What Accuracy for 3D Measurements with Cameras?

Giannoula Florou, Roger Mohr

What Accuracy for 3D Measurements with Cameras?

Giannoula FLOROU and Roger MOHR

GRAVIR-INPG-INRIA
655 avenue de l'Europe
F-38330 Montbonnot FRANCE Cedex
E-mail: florou@macedonia.uom.gr mohr@imag.fr

Abstract

We estimate the internal and external parameters of the camera, and simultaneously the distortion's parameters. Our aim is the selection of the best distortion model, using stiatistical test for the importance of distortion parameters. Also we examine the accuracy in camera parameter estimation and 3D reconstruction, in relation with the noise in the image. We answer the question "until which level of noise, is it possible to obtain a good camera parameter estimation, and from there a good reconstruction?" Experiments are evalueted on simulated data and finally a test is performed with real data.

1. Introduction

1.1. The problem

We have one or more images, taken by the same camera, of an objet in the 3D space. We know the 3D coordinates of some points in the objet, but we do not know the position of the camera in the 3D coordinates system. Camera calibration is the estimation of the parameters in the model for the transform between the world (3D) and camera coordinate (2D) frame. It is a key issue for deriving metric information from images. We assume that the 2D points (points in image) are known, with an accuracy modeled as noise in the image. We want to find the relationship between image coordinates and 3D coordinates, i.e. the complete model for the transformation between the world and camera coordinate frames. Our aim is the link of the accuracy in the estimation of camera parameters and the quality of reconstruction.

1.2. Related works

Many methods were proposed for this calibration process; the reader is refered to [10] for a review. Methods can be divided into the following categories:

1. Methods that use geometric properties. They use objects whose images have some characteristic that are invariant to the actual position of the object in space and can be used to calibrate some of the internal camera parameters [4].
2. Methods that don't require known calibration points but they require camera motion and more than one image [7].
3. Methods that use known world points. This is the standard calibration method and it works with well designed calibration frame [9]. We will use such a method in section 3. The difference of our approach, is the distortion parameters estimation done simultaneously with the camera calibration.

Works about the importance and precision of distortion parameters are refered in [2].

2. Modeling the Camera

2.1. The pin hole model

The standard camera model is the pin hole model of perspective projection. Given the position of a point in 3D world coordinates, the model predicts the position of the point's image in 2D pixel coordinates.

Calibration data for the model consists of 3D (x, y, z) world coordinates of a feature point and corresponding 2D (u, v) coordinates (in pixels) of the feature point in one or more images. The image plane is parallel to the (X,Y) plane and the Z axis coincides with the optical axis. Then the transform between the world and camera coordinate frames, is represented by a 3x4 matrix M, and it can be writen in homogeneous coordinates, as:

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = M \begin{pmatrix} x \\ y \\ z \\ s \end{pmatrix} \quad (1)$$

1015-4651/96 $5.00 © 1996 IEEE
Proceedings of ICPR '96

where: M=

$$\begin{pmatrix} a_u & 0 & u_o \\ 0 & a_v & v_o \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} R_{11} & R_{12} & R_{13} & t_1 \\ R_{21} & R_{22} & R_{23} & t_2 \\ R_{31} & R_{32} & R_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The model (without distortion) for this transormation, has 10 parameters. Four internal for image scaling constants (a_u,a_v) and the image centre position (u_o,v_o), and six external parameters for the camera coordinate system's position and orientation relative to the world system. The external parameters are :
e_1, e_2, e_3 the rotation angles if we parametrise the rotation matrix R between world coordinate system and camera coordinate system, with Euler angles.
t_1, t_2, t_3 the translation components for the transform between the world and camera coordinate frames.

2.2. Modeling the distortion

Distortion provide systematic correlated errors in the image position. It is usually modeled with 5 parameters [4], [?]. These parameters are:

k_1, k_2, k_3 for radial lens distortion coefficients

P_1, P_2 for decentering distortion coefficients

The formula is

$$u = a_u \frac{R_{11}x + R_{12}y + R_{13}z + t_1}{R_{31}x + R_{32}y + R_{33}z + t_3} +$$
$$u_o + k_1\bar{u}r^2 + k_2\bar{u}r^4 + k_3\bar{u}r^6 + P_1(2\bar{u}^2 + r^2) + 2P_2\bar{u}\bar{v}$$
$$v = a_v \frac{R_{21}x + R_{22}y + R_{23}z + t_2}{R_{31}x + R_{32} * y + R_{33}z + t_3} +$$
$$v_o + k_1\bar{v}r^2 + k_2\bar{v}r^4 + k_3\bar{v}r^6 + 2P_1\bar{v}\bar{u} + P_2(2\bar{v}^2 + r^2) \quad (2)$$

where:
$\bar{u} = u - u_o \quad \bar{v} = v - v_o \quad r^2 = \bar{u}^2 + \bar{v}^2$

There are alternative non parametric methods for correcting the distortion [3]. They correct the distortion of points in the image, before the proccess of calibration, and they use a simple geometric property: straight lines in the space have to be projected as straight lines in the images.

