PANORAMA-BASED CAMERA CALIBRATION

Bertrand Cannelle, **Nicolas Paparoditis**, **Olivier Tournaire**

Université Paris-Est, Institut Géographique National, MATIS Laboratory
73 Avenue de Paris, 94165 Saint-Mandé Cedex, France
firstname.lastname@ign.fr

Commission III/1

KEY WORDS: Camera calibration, panorama, self-calibration, bundle adjustment

ABSTRACT:

This paper presents a method to calibrate a camera from panoramas. Camera calibration using panoramas has two main advantages: on the one hand it requires neither ground control points or calibration patterns and on the other hand the estimation of intrinsic and distortion parameters is of higher quality due to the loop constraint and to a decorrelation of tied parameters due to the fixed perspective center.
The paper is organised as follow. The first part presents the acquisition process and our mathematical estimation framework. The second part explores with simulated data sets the impact of noisy measures, of geometry of acquisition and of unmodelled parallaxes on the calibration results. A comparison with a traditional calibration method (i.e by using a 3D target network) is then studied. The final section presents results in a real case and compares the results obtained with our panorama approach against the classical calibration. The results are very promising.

INTRODUCTION

In photogrammetric surveys, the camera calibration is most of the time performed prior to the survey and the extrinsic parameters of the poses of the survey are obtained by bundle adjustment. A "classical" photogrammetric camera calibration methods (Tsai, 1986, Zhang, 2000) consists in taking images of a topometrically surveyed 2D or 3D target network, in measuring manually or automatically the positions of the projection of the targets (the observations) in image space and finally in estimating the set of parameters of a mathematical projection model (usually the collinearity equation) minimising the distance in image space between the observations and the projection of the corresponding targets given the set of parameters. In "classical" surveys where images are parallel to the surfaces of the objects or landscapes to be surveyed, the extrinsic parameters determined through a bundle adjustment can absorb/compensate errors of the camera calibration. In image sets with loops, like in panoramas or when turning around objects, these errors can unfortunately not be compensated. In order to perform a better estimation and decorrelation of intrinsic (and distortion parameters) and extrinsic parameters, some other techniques have been developed using rotating images (Hartley, 1994), or using panoramas (Ramalingam et al., 2010). Some woks using the same acquisition framework already exist ((Agapito et al., 2001, Tordoff and Murray, 2004)). However the distortion modeling is different than ours.

Our calibration approach consists in carrying out a self-calibration from panoramas, i.e. to estimate intrinsic and extrinsic parameters at the same time while closing a loop and with a fixed perspective center to decorrelate some tied parameters and limit the number of unknowns to estimate (we only need to estimate a rotation between our different images). This approach has many advantages: it is fully automatic, it does not need a qualified operator to acquire images with a "good geometry" (with targets in the corner, etc.), it does not need any ground control point and calibration patterns (any detail or texture of a scene becomes a tie point) and it is thus ultra-portable. Indeed, the calibration can be realised close to the survey thus for example in the same thermal conditions knowing that temperature has a relatively strong impact on the intrinsic and the distortion parameters.

Our panoramas are acquired with a low cost motorised pan-tilt device thus with a gross angular accuracy (around 0.1) which is insufficient to measure the perspective bundle in a direct way (ray by ray by observing a point while turning the camera) but which is sufficient enough to provide initial solutions for rotations and limit the search space for homologous points.

Our work present a method to calibrate camera without ground point. One of the main advantage to work in a panoramic geometry is that we only needs to estimate a rotation between images. Another interesting particularity is that it requires neither ground points nor geometric information extracted from the images.
We will first start by presenting our acquisition process, our geometric camera model, and our mathematical estimation framework (Section 1). Then we will present some experiments with synthetic data to study the influence of noise on the estimation of intrinsic parameters and distortion (Section 2). A comparison with a "classical" calibration with ground control points will then be presented in Section 3. Finally, Section 4 presents results on a real dataset.

1 OUR METHODOLOGY

In this section we present our calibration framework. We first discuss the acquisition process and the matching of tie points between images. Then, we present our camera geometrical model and introduce the mathematical bundle adjustment framework in which the calibration is embedded, and we explain how to solve it.

