

Calibration of Omnidirectional Stereo for Mobile Robots

Yoshiro Negishi, Jun Miura, and Yoshiaki Shirai

Department of Computer-Controlled Mechanical Systems, Osaka University
 {negishi,jun,shirai}@cv.mech.eng.osaka-u.ac.jp

Abstract—This paper describes a calibration method of an omnidirectional stereo system. The system uses a pair of vertically-aligned catadioptric omnidirectional cameras, each of which is composed of a perspective camera and a hyperboloidal mirror, thus providing a single projection point. We divide the calibration into two steps. The first step estimates the image center and the aspect ratio by fitting an ellipse to the mirror boundary in the image. The second step estimates the focal length and the camera pose (position and orientation) including scale by using a calibration pattern and epipolar geometry. Experimental results show the effectiveness of the proposed calibration method.

I. INTRODUCTION

Detection of obstacles and free spaces is an essential function of the vision system for mobile robots. Even if a robot is given a map of the environment, this function is indispensable to cope with unknown, possibly dynamic, obstacles or map errors. Omnidirectional stereo is a suitable sensing method for such vision systems, because it can acquire images and ranges of surrounding areas simultaneously. We have been developing an omnidirectional stereo system and applying it to mobile robot navigation [7], [8], [10], [13], [14].

Our stereo system uses a pair of vertically-aligned omnidirectional cameras (see Fig. 1). Each omnidirectional camera, called HyperOmni Vision [18], uses a hyperbolic mirror with a perspective camera, thus providing a single effective viewpoint at the focal point of the mirror [1]. An input image of an omnidirectional camera is projected onto a cylindrical image plane to generate a panoramic image. A pair of such panoramic images has a nice property that epipolar lines are vertical and in parallel; thus, efficient stereo matching algorithms for the conventional stereo configuration can be applied [5].

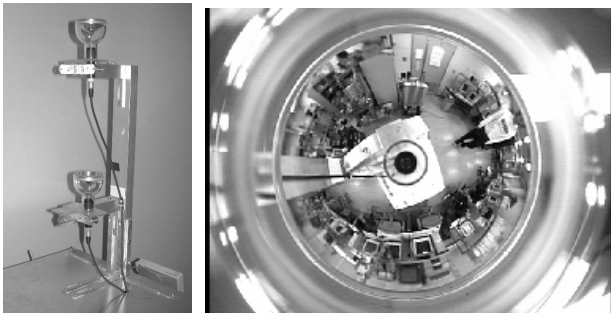


Fig. 1. Stereo setup and an example input image.

In using the omnidirectional stereo system, its calibration is important, as in the case of conventional stereo systems [9], [19]. This paper describes a calibration method of an

omnidirectional system and examines its effectiveness in robot navigation, especially in free space map generation.

There have been many works on the calibration of omnidirectional cameras. Some of them are for estimating intrinsic parameters (including the mirror-camera relationship in catadioptric cameras) [3], [6], [16], while others for estimating extrinsic parameters [2] or epipolar geometry [17]. Mičušík and Pajdla developed methods of calibrating both intrinsic and extrinsic parameters [11], [12]. Geyer and Daniilidis [4] developed a method for rectifying omnidirectional image pairs, generating a rectified pair of normal perspective images.

In this paper, we develop a calibration method which eventually generates a pair of *rectified* cylindrical images whose epipolar lines are vertical and in parallel. The proposed method calibrates both intrinsic parameters (aspect ratio, image center, and focal length) and extrinsic parameters (6D pose) of omnidirectional cameras. Like [11], we first estimate the image center and the aspect ratio by fitting an ellipse to the mirror boundary in the image. We then estimate the focal length and the camera pose including scale using a calibration pattern as well as epipolar geometry.

