CALIBRATION AND SELF-CALIBRATION OF MULTI-OCULAR CAMERA SYSTEMS

Federico Pedersini[*] , *Danielle Pele*[**] , *Augusto Sarti*[*], *Stefano Tubaro*[*]

* Dipartimento di Elettronica ed Informazione, Politecnico di Milano, P.zza L. da Vinci 32, I-20133 Milano, Italy. Tel.: +39 2 2399 3647 Facsimile: +39 2 2399 3413
E-mail: pedersin, sarti, tubaro@elet.polimi.it

** France Telecom CNET/DIH, CCETT Laboratory, “Advanced image coding and processing” Department - 4 rue du Clos Courtel, BP 59, 35512 Cesson Sevigne Cedex, France, phone: +33-2-99-124-184, fax: +33-2-99-124-098, e-mail: pele@ccett.fr

ABSTRACT

In this paper we present a technique for the accurate calibration of a multiple-camera image acquisition systems used for applications to 3D reconstruction of scenes. The main characteristic of this technique is the fact that it allows us to obtain highly accurate calibrations even when using a very simple and inexpensive calibration target such as a planar pattern. This has been made possible by exploiting the geometry constraints of a multiple-camera acquisition system and by applying a self-calibrating approach that estimates the camera parameters and, at the same time, refines the estimation of the exact position of the calibration pattern's fiducial points. Experiments with rather different setups have been carried out in order to test the proposed calibration techniques in a variety of conditions. The calibration performance, in terms of accuracy of the measured camera parameters, have been predicted through error propagation analysis and verified experimentally through measurement error estimation.

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, for instance, 3D scene reconstruction from digital perspective views, captured by such camera systems.

In the past few years several approaches to the calibration problem have been proposed. Such methods apply to electronic cameras the same techniques that were traditionally used for the calibration of photogrammetric cameras [1,2,3]. The camera characteristics are, in fact, computed through a proper processing of the image of a test object (calibration pattern) placed in the scene. The accuracy of the camera model can be arbitrarily improved by employing an adequate number of parameters [3] therefore, when the goal is that of improving the calibration accuracy as much as possible, the pattern's accuracy becomes the major bottleneck. For this reason, several photogrammetric methods [4] that jointly estimate the camera parameters and the calibration pattern's geometry in a more accurate way have been proposed. These methods are called self-calibration techniques. Because of the increased number of unknowns, self-calibration methods can be successfully applied only when the available data is highly informative, i.e. known with great accuracy, which is the case of photogrammetric images.

In this paper we propose a technique for calibrating multiple-view CCD camera systems. The method we propose is able to achieve very accurate results thanks to the adoption of an appropriate camera model for the description of standard TV-resolution CCD cameras and lenses. The technique is based on a *multi-camera, multi-view* calibration approach, that performs the combined calibration of all the cameras at the same time and by using multiple acquisitions of a simpler calibration pattern, like a planar plate with the fiducial marks attached on the surface, placed in different positions, with respect to the acquisition system. Moreover, a self-calibration module has been developed, that is able to estimate not only the camera parameters, but also the 3-D position of the fiducial points on the calibration pattern, that are only partially known. In order to do that, the calibration problem has been treated as an inverse problem. In fact, inverse problem theory [6] allows us to treat calibration and self-calibration as the same problem, but with different input data. Several experiments have been carried out with different calibration setups, in order to evaluate the performance, the accuracy and the robustness of the calibration algorithm. Some experiments are also presented in order to show the capability of the self-calibration approach, in computing both the camera parameters and to refine the a-priori given input data.

THE CALIBRATION PROBLEM

The camera model

As already said in the Introduction, the calibration problem consists of the identification of the *camera model*, which describes all the characteristics of the acquisition system. 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 many different ways, as the literature shows. Among them, a distinction could be made between those that define an operator (projection matrix) that links the coordinates of a 3D point to the coordinates of its projection in the image [3,7], and those that define a model which involves all optical and geometric parameters of the camera [1,5]. The camera model that we adopted, (see fig. 1) belongs to the latter category.

