Auto-calibration of Pan-Tilt Cameras including Radial Distortion and Zoom

Ricardo Galego, Alexandre Bernardino, and José Gaspar

Institute for Systems and Robotics, Instituto Superior Técnico / UTL, Lisboa, Portugal
{rgalego,alex,jag}@isr.ist.utl.pt

Abstract. Although there are intrinsic advantages of using pan-tilt-zoom cameras their application in automatic surveillance systems is still scarce. The difficulty of creating background models for moving cameras and the difficulty of keeping fitted pose and optical geometrical projection models are key reasons for the limited use of pan-tilt-zoom cameras. Geometric calibration is a useful tool to overcome these difficulties.

In this work we propose a method for PTZ camera auto-calibration over the camera's zoom range. A new approach based on a *division distortion model*, which allows designing linear algorithms, is followed to solve the radial distortion when images are captured in different zooms. Results obtained over both synthetic and real data show that a full zoom range, complete field of view, pan-tilt-zoom camera calibration is possible.

1 Introduction

Despite the high versatility and potential of pan-tilt-zoom cameras, their use in current surveillance systems is still much less frequent than the use of fixed cameras. The limited dissemination of pan-tilt-zoom cameras can be justified by the slightly higher costs of the hardware and the higher risk of failure due to the mechanical components. This is however just a partial justification, since the hardware costs and failures can decrease significantly with mass production. Other, more compelling, justifications arise from the operation of the cameras and the difficulty of developing surveillance methodologies. Aiming to have mostly high quality employment implies that installations with numerous pan-tilt-zoom cameras cannot involve many human operators. This motivates developing automatic control and surveillance methodologies in opposition to the currently utilized manual control. The difficulty of creating background models for moving cameras and the difficulty of keeping fitted pose and optical geometrical projection models are also key reasons for the limited availability of automatic surveillance methodologies provided by the industry. Geometric calibration is a useful tool to overcome these difficulties.

There are several methods documented in the literature for calibrating cameras. The method proposed by Bouguet [1] is nowadays one of the most used calibration methods. Bouguet's method allows estimating intrinsic and radial distortion parameters based on imaging a planar chess pattern placed at various orientations. In the case of mobile cameras, in particular the pan-tilt-zoom cameras, there has been shown that they can be auto-calibrated using natural features of a static scenario. Hartley [2] presented an

auto-calibration method for rotating cameras and later Agapito et al. [3, 4] introduced a self-calibration method for rotating and zooming cameras. These works showed that geometric calibration of intrinsic and extrinsic parameters can be achieved without requiring non-linear optimization. Estimating of the radial distortion still implied non-linear optimization. Sinha and Pollefeys [5] also recurred to the non-linear minimization of reprojection errors in order to consider the estimation of radial distortion in pan-tilt-zoom cameras. Their approach is based on estimating intrinsic and radial distortion coefficients using imagery taken within a small range of the camera's FOV. In all methods presented till here, there is no approach which achieves an accurate estimation of both intrinsic and radial distortion parameters with low (time) computational costs.

One of the key aspects justifying the complexity of including the estimation of radial distortion in the calibration process is that the distortion model cannot be easily inverted. To overcome this aspect, Fitzgibbon [6] proposed the *division distortion model*, which directly maps the radially distorted points into undistorted points. This direct mapping allows estimating the radial distortion and other calibration parameters in a manner very similar to the conventional calibration of pan-tilt cameras [2]. Steele and Jaynes [7] improved Fitzgibbon's algorithm by making it more stable. Kukelova and Pajdla [8] solve the same problem with the use of the fundamental matrix.

In this work, we propose using a linear algorithm [3] for the calibration of the intrinsic parameters of a pan-tilt-zoom camera, including the estimation of radial distortion. A new approach is followed for solving the radial distortion when images are captured at different zooms.

This paper is organized as follows: Section 2 introduces the pan-tilt-zoom camera model, Section 3 presents several calibrations methodologies, Section 4 proposes a methodology for the calibration of radial distortion when in presence of zoom, Section 5 shows some results with real and synthetic data, and finally, in Section 6 some conclusions are drawn and the future work is stated.

