<table>
<thead>
<tr>
<th>Title</th>
<th>Calibration of omnidirectional stereo for mobile robots</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Negishi, Yoshiro; Miura, Jun; Shirai, Yoshiaki</td>
</tr>
<tr>
<td>Issue Date</td>
<td>2004-09</td>
</tr>
<tr>
<td>oaire:version</td>
<td>VoR</td>
</tr>
<tr>
<td>URL</td>
<td><a href="https://hdl.handle.net/11094/14088">https://hdl.handle.net/11094/14088</a></td>
</tr>
<tr>
<td>rights</td>
<td>©2004 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.</td>
</tr>
<tr>
<td>Note</td>
<td></td>
</tr>
</tbody>
</table>
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 [17].

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 [11]. 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.](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) [3], [6], [16], while others for estimating extrinsic parameters [2] or epipolar geometry [17]. Miščik 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 \((x, y)\):

\[
\begin{align*}
X &= \frac{xf(b^2 - c^2)}{(b^2 + c^2) \cdot (Z - c) - 2bc \sqrt{(Z - c)^2 + X^2 + Y^2}}, \\
y &= \frac{yf(b^2 - c^2)}{(b^2 + c^2) \cdot (Z - c) - 2bc \sqrt{(Z - c)^2 + X^2 + Y^2}}, \\
c &= \sqrt{a^2 + b^2}
\end{align*}
\]

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:

\[
x = F(X; f),
\]

where \(x = (x, y)\) is a 2D image position and \(X = (X, Y, Z)\) 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.
III. CALIBRATION OF A SINGLE OMNIDIRECTIONAL CAMERA

This section describes a method of estimating some of the 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, \(A\).
- 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. 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):

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

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 \(F_-(C_x, C_y - \sqrt{B^2 - A^2})\).

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

\[
\sqrt{(C_x - x)^2 + (C_y - \sqrt{B^2 - A^2} - y)^2} + \sqrt{(C_x - x)^2 + (C_y + \sqrt{B^2 - A^2} - y)^2} = 2B. \tag{4}
\]

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

\[
\chi^2 = \sum_{i=1}^{N} d_i^2, \tag{5}
\]

\[
d_i = \sqrt{(C_x - x_i)^2 + (C_y - \sqrt{B^2 - A^2} - y_i)^2} + \sqrt{(C_x - x_i)^2 + (C_y + \sqrt{B^2 - A^2} - y_i)^2} - 2B. \tag{6}
\]

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 \(F_i(x_i, y_i) (i = 1, \ldots, M)\) be a set of feature points on the calibration pattern, and let \(p_i^* (at\ ax^*_i)\) and...
Let \( p_i \) be the predicted and the observed image position of the \( i \)th feature, respectively.

We minimize the following squared error \( \chi^2_{fp} \):

\[
\chi^2_{fp} = \sum_{i=1}^{M} \| x^*_i - x_i \|^2. 
\]  

(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 \( X \) in the camera coordinates and its position in the global coordinates \( X_G \) is given by:

\[
X_G = T(X_0, Y_0, Z_0)R(X, \theta_X)R(Y, \theta_Y)R(Z, \theta_Z)X \quad (8)
\]

\[
= \begin{pmatrix} R & T \\ 0 & 1 \end{pmatrix} X = HX. 
\]  

(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 \( \chi^2_{fp} \) becomes the function of seven parameters of the focal length and the camera pose:

\[
\chi^2_{fp}(f, X_0, Y_0, Z_0, \theta_x, \theta_y, \theta_z) \]

\[
= \sum_{i=1}^{M} \| F(H^{-1}(X_0, Y_0, Z_0, \theta_x, \theta_y, \theta_z)X_i; f) - x_i \|^2. 
\]  

(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 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 \( O_{Mu}, O_{Ml}, P_u, P_l \) are coplanar, we obtain the following epipolar constraint:

\[
X_u^T E X_l = 0, 
\]  

(11)

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

\[
E = E(X_u^0, Y_u^0, Z_u^0, \theta_X^0, \theta_Y^0, \theta_Z^0, X_l^0, Y_l^0, Z_l^0, \theta_X^1, \theta_Y^1, \theta_Z^1), 
\]  

(12)

where superscripts indicate the camera.