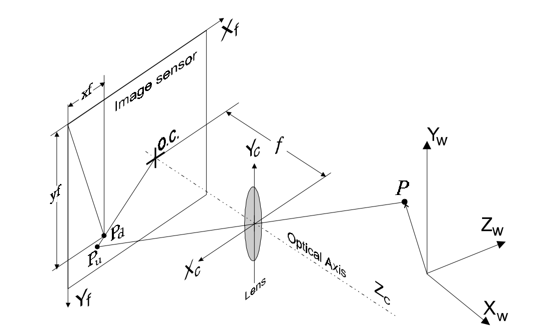

Figure 1. The adopted camera model

The camera model of Fig. 1 provides us with a direct way of calculating the image coordinates of the projection of *P*, given its position in the scene space. This can be represented as a function *g* that maps a point *m* in the model space into a point *d* in the observation space. This set of equations is called *direct model*. For each fiducial point P_i and for each camera C_j, the direct model provides us image coordinates in the form

$$d_{i,j} = g\left(P_i \;, C_j\right) = g\left(m_{i,j}\right);$$

$$d_{i,j} = \left[x_i \quad y_i\right]; \qquad m_{i,j} = \left[P_i \quad C_j\right];$$

where C_j is the set of the 11 parameters which define the model of the *j*-th camera

$$C_j = \left[\varphi, \vartheta, \psi,\; t_X, t_Y, t_Z,\; f,\; k_3, k_5,\; c_X, c_Y\right],$$

f, q and y are the 3 angles that characterize the rotation matrix *R*, $T=\{t_X,t_Y,t_Z\}$ is the translation vector, *f* is the focal length of the lens, k_3 and k_5 are the radial lens distortion coefficients, $OC=\{c_X,c_Y\}$ is the optical center on the image plane. *R* and *T* are usually referred to as *extrinsic parameters*, while the other five are called *intrinsic parameters*, as they characterize the camera. Rewriting the above equations in matrix form and extending them to all the considered fiducial points and all the cameras, the direct model becomes

$$\begin{aligned} \mathrm{g}(\cdot):\quad & \Re^{3N+11\cdot V} && \rightarrow && \Re^{2N\cdot V} \\ & \mathbf{m} && \mapsto && \bar{\mathbf{d}} = \mathrm{g}(\mathbf{m}) \end{aligned}$$

g(.) being the nonlinear function that maps the 3D world coordinates of *N* fiducial points, given the camera parameters of *V* cameras, into the *V* sets of *N* two-dimensional image coordinates of the perspective projection in each camera.

Considering this formalization of the camera model, the problem of camera calibration is that of computing the model vector **m** by exploiting the knowledge about the observed data **d** and the direct model **g**(.). In other words, the solution of the problem is given by the inverse of the model function:

$$\mathbf{m} = \mathbf{g}^{-1}(\mathbf{d})$$

where **d** and **g**() are known. This corresponds to the classical formulation of an inverse problem. In this sense, the camera calibration problems is a typical inverse problem.

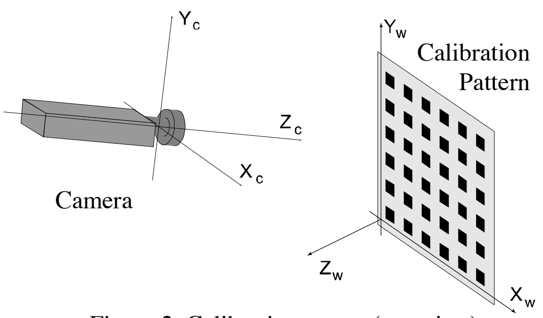

Figure 2. Calibration set-up (one view)

