Accurate Camera Calibration with simplified Patterns with a Self-calibration Approach

Federico Pedersini, Augusto Sarti, Stefano Tubaro

Image and Sound Processing Group – D.E.I. – Politecnico di Milano
Piazza L. da Vinci 32, I-20133 Milano, Italy
Phone: +39-2-2399-3444 - Fax: +39-2-2399-3413
E-mail: pedersini, sarti, tubaro@elet.polimi.it

ABSTRACT

In this paper we present a unified approach for calibration and self-calibration of multi-camera systems. The technique is based on an algorithm for the resolution of the self-calibration problem, with which it is possible to find the camera geometry up to an arbitrary perspective transformation of the calibrated scene space. The correct geometry is then obtained by properly exploiting the dimensional constraints introduced with the observation of a very simple calibration pattern. Experimental results have been carried out with many different 3D configurations of the fiducial points, in order to calculate both the minimum number of points, necessary to reach the desired accuracy, and their best disposition in the scene space. Moreover, the accuracy of the proposed calibration method is evaluated and experimentally validated.

INTRODUCTION

Camera calibration is that set of operations with which geometric, optical and electrical characteristics of a camera system are determined. High accuracy in the calibration results is normally highly desirable, especially when performing 3D scene reconstruction from digital perspective views, captured by such camera systems.

There are several calibration approaches, which differ depending on the available *a-priori* information in the calibration scene. One common approach is to use a *calibration pattern* that occupies part of the 3D scene space. The calibration pattern has to be built in such a way to allow a simple and accurate recovery of its geometrical characteristics. Typically, it contains a sufficient number of *fiducial marks*, whose shape allows their accurate localization in the images. When the 3D coordinates of such points, in the external reference frame, are known beforehand, it is possible to exploit their relationship with the corresponding image coordinates in order to estimate the camera parameters. This approach is traditionally known as *camera calibration* [1]. Conversely, if the 3D coordinates of the fiducial points are unknown, the problem is known as a *self-calibration* problem. Without any *a-priori* information on the pattern, the increased number of unknowns makes this problem undetermined, in the sense that it can be solved for only a part of the unknowns. The problem, however, can be made determined if some *a-priori* information about the pattern is available.

Aim of this work is to develop a camera calibration procedure in which the necessary *a-priori* information is introduced in such a way to simplify as much as possible the procedure of acquisition of the calibration pattern. In the proposed calibration scheme, a self-calibration approach is employed for the estimation of the linear part of the camera model. The 3D information provided by the pattern allows then to solve for the residual uncertainty, as well as to estimate the nonlinear part of the model. The key feature of such an approach is that it works with any kind of calibration pattern and any degree of knowledge on the pattern itself (once the minimum information necessary to solve the problem is provided). This freedom can be exploited for making the calibration pattern and the acquisition procedure maximally simple. The importance of a simple calibration pattern is further emphasized by the consideration that one crucial problem of calibration strategies is their *range of validity*. As the calibration pattern plays the role of a sort of "training set", it should be designed in such a way to be statistically representative of the scene to be reconstructed. Accurate results are, in fact, obtained when the pattern properly "fills up" the volume of interest, which means that the pattern should be as large as the scene. For this reason, it is very desirable to calibrate a scene with a pattern of modest size, which is moved an acquired by the cameras in several different positions.

A quite exhaustive set of experiments have been carried out, in order to evaluate the accuracy of the estimation by using very simple calibration patterns. The experiments have been carried out on a trinocular camera system, and two patterns have been tested: a planar surface and a rigid bar, with two fiducial points, situated at the extremities, whose distance is known [6].