2   Pan-Tilt-Zoom Camera Model

The pin-hole camera model for the perspective pan-tilt-zoom camera consists of a mapping from 3D projective space to 2D projective space. This is represented by a 3x4 rank-3 perspective matrix, P. The mapping from 3D to the image plane takes a point $X = [X\ Y\ Z\ 1]^T$ to a point $u = PX$ in homogeneous coordinates. The projection matrix is usually decomposed as:

$$P = K^z[R\ t] \tag{1}$$

where t is a 3x1 vector that represents the camera location, R is a 3x3 rotation matrix that represents the orientation of the camera with respect to an absolute coordinate frame and K^z is a 3x3 upper triangular matrix encompassing the intrinsic parameters of the camera:

$$K^z = \begin{bmatrix} k_u & s & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

where k_u and k_v are the magnifications in the respective u and v directions, u_0 and v_0 are the coordinates of the principal point of the camera and s is a skew parameter (in this work we assume $s = 0$).

2.1 Specific aspects of pan-tilt-zoom cameras

Contrarily to cameras with fixed optics, pan-tilt-zoom cameras allow zoom in/out, and thus allow varying the K^z matrix.

In addition, note that the pan and tilt movements are simply rotations about the projective center $O = -R^{-1}t$, which is usually chosen to be the world origin and thus $t = [0\ 0\ 0]^T$. The pan and tilt movements are included in R, and thus R is also a time varying matrix.

2.2 Radial distortion

Most cameras deviate from the pin-hole model due to radial distortion. This effect decreases with increasing focal length. Due to radial distortion a 3D point X is projected to a point $x_d = [x_d\ y_d]^T$. This point is deviated from the point $x = [x\ y]^T$ according to the radial distortion function, $\mathfrak{R}^z$:

$$[x_d\ y_d]^T = \mathfrak{R}^z\left([x\ y]^T;\ k_1,\ k_2\right) = L(r)[x\ y]^T \tag{3}$$

where $L(r) = (1 + k_1 r^2 + k_2 r^4)[x\ y]^T$ and $r = \sqrt{x^2 + y^2}$. This radial distortion model corresponds to a simplified two coefficient version of the model proposed by Heikkila [9] where r is the radial distance (distance from point x to the center of distortion (x_c, y_c)), $L(r)$ is a radially symmetric distortion factor and k_1 and k_2 are the two radial distortion coefficients considered. For every zoom level z, $\mathfrak{R}^z$ is parameterized by $(x_c^z, y_c^z, k_1^z, k_2^z)$. In this work we assume that the principal point (u_0, v_0) is constrained to be the center of distortion and therefore the radial distortion function is only parameterized by coefficients k_1^z and k_2^z.

In summary, the goal of the calibration process of a pan-tilt-zoom camera is to estimate the unknown parameters of a model $(K^z, R$ and $\mathfrak{R}^z)$ that provides the intrinsic parameters for any pan, tilt and zoom admissible configuration. Since R is likely to change continuously due to the operation of the camera, in general the calibration of pan-tilt-zoom cameras refer just to estimating K^z and $\mathfrak{R}^z$.

3 Calibration Methodologies

This section contains a comprehensive review of auto-calibration methodologies for pan-tilt-zoom cameras.

3.1 Calibration Method by Agapito et al.

Agapito et al. in [10] introduced an image based auto-calibration method that estimates the intrinsic parameters of a rotating camera. Given a 3D world point X, it has different

projections on different images, $x_i = P_i X$. Since X is the same on all the images and $P = KR$, we can calculate the relation between the images

$$x_i = P_i P_j^{-1} x_j. \tag{4}$$

The relation between images is a homography (projective transform) $H_{ji} = P_i P_j^{-1} = K_i R_{ij} K_j^{-1}$. Since R is a rotation matrix, it satisfies $R = R^{-T}$, and thus

$$H_{ji} K_j K_j^T H_{ji}^T = K_i K_i^T. \tag{5}$$

In the cases that $K_i = K_j$, i.e. images with constant zoom, Eq. 5 can be solved by a set of linear equations in the entries of KK^{-T}.

