CYLINDRICAL SENSOR CALIBRATION USING LINES

L. Smadja, R. Benosman, J. Devars

Université Paris VI
Laboratoire des Instruments et Systèmes d'Ile-de-France
BP 164, Tour 22-12 2ème étage
4 Place Jussieu, 75252 Paris Cedex 05, France

ABSTRACT

This paper deals with the calibration of a cylindric omnidirectional imaging system, based on a rotating 2048 pixels linear camera, which provides high definition panoramas.

The two-steps algorithm relies on line segments projections, which are sinusoidal curves. Moreover, Plucker formalism makes it possible to linearize the full cylindrical model.

Calibration data are provided by a sinusoid detection step, based on a "dual" Hough transform. Results obtained from the first process are used to evaluate one of the intrinsics, the other one being determined by a linear minimization criterion in the dual space.

1. INTRODUCTION

The significant development of omnidirectional vision has led to two different directions [1]. The first one relies on the use of classic cameras coupled with curved mirrors [2, 3], providing real time panoramas, whereas the second relies on setting linear cameras on rotation [4, 5, 6], generating high definition images. The approach relies on such a sequential sensor.

Calibration process constitues the first step in any vision algorithm. It provides the relation between image points, expressed in pixels, and $3D$ metric directions which they are related to. It has be shown in [7], that cylindrical sensors can hardly be calibrated by usual methods, using points or unconstrained lines projections.

We present here the use of Plucker formalism in order to linearize the cylindrical full projection of $3D$ real lines, which are sines combinations on the image cylinder. A sinusoid detection is also proposed, based on a "dual" Hough Transform, which allows an estimation of one of the intrinsics and provides data for the full calibration.

2. CAMERA MODEL AND CALIBRATION

2.1. Sensor Architecture

The sensor is based on a 2048 pixels Tri-CCD bar, put on rotation on its optical axis by a DC engine. During the rotation, the engine sends a signal to the camera *via* an optical coder, typically every tenth of degree, to allow the acquisition column by column of the environment around the sensor. After the whole rotation, a high definition *RGB* cylindrical panoramas, of dimensions $3600 \times 2048 \times 3$, is obtained. For further details, the reader should refer to [8].

2.2. Point Projection

Let us consider the *Camera Coordinates System* R_C, centered on C, whose $z - axis$ is aligned with the rotation axis. The coordinates of the projection p of a physical point $P = (X_c, Y_c, Z_c)^T$ on the image cylinder are simply given by :

$$p \begin{cases} x = f\frac{X_c}{Z_c} \\ y = f\frac{Y_c}{Z_c} \\ z = f\frac{Z_c}{R} \end{cases} \quad (1)$$

where $R = \sqrt{X_c^2 + Y_c^2}$(Cf Fig. 1).

Image coordinates m on the cylinder can be expressed from the P coordinates by :

$$m \begin{cases} u = K_u \arctan(\frac{Y_c}{X_c}) \\ v = v_0 - K_v f \frac{Z_c}{X_c} \cos(\frac{u}{K_u}) \end{cases} \quad (2)$$

where K_u and K_v are the image discretization steps and v_0 the projection of C on the bar. The K_u parameter corresponds more precisely to the inverse of the vertical pixel size, and K_v is the angular step (*ie* 2π divided by the number of acquired columns during a complete rotation).

Classic calibration technics generally use the point as a starting feature. In our case however, the Z_c/X_c ratio in equation (2) induces a singularity for the (Y, Z) plane points. Moreover, point P has to be expressed in a proper

0-7803-8554-3/04/$20.00 ©2004 IEEE.

coordinates system, known as the *World Coordinates System* R_W linked to R_C by the rigid transform $(R, \mathbf{t})$. R is a rotation matrix and $\mathbf{t}$ corresponds to the translation between the two coordinates systems origins. It is quite difficult to

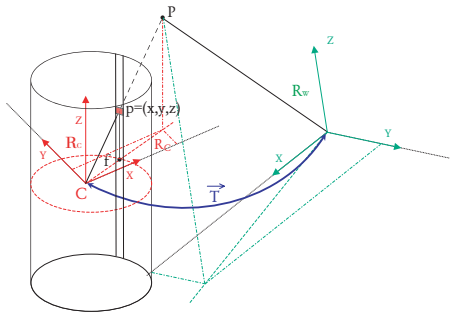

Fig. 1. Point P expressed in the two coordinates systems R_W and R_C, connected by the rigid transform (R, t).

inject this last transformation in equation (2), in order to obtain a global expression. The point projection cannot thus be fully modelized in a formal scheme, grouping intrinsics and extrinsics, linking metric R_W points to their pixels coordinates.

