View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by DSpace at University of West Bohemia

Camera Calibration Methods Evaluation Procedure for Images Rectification and 3D Reconstruction

Rachid Guerchouche and François Coldefy
Orange Labs – France Telecom R&D
TECH/IRIS/VIA
2, Av. Pierre Marzin
France 22307, Lannion Cedex
{rachid.guerchouche, françois.coldefy}@orange-ftgroup.com

ABSTRACT

Camera calibration and images rectification are two necessary steps in most 3D reconstruction methods using image acquisition. This paper proposes an evaluation procedure for camera calibration methods for the case of 3D reconstruction using rectified multi-stereo images. The evaluation is based on the accuracy of the rectification and of the 3D reconstruction which are directly related to the calibration precision. Three methods are thus compared: Faugeras-Toscani, Zhang and a robust calibration algorithm. The procedure can be applied for computer vision systems with an arbitrary number of cameras and for any other calibration method. We show that, although the three methods provide significantly different intrinsic and stereo system parameter estimations, the rectified images of the planar target that we use for evaluation are relatively coherent and lead to close 3D reconstruction errors.

Keywords

Camera calibration, images rectification, 3D reconstruction.

1. INTRODUCTION

3D reconstruction of real objects is one of the most widely known computer vision applications. Many techniques are described in the literature. In this paper, we are interested in 3D reconstruction with several images, using a technique known as multi-stereo reconstruction. Two steps are necessary to achieve the reconstruction: calibration of the cameras [Fau01] and rectification of the image pairs [Sha01]. Camera calibration aims at estimating the parameters of the relationship binding the 3D world reference space and the 2D camera coordinates system. It consists of estimating the intrinsic and extrinsic parameters representing respectively the internal camera characteristics and the camera pose in the world reference.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Copyright UNION Agency – Science Press, Plzen, Czech Republic.

Rectifying a stereo image pair consists of finding two transformations projecting both images onto the same plane so that each pair of conjugate epipolar lines becomes collinear and parallel to one of the image axes, generally the horizontal one. The main advantage of rectification in stereo applications is that the matching is performed in 1D along the same line instead of in 2D.

This paper proposes an evaluation procedure for camera calibration methods based on images rectification and 3D reconstruction in the case of multi-stereo. Among the camera calibration algorithms proposed in the past years, we decide to focus on three methods: Faugeras-Toscani [Fau86, Fau87], Zhang [Zha99, Zha00] and a robust calibration method [Gue06].

2. CAMERA CALIBRATION

In this section, a brief description of the adopted camera model is given. We also give a short description of the three compared and evaluated camera calibration methods.

2.1 Camera Model

We focus on the pinhole camera model, which is widely used in computer vision. It assumes that the camera performs a perfect perspective transformation

 ISBN 978-80-86943-16-9

P from the 3D scene coordinates $[x \quad y \quad z]$ to image plane coordinates $(u \quad v)$:

$$[\eta u \quad \eta v \quad \eta]^T = \mathrm{P} \quad [x \quad y \quad z \quad 1]^T \qquad (1)$$

where η is an homogeneous factor and $[.]^T$ is the transpose operator. The PPM P, a 3×4 matrix, is defined as the product $\mathrm{P} = Q \quad [R \mid t]$ where Q and $[R \mid t]$ are respectively the intrinsic and extrinsic matrices. R is an 3×3 orthogonal matrix representing the camera orientation and t the position vector of the camera in the 3D space. More details on this model can be found in [Fau01].

2.2. The Calibration Methods

Faugeras and Toscani [Fau86, Fau87] proposed a calibration method based on an estimation of the Perspective Projection Matrix (PPM) using an image of a non-planar pattern. It uses both linear and nonlinear approaches. One image and at least six non-coplanar feature 3D points, manually selected or automatically detected on the acquired image, are needed.

Zhang [Zha99, Zha00] described an algorithm which requires at least two different views of a planar pattern. An even more accurate calibration is obtained using a large number of views (twenty or more). The displacements of the pattern between the views are not necessarily known.