II. CAMERA MODEL

The omnidirectional camera we use is composed of a hyperboloidal mirror with a perspective camera (see Fig. 2). From this geometry, we obtain the following relationship between scene position (X, Y, Z) and image position (x, y) :

$$\begin{aligned} x &= \frac{Xf(b^2 - c^2)}{(b^2 + c^2) \cdot (Z - c) - 2bc\sqrt{(Z - c)^2 + X^2 + Y^2}}, \\ y &= \frac{Yf(b^2 - c^2)}{(b^2 + c^2) \cdot (Z - c) - 2bc\sqrt{(Z - c)^2 + X^2 + Y^2}}, \quad (1) \\ &\quad (c = \sqrt{a^2 + b^2}) \end{aligned}$$

where a and b are the parameters of the mirror shape; c is the half of the distance between the focal points of the mirror and the camera; f is the focal length of the camera. The origin of the camera coordinates is set at the right midpoint of the two focal points. We hereafter represent this projection as:

$$\mathbf{x} = F(\mathbf{X}; f), \quad (2)$$

where $\mathbf{x} = (x, y)$ is a 2D image position and $\mathbf{X} = (X, Y, Z)$ is a 3D scene position. Since we assume the mirror shape parameters, a and b (and thus c), are known and correct, projection F is parameterized only by the focal length f , which will be estimated.

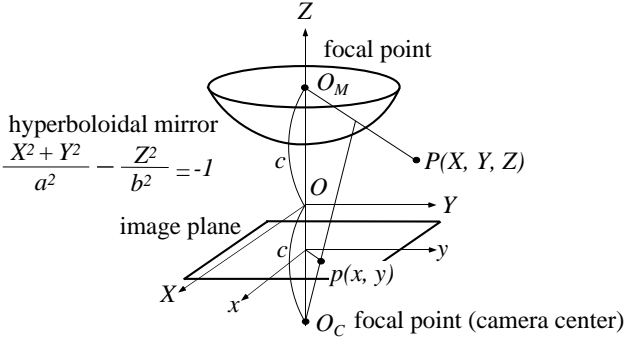


Fig. 2. Geometry of hyperboloidal projection [18].

III. CALIBRATION OF A SINGLE OMNIDIRECTIONAL CAMERA

This section describes a method of estimating some of intrinsic parameters and extrinsic parameters of a single omnidirectional camera.

A. Parameters to be estimated

We assume that the symmetry axis of the mirror and the optical axis of the camera are perfectly aligned, and that the mirror shape parameters are known and correct. Under these assumptions, we estimate the following parameters:

- image center, (C_x, C_y) .
- aspect ratio, k .
- focal length, f .
- camera pose, R and t .

The calibration of an omnidirectional camera consists of two steps. We first estimate the image center and aspect ratio of the camera using the image of the mirror boundary [11]. Once these parameters are estimated, we then estimate the focal length and the camera pose using a calibration pattern with known metric information.

B. Estimating Image Center and Aspect Ratio

The image center (C_x, C_y) is the intersection of the optical axis with the image plane. We assume that the image plane is perpendicular to the axis, and that the only cause that makes the aspect ratio deviate from one is the digitization error of the CCD array and the image capture board (i.e., the difference of the horizontal and the vertical digitization frequency). So the boundary of the mirror can be represented by (see Fig. 3):

$$\frac{(x - C_x)^2}{A^2} + \frac{(y - C_y)^2}{B^2} = 1. \quad (3)$$

Four parameters, A, B, C_x, C_y will be estimated. By putting a black plate on the top of the backside of the mirror, the mirror boundary in the image is at the edge points changing from black to gray in the radial direction. The above ellipse equation is fitted to the extracted boundary points.

In the case of our system, the vertical axis of the ellipse becomes longer than the horizontal one (i.e., $B > A$). Let F_+ and F_- be the two foci of the ellipse; their positions are $F_+(C_x, C_y + \sqrt{B^2 - A^2})$ and $F_-(C_x, C_y - \sqrt{B^2 - A^2})$.

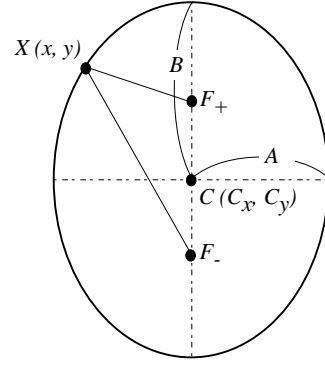


Fig. 3. Ellipse and its two foci.

