

# Calibration Method Evaluation via Depth Reconstruction

Lucile Martin<sup>1,2</sup>, Christophe Leroux<sup>1</sup>, Edwige Pissaloux<sup>2</sup>

<sup>(2)</sup>CEA/LIST, BP 6, 92 256 Fontenay-aux-Roses, France  
martinl@zoe.cea.fr, christophe.leroux@cea.fr

<sup>(2)</sup>Université Paris 6, Laboratoire de Robotique de Paris, CNRS/FRE 2507, BP 61, 92 256 Fontenay-aux-Roses, France {martin, pissaloux}@robot.jussieu.fr

## Abstract

This paper presents a fast method to evaluate the accuracy of a calibration method of a stereovision rig built with commercially available cameras. The method allows by retro projection to quantify very precisely errors in terms of depth reconstruction. The proposed method has been evaluated on the depth reconstruction of real scenes from images taken with an unmanned aerial vehicle.

**Keywords:** calibration, depth reconstruction, distortion, camera model.

## 1.1. Introduction

3D scene reconstruction is one of the most important basic operations for intelligent vision systems and especially for autonomous robotic systems such as unmanned aerial vehicle (UAV).

Moreover, the precision of 3D scene reconstruction depends on the expected behavior of an autonomous robot in its nearest environment. Many 3D reconstruction processes use active sensors (telemeter, sonar etc); but passive sensors, such as optical cameras are often more suitable in terms of price, accuracy, non invasivity of the environment etc...

However, the performance of vision depth estimators strongly depends on their calibration accuracy.

The choice of a calibration method depends on:

- the requested camera parameters
- the conditions of experimentation
- the aimed accuracy of the depth reconstruction.

Camera calibration is the process of identifying the camera's optical parameters and its position and orientation with respect to a pattern frame.

Many calibration methods have been initiated by the photogrammetry community [2], [8], using a calibration object which geometry is known with a very good precision [9]. Classic approach solves usually the problem by minimizing a nonlinear error function. The processing time and computational burden of these classic techniques

have pushed the vision community to search other solutions such as the determination of the whole set of (11) camera parameters at the same time or auto-calibration [3], [4], [6], [11], [9].

However, these calibration methods, due to the correlation between intrinsic and extrinsic parameters, usually lead to errors on each determined camera parameter that results in important errors in 3D scene reconstruction.

This paper proposes an approach which allows to select a calibration method by quantifying its errors in terms of depth reconstruction accuracy.

The proposed method has been defined and validated on real images in the frame of the CNRS RobVollnt project (of the ROBEA program), a prototype of a UAV under design by the French Atomic Energy Commission (CEA), in collaboration with the IRISA, the I3S, Nice University, and the LRP, University of Paris 6.

Subsequent sections outline the proposed approach for a calibration method evaluation (section 2), the context of application of the method (section 3), the obtained results (section 4) and a short conclusion (section 5).

## 2. Novel approach for a calibration method evaluation

This section addresses the principle of the proposed method for calibration algorithm evaluation and its implementation consideration.

### 2.1. Principle of the proposed method.

The calibration method can be evaluated via the error quantification between 3D scene real points and their 3D estimates obtained by triangulation from real points' projections in images [5].

If the coordinates of a real point P are (X, Y, Z) and its reconstructed coordinates are (X', Y', Z'), therefore the calibration method error is defined by a classic definition (1):

$$e_z = (Z - Z') \quad (1)$$

The sign of the error allows positioning the reconstructed points with respect to the real points

and thus indicates how the calibration parameters should be adapted.

## 2.2. Proposed method implementation

3D scene reconstruction is one of the most important basic operations for intelligent vision systems and especially for autonomous robotic systems such as unmanned aerial vehicle (UAV).

## 2.3. Appropriate calibration method selection

Several parameters have to be taken into account in order to select a calibration method; among them there are:

- the number of cameras
- the cameras' pertinent parameters (extrinsic, intrinsic, their subset(s))
- the vision systems calibration frequency (a frequent calibration should be easy to set up and use)
- The calibration accuracy necessary to fulfil the goal of the vision system.

