

Cylindrical Sensor Calibration using Lines

L. Smadja, R. Benosman, and J. Devars

Laboratoire des Instruments et Systèmes d'Ile-de-France
Université Paris VI
BP 164, Tour 22-12 2ème étage
4 Place Jussieu, 75252 Paris Cedex 05, France
`laurent.smadja@lis.jussieu.fr`

1 Introduction

Calibration process is fundamental in any vision algorithm. It provides the relation between image points, expressed in pixels, and $3D$ metric directions which they are related to. This step has been largely studied within the classical frame [5, 7].

The significant development of omnidirectional sensors [2] has led to a general adaptation of the calibration process to any panoramic system. Two different ways have been explored, the first one based on the use of curved mirrors [1, 6, 17] whereas the second one relies on setting linear cameras on rotation [3, 10, 12, 14]. This second kind of panoramic sensors generally provides much higher definition panoramas, but can hardly be calibrated in the usual way without any kind of physical constraint [9].

Our approach relies on such a sequential system, which can be modeled by a *Single View Point* cylinder. Usual calibration patterns or methods cannot be applied in this case, because of the projection non-linearity. A different kind of features, as the line segments, can then be used in order to linearize the global projection, through the use of Plucker coordinates. Imaged as sines combinations on the cylindrical projection, these features have to be detected in order to provide calibration data. We present a sines detection, based on a “dual” Hough transform, which allows moreover an estimation of one of the intrinsics, using two dual methods. The full calibration process is then achieved through a “dual” criterion minimization, using the linear process.

2 Camera Model and Calibration

2.1 Sensor Architecture

The sensor used is presented in [15]; It is based on a 2048 pixels Tri-CCD bar, put on rotation on its optical axis by a DC engine. During the rotation, this 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. The rotation speed is directly linked to the acquisition time and is strongly influenced by the experimental conditions; Indeed, for an indoor scene,

a 250 ms exposure time per column is needed, whereas 15 ms are sufficient for an outdoor scene. The total acquisition time can thus vary from 1 to 15 minutes. We obtain, after the whole rotation, high definition *RGB* cylindrical panoramas, whose dimensions are $3600 \times 2048 \times 3$.

The optical center C of the CCD bar must be aligned with the engine axis, to verify the Single Viewpoint model, this constraint being assured by a simple mechanism [15]. The camera model can then be defined by a discrete cylinder, centered on C , whose radius f corresponds to the focal length of the camera.

2.2 Point Projection

Let us consider the *Camera Coordinates System* R_C , centered on C , whose z -axis is aligned with the engine one. The coordinates of the projection $p = (x, y, z)^T$ of a physical point $P = (X_c, Y_c, Z_c)^T$ on the image cylinder (Cf Fig. 1) 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 $R = \sqrt{X_c^2 + Y_c^2}$.

Image coordinates m are directly expressed from the cylindrical coordinates (R, θ, Z_c) , by the transformation :

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

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).

Classical calibration technics generally use the point as a starting feature. Besides, point P has to be expressed in a proper 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 seems to be quite difficult to inject this last transformation in equation (2), in order to obtain a global linear expression. The point projection cannot thus be easily modeled 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 $P = (X_c, Y_c, Z_c)^T$ lies on a line segment (Fig. 2). This line defines with C a plane, called the *view plane*, whose equation is given by :

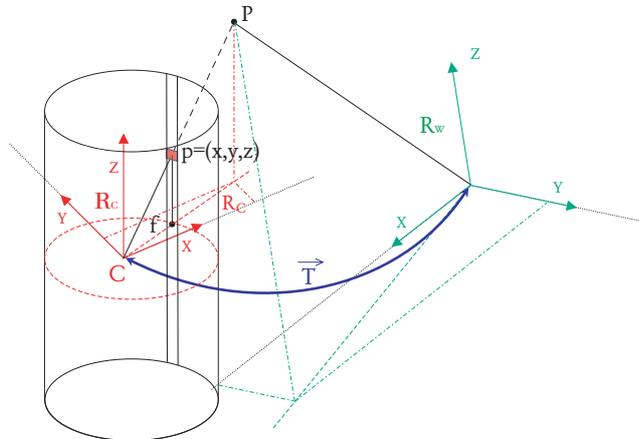


Fig. 1. Point P expressed in the two coordinates systems R_W and R_C , connected by the rigid transform (R, t) .

