Determining Epipolar Constraint on Cylindrical Images and Using it for 3D Reconstruction

Laurent Smadja Ryad Benosman Jean Devars
Laboratoires des Instruments et Systèmes d'Ile de France
Université Pierre et Marie Curie
75252 Paris Cedex 05
{laurent.smadja | rbo@lis.jussieu.fr,devars@ccr.jussieu.fr}

Abstract

In this paper, we describe the design of a panoramic vision system. We also present a generalization of epipolar geometry for cylindrical stereo images. Knowing this constraint, minimization processes are computed to recover some experimental parameters. Furthermore, epipolar constraint optimizes the matching stage and thus, allows to achieve 3D reconstructions of outdoor static scenes from high definition images.

1. Introduction

Generating panoramic images can be done using a variety of sensors mainly mirrors, lenses and rotating cameras. Omnidirectional sensors based on lenses and mirrors [1][2] are usually very used in real time imaging as they do not rely on mechanical parts for their acquisition. They are mainly used to video conferencing and in robotics applications [3][4][5]. Although they allow real time applications, they suffer from a lack of resolution that prevent them from generating high resolution images. Omnidirectional sensors based on rotating cameras [6], on the other hand can not be used in real time applications, due to the time of exposure of the cameras and also due to the mechanical constraints. Compared to mirrors and lenses, rotating sensors provide very high resolution images using on linear or matrix cameras. This high amount of pixels allow them to generate very precise and thin reconstructions of surfaces. These properties are very useful in modelisation applications like cinematographical ones to generate the set, which is the scope of the presented work. This work can be seen as the follow up of the work presented in [7][8], by generalizing the concept of cylindrical acquisition, allowing free positionning and orientation of view points.

1.1. Architecture of the Sensor

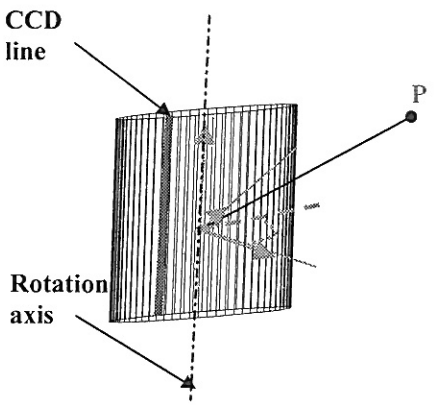

Figure 1: Geometry of the sensor

The developped sensor is based on a 3 CCD linear RGB camera. of 2048 pixels. The camera is

rotating on its optical axis by a DC engine, which sends a regular signal every angular position, typically differing by a tenth of degree between signals, to grab a column. By concatenating these columns, we can obtain a 2048*3600 a panoramic image.

The rotating device relies on:

- a command card integrating an AVR microchip
- a DC motor with an optical encoder, and a reducer of which the reduction rate is 1152.

The card is configured via the serial port, it controls the rotation speed.The engine control card provides a width of pulse modulated signal from the microchip, via a MOSFET commutating stage. The optical encoder supplies 500 pulses per revolution, then at the reducer's output, we have 500 * 1152 = 576000 pulses. Thanks to this signal, the DC motor is speed servo controlled.

The camera used is a tri CCD *JAI CV L-103*, which provides 1*2048*3 RGB images at 2400 fps, much more than needed. It is coupled to an *Imaging PC-DIG* acquisition card. The lens is a *Nikon F-Or 28/80 mm* usually set on the 28 mm focal length. This assembly can be controlled by an external signal sent by the DC engine at a constant rate that we choose according to the experimental conditions. This rate determines the exposure time on the CCD sensor. As an example for bright scenes, a short time will be preferred, 10 ms for exposure time, thus the engine runs quickly and the total acquisition takes 36 seconds, whereas for dark environments, the exposure time can slow down to 255 ms a column (about 15 minutes for an image).

The camera is rotating on its own optical axis, one may wonder how to locate the optical center. An optical set has been developped which principle is illustrated by ***Figure 2***. Two boards are needed, each presenting three blue lines. One board carries a planar mirror. The boards face each other and are perfectly parallel. The camera is placed to see both the mirror and the lines. The mirror will then reflect its facing board. By locating the reflected lines, we can recover the angles $\alpha=(P_{21}\ O\ P_{22})$ and $\beta=(P_{11}\ O\ P_{12})$.