The algorithm detailed in [Gue06] for camera calibration is based on a robust estimation of the PPM. The target used is a 3D cube with different colored faces. A manual selection of two adjacent faces on the acquired image allows the system to automatically detect the six vertices associated to these faces and thus initialize the PPM using the Faugeras-Toscani algorithm. A refinement of the PPM estimation is then achieved by minimizing the distance between the projected cube edges and the image contours. The estimation of the camera parameters can be improved by acquiring additional images taken at different positions in the camera field of view, provided the same two adjacent faces are seen. This estimation involves a nonlinear optimization technique (Levenberg-Marquardt).

2.3. Calibration Methods Comparison: Related Works

Objective evaluation of camera calibration algorithms is affected by the lack of criteria to compare the final estimation of camera parameters obtained by the different methods. Few authors compare existing camera calibration methods and examine this problem in particular.

Zollner and Sablatnig [Zol04] investigate the performances of the three most widely used plane-based calibration algorithms.

González *et al.* [Gon05] presente a comparative analysis of eight camera calibration methods in which they focus specifically on the stability of the camera parameters: *(i)* the stability of the intrinsic parameters when the camera setup is constant and the calibration pattern is displaced *(ii)* the stability of the extrinsic parameters when the pattern is still and the configuration of the camera varies. The conclusion of their study is that the result of the camera parameters estimation depends on the location of the calibration pattern in the acquired images. In the case of constant camera setup, intrinsic parameters values should theoretically not change. However, in practice, these values vary from one calibration process to another, which constitutes a problem. In the case of a fixed camera and pattern, with variation of the camera configuration (focus and/or zoom) the extrinsic parameter values are not constant as they should be.

Salvi *et al*. [Sal02] present a detailed review of five of the most known calibration techniques. The authors regroup some criteria to evaluate and to compare calibration methods. These criteria are essentially based on the measurement accuracy of 3D and 2D points. A set of 3D points in the reference scene, with known coordinates, are reconstructed using a stereo system. The dispersions between the estimated 3D positions (respectively estimated 2D positions on image plane) and the real known positions (respectively projections of the real positions of the 3D points detected on the image plane) are calculated. The proposed criteria constitute a good evaluation and comparison of the calibration methods because the accuracy of the calibration is directly related to the 3D reconstruction one.

In this paper we propose a new procedure to evaluate and compare camera calibration methods. Our procedure uses a planar pattern with known dimensions, however, it's position in the reference scene is not necessarily known. The evaluation and comparison is performed with new and accurate criteria based on images rectification and 3D reconstruction of a set of coplanar points on the planar target.

3. EXPERIMENTS

3.1. Calibration Procedure

In order to compare the three methods, Zhang, robust calibration and Faugeras-Toscani, we install a vision system composed of three horizontally fixed

 ISBN 978-80-86943-16-9

mvBlueFOX® cameras[1], with image resolution 1024×768. The three cameras are placed on an approximate arc circle in order to converge on a focus zone. We do not assume that the focal axes of the three cameras are coplanar (see Figure 1).

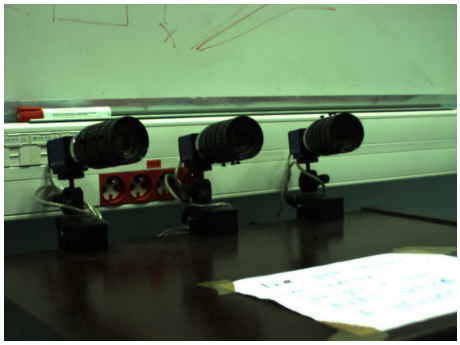


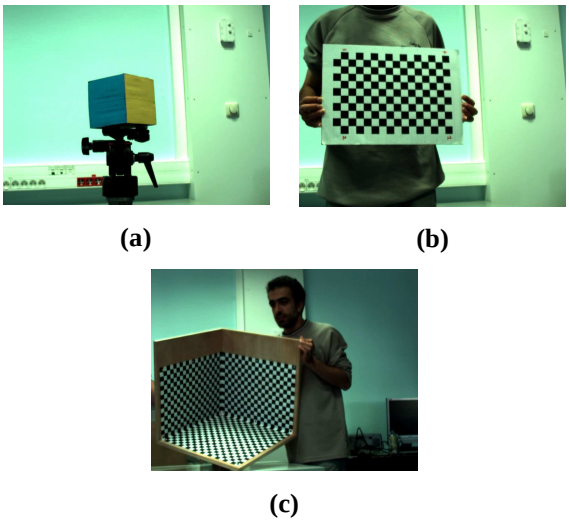
Figure 1. The vision system composed of three fixed cameras.