In our estimation, we do not know relations between points, and we estimate the model for camera parameters and distortion parameters, simultaneously. As a result this provide a complete camera model for the transformation between the world (3D) and camera coordinate (2D) frames.

3. Estimation method

We estimate the model by a standard least suare fitting which minimizes the error in the image location of points with cost function:

$$\sum_{i=1}^{n}[(u_i - u_i')^2 + (v_i - v_i')^2] \quad (3)$$

where:
n the number of points,
u_i, v_i are the predict coordinates of the point i in image frame (pixel),
u_i', v_i' are the mesured coordinates of the point i in image frame (pixel).

As such a minimization is non linear problem, we use a non linear optimisation program based on Leverberg-Marquardt algorithm [8].

Parameter estimation is done by minimising (3). This can be performed using the full model given by equations (2).

From (2), we can estimate the distortion model and the camera parameters, simultaneously. Such a method requires however a non linear minimization and therefore it relies on an initial estimate which should be close enough to the final solution.

We use the Faugeras-Toscani [5] method for computing the matrix M and then, using QR decomposition of this matrix, we have an initial estimate for the camera parameters. Assuming no distortion, the distortion parameters are set to zero. After this step, Leverberg-Marquardt algorithm is used to find the optimal solution minimizing the image-coordinate error (cost function (3)).

4. Experiments with simulated data

For the synthetic data, they are 60 point in 3D space. The 60 points were randomly scatted in a sphere of radius 14 unit. The cameras were given random orientations and were placed at various distances from the centre of the sphere with a mean distance 70 units. The focal length is 15mm and the scene size is almost 45 units. These 3D points are projected in the image, and theirs locations are perturbed with different gaussien noises with mean 0 and standard deviation of 0.05, 0.1, 0.2 or 1.0 pixel. Furthemore, we added to each point (u, v), radial distortion (δ_u, δ_v) obtained from the following formulaes:

$$\delta_u(u, v) = k_1(u - u_o)r^2$$
$$\delta_v(u, v) = k_1(v - v_o)r^2 \quad (4)$$

where (u_o, v_o) is the coordinates of the principal point,
r^2 is the distance between the principal point and (u, v), and
k_1 is the distortion parameter.

For each level of noise 10 simulated images were generated. From them the parameters and the mean and standard deviation for each parameter were estimated, from the results.

4.1. Case of a single camera with distortion

noise	a_u	a_v	u_o	v_o	e_1	e_2	e_3	t_1	t_2	t_3
–	-1009	728	300	400	124.5	-3.7	15.9	1.5	-15.6	70.0
0.05	-985.5	709.2	284.5	425.5	124.7	-4.9	13.9	0.4	-18.1	67.7
std	0.58	0.43	0.39	0.30	0.004	0.04	0.01	0.03	0.03	0.04
0.1	-985.1	708.9	284.0	425.3	124.7	-4.9	13.9	0.4	-18.0	67.7
std	2.01	1.40	1.13	0.69	0.008	0.07	0.04	0.08	0.07	0.14
0.2	-986.8	710.2	283.4	425.1	124.7	-4.9	13.9	0.4	-18.0	67.8
std	3.76	2.58	1.69	1.03	0.012	0.13	0.05	0.12	0.10	0.25
1.0	-991.3	712.8	281.2	422.5	124.7	-4.7	13.9	0.2	-17.8	68.1
std	12.14	8.89	5.86	6.33	0.061	0.49	0.24	0.40	0.62	0.85

Table 1. Estimation of camera parameters without distortion model for synthetic data having a maximum distortion of 5.0 pixels.