2.3. Line Intrinsic Projection

Let us consider P lying on a line segment (Fig. 2). This line defines with C a plane, called the *view plane*, whose equation is given by :

$$N_1X_c+N_2Y_c+N_3Z_c=0 \Longrightarrow \frac{Z_c}{X_c}=-\frac{N_2}{N_3}-\frac{N_1}{N_3}\frac{Y_c}{X_c} \quad (3)$$

where $\mathbf{N} = (N_1, N_2, N_3)$ corresponds to the view plane normal vector.

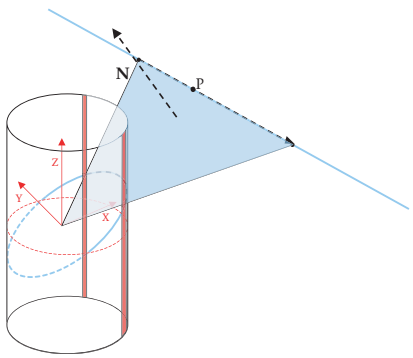

Fig. 2. Line projection on the cylinder :The line image depends on the corresponding view plane **N**.

Combining equations (3) and (2) leads to :

$$m \begin{cases} u=K_u\theta \\ v=v_0+K_vf\frac{N_1}{N_3}\cos\left(\frac{u}{K_u}\right)+K_vf\frac{N_2}{N_3}\sin\left(\frac{u}{K_u}\right) \end{cases} \quad (4)$$

Considering equations (4), we deduce that the line image on the cylindrical sensor is a *sinusoids combination*, whose magnitudes are given by ratios of the corresponding view plane components.

This expression is however restricted to R_C and should be extended to R_W. The use of Plucker formalism allows to express the view plane components in the world coordinates system. Indeed, it can be shown [9] that the Plucker coordinates $L = [\mathbf{m}, \mathbf{n}]^T$ in R_W can be expressed in another coordinates system such as :

$$L' = [\mathbf{m}', \mathbf{n}'] = [R\mathbf{m}, R(\mathbf{m} \times \mathbf{t} - \mathbf{n})]^T \quad (5)$$

where R and $T_\times$ are the extrinsics, parameters between the two coordinates systems.

2.4. Full Model

Let us now consider a R_W line, described by $L = [\mathbf{m}, \mathbf{n}]^T$. Using expression (5), the view plane normal $\mathbf{N}$ corresponding to this line can be expressed in R_C, through a linear way :

$$\mathbf{N}_{R_C} = \begin{pmatrix} N_1 \\ N_2 \\ N_3 \end{pmatrix}_{R_C} \sim \begin{bmatrix} -RT_\times | R \end{bmatrix} \begin{bmatrix} \mathbf{m} \\ \mathbf{n} \end{bmatrix}_{R_W} \quad (6)$$

where $T_\times$ is the skew-symmetric matrix corresponding to the $\mathbf{t}$ cross product.

Moreover, an image sine $v = v(u)$ is described by two magnitudes A and B, which can easily be identified to the view plane components ratios $K_vf\frac{N_1}{N_3}$ and $K_vf\frac{N_2}{N_3}$ in equation (4).

From equations (4) and (6), we obtain a model grouping both extrinsics and one of the intrinsics, K_vf. This model links Plucker vector $L = [\mathbf{m}, \mathbf{n}]^T$ of any R_W line to its associated (A, B) image sine magnitudes :

$$\begin{pmatrix} A \\ B \\ 1 \end{pmatrix} = K_vf \begin{pmatrix} \frac{N_1}{N_3} \\ \frac{N_2}{N_3} \\ 1 \end{pmatrix} \sim K_vf \cdot \begin{bmatrix} -RT_\times | R \end{bmatrix} \begin{bmatrix} \mathbf{m} \\ \mathbf{n} \end{bmatrix} \quad (7)$$

This complete model implies nine parameters : 3 Euler angles for R, 3 components for $T_\times$, K_vf, the sensor focal length, expressed in pixels, K_u, implicitly given by the number of acquired columns and v_0, corresponding to the optical center projection. These last two values being already known, seven parameters are to be determined.

Each detected image line provides two equations related to the associated sine magnitudes (Cf equation (7)). Four image sines/Plucker vectors correspondences are then needed in the minimal case to determine all the parameters.

3. LINE DETECTION

Line segments detection in classical imaging is an understood topic. One of the most used methods is based on an accumulation chart in a dual space corresponding to the $\mathbb{R}^2$ lines space. This transform, presented by Hough [10], links points on image lines to dual sinusoids, which are stacked in a map. Maxima extracted from this accumulation map correspond to the *slope* and *y-intercept* of the detected lines in the edge source image.

