ISPRS Journal of Photogrammetry and Remote Sensing 108 (2015) 72–79

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Improved wide-angle, fisheye and omnidirectional camera calibration

Steffen Urban *, Jens Leitloff, Stefan Hinz

Karlsruhe Institute of Technology (KIT), Institute of Photogrammetry and Remote Sensing, Englerstr. 7, 76131 Karlsruhe, Germany

A R T I C L E I N F O

Article history:
Received 24 June 2014
Received in revised form 24 March 2015
Accepted 22 June 2015

Keywords:
Camera
Calibration
Accuracy
Precision
Geometric
Robotics
Bundle
Fisheye
Omnidirectional

A B S T R A C T

In this paper an improved method for calibrating wide-angle, fisheye and omnidirectional imaging systems is presented. We extend the calibration procedure proposed by Scaramuzza et al. by replacing the residual function and joint refinement of all parameters. In doing so, we achieve a more stable, robust and accurate calibration (up to factor 7) and can reduce the number of necessary calibration steps from five to three. After introducing the camera model and highlighting the differences from the current calibration procedure, we perform a comprehensive performance evaluation using several data sets and show the impact of the proposed calibration procedure on the calibration results.

1. Introduction

In this paper we extend an existing method to calibrate wide-angle, fisheye and omnidirectional cameras based on the generalized camera model introduced in Scaramuzza et al. (2006a). In the following, we propose contributions that not only enhance the methods implemented in Scaramuzza (2014), but also the underlying calibration procedure. Our main contribution will be the reformulation of the two step non-linear optimization to jointly refine all calibration parameters. We then perform various calibrations on data sets provided by the authors (Scaramuzza, 2014) and own data sets. Afterwards, we elaborately analyze the results and detail the significant gain in performance and accuracy. In addition, we calibrated all cameras with another calibration method proposed by Mei and Rives (2007) which is also available as a Matlab toolbox (Mei, 2014). The Matlab implementation of all proposed contributions as well as test data sets and evaluation scripts will be made available online.[1]

1.1. Related work

Wide-angle, omnidirectional and fisheye cameras are nowadays popular in many fields of computer vision, robotics and photogrammetric tasks such as navigation, localization, tracking, mapping and so on. As soon as metric information needs to be extracted, camera calibration is the first necessary step to determine the relation between a 3D ray and its corresponding mapping on the image plane.

In general, camera calibration consists of two steps. The first involves the choice of an appropriate model that describes the behavior of the imaging system. In recent years various such models for dioptric (fisheye) and catadioptric (omnidirectional) cameras have been proposed (Scaramuzza et al., 2006a,b; Mei and Rives, 2007; Micušık, 2004; Kannala and Brandt, 2006; Geyer and Daniilidis, 2000). In Puig et al. (2012) a comprehensive overview is given. The second step is the estimation of all parameters that a certain model incorporates, i.e. interior and exterior orientation of the system as well as distortion parameters that model the differences from an ideal imaging process. Several methods for calibration have been proposed, that can be classified into three categories (Zhang, 1999) starting with the most general one.

The first is auto- or self-calibration, where epipolar-constraints and point correspondences between multiple views are used (Micušık, 2004; Micušık and Pajdla, 2006; Faugeras et al., 1992).

* Corresponding author.
E-mail address: steffen.urban@kit.edu (S. Urban).
URL: http://www.ipf.kit.edu (S. Urban).
[1] http://www.ipf.kit.edu/code.php.

http://dx.doi.org/10.1016/j.isprsjprs.2015.06.005

In Micušík and Pajdla (2006) the auto-calibration for fisheye and catadioptric systems is based on the robust establishment of the epipolar geometry from few correspondences in two images by solving a polynomial eigenvalue problem. Using this model, Ransac is employed to remove outliers and no prior knowledge about the scene is necessary. Subsequently, bundle block adjustment can be employed, to estimate all camera parameters and reconstruct the scene except for a similarity transformation.

The second is space resection of one image from a non-planar object with known 3D world coordinates (Schneider et al., 2009; Puig et al., 2011). Schneider et al. (2009) investigate the validity of geometric projections for fisheye lenses. The camera is calibrated using one image of a room prepared with 3D-passpoints, that cover the field of view of the fisheye lens. To enhance the model quality and compensate for real world deviations from the geometric model, radial as well as tangential distortion parameters are subsequently added to the projection model. A camera calibration approach for catadioptric cameras using space resection is proposed in Puig et al. (2011). Lifted coordinates are used to linearily project scene points to the image plane. After the projection matrix is estimated from scene points distributed over a 3D-calibration object a final non-linear refinement is applied.

