THE ANNALS OF "DUNAREA DE JOS" UNIVERSITY OF GALATI
FASCICLE III, 2009, Vol.32, No.2, ISSN 1221-454X
ELECTROTECHNICS, ELECTRONICS, AUTOMATIC CONTROL, INFORMATICS

SINGLE CAMERA CALIBRATION IN 3D VISION

Caius Suliman, Dan Puiu, Florin Moldoveanu

Department of Automation
Transilvania University of Brasov
Eroilor. 29,69121 Brasov, Romania
{*caius.suliman, puiudan, moldof*}*@unitbv.ro*

Abstract: Camera calibration is a necessary step in 3D vision in order to extract metric information from 2D images. A camera is considered to be calibrated when the parameters of the camera are known (i.e. principal distance, lens distorsion, focal length etc.). In this paper we deal with a single camera calibration method and with the help of this method we try to find the intrinsic and extrinsic camera parameters. The method was implemented with succes in the programming and simulation environment Matlab.

Keywords: camera calibration, 3D vision, calibration pattern, intrinsic parameters, extrinsic parameters.

1. INTRODUCTION[1]

Camera calibration is a necessary step in 3D vision in order to extract metric information from 2D images. A camera is considered to be calibrated when the parameters of the camera are known (i.e. principal distance, lens distorsion, focal length etc.). For this purpose, in the last twenty years, many calibration algorithms have been developed in the computer vision community. This algorithms are generally based on the perspective camera models. Among the most popular is Robert Tsai's calibration algorithm (Horn, 2000), (Tsai, 1987). His algorithm is based on the pinhole model of perspective projection. The model proposed by Tsai assumes that some of the camera's parameters are given by the are manufacturer, in order to reduce the initial guess of estimation. The algoritm requires n feature points ($n > 8$) per image and solves the calibration problem with a set of n linear equations based on the radial alignment constraint. A second order radial distorsion model is used while no decentering distorsion terms are considered. This method can be used with either a single image or multiple images of a planar or 3D calibration grid.

Another important and very popular calibration method has been developed by Zhenyou Zhang (Zhang, 1998). His method requires a planar checkerboard grid to be placed in front of the camera at different orientations. The algorithm uses the extracted corner points of the checkerboard to calculate a projective transformation between the image points of the different images. The camera's intrinsic and extrinsic parameters are recovered using a closed-form solution. The radial distorsion terms are recovered within a linear least-squares solution. The final step is the use of a non-linear minimization of the reprojection error that refines all the recovered parameters. Zhang's method is similar to the one proposed by Triggs (Triggs, 1998). Zhang's algorithm is the basis behind some popular open source implementations of camera calibration (i.e. Intel's OpenCV and Matlab's calibration toolkit).

In this paper we present the implementation of a camera calibration method based on the calibration methods presented by Trucco (Trucco and Verri, 1998).

[1]This paper is supported by the Sectoral Operational Programme Human Resources Development (SOP HRD), financed from the European Social Fund and by the Romanian Government under the contract number POSDRU/6/1.5/S/6.

This paper was recommended for publication by Rustem Popa

2. CAMERA CALIBRATION

The key idea behind calibration is to write the projection equations linking the known coordinates of a set of 3D points and their projections, and solve for the camera parameters. In order to get to know the coordinates of some 3D points, camera calibration methods rely on one or more images of a calibration pattern (a 3D object of known geometry, possibly located in a known position in space and generating image features which can be located accurately). In the aborded method we use the perspective camera model, also known as the pinhole camera model.

The method presented here consists in two stages:

- Estimation of the projection matrix by linking world and image coordinates;
- Computation of the camera's parameters as closed-form functions of the entries of the projection matrix.

2.1. Estimation of the projection matrix

The relation between the 3D coordinates (X_i^w, Y_i^w, Z_i^w) of a point in space and the 2D coordinates (x, y) of its projection on the image plane can be written by means of a 3x4 projection matrix, M, as follows:

$$(1)\quad \begin{pmatrix} u_i \\ v_i \\ w_i \end{pmatrix} = M \begin{pmatrix} X_i^w \\ Y_i^w \\ Z_i^w \end{pmatrix},$$

with:

$$(2a)\quad x = \frac{u_i}{w_i} = \frac{m_{11}X_i^w + m_{12}Y_i^w + m_{13}Z_i^w + m_{14}}{m_{31}X_i^w + m_{32}Y_i^w + m_{33}Z_i^w + m_{34}}$$

$$(2b)\quad y = \frac{v_i}{w_i} = \frac{m_{21}X_i^w + m_{22}Y_i^w + m_{23}Z_i^w + m_{24}}{m_{31}X_i^w + m_{32}Y_i^w + m_{33}Z_i^w + m_{34}}$$

The matrix M is a scaling factor and its entries can be determined through a homogenous linear system formed by writing (2) for at least 6 world image points. With the help of a calibration pattern like the one presented in fig.1, many more correspondences and equations can be obtained and M can be estimated through least squares techniques.