$$N_1 X_c + N_2 Y_c + N_3 Z_c = 0 \implies \frac{Z_c}{R} = -\frac{N_1}{N_3} \frac{X_c}{R} - \frac{N_2}{N_3} \frac{Y_c}{R} \quad (3)$$

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

Combining equations (2) and (3) – noticing that $\frac{X_c}{R} = \cos \theta$ and $\frac{Y_c}{R} = \sin \theta$ – leads to :

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

Considering particularly the second of these last two equations, 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.

Expression (4) is restricted to R_C and should be expressed in R_W , using the extrinsics and associated with the sinusoids magnitudes.

2.4 Plucker Formalism

Plucker vectors [13] constitute a simple representation of the 3D lines, which allows many computing facilities.

Definitions

1. Any six coordinates homogenous vector (*ie* defined up to a scale factor) $L = [\mathbf{m}, \mathbf{n}]^T$, such as $\mathbf{n} \cdot \mathbf{m} = 0$ corresponds to a \mathbb{R}^3 line segment.
2. The plane (four homogenous coordinates) defined by $L = [\mathbf{m}, \mathbf{n}]^T$ and point $P = (\mathbf{P}, w)$ is given by $(\mathbf{m} \times \mathbf{P} - w \cdot \mathbf{n}, \mathbf{n} \cdot \mathbf{P})$

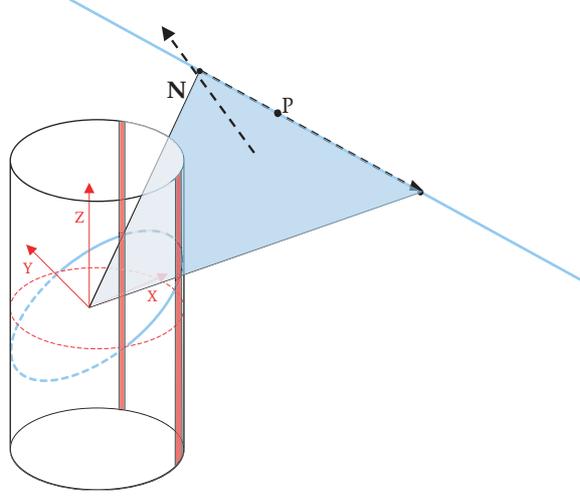


Fig. 2. Line projection on the cylinder :The line image depends on the corresponding view plane \mathbf{N} .

3. The line joining P and Q is given by $L = [\mathbf{P} - \mathbf{Q}, \mathbf{P} \times \mathbf{Q}]^T$.

This last point is particularly interesting, allowing to give a physical sense to the two Plucker components. \mathbf{m} is clearly the line direction vector, whereas \mathbf{n} corresponds to the normal of the plane formed by the line and the origin of the coordinates system, *ie* the view plane. Plucker formalism is then particularly well adapted to our sensor for it encodes almost directly the image of a line segment on a cylindrical panoramic image.

Rigid Motion $L = [\mathbf{m}, \mathbf{n}]^T$ is modified by a rigid transform.