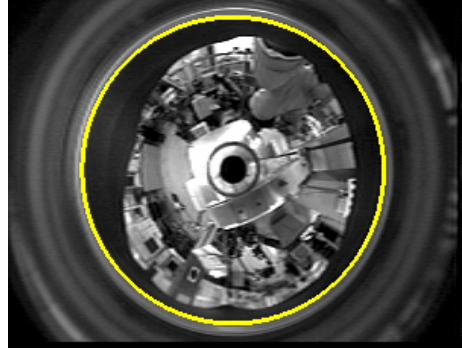


Fig. 4. Ellipse fitting to the lower camera image.

Since a point $X(x, y)$ on the ellipse satisfies $\overline{XF_+} + \overline{XF_-} = 2B$, we obtain

$$\begin{aligned} & \sqrt{(C_x - x)^2 + (C_y - \sqrt{B^2 - A^2} - y)^2} \\ & + \sqrt{(C_x - x)^2 + (C_y + \sqrt{B^2 - A^2} - y)^2} = 2B. \end{aligned} \quad (4)$$

For a set of extracted boundary points $\{(x_i, y_i)\}$ ($i = 1, \dots, N$), the squared error χ_c^2 is defined as:

$$\chi_c^2 = \sum_{i=1}^N d_i^2, \quad (5)$$

$$\begin{aligned} d_i &= \sqrt{(C_x - x_i)^2 + (C_y - \sqrt{B^2 - A^2} - y_i)^2} \\ & + \sqrt{(C_x - x_i)^2 + (C_y + \sqrt{B^2 - A^2} - y_i)^2} - 2B. \end{aligned} \quad (6)$$

We then apply the Levenberg-Marquardt method [15] to search for the parameter values which minimizes this squared error.

Fig. 4 shows the result of ellipse fitting to the lower camera image. The calibration took about 0.03 [sec] in 15 iterations with about 800 edge points. The estimated parameters in this case are: $A = 97.735$, $B = 100.590$, $(C_x, C_y) = (174.415, 121.805)$ and the aspect ratio $A/B = 0.9716$.

C. Estimating Focal Length and Camera Pose

We use a calibration pattern to estimate the focal length and the camera pose such that the squared error of the feature positions predicted from these parameters and the observed feature positions in the image is minimized (see Fig. 5). Let P_i (at \mathbf{X}_i) ($i = 1, \dots, M$) be a set of feature points on the calibration pattern, and let p_i^* (at \mathbf{x}_i^*) and

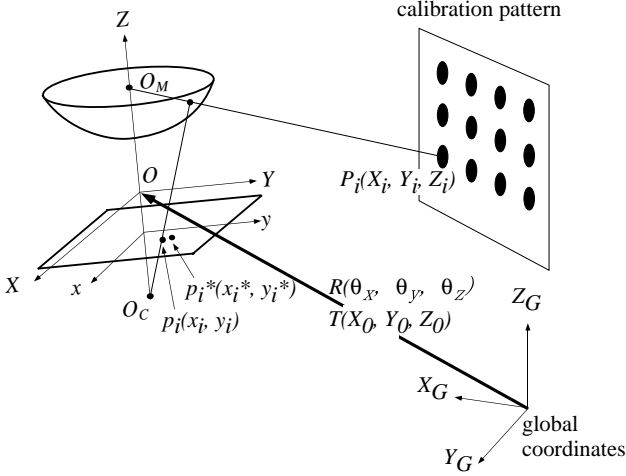


Fig. 5. Calibration using a pattern.

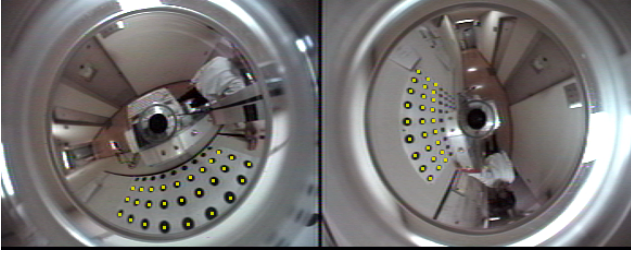


Fig. 6. The calibration pattern and the calibration result.

p_i (at \mathbf{x}_i) be the predicted and the observed image position of the i th feature, respectively.

