

OPTIMAL CONDITIONS FOR CAMERA CALIBRATION USING A PLANAR TEMPLATE

Carlos Ricolfe-Viala, Antonio-Jose Sanchez-Salmeron

Instituto de Automática e Informática Industrial
 Universitat Politècnica de València, Spain
 {cricolfe,asanchez}@isa.upv.es; <http://robotica.ai2.upv.es/>

ABSTRACT

Different techniques exist for camera calibration based on photogrammetry or self-calibration. Much work has been done to obtain intrinsic and extrinsic camera parameters and also in data pre- and post-processing techniques. From a practical viewpoint, it is quite difficult to decide which calibration method produces the best results and even whether any data processing is necessary.

This paper proposes a set of optimal performance conditions to resolve the calibration process accurately. The calibration method based on a 2D pattern has been chosen since camera is calibrated using several images of a planar template easily. The proposed optimal conditions define the number of points to design the calibration pattern, the number of images to resolve the calibration accurately, as well as positions and orientations from where images should be taken. The recommend rules will help to obtain a well-conditioned calibration process.

Index Terms— Camera calibration, 2D pattern, non-linear calibration, optimal conditions

1. INTRODUCTION

In the calibration process, it is essential to consider the quality of results, which obviously depends on data accuracy. Some studies have been made on calibration errors committed due to imprecise measurement of the image plane or calibration template [1, 2, 3, 4]. The calibration process is also affected by erroneous association of one point in reality with a point in the image. Some authors use statistical tools to detect these anomalies [5]. Computing errors due to instabilities of the mathematical tool should also be taken into account [6]. In these cases, data normalization improves the robustness of the algorithm and gives more accurate results [7]. Another significant case is the computed camera model. A very complex model can complicate the algorithm and the result obtained will not be much better than that obtained with a simpler model. A very complicated model may produce instabilities in the process of questing and produce absurdities [8].

The state of the art of calibration provides some guide about efficiency of camera calibration in all situations. The Tsai method [9] represents a classical calibration process based on the measurements of the 3D points in the template taking a fixed reference. This method has been widely used in the past. Salvi [10] compares the calibration methods developed between 1982 and 1998, with the Tsai method showing better performance, despite the fact that it requires high precision in input data. On the other hand, Zhang's method [3], which is not included in Salvi's comparison [10], represents a new era in the camera calibration process. This method uses images of a 2D template taken from different camera positions and orientations. In this way, the advantages of camera self-calibration are combined with the point coordinate-based calibration. This calibration method is highly flexible, since the camera and the template can be moved freely and also as many images as are required can be taken without measuring any position of the template. Sun [11] compares the Tsai method with Zhang's method. On one hand, Tsai produces a precise estimation of camera parameters if the input data have not been corrupted. Since 100 points in the template are necessary and the coordinates should be referred to a fixed origin, careful design of the calibration template and a very accurate coordinate measurement are necessary. Nevertheless, errors are too easily committed and in practice these results are not as accurate as expected, as shown by Sun [11]. On the other hand, Zhang's method based on a 2D template requires neither a special design nor precise point measurement. Sun obtains camera calibration with a hand-made template, and better results are computed using Zhang's method. Furthermore, the sensibility of the calibration algorithm to errors in the measures can be improved by increasing the spotted number in the template, by simply printing a chessboard with more corners. The results of the comparison show the flexibility and adaptability of Zhang's method, as it can be performed on any scene. Considering the results of these two authors, Zhang's method is used as a reference for camera calibration.

This paper proposes the optimum conditions from the viewpoint of number of points, number of images and location of the camera for taking the images, to improve on Zhang's camera calibration method. These definitions obtain

a well-conditioned calibration process which compute better results. The paper is structured as follows: Section 2 defines optimal conditions for camera calibration from the viewpoint of number of points, number of images and location of the camera taking the images. In section 3 experimental results are presented. Paper ends with some conclusions.