Considering the geometry of hyperbolic projection, the points on the mirrors, \( P_u \) (at \( X_u \)) and \( P_l \) (at \( X_l \)), are
represented as the function of their image points, \( p_u \) (at \( x_u \)) and \( p_l \) (at \( x_l \)), respectively. Let the mapping from an image point to a mirror point, which is parameterized by the focal length, be:

\[
X = I(x; f). 
\]

Then we have the following epipolar geometry constraint on a pair of image points, \( x_u \) and \( x_l \):

\[
d_e(x_u, x_l) = I(x_u^f; f^u)E(X_0^u, \ldots, \theta_2^u, X_0^l, \ldots, \theta_2^l)I(x_l^f; f^l). 
\]

For \( L \) pairs of matched points \( \{x_u^i, x_l^i\} \) (\( i = 1, \ldots, L \)), we define the squared error \( \chi^2 \) as:

\[
\chi^2 = \sum_{i=1}^{L} \{d_e(x_u^i, x_l^i)\}^2. 
\]

In order to simultaneously calibrate the whole stereo system, we use the following squared error \( \chi^2 \):

\[
\chi^2 = \chi^2_p + \chi^2_f + w_e \chi^2_e, 
\]

where \( \chi^2_p \) and \( \chi^2_f \) 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.
simultaneously calibrated stereo is about 195 [mm], but that
by the uncalibrated stereo is about 150 [mm]. 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 mea-
sured 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
by the proposed method to our map generation method
scheme (i.e., Lweuberg-Marquardt) without any special
motions are done by an ordinary non-linear minimization
focal lengths and the poses of both cameras using both
omnidirectional stereo system. We first estimate the image
center and the aspect ratio of each camera using the

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 esti-
mations are done by an ordinary non-linear minimization
scheme (i.e., Levenberg-Marquardt) without any special
settings; nevertheless, the performance is satisfactory for
the purpose of robot navigation.

REFERENCES

tric Image Formation. Int. J. of Computer Vision, Vol. 35, No. 2,


directional Stereo Pairs. In Proceedings of IEEE Workshop on


for Obstacle Detection and Tracking in Dynamic Environments.
In Proceedings of the 2001 IEEE/RSJ Int. Conf. on Intelligent Robots

Dynamic Environments using Omnidirectional Stereo. In Proceedings of

17, No. 1, pp. 43–76, 1996.

by Integrating Omnidirectional Stereo and Laser Range Finder.
In Proceedings of 2002 IEEE/RSJ Int. Conf. on Intelligent Robots


[12] B. Mlčič and T. Pajdla. Omnidirectional Camera Model and
Epipolar Geometry Estimation by RANSAC with Binning. In
Proceedings of the 13th Scandinavian Conf. on Image Analysis, pp.

by Considering Map and Localization Uncertainty. In Proceedings
of the 8th Int. Conf. on Intelligent Autonomous Systems, pp.

Unknown Environments Using Omnidirectional Stereo and Laser
Range Finder. In Proceedings of 2004 IEEE/RSJ Int. Conf. on

Numerical Recipes in C: The Art of Scientific Computing, 2nd

[16] D. Szelow, J. Mishekr, D. Koos, and S. Singh. Precise Omni-
on Computer Vision and Pattern Recognition, Vol. 1, pp. 689–694,

Panoramic Cameras. In Proceedings of 5th European Conf. on


[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.

Table I
ESTIMATED PARAMETERS.

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

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>upper</td>
<td>364.49</td>
<td>-322.05</td>
<td>1554.36</td>
<td>-3.26</td>
<td>-1.90</td>
<td>78.93</td>
<td>179.235</td>
</tr>
<tr>
<td>lower</td>
<td>367.59</td>
<td>-320.47</td>
<td>1255.15</td>
<td>6.02</td>
<td>-0.80</td>
<td>-25.76</td>
<td>179.856</td>
</tr>
</tbody>
</table>

(b) Parameters obtained by calibration using epipolar geometry.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>upper</td>
<td>375.86</td>
<td>-321.65</td>
<td>1554.36</td>
<td>-2.71</td>
<td>-2.86</td>
<td>77.47</td>
<td>179.235</td>
</tr>
<tr>
<td>lower</td>
<td>364.48</td>
<td>-318.82</td>
<td>1255.15</td>
<td>5.88</td>
<td>-0.68</td>
<td>-25.25</td>
<td>179.856</td>
</tr>
</tbody>
</table>
with calibration using epipolar geometry

(a) panoramic images

(b) disparity images

(c) scan of laser range finder

Fig. 9. Evaluation of the calibration result.