We minimize the following squared error χ_{fp}^2 :

$$\chi_{fp}^2 = \sum_{i=1}^M \|\mathbf{x}_i^* - \mathbf{x}_i\|^2. \quad (7)$$

We represent the camera pose by position vector (X_0, Y_0, Z_0) of the camera coordinates and rotation angles $(\theta_X, \theta_Y, \theta_Z)$; the relationship between a scene point \mathbf{X} in the camera coordinates and its position in the global coordinates \mathbf{X}_G is given by:

$$\mathbf{X}_G = T(X_0, Y_0, Z_0)R(X, \theta_X)R(Y, \theta_Y)R(Z, \theta_Z)\mathbf{X} \quad (8)$$

$$= \begin{pmatrix} R & T \\ 0 & 1 \end{pmatrix} \mathbf{X} = H\mathbf{X}. \quad (9)$$

Since the predicted position p_i^* of each feature is calculated from its scene position P_i , the above global transformation from the world to the camera coordinates (eq. (9)), and from the projection function (eq. (2)), the squared error χ_{fp}^2 becomes the function of seven parameters of the focal length and the camera pose:

$$\chi_{fp}^2(f, X_0, Y_0, Z_0, \theta_x, \theta_y, \theta_z) = \sum_{i=1}^M \|F(H^{-1}(X_0, Y_0, Z_0, \theta_X, \theta_Y, \theta_Z)\mathbf{X}_i; f) - \mathbf{x}_i\|^2. \quad (10)$$

We again apply a Levenberg-Marquardt method to search for the parameter values which minimizes the above squared error. Fig. 6 shows a result obtained by independently calibrating two cameras. Black circles put on a wall

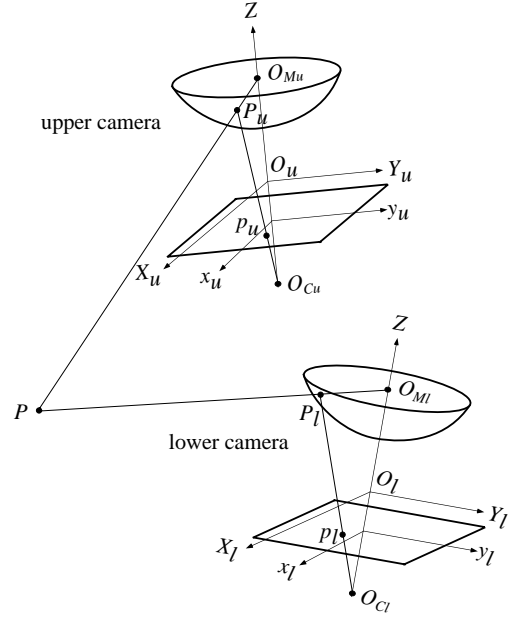


Fig. 7. Epipolar geometry in omnidirectional stereo.

are used as features. The position of each circle in the image is obtained by binarizing the input image and then calculating the centroid of the corresponding black region. In the figure, the predicted positions of the features are backprojected on the input images.

IV. SIMULTANEOUS CALIBRATION OF A PAIR OF OMNIDIRECTIONAL CAMERAS USING EPIPOLAR GEOMETRY AND A CALIBRATION PATTERN

Estimation of the relative pose between two cameras from their independently-calibrated (absolute) poses is possible but the estimation result may not be accurate enough for stereo calculation, due to the additive characteristic of the estimation errors. So we additionally use the epipolar geometry for the cameras [2], [11], [12], [17] to improve the accuracy.