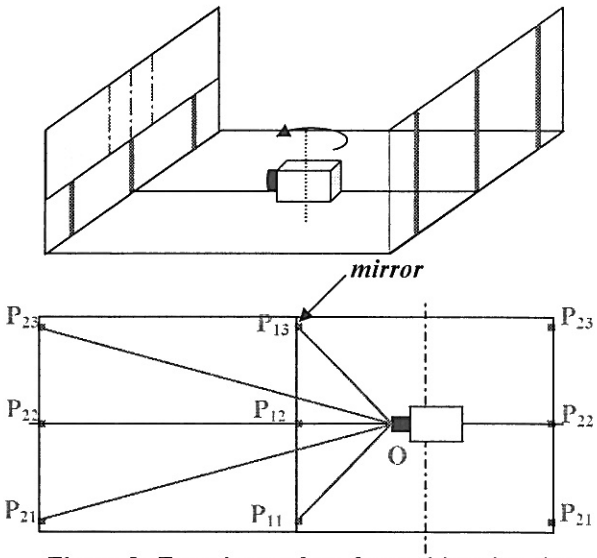

Figure 2: Experimental set for positionning the camera

If the optical center is exactly placed between the two planes we verify:

$$\alpha = Arctg(\frac{1}{3}tg\beta) \qquad (1)$$

This method is highly accurate, it can precisely determine the location of the center of projection and thus estimate our cylindrical image's radius

Following all its main caracteristics previously explained the developped sensor is naturally dedicated to static environments. Due to its single linear camera we can carry out any displacements between the two views, giving us a total freedom in our manipulations.

The size of the acquired images being huge it appears the importance of a computing very accurate epipolar geometry. That is the theorical purpose of this study.

1.2. Determining epipolar lines

Practically all of the existing stereo systems use a geometrical constraint to reduce the complexity in the matching stage, more specifically the *epipolar constraint*. This approach allows us to reduce the search of matching points from a 2D space to a 1D space. The epipolar constraint is well known for standard images, but few authors have considered

unusual cases, such as hyperbolic mirrors [9] or cylindrical images[10][11].

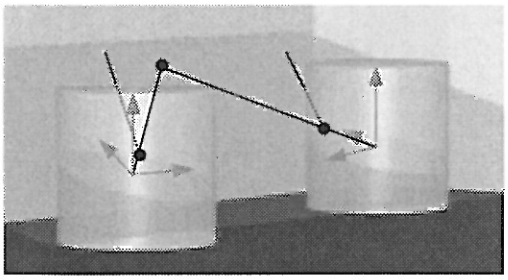

Figure 3: Epipolar geometry general scheme

The presented approach can be used to determine the equation of the epipolar line on cylindrical images resulting from an arbitrary couple acquired by the sensor. Let **P** be a physical point, $\mathbf{p_1}$ and $\mathbf{p_2}$ respectively its two projections on both cylindrical images. $\mathbf{O_1}$, $\mathbf{O_2}$ the two optical centers of the images. $\mathbf{T_{12}}$, the rigid transform between $\mathbf{R_1}$ and $\mathbf{R_2}$ the coordinates systems associated to $\mathbf{O_1}$ and $\mathbf{O_2}$. For the classical planar case, the epipolar line is the geometric intersection between the second image plane and a specific plane, called the *view plane*, formed by **P**, $\mathbf{O_1}$ and $\mathbf{O_2}$

Figure 3 shows that in our case, this intersection will not be a line but an ellipse, or sometimes a circle if the observed point is at the same height as the optical center

The main idea was not to consider the physical point **P** in our calculations but its image positions, $\mathbf{p_1}$ and $\mathbf{p_2}$. The second idea was to work on a metric scale rather than a pixel scale to compute this constraint.

Our goal is to determine in a chosen coordinates system the intersection between the *view plane* and the image cylinder. Let us compute the coordinates of **P** in $\mathbf{R_1}$. Since the image is cylindrical, we can assume that :

$$p1O_1 \begin{cases} x_{p1} = focal * \cos(\alpha) \\ y_{p1} = focal * \sin(\alpha) \\ z_{p1} = (v_{center} - v_1) * pixelsize \end{cases} \quad (2)$$

Where:

$$\alpha = u_1 * 2\pi / Ncolumn\ , \quad (3)$$

focal is the radius of the cylinder and v_{center} is the position of the optical center in the image line.