3.1. Dual Hough Transform

We proved in section 2.3 that any $3D$ space line is projected onto the cylindrical sensor through a sines combination centered on v_0 (Cf. equation (4)). We can associate points on image sines to dual lines, using a dual reverse Hough Transform. Image coordinates are centered on an arbitrary v_0 value, $offset_v$, and the following algorithm is applied :

For every (u, v) points in the source edge image:

- Compute $a = \frac{-1}{\tan\frac{u}{K_u}}$ and $b = \frac{v - offset_v}{\sin\frac{u}{K_u}}$.
- By varying x from X_{min} to X_{max}, trace $y = ax + b$ in the Hough dual space.

Extract the local maxima from the obtained accumulation map, whose coordinates respectively correspond to

$$A = K_v f \frac{n_1}{n_3} \text{ and } B = K_v f \frac{n_2}{n_3}$$

3.2. Error Relative to a Wrong v_0 Offset

The $offset_v$ choice strongly influences the precision on the results, which can be used in order to determine the v_0 parameter. Two dual approches have been developed, the first one relying on the accumulation map itself, whereas the second depends on an image process, using a voting points identification. These two algorithms are based on an iterative process relying on the computation of many dual maps, built around numerous values of v_0.

v_0 Influence on the Accumulation Map : Dual Criterion

For a δv_0 error, the dual coordinates (A, B) corresponding to the sines intersection are modified such as :

$$\begin{cases} A' = A + \delta v_0 \left(\frac{\sin\theta_1 - \sin\theta_2}{\sin\theta_1\cos\theta_2 - \sin\theta_2\cos\theta_1}\right) \\ B' = B - \delta v_0 \left(\frac{1+\cos\theta_2}{\sin\theta_1}\frac{\sin\theta_2 - \sin\theta_1}{\sin\theta_2\cos\theta_1 - \sin\theta_1\cos\theta_2}\right) \end{cases} \quad (8)$$

The intersection positions move along an additive term, proportional to the committed error δv_0. There is then a dispersion phenomenon around the optimal accumulation positions. A decrease of the accumulation peaks height, proportional to the δv_0 error, can logically be observed. It is thus possible to estimate an optimal value for v_0 when the local maxima reach their highest level. The dispersion phenomenon is indeed minimal when the vertical offset is close to the v_0 real value.

Using an edge image of a lines calibration pattern, whose dual map present many local maxima, our first criterion relies on the mean height value of these peaks. It is solely based on the dual accumulation charts, that is why we define it as the *dual criterion*.

Voting Points Identification : Image Criterion

The second method is based on an identification of the voting points. We have to recognize in the dual accumulation map any edge points, *ie* any dual line, which has voted for a specific sinusoid curve, associated to a local maximum. We obtain, by computing the square sum of image distances between every voting point and their associated sine, an error relative to a map centered on a v_0 specific value :

For every (u, v) points in the source image:

- Recompute a et b and estimate the dis- between the obtained line to the local maxima in the dual space.
- Associate this feature point to a sine if d is below a fixed threshold.
- Compute the square distance between the point and the sine for which it has vot -ed.
- Update the Image criterion.

We obtain an image error for all the voting points, whose mean value correspond to the error associated to a v_0 value. One finally assigns to the first intrinsic the offset value returning the lowest image error.

4. RESULTS

4.1. Line Detection and v_0 Determination

A specific calibration pattern containing lines, made of two perpendicular planes, has been used. An edge image is then generated, providing data to build the accumulation charts around many v_0 values. We chose a v_0 interval from 980 to 1200, which gives 221 different dual cards.

The dual map returning the lowest normalized image error then corresponds to the best image estimation of v_0. The image criterion returns an optimal v_0 value of **1067**, which corresponds to the mean error value 0.4059 pixel per voting point (Fig.3).

The peak height mean value, as its median, return a value close to **1067**, corresponding to a mean peak value of 140 voting points. This second result enforces the image criterion results (Fig. 3). The finally chosen value for the first intrinsic v_0 is **1067**, which verifies both of the two presented criteria. It corresponds indeed to the highest peak

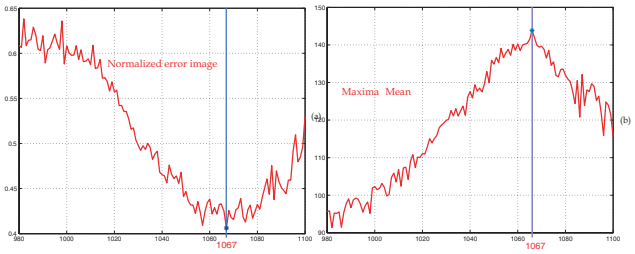

Fig. 3. (a) Highest dual criterion is reached at $v_0 = 1067$. (b) Lowest image error is also reached at $v_0 = 1067$.

mean value, which means the smallest dispersion and minimizes the image criterion, *ie* returns the lowest normalized image error.