We here follow the discussion by Chang and Hebert [2]. Let 3D point P be projected to p_u and p_l in the upper and the lower image through point P_u and P_l on the corresponding mirror, respectively (see Fig. 7). Since $O_{Mu}, O_{Ml}, P, P_u, P_l$ are coplanar, we obtain the following epipolar constraint:

$$\mathbf{X}_u^t E \mathbf{X}_l = 0, \quad (11)$$

where \mathbf{X}_u and \mathbf{X}_l are the position of P_u and P_l in the local coordinate systems centered at O_{Mu} and O_{Ml} , respectively, and $E = [\mathbf{t}]_{\times} R$ is the essential matrix encoding the relative pose (rotation R and translation \mathbf{t}) between the two local coordinate systems. Since the relative pose is parameterized by a pair of the absolute poses, we obtain:

$$E = E(X_0^u, Y_0^u, Z_0^u, \theta_X^u, \theta_Y^u, \theta_Z^u, X_0^l, Y_0^l, Z_0^l, \theta_X^l, \theta_Y^l, \theta_Z^l), \quad (12)$$

where superscripts indicate the camera.

Considering the geometry of hyperbolic projection, the points on the mirrors, P_u (at \mathbf{X}_u) and P_l (at \mathbf{X}_l), are

represented as the function of their image points, p_u (at \mathbf{x}_u) and p_l (at \mathbf{x}_l), respectively. Let the mapping from an image point to a mirror point, which is parameterized by the focal length, be:

$$\mathbf{X} = I(\mathbf{x}; f). \quad (13)$$

Then we have the following epipolar geometry constraint on a pair of image points, \mathbf{x}^u and \mathbf{x}^l :

$$d_e(\mathbf{x}^u, \mathbf{x}^l) = I(\mathbf{x}^u; f^u)^t E(X_0^u, \dots, \theta_Z^u, X_0^l, \dots, \theta_Z^l) I(\mathbf{x}^l; f^l). \quad (14)$$

For L pairs of matched points $\{\mathbf{x}_i^u, \mathbf{x}_i^l\}$ ($i = 1, \dots, L$), We define the squared error χ_e^2 as:

$$\chi_e^2 = \sum_{i=1}^L \{d_e(\mathbf{x}_i^u, \mathbf{x}_i^l)\}^2. \quad (15)$$

In order to simultaneously calibrate the whole stereo system, we use the following squared error χ^2 :

$$\chi^2 = \chi_{fp^u}^2 + \chi_{fp^l}^2 + w_e \chi_e^2, \quad (16)$$

where $\chi_{fp^u}^2$ and $\chi_{fp^l}^2$ are the squared error for features on the calibration pattern for the upper and the lower camera, respectively, and w_e is a weight to balance the first two and the last squared errors; actually, the weight is determined such that it converts the error in the epipolar constraint (d_e) in eq. (14) into the distance in the image at an average position of the feature points used.

We first tried to estimate all 14 parameters (one focal length and six pose parameters for each camera) using this squared error; but the estimation result was not satisfactory. This is probably due to a large uncertainty in the vertical position and the focal lengths when estimating them from the epipolar geometry constraint, which arises from the specific camera placement of our system. That is, if the two axes of the omnidirectional cameras are completely aligned, all epipolar lines become radial lines in the image and, therefore, we cannot determine the relative vertical distance and the focal lengths from the epipolar constraint (eq. (11)). Although the cameras are not completely aligned actually, the uncertainty of these values could be large.

We, therefore, take the following two step approach. In the first step, we estimate the focal length and the pose of each camera independently, using the method described in Sec. III-C. Then we fix the values of the focal lengths and the vertical positions of both cameras and estimate the rest 10 parameters (i.e., $X_0, Y_0, \theta_X, \theta_Y, \theta_Z$ for each camera) by a Levenberg-Marquardt method using the squared error in eq. (16).

V. PANORAMIC PROJECTION

To adopt standard stereo matching algorithms, we would like to have a pair of panoramic images which has only parallel epipolar lines. Usually we set a cylindrical image plane for each camera to get a pair of desired panoramic images. This works well as long as the two axes of the cameras are aligned.