CAMERA MODELS AND CALIBRATION PROBLEMS

With *camera model*, we mean the set of mathematical relationships that link the 3D coordinates of a point in the scene space, to the 2D coordinates of its projection on the acquired image. Such relationships can be defined in different ways, as the literature shows. Among them, a distinction could be made between:

- The *physical model*, which describes the projection process in terms of all the optical and geometric parameters of the camera, like its position, orientation, the focal length, the position of the optical axis, etc.
- The *projective model*, which defines an operator **P** (*projection matrix*) that links the coordinates of a 3D point to the coordinates of its projection in the image. In this model, both the 2D image coordinates and the 3D scene coordinates are expressed as homogeneous coordinates. Although such a parametrization is far not so intuitive as the former one, it has the advantage that the relationship between scene points and image points is linear. For each point, the scene coordinates P_W and the corresponding image coordinates P_I, are related as follows:

$$\lambda \cdot P_I = \begin{bmatrix} \lambda\, x_d \\ \lambda\, y_d \\ \lambda \end{bmatrix} =$$

$$= \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \mathrm{P} \cdot P_W$$

Considering the above equation, all the calibration problems can be seen as the problem of determining the projection matrix **P**, given the *image coordinates* P_I of the calibration points and some information about their *world coordinates* P_W in the external coordinate frame. Depending on the degree of knowledge about the world coordinates P_W, the calibration techniques can be actually classified into:

- *Traditional calibration* approaches, where the calibration pattern is completely known, that is, the 3D coordinates of all the fiducial points P_W are given. The only unknown is the projection matrix.
- *Self-calibration* approaches, the complementary case, as in such problems no information is given about the position of the fiducial points P_W. Given the image points, both the projection matrices and the scene structure have to be determined. In the general case, such a problem can be solved up to a scale factor on the reconstruction.

The idea of this work is therefore to develop a calibration scheme that, at first, adopts a self-calibration approach to solve the problem in the projective space, using no knowledge about the world-coordinates of the fiducial marks. After then, the remaining uncertainty is solved for, by properly exploiting the constraints introduced by the available 3-D information of the fiducial marks. The main advantage of this approach is the total flexibility of the algorithm with respect to the knowledge about the fiducial mark. The algorithm is capable to solve the calibration problem even with the minimum degree of knowledge about the fiducial marks.

Aim of this work is to exploit this flexibility in order to simplify as much as possible the structure of the calibration pattern and the procedure for its acquisition.

THE PROPOSED ALGORITHM

The flow diagram of the proposed algorithm is shown in figure 1. As the diagram shows, the estimation procedure is divided in two parts. The linear part of the model, which is determined in closed form, and the non-linear part which is iteratively estimated, through minimization of a cost function. This algorithmic structure has the advantage that the search space through iterative minimization is greatly reduced, while most of the parameters are estimated in closed form. This leads the necessary computation time to be reduced up to 2 orders of magnitude, with respect to traditional calibration procedures [2].

The estimation of the linear part of the model is based on a self-calibration approach [3,4] and has been optimized in order to maximize the computational efficiency.

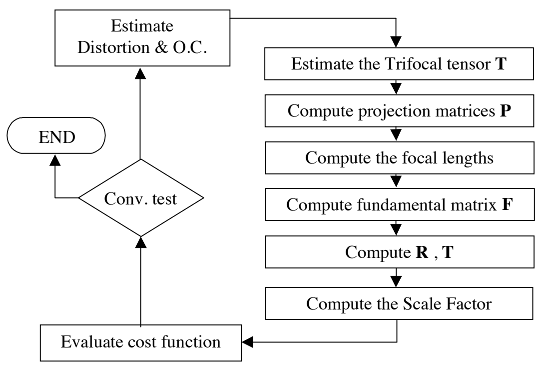

Figure 1. Flow diagram of the proposed technique.

