p &= [\ u_p\quad v_p\quad 1\ ]^T \\ \tilde{\boldsymbol{X}}_w &= [\ X_w\quad Y_w\quad Z_w\quad 1\ ]^T \end{aligned}$$

They are related as follows.

$$\tilde{\boldsymbol{m}}_p = \boldsymbol{P}_p \tilde{\boldsymbol{X}}_w \quad (14)$$

$\boldsymbol{P}_p$, a 3×4 matrix, is called a perspective projection matrix. $\boldsymbol{P}_p$ includes extrinsic parameters. Let $\boldsymbol{p}_i$ be the i-column element of the matrix $\boldsymbol{P}_p$. A projected plane is $Z_w = 0$. Therefore eq.(14) is calculated as follows.

$$\begin{aligned} \tilde{\boldsymbol{m}}_p = \boldsymbol{P}_p \tilde{\boldsymbol{X}}_w &= [\ \boldsymbol{p}_1\quad \boldsymbol{p}_2\quad \boldsymbol{p}_3\quad \boldsymbol{p}_4\ ] \begin{bmatrix} X_w \\ Y_w \\ 0 \\ 1 \end{bmatrix} \\ &= [\ \boldsymbol{p}_1\quad \boldsymbol{p}_2\quad \boldsymbol{p}_4\ ] \begin{bmatrix} X_w \\ Y_w \\ 1 \end{bmatrix} = \boldsymbol{H}_p \tilde{\boldsymbol{x}}_w \end{aligned} \quad (15)$$

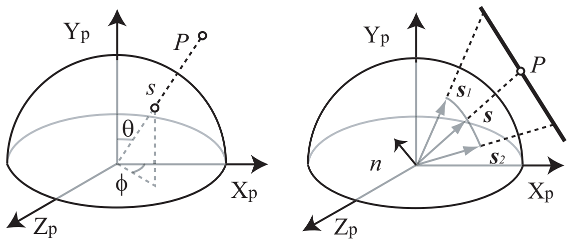
$\boldsymbol{H}_p$ is a 3×3 matrix. Meanwhile, $\tilde{\boldsymbol{m}}_i = [u_i\ v_i\ 1]^T$ is a point in a overhead view image coordinate system. Similarly,

$$\tilde{\boldsymbol{m}}_i = \boldsymbol{H}_i \tilde{\boldsymbol{x}}_w. \quad (16)$$

From eq.(15) and (16), the following equation is satisfied.

$$\tilde{\boldsymbol{m}}_p = \boldsymbol{H}_p {\boldsymbol{H}_i}^{-1} \tilde{\boldsymbol{m}}_i \quad (17)$$

This equation shows that $\boldsymbol{m}_p$ is a function of $\boldsymbol{m}_i$.



(a) Point (b) Line

Figure 2: Sphere model

3.3 Integration of overhead view images

Multiple overhead view images are generated by the method described in section 3.1 and 3.2. They are images captured through a same virtual camera. A point which is obtained on each image comes to a same pixel. Therefore, an integrated overhead view image is generated by calculating an average pixel value of each pixel position on each overhead view image.

4 Estimating Camera Parameters

4.1 Method for estimating intrinsic parameters

We estimate intrinsic parameters of fish-eye cameras using the method proposed by Nakano et al. Linearity of stripe patterns is used in this method[6].

We consider a sphere model whose origin is optical center of a camera. This model and each parameter are shown in Fig.2. A unit position vector of the point in a scene P is given as follows.

$$\boldsymbol{s} = [\ \sin\theta\cos\phi \quad \cos\theta \quad \sin\theta\sin\phi\ ]^T \tag{18}$$

where θ is a projection angle and ϕ is a direction angle. A unit normal vector which crosses perpendicularly with position vectors $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ is represented as

$$\boldsymbol{n} = [\ n_x \quad n_y \quad n_z\ ]^T. \tag{19}$$

From eq.(18) and (19),

$$\boldsymbol{n}\cdot\boldsymbol{s} = n_x\sin\theta\cos\phi + n_y\cos\theta + n_z\sin\theta\sin\phi = 0. \tag{20}$$

The evaluation function is given as follows.

$$\xi_i = \sum_{l=1}^{L}\sum_{i=1}^{P_l}(\boldsymbol{n}_l\cdot\boldsymbol{s}_i)^2 \tag{21}$$

where L is the number of observed straight lines and P_l is the number of points on l-th line. Intrinsic parameters $\boldsymbol{I}$ are estimated by minimizing this function. Gauss-Newton method is used in this estimation. The angle θ is calculated from eq.(5) using Brent's method[9] for observed points $\boldsymbol{m}_{fi} = [u_{fi}\ v_{fi}]^T$ in a fish-eye image coordinate system. By substituting θ for eq.(18), a projection coordinate point $\boldsymbol{s}_i$ is calculated.