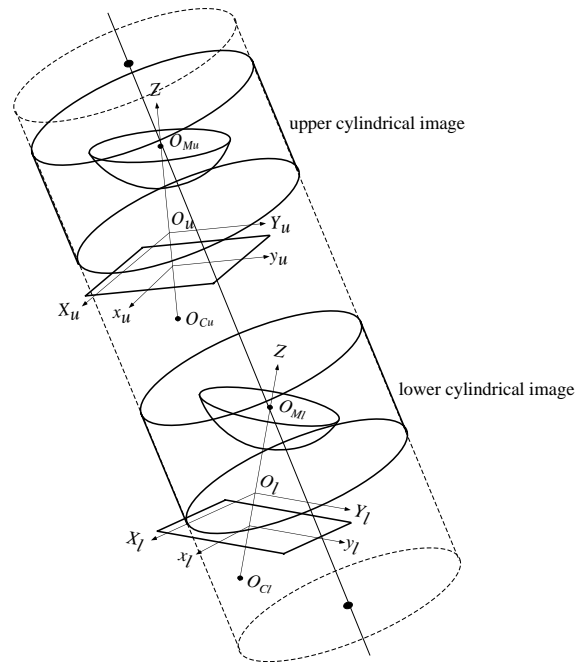


Fig. 8. Cylindrical projection.

In the case where the axes of two cameras are not aligned, however, this cylindrical projection should not be done independently for each camera. To get a *rectified* pair of cylindrical images we need to set a cylindrical image plane whose axis is connecting the two effective focal points of the cameras as shown in Fig. 8. The geometry of this cylindrical image plane can be calculated by the estimated extrinsic parameters (the poses of the cameras).

VI. EXPERIMENTAL RESULTS

We used 27 feature points on the calibration pattern (shown in Fig. 6) for each camera, and other 26 feature points for epipolar geometry evaluation. Starting from the parameters for the ideal camera placements, the calibration of seven parameters for each camera using the calibration pattern took a few seconds, and the calibration using both the pattern and epipolar geometry took another few seconds using Athlon 2200+.

To test the feasibility of the calibration method, we intentionally misaligned the cameras and performed the calibration. The estimated parameters are listed in Table I. Fig. 9 shows a typical calibration result. Fig. 9(a) compares three pairs of panoramic images; the upper pair is generated by assuming the complete alignment of the cameras, the middle pair is generated using the parameters estimated by the *independent* calibration method (Sec. III-C), and the lower pair is generated using the parameters estimated by the *simultaneous* calibration method (Sec. IV).

We first compare the upper (no calibration) and the lower (simultaneous calibration) pairs. It is clear that at many places (e.g., direction A), the vertical lines are slanted in the upper pair. We also compared the measured distance by the calibrated stereo with that by a laser range finder (see Fig. 9(c)). For example, in the direction B, the distance by the range finder is 190 [cm] and that by the

Table I
ESTIMATED PARAMETERS.

(a) Parameters obtained by independently calibrating the two cameras.

| | X_0 [mm] | Y_0 [mm] | Z_0 [mm] | θ_X [deg] | θ_Y [deg] | θ_Z [deg] | f [pixel] |
|-------|------------|------------|------------|------------------|------------------|------------------|-----------|
| upper | 364.49 | -322.05 | 1554.36 | -3.26 | -1.90 | 78.93 | 179.235 |
| lower | 367.59 | -320.47 | 1255.15 | 6.02 | -0.80 | -25.76 | 179.856 |

(b) Parameters obtained by calibration using epipolar geometry.

| | X_0 [mm] | Y_0 [mm] | Z_0 [mm] | θ_X [deg] | θ_Y [deg] | θ_Z [deg] | f [pixel] |
|-------|------------|------------|------------|------------------|------------------|------------------|-----------|
| upper | 375.88 | -321.65 | 1554.36 | -2.71 | -2.86 | 77.47 | 179.235 |
| lower | 364.48 | -318.82 | 1255.15 | 5.88 | -0.68 | -25.25 | 179.856 |

simultaneously calibrated stereo is about 195 [cm], but that by the uncalibrated stereo is about 150 [cm]. We compared the distances at many directions and confirmed that the calibrated stereo can generate reasonable range information for free space modeling.

We then compare the middle (independent calibration) and the lower (simultaneous calibration) pairs. The measured distances are mostly correct in both cases, but due to a larger directional difference (i.e., difference in the horizontal position in panoramic images), the number of pixels whose disparities are obtained is smaller in the independent calibration than in the simultaneous one (3520 and 4453 in this case).