3.2 Bundle Adjustment, Sinha and Pollefeys' Calibration Method

Sinha and Pollefeys [5] proposed an algorithm to calibrate simultaneously the intrinsic parameters of a camera as well its radial distortion. Their method is based on multiple bundle adjustments. Given a set of images, their algorithm estimates the radial distortion, $\mathfrak{R}$, and the homographies among the images, T, by minimizing the reprojection error, D, on the surface of a cube:

$$min_{T_i, \mathfrak{R}(z_{min}), X^j} \sum_{j=1}^{m} \sum_{i=1}^{n} D(x_i^j, \mathfrak{R}(T_i X^j))^2. \tag{6}$$

Having an initial estimate of the radial distortion and of the homographies, the algorithm uses a method like Hartley and Agapito's to solve for the intrinsic parameters. Afterwards the method encompasses another bundle adjustment to refine the radial distortion parameters and the intrinsic parameters:

$$min_{K(z_{min}), R_i, \mathfrak{R}(z_{min}), X^j} \sum_{j=1}^{m} \sum_{i=1}^{n} D(x_i^j, K(\mathfrak{R}(R_i X^j)))^2 \tag{7}$$

As it has several bundles this method can be slow, depending on the number of images and corresponding points.

3.3 Division distortion model, Fitzgibbon's Calibration Method

Fitzgibbon proposed an auto-calibration algorithm that finds the homographies among images taken at various pan and tilt poses and, simultaneously, finds the radial distortion characterizing the camera at a fixed zoom level [6]. In order to estimate simultaneously the homographies and the radial distortion, the radial distortion is represented in the so called *one-parameter division distortion model*:

$$x = \frac{1}{1 + \lambda \|x_d\|^2} x_d \tag{8}$$

where x is the undistorted pixel position, x_d is the distorted pixel position and λ is the radial distortion parameter. The calibration methodology estimates the radial distortion from a pair of images. From Eq. 4 one has that x_i is equal to Hx_j, which leads to:

$$x_i \otimes Hx_j = 0 \tag{9}$$

where $\otimes$ denotes the Kronecker product. If x is replaced using Eq.8 then one obtains:

$$(x_i + \lambda z_i) \otimes H(x_j + \lambda z_j) = 0 \tag{10}$$

where $x_i = [u_i \ v_i \ 1]^T$ and $z_i = [0 \ 0 \ u_i^2 + v_i^2]^T$. The expansion of Eq. 10, based on distorted points directly observed on the images, leads to a quadratic eigenvalue problem (QEP):

$$(D_1 + \lambda D_2 + \lambda^2 D_3)h = 0 \tag{11}$$

where the D_1, D_2 and D_3 matrices are composed from the data, and h is the homography, H vectorized. The solution proposed by Fitzgibbon involves solving a number of times Eq. 11, each one for a pair of images, and then computing the median of the solutions. In addition, Eq. 11 is modified to be written with squared matrices. The matrices in Eq. 11 are made square by multiplying all the terms of the equation by D_1^T:

$$(D_1^T D_1 + \lambda D_1^T D_2 + \lambda^2 D_1^T D_3)h = 0. \tag{12}$$

The solution of the problem described by Eq.12, i.e. λ and h, can be obtained numerically using e.g. the function `polyeig` of Matlab.

3.4 Division distortion model, Steele and Jaynes' Calibration Method

Steele and Jaynes [7] improved Fitzgibbon's algorithm by making it lesser biased by the noise existing in the correspondences found among images. Their method takes advantage of an QEP algorithm working directly on rectangular matrices, and thus eliminates the need to enforce square matrices as Fitzgibbon proposed. In their method they add a new variable to Eq. 11:

$$D_1 h + \lambda(D_2 h + D_3 u) = 0 \tag{13}$$
$$u - \lambda h = 0 \tag{14}$$

Solving the equation for h and u one gets $(A - \lambda B)v = 0$ with:

$$A = \begin{bmatrix} D_1 & 0 \\ 0 & I \end{bmatrix}; \ B = \begin{bmatrix} -D_2 & -D_3 \\ I & 0 \end{bmatrix}; \ v = \begin{bmatrix} h \\ u \end{bmatrix} \tag{15}$$

The solution to this problem is obtained iteratively. The algorithm starts with $\lambda = 0$ and then v is computed through single value decomposition. The update to λ is done by solving the quadratic equation

$$v^T (B^T + \lambda A^T)(A - \lambda B)v = 0 \tag{16}$$

and choosing the positive root[1].