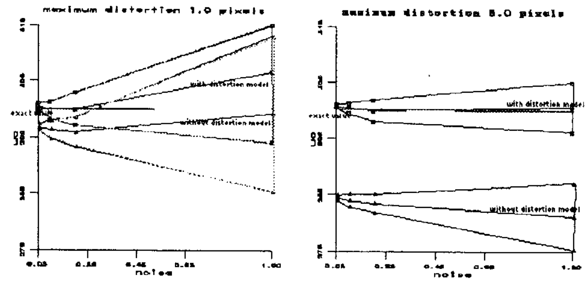

Figure 1. Estimation of uo with and without model for data with maximum distortion 1.0 (left) and 5.0 (right) pixels

We simulated camera with radial distortion 0.5, 1.0 or 5.0 pixel at the image corner (cf. equations (4)). We have estimated the camera model without distortion parameters. The results of camera parameters estimation, are summurized in the table 1. Notice that the camera parameters are not correctly estimated. The confidence interval of parameters (mean values $\pm$ 3 standard deviation) doesn't include the true value. This because the distortion has an influence in the model. Distortion parameters and camera parameters are also highly correlated.

Introducing a model including the distortion parameters, the estimations are slightly better. However, we can find smaller errors with smaller standard deviation for the parameters u_o, v_o, without using model of distortion, when the noise (1.0 pixel) is larger than the distortion 1.0 pixel at the image corner.

Figure 1 displays the estimation for u_o with its standard deviation, in the case of small (1 pixel maximum) and median distortion (5 pixels maximum). Considering the case of small distortion, it has to be noticed that, when estimating u_o with a model without distortion parameters, the result is better than with distortion parameters, as the noise reaches at $\sigma = 1$ pixel (cf. left part of figure 1). The reason is that distortion effects are absorbed by the noise.

But when the distortion is larger,observing the right part of figure 1, we see that the parameter estimation using distortion model is much better.

4.2. Number of distortion parameters in the model

It is well known that radial distortion is the major distortion source; the question is to derive how much parameters are needed as we can have a model for distortion with 1, 2, 3, 4 or 5 parameters. If we have more parameters in the model, we obtain smaller error of estimation, but larger confidence interval. We will discuss this in section 7. We need to select between nombre of parameters in the model and accuracy of the estimation. For this reason, we use statistical tests.

5. Selection of right camera model

As we have seen, we need to select the parameters of distortion, which are significant. For this, we propose to validate them, using statistical tests.

We suppose that errors are normally distributed with mean 0. We are interested in finding the importance of the parameters of distortion in the model. For every parameter of distortion we make the folowing hypothesis:

Ho: we considere that the parameter i is not significant and is set to 0
H1: we considere that the parameter i is significant

In order to validate these hypothesis, we compute the confidence regions. A confidence region is a region of parameter values that contains a certain percentage of the total probability distribution. For example: there is a confidence level 95% that the true parameter values fall within this region around the measured value. The more the region is large, the more this percentage is large too. Regions are designed to have the smallest area (volume), for a given level of confidence. In case of gaussien distribution they are limited by confidence ellipses.

If we have only one estimation of parameters (real data) we can find the confidence intervals, for every parameter in the model, using the obtained variance-covariance matrix C, which is provided by the Leverberg-Marquardt algorithm, as an additional result.

For the parameter λ_i, the confidence interval $\alpha\%$ is given by the formula [1] :

$$\lambda_i \pm \delta\lambda_i$$

where:
$\delta\lambda_i = \sqrt{x^2_\alpha}\sqrt{\frac{f}{n}}\sqrt{C_{ii}}$
x^2_α is the value of χ distribution for the significant level $\alpha\%$
f the value of objective function with the estimated parameters
n the number of points
$\frac{f}{n}$] is the variance of the residuals
C the variance-covariance matrix $(D'D)^{-1}$ and
D is the jacobi matrix of the estimated parameters

In the results we computed the $\delta\lambda_i$ value for $\alpha = 90\%$ and we can see if the parameters of the distortion model are significant or not. If this value is bigger than the estimated value (that is 0 is included in the confidence region), the parameter is not significant (there is a probability only 10% that it is not equal to 0).

6. Quality of reconstructed points

If we estimate the parameters for at least two cameras, we can reconstruct the 3D points by triangulation [6]. We investigate here the accuracy of the 3D reconstruction from the accuracy of the camera parameters. We assume that only camera parameters are source of errors; no error is assumed in the image measurements for these steps. So when we refer to noise, this is the noise with which camera parameters were estimated. After correcting the distortion at every point, we find the line which is supposed that join the point in the image (u,v) with the correspondant 3D point (X,Y,Z).