Since a trinocular camera systems is considered, it is particularly efficient to exploit the trifocal constraint, applied to the matched triplets of

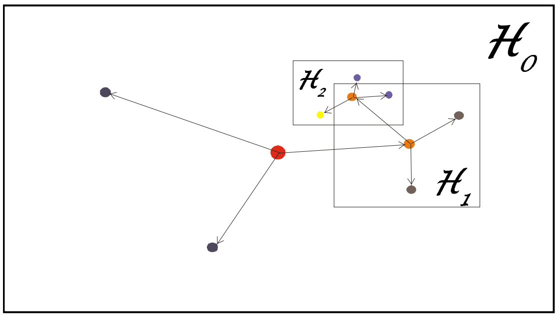

Figure 2 The evolutive strategy of the optimization algorithm. As the estimate becomes more accurate, the search space is correspondingly reduced.

calibration points, in order to estimate the *trifocal tensor* of the system, **T**. From **T**, the *projection matrices* $\mathbf{P_i}$ can be computed by exploiting the invariance of the absolute conic with respect to the perspective projection [5]. For each couple of projection matrices $\{\mathbf{P_i},\mathbf{P_k}\}$, the corresponding *fundamental matrix* $\mathbf{F_{ik}}$ is obtained and, from $\mathbf{F_{ik}}$, it is possible to compute the effective focal lengths f_i and f_j. This allows to derive from **F** the extrinsic and the intrinsic calibration parameters through singular value decomposition, by exploiting the following relationships:

$$\mathbf{F} = \mathbf{K}_2^{-T}\,\mathbf{R}\left[\mathbf{T}\right]_\wedge \mathbf{K}_1^{-1}; \qquad \mathbf{K} = \begin{bmatrix} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

where **R** is the rotation matrix, **T** the translation vector and $\mathbf{K_i}$ is the matrix containing the intrinsic parameters of camera i, being f the focal length and $\{c_x,c_y\}$ the image position of the optical center.

At this point, the calibration problem is solved, up to an unknown scale factor that affects f, **T** and, of course, the reconstructed scene. In order to solve for the scale, the available 3D information of the calibration pattern is exploited. Through the constraints introduced by the given 3D coordinates, or just with given distances among points, the scale factor can be easily determined by solving an over-determined linear system.

The estimation of the non-linear part of the model is performed by iterative minimization of a cost function, which measures both the 3D reconstruction error in the scene and the projection error on the image plane. The estimated parameters are the position of the optical center and two coefficients of the radial lens distortion. In order to obtain a high computation speed, the minimization strategy is based on an optimized genetic algorithm, which has provided comparable robustness but much higher convergence speed, with respect to traditional deterministic approaches.

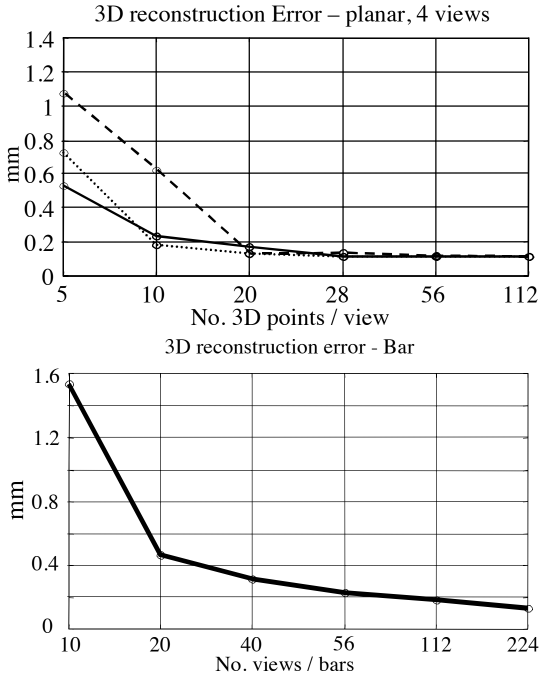

Figure 3 The 3D reconstruction errors, as function of: a) the number of points on a planar pattern (top); b) the number of views of a rigid bar (bottom).

In order to increase the computation efficiency, the minimization is applied to a reduced data-set, which is obtained through random selection of the available fiducial points. The random selection is repeated at each iteration, in order to avoid any polarization of the estimate. Moreover, as long as the estimate gets more and more accurate, the search space is progressively reduced, in order to further increase the computational efficiency (see figure 2).