---

[1] For the quadratic equation $ax^2 + bx + c = 0$ the positive root is given by $(-b + \sqrt{b^2 - 4ac})/(2a)$

3.5 Comments

The pioneer work by Agapito et al. already focuses on the geometric auto-calibration of pan-tilt-zoom cameras. However, it aims solely to estimate efficiently the intrinsic parameters, not including radial distortion. Sinha and Pollefeys's approach to auto-calibrate these cameras, addresses estimation of both intrinsic and radial distortion parameters, but relies on multiple bundle adjustment processes which make the process computationally complex (time). The methods of Fitzgibbon, and Steele and Jaynes, allow a simple computation of both radial distortion and homographies, but do not provide calibration for multiple zoom levels.

Table 1 summarizes the features of the calibration methodologies presented in this section. One observes that no single methodology has all the presented three features, namely estimating radial distortion, allowing multiple zoom levels, and having at the same time low computational demands (more precisely, low computational time).

Calibration methodology	Multiple zoom levels	Radial distortion	Low Comp. resources
Agapito et al. [10]	yes		yes
Sinha and Pollefeys [5]	yes	yes	
Fitzgibbon [6]		yes	yes
Steele and Jaynes [7]		yes	yes

Table 1. Comparing calibration methodologies

4   Calibration considering both Radial Distortion and Multiple Zoom Levels

In this section we propose an auto-calibration methodology that estimates radial distortion, allows multiple zoom levels, and can be implemented efficiently (processing time). This methodology builds on the methodologies available in the literature that were described in the previous section. Before presenting the complete calibration algorithm, we will firstly introduce the calibration procedure for one zoom level given the calibration of the preceding zoom level.

Considering that the radial distortion varies with the zoom level, one has to modify Eq. 10 to represent both distortion levels, λ_1 and λ_2:

$$(x_i + \lambda_1 z_i) \otimes H(x_j + \lambda_2 z_j) = 0 \tag{17}$$

Expanding Eq. 17 one obtains:

$$x_i \otimes Hx_j + \lambda_1 z_i \otimes Hx_j + \lambda_2(x_i \otimes Hz_j + \lambda_1 z_i \otimes Hz_j) = 0 \tag{18}$$

here factorized for λ_2. The equation can be further expanded to factorize the homography as a column vector, $h = [h_{11} \ h_{21} \ \cdots h_{33}]^T$, pre-multiplied by matrices D_1 and

D_2:

$$(D_1 + \lambda_2 D_2)h = 0 \qquad (19)$$

where

$$D_1 = \begin{bmatrix} 0 & 0 & 0 & -x' - \lambda_1 r'^2 x & -y' - \lambda_1 r'^2 y & -1 - \lambda_1 r'^2 & yx' & yy' & y \\ x' + \lambda_1 r'^2 x & y' + \lambda_1 r'^2 y' & 1 + \lambda_1 r'^2 & 0 & 0 & 0 & -xx' & -xy' & -x \end{bmatrix} \qquad (20)$$

and

$$D_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & -r^2 - \lambda_1 r^2 r'^2 & 0 & 0 & y'r^2 \\ 0 & 0 & \lambda_1 r^2 r'^2 + r^2 & 0 & 0 & 0 & 0 & 0 & -x'r^2 \end{bmatrix}. \qquad (21)$$

Note that Eq. 19 can be once more solved using the `polyeig` function of Matlab, or using the iterative method proposed by Steele and Jaynes (see equations 15 and 16).

Now that one has a methodology to calibrate a zoom level given the calibration of another zoom level, it is possible to design a complete calibration algorithm. Our proposal of a complete calibration methodology encompasses the following main steps:

1. Calibration at minimum zoom using the methodology of Steele and Jaynes (section 3.4)
2. Calibration of one zoom level after another using the methodology introduced in this section (Eq. 19).