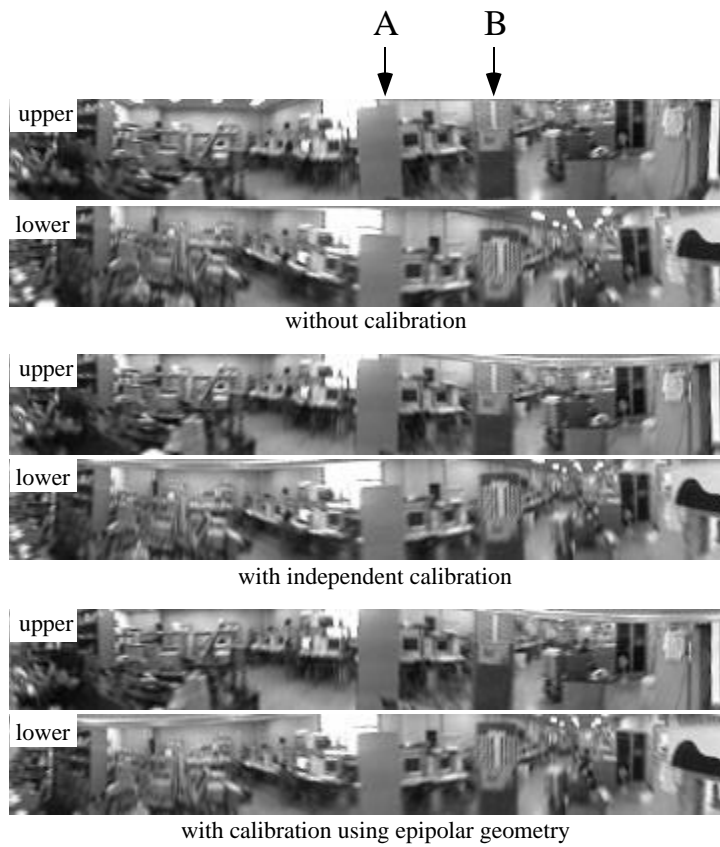
We applied the omnidirectional stereo system calibrated by the proposed method to our map generation method [10], [14] and succeeded in safe navigation in an unknown environment.

VII. SUMMARY

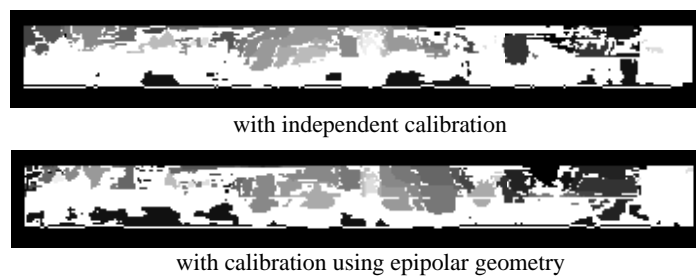
This paper has presented a method of calibrating an omnidirectional stereo system. We first estimate the image center and the aspect ratio of each camera using the mirror boundary in the image. Then, we estimate the focal lengths and the poses of both cameras using both a calibration pattern and epipolar geometry. These estimations are done by an ordinary non-linear minimization scheme (i.e., Levenberg-Marquardt) without any special tunings; nevertheless, the performance is satisfactory for the purpose of robot navigation.

