Extrinsic Parameters Calibration of Multiple Fisheye Cameras in Manhattan Worlds

Weijie Chen*[a], Sarthak Pathak[a], Alessandro Moro[a], Atsushi Yamashita[a], and Hajime Asama[a]

[a]Department of Precision Engineering, The University of Tokyo, Tokyo, Japan

ABSTRACT

With the advantage of having a large field of view, fisheye cameras are widely used in many applications. In order to generate a precise view, calibration of the fisheye cameras is very important. In this paper, we propose a method of extrinsic parameters calibration of multiple fisheye cameras working in man-made structures. A Manhattan Worlds space assumption is used, which describes man-made structures as sets of planes that are either orthogonal or parallel to each other. The orientation of the cameras is obtained by extracting vanishing points that denote orthogonal principal directions in different images captured by the cameras at the same time. With the proposed method, the calibration of extrinsic parameters is very convenient and the system can be recalibrated remotely.

Keywords: fisheye camera, calibration, vanishing points

1. INTRODUCTION

Thanks to the continuing advancements in robotics and automation, applications of mobile robots are rapidly increasing. Mobile robots are very useful in scenes that are dangerous or harmful for human, where we can use the mobile robots rather than risking human's life. Recently, the use of mobile service robots in indoor spaces like offices, homes, etc. has also become popular. Fisheye cameras are more and more widely used on robots with their advantage of a large field of view. For robots that carry multiple fisheye cameras in indoor spaces, it is a common task to generate a bird's-eye view for navigation. In order to generate a precise view, calibration of the cameras is very important. The calibration of the fisheye cameras includes intrinsic and extrinsic parameters calibration. The intrinsic parameters are the property of the camera itself and are independent of the environment or the location. The extrinsic parameters describe the position and orientation between the cameras and determine how the environment is projected on the camera. They will change as long as the relative position or orientation between the cameras changes.

There have been many different types of intrinsic model of fisheye cameras. The most widely used model and calibration method is that of OCamCalib (Omnidirectional Camera Calibration) toolbox, which uses black-and-white checkboard like the calibration of normal cameras.[1] The OCamCalib allows the mapping of every pixel of the 2D fisheye image to a sphere in order to obtain a spherical image. As the intrinsic parameters are independent of the environment or the location, they won't change once calibrated.

As for the extrinsic parameters, there has been research using checkerboards with black and white squares to calibrate the cameras.[2] Such method has a significant shortcoming. If the orientation or position of a camera changes during operation, the operation has to be interrupted and the system needs to be calibrated again. For robots working in indoor spaces, as there are usually many orthogonal or parallel lines in man-made structure, we can take advantage of this to obtain the orientation of the cameras. To solve the problem that has been mentioned, this research proposes a new method of extrinsic parameters calibration of multiple fisheye cameras by extracting vanishing points that denote orthogonal principal directions. With the proposed method, the system can be recalibrated remotely without interruption if the extrinsic parameters change during operation. The vanishing points have useful properties that their directions are independent of the position of the cameras and they are always orthogonal or opposite to each other. Previous research has used the vanishing points in localization.[3] A Manhattan World space assumption is used, which describes the world as a set of planes that

*Email: chen@robot.t.u-tokyo.ac.jp

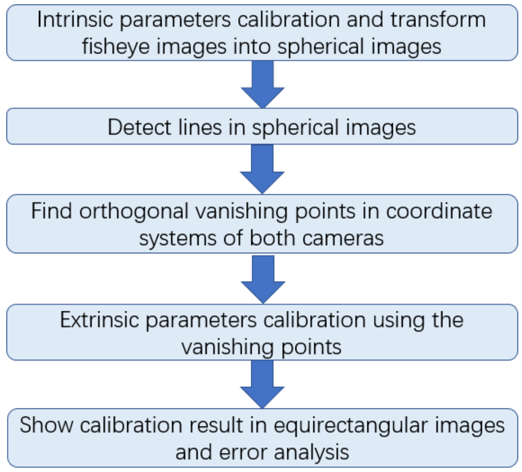


Figure 1. Overview of proposed method

are either orthogonal or parallel to each other.[4] Under that assumption, it's easy to obtain the orientation of the cameras by using information from the environment.