The rigid transform $\mathbf{T_{12}}$ is composed of three rotations φ, θ and ξ around the three axes and of a translation between $\mathbf{O_1}$ and $\mathbf{O_2}$:

$$T_{12} = \begin{pmatrix} & \Re_{12} & & tx \\ & & & ty \\ & & & tz \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (4)$$

We can obtain the same coordinates for $\mathbf{p_2}$ in the coordinates system $\mathbf{R_2}$. We can then write:

$$p1O_2 = T_{12} * p1O_1 \quad (5)$$

where $\mathbf{p_{01}}$ is the homogeneous coordinates of $\mathbf{p_1}$ in the first coordinate system:

$$p1O_1 = \begin{pmatrix} x_{p1} \\ y_{p1} \\ z_{p1} \\ 1 \end{pmatrix} \quad (6)$$

We are now able to express the $\mathbf{p_1}$, $\mathbf{p_2}$, $\mathbf{O_1}$ and $\mathbf{O_2}$ coordinates in a single coordinate system, the one associated to $\mathbf{O_2}$ for instance. Therefore, we can compute the *view plane*'s equation by:

$$Plane : \begin{vmatrix} x & y & zcalcul & 1 \\ xO_1 & yO_1 & zO_1 & 1 \\ xO_2 & yO_2 & zO_2 & 1 \\ xp_1 & yp_1 & zp_1 & 1 \end{vmatrix} = 0 \quad (7)$$

The epipolar line is the intersection between the computed plane and the cylinder centred on $\mathbf{O_2}$. The intersection gives us a relatively complex equation that can factorized giving the cylinder height called z_{calcul} according to the angular position α varying from 0 to 2π. We can then retrieve the ellipse equation $z_{calcul}(\alpha)$. The final stage consists of transcribing these values into image coordinates, using ***equations (2)*** and ***(3)***.

To summarize, the cylindrical geometrey relies on the estimation of v_{center}, *focal* and *pixelsize*, which are the intrinsics parameters, and $\mathbf{T_{12}}$ the extrinsics parameters. Several cases can be determined:

1. $\mathbf{T_{12}}$ unknown, intrinsics parameters unknown.
2. $\mathbf{T_{12}}$ unknown, intrinsics parameters known.

Case 1 is unstable as too many parameters must be determined, minimization diverge. Case 2 is the retained one, it gives reliable results if the starting $\mathbf{T_{12\ 0}}$ of the minimization process is not too far from the optimal solution. $\mathbf{T_{12\ 0}}$ in our case is estimated during the acquisition stage.

1.3. Estimating the calibration parameters

As explained previously, the entry of the minimization relies on two uncalibrated cylindrical image points to be matched.We used a matching algorithm as presented in [12] on our images to find reliable matches. This algorithm is quite robust and deals with scaling, intensity and orientation changes, which seems to fit with the type of image we have. The **N** obtained matching points are used to build a learning base, on which computations will be carried out. We have image coordinates of each of these points, with which we can recover their cylinder coordinates α_i and $Z_i(\alpha_i)$, and compute for each α_i the corresponding $Z_{calcul}(\alpha_i)$ with our epipolar equation. The criterion Q becomes evident; it has to be derived for all extrinsic parameters, which are the values to be determined, φ, θ, ξ, t_x, t_y and t_z :

$$Q = \sum_{i=1}^{N} \left| zi - zcalculi \right|^2 , N \in LearningBase \qquad (8)$$

We tried out three different iterative minimization algorithms, the first of them uses the MATLAB function leastsq.m, which combines many ways to achieve a correct result.The second one is a simple gradient descent based on equation (8) where the gradients $\nabla Q(X_i)$ are calculated using MAPLE.
The third minimization algorithm is based on a Levenberg-Marquardt method, to minimize the criterion.

2. Experimental results

The stereo pair presented in Figure 5 has been shot in the city of Angoulême. The chosen focal length was 28 mm and the exposure time between 15 and 45 ms a column, depending on the weather conditions. We can observe two particular points, called *epipolar centers*, which are the intersections between $(O_1\ O_2)$ and the image cylinder.
For 45 matched points, we applied the three different methods.