An autonomous light weighted UAV evolving in a classical indoor environment (such as buildings' corridors) should be equipped with a stereo rig. Intrinsic and extrinsic parameters of the stereo rig, should be obtained with a simple method which will occasionally be applied. Therefore, flexibility, robustness, calibration time and low computing cost are important criteria for calibration method selection.

Zhang's calibration method [12] satisfies all these constraints; it only requires to know the different orientations of at least one couple of images of a one planar pattern acquired with the same camera.

The pattern can be printed on a laser printer and attached to a reasonably planar surface.

A pattern reconstruction (used also for system's calibration) has been chosen as an example of images of real scenes (Figure 4).

## 2.4. Stereo rig calibration

The stereo rig calibration includes the evaluation of its intrinsic and extrinsic parameters.

### Intrinsic parameters estimation.

The stereo rig calibration with the chosen method, gives the rig intrinsic parameters (2) :

$$\mathbf{I}_c = \begin{pmatrix} \alpha_u & 0 & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (2)$$

where:  $\alpha_u = -k_u f$ ,  $\alpha_v = k_v f$  being the scale factors (in pixels) and  $v_0$ ,  $u_0$  being the principal point coordinates (in pixels) in the image frame.

The matrix  $\mathbf{I}_c$  express the transformation between the image frame and the camera frame (Figure 1). The same transformation, in homologous coordinates takes the form of expression (3):

$$\begin{pmatrix} s u \\ s v \\ s \end{pmatrix} = \mathbf{I}_c \begin{pmatrix} x \\ y \\ z \\ s \end{pmatrix} \quad (3)$$

where  $(u \ v)$  are the 2D image point coordinates (in pixels) of the 3D point  $(x \ y \ z)$ , both in the camera frame.

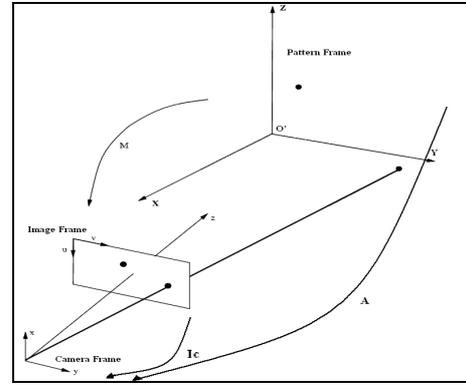


Fig. 1: Frames transformations obtained by the camera calibration

### Extrinsic parameters estimation.

The extrinsic parameters of the stereo rig can be estimated from the orientation of a couple of images.

As the orientation of a pair of pattern's images is known, the translation and rotation expressing the camera frame position with respect to the pattern frame have to be determined with the equation (3)

$$\mathbf{A} = \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 1 \end{pmatrix}, \quad (3)$$

where  $\mathbf{t} (t_x \ t_y \ t_z)$  is the translation vector between these two frames.

Therefore, if  $(X \ Y \ Z)$  are the coordinates of a pattern point in 3D real world, its corresponding coordinates in the camera frame are given by equation (4).

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} \quad (4)$$

## 2.5. 3D points reconstruction error estimation.

The appropriateness of a calibration method to provide the final application's acceptable 3D scene reconstruction is evaluated from error on depth estimation (cf. equation (1)).

The proposed method estimates the coordinates of the (real) corners of the pattern in a 3D "virtual world". These latter are obtained by projection of the recomputed pattern corners (estimated corners) in 3D space (virtual space) (figure 2).

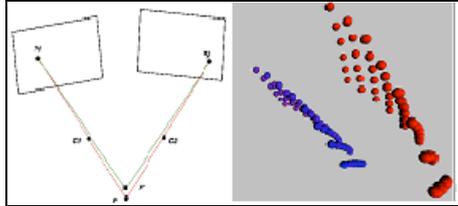


Fig. 2: Accuracy determination process. The dot point represents the real point (in red on the right side) and the square point represents the retro projected points in blue on the right image)

Equation (5) links the 3D coordinates of a point with its 2D projection in the camera reference frame :

$$\begin{pmatrix} su \\ sv \\ s \end{pmatrix} = \mathbf{M} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (5)$$

where  $\mathbf{M} = \mathbf{I}_c \mathbf{A}$  .