In this research, we work on a system of two fisheye cameras which are assembled to the opposite direction and form a spherical camera. For the cameras we work on, the translation distance between the two fisheye cameras is fixed and very small, so the translation distance is negligible and the extrinsic parameters only include the rotation.

2. OVERVIEW OF PROPOSED METHOD

The overview of the proposed method is shown in Figure 1. In the first step, the intrinsic parameters of both cameras are calibrated by OCamCalib. Both the fisheye images are converted into their respective half-spherical images. In the next step, lines in spherical images are detected and the vanishing points in both images will be extracted from orthogonal sets of lines. Then the extrinsic parameters will be calculated using the vanishing points. In the final step, the images from both cameras are stitched into a full spherical image, and the result will be shown in an equirectangular image.

3. METHODOLOGY

3.1 Mathematical model and intrinsic parameters calibration

The mathematical model of a spherical camera is shown in Figure 2. The raw image output consists of two fisheye images. These images need to be converted into their respective half-spherical images and then form a

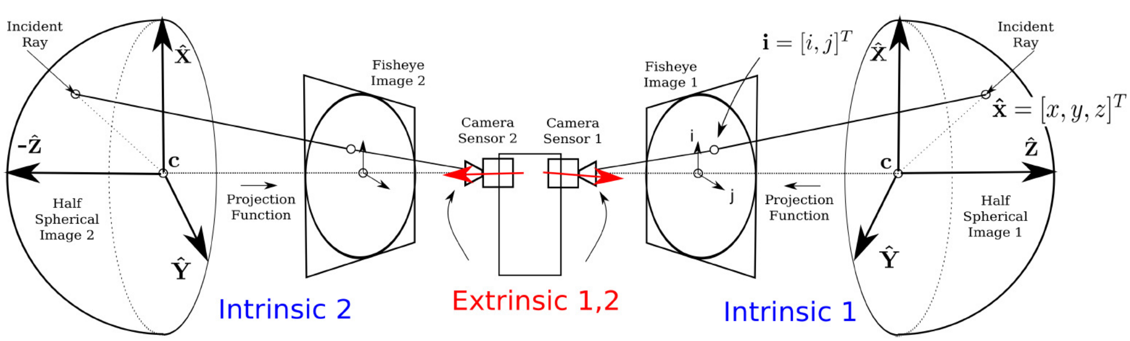


Figure 2. Mathematical model of the spherical camera

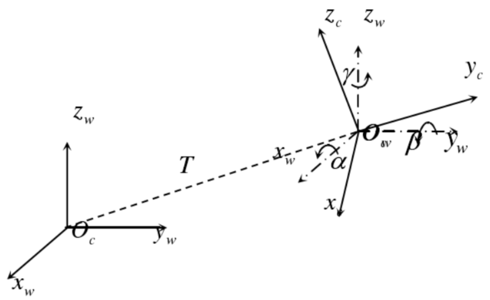


Figure 3. The translation and rotation between two coordinate systems

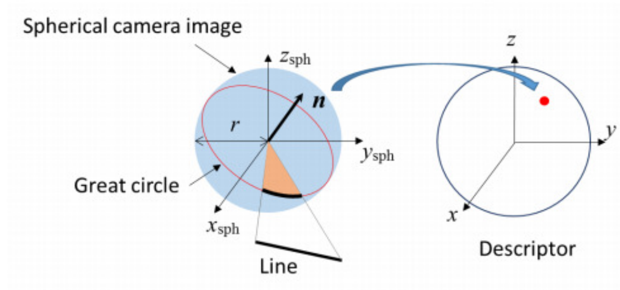


Figure 4. Great circles in the spherical image

full spherical image. The intrinsic parameters of both the cameras and the extrinsic parameters between the two camera coordinate systems need to be calibrated.