2. OPTIMAL CONDITIONS FOR CAMERA CALIBRATION

To calibrate the linear camera model, Zhang [3] describes a method based on the homographies between a planar calibration pattern and its images from several camera positions. For each homography two homogeneous equations arise as:

$$\begin{bmatrix} v_{11}^T & -v_{22}^T \\ v_{12}^T \end{bmatrix} \mathbf{b} = 0 \quad (1)$$

where:

$$v_{ij}^T = \begin{bmatrix} h_{i1}h_{j1} & h_{i1}h_{j2} + h_{i2}h_{j1} & h_{i2}h_{j2} \\ h_{i3}h_{j1} + h_{i1}h_{j3} & h_{i3}h_{j2} + h_{i2}h_{j3} & h_{i3}h_{j3} \end{bmatrix} \quad (2)$$

h represents elements of the homography \mathbf{H} and $\mathbf{b} = [b_{11}, b_{12}, b_{22}, b_{13}, b_{23}, b_{33}]$, where b_{ij} represents the element ij of the symmetric matrix $\mathbf{K}^T \mathbf{K}^{-1}$. \mathbf{K} contains intrinsic parameters of pin-hole model. If m images of the calibration object are observed, m equations such as (1) arise, giving $\mathbf{V}\mathbf{b} = \mathbf{0}$, where \mathbf{V} is a $2m \times 6$ matrix. At least three images are necessary $m \geq 3$ in order to obtain a unique solution. The closed-form solution is given by the eigenvector of $\mathbf{V}^T \mathbf{V}$ associated with the smallest eigenvalue. Once \mathbf{b} is estimated, camera internal parameters can be computed. When \mathbf{K} is known, external parameters for each image can be computed when the corresponding homography is known. See [3] for details.

2.1. Camera calibration scene

To define the optimal positions for image capture, a calibration scene is defined. $(\mathbf{o}, \{x_w, y_w, z_w\})$ is 3D scene coordinate system and $(\mathbf{o}', \{x_c, y_c, z_c\})$ is camera coordinate system located as shown in figure 1. The centre of the template is situated in the origin of coordinates of the scene and the camera is always translated with a negative z-coordinate. The orientation of the camera is defined using the position. This means that the camera optical axis crosses the centre of the template always as shown on figure 1. Moreover, X_c axis of the camera coordinate system is always parallel to plane X_w-Y_w of the scene. From a practical point of view, the template is located on the floor of the calibration scene and the camera is located on a tripod, which keeps the upper border of the image parallel to the floor. The origin of the scene coordinate system is the template centre.

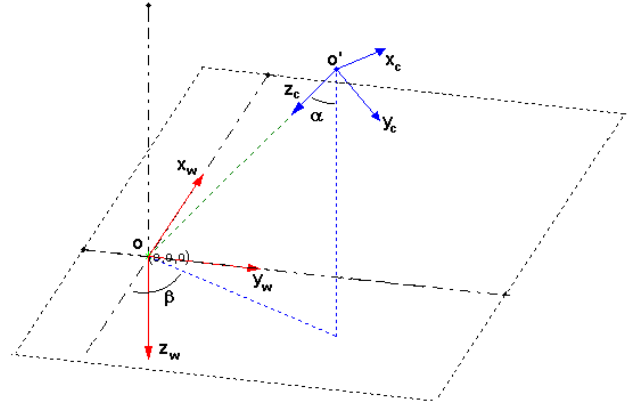


Figure 1. Scene and camera coordinate systems. Orientation is computed starting with camera location. Optical axis goes through origin of coordinates in the scene.

To relate orientation with position of the camera, two rotations are defined. First camera and scene coordinate systems coincide. Then, the camera is rotated at angle β with respect to the Z_c axis as shown in figure 1. Second, in relation to the X_c axis, the camera is rotated an angle α as shown on figure 1. These two rotations keep the X_c axis parallel to plane $X-Y$ of the calibration scene. These angles are a function of the camera position:

$$\begin{aligned} \cos \beta &= \frac{t_y}{m_1} & \sin \beta &= \frac{t_x}{m_1} \\ \cos \alpha &= \frac{|t_z|}{m_2} & \sin \alpha &= \frac{m_1}{m_2} \\ m_1 &= \sqrt{t_x^2 + t_y^2} & m_2 &= \sqrt{t_x^2 + t_y^2 + t_z^2} \end{aligned} \quad (3)$$

where m_1 represents the module of camera translation in plane X_w-Y_w , symbolizing the floor, and m_2 is the distance of the camera to the origin of coordinates in the scene. The origin of coordinates of the image is in the centre and will therefore be pixels with negative coordinates. Another supposition refers to the scale factor k_u and k_v in image axes. It is assumed that pixels are squared and therefore the scale factor is the same in both axes and we rename with k .

