A&QT-R 2002 (THETA 13)
2002 IEEE-TTTC International Conference on Automation, Quality and Testing, Robotics
May 23 – 25, 2002, Cluj-Napoca, Romania

Camera Calibration Method for Stereo Measurements

Sergiu Nedevschi, Tiberiu Marita, Mircea Vaida, Radu Danescu,
Dan Frentiu, Florin Oniga, Ciprian Pocol

Technical University of Cluj-Napoca, 15 C. Daicoviciu Str.
E-mail:Sergiu.Nedevschi@cs.utcluj.ro, Marita.Tiberiu@cs.utcluj.ro, rdanescu@vision.utcluj.ro

Abstract: Camera calibration is an important and sensitive step in 3D environment reconstruction by stereovision. Small errors in the initial estimation of the camera parameters (especially in the estimation of the principal point) could rise to high errors in the 3D measurements, which are increasing with the working distance. There are many general-purpose calibration methods for calibration of individual cameras. We propose a method for refining the results of such methods by inferring stereo information from a controlled 3D scene. The parameters of two cameras composing a stereovision system are recalibrated together by minimizing the depth reconstruction error of the control points from the reference scene.

Key words: camera calibration, camera model, stereovision, depth estimation, 3D reconstruction.

1. INTRODUCTION

Camera calibration in the context of three-dimensional machine vision is the process of determining the internal geometric and optical characteristics (intrinsic parameters) of the camera and/or the 3D position and orientation of the camera coordinate system relative to a certain world coordinate system (extrinsic parameters).

The correct estimation of these parameters is crucial for applications based on artificial vision in order to perform a correct 3D reconstruction from 2D images of the environment for far working ranges. The performances of these applications are strongly affected by their precision especially when stereovision algorithms are used.

Several methods for general-purpose geometric camera calibration are presented in the literature. Some representative ones are: Direct Linear Transformation (DLT) [1], Tsai's algorithm [2], Zhang's method [3], Heikkila's method [4] or the CALTECH **(**California Institute of Technology**)** method [5]. Some of their assessment criteria are: autonomy, accuracy, efficiency and versatility [1].

A camera calibration method is good as long as it permits the stereo algorithm to establish good 3D parameters, given the correspondence of features in stereo image pairs. The general-purpose calibration methods are used to calibrate each camera individually and do not aim to infer stereo information. The only success criterion for the calibration is that the reprojected image of some control points (grid pattern of known size) should coincide with their actual position in the image.

It was demonstrated however that even if the projection errors were near zero, the 3D reconstruction from the stereo algorithm was not precise enough for measurements at far ranges. A solution oriented towards stereo must be developed, but this requires a lot of time and work. Thus we propose a combined solution which estimates the camera parameters in the classical way and then refines them using stereovision information:

a. The initial estimation is done using an adapted version of the original Caltech Calibration Toolbox. The main drawback of this method is that is highly dependent on the calibration image set (size and position of calibration pattern, lighting conditions, camera quality, number of images) and errors are encountered in the estimation of the principal point coordinates.

b. Then these initial parameters are refined/corrected by inferring stereo information. For this second step we propose methods named **the brute force calibration method** that performs 3D reconstruction on a set of feature pairs of known distance (depth), and refines the initial calibration parameters by minimizing the estimated distance error to those control points.

2. THE CAMERA MODEL

The pinhole camera model is used. [6]. The model is augmented considering the lens distortions [4]:

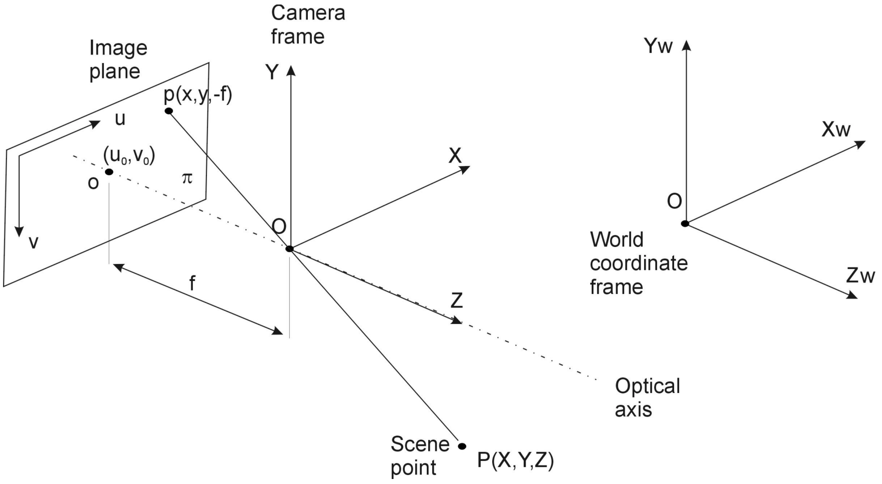


Fig. 1. The pinhole camera model

The following camera parameters are considered:

Intrinsic parameters – describing the internal geometric and optical characteristics of the camera (those that specify the camera itself):

- focal length, that is, the distance between the camera lens and the image plane: ***f***;
- the principal point - location of the image center in pixel coordinates: $(\mathbf{u_0},\mathbf{v_0})$;
- the radial and tangential distortion coefficients of the lens.

Extrinsic parameters - the 3-D position and orientation of the camera frame relative to a certain world coordinate frame (are needed to transform world coordinates to a camera centered coordinate frame or vice versa):

- the translation vector $\mathbf{T} = [\ Tx, Ty, Tz\ ]^T$;
- the rotation vector $\mathbf{r} = [\ Rx,.Ry, Rz\ ]^T$ or its equivalent rotation matrix **R**.

The projection of a 3D point P of coordinates [X,Y,Z] in the camera reference frame on the image plane is computed using the intrinsic parameters as follows:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = A \cdot \begin{bmatrix} x_D \\ y_D \\ 1 \end{bmatrix} \tag{1}$$

where:

$[u\ v\ 1]^T$ – is the pixel coordinate vector in the image frame;
$[x_D\ y_D\ 1]^T$ – is the distorted image point vector in camera coordinate frame;
$[x\ y]^T = [X/Z\ \ Y/Z]^T$ are the normalized coordinates of P in the camera frame;
A – is the internal camera matrix:

$$A = \begin{bmatrix} f_X & 0 & u_0 \\ 0 & f_Y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

f_X, f_Y – are the focal lengths expressed in units of horizontal and vertical pixels.

The transformation between the coordinates of the point P expressed in the world reference frame $\mathbf{XX_W} = [\,X_W\,,\ Y_W\,,\ Z_W\,]^T$ and its coordinates in the camera reference frame $\mathbf{XX_C} = [\,X\,,\ Y\,,\ Z\,]^T$ can be performed trough the following matrix equation:

$$XX_C = R_{WC} \cdot XX_W + T_{WC} \tag{3}$$

3. INITIAL ESTIMATION OF THE CALIBRATION PARAMETERS

The initial estimation is done using an adapted version of the original Caltech Calibration Toolbox. The CALTECH method [5, 7] uses the intrinsic camera model described in chapter 2. In the initialization phase the homographies for all control points on series of images are found and the intrinsic parameters are initialized by setting the distortion to 0. Then the extrinsic parameters are found for each image of the pattern. The final nonlinear optimization using Maximum Likelihood estimation is identical to Zhang,s method [3], augmented with the tangential distortion coefficients.

As calibration pattern, a plane plate with black and white squares of known sizes is used. Each camera is calibrated individually. For intrinsic parameters a set of at least 20 images must be taken at several orientations and distances, with the pattern covering the maximum area from the image, as possible (fig. 2). A good set of input images gives smaller uncertainties (less then 5 pixels) in the estimation of the parameters and determines the algorithm to converge in fewer steps.