The OCamCalib toolbox based on Matlab is used for intrinsic parameters calibration. Determination of the intrinsic parameters allows the mapping of every pixel of the 2D fisheye image to the unit sphere.

Point $[i,j]^T$ on the fisheye image is transformed into a unit vector $[x,y,z]^T$on the unit sphere. The following equation can be written:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} i \\ j \\ f(\rho) \end{bmatrix}, \tag{1}$$

where polynomial function $f(\rho)$ is assumed to be:

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

The degree of the polynomial is adjustable. In this research the degree is chosen as 4 for both the cameras. Parameters a_0, a_1, a_2, a_3, a_4 are the intrinsic parameters and they are calculated by OCamCalib.

3.2 Extrinsic parameters calibration

As shown in Figure 3, $o_c - x_c y_c z_c$ and $o_w - x_w y_w z_w$ are the coordinate systems of the two cameras. The extrinsic parameters include the relative position and orientation between the cameras. As the distance between the two fisheye cameras we work on is fixed and very small, we do not consider the translation distance and the translation vector $\vec{T} = \vec{0}$. The extrinsic parameters only include rotation matrix R.

3.2.1 Line detection in fisheye image

In previous research, it has been shown that 3D lines in the environment are projected as 'great circles' in the spherical image.[5] Each great circle can be represented by a unit vector perpendicular to its plane from the center of the sphere, as shown in Figure 4. The direction of this vector can define a line uniquely. The normal vectors of great circles form a descriptor of the line information inside a spherical image.

Thus, to detect a 3D line uniquely, it only needs to obtain two vectors located on the great circle of that 3D line. The normal vector of the great circle can be calculated by the cross product of the two vectors. Assume vectors $\vec{u}$ and $\vec{w}$ are two vectors located on a great circle, the normal vector $\vec{n}$ of that great circle can be calculated by:

$$\vec{n} = \frac{\vec{u} \times \vec{w}}{||\vec{u} \times \vec{w}||}. \tag{3}$$

In the proposed method, we firstly detect feature points in the fisheye image, and then choose two feature points on the same 3D line by clicking the points with the mouse. This can be done remotely by the operator. Every time the mouse is clicked, the feature point closest to the clicking point will be chosen. The pixel coordinates of the chosen points are $u_0(i_u, j_u)$ and $w_0(i_w, j_w)$. They are transformed into vector $\vec{u}$ and $\vec{w}$ in the unit sphere with (1), thereafter $\vec{n}$ can be calculated. The line that normal vector $\vec{n}$ defines in the fisheye image is shown in Figure 5.

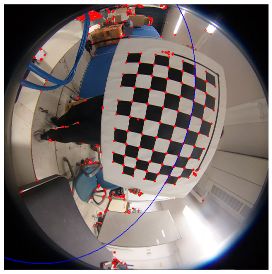


Figure 5. Line detection in fisheye image

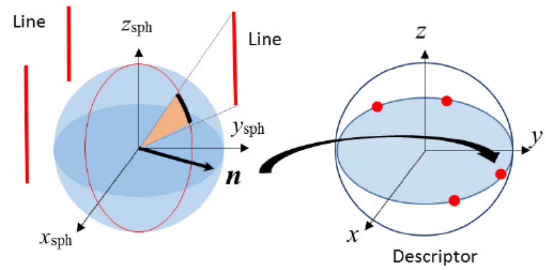


Figure 6. Parallel lines are transformed on the same plane in the descriptor

3.2.2 The vanishing points

In spherical projection, parallel lines converge at vanishing points as they do in perspective projection. As shown in Figure 6, we make use of the fact that unit normal vectors obtained from parallel lines lie on the same plane, and the vector to the vanishing point is the normal vector of that plane. Under the Manhattan world assumption, there should be three such orthogonal planes. The coordinates of a vanishing point are calculated by the cross product of the normal vectors obtained from two lines. Vanishing point $\vec{v}$ is calculated by:

$$\vec{v} = \frac{\vec{n_1} \times \vec{n_2}}{||\vec{n_1} \times \vec{n_2}||}, \tag{4}$$

where $\vec{n_1}$ and $\vec{n_2}$ are normal vectors obtained from two parallel lines.