1.1 Acquisition process

Our pan-tilt device which is controlled by a computer is presented on Fig. 1. Mechanically, it provides two rotation axes (from left to right and top to bottom) and can support any reasonable weight camera. In our framework, images are taken with an overlap ratio around 50% in order to have enough measures of homologous

points on the neighbouring images. For instance, for a camera with a field of 90°, there are 8 images at the equator, 6 images at ± 45° and one at each pole, a total of 22 images.

Figure 1: The panoramic images acquisition system. It can support any kind of camera.

For our calibration process, we need to have all the images acquired from the same point of view. To minimize manually the parallax due to the mounting of the camera on the acquisition device, one takes a pair of rotated images with near and far objects. If the two objects remain in the same order in both images, there is no or only a small parallax. If the order is changed, there is some parallax. Fig. 2 shows two images taken with our system.

Figure 2: Rotation with a decentered camera around pan-tilt axes

We can see that the order of the wire in the foreground and the TV antenna changes between the two images. Thus, the position of the camera on the device is shifted manually until this effect disappears.

1.2 Description of our matching process

To compute homologous points, you can use different strategies. For example by extracting and matching interest points such as SIFT point (Lowe, 2004). We have used also a matching of homologous points in two neighbouring images based on anonymous features (Craciun et al., 2009), following a two steps approach. In a similar way to the process describe in (Coorg and Teller, 2000), the first step of our pose estimation method consists in finding for each pair of overlapping images the rotation which optimises the Normalised Cross Correlation similarity score computed on the overlap of the first image with a rigid transform of the second image to put it in the geometry of the first. The optimisation is performed in a greedy way within a multi-resolution framework.

In our case, the acquisition system directly gives an initial solution and bounds for the rotation. Then, we use the refined rotation to initialise the matching of anonymous points, i.e. points on a grid covering the overlap, based on the similarity of image patches centred at the considered points. Indeed, these homologous points are necessary to feed a photogrammetric bundle adjustment to estimate accurately the relative pose of all the images within the panorama. The main advantage of this method is that it finds corresponding points in any situation, even if the surface is uniform or regular. The second advantage is that it chooses the number of corresponding points per images as well as the repartition of these points.

1.3 Our mathematical calibration model

Most calibration techniques try to minimize the residuals in the image space of known points, or geometrical characteristics extracted from images. In our case, we only want to minimize the angle between the corresponding photogrammetric rays of homologous points directly in panoramic space (see eq. 1). This explains why our process does not require ground points.
Each image is fixed in the panoramic space by a rotation noted $R_{i,p}$. Our perspective model contains the Principal Point of Autocollimation (the intersection of the focal plane with the optical axis) of coordinates (c_{PPA}, l_{PPA}) and the focal length denoted f.

1.3.1 Ray in 3D space To transform a point (c, l) in image coordinates to a ray (x', y', z') in panoramic space, we use a function g which depends on $R_{i,p}$, f and (c_{PPA}, l_{PPA}) (see eq. 1).

$$g\,(c,l) = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \frac{R_{i,p} \begin{pmatrix} x \\ y \\ z \end{pmatrix}}{\sqrt{x^2+y^2+z^2}} \tag{1}$$

where:

$$\begin{cases} x &= c - c_{PPA} \\ y &= l_{PPA} - l \\ z &= -f \end{cases}$$

1.3.2 Distortion modeling We consider an additive radial distortion model which amplitude is modelled by a polynomial of the form $p(r) = a.r^3 + b.r^5 + c.r^7$ where r is the radius centred on the Principal Point of Symmetry (c_{PPS}, l_{PPS}) which is different from the PPA.
Eq. 2 shows how to compute a corrected measure from a real measure, where (c_b, l_b) is the measure directly taken in image space, and (c_c, l_c) is the corrected measurement:

$$\left.\begin{aligned} r &= \sqrt{(c_b - c_{PPS})^2 + (l_b - l_{PPS})^2} \\ dr &= ar^2 + br^4 + cr^6 \\ c_c &= c_b + (c_b - c_{PPS})dr \\ l_c &= l_b + (l_b - l_{PPS})dr \end{aligned}\right\} \tag{2}$$

In this process, we must ensure that the model is not correlated with intrinsic parameters. For example, a model of purely linear radial distortion shows that an increase of the focal length has the same effect as reducing the coefficient of distortion. The system thus cannot converge or converges to a mathematical minima which is physically incorrect.