- Pure Rotation R : Both vectors \mathbf{m} and \mathbf{n} are modified by a simple left product :

$$L' = [R\mathbf{m}, R\mathbf{n}]^T$$

- Pure Translation \mathbf{t} : If \mathbf{m} is logically invariant to the translation, \mathbf{n} is in the opposite largely modified. Indeed, \mathbf{n}' corresponds to the *new* normal plane, formed by the line L and the new origin $P = (\mathbf{t}, 1)$. Using definition 2, we obtain for a pure translation along \mathbf{t} :

$$L' = [\mathbf{m}, \mathbf{m} \times \mathbf{t} - \mathbf{n}]^T$$

- Rigid Motion R, \mathbf{t} Combining the last two results, we can express the modification of a Plucker vector L after a rigid motion (R, \mathbf{t}) (Fig. 3) :

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

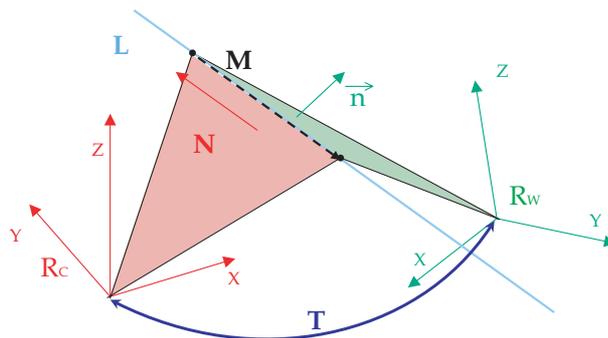


Fig. 3. Rigid motion for a Plucker vector.

2.5 Full Model

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

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

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

Moreover, an image sine (u, v) is characterized by the equation :

$$v = v_0 + A \cos \frac{u}{K_u} + B \sin \frac{u}{K_u} \quad (7)$$

whose magnitudes A and B can easily be identified to the view plane components ratios $K_v f \frac{N_1}{N_3}$ and $K_v f \frac{N_2}{N_3}$ in equation (4) :

$$\begin{pmatrix} A \\ B \\ 1 \end{pmatrix} = K_v f \cdot \begin{pmatrix} \frac{N_1}{N_3} \\ \frac{N_2}{N_3} \\ 1 \end{pmatrix}_{R_C} \sim K_v f \cdot \begin{pmatrix} N_1 \\ N_2 \\ N_3 \end{pmatrix}_{R_C} \quad (8)$$

By grouping equations (6) and (8), we obtain a model absorbing both extrinsics and one of the intrinsics, $K_v f$. This model links Plucker vector $L = [\mathbf{m}, \mathbf{n}]_{R_W}^T$ of R_W to its associated (A, B) image sine magnitudes :

$$\begin{pmatrix} A \\ B \\ 1 \end{pmatrix} \sim K_v f \cdot [-RT_{\times}|R] \begin{bmatrix} \mathbf{m} \\ \mathbf{n} \end{bmatrix}_{R_W} \quad (9)$$

This relation will be very useful within the frame of calibration. The full model implies nine parameters :

- ξ, θ, ϕ , the three Euler angles characterizing extrinsic rotation R .

- t_x, t_y, t_z , cartesian components of the extrinsic translation \mathbf{t} .
- $K_v f$, 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.

Each detected image line provides two equations related to the associated sine magnitudes :

$$\begin{cases} A = K_v f \frac{N_1}{N_3} \\ B = K_v f \frac{N_2}{N_3} \end{cases} \quad (10)$$

where $\mathbf{N} = [N_1, N_2, N_3] = -[RT_{\times}|R] L$.

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 a quite 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 [8] has been constantly improved, and the generalized algorithm is able to handle with any parametrized curve [11]. Each type of curves can indeed be expressed in a specific dual space, in which image points are mapped.