The relative error of reconstruction is the mean of the difference of distance between two 3D points and the two points 3D reconstructed. We find the mean distance for all couple of points, and we divide it by the maximun distance. The formula is:

$$d = \frac{\frac{1}{\frac{n(n-1)}{2}} \sum_{i=1}^{n} \sum_{j=i+1}^{n} \left| \sqrt{d_e(P_i - P_j)} - \sqrt{d_e(P_i' - P_j')} \right|}{\max\limits_{i,j} \sqrt{d_e(P_i - P_j)}} \tag{5}$$

where:
P_i, P_j are the known 3D points
P_i', P_j' are the recontructed 3D points, and
$d_e(P_i - P_j)$ is the euclidean distance between the 3D points P_i, P_j.

The reconstruction depends on the relative position of the two cameras. If the angle between the two camera axis is too small (we call this angle, rotation angle), it is well known that the reconstruction degenarates. We made experiments using rotation angle between the two camera position, of 10, 20, 30, and 45 degrees. The Z-axis coincides with the optical axis and we rotate the camera around the Y-axis. The camera is also translated in order to observe the same scene. The reconstruction error is very small for angle 40 or 45 degrees, if the noise is 0.05, 0.1 or 0.2 pixels.

7. Real data

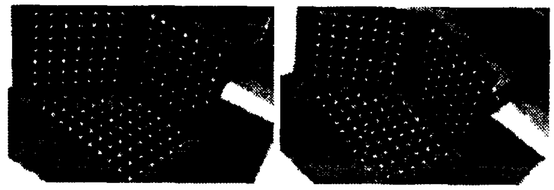

Figure 2. images of the calibrated reference frame

In figure 2, we see two images of a calibration frame. We find the locations of 3D points in the image with subpixel accuracy 0.05 pixel [3].

The results with real data are seen in the table 2, for real images (figure 2) with distortion. We don't know the true values of parameters. In order to validate the results, we compute only the reconstruction error, since we know the 3D points coordinates. We use the half part of the data to estimate the parameters, and the other half part to reconstruct the scene. Thus we have no influence of estimation process in the reconstruction.

We estimate first, a model without parameters of distortion. Secondly we consider the 5 distortion parameters, and finally only the significant distortion parameters are put into the model.

We did these estimations separately, for each image and also simultaneously for two images. In the estimation with both images, the parameter estimation is more accurate, as it relies on more data.

In table 2 we have two images taken with a lens having distortion. Remember that there is a * before a number if the correspondant parameter is considered as being signifiacant. In practice, we consider only parameters for which the $\delta\lambda_i$ is close to its estimated value. For the model without parameters, the reconstruction error is 0.70 10^{-3}. If we use model with 5 parameters of distortion, the error is only 0.55 10^{-3} and if we use model with only k_1, P_1, P_2 parameters, the error is 0.56 10^{-3}. Using the distortion model, we divide error by a factor close to 2. We have to select the model with k_1, P_1, P_2, where we have the smaller reconstruction error, and the smaller confidence interval too.

8. Conclusion

We have described the process of camera calibration and distortion model estimation, when we want to find an accur-

model	image	au	av	uo	vo	k1	k2	k3	P1	P2	error x10^{-3}
model	image 1	-1644.7	1117.2	280	247	0	0	0	0	0	
	$\delta\lambda_i$	29.6	20	16.1	7.3	0	0	0	0	0	0.71
without	image 2	-1614.7	1096.3	285.9	220	0	0	0	0	0	
	$\delta\lambda_i$	32	21	16.9	14.7	0	0	0	0	0	0.71
parameters	image 1&2	-1620.6	1101.3	281.7	239.2	0	0	0	0	0	
	$\delta\lambda_i$	34.2	23.1	19.8	11.5	0	0	0	0	0	0.8
model	image 1	-1636.2	1111.4	280.1	246.9	-3.99 e-08	5.23 e-13	-1.05 e-17	-2.56 e-06	-4.06 e-07	
	$\delta\lambda_i$	54.7	36.9	99.6	32	4.11 e-07	1.15 e-11	9.40 e-17	1.73 e-05	1.03 e-05	0.55
with 5	image 2	-1605.6	1089.6	285.9	220	-2.24 e-08	6.17 e-13	-2.36 e-17	-9.76 e-07	-2.99 e-07	
	$\delta\lambda_i$	43.7	28.7	80.7	40.5	4.71 e-07	1.46 e-11	1.37 e-16	1.23 e-05	1.21 e-05	0.55
parameters	image 1&2	-1612.4	1095.5	281.8	239.3	-2.83 e-08	3.40 e-13	-1.47 e-17	-9.22 e-07	5.70 e-07	
	$\delta\lambda_i$	32.6	21.9	72.5	25.5	3.51 e-07	9.29 e-12	7.51 e-17	1.17 e-05	8.36 e-06	0.73
model	image 1	-1638.3	1112.7	280.2	247	-3.44 e-08	0	0	0	0	
	$\delta\lambda_i$	47.6	32.4	23	7.1	7.54 e-08	0	0	0	0	0.56
with 1	image 2	-1605.3	1089.2	286	220	-9.01 e-08	0	0	0	0	
	$\delta\lambda_i$	29	19.3	12.9	11.8	*4.16 e-08	0	0	0	0	0.56
parameter	image 1&2	-1611.9	1095.1	281.9	239.3	-7.61 e-08	0	0	0	0	
	$\delta\lambda_i$	34.5	23.5	18.5	9.5	*5.63 e-08	0	0	0	0	0.72