1.4 System to solve

To compute the pose of images in a panorama and calibrate a camera, one must find a set of parameters: rotation of each image in the panoramic system $(R_{1,p}, \cdots, R_{N,p})$ and intrinsic parameters (f and the PPA image coordinates (c_{PPA}, l_{PPA})). This can

be achieved by minimising eq. 3:

$$\underbrace{argmin}_{R_{1,p}\cdots R_{N,p},f,PPA}\left(\sum_{c=0}^{N}\left(g_i\left(c,l\right)-g_j\left(c,l\right)\right)^2\right) \quad (3)$$

where:

- N is the number of couples of homologous points
- $f_i\left(c,l\right)$ (*resp.* $f_j\left(c,l\right)$) is the function defined in eq 1 applied to image i (*resp.* j).

Of course, if there is some distortion, we must also find the parameters of the distortion function

$$d_{(c_{PPS},l_{PPS}),a,b,c}(c_b,l_b)=(c_c,l_c) \quad (4)$$

The cost function is minimized with a least squares optimization. Indeed, we are in a favourable case, since we can eliminate outliers using our initial solution provided by the pan-tilt system. As outliers are the main drawback of this kind of optimization, this choice is suitable for our application. Furthermore, as the calibration process does not need to be performed often, computation time is a minor concern.
For example, with 3 images and some homologous points between images $\{1,2\}$, $\{2,3\}$ and $\{1,3\}$, the system is defined by the matrix A with the unknown X in eq 5.

$$h=g_i\left(c,l\right)-g_j\left(c,l\right)$$

$$^tA=\begin{pmatrix}
\frac{\partial h}{\partial f} & \frac{\partial h}{\partial f} & \frac{\partial h}{\partial f}\\
\frac{\partial h}{\partial PPA} & \frac{\partial h}{\partial PPA} & \frac{\partial h}{\partial PPA}\\
\frac{\partial h}{\partial PPS} & \frac{\partial h}{\partial PPS} & \frac{\partial h}{\partial PPS}\\
\frac{\partial h}{\partial a} & \frac{\partial h}{\partial a} & \frac{\partial h}{\partial a}\\
\frac{\partial h}{\partial b} & \frac{\partial h}{\partial b} & \frac{\partial h}{\partial b}\\
\frac{\partial h}{\partial c} & \frac{\partial h}{\partial c} & \frac{\partial h}{\partial c}\\
\frac{\partial h}{\partial R_{1,p}} & 0 & \frac{\partial h}{\partial R_{1,p}}\\
\frac{\partial h}{\partial R_{2,p}} & \frac{\partial h}{\partial R_{2,p}} & 0\\
0 & \frac{\partial h}{\partial R_{3,p}} & \frac{\partial h}{\partial R_{3,p}}
\end{pmatrix}
X=\begin{pmatrix}
df\\ dPPA\\ dPPS\\ da\\ db\\ dc\\ dR_{1,p}\\ dR_{2,p}\\ dR_{3,p}
\end{pmatrix} \quad (5)$$

The cost function defined by eq. 3 leads to solve the system of eq. 6.

$$A_nX_n=0 \quad (6)$$

In eq. 6, n denotes the iteration index. Solving this system is thus done iteratively using a least squares approach. For the first iteration ($n=0$), we use a rough estimate of the parameters: f can take any arbitrary value, PPS and PPA are initialized at the center of the image, distortion coefficients are null and rotations are given by the pan-tilt system. The convergence of the minimisation process is detected when there are no changes of the parameters between two consecutive iterations.

2 EXPERIMENTS WITH SYNTHETIC DATA

2.1 Simulation protocol

To validate our calibration process regardless of possible mounting errors on the acquisition system, we worked with synthetic data. This provides a ground truth to evaluate the calibration parameters that we have computed with our minimisation scheme. We have thus simulated a 3000×2000 pixels camera[1] with 2 different focal lengths:

- a 1000 pixels short focal length (90° per 112°).
- a 3000 pixels long focal length (36° per 53°).

A points dataset is then created by regularly sampling points on a 10 meters radius sphere (see Fig. 3). Those sampled points are then back-projected in all the images to simulate a real acquisition (c_i,l_i).
The camera is positioned at the center of the sphere to simulate

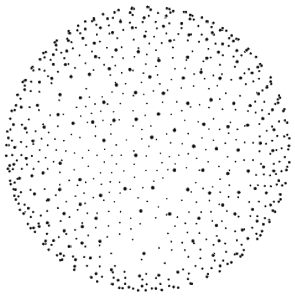

Figure 3: Sample points on sphere

images. We used a 50% overlap in each dimension. It leads to an angle of 61° between horizontal images and 45° between vertical images (see Fig. 4).