The depth can be computed (via the triangulation, for example) only if all image points coordinates (from right and left images) are expressed in the same frame.

Let  $\mathbf{A}_s$  be left-to-right image transfer matrix;  $\mathbf{A}_s$  can be computed from the equation (6).

$$\mathbf{A}_s = \mathbf{A}' \mathbf{A}^{-1} \quad (6)$$

Consequently, left image points' coordinates in the right camera frame are expressed by the equation (7) :

$$\begin{pmatrix} X' \\ Y' \\ Z' \\ 1 \end{pmatrix} = \mathbf{A}_s \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (7)$$

with  $\mathbf{M}'$  ,  $\mathbf{A}'$  , the right camera's extrinsic and intrinsic parameters matrices and  $\mathbf{M}$  ,  $\mathbf{A}$  , the left camera's ones.

Figure 3 illustrates the transformations process which links two images

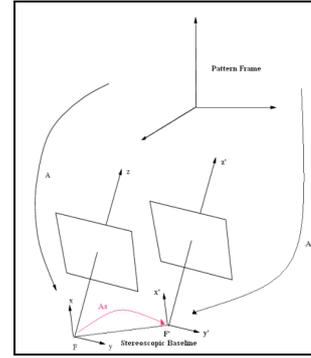


Fig. 3 : Stereo rig to pattern frame transformation

## 2.6. Accuracy estimation.

The accuracy of the calibration method is estimated from the equation (1) applied to the depth only, but it can be computed from the errors (on each axes of the pattern frame) between the real points' coordinates and the estimated (via calibration results) points' coordinates.

This accuracy computation indicates directly if the chosen calibration method accuracy obtained is good enough to meet the expected depth-reconstruction accuracy condition.

## 3. Context of the proposed method evaluation.

The method has been tested both with virtual and real data (indoor scene and images only).

Grey-level images were taken under usual inside lighting unchanged condition.

The distance between the rig and the pattern was no more than 3 meters. There was no occultation of the pattern or a part of it.

All algorithms have been coded using C++ with Qt graphical user interfaces on a PC platform.

## 4. Results and their discussion.

The experimental virtual data tests show that the depth reconstruction error is proportional to the errors done on the determination of the stereo rig extrinsic parameters. This corresponds to a very small error, a negligible one, when compared to these induced by the errors done on the estimation of the optical centers coordinates (especially if these errors are different for each camera of the rig). For instance, a 5 pixel error on optical center coordinates estimation leads to 1.5mm error at 1m distance. The lenses we use for our vision system have a focal length of 6 mm and induce a distortion which translates image points less than 5 pixels away from their real position.

In fact, the accurate determination of the coordinates of each camera optical center is a

crucial point, and the main difficulty, in stereo rig calibration.

Considering the computing time takes to correct the distortion and the frequency of analyses of in-flight taken pictures, the images aren't distorted enough to make the whole system benefit from a correction of this geometrical anomaly.

Moreover, the distortion effect on the pixels' coordinates has a negligible effect on the depth reconstruction accuracy compared to the one of the noises induced by the UAV and the vision system.

Under real conditions, the pattern used is shown a checkerboard (Figure 4).

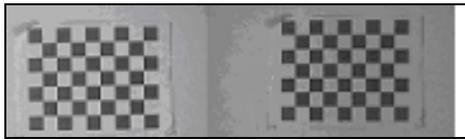


Fig. 4 : Left and Right pictures of the pattern

The pattern was systematically about 1m from the stereo rig. The orientation of this pattern was changing over time and its images in left and right cameras never suffered from occlusion.

The cameras calibration gave apparently good results for the intrinsic parameters. The optical centers seemed correctly centered in the images and the scaling factors were what one could expect. In order to evaluate its effect, the distortion correction was done only using the first order parameter of radial distortion, as it is usually done in literature [4].

Figure 5 shows both real and estimated pattern points in virtual world. A part of them seem to be totally superposed (~40%) while some are slightly shifted.

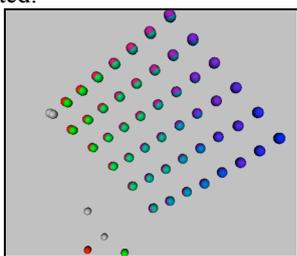


Fig. 5 : Accuracy determination on real data using cameras nearly without distortion.