For the Zhang calibration algorithm, two calibrations are performed, using respectively 10 and 30 images per camera. *Camera Calibration Toolbox for Matlab* is used[2]. In the case of the robust calibration, 9 images per camera are acquired. A first calibration is performed using a single image per camera with the coloured cube roughly placed in the centre of the field of view of the three cameras. A second calibration is performed with 8 additional images of the cube positioned in order to cover the whole field of view of the camera. A minimization is then performed with the 9 images [Gue06] from each of the three cameras. Faugeras-Toscani calibration method is performed using a single image per camera.

In the case of multi-image calibration (Zhang and the robust methods), the pattern is relocated for each view and a PPM is obtained for each position. Figure 2 shows the patterns used for each method. Both the Zhang and the Faugeras-Toscani calibration reference points were obtained using the corner detector proposed in the *Matlab Toolbox*, which automatically detects the corners of the black and white squares on the patterns. We used the non-linear method of Faugeras-Toscani proposed by Gonzàlez[3] to compute the intrinsic and extrinsic parameters.

3.2. Rectification Procedure

We use a planar pattern containing 21×21 black and white squares to test the obtained calibrations.



(a) **(b)**

(c)

Figure 2. Sample images of the calibration patterns. (a) a cube for the robust calibration, (b) a planar target for Zhang calibration, and (c) a 3D target for Faugeras-Toscani calibration.

Figure 3 shows the three acquired images of our test pattern for the rectification comparison.

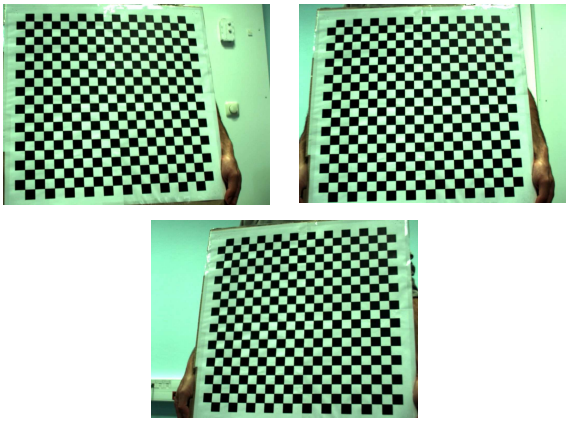


Figure 3. The rectification test images.

Note that these images were not involved in the calibration process. The test pattern rectification is computed for each camera pair $(i, j)_{i,j=1,2,3}^{i<j}$ using the Fusiello *et al.* [Fus00] compact algorithm.

The rectified images using this method are acquired virtually by a new computed stereo rig, in which the original cameras are rotated.

The rectification using this algorithm requires cameras calibration of the original stereo rig in order to obtain the two image transformations. We think that this is a good candidate to asses the quality of camera calibration algorithms: if the PPMs given by the tested algorithms are accurate enough, the same

[1] http://www.matrix-vision.com

[2] http://www.vision.caltech.edu/bouguetj/calib_ doc/

[3] http://mozart.dis.ulpgc.es/Gias/josep/source_ code.htm

point on the left and right images will be projected onto the same line in the rectified images; if not, it will be projected onto different lines.

The same automatic detector proposed in the *Matlab Toolbox* is used to detect the intersection between the black and the white squares. The inner 19×19 squares produce, for each image acquired by a camera $i, (i=1,2,3)$ a set of *n=400* organized points $p_k^i = \begin{pmatrix} u_k^i & v_k^i \end{pmatrix}$. After rectification, all points p_k^i and p_k^j on the images acquired by a camera pair (i, j) should be on epipolar lines parallel to the second axis, *i.e.* $v_k^i = v_k^j$.