Doing calibration at minimum zoom (i.e. maximum *zoom out*) involves moving (panning and tilting) the camera with a fixed zoom, detecting and matching features. In our case we use SIFT features [11]. The calibration for various zoom levels involves fixating the pan and tilt angles while increasing the zoom level. Between each pair of zoom levels one has to find once more corresponding SIFT features.

5 Experiments and Results

This section describes two experiments: (i) comparing the performance of Fitzgibbon's calibration algorithm against the Steele and Jaynes' in a pan-tilt sequence created from a synthetic scenario, and comparing the performance of our zoom algorithm combined with both algorithms; (ii) qualitatively assessment of the performance of our methodology on real data, namely mosaics built from pan-and-tilt and from zoom sequences.

5.1 Noise Analysis

This experiment is based on synthetic data, generated from random $3D$ points imaged by a virtual pan-tilt-zoom camera with radial distortion. White Gaussian noise is added to the position of the features observed by the camera. The noise varies from zero pixels of standard deviation up to two pixels.

The first part of the experiment involves comparing the performance of Fitzgibbon's and Steele and Jaynes' algorithms in estimating the radial distortion from a set of images with constant zoom. The Fitzgibbon's algorithm is tested with a set of 50 pair of images and the Steele and Jaynes' algorithm is tested with just 6 pairs of images. Both

algorithms performed well with no noise, as they both found the correct radial distortion parameter $\lambda = 0.1278$ (see fig. 1, green line). However as the noise increases the Fitzgibbon algorithm (left plot, A) tends to go away from the real value of the parameter. On the other hand Steele and Jaynes' algorithm (left plot, B) stays close to the real solution even in a presence of a considerable noise. Despite using a lesser number of image pairs, Steele and Janeys' algorithm performs better.
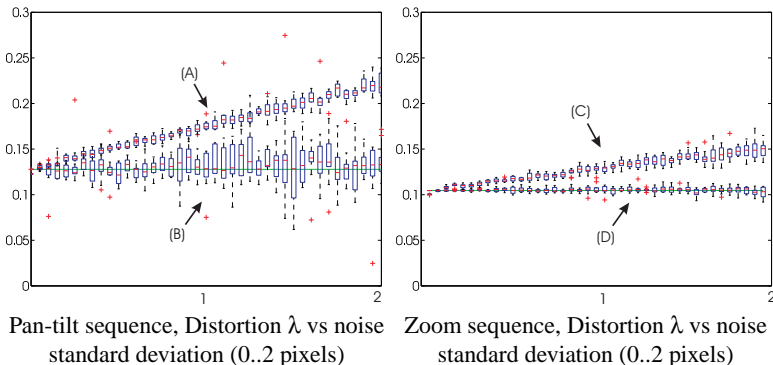

Pan-tilt sequence, Distortion λ vs noise standard deviation (0..2 pixels) Zoom sequence, Distortion λ vs noise standard deviation (0..2 pixels)

Fig. 1. Estimated distortion vs noise in the data. Comparing two estimation methods in a pan and tilt sequence (left plot) and in a zoom sequence (right plot). The green line represents the true distortion value in both plots. Noise is white Gaussian noise varying from 0 to 2 pixels (standard deviation).
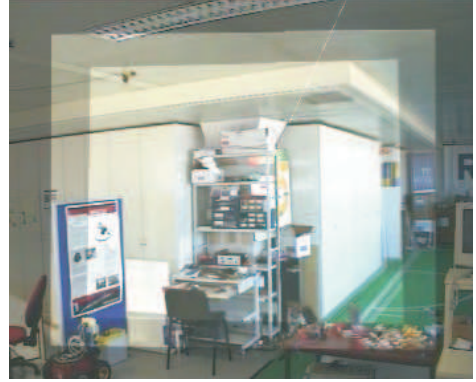
In the second part of the first experiment we assess our methodology for calibrating the radial distortion given images taken at different zoom levels. More precisely, this experiment involves generating random $3D$ points and capturing two images with different zooms. One of the two images has known radial distortion. The imaged points, matched between the two images, are corrupted with white Gaussian noise with 0 to 2 pixels pixels of standard deviation. Figure 1(right plot) compares using `polyeig` once (C) with using `SVD` iteratively (D), as proposed by Steele and Jaynes' for the constant zoom case (the horizontal green line denotes the ground truth value). One observes that the iterative algorithm is beneficial for estimating the radial distortion in the case of multiple zoom levels, as it was in the single zoom level.

