Computer Vision and Image Understanding 113 (2009) 11–28

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

Estimating parameters of noncentral catadioptric systems using bundle adjustment

Nuno Gonçalves *, Helder Araújo

Institute of Systems and Robotics—Coimbra, Department of Electrical and Computers Engineering, University of Coimbra, Pinhal de Marrocos—Polo II, 3030 Coimbra, Portugal

ARTICLE INFO

Article history:
Received 2 October 2006
Accepted 23 June 2008
Available online 9 July 2008

Keywords:
Noncentral catadioptric
Calibration
Bundle adjustment
Apparent contour
Projection model
Ray calibration

ABSTRACT

This paper describes a new method to calibrate the intrinsic and extrinsic parameters of a generalized catadioptric camera (central or noncentral). The algorithm has two steps. The first one is the estimation of correspondences between incident lines in space and pixels (black box model calibration) in an arbitrary world reference frame. The second step is the calibration of the intrinsic parameters of the pinhole camera, the coefficients of the mirror expressed by a quadric (quadric mirror shape and the pose of the camera in relation to it), the position of the optical center of the camera in the world reference frame and its relative orientation (pose of the camera in world reference frame). A projection model relaxing Snell's Law is derived. The deviations from Snell's Law and the image reprojection errors are minimized by means of bundle adjustment. Information about the apparent contour of the mirror can be used to reduce the uncertainty in the estimation by introducing a new term in the cost function of the second step minimization process. Simulations and real experiments show good accuracy and robustness for this framework. However, the convergence is dependent on the initial guess as expected. A well-behaved algorithm to automatically generate the initial estimate to be used in the bundle adjustment is also presented.

1. Introduction

Panoramic vision is a relatively recent chapter of computer vision which conjugates characteristics such as wide field of view and poor space resolution, due to the nature of the image formation. Several other important characteristics often depend on the type of the projection: central or noncentral. These designations distinguish between those systems where all incident light rays intersect in one single point called center of projection (central systems) and those where such a point does not exist (noncentral ones). Nowadays, central systems, specially central catadioptric cameras (made up by one camera and a reflecting mirror), are very popular and their complexity has been reduced with extensive work in calibration, 3D-reconstruction, motion, structure from motion and applications.

Noncentral vision systems in general, and catadioptric ones in particular, present some advantages over central systems, since the designer is able to place camera and mirror in unconstrained positions, allowing zooming and resolution enhancing in some selected regions of the image, for instance. However, noncentral catadioptric systems are much less used than central ones, mainly in applications where some accurate measurements are needed. The relative complexity of such systems is still high. There is no robust linear calibration algorithm for them and thus their applications often require less accuracy. The reason for such situation is mainly the following one: the non-existence of a projection model relating 3D points in world to 2D image points. The projection of a given world point to the image can be calculated by back-projecting image pixels until the incident ray pass over the given point. There are some alternative methods, but all of them are non linear and implicit. In other words, there is no closed-form expression for the camera projection, mapping 3D world points to image. For this reason the accurate calibration of these systems is difficult and also their use in applications requiring high precision measurements.

Some recent calibration methods provide the correspondence between pixels and a particular line in space. They are based on the black box model introduced by Grossberg and Nayar [1] for which the calibration is no more than this list of correspondences (pixel $\longleftrightarrow$ 3D line). The vision system is therefore considered a black box model and the path of light rays is unknown as well as the reflection model. Whether it is central or noncentral it is not relevant for the method.

Grossberg and Nayar [1] presented a method to calibrate in this sense general vision systems with structured light patterns and Sturm and Ramalingam [2] have also formulated a method to calibrate generalized cameras. Once the calibration is performed the results can be applied to 3D reconstruction, motion analysis and several other applications such as described by Pless [3], Ramalingam et al. [4] and others.

In the method of Grossberg and Nayar [1] motion is considered to be known and binary encoding of image pixels is used to calibrate the vision system. On the other hand, the generalized method presented by Sturm and Ramalingam in [2] uses no more than

* Corresponding author. Fax: +351 239 406672.
E-mail address: nunogon@isr.uc.pt (N. Gonçalves).

1077-3142/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.cviu.2008.06.004

three images acquired by the vision system in three different positions and the local geometrical description of a calibration object. The knowledge of the motion between the camera in the three different positions is hence not required. The method provides the correspondence between each pixel and a 3D ray in space, expressed in world coordinates (made coincident with the local reference frame of the first image). No information on the projection itself is thus present in the calibration and therefore nothing is known about the intrinsic parameters and type of the camera. Ramalingam et al. [5] presented an extension of this method to allow the use of more than three images. Overlapping calibration grids are used to produce an initial estimate of the calibration parameters and then pose estimation and bundle adjustment algorithms are used to improve the estimation results. This proved to be applicable to a wide variety of camera types expressed as generalized models.