THE ALGORITHM

a) Multi-view, multi-camera calibration

The technique that we propose is able to calibrate a multi-camera system, through multiple acquisitions, with different positions of the calibration pattern with respect to the camera system. The joint calibration of several cameras that point at the same scene allows us to take advantage of the geometrical constraints that the reciprocal position and orientation between the different cameras must satisfy. It is well-known that, in order to obtain satisfactory and reliable calibration results, it is necessary for the fiducial points to "fill up" the entire scene space. This fact usually forces to construct 3D calibration patterns that are as large as the object to be reconstructed. In this case, calibration is possible only for scenes of limited spatial extension. When scenes of larger size are considered (e.g. a video-conferencing scene), 3D pattern cannot be employed anymore. In fact, an accurate 3D pattern is very expensive to build and quite cumbersome to handle. For this reason a *multi-view* calibration set-up has been developed, that, conversely, allows us to generate the desired 3D set of fiducial points through multiple acquisition of a smaller and simpler calibration pattern, such as a planar target. In fact, the pattern can be placed in several different positions, so that the entire 3D volume of interest will be "scanned". The relative motion between different positions of the pattern is not measured, therefore the position of the fiducial points in the scene is only partially known. This, of course, makes the calibration problem much more difficult to approach, as the relative motion of the pattern must be estimated as well. This generates six extra unknowns for each new position, as shown in figure 3a.

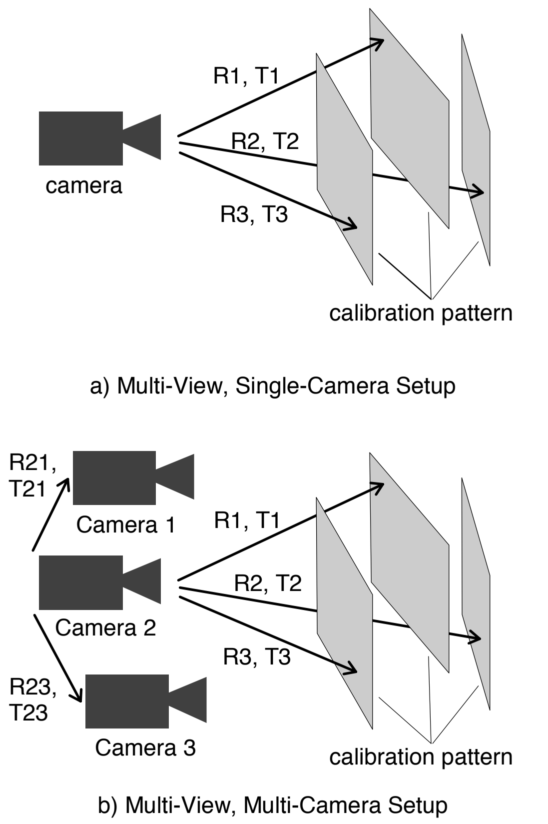

Figure 3

In applications of 3-D reconstruction, a multiple camera system is generally adopted, as any stereoscopic technique needs more views of the same scene object. If the acquisition system to be calibrated is made of more cameras, the *multi-view* calibration should be performed for each one of the cameras. However, if we consider that the pattern motion, from view to view, is the same for all the cameras, this means that each camera gives its contribution to the estimate of the pattern motion. For this reason the calibration, at a same time, of all the cameras of the acquisition system would make the calibration problem more constrained and, consequently, the estimation more reliable. In fact, with respect to the case of calibrating one camera with V pattern views, each new camera to calibrate would lead to 11 more unknowns, while providing approx. $2N \cdot V$ equations (the image coordinates of NV fiducial points), as shown in figure 3b. The final approach is then a *multi-view, multi-camera* calibration, where the parameters of all the cameras composing the acquisition system are estimated together, within the same error minimization process.

b) The self-calibrating approach

Thanks to the large number of constraints and to the fact that, through multiple pattern positioning, the fiducial points end up covering the whole scene space, the self-calibrating approach can lead to the best results that can be obtained with such low-cost calibration setups, in terms of global accuracy throughout the scene space. In fact, experimental results have shown that, under proper conditions, the achieved calibration accuracy, with this approach, reaches the limit imposed by the accuracy with which the position of the fiducial points is known, with respect to the pattern frame.