The third calibration method involves the observation of a planar object (e.g. checkerboard) with known 3D object coordinates (Scaramuzza et al., 2006a; Mei and Rives, 2007; Zhang, 1999) where at least two images of the planar object are necessary (Zhang, 1999). The calibration procedure of Mei and Rives (2007) involves the manual selection of mirror border and center, respectively. For fisheye lenses the image center is used. The generalized focal length, which is part of the exact projection model, is estimated from at least three points on a line. To automatically extract all checkerboard corners, a planar homography is estimated from four manually selected points. Then the remaining checkerboard points are projected to their estimated location and a local subpixel corner search is performed. The initially estimated homography is based on start values such as the generalized focal length and the center of distortion and thus already imposes model assumptions. If the initial assumptions are highly biased the subsequent procedure can hardly recover. The projection model of Mei and Rives (2007) include tangential distortion parameters, whereas Scaramuzza et al. (2006a) models only radial symmetric effects. Scaramuzza et al. (2006a) initially extracts all checkerboard corners using the algorithm of Rufli et al. (2008). The calibration procedure is based on multiple observations of a planar calibration object (Zhang, 1999), is fully automatic and does not require further user interaction or prior knowledge. A potential disadvantage are wrong point extractions, that can bias the calibration result. Also the two-step refinement of interior and exterior orientation is critical, as all parameters are strongly correlated. Our proposed approach tries to avoid both possible disadvantages, but does not change the underlying projection model.

1.2. Contributions

In this paper we extend and improve the widespread calibration procedure of Scaramuzza et al. (2006a) and its online available implementation (Scaramuzza, 2014) as this does not provide an optimal solution. The camera model proposed in Scaramuzza et al. (2006a) is general and does not require prior knowledge about the imaging function. The calibration procedure is almost fully automatic, online available (Scaramuzza, 2014) and is conducted in five consecutive steps.

We analyzed the two step refinement of interior and exterior orientation and replaced the employed residual function in order to be able to jointly refine all parameters. Thereby we obtain better convergence behaviors of the non-linear least squares optimization and it shows that the original five step calibration procedure can be conducted in three steps. In addition, we change the original implementation to a consistent use of subpixel accurate point measurements. All changes are available online and can be used as an add-on to the original toolbox at hand.

Previous work for improving (Scaramuzza et al., 2006a) was conducted by Frank et al. (2007). They extended the camera model by replacing the affine transformation between image and sensor plane coordinates with a perspective transformation and also proposed a new method to estimate the center of distortion. They indicate that a bundle adjustment was performed in the end without providing details. Then they conducted calibrations on several data sets and compared the results to the original implementation. They observed slight improvements of the mean squared reprojection error (MSE in pixel) but did not discuss why the standard deviation of the reprojection error was partly larger than the MSE and even larger than the standard deviation of the original implementation. Further, only the reprojection error was investigated to assess the calibration quality. This paper additionally examines the quality of the estimated calibration parameters.

In the remainder of this paper, we first recapitulate the camera model as well as the calibration procedure as proposed in Scaramuzza et al. (2006a). Then, we detail our contributions to the existing approach and perform a comprehensive evaluation using data sets provided by the authors (Scaramuzza, 2014) as well as own data sets and show that we achieve a significant accuracy increase. The improved results are subsequently compared to a different calibration method proposed by Mei and Rives (2007).

2. Camera model and calibration

In the following and for the sake of the self-containedness of this paper, we include a condensed compilation of the generalized camera model of Scaramuzza et al. (2006a) and the five consecutive steps that it takes to estimate all model parameters. For a detailed description and insight we refer to Scaramuzza et al. (2006a,b). We use a similar notation as Scaramuzza et al. (2006a) to facilitate the comparison, i.e. vectors and matrices are typed in boldface, e.g. $\mathbf{R}, \mathbf{m}$, thereby vectors are lowercase and matrices are uppercase. In addition parameters to be estimated are indicated by a hat, e.g. $\hat{\mathbf{R}}, \hat{a}$.

Subsequently, we detail the procedure that is employed to estimate all calibration related parameters. In Section 3 we propose our improvements to the given procedure and their implementation.

2.1. Generalized camera model

Let $\mathbf{m}' = [u', v']^T$ be a point on the image plane, i.e. the coordinates are measured from the upper left corner of the image. Its corresponding point on the sensor plane, i.e. measured from the image center, is denoted by $\mathbf{m} = [u, v]^T$. Those two points, are related by an affine transformation $\mathbf{m}' = \mathbf{A}\mathbf{m} + \mathbf{O}_c$:

$$\begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} c & d \\ e & 1 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} + \begin{bmatrix} o_u \\ o_v \end{bmatrix} \tag{1}$$

The transformation matrix $\mathbf{A}$ accounts for small misalignments between sensor and lens axes and the digitization process. The translation $\mathbf{O}_c = [o_u, o_v]^T$ relates all coordinates to the center of distortion. Now let $\mathbf{X}_c = [X_c, Y_c, Z_c]^T$ be a scene point in the camera coordinate system. Then, the following forward mapping from a 2D image point to its corresponding 3D ray is described through the imaging function g.

$$\lambda g(\mathbf{m}) = \lambda(u, v, f(u, v))^T = \lambda(u, v, f(\rho))^T = \mathbf{X}_c \tag{2}$$