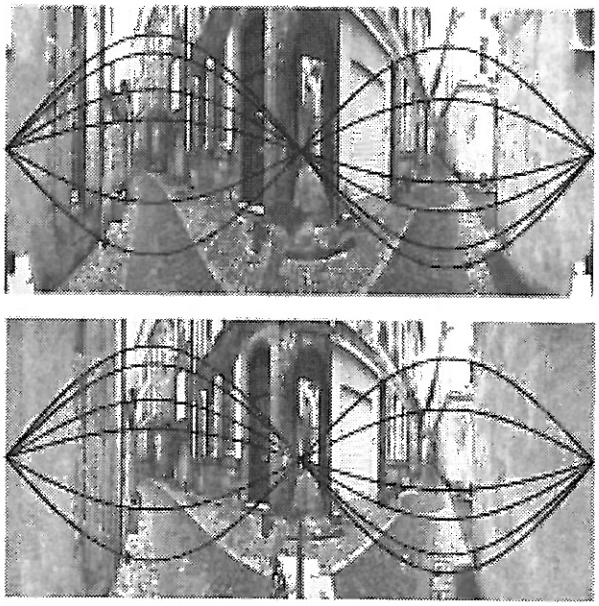

Figure 5: Determination of the epipolar lines

The obtained results are represented in Table 1 :

	$T_{12\,0}$	Leastsq	Gradient descent	Levenberg Marquardt
φ	0	0.036	- 0.015	- 0.042
θ	0	- 0.019	- 0.004	0.000
ζ	0	- 0.099	- 0.003	0.000
t_x	-1	0.153	- 0.970	- 0.910
t_y	0	0.063	- 0.324	- 0.042
t_z	0	0.141	0.039	0.003

Table 1: Minimization results.

Figures 6, 7 and 8 illustrate the distance between the calibration points and their respective computed epipolar lines (as shown in ***Figure 8***) expressed in pixels

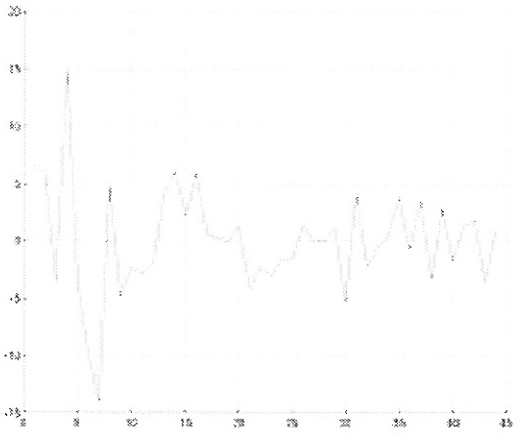

Figure 6 :Pixel error with leastsq.m

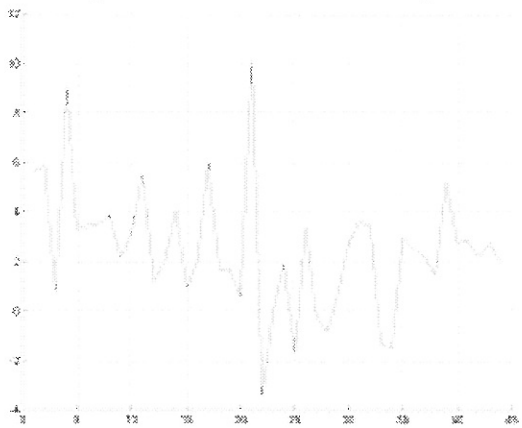

Figure 7 :Pixel error with gradient descent

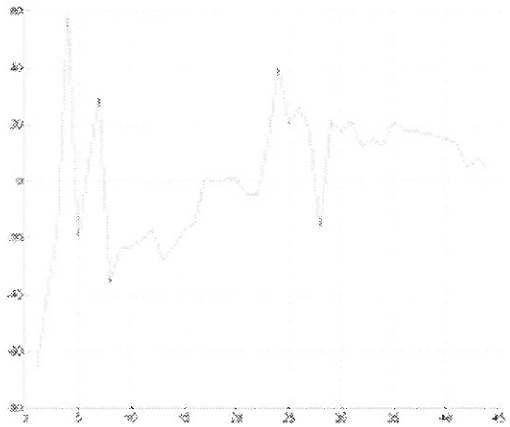

Figure 8 :Pixel error with Levenberg Marquardt

Some simple sparse 3D reconstructions have been achieved, using correct matches, one of them is presented in ***Figure 9***.