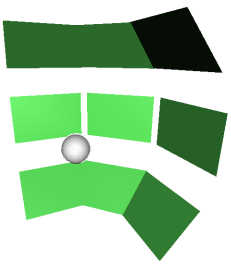

Figure 4: Illustration of the simulation protocol with nine images (in green) and the camera (in white).

In all experiments presented in this section, we have added some noise following a normal distribution. The noise is centred on the true measure and its standard deviation is chosen in the set $\{0.3;0.5;1.0;2.0\}$ pixels.

2.2 Intrinsic parameters

We have first examined the influence of noise on the estimation of the focal length and of the PPA. The unknowns to be estimated are given by eq.7:

$$\left[\ R_{1,p},\ldots,R_{n,p},f,c_{PPA},l_{PPA}\ \right] \quad (7)$$

[1] $(c_{PPA},l_{PPA})=(1470,980)$, $(c_{PPS},l_{PPS})=(1530,1020)$ and $a=10^{-8}$, $b=10^{-15}$, $c=10^{-21}$

		noise (in pixel)			
		0.3	0.5	1.0	2.0
Cam. 1	Δ f	0.004	0.006	0.005	0.04
	$\Delta\ c_{PPA}$	0.010	0.020	0.001	0.20
	$\Delta\ l_{PPA}$	0.022	0.036	0.014	0.128
Cam. 2	Δ f	0.013	0.022	0.052	0.021
	$\Delta\ c_{PPA}$	0.046	0.077	0.215	0.308
	$\Delta\ l_{PPA}$	0.166	0.276	0.174	0.348

Table 1: Influence of noise on the intrinsic parameters on simulated data.

Tab. 1 shows the results that we obtained in our experiments. They show that our method can efficiently estimate the intrinsic parameters of the camera and that the noise only has a very slight influence on the calibration process. The results are quite similar for both short and long focal lengths. Even if there is an important difference between the results for the two cameras, the calibration can be considered very good since the parameters are estimated with an error bounded by 0.2 pixel.

2.3 Distortion function

Most of the cameras have a radial distortion which can be quite well modelled by a polynomial (see eq. 2). The unknowns to be estimated are thus:

$$\left[\ R_{1,p}, \ldots, R_{n,p}, f, c_{PPA}, l_{PPA}, c_{PPS}, l_{PPS}, a, b, c\ \right] \quad (8)$$

Tab. 2 summarizes the results of the tests that were conducted for Camera 1. The intrinsic parameters (f and PPA) are very close to real values (about 0.1 pixel with 1 pixel noise). The error on the distortion parameters between the distortion function and the estimated model is around 1 pixel in the image corners for a noise of 1 pixel.

	noise (in pixel)			
	0.3	0.5	1.0	2.0
Δ f	0.01	0.02	0.03	0.63
$\Delta\ c_{PPA}$	0.04	0.09	0.12	0.21
$\Delta\ l_{PPA}$	0.06	0.11	0.09	0.14
$\Delta\ c_{PPS}$	1.27	2.36	2.64	9.18
$\Delta\ l_{PPS}$	1.59	2.67	3.75	0.70
Δ_{pixels} (image border)	0.15	0.23	1.08	4.7

Table 2: Influence of noise on the intrinsic parameters and distortion parameters on simulated data (camera with f=1000).

Tab. 3 summarizes the results for Camera 2. We can note that the intrinsic parameters (f and PPA) are very close to real values. The error on the distortion parameters between the distortion function and the estimated model is around 0.5 pixel in the image corners for a noise of 1 pixel.

2.4 Influence of parallax

To study the influence of the the parallax, we have simulated a failure in the position of the camera nodal point compared of the center of the pan-tilt system. The magnitude of this defect is in the range $[-5; 5]$ centimetres on each of the 3 X, Y and Z components. Fig. 5, 6 and 7 represent the variations of f, c_{PPA} and l_{PPA} when X, Y or Z evolve. These results are just for Camera 2.

	noise (in pixel)			
	0.3	0.5	1.0	2.0
Δ f	0.07	0.11	0.02	0.91
$\Delta\ c_{PPA}$	0.02	0.04	0.47	1.80
$\Delta\ l_{PPA}$	0.68	1.14	1.59	2.96
$\Delta\ c_{PPS}$	0.04	0.08	0.52	3.24
$\Delta\ l_{PPS}$	1.39	2.33	3.81	6.98
Δ_{pixels} (image border)	0.14	0.24	0.48	0.72

Table 3: Influence of noise on the intrinsic parameters and distortion parameters on simulated data (camera with f=3000).