The projection matrix M can be estimated by solving the following homogenous linear system:

$$(3)\quad Am = 0,$$

where A is presented in (3) and m is:

$$m = [m_{11}, m_{12}, ..., m_{33}, m_{34}]^T$$

The non-trivial solution of the homogenous equation $Am = 0$ is found by recovering the vector m from singular value decomposition (SVD) techniques of matrix A as the last column of V. In agreement with the above definition of M this means that the entries of M are obtained up to an unknown scale factor.

2.2. Camera parameters from the projection matrix

In practical solutions it is not always sufficient to have estimated the projection matrix M, the camera intrinsic (f_x, f_y, o_x, o_y) and extrinsic (R, T) parameters are also needed. We assume that the projection matrix has been estimated with the procedure from the previous section. The estimated projection matrix is denoted as $\hat{M}$:

$$\hat{M} = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix}.$$

First we rewrite the full expression for the entries of the projection matrix M:

$$(4)\quad M = \begin{pmatrix} -f_x r_{11} + o_x r_{31} & -f_x r_{12} + o_x r_{32} & -f_x r_{13} + o_x r_{33} & -f_x T_x + o_x T_z \\ -f_y r_{21} + o_y r_{31} & -f_y r_{22} + o_y r_{31} & -f_y r_{23} + o_y r_{31} & -f_y T_y + o_y T_z \\ r_{31} & r_{32} & r_{33} & T_z \end{pmatrix}$$

In this case we are trying to find the camera parameters f_x, f_y (where f represents the focal length in horizontal pixels size units), o_x, o_y (where o_x and o_y are the image center coordinates), R (the rotation matrix) and the translation vector T (T_x, T_y, T_z beeing its elements). For this purpose we also need the following 3D vectors:

$$q_1 = [\hat{m}_{11}, \hat{m}_{12}, \hat{m}_{13}]^T$$

$$q_2 = [\hat{m}_{21}, \hat{m}_{22}, \hat{m}_{23}]^T$$

$$q_3 = [\hat{m}_{31}, \hat{m}_{32}, \hat{m}_{33}]^T$$

$$q_4 = [\hat{m}_{14}, \hat{m}_{24}, \hat{m}_{34}]^T$$

Now, the estimated projection matrix $\hat{M}$ can be written as $\hat{M} = \gamma M$. Here $|\gamma|$ is $\sqrt{q_3^T q_3}$. q_3 is the last row of the R matrix. The next step is the division of the $\hat{M}$ matrix by $|\gamma|$. From the last row of (4) we have:

$$T_z = \sigma\hat{m}_{34} \quad \text{and} \quad r_{3i} = \sigma\hat{m}_{3i}, i = 1,2,3.$$

with $\sigma = \pm 1$. By taking the dot product of q_3 with q_1 and q_2 it results that:

$$o_x = q_1^T q_3 \text{ and } o_y = q_2^T q_3.$$

Then we can compute f_x and f_y as:

$$f_x = \sqrt{q_1^T q_1 - o_x^2} \text{ and } f_y = \sqrt{q_2^T q_2 - o_y^2}.$$

Until now we have computed the intrinsic parameters of the camera. Now we can compute the extrinsic parameters:

$$r_{1i} = \sigma(o_x\hat{m}_{3i} - \hat{m}_{1i})/ f_x, i = 1,2,3,$$

$$r_{2i} = \sigma(o_y\hat{m}_{3i} - \hat{m}_{2i})/ f_y, i = 1,2,3,$$

$$T_x = \sigma(o_x T_z - \hat{m}_{14})/ f_x,$$

$$T_y = \sigma(o_y T_z - \hat{m}_{24})/ f_y.$$

The estimated rotation matrix $\hat{R}$ obtained by this procedure is not really orthogonal. Therefore we must compute the rotation matrix that is the closest to the estimated matrix $\hat{R}$. By using SVD we have:

$$\hat{R} = UDV^T \text{ and } R = UV^T.$$

Now we have all intrinsic and extrinsic parameters. The only thing that bothers us now is the sign of σ. It can be obtained very easily from $T_z = \sigma\hat{m}_{34}$. If the origin of the worl frame is in front of the camera then the sign of σ is "+", else the sign of σ is "-".

3. EXPERIMENTAL EVALUATION

For the experimental evaluation of the above method we have used a wireless CMOS camera. Because the images taken by the camera were very noisy, we needed a filter to remove that noise. The filter used for this purpose was a simple one, the median filter. The calibration object used for the process consists of two perpendicular planes, in our case we have used two sides of a box. On the two sides we added a calibration pattern consisting of square tiles. The sides of the tiles are 1 cm. Another important step in the calibration process is to chose a convenient world coordinate frame (see Fig.1).

The first step of the calibration process is to set the 3D matrix *XYZ*, that containts the 3D coordinates of the calibration points. We have chosen those points as follows: as we know the sides of the tiles are 1 cm, so the *XYZ* coordinates for a point near the origin of the world system are (-1, 0, 1). Then we need to set up the 2D matrix *xy*, matrix that contains the 2D coordinates of the calibration points. For this purpose we have developed an interactive program that colects the 2D coordinates of these points (see Fig.2).