To evaluate the rectification errors, we compute the Rectification Mean Square Error (*RMSE*) $e^{i,j}$ for each image pair (i, j) as:

$$e^{i,j} = \sqrt{\frac{1}{n}\sum_{k=0}^{k=n-1}\left(v_k^i - v_k^j\right)^2} \qquad (2)$$

In the case of multi-images calibration, the RMSE is computed over all the couples (p_k^i, p_k^j) over all the images.

To compare the accuracy of each calibration method, we compute the global mean error *M* associated with the three cameras system as:

$$M = \frac{1}{3}\sum_{i,j=1,2,3}^{i<j} e^{i,j} \qquad (3)$$

3.3. Rectification Results and Discussion

The *RMSE* errors are shown in Table 1. **Z**, **F**, and **R** respectively stand for Zhang, Faugeras-Toscani, and Robust methods; *nb* is the number of image pairs for a given pair of cameras used for parameters estimation.

All the experiments show large error values except for the calibrations using one image per camera (the robust method and Faugeras-Toscani one), and for one of the stereo pair, $(1,3)$ with Zhang method.

This is a quite surprising result since the camera parameters estimation using several images is known to improve the calibration accuracy and thus minimize the *RMSE*. In fact, both the robust method and Zhang's method (except for the $(1,3)$ camera pair) failed to properly estimate the PPM for some image pairs. The global estimation of the intrinsic parameters is then biased.

	nb	$e^{1,2}$	$e^{1,3}$	$e^{2,3}$	*M*
Z	10	27.30	3.22	19.61	16.71
	30	5.84	2.52	6.67	5.01
F	1	2.83	3.32	5.08	3.74
R	1	1.69	2.64	3.75	2.69
	9	5.30	11.21	7.33	7.94

Table 1. Rectification Mean Square Error of the three calibration methods: Z, F, and R respectively stand for Zhang, Faugeras-Toscani, and Robust methods; *nb* is the number of image pairs for a given pair of cameras used for parameters estimation.

Rectification errors are usually due to an inadequate position of the calibration pattern out of the camera focus, which leads to blurred images.

We apply a robust algorithm to detect and remove those image pairs. For each stereo pair (i, j), and all the *m* calibration image pairs, we estimate the *Baseline* $(B_{i,j})_{k=1,\ldots,m}$. Theoretically, the baseline is constant. We look for outliers of the baseline estimates, *e.g.* values which are more than 1.5σ away from the mean (σ is the estimated standard deviation of the baseline). We then suppress the corresponding calibration image pairs, and estimate again the extrinsic and intrinsic parameters. We reiterate until no image pair is removed.

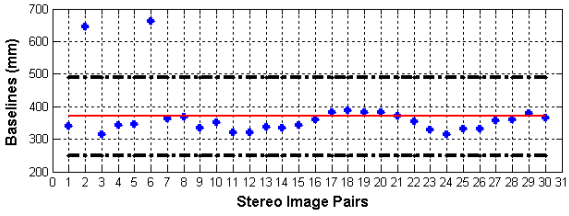


Figure 4: Plot of the baseline values $B_{1,2}$ for the 30 image pairs estimated with Zhang method; points: baseline values; continuous line: the baseline mean; discontinuous lines: the *mean* $\pm 1.5\sigma$ deviations; the two outlier values correspond to the pairs 2 and 6.

Figure 4 shows the plot of the baseline values (*points*), of the baseline *mean* (*continuous line*), and of the *mean* $\pm 1.5\sigma$ deviations (*discontinuous lines*) found at the first iteration for the Zhang method using 30 images for camera pair $(1,2)$.