Table 2. Estimation of parameters and relative reconstruction error for real data

ate model for the transformation world-image.

If we use a model with many distortion parameters, we have a big confidence interval for every parameter in the camera and distortion model. When we want accuracy in the model estimation, we need to diminuer this interval. So, we made statistical test for the significance of parameters, to find only the important distortion parameters in the model.

If the distortion is much smaller than the noise in the image, the use of distortion model, causes an error in the camera parameters estimation, because the effect of noise is more important than the effect of distortion. In this case, it is better to use a camera model without distortion. If we want parameters estimation with accuracy smaller than 0.1 units, we need to have image location with accuracy until 0.1 pixel.

If the distortion is greater, it is worth to introduce distortion model, particulary if images measurements are accurate (with subpixel accuracy). We have seen that using a good distortion model (having only the important parameters), we have a better camera parameters estimation and a better image location after correction of distortion. Hence we can obtain a good accuracy in the reconstruction. For real data with large distortion, the gain in 3D reconstruction is better about a factor 2 or more, if we use the right distortion model.

Also, we have seen that the z-axis rotation angle e_1, is estimated well, because is not correlated with the others camera parameters. But the other camera parameters are correlated and a small error in camera parameters estimation has an large influence in the 3D reconstruction error.

References

[1] A. Antoniadis, J.Berruer, and R.Carmona. *Regression non lineaire et applications*, volume 1 of *economie et statistiques avancees*. Economika, 1992.

[2] H.A. Beyer. *Geometric and Radiometric Analysis of a CCD-Camera Based Photogrammetric Close-Range System*. PhD thesis, ETH-Zurich, 1992.

[3] P. Brand. *Reconstruction Tridimensionnelle d'une Scène à partir d'une Camèra en Mouvement: de l'Influence de la Précision*. PhD thesis, Université Claude Bernard, Lyon I, October 1995. ftp://ftp.imag.fr/pub/MOVI/theses/brand.ps.gz.

[4] D.C. Brown. Close-range camera calibration. *Photogrammetric Engineering*, 37(8):855–866, 1971.

[5] O.D. Faugeras and G. Toscani. Camera calibration for 3D computer vision. In *Proceedings of International Workshop on Machine Vision and Machine Intelligence, Tokyo, Japan*, 1987.

[6] R. Hartley and P. Sturm. Triangulation. In *Proceedings of* ARPA *Image Understanding Workshop, Monterey, California*, pages 957–966, November 1994.

[7] R.I. Hartley. Self-calibration from multiple views with a rotating camera. In *Proceedings of the 3rd European Conference on Computer Vision, Stockholm, Sweden*, pages 471–478. Springer-Verlag, 1994.

[8] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. *Numerical Recipes in C*. Cambridge University Press, 1988.

[9] C.C. Slama, editor. *Manual of Photogrammetry, fourth edition*. American Society of Photogrammetry and Remote Sensing, Falls Church, Virginia, USA, 1980.

[10] G.P. Stein. Accurate internal camera calibration using rotation, with analysis of sources of error. In E. Grimson, editor, *Proceedings of the 5th International Conference on Computer Vision, Cambridge, Massachusetts, USA*, pages 230–236. IEEE Computer Society Press, June 1995.