In order to further improve the accuracy of the calibration, it is either necessary to use calibration patterns of higher precision, for which the fiducial point coordinates have been determined with high accuracy (e.g. with photogrammetric techniques) or to adopt a self-calibrating approach, which, besides estimating the camera parameters, refines the estimates of the a-priori given coordinates of the fiducial points. The complexity of the former solution is the same as in the previous case, but it requires expensive calibration patterns. As the aim of this work is that of obtaining high performance at low cost, we focused on the latter solution.

The self-calibration problem is much more undetermined than the previously considered one, because the calibration points coordinates WP are considered only approximately known. In other words, also the data points become, to some extent, unknowns to be estimated. The a-priori knowledge about the data generally consists of a rough estimate of the world-coordinates. The proposed technique is able not only to further improve the accuracy of the estimated camera parameters, but to refine the a-priori given estimate of the world coordinates as well.

EXPERIMENTAL RESULTS

Several experiments have been carried out with rather different calibration setups, in order to test the proposed calibration techniques in many different conditions. In particular, a series of calibration tests have been performed within the EC Project "ACTS-PANORAMA", for calibrating a trinocular camera system. For such experiments, the adopted calibration pattern was a set of circular points applied on a planar rigid structure (honeycomb aluminum plane), shown in fig. 3. The image coordinates were computed through image localization. A template matching technique is adopted, in order to extract the image coordinates with sub-pixel accuracy. After detection of each visible fiducial point, the luminance function in the neighborhood of the point is compared to a synthetic luminance template; through a mean-squares optimization, the parameters of the ellipse are estimated. Evaluation tests on this algorithm estimated an accuracy of about 0.1 pixel along both axes. The localized fiducial points were then sorted according to their position in the scene, in order to be matched to their corresponding 3D world coordinates. The position of the fiducial points, relative to the pattern's frame, was known with an accuracy that was better than 0.1 mm (measured with photogrammetric methods). By using this knowledge, the estimated camera models resulted in a standard deviation that was better than 0.2 pixel on the image plane, corresponding to less than 0.4 mm in the scene space.

In the case of self-calibration, it was possible to achieve about the same accuracy as above, without using the exact world coordinates of the fiducial points. The only available a-priori information about the world coordinates was the nominal step size of the rectangular grid of fiducial points, where the points were attached manually to the plate. The last multi-camera, multi-view estimation phase, also the exact position of the fiducial points has been considered an unknown while the nominal position has been used as initial estimate. The exact coordinates of the fiducial points resulted, in fact, to be recovered with an accuracy that was comparable with that of the photogrammetric measurements.

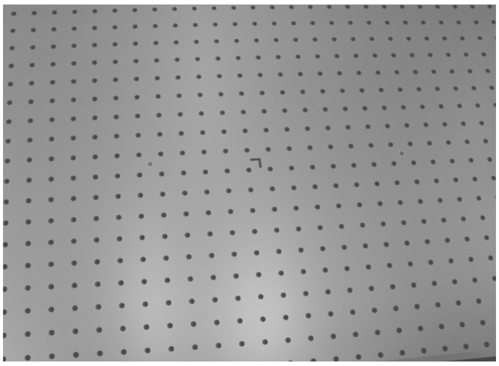

Figure 3. A view of the planar calibration pattern

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 J. 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 PAMI*, Oct 1992, Vol. 14, No 10, 965-980

[3] Y. Yakimowsky, R. Cunningham, "A System for Extracting Three-Dimensional Measurements from a Stereo Pair of TV Cameras", *Computer Graphics and Image Processing*, No 7, pp. 195-210, 1978

[4] W. Faig, *Manual of Photogrammetry*, 4th ed., Am. Soc. of Photogrammetry, 1990

[5] F. Pedersini, S. Tubaro, F. Rocca, "Camera Calibration and Error Analysis. An Application to Binocular and Trinocular Stereoscopic Systems", *4th Int. Workshop on Time-Varying Image Processing and Moving Object Recognition*, Florence, 1993

[6] A. Tarantola, *Inverse Problem Theory*, Elsevier, 1987

[7] N. Ayache, *Artificial vision for Mobile Robots*, MIT Press, 1991

[8] M. Born, E. Wolf, *Principles of Optics*, Pergamon Press, 1959