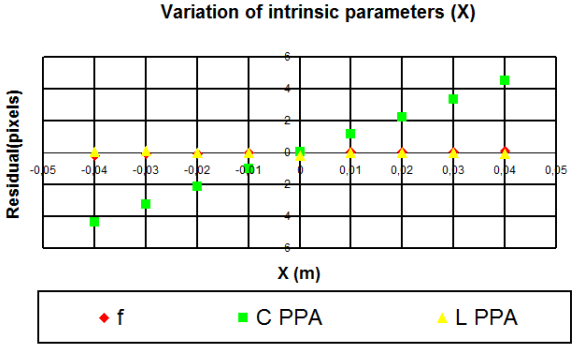

Figure 5: $f(X), c_{PPA}(X), l_{PPA}(X)$

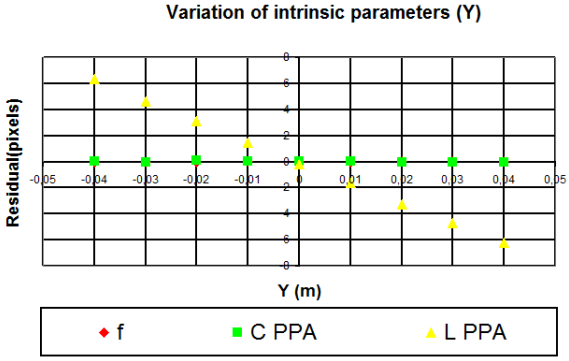

Figure 6: $f(Y), c_{PPA}(Y), l_{PPA}(Y)$

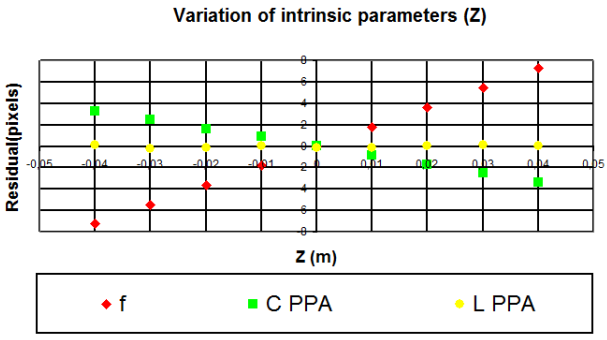

Figure 7: $f(Z), c_{PPA}(Z), l_{PPA}(Z)$

		Δ f	$\Delta\ c_{PPA}$	$\Delta\ l_{PPA}$
Camera 1	min	-9.19	-8.97	-6.57
	max	7.39	8.61	8.10
Camera 2	min	-1.18	-1.23	-1.03
	max	1.18	1.09	1.26

Table 4: Influence of parallax on the intrinsic parameters on simulated data.

When we vary the parallax on the 3 axes simultaneously, we get

errors that are reported in tab 4.

Through these experiments, we find that over the length is long, at least the parallax effect on the calibration.

2.5 Conclusion

In the various tests that we have performed, we have found that the calibration of a camera through a panorama acquisition process is relatively stable and reliable. We will now compare our method with a traditional calibration method.

3 COMPARISON WITH A TRADITIONAL METHOD

To evaluate our algorithm, we have compared it with a more traditional method on the same simulated data. In this section, the camera is calibrated with the same set of observations by a traditional method by estimating the parameters R and S (respectively rotation and translation) for each image by using points on the sphere as ground control points noted M. The cost function to minimize is:

$$\begin{pmatrix} c \\ l \end{pmatrix} - \left(\begin{pmatrix} c_{PPA} \\ l_{PPA} \end{pmatrix} - \frac{fR(M-S)}{(0,0,1).R(M-S)} \right) \quad (9)$$

3.1 Influence on parameters

Camera 1 Tab. 5 shows the difference between the true parameter and the calculated parameters. There is no significant difference between intrinsic parameters and distortion parameters for the calibration of a short focal camera.