4.2. Full Calibration

We have dual coordinates extracted from $v_0 = \mathbf{1067}$, associated to real $3D$ Plucker vectors, expressed in a metric world coordinates system. Minimization process relies on a *Levenberg-Marquardt* algorithm, included in the *Matlab* optimization toolbox. It is based on an exhaustive search along the whole directions in the minimization space, here $\mathbb{R}^7$. The seven retrieved calibration parameters corresponding to $v_0 = \mathbf{1067}$ are :

$$\left\{ \begin{array}{ccc} \xi = -0.0077 & \theta = 3.1374 & \varphi = 4.7017 \\ tx = 546.98 & ty = 693.75 & tz = 212.22 \\ & K_v f = 3274.7 & \end{array} \right\} \quad (9)$$

with an error, which can be defined as dual, of **2.5917**. This is a dimensionless measure, as it corresponds to a dual distance, *ie* a distance in the sines magnitudes space.

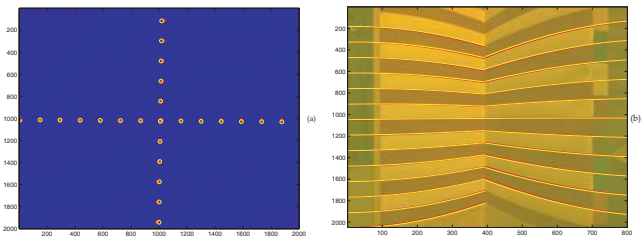

Fig. 4. (a) Obtained maxima coordinates.(b) Reprojected calibration pattern

The calibration precision can be observed on figure (4), where are illustrated reprojected local maxima in the dual map and corresponding image sines.

5. CONCLUSION

We have developed in this work a global formalization of the projection relative to a cylindrical panoramic sensor.

$3D$ line segments are proved to be imaged as sines combinations, and introducing Plucker vectors linearizes the full cylindrical projection model. This new kind of formalism is very useful and allows many computing facilites for panoramic images.

A image sines detection is also proposed, in order to provide one of the intrinsics, v_0 and the necessary data for the full calibration process. This algorithm can be adapted on many omnidirectional systems, for it relies on cylindrical panoramic projections, often used in this type of applications.

6. REFERENCES

[1] R. Benosman and S.B. Kang, *Panoramic Vision: Theory, Sensors and Applications*, Springer Verlag, 2001.

[2] S. Baker and S. Nayar, "A theory of single-viewpoint catadioptric image formation," *International Journal of Computer Vision*, pp. 175–196, November 1999.

[3] C. Geyer and K. Daniilidis, "Catadioptric projective geometry," *Proc. of IEEE Workshop on Omnidirectional Vision Omnivis' 00*, pp. 17–30, 2001.

[4] R. Benosman, T. Manière, and J. Devars, "Multidirectionnal stereovision sensor, calibration and scenes reconstruction," *13th International Conference on Pattern Recognition*, vol. A, pp. 161–165, August 1996.

[5] S. Peleg, M. Ben Ezra, and Y. Pritch, "Omnidirectional panoramic stereo imaging," *IEEE transactions on PAMI*, vol. 23, no. 3, pp. 279–290, March 2001.

[6] H. Y. Shum and R. Szeliski, "Stereo reconstruction from multiperspective cameras," *Proc. of IEEE ICCV*, vol. 1, pp. 14–21, September 1999.

[7] F. Huang, S. K. Wei, and R. Klette, "Comparative studies of line-based panoramic camera calibration," *IEEE Workshop on Omnidirectional Vision Omnivis'03*, June 2003.

[8] L. Smadja, R. Benosman, and J. Devars, "Determining epipolar constraint on cylindrical images and using it for 3d reconstruction," *Proc. ICAR*, August 2001.

[9] L. Smadja, *Génération d'environnements 3D denses à partir d'images panoramiques cylindriques*, Ph.D. thesis, UPMC, 2003.

[10] P. V. C. Hough, "Methods and means for recognizing complex patterns," *U.S. Patent 3 069 654*, December 1962.