3.1 “Classical” Hough Transform

A fundamental evolution consists of the normal representation of lines, presented by Duda and Hart [4]. It can be noted that a line defined by $y = ax + b$, where a is the *slope* and b the *y-intercept*, can be expressed in polar coordinates. Indeed, let M_0 be the line point closest to the system origin, ρ_0 the OM_0 distance, and θ_0 the oriented angle between the x -axis and $[OM_0]$, a and b values check :

$$\begin{cases} a = \tan(\frac{\pi}{2} - \theta_0) = \frac{-1}{\tan \theta_0} \\ b = \frac{\rho_0}{\sin \theta_0} \end{cases} \quad (11)$$

A plane line can thus be represented by the (ρ_0, θ_0) couple, in comparison to the “traditional” (a, b) couple :

$$y = \frac{-1}{\tan \theta_0} x + \frac{\rho_0}{\sin \theta_0} \implies \rho_0 = x \cos \theta_0 + y \sin \theta_0 \quad (12)$$

Any (x, y) plane point is then associated with a dual sine in (ρ, θ) Hough space, when θ varies from 0 to 2π . The set of points, belonging to a particular (a_i, b_i) line, correspond to many sines, passing by this characteristic point, in the dual space. Maxima extraction, on the accumulation charts, give the couples corresponding to the origin image lines.

Reminding that any panoramic image can be reprojected onto a cylindrical image, and that any real line is imaged as a sines combination on such a surface, we present here a detection algorithm based on the dual Hough theory.

3.2 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)). Recall that in the classical case, Hough Transform links image lines to dual sines, we can associate image sines to dual line, using a dual (reverse) process (Fig. 4). Image coordinates are centered on an $offset_v$ value, and the following algorithm is applied :

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

- Compute $a = \frac{-1}{\tan \frac{\theta}{K_u}}$ and $b = \frac{v - offset_v}{\sin \frac{\theta}{K_u}}$ (Cf equation 12).
- 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 $A = K_v f \frac{n_1}{n_3}$ and $B = K_v f \frac{n_2}{n_3}$

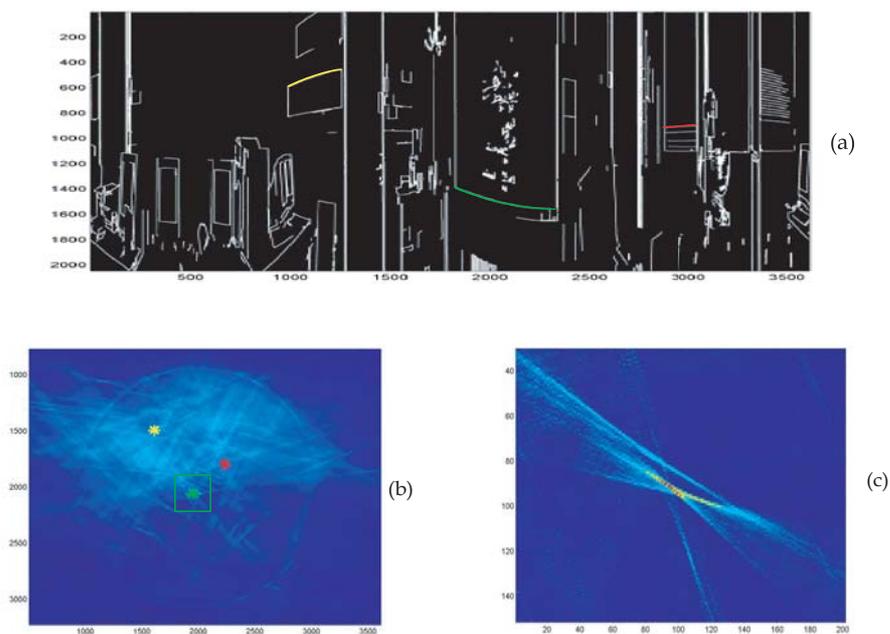


Fig. 4. (a) A panoramic edge image. (b) The associated dual accumulations map. (c) Zoom on the dual map.