	noise (in pixel)			
	0.3	0.5	1.0	2.0
Δ f	0.04	0.07	0.16	1.,55
$\Delta\ c_{PPA}$	0.09	0.14	0.27	0.95
$\Delta\ l_{PPA}$	0.34	0.54	0.89	15.93
$\Delta\ c_{PPS}$	0.48	0.84	1.74	3.84
$\Delta\ l_{PPS}$	0.92	1.6	3.15	0.54
Δ_{pixels} (image border)	0.83	1.05	1.68	34.4

Table 5: Influence of noise on the intrinsic parameters and distortion parameters

Tab. 6 consolidates the results obtained with our method and with the traditional method. Our method is more accurate in the estimation of intrinsic parameters than the traditional method. It is more difficult to compare the estimation of distortion parameters.

	noise (in pixel)			
	0.3	0.5	1.0	2.0
f	+	+	+	+
ppa	+	+	+	+
distortion	+	+	+	+

Table 6: Comparison between estimation with our method and the traditional method. A "+" (*resp.* "-") indicates that our method is more (*resp.* less) accurate than the traditional method.

	noise (in pixel)		
	0.3	0.5	1.0
$\Delta\ f$	31.88	47.48	71.81
$\Delta\ c_{PPA}$	1.35	2.09	3.39
$\Delta\ l_{PPA}$	2.21	3.59	6.75
$\Delta\ c_{PPS}$	2.46	4.33	11.74
$\Delta\ l_{PPS}$	1.47	2.46	4.99
Δ_{pixels} (image border)	4.23	6.47	10.21

Table 7: Influence of noise on the intrinsic parameters and distortion parameters

Camera 2 Tab. 7 shows the difference between actual parameters and the estimated parameters. Note that with little noise (0.3 pixel) there is a significant error on intrinsic parameters (1% of error on the focal). One can note in tab.8 that our method is more accurate than the traditional method. This table is just a qualitative summary but when you take a look on tab.7, you can see traditional method is not very accurate to calibrate long focal!

	noise (in pixel)		
	0.3	0.5	1.0
f	+	+	+
ppa	+	+	+
distortion	+	+	+

Table 8: Comparison between estimation with our method and traditional method. See Tab. 6 caption for an explanation.

3.2 Impact on the projection center

The main difference between a traditional calibration method and our method is that the estimation of the position and rotation of each image is done separately. This section discusses the difference between the simulated position of the camera at (0,0,0) and the position estimated in the calibration process.

Camera 1 The figure 8 represents the projection centres of each image after calibration (top view). We found that when increasing the noise, the calculated positions differ from the true position. The same behaviour can be noted for the third component on fig. 9

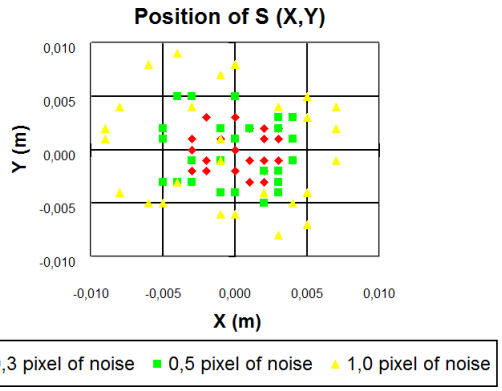

Figure 8: Projection center of images in plan (X,Y) after calibration and noise on measures

Camera 2 The same behaviour can be observed for the short and the long focal, but the difference between estimated position and true position is more important (Fig. 10 and Fig.11).

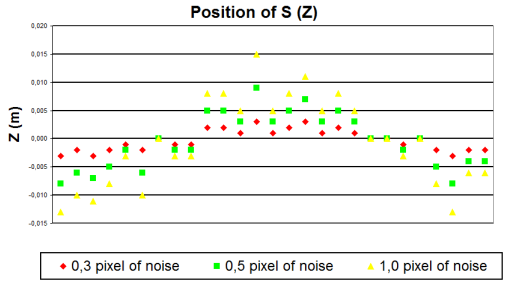

Figure 9: Projection center of images in Z after calibration and noise on measures