2.2. Condition of camera calibration

As known in a general framework, a matrix \mathbf{M} has r eigen vectors and r eigen values, where r is the rank of matrix \mathbf{M} . Eigen vectors are orthogonal to one another. Eigen vector modules represent the number of vectors in the matrix \mathbf{M} oriented in the direction defined by each eigen vector. The module of one eigen vector increases if there are more vectors in matrix \mathbf{M} oriented in this eigen vector direction. The ratio between the bigger module and the smaller module is called the condition number of the matrix. To obtain a well-conditioned system, the condition number should be

one. This means there are vectors in \mathbf{M} which cover all R -dimensional space. Therefore, if a parameter vector is estimated with this matrix \mathbf{M} , information from the data affects all dimensions of the parameter vector. In the event of a badly conditioned matrix, some parameters are not influenced by the information contained in matrix \mathbf{M} and therefore such estimation will be erroneous.

If this theory is particularized to linear camera calibration, the camera parameters are in the eigen vector associated with the smaller eigen value of matrix \mathbf{V} defined by expression (1). Matrix \mathbf{V} is formed from the elements of homographies ${}^l\mathbf{H}$. Since the condition number of matrix \mathbf{V} should be one, all vectors of matrix \mathbf{V} should have the same module, and they should be orthogonal to one another. Vectors of \mathbf{V} depend on the homographies ${}^l\mathbf{H}$, whereas ${}^l\mathbf{H}$ depends on the locations from which images are taken. Therefore, vectors of \mathbf{V} depend indirectly on positions in which the camera is located. Consequently, we can say that positions from which images are taken, influence the condition of matrix \mathbf{V} . As a result, to obtain a well-conditioned matrix \mathbf{V} several locations for the camera should be defined.

Dimension of \mathbf{V} is $2m \times 6$. Therefore the minimum number m of homographies to resolve the system is $m \geq 3$. Within this framework, a well-conditioned matrix \mathbf{V} will arise with vectors ${}^l\mathbf{v}_{ij}$, which are orthogonal to one another, with equal modules. Sub index ij represents vector \mathbf{v} formed from the columns i -th and j -th of homography l . Thus, homographies forming a set of vectors ${}^l\mathbf{v}_{ij}$ orthogonal to one another and with identical modules are required. In consequence care must be taken when dealing with homography elements.

To obtain a well-conditioned matrix \mathbf{V} , the following expressions must be true for all i, j, l .

$$\left({}^l\mathbf{v}_{11}^T - {}^l\mathbf{v}_{22}^T \right) {}^l\mathbf{v}_{12} = 0 \quad (4)$$

$$\left({}^l\mathbf{v}_{11}^T - {}^l\mathbf{v}_{22}^T \right) \left({}^l\mathbf{v}_{11} - {}^l\mathbf{v}_{22} \right) - {}^l\mathbf{v}_{12}^T \cdot {}^l\mathbf{v}_{12} = 0 \quad (5)$$

The first expression represents an orthogonal condition and the second stands for module equality. i, j , have values 1 and 2 since they correspond with the two left-hand columns of the homography l . The number of homographies l should be at least 3. If 3 homographies are established whose vectors are orthogonal and with equal modules, a well-conditioned matrix \mathbf{V} will be obtained with only a minimum amount of information.

2.3. Locations for image capturing

Analysing expressions (4) and (5) they are satisfied when a camera is located with null coordinates ${}^l t_x$ or ${}^l t_y$. If ${}^l t_x$ is zero, ${}^l t_y$ and ${}^l t_z$ will not be zero. Also, if ${}^l t_y$ is zero, ${}^l t_x$ and ${}^l t_z$ will not be zero. ${}^l t_x$ or ${}^l t_y$ are set different to zero moving the camera along the X or Y scene axis. The camera position in the Z axis of the scene ${}^l t_z$ is computed with the following

expressions depending on whether the camera has been located with coordinate ${}^l t_x$ or ${}^l t_y$ different to zero:

$$\begin{aligned} {}^l t_z &= \pm \sqrt{-\frac{2 \cdot {}^l t_x^2 (-1 - 2 \cdot {}^l t_x^2 - 2 \cdot {}^l t_x^4 + \sqrt{(-3 - 4 \cdot {}^l t_x^2 + 12 \cdot {}^l t_x^4)})}{1 + 3 \cdot {}^l t_x^2 + {}^l t_x^4}} \\ {}^l t_z &= \pm \sqrt{-\frac{2 \cdot {}^l t_y^2 (-1 - 2 \cdot {}^l t_y^2 - 2 \cdot {}^l t_y^4 + \sqrt{(-3 - 4 \cdot {}^l t_y^2 + 12 \cdot {}^l t_y^4)})}{1 + 3 \cdot {}^l t_y^2 + {}^l t_y^4}} \end{aligned} \quad (6)$$

From a practical point of view, the camera is moved along one of the axes representing the X or Y axis of the coordinate system of the scene and then the altitude of the camera in the Z axis is defined using the expression (6). After this, the camera must be oriented bearing in mind that the optical axis passes through the origin of coordinates in the scene. The camera must be moved along the X or Y axis exactly the same distance. It is important to note that expressions (6) do not depend on intrinsic parameters. This means that such camera locations do not depend on camera features and can be used with any camera.

3. EXPERIMENTAL RESULTS

To test the performance of the camera calibration conditions a simulated camera is similar to Zhang in [3] with parameters $k_u=1250$, $k_v=900$, $u_0=250$, $v_0=250$ and $\theta=1.09083$ equivalent to 89.95° is used. Image resolution is 512×512 pixels. The calibration template is a chessboard of $10 \times 14 = 140$ corners. Template size is 180×250 mm. In this case, it is situated on the floor and all corners have coordinates $z_w=0$. The camera is oriented so as to view the origin of coordinates in the centre of the image. Also, camera orientation is computed to ensure the x -coordinate axis of the image is parallel to the template plane. This simulates the effect of taking an image of a template situated on the floor using a tripod.

To calibrate the camera under optimal conditions, the simulated camera is located at $\mathbf{t}_1=(200, 0, 401.8)$, $\mathbf{t}_2=(-100, 0, 200.9)$ and $\mathbf{t}_3=(0, 300, 602.1)$. Also, camera is calibrated using images from random positions $\mathbf{t}_1=(150, 200, 580)$, $\mathbf{t}_2=(-50, 250, 880)$ and $\mathbf{t}_3=(100, -20, 820)$. Both calibration results changing the number of points in the template are compared. Results are shown in figure 2 for intrinsic camera parameters k_u , u_0 only. Similar performance has been computed for the remaining camera parameters. Camera calibration is solved using the linear and the non-linear process. In all cases, computed camera parameters are improved if optimal conditions are used.

The number of points necessary for the calibration process can be defined analysing experimental results. If images are taken taking into account optimal conditions, more than 70 does not reduce parameters errors significantly. Obviously, since constructing a calibration template is an easy task, more than 70 points can be used to improve the results.