4.2 Method for estimating extrinsic parameters

Extrinsic parameters of fish-eye cameras are estimated by the correspondence of $\boldsymbol{m}_{fi} = [u_{fi}\ v_{fi}]^T$ and $\boldsymbol{X}_{wi} = [X_{wi}\ Y_{wi}\ Z_{wi}]^T$. The former is the point P_i which is projected in a fish-eye image coordinate system and the latter is one in a world coordinate system. In a world coordinate system, θ_{cam} is a camera direction, α_{cam} is an elevation angle and β_{cam} is a roll angle. And $[X_{cam},\ Y_{cam},\ Z_{cam}]^T$ is a camera position. Extrinsic parameters are represented as

$$\boldsymbol{E} = [X_{cam}\ Y_{cam}\ Z_{cam}\ \theta_{cam}\ \alpha_{cam}\ \beta_{cam}]^T. \tag{22}$$

These are included in $\boldsymbol{P}_p$ of eq.(14). Let $\boldsymbol{m}_{wi}$ be the back projected point of $\boldsymbol{X}_{wi}$ in a fish-eye image coordinate system. From eq.(13) and (14), $\boldsymbol{m}_{wi}$ is represented as a function of $\boldsymbol{E}$ and $\boldsymbol{X}_{wi}$ as follows.

$$\boldsymbol{m}_{wi} \equiv \boldsymbol{m}_{wi}(\boldsymbol{E},\ \boldsymbol{X}_{wi}) \tag{23}$$

Therefore, the evaluation function becomes as follows.

$$\xi_e = \sum_{i=1}^{N}\{\boldsymbol{m}_{fi} - \boldsymbol{m}_{wi}(\boldsymbol{E},\ \boldsymbol{X}_{wi})\}^2 \tag{24}$$

where N is the number of observed points. Extrinsic parameters $\boldsymbol{E}$ are estimated by minimizing this function. Gauss-Newton method is used in this estimation.

4.3 Simultaneous optimization of camera parameters

As intrinsic parameters $\boldsymbol{I}$ are included in eq.(23), eq.(23) can be represented as

$$\boldsymbol{m}_{wi} \equiv \boldsymbol{m}_{wi}(\boldsymbol{E},\ \boldsymbol{I},\ \boldsymbol{X}_{wi}). \tag{25}$$

In the section 4.1 and 4.2, intrinsic and extrinsic parameters are estimated. By using them as the initial values, camera parameters are updated by minimizing the following equation, which means the simultaneous optimization.

$$\xi_c = \sum_{i=1}^{N}\{\boldsymbol{m}_{fi} - \boldsymbol{m}_{wi}(\boldsymbol{E},\ \boldsymbol{I},\ \boldsymbol{X}_{wi})\}^2 \tag{26}$$

5 Experiment of Generating Overhead View Images

Using the proposed method, we generated overhead view images. We used PointGreyResearch DRAGONFLY2 CCD camera and SPACE TV1634M fish-eye lens. The projection system of this lens is orthogonal projection. Image size is 512[pixel]×384[pixel]. The calibration environment and the world coordinate system are shown in Fig.3. The head of triangle is a camera position and the acute-angled direction is a camera direction θ_{cam}. θ_{cam} was the counter-clockwiseangle from the direction of Y_w axis. The camera position Z_w was set to 200[mm]. Observed points on fish-eye images for estimating extrinsic parameters are cross points on the grid pattern. These were obtained with pixel accuracy.

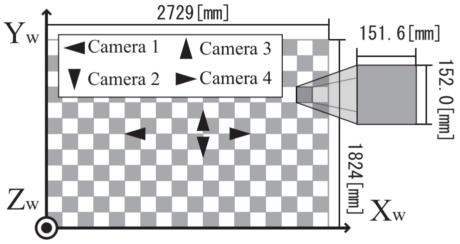


Figure 3: World coordinate system and camera position

5.1 Evaluation of simultaneous optimazation

We evaluated generated images quantitatively to show effectiveness of simultaneous optimization. We captured fish-eye images as shown in Fig.4. We generated an overhead view image with and without the simultaneous optimization in section 4.3. The camera position is camera1 in Fig.3. The elevation angle α_{cam} was set to -60[deg], where minus means the downward direction against the ground level. The roll angle β_{cam} was set to 0[deg].