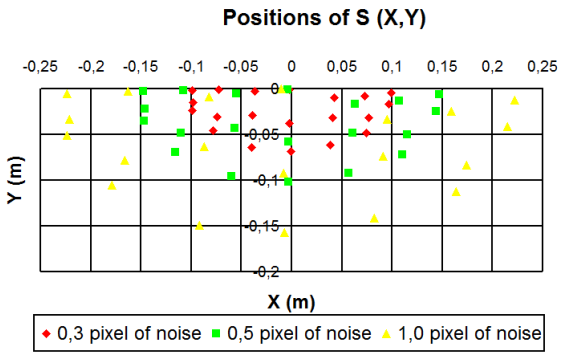

Figure 10: Centres of projection of images in plan (X,Y) after calibration and noise on measures

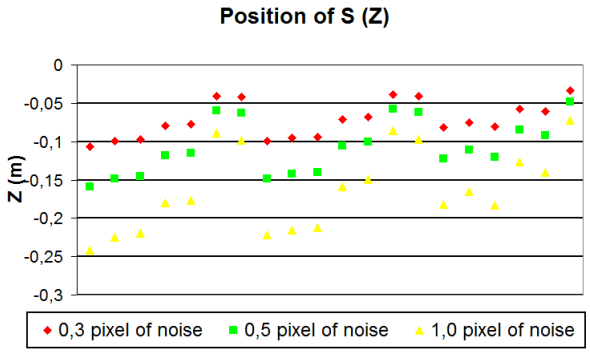

Figure 11: Centres of projection of images in Z after calibration and noise on measures

3.3 Conclusion

In this section, we exhibited the limits of the polygon based calibration and the difficulty to estimate the intrinsic and extrinsic parameters in the same process. The geometry of the panoramic avoids calculating the position of each image. Furthermore, geometry of panorama constraint naturally and geometrically intrinsic parameters.

4 EXPERIMENTS WITH REAL DATA

After working with simulated data, we have calibrated a real camera. The interest of this acquisition is to validate the convergence of the process on real data. We also compare our results with a calibration on a topometrically surveyed network targets.
After a bundle adjustment by estimating only rotation and the camera calibration obtained on targets network, we have a value of RMS equal to 0.021. With our method, we have 0.015 RMS, i.e. 28% better. The focal length difference between two techniques is more important than 18 pixels. The difference between 2 models can be more than 10 pixels (in the corners). It means an difference on field of camera of 0.7123° in column and per 0.5989° in line.

Figure 12: Sample of panorama

5 CONCLUSIONS AND FUTURE WORK

This paper has presented a method to calibrate a camera's intrinsic parameters but also to estimate a distortion pattern. We have shown that this method is not sensitive to noise and it is applicable with both short and long focal. Future work will consist on looking for more complex and accurate model. We also look for a framework to estimate the mechanical residual parallax.

REFERENCES

Agapito, L., Hayman, E. and Reid, I., 2001. Self-calibration of rotating and zooming cameras. International Journal of Computer Vision 45(2), pp. 107–127.

Coorg, S. and Teller, S., 2000. Spherical mosaics with quaternions and dense correlation. In: International Journal of Computer Vision 37(3), pp. 259–273.

Craciun, D., Paparoditis, N. and Schmitt, F., 2009. Automatic gigapixel mosaicing in large scale unstructured underground environments. In: IAPR Machine Vision Applications, Yokohama, Japan.

Hartley, R. I., 1994. Self-calibration from multiple views with a rotating camera. In: ECCV '94: Proceedings of the third European conference on Computer vision (vol. 1), Secaucus, NJ, USA, pp. 471–478.

Lowe, D. G., 2004. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision 60(2) pp. 91–110.

Ramalingam, S., Sturm, P. and Lodha, S. K., 2010. Generic self-calibration of central cameras. Computer Vision and Image Understanding 114(2), pp. 210–219. Special issue on Omnidirectional Vision, Camera Networks and Non-conventional Cameras.

Tordoff, B. and Murray, D. W., 2004. The impact of radial distortion on the self-calibration of rotating cameras. International Journal of Computer Vision 96(1), pp. 17–34.

Tsai, R., 1986. An efficient and accurate camera calibration technique for 3d machine vision. In: IEEE Computer Vision and Pattern Recognition, Miami, FL, USA, pp. 364–374.

Zhang, Z., 2000. A flexible new technique for camera calibration. In: IEEE Transactionson Pattern Analysis and Machine Intelligence 22(11), pp. 1330–1334.