Catadioptric vision systems made up of cameras and specular surfaces are specially suitable to these type of calibration methods. Indeed, in those cases, the projection model is described by several parameters which are not considered if the calibration is performed in the sense of the black box model. Furthermore, as mentioned above, a closed-form projection model exists only when the projection is central (central catadioptric cameras, see [6–9]) and no explicit projection model is known for noncentral catadioptric cameras ([10]). For central catadioptric vision systems there is an Omnidirectional Calibration Toolbox by Christopher Mei [11] based on the Camera Calibration Toolbox by Jean-Yves Bouguet [12].

The alternative class of bundle adjustment methods for camera calibration requires the knowledge of the projection model that maps 3D points to the image plane such that the Jacobian of the projection equations can be evaluated. Since there is no closed-form projection model for noncentral catadioptric cameras and hence for its Jacobian, this class of methods has not been used for their calibration. Numerical evaluations of the projection function and its Jacobian can however be used to perform the wanted calibration. Lhuillier [13] recently presented a model for catadioptric camera calibration using numerical evaluations for bundle adjustment methods.

We are interested in the estimation of the intrinsic and extrinsic parameters of general catadioptric cameras with quadric mirrors regardless of being central or not. The method is composed by two steps. In the first one the system is calibrated in the sense of a black box model, that is, we assume that the correspondence pixel $\longleftrightarrow$ 3D line is provided by using Grossberg and Nayar method [1], Sturm and Ramalingam method [2] or by some other method. We opted to use known motion between dense calibration grids to perform a stable ray calibration.

The second calibration step presented in the paper proposes the application of the class of bundle adjustment methods for camera calibration to general (central or not) catadioptric cameras. The explicit computation of the Jacobian of the projection equations is possible due to the relaxation of Snell's law constraint. The non existence of closed-form equations for the projection (and hence the non existence of a means to provide an estimate for the coordinates of the reflection point on the mirror surface) is circumvented by the fact that there are available correspondences between pixels and lines in space and not between pixels and points in space. The intersections between the direction rays and the mirror surface thus provide the reflection points.

Bundle adjustment is then applied to the projection model by using the following parameterization: intrinsic parameters of the pinhole camera (5 parameters), position and orientation of the camera in the world reference frame (three rotation angles and three displacements—6 parameters), the quadric mirror shape parameters in canonical form (3 parameters) and the position and orientation of the camera in relation to the mirror (three rotation angles and three displacements—6 parameters). The total number of parameters of the state vector is 20. We show that bundle adjustment methods are suitable for the calibration of general catadioptric cameras and that the convergence is generally achieved both in experiments with simulated data and in experiments with real images. Since bundle adjustment methods require an initial guess for the state vector, we also provide an automatic algorithm to compute the initial estimates.

Rotations are parameterized by Euler angles. As is widely known, usually Euler angles present stability and numerical problems due to their high non-linear nature. This problem is solved by using frozen (or cumulative) and update rotation matrices in the bundle adjustment optimization algorithms. The current estimate in each iteration is frozen in a rotation matrix and the derivatives are evaluated in the update angles rather than in the accumulated angles. This strategy provides very simple Jacobian expressions. After the update state vector is computed, the next iteration starts by accumulating the last update in the frozen angles. See for instance [14,15].

Our method was also previously presented in [16]. As the approach is similar, some differences exists between the current improved method and the previously presented. The main differences are the parameterization of the rotation angles that were described by quaternions and the parameterization of the quadric mirror that was previously described by a full rank quadric matrix. The method is now able to calibrate the shape of the mirror independently of the pose of the camera in relation to it. This showed to provide better calibration results. Additionally, a more comprehensive experimental set of tests is now presented.

In the next section the problem is described and discussed. Some considerations regarding the mathematical tools used are discussed and the notation is also presented. Next, Section 3 presents the projection model relating, in closed-form, the 3D lines in space with a point in the image plane. The first step of the calibration method is then discussed in Section 4. In Section 5 we discuss the minimization of the cost function by means of a bundle adjustment method and we also discuss and present an automatic algorithm to compute the initial estimates from where the bundle adjustment minimization should start. In Section 6 the apparent contour of the mirror quadric is used to enhance the convergence of the algorithm, by presenting a new constraint to be added to the cost function. Section 7 presents the experiments and the results obtained. Finally, Section 8 contains the main conclusions and the future directions of the work. In Appendix, we derive the explicit expressions of the cost function and its Jacobian.

2. Problem statement

Consider a catadioptric vision system made up of a pinhole camera whose intrinsic parameters are given by matrix **K**:

$$\mathbf{K} = \begin{bmatrix} f_u & \nu & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1}$$

and a specular mirror surface given by the quadric in its canonical form:

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & A_s & B_s \\ 0 & 0 & B_s & -C_s \end{bmatrix} \tag{2}$$

The camera is positioned in the center of the main reference frame and its poses (position and orientation) relative to the quadric mir-