3.3 Error Relative to a Wrong v_0 Offset

The value of $offset_v$ strongly influences the precision of the results. This can be used in order to find an optimal value of $offset_v$, ie to determine the v_0 parameter. Two dual approaches have been developed, the first one relies 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 An error on the v_0 value only influences $\rho = v - v_0$ and modifies the dual lines y -intercepts. For a δv_0 error, the dual coordinates (A, B) corresponding to the sines intersection are modified such as :

$$\begin{cases} A' = A + \delta v_0 \left(\frac{\sin \theta_1 - \sin \theta_2}{\sin \theta_1 \cos \theta_2 - \sin \theta_2 \cos \theta_1} \right) \\ B' = B - \delta v_0 \left(\frac{1 + \cos \theta_2}{\sin \theta_1} \frac{\sin \theta_2 - \sin \theta_1}{\sin \theta_2 \cos \theta_1 - \sin \theta_1 \cos \theta_2} \right) \end{cases} \quad (13)$$

Equation (13) shows that the intersection positions move along an additive term, proportional to the committed error δv_0 , but also depending on the angular position of the voting points $\theta_i = \frac{u_i}{K_u}$. There is then a dispersion phenomenon around the optimal accumulation positions, which is presented in Fig. 5, during a simulation stage.

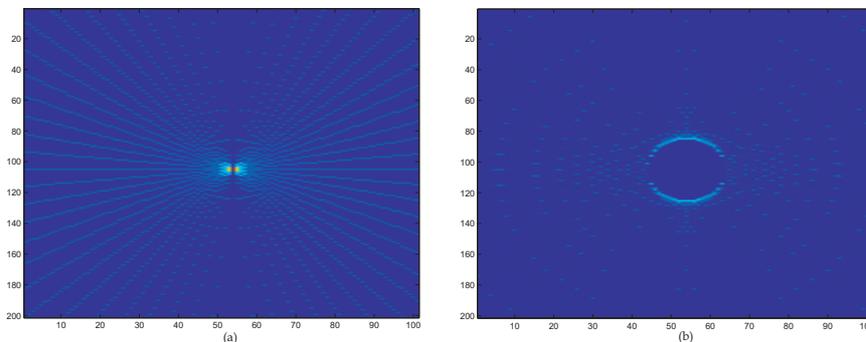


Fig. 5. (a) One can observe a perfect accumulation for a correct v_0 value. (b) There is a dispersion of the votes, relative to the bias on v_0 .

A decrease of the accumulation peaks height, proportional to the δv_0 error, is the logical consequence of the dispersion. It is then possible to estimate an optimal value for v_0 when the local maxima reach their highest points. The dispersion phenomenon is indeed minimal when $offset_v$ is the closest from the v_0 real value.

Using an edge image of a calibration pattern containing lines, whose dual map present many local maxima, our first criterion relies on the mean height value of these maxima. 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 point, *ie* any dual line, which has voted for a specific sinusoid curve, associated to a local maximum.

An error relative to a v_0 centered map is then computed from the square sum of image distances between each voting point and its associated sine. This process is realized through a second step on all the accumulation maps, obtained from the preceding algorithm :

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

```
-- Recompute  $a$  et  $b$  and estimate the distance between
the obtained line to the local maxima in the dual space.
Recall that the distance of a  $[x_0, y_0, 1]$  point to a
 $[a, -1, b]$  line is given by :
```

$$d = \frac{ax_0 - y_0 + b}{\sqrt{a^2 + 1}}$$

```
-- Associate this feature point to a sine if  $d$  is below a fixed
threshold, empirically chosen, and strongly depending on the
discretization step and global chart dimension.
-- Compute the square distance between the point and the sine for which it
has voted.
-- Update the Image criterion.
```