Two outlier values are detected for the stereo image pairs 2 and 6. After aberrant image pairs removal, Zhang calibration is performed with only 28 images

per camera, Zhang calibration with 10 images is performed with 10 images, replacing the aberrant pairs by other unaberrant pairs, and the robust calibration with 9 images is performed with 7 images (also two aberrant pairs are removed). The obtained results after outliers' removal are shown in Table 2.

	nb	$e^{1,2}$	$e^{1,3}$	$e^{2,3}$	*M*
Z	10	3.33	3.38	1.95	2.88
	28	3.27	2.93	1.63	2.61
R	7	1.59	2.41	3.44	2.48

Table 2: RMSE values after removal of the aberrant image pairs.

The *RMSE* values (about 3 pixels for 768 pixels image height) are now 2 to 8 times smaller. The Zhang method using 28 images (M = 2.61), the robust method using 7 images (M = 2.48) and the robust method using a single image (M = 2.69) show similar accuracy whereas the Faugeras-Toscani method provides greater errors (M = 3.74). In case of multi-image calibration (the Zhang and the robust method), we have as many PPM estimations as image pairs. For each calibration method and for each camera pair, we select the PPM corresponding to the smallest *RMSE* to perform the 3D reconstruction of our test pattern grid (Figure 3). Table 3 shows the observed minimum *RMSE* for each camera among all the calibration image pairs. Zhang and the robust methods show similar accuracy.

	$e^{1,2}$	$e^{1,3}$	$e^{2,3}$	*M*
Z	1.53	1.96	1.03	1.50
R	1.06	1.88	1.83	1.59

Table 3. Minimum *RMSE* among all the calibration image pairs for each camera pair.

3.4. 3D Reconstruction

Table 4 summaries the estimations of the camera focal lengths f and of the baselines B of the 3 camera pairs for each calibration method. Note that, although the *RMSE* are low and the rectifications are coherent, the 3 methods exhibit significant differences in intrinsic (focal length) and extrinsic (baseline) parameter estimations.

The question arises: what is the influence of these differences on the 3D reconstruction?

For a camera pair (i, j), 3D reconstruction of the 400 points of our grid pattern, obtained from the rectification applying Fusiello *et al.* algorithm [Fus00], is expressed in a global reference system where the origin is the middle of the baseline, the x axis is parallel to the baseline and the focal axis is given by the mean of the focal axis of each camera (see [Fus00] for more details).

Method	**F**	**Z**	**R**
f_1 (mm)	7.83	8.13	8.40
f_2 (mm)	7.89	8.03	8.32
f_3 (mm)	7.62	8.03	8.12
$B_{1,2}$ (mm)	206.33	238.11	226.03
$B_{1,3}$ (mm)	406.64	475.32	445.32
$B_{2,3}$ (mm)	196.89	250.26	232.88

Table 4: Focal and baseline values (Z: Zhang, F: Faugeras-Toscani, R: Robust).

Up to the variation in estimation of the baseline orientation and focal axis of each camera, the 3D reconstruction is thus expressed in a same global reference system, independently of the chosen camera calibration method.

Methods		**(R,Z)**	**(R,F)**	**(Z,F)**
(1,2)	x	2.89	25.56	28.90
	y	1.33	71.20	70.02
	z	7.89	88.33	79.33
(1,3)	x	1.05	29.26	25.15
	y	0.30	78.12	74.50
	z	1.88	88.33	76.56
(2,3)	x	0.58	29.26	30.12
	y	0.31	78.12	68.15
	z	3.86	81.74	90.22

Table 5. The mean absolute difference of 3D reconstruction obtained from the three different parameters estimation, for each camera pair.

Table 5 shows, for each camera pair, the mean of the absolute difference in position, expressed in millimeters between the 3D reconstruction of the grid pattern points obtained from Zhang (calibration with 28 images pairs), Faugeras-Toscani and the robust method (calibration with 7 image pairs). Faugeras-Toscani reconstruction is far from the two others. The robust method and Zhang's reconstructions essentially differ on the z axis. The smallest difference occurs for the $(1,3)$ pair which corresponds to the most convergent camera pairs.