We evaluated generated images by comparing ideal and obtained cross points on each overhead view image. We obtained 65 points with sub-pixel accuracy.

Images without and with simultaneous optimization are shown in Fig.5 (a) and (b) respectively. Ideal points are shown as dots and obtained points are connected to them by lines. The colors indicate the error size, from red (maximum) to blue (minimum). RMS, maximum and minimum errors are shown in Table 1. Fig.5 (b) with optimization has smaller errors. From Fig.5 and Table 1, it is shown that the simultaneous optimization is effective.

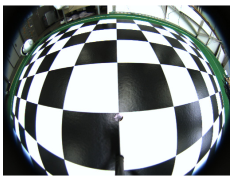
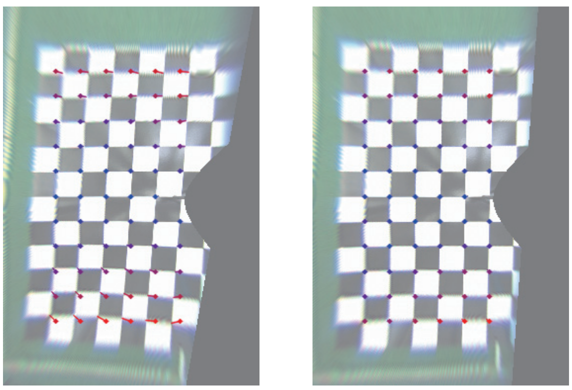


Figure 4: Fish-eye image of Camera1



(a) Without optimization (b) With optimization

Figure 5: Effect of simultaneous optimization

Table 1: Error value of each image [pixel]

	RMS	max	min
without optimization	5.45	9.78	0.09
with optimization	1.80	3.71	0.13

5.2 Generation of an integrated overhead view image

We generated an integrated overhead view image of all surrounding area using four fish-eye cameras. The camera positions are Camera1-4 in Fig.3. Each elevation angle α_{cam} was set to -60[deg]. Simultaneous optimization was conducted.

Overhead view image after integration of four images is shown in Fig.6. We could generate an image of all surrounding area with small errors using the proposed method.

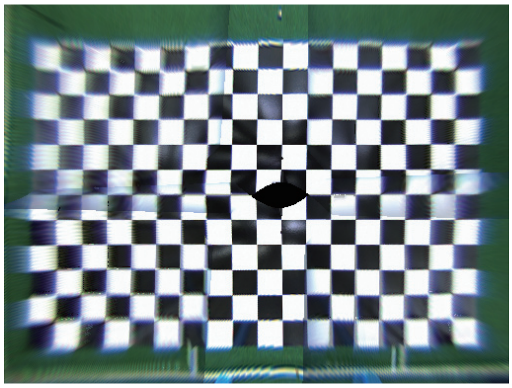


Figure 6: Integrated overhead view image

6 Conclusion

In this paper, we proposed a method for generating overhead view images with small errors by estimating intrinsic and extrinsic parameters of fish-eye cameras. Experimental results showed quantitative evaluation and effectiveness of our method.

As future works, we will consider a fish-eye camera model including tangential distortions and a tilt of optical axis.

References

[1] K.Asari, et al.: "A Practicable Calibration Method for Top View Image Generation," SSII07, IN1-13, 2007.(in Japanese)

[2] K. Oizumi, et al.: "Development of"All-Around View" System," SAE Technical Paper Series, 2003.

[3] Nissan Motor Co. Ltd.: http://www.nissan-global.com /EN/NEWS/2007/_STORY/071012-01-e.html

[4] F.Devernay, et al.: "Straight lines have to be straight," Machine Vision and Applications, Vol.13, pp.14-24,2001.

[5] J. Kannala, et al.: "A Generic Camera Calibration Method for Conventional, Wide-Angle, and Fish-Eye Lenses," IEEE TPAMI, Vol.28 No.8, pp.1335-1340, 2006.

[6] M.Nakano, et al.: "Calibrating Fisheye Camera by Stripe Pattern Based upon Spherical Model," IEICE D, Vol.J90-D, No.1, pp.73-82, 2007.(in Japanese)

[7] Z. Zhang: "A Flexible New Technique for Camera Calibration," IEEE TPAMI, Vol.22, No.11, pp.1330-1334, 2000.

[8] R.Swaminathan, et al.: "Nonmetric Calibratin of Wide-Angle Lenses and Polycameras," IEEE TPAMI, Vol.22, No.10, pp.1172-1178, 2000.

[9] W.H.Press, et al.: "Numerical Recipes in C," Cambridge University Press, Cambridge, 1992.