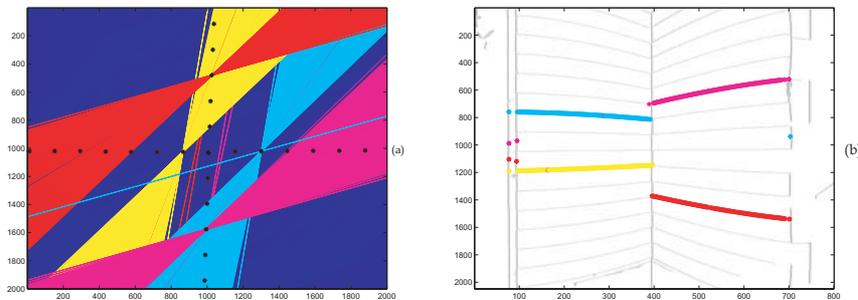


Fig. 6. (a) Some local maxima and duale lines having voted for them in the dual space. (b) The corresponding points in the calibration pattern edge image.

An image error is then estimated for all the voting points, whose mean value correspond to the error associated to a v_0 value. One finally assigns to v_0 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.

Image Criterion Fig. 6 presents an exemple of edge voting points, whose dual lines are identified in the accumulation map and associated to a specific local maximum. This step allows, through an image square difference between every voting point and the sine for which it has voted, to estimate the image criterion associated to an accumulation chart, *ie* to a v_0 value. The dual map returning the lowest normalized image error then corresponds to the best estimation. 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.7).

Dual Criterion Each height maximum evolution shows an increase when v_0 gets close to 1065. It is clear that the whole sines does not indicate the same optimal value, but the mean peak height value, as its median, return a value close to **1067**, corresponding to a peak mean value of 140 voting points. This second result enforces the image criterion results (Fig. 7).

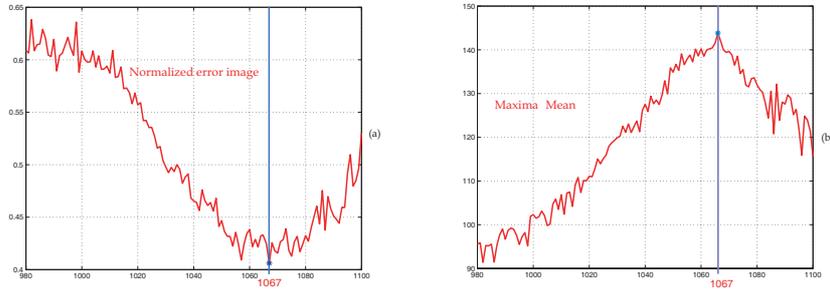


Fig. 7. (a)Mean maxima value evolution function of v_0 : the highest mean value is reached at $v_0 = 1067$.(b)Image criterion function of v_0 : the lowest error is reached at $v_0 = 1067$.

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 mean value, which means the smallest dispersion. Moreover, it minimizes the image criterion, *ie* the lowest normalized image error. This value, also the maxima positions in the dual space, allow us to achieve the calibration second step, where we determine all the parameters left.

4.2 Full Calibration

The full calibration process is achieved, using dual coordinates extracted from $v_0 = \mathbf{1067}$, associated to real $3D$ Plucker vectors, data linked by the relation (9). 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 calibration parameters corresponding to $v_0 = 1067$ are :

$$\left\{ \begin{array}{lll} \xi = -0.0077 & \theta = 3.1374 & \varphi = 4.7017 \\ tx = 546.98 & ty = 693.75 & tz = 212.22 \\ & K_v f = 3274.7 & \end{array} \right\} \quad (14)$$

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.

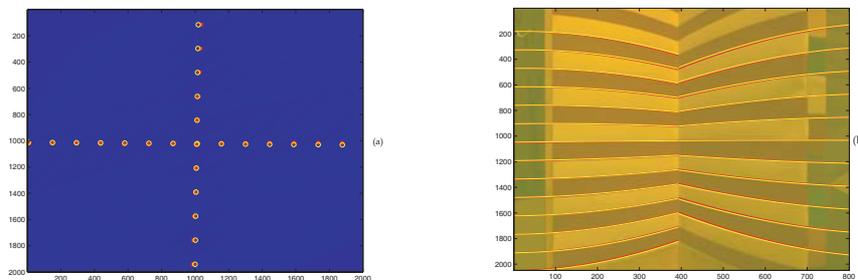


Fig. 8. After the dual space minimization : (a) Obtained maxima coordinates. (b) Reprojected calibration pattern lines after calibration : In red, detected lines (data), in yellow rebuilt lines.