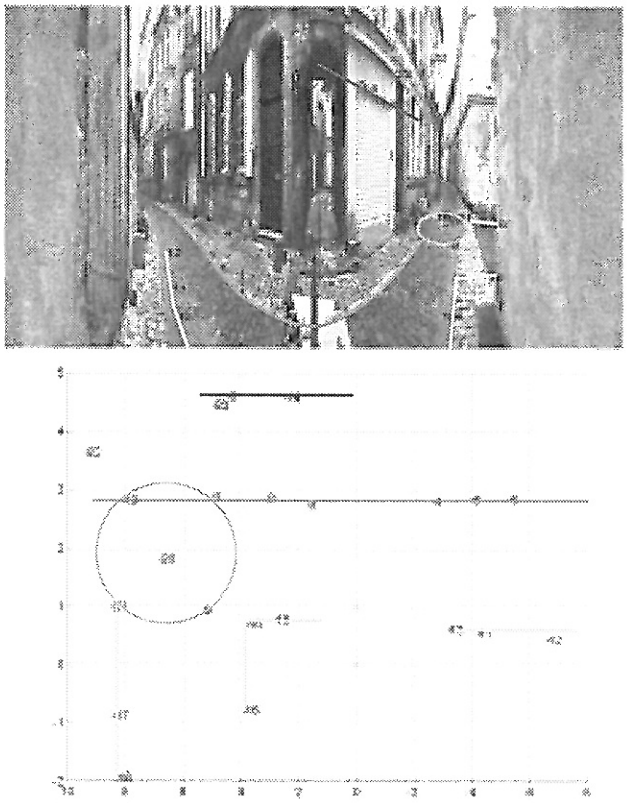

Figure 9: Sparse reconstruction of correct matches

3. Conclusion

In this paper we presented a generalized concept for panoramic imaging acquired by rotating cameras. A new sensor has been shown and allows to acquire at a reasonable speed, high resolution images. A method to compute the epipolar constraint has been introduced. Several minimizations methods have been presented, showing for each its performance. So far the study carried out ensures that the epipolar constraint is retrieved from two semi-calibrated panoramic stereo pair. This work has been used to generate a cinematographical set.

BIBLIOGRAPHY

[1]D.Southwell et al, "Panoramic stereo", I.C.P.R, pages 378-382, August 1996, Vienna, Austria.

[2]T.Sogo, H.Ishiguro, "Real-time tagert localization and tracking by N-ocular stereo", pages 153-160, I.E.E.E Workshop on Omnidirectional Vision, Hilton Head Island, June 2000.

[3]E.Brassart et al., "Experimental, results got with the omnidirectional vision sensor: SYCLOP", pages 145-152, I.E.E.E Workshop on Omnidirectional Vision, Hilton Head Island, June 2000.

[4]C.Drocourt et al., "Mobile robot localization based on a omnidirectional stereoscopic vision perception system", pages 1329-1334, I.E.E.E International Conference on Robotics and Automation, Detroit,Michigan, May 1999.

[5]H.Ishigureo, M.Yamamoto,S.Tsuji, "Omni-directional stereo", I.E.E.E Transactions on P.AM.I, vol. 14, No.2, pages 257-262, 1992.

[6]Y.Pritch, M.Ben-Ezra, S.Peleg, "Automatic disparity control in stereo panoramas (Omnistereo)",pages 54-61, I.E.E.E Workshop on Omnidirectional Vision, Hilton Head Island, June 12, 2000.

[7]R.Benosman, J.Devars, "Panoramic stereovision sensor", I.C.P.R, pages 767-769, August 1998,Vienna, Austria.

[8]R.Benosman, T.Manière, J.Devars, "Multi directional stereovision sensor, calibration and scene reconstruction", ICPR, Track A: 161-165, August 1996.

[9]T.Svoboda, T.Pajdla, V.Hlavac, "Epipolar geometry for panoramic cameras", 5th European Conference on Computer Vision, 1998.

[10]L.Mc Millan, "Acquiring immersive virtual environments with an uncalibrated camera", UNC Technical report 95-006

[11]S.B Kang, R. Seliski, "3d scene recovery using multidirectional multibaseline stereo", I.J.C.V, 25(2), November 1997.

[12]E.Bigorgne, C.Achard, J.Devars, "Local Zernike moments vector for content-based queries in image data base", Machine Vision Application, December 2000.