5.2 Mosaics

The second experiment involves testing the proposed calibration methodologies with real data acquired with an Axis 215 PTZ camera. The calibration of the intrinsic parameters and the radial distortion is performed using a set of 16 images, captured at different orientations, with the same zoom level. First the radial distortion is estimated using Steel and Jaynes' algorithm. After correcting the images the algorithm of Agapito et al. is used to estimate the intrinsic parameters, K^z. Figure 2 (a) shows that both algorithms work well together since the texture is qualitatively consistent at stitching image seams.

(a) Pan and tilt mosaic   (b) Zooming on mosaic

Fig. 2. (a) Mosaic build from sequence of 16 pan and tilt images. (b) Mosaic from a sequence 4 images with different levels of zoom.

The experiment on real data continues with a sequence of images with varying zoom. Four images were acquired with different zoom levels. The radial distortion parameters are known (estimated in the previews step) just for the first image. The parameters of the other images are estimated using our methodology. The result of the radial distortion correction and pasting of the images is shown on fig. 2 (b).

6   Conclusions and Future Work

In this paper we approach the problem of auto-calibrating pan-tilt-zoom cameras. The auto-calibration comprises estimating intrinsic and radial distortion parameters over the camera's full zoom range.

Fitzgibbon [6] proposed a method for calibrating the radial distortion without knowing anything besides matching features between images. Steele and Jaynes [7] improved the method by making it more robust. In both works the zoom level was considered to be fixed, and thus the radial distortion was constant for all images acquired by the camera. In our methodology we show that auto-calibration can be generalized to the case of multiple zoom levels while still effectively including the estimation of radial distortion. We propose solving the problem of different zoom levels by estimating the radial distortion at one zoom level given the radial distortion of another level.

Future work will focus on investigating even more complex modelings of pan-tilt-zoom cameras, while keeping computationally time-efficient solutions. Recently work by Kukelova and Pajdla [8] is already pointing towards this direction.

Acknowledgments

References

1. Bouguet, J.Y.: Camera calibration toolbox for matlab. `http://www.vision.caltech.edu/bouguetj` (2008)
2. Hartley, R.I.: Self-calibration of stationary cameras. International Journal of Computer Science **22** (1997) 5–23
3. Agapito, L.D., Hayman, E., Reid, I.: Self-calibration of a rotating camera with varying intrinsic parameters. In: Proc 9th British Machine Vision Conf, Southampton. (1998) 105–114
4. Agapito, L.D., Hayman, E., Reid, I.: Self-calibration of rotating and zooming cameras. Int. J. Comput. Vision **52** (2001) 107–127
5. Sinha, S.N., Pollefeys, M.: Pan-tilt-zoom camera calibration and high-resolution mosaic generation. Comput. Vis. Image Underst. **103** (2006) 170–183
6. Fitzgibbon, A.W.: Simultaneous linear estimation of multiple view geometry and lens distortion (2001)
7. Steele, R.M., Jaynes, C.: Overconstrained linear estimation of radial distortion and multi-view geometry. In: Computer Vision ECCV 2006, Springer (2006)
8. Kukelova, Z., Pajdla, T.: A minimal solution to radial distortion autocalibration. IEEE Transactions on Pattern Analysis and Machine Intelligence **33** (2011)
9. Heikkila, J., Silven, O.: A four-step camera calibration procedure with implicit image correction. In: 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'97). (1997)
10. Agapito, L., Hartley, R., Hayman, E.: Linear calibration of a rotating and zooming camera. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. (1999) 15–21
11. Lowe, D.: Object recognition from local scale-invariant features. In: Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on. Volume 2. (1999) 1150–1157 vol.2