The calibration precision can be observed on Fig. 8, where the reprojected local maxima in the dual map, as well as the corresponding image sines, are illustrated.

If the extrinsics, rotation and translation parameters between the pattern coordinates system and the sensor one, are not significant within this study, $K_v f$ parameter is very important. Pixel focal length is indeed the second intrinsic, as we defined it in the section 2.3.

5 Conclusion

This paper presented a global formalization of the projection relative to a high definition sequential sensor, which provides $2048 \times 3600 \times 3$ panoramas.

This kind of non-linear sensor cannot be calibrated by the usual methods, for the point projection induces many singularities. $3D$ line segments are proved

to be imaged as sines combinations, and introducing Plucker vectors linearizes the full projection model. This new kind of formalism is very useful and allows many computing facilities for panoramic images.

An 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, as it relies on cylindrical panoramic projections, often used in this type of applications.

The calibration stage constitutes the first step for any further vision algorithm. This sensor has indeed been used in a stereoscopic system, for rebuilding existing scenes through high definition disparity maps [16].

References

1. S. Baker and S. Nayar. A theory of single-viewpoint catadioptric image formation. *International Journal of Computer Vision*, pages 175–196, November 1999.
2. R. Benosman and S.B. Kang. *Panoramic Vision: Theory, Sensors and Applications*. Springer Verlag, 2001.
3. R. Benosman, T. Manière, and J. Devars. Multi-directionnal stereovision sensor, calibration and scenes reconstruction. *13th International Conference on Pattern Recognition*, A:161–165, August 1996.
4. R. O. Duda and P. E. Hart. Use of the hough transform to detect lines and curves in pictures. *Communications of the ACM*, 15:11–15, 1972.
5. O. Faugeras. *Three-Dimensional Computer Vision*. MIT Press, 1993.
6. C. Geyer and K. Daniilidis. Catadioptric projective geometry. *Proc. of IEEE Workshop on Omnidirectional Vision Omnivis' 00*, pages 17–30, 2001.
7. R. I. Hartley and A. Zissermann. *Multiple view geometry in computer vision*. Cambridge University Press, 2000.
8. P. V. C. Hough. Methods and means for recognizing complex patterns. *U.S. Patent 3 069 654*, December 1962.
9. 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.
10. I. Ishiguro, M. Yamamoto, and S. Tsuji. Omnidirectional stereo. *IEEE Trans. on PAMI*, 14:257–262, February 1992.
11. C. F. Olson. Constrained Hough transforms for curve detection. *Computer Vision and Image Understanding*, pages 329–345, 1999.
12. S. Peleg, M. Ben Ezra, and Y. Pritch. Omnidirectional panoramic stereo imaging. *IEEE transactions on PAMI*, 23(3):279–290, March 2001.
13. K. Shoemake. Plucker coordinate tutorial. *Ray-Tracing News vol 11*, July 1998.
14. H. Y. Shum and R. Szeliski. Stereo reconstruction from multiperspective cameras. *Proc. of IEEE ICCV*, 1:14–21, September 1999.
15. L. Smadja, R. Benosman, and J. Devars. Determining epipolar constraint on cylindrical images and using it for 3d reconstruction. *Proc. ICAR*, August 2001.
16. L. Smadja, E. Bigorgne, R. Benosman, and J. Devars. Génération de cartes de disparité denses partir d'images panoramiques cylindriques haute définition. *Proc. GRETSI*, September 2003.
17. T. Svoboda, T. Pajdla, and V. Lhavac. Central panormaic cameras : Geometry and design. Technical report, Center of Machine Perception, 1998.