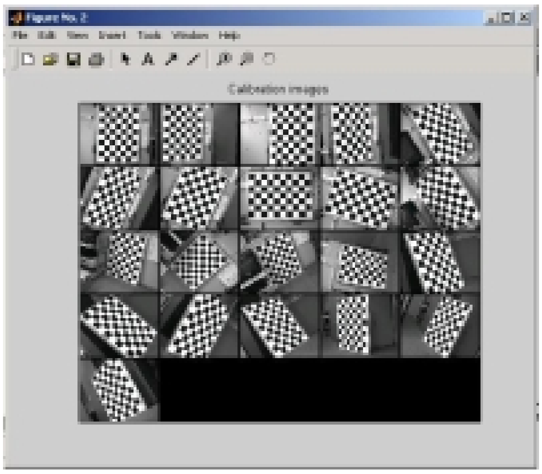


Fig. 2. Set of images for intrinsic calibration

Then the extrinsic parameters are estimated. For this task the cameras are mounted on a stereo rig, and the pattern is positioned to a known position relative to the cameras. For each camera a single image is shot and the extrinsic parameters of each camera, relative to a world reference frame related to the calibration pattern, are computed.

The main drawback of this method is that is highly dependent on the calibration image set (size and position of calibration pattern, lighting conditions, camera and lens quality, lens focusing and aperture, number of images). If the proper intrinsic calibration methodology is not kept, errors could be encountered especially in the estimation of the principal point coordinates. The principal point estimation errors can degrade drastically the quality of the extrinsic parameters and the stereo reconstruction process.

4. THE BRUT FORCE CALIBRATION METHOD

The method is based on the hypothesis that the main source of errors is the position of the principal point of the two cameras. Finding the right position of this point has proven a difficult task, and the oscillations of this position in different calibration sessions had lead us to believe that this position is the main source for errors.

The main steps of the method are:

i. Initial intrinsic parameters calibration using the method described in chapter 3.

ii. Initial extrinsic parameters calibration using the method described in chapter 3. The pair of shouted images and the distance from the cameras to the calibration pattern (the distance estimation can be considered accurate enough for this purpose) is kept for the next steps.

iii. Take stereo image pairs with objects at known positions. The distance to these objects should be accurately measured using a laser device. The origin of this distance must be the world reference system, that is the position of the pattern used for extrinsic calibration. The distance distribution must be as uniform as possible.

iv. Matching corresponding features of the objects (control points) in the image pairs and providing distances to the objects. A database is build, containing:

- Coordinates of the point in the left image;
- Coordinates of the point in the right image;
- Distance of the point from the world coordinate system.

v. Recalibrating the camera parameters: for any possible position of the principal point of the two cameras in a given search interval depending on the current search step (serch_interval = search_magnitude * search_step) around the original position (given by the classical method) the following operations are done:

- Perform extrinsic calibration using the shouted extrinsic images in step **ii.** ;
- For any point in the database, compute its distance using the stereo reconstruction algorithm [6], with the parameters obtained from the current extrinsic calibration;
- Sum the relative errors of the points. Keep the minimum total relative error, its principal point coordinates and the extrinsic parameters that lead to it.

vi. Diminish the search interval by halving the search step and continue the search around the new principal point providing minimum error by repeating step **v.** until the search step is small enough (about 0.05).

vii. The last found principal points coordinates providing minimum error are kept and the intrinsic and extrinsic parameters are updated.

5. EXPERIMENTAL REZULTS

A comparison of the calibration results using a classical method based on the Caltech Calibration Toolbox and using the brut force calibration method is presented. The estimated distance by stereo reconstruction to some control points were the assessment criteria.

The set of control points were the four visible corners of a cubical object, placed at several distances, uniformly distributed in a range from 14 to 107 m (14.160, 37.860, 61.760, 85.250, 107.460) measured in the world reference frame coordinates. These control point were also used in the brut force calibration method for building the database with the needed data for inferring stereo information in the recalibration process. The distances to the control points were measured using a laser measurement device. The origin of the world reference frame (related to the pattern position in the extrinsic calibration process) was at about 6.8m in front of the cameras.