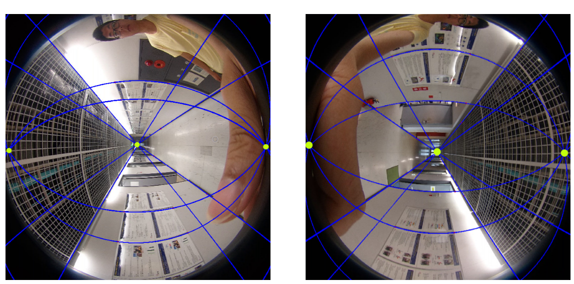
As shown in Figure 7, blue curves are the detected lines in fisheye images and the green points are the vanishing points that parallel lines converge. We can obtain two opposite vanishing points from a set of parallel lines. If there are three sets of orthogonal lines in the image, all three orthogonal vanishing points can be obtained. In case there are only two sets of orthogonal lines, the third can be calculated by the cross product of the two detected vanishing points.

3.2.3 Calculation of rotation matrix

Coordinate systems $o_w - x_w y_w z_w$ and $o_c - x_c y_c z_c$ are the two camera coordinate systems. By extracting the vanishing points that denote the three orthogonal directions, we can obtain their coordinates in both coordinate systems. w_1, w_2 and w_3 are their coordinates in $o_w - x_w y_w z_w$ and c_1, c_2 and c_3 are their coordinates in $o_c - x_c y_c z_c$. w_i and c_i $(i = 1, 2, 3)$ denote the same direction. Thus, the transformation matrix R from $o_w - x_w y_w z_w$ to $o_c - x_c y_c z_c$ can be calculated using the vanishing points via a method of Singular Value Decomposition (SVD).

4. EXPERIMENT RESULT AND ANALYSIS

Camera coordinate systems $o_w - x_w y_w z_w$ is used as the world coordinate system. The three orthogonal vanishing points are w_1, w_2, w_3 in $o_w - x_w y_w z_w$ and they are c_1, c_2, c_3 in $o_c - x_c y_c z_c$. The coordinates of w_1, w_2, w_3 are shown in Table 1 and the coordinates of c_1, c_2, c_3 are shown in Table 2.



(a) Fisheye image 1 (b) Fisheye image 2

Figure 7. Vanishing points in both fisheye images

Table 1. Coordinates of three orthogonal vanishing points in $o_w - x_w y_w z_w$

	$\mathbf{x_w}$	$\mathbf{y_w}$	$\mathbf{z_w}$
w_1	0.996036	0.0811818	0.0362806
w_2	-0.0285451	0.746666	-0.66454
w_3	0.120418	-0.674861	-0.728053

Table 2. Coordinates of three orthogonal vanishing points in $o_c - x_c y_c z_c$

	$\mathbf{x_c}$	$\mathbf{y_c}$	$\mathbf{z_c}$
c_1	-0.992663	0.0916338	-0.0782038
c_2	-0.020476	-0.74552	-0.666169
c_3	0.100764	0.67449	-0.73137

With the SVD method, the rotation matrix R from coordinate system $o_w - x_w y_w z_w$ to $o_c - x_c y_c z_c$ is calculated as:

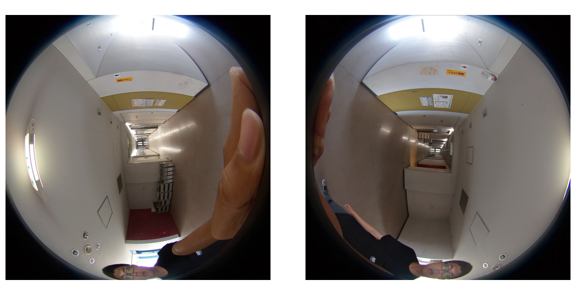
$$R = \begin{bmatrix} 0.999855 & -0.00421668 & 0.0164685 \\ 0.00424719 & 0.999989 & -0.001818 \\ -0.0164606 & 0.00188768 & 0.999863 \end{bmatrix}, \tag{5}$$