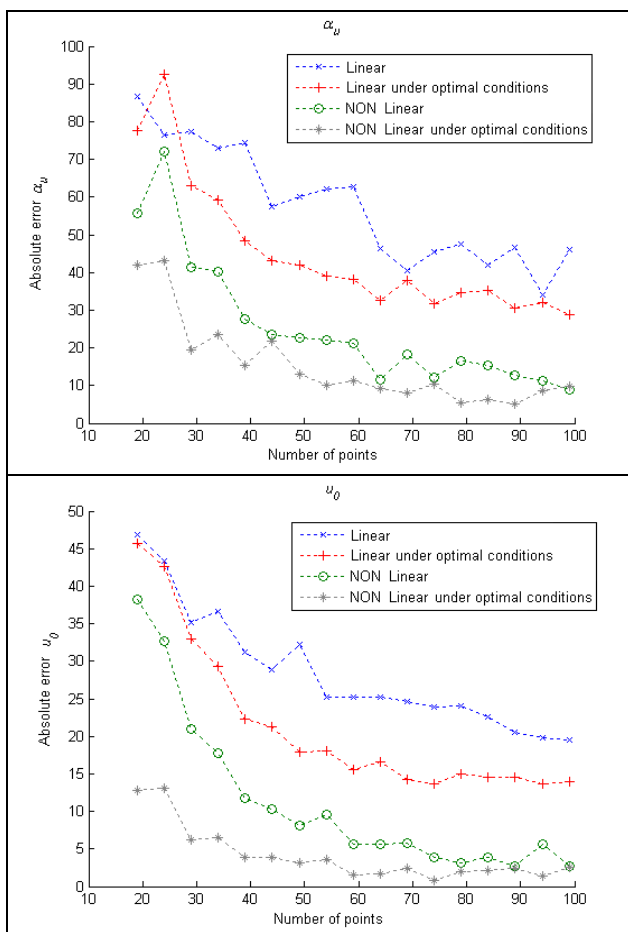


Figure 2. Comparison of intrinsic camera parameters calibration errors under optimal and non- optimal calibration process conditions changing the number of points

4. CONCLUSIONS

Optimal conditions for camera calibration using a 2D pattern have been defined. Camera parameters are computed the eigen vector associated with the smaller eigen value of the matrix composed with elements of several homographies. The calibration process will therefore be more stable if the condition number of this matrix is close to 1. To obtain this condition number closer to 1, elements of the homographies should be taken into account. Therefore, images of the template should be captured from specific locations to obtain a well-conditioned calibration process. To define the optimal locations from where to take images of the template, the condition number of this matrix has been analysed. The camera is located taking into account that the altitude should be twice its separation from the origin of coordinates in the scene. Also, the camera should be located along the X or Y axes of the scene coordinate system. The camera orientation is defined assuming that the optical axis goes through the calibration template. Finally, coordinates in the image and

the template should be referred to the centre of the image and the template, respectively.

Although Zhang's method was intended for self-calibration in which images are taken from anywhere, here we propose a useful guide to improve the calibration results.

Acknowledgements

This work was partially funded by the Universidad Politécnic de Valencia (PAID 2010-2431 and PAID 10017), Generalitat Valenciana (GV/2011/057) and by Spanish government and the European Community under the project DPI2010-20814-C02-02 (FEDER-CICYT) and DPI2010-20286 (CICYT)

5. REFERENCES

- [1] S. Koppurapu, P. Corke, "The effect of noise on camera calibration parameters" *Graphical Models* 63, pp. 277-303, 2001.
- [2] J. Lavest, M. Viala, M. Dhome, "Do we really need accurate calibration pattern to achieve a reliable camera calibration". *European Conference on Computer Vision* 1, pp. 158-174, 1998.
- [3] Z. Zhang, "A flexible new technique for camera calibration", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11), pp. 1330-1334, 2000.
- [4] C. Ricolfé-Viala, A.J. Sánchez-Salmerón, "Correcting non-linear lens distortion in cameras without using a model" *Optics & Laser Technology*, Vol. 42, pp. 628-639, 2010.
- [5] D. Huynh, R. Hartley, A. Heyden, "Outlier correction in image sequences for the affine camera", *9th IEEE International Conference on Computer Vision*, vol. 1, pp.585, 2003
- [6] G. Stewart, "Perturbation theory for the singular value", *Tech. Report TR90 -124*, University of Maryland 1990.
- [7] R. Hartley, "In defence of the eight point algorithm", *IEEE Transactions on pattern analysis and machine intelligence*. 19, pp 580 – 593, 1997.
- [8] J. Weng, P. Cohen, M. Herniou, "Camera calibration with distortion models and accuracy evaluation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14, pp. 965-980, 1992
- [9] R. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-self TV camera lenses", *IEEE Journal of Robotics and Automation*. RA-3, pp. 323-344, 1997.
- [10] J. Salvi, X. Armangué, J. Batlle, "A Comparative review of camera calibrating methods with accuracy evaluation", *Pattern recognition* 35, pp. 1617-1635, 2002.
- [11] W. Sun, J. Cooperstock, "An empirical evaluation of factors influencing camera calibration accuracy using three publicly available techniques", *Machine Vision and Applications* 17, pp. 51-67, 2006.