EXPERIMENTAL RESULTS

The presented technique has been tested for calibrating a trinocular camera system. Two different calibration patterns have been adopted and tested: a planar surface with a set of fiducial marks on the surface, whose coordinates are known in the frame of the pattern, and a rigid bar, with two fiducial points, of known distance, situated at the two extremities. The experiments have been carried out with the following aims: on one hand, the evaluation of the performance of the calibration procedure, in terms of accuracy and computation time. On the other hand, an analysis devoted to find the best configuration for the calibration pattern, that is, to find the minimum number of fiducial points necessary to obtain accurate results and their best distribution in the scene space, in order to maximize the accuracy.

The plot in Figure 3a shows the mean 3D-reconstruction error, as a function of the number of fiducial points employed for the calibration. The different lines in the diagram correspond to different spatial distribution of the points: next to the image center (dashed line), far away from it (dotted) and randomly chosen (solid). The plot in fig. 3b shows the performance obtained using a rigid bar as pattern. The results show that the procedure is able to work with very simple patterns, providing the same accuracy.
From the above plots, it can be derived the following:

- The algorithm allows to reach an accuracy of 0.1 mm in the scene space, corresponding to a relative precision of 150 parts per million, that is the same accuracy provided by calibration with perfectly known 3D calibration patterns.
- A good accuracy can be already obtained with a total number of fiducial points of 100 (4 views with 28 points on the planar pattern; 56 views of the bar).
- The comparison of the two plots in figure 2 shows that the algorithm has been able to obtain, for the same total number of calibration points, the same accuracy, with the bar, as with the planar pattern. This confirms that the proposed technique is effectively able to obtain accurate calibrations with any amount of a-priori 3D information, up to the minimum necessary, thus allowing easy acquisition procedures and the use of very simple and non-critical calibration patterns.

CONCLUSIONS

This paper presents an efficient technique for camera calibration, characterized by fast computation times and maximal simplification of the structure of the calibration pattern and the procedure for its acquisition.
The experiments have shown the proposed technique to be able to generate, with very simple calibration patterns, as accurate results as those obtained by acquiring big ("scene-sized") 3-D patterns.
In particular, quite accurate results and a very simple acquisition procedure have been achieved using the bar as calibration pattern, even if slightly poorer results have been obtained in the estimation of the radial distortion.
A high computational efficiency (up to 2 orders of magnitude, with respect to traditional techniques) has been reached. This is mainly due to the following reasons:

1. The high efficiency of the genetic strategy in the non-linear optimization algorithm.
2. The non-linear part of the problem is limited to its minimum dimensionality, solving the other unknowns in a linear way.

REFERENCES

[1] R. Y. Tsai, *A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using off-the-shelf TV Cameras and Lenses*, IEEE Journal on Robotics and Automation, Vol. RA-3, No. 4, Aug. 1987, pp. 323-344

[2] J. Weng, P. Cohen, M. Herniou, *Camera Calibration with distortion model and accuracy evaluation*, IEEE Trans. On Pattern Analysis and Machine Intelligence, Vol.14, No.10, 1992, pp. 965-980.

[3] Q.T. Luong, O. Faugeras, *Self-calibration of a moving camera from point correspondences and fundamental matrices*, International Journal of Computer vision, Vol.22, No. 3, 1997, pp. 261-289.

[4] R. Mohr, B. Triggs, *Projective Geometry for Image Analysis*, tutorial given at: International Conference on Photogrammetry and Remote Sensing – ISPRS-96, Vienna, July 1996.

[5] A. Heyden, *A common framework for multiple-view tensors*, Proc. European Conference on Computer Vision – ECCV '98, Freiburg, Germany, 1994, pp. 237-256.

[6] N.A. Borghese, P. Cerveri, *Calibrating a Video Camera Pair with a Rigid Bar*, to be published on: Pattern Recognition.