Our experiments results confirm that the correction of the distortion has a negligible effect on the depth reconstruction accuracy.

The depth reconstruction error mainly comes from the optical centers determination errors. Moreover, the more disparate these errors are between the left and right cameras, the more the depth reconstruction error is important.

With the considered calibration method [12], we managed to come to a set of camera parameters

leading to an accuracy of about 6 cm at 1m distance.

The problem of this method is the instability of the optical center estimation (what can be solved by several calibrations of the same system with the same images).

Depending on the depth accuracy required (more or less than 6% error) one can choose to keep the considered here calibration method or choose another one, giving better principal point's coordinates estimation. For approximate navigation of an UAV the accuracy obtained with the considered method could be judged good enough, but it is probably inappropriate for visual servoing.

## 5. Conclusion

This paper has addressed the accuracy and suitability estimation method of calibration algorithm. The calibration method is evaluated via the error quantification between 3D scene real points and their 3D estimates obtained by triangulation from real points' projections in images. Both computer simulation and real data have been used to show the interest of this technique. The test results of the investigated calibration method applied to indoor scene images taken with a stereo rig onboard a UAV allows to estimate the depth of the real scene objects with a precision of about 6%.

The proposed approach can be used with any kind of camera calibration method which aims to recover metric information from images. This technique only requires to have previously determined camera parameters and to have a pair of images taken at the same instant of time. It is simpler and faster than evaluating the errors on the 22 calibration parameters of a stereo rig and combining them to obtain the depth error.

Compared to usual accuracy determination techniques, and considering the simplicity of the operating mode, this method is immediately exploitable and applicable to the calibration of the vision systems of depth reconstruction applications.

## 6. Acknowledgement

We are most grateful for the support from the DGA. The support of the LISRA is gratefully acknowledged especially M. Laurent ECK and also the programming support of M. Jean-Pierre MARTIN, Systémique SA.

## 7. References

- [1] Beardsley, P., Murray, D., Zisserman, A., Camera Calibration using Multiple Images, ECCV92, 312-320

- [2] Brown, D., C., Calibration of Close Range Camera, 12th Congress of the Int. Soc. Of Photogrammetry, Ottawa, Canada, august 1972
- [3] Devernay F., Faugeras O., Straight lines have to be straight, Automatic Calibration and Removal of Distortion from Scenes of Structured Environments, Machine Vision and Applications, 1, 2001, pp. 14-24
- [4] Fauregas, O. et al., Real-time correlation-based stereo: algorithm, implementations and applications, RR 2013, INRIA,1993
- [5] Faugeras O., the dimentionnal computer vision : a geometrical viewpoint ; MIT Press 1993
- [6] Guermeur Ph., Vision robotique monoculaire : reconstruction du temps-avant-collision et de l'orientation des surfaces à partir de l'analyse de la déformation, Thèse de Doctorat, Université de Rouen, 2002
- [7] Heikkilä J., Silvén O., A four-step camera calibration procedure with implicit image correction, IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'97), San Juan, Puerto Rico, p. 1106-1112.
- [8] Karara, A., Y., I., Direct Linear Transformation into object space coordinate in close-range photogrammetry, Proc. ASP Symp. On Close-Range Photogrammetry, Urbana, Illinois, 1971, 1-18
- [9] Lenz, R. K., Tsai, R., Y. , Techniques for calibration of the Scale Facto rand Image Center for High Accuracy 3D Machine Vision Metrology, IEEE ICRA, Raleigh, vol. 1, 68-75, 1987
- [10] Pissaloux, E., E., Le Coat, F., Tissot, A., Durbin, F., An adaptive parallel system dedicated to projective image matching, IEEE ICIP2000 (Int. Conf. on Image Processing), Vancouver, Canada, 10—13 September, 2000 pp. 184—187
- [11] Roger Y. Tsai, IEEE Journal of Robotics and Automation, "A versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses", Vol. RA-3, No. 4, August 1987, pages 323-344
- [12] Z. Zhang. A Flexible New Technique for Camera Calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11):1330-1334, 2000.