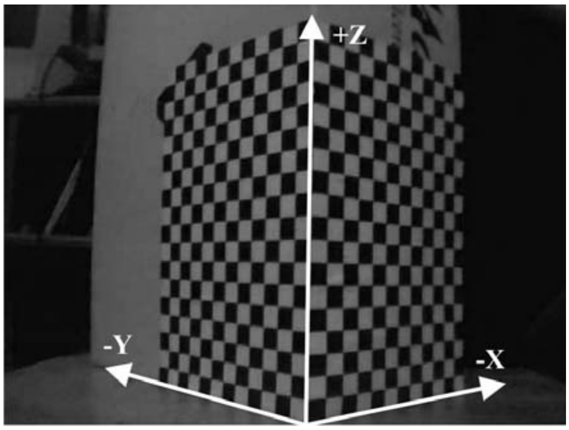

Fig.1. The calibration object and the chosen world coordinate frame.

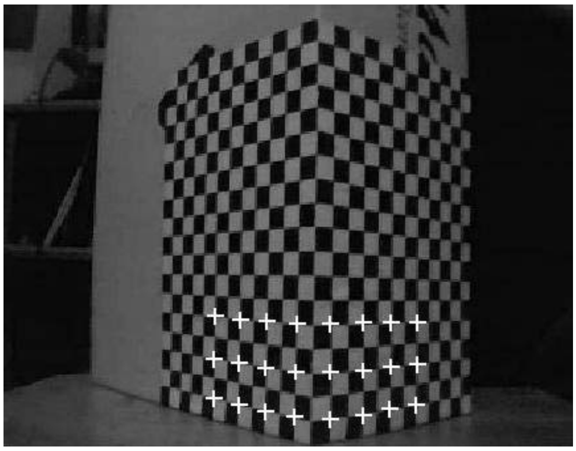

Fig.2. The calibration object and resulting calibration points.

For an accurate calibration process we need between 20 and 30 points. After this step is completed, the estimation of the projection matrix *M* can be done by applying the method presented early in this work.

To show that the calibration process was accurate we have considered some cubes with the sides equal to the sides of the tiles and we have put them over the calibration pattern. It can be easily observed the correction of the projection matrix estimation (see Fig.3).

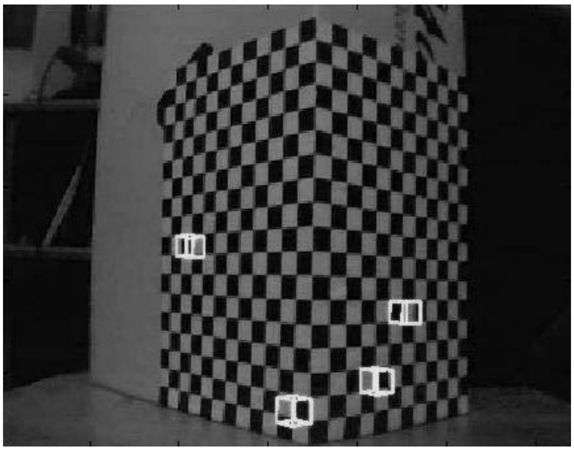

Fig.3. The calibrated camera allows for 3D objects to be drawn in the scene.

In the final step we have computed the estimates for the camera's intrinsic and extrinsic parameters from the projection matrix. The resulting intrinsic parameters are:

$$f_x = 321.0655\,px;\ f_y = 329.5092\,px,$$

$$o_x = 156.8256;\ o_y = 164.7667.$$

And the camera's extrinsic parameters obtained from the calibration process are:

$$R = \begin{pmatrix} 0.0192 & -0.0070 & 0.9998 \\ 0.7375 & -0.6751 & -0.0189 \\ -0.6750 & -0.7377 & 0.0078 \end{pmatrix},$$

$$T = \begin{pmatrix} -6.2695 \\ -0.6508 \\ 23.9378 \end{pmatrix}.$$

The entire calibration process has been implemented in the programming and simulation environment Matlab.

4. CONCLUSIONS

The method presented here is a simple calibration method that gets the job done. The precision of the calibration depends on how accurately the image and world reference points are located. The errors on the parameter estimates propagate to the result of the application. The calibration process ultimately depends on the accuracy requirements of the target application. For example, in industry accuracies of submillimeter are required. In other application are accepted even errors of centimeters.

5. REFERENCES

Horn, B.K.P. (2000). Tsai's camera calibration method revisited. *Technical report*, *MIT* Artificial Intelligence Laboratory website.

Triggs, B. (1998). Autocalibration from planar scenes. In: *Proc. of Fifth European Conference on Computer Vision*, pp.89–105.

Trucco, E., Verri, A. (1998). *Introductory Techniques for 3D computer vision*. Prentice-Hall, Inc., Upper Saddle River, New Jersey, ISBN 0-13-261108-2.

Tsai, R.Y. (1987). A Versatile Camera Calibration for 3D Machine Vision. In: *IEEE Journal of Robotics and Automation*, **Vol. RA-3**, **No. 4**, p.323-344.

Zhang, Z. (1998). A Flexible New Technique for Camera Calibration. *Technical Report MSR-TR-98-71*, Microsoft Research.