REFERENCES

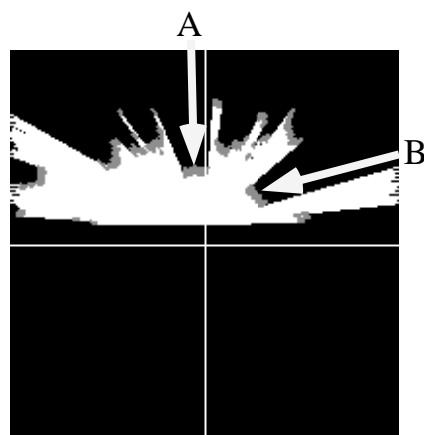
- [1] S. Baker and S.K. Nayar. A Theory of Single-Viewpoint Catadioptric Image Formation. *Int. J. of Computer Vision*, Vol. 35, No. 2, pp. 175–196, 1999.
- [2] P. Chang and M. Hebert. Omni-directional Structure from Motion. In *Proceedings of IEEE Workshop on Omnidirectional Vision*, pp. 127–133, 2000.
- [3] C. Geyer and K. Daniilidis. Catadioptric Camera Calibration. In *Proceedings of the 7th Int. Conf. on Computer Vision*, Vol. 1, pp. 398–404, 1999.
- [4] C. Geyer and K. Daniilidis. Conformal Rectification of Omnidirectional Stereo Pairs. In *Proceedings of IEEE Workshop on Omnidirectional Vision and Camera Networks*, 2003.
- [5] J. Gluckman, S.K. Nayar, and K.J. Thoresz. Real-Time Omnidirectional and Panoramic Stereo. In *Proceedings of 1998 DARPA Image Understanding Workshop*, 1998.
- [6] S.B. Kang. Catadioptric Self-Calibration. In *Proceedings of 2000 IEEE Conf. on Computer Vision and Pattern Recognition*, Vol. 1, pp. 201–207, 2000.
- [7] H. Koyasu, J. Miura, and Y. Shirai. Realtime Omnidirectional Stereo for Obstacle Detection and Tracking in Dynamic Environments. In *Proceedings of the 2001 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 31–36, 2001.
- [8] H. Koyasu, J. Miura, and Y. Shirai. Mobile Robot Navigation in Dynamic Environments using Omnidirectional Stereo. In *Proceedings of 2003 IEEE Int. Conf. on Robotics and Automation*, pp. 893–898, 2003.
- [9] Q.T. Luong and O.D. Faugeras. The Fundamental Matrix: Theory, Algorithms, and Stability Analysis. *Int. J. of Computer Vision*, Vol. 17, No. 1, pp. 43–76, 1996.
- [10] J. Miura, Y. Negishi, and Y. Shirai. Mobile Robot Map Generation by Integrating Omnidirectional Stereo and Laser Range Finder. In *Proceedings of 2002 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 250–255, 2002.
- [11] B. Mičušík and T. Pajdla. Estimation of Omnidirectional Camera Model from Epipolar Geometry. In *Proceedings of 2003 IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 485–490, 2003.
- [12] B. Mičušík and T. Pajdla. Omnidirectional Camera Model and Epipolar Geometry Estimation by RANSAC with Bucketing. In *Proceedings of the 13th Scandinavian Conf. on Image Analysis*, pp. 83–90, 2003.
- [13] Y. Negishi, J. Miura, and Y. Shirai. Adaptive Robot Speed Control by Considering Map and Localization Uncertainty. In *Proceedings of the 8th Int. Conf. on Intelligent Autonomous Systems*, pp. 873–880, 2004.
- [14] Y. Negishi, J. Miura, and Y. Shirai. Mobile Robot Navigation in Unknown Environments Using Omnidirectional Stereo and Laser Range Finder. In *Proceedings of 2004 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2004. (to appear).
- [15] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing, 2nd Edition*. Cambridge University Press, New York, NY, 1992.
- [16] D. Strelow, J. Mishler, D. Koes, and S. Singh. Precise Omnidirectional Camera Calibration. In *Proceedings of 2001 IEEE Conf. on Computer Vision and Pattern Recognition*, Vol. 1, pp. 689–694, 2001.
- [17] T. Svoboda, T. Pajdla, and V. Hlaváč. Epipolar Geometry for Panoramic Cameras. In *Proceedings of 5th European Conf. on Computer Vision*, pp. 218–232, 1998.
- [18] K. Yamazawa, Y. Yagi, and M. Yachida. Omnidirectional Imaging with Hyperboloidal Projection. In *Proceedings of 1993 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 1029–1034, 1993.
- [19] Z. Zhang, O. Faugeras, and R. Deriche. An Effective Technique for Calibrating a Binocular Stereo Through Projective Reconstruction Using Both a Calibration Object and the Environment. *Videre: J. of Computer Vision Research*, Vol. 1, No. 1, pp. 58–68, 1997.



(a) panoramic images



(b) disparity images



(c) scan of laser range finder

Fig. 9. Evaluation of the calibration result.