Results of the classical calibration procedure are presented below. The principal point position was wrong estimated (due to an improper set of input images for the intrinsic calibration procedure):

Table 1. Intrinsic parameters from the classical calibration method:

	fx [pixels]	**fy [pixels]**	**u_0 [pixels]**	**v_0 [pixels]**
Left camera	2387.8	2387.3	353.2960	209.9128
Right camera	2.3916	2.3905	423.6890	218.8417

In figure 3 are presented the results of the depth estimation by stereo reconstruction for the previously considered 20 control points. Because the principal point was wrong estimated, also the extrinsic parameters were wrong estimated. This lead to high errors in the depth estimation, which are increasing linearly with the distance (fig. 3.b):

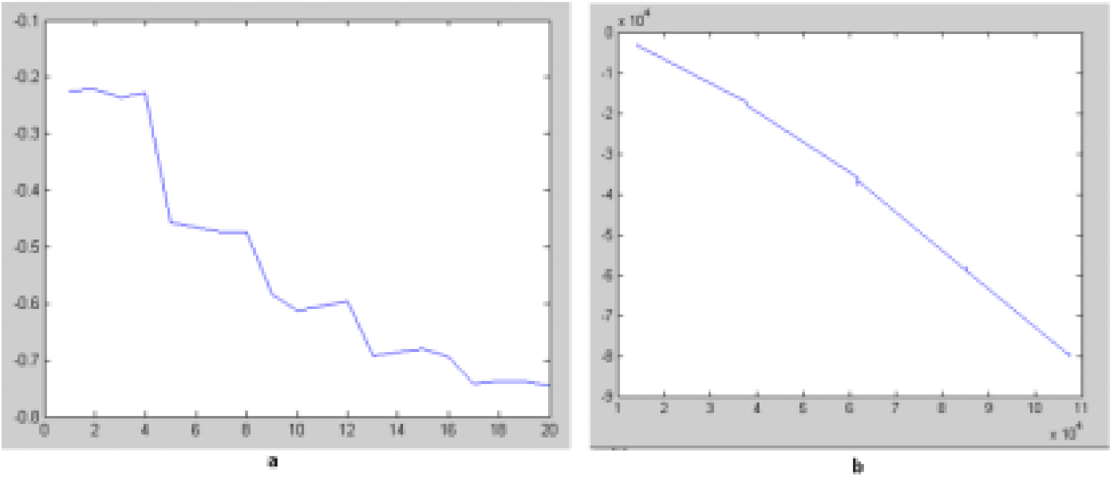


Fig. 3. Error of the depth reconstruction using parameters from the classical calibration method: a). relative error vs. control point; b). absolute error vs. depth [mm]

Results of the brut force calibration procedure are presented next. The principal point position was corrected by inferring stereo information. The correction was done only for the ***u*** (horizontal coordinate) due to the configuration of the stereovision system (with cameras displaced on ***x*** direction - horizontal baseline) and therefore in the reconstruction process only the horizontal disparity was used. For the correction of the principal point on vertical direction a vertical displacement of the two cameras should be used, the principle of the method remaining the same.

The corrected coordinates of the principal point are:

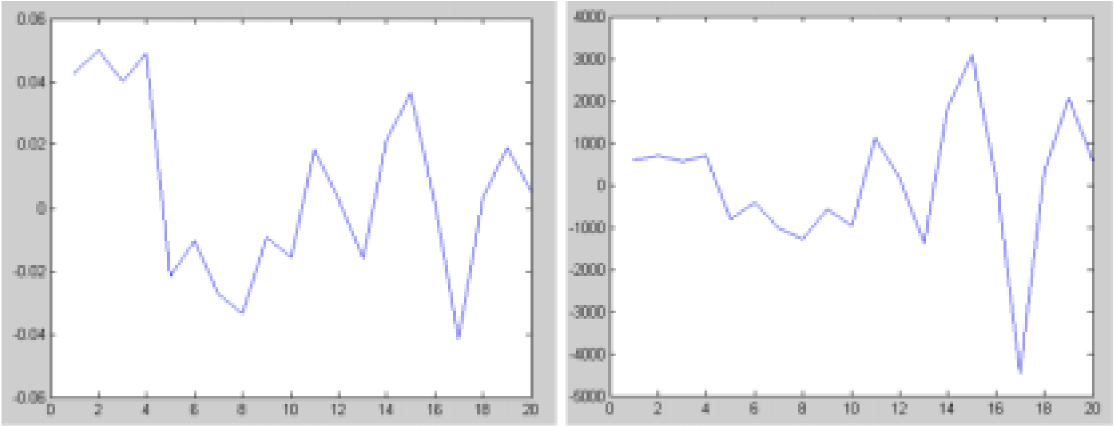
- Left camera: $(u_0, v_0) = (\mathbf{444.11628}, 209.91278)$
- Right camera: $(u_0, v_0) = (\mathbf{467.34139}, 218.8417)$

The corrected extrinsic parameters for 3D reconstruction are:

$$T_{CW\_left} = \begin{bmatrix} 209.748 \\ -493.988 \\ -6849.142 \end{bmatrix}; T_{CW\_right} = \begin{bmatrix} 584.437 \\ -493.516 \\ -6855.708 \end{bmatrix}; \tag{4}$$

$$R_{CW\_left} = \begin{bmatrix} 0.999876 & 0.001427 & 0.015696 \\ -0.002655 & 0.996908 & 0.078534 \\ -0.015535 & -0.078566 & 0.996788 \end{bmatrix}; R_{CW\_right} = \begin{bmatrix} 0.999562 & 0.007656 & 0.028590 \\ -0.009748 & 0.997228 & 0.073763 \\ -0.027946 & -0.074009 & 0.996866 \end{bmatrix} \tag{5}$$

In figure 4 are presented the results of the depth estimation by stereo reconstruction for the same 20 control points. The absolute error is under 2m (with two exceptions at point 15: 3 m and point 17: -4m), and does not depend on depth.



a

b

Fig. 4. Error of the depth reconstruction using parameters from the brut force calibration method: a). relative error vs. control point; b). absolute error vs. control point [mm].

6.CONCLUSIONS

The proposed calibration method can be used to refine the parameters obtained with general-purpose calibration algorithms. Their correct estimation is essential in the stereo reconstruction process, and thus stereo information from a controlled 3D scene was used to perform the calibration improvements. The experimental results proved that even a set of poor initial calibration parameters (in terms of principal point) can be corrected using this method in order to perform precise 3D reconstruction of the working environment for far depth ranges.

The limitation of this method is when significant errors are encountered in the initial estimation of the other camera parameters (focal length, distortion coefficients). In this case a solution that estimates all parameters of the cameras in a stereo fashion, similar to the method presented here, should be used.

REFERENCES

[1] Abdel-Aziz, Y. I. & Karara, H. M. (1971), “Direct linear transformation into object space coordinates in close-range photogrammetry”. *Proc. Symposium on Close-Range Photogrammetry*, Urbana, Illinois, pp. 1-18.

[2] Tsai, R. Y.(1987), “A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses”. *IEEE Journal of Robotics and Automation*, RA-3(4)/1987: pp.323-344.

[3] Zhang Z (1999), “Flexible Camera Calibration By Viewing a Plane From Unknown Orientations.” *International Conference on Computer Vision (ICCV'99)*, Corfu, Greece, September 1999, pp. 666-673

[4] Heikkila, J.; Silven, O. (1997), "A four-step camera calibration procedure with implicit image correction," *Proc. IEEE Computer Society Conf.*, 1997 pp.1106 –1112.

[5] Caltech Vision Group offical site: http://www.vision.caltech.edu/.

[6] Trucco E., Verri A. (1998), “Introductory techniques for 3D Computer Vision”, Prentice Hall, 1998.

[7] Intel Corporation (2001), “Open Source Computer Vision Library.” *Reference Manual*, http://www.intel.com/research/mrl/research/opencv/