An accurate reconstruction should also preserve the planarity of the grid pattern points. To compare the 3 reconstructions, for each camera pair, we compute by *Principal Component Analysis* the best plane fitting the reconstructed points. The *mean* of the *distances* of reconstructed points to the *plane* (*mdp*), expressed

in millimetres and shown in the first line of Table 6, is a measure of the planarity of the reconstruction. No noticeable difference can be found between the three obtained reconstructions.

A last criterion to evaluate the accuracy of the calibration estimation is the estimated area of the reconstructed squares, each with known value equal to 900 mm^2.

Methods	**F**	**Z**	**R**
mdp	2.69	2.39	2.49
mds	28.94	21.18	21.26

Table 6. Values of *mdp* and *mds* for the three methods and for the three pairs.

The second line of Table 6 shows for each calibration method the *mean difference* between real *square* areas and corresponding estimations over the entire grid (*mds*). Let r_i be the real area of the i^{th} square and e_i the corresponding estimated one, after 3D reconstruction:

$$mds = \frac{1}{19 \times 19} \sum_{i=1}^{i=19\times19} (r_i - e_i) \qquad (4)$$

The *mds* values show that Zhang and the robust methods are more accurate.

4. CONCLUSION

In this paper, we proposed an experimental procedure to evaluate calibration methods for image rectification and multi-stereo 3D reconstruction based on the accuracy of the rectification and of the 3D reconstruction of a known object. This procedure can be applied for all stereo systems with unlimited number of cameras.

Three calibration algorithms have been evaluated: Zhang, Faugeras-Toscani, and a robust method. We demonstrated that, in the case of multi-image calibration, the existence of aberrant images considerably affects the accuracy of calibration. We then developed an efficient technique to detect and remove those image pairs in order to exclude them from the calibration process.

Experimental results show that although the three methods provide significantly different estimations of the camera's intrinsic and the stereo system parameters (±6% for the focal lengths, and ±8% for the baseline), the rectification and 3D reconstruction errors remain close.

5. REFERENCES

[Fau01] Faugeras, O. D., Luong, Q. T., and Papadopoulo T. The Geometry of Multiple Images: The Laws That Govern the Formation of Multiple Images of a Scene and Some of Their Applications. The MIT Press, 2001.

[Fau86] Faugeras, O. D., and Toscani, G. The calibration problem for stereo. Proc. of the International Conference on Computer Vision and Pattern Recognition (CVPR), pp. 15–20, June 1986.

[Fau87] Faugeras, O. D., and Toscani, G. Camera calibration for 3d computer vision. International Workshop on Machine Vision and Machine Intelligence, pp. 240–247, 1987.

[Fus00] Fusiello, A., Trucco, E., and Verri, A. A compact algorithm for rectification of stereo pairs. Machine Vision Applications, Vol. 12, No. 01, pp. 16–22, 2000.

[Gue06] Guerchouche, R., Coldefy, F., and Zaharia, T. Accurate camera calibration algorithm using a robust estimation of the perspective projection matrix. Proc. of SPIE Mathematics of Data/Image Pattern Recognition, Compression, and Encryption with Applications IX, Vol. 6315, 2006.

[Gon05] Isern-González, I., Cabrera-Gámez, J., Guerra-Artal, C., and Naranjo-Cabrera, Á. Stability study of camera calibration methods. VI Workshop en Agentes Físicos, WAF'2005, I Congreso Español Informática, 2005.

[Sal02] Salvi, J., Armangué, X., and Batlle, J. A comparative review of camera calibrating methods with accuracy evaluation. Pattern Recognition, Vol. 35, No. 7, pp. 1617–1635, 2002.

[Sha01] Shapiro, L. G., and Stockman, G. C. Computer Vision, Prentice Hall, 2001.

[Zha99] Zhang, Z. Flexible camera calibration by viewing a plane from unknown orientations. Proc. of International Conference on Computer Vision (ICCV), Vol. 1, pp. 666–673, 1999.

[Zha00] Zhang, Z. A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 11, pp. 1330–1334, 2000.

[Zol04] Zollner, H., and Sablatnig, R. Comparison of methods for geometric camera calibration using planar calibration targets. Digital Imaging in Media and Education, Vol. 179, pp. 237–244, 2004.