We use equirectangular images to represent the result of spherical images in this research. The two original fisheye images are shown in Figure 8. Figure 8 (a) is captured by fisheye camera whose coordinate system is $o_w - x_w y_w z_w$ and Figure 8 (b) is captured by fisheye camera whose coordinate system is $o_c - x_c y_c z_c$. They are transformed into their respective half-spherical images with their intrinsic parameters calibrated. The result of directly stitched spherical image is shown in Figure 9 (a). Obvious shear can be seen in the area where the images are stitched. In Figure 9 (b), the result of stitched spherical image with its half image from Figure 8 (b) rotated by rotation matrix R is shown. It can be seen that with calibration, the stitching is well improved.

The error of each pixel is evaluated by:

$$error_{i,j} = \sqrt{\frac{\Delta R_{i,j}^2 + \Delta G_{i,j}^2 + \Delta B_{i,j}^2}{3}}. \tag{6}$$

The error of directly stitched result without calibration is shown in Figure 10 (a), while the error of stitched result with calibration is shown in Figure 10 (b). The average error with and without calibration is shown in Table 3. It can be seen that with our calibration method, the average error is significantly reduced from 11.16 to 3.41.



(a) Fisheye image 1 (b) Fisheye image 2

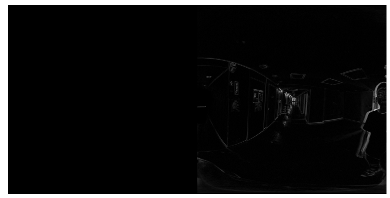
Figure 8. Original fisheye images

(a) Without calibration

(b) With calibration

Figure 9. Stitched equirectangular images

(a) Without calibration



(b) With calibration

Figure 10. Error with and without calibration

5. CONCLUSIONS AND FUTURE WORKS

In this research, a remote calibration method of extrinsic parameters of two fisheye cameras is proposed. A Manhattan Worlds space assumption is used and the orientation of the cameras is obtained by extracting orthogonal vanishing points. Using the vanishing points has the advantage that their directions are independent of the position of the cameras and they are always orthogonal or opposite to each other. With the proposed method, the extrinsic parameters calibration can be done remotely during operation.

As for future work, a more complicated system of four fisheye cameras carried by a mobile robot will be worked on. Since the distance between the cameras is no longer negligible, the extrinsic parameters calibration should include the calibration of both orientations and positions.

Table 3. Average error with and without calibration

	Without calibration	**With calibration**
Average error	11.16	3.41

REFERENCES

[1] Scaramuzza, D., Martinelli, A., and Siegwart, R., "A toolbox for easily calibrating omnidirectional cameras," *Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 5695–5701 (2006).

[2] Li, S., "Full-View Spherical Image Camera," *Proceedings of the 18th International Conference on Pattern Recognition*, Hong Kong, 386–390 (2006).

[3] Goto, T., Pathak, S., Ji, Y., Fujii, H., Yamashita, A., and Asama, H., "Line-based Global Localization of a Spherical Camera in Manhattan Worlds," *Proceedings of the 2018 IEEE International Conference on Robotics and Automation*, Brisbane, 2296–2303 (2018).

[4] Coughlan, J. M. and Yuille, A. L., "The manhattan world assumption: Regularities in scene statistics which enable bayesian inference," *Proceedings of the 13th International Conference on Neural Information Processing Systems*, 809–815 (2000).

[5] Torii, A. and Imiya, A. "The randomized-hough-transform-based method for great-circle detection on sphere," *Pattern Recognition Letters* , 28 (10), 1186–1192 (2007).