PRACTICAL METHOD FOR AUTO-CALIBRATION OF ZOOMING CAMERAS FROM THE VIEWS CONTAINING A RECTANGLE OF UNKNOWN SIZE

JaeChul Kim, BonKi Koo, ChangWoo Chu, and ByeongTae Choi

Electronics and Telecommunications Research Institute, 161 Gajeong-dong, Yuseong-gu, Daejon, Korea
{jaechul, bkkoo, cwchu, btchoi}@etri.re.kr

ABSTRACT

In this paper, we propose a practical auto-calibration algorithm of zooming cameras from the multiple images containing a rectangle of unknown size. The proposed algorithm need not solve the complex multivariate non-linear problem, since the equations for calibration are formulated as the polynomials in a single variable. In addition, the proposed algorithm is able to provide reliable results under the significant change of zooming where most multi-view/mono-plane (auto-) calibration algorithms often fail to produce reliable results. Experimental results validate the effectiveness of the proposed algorithm. Applications to image-based modeling are demonstrated as well.

1. INTRODUCTION

Camera auto-calibration is the process of computing the intrinsic parameters of camera directly from multiple images without using any *a priori* Euclidean structure and has been widely investigated in computer vision field during last decade. To calibrate cameras directly from images, most auto-calibration techniques employ some kinds of constraints. For example, some knowledge on the intrinsic parameters (e.g. skew zero, unit aspect ratio etc.) can be used [1], or specific camera motions can be applied [2]. Constraints for the observed scene such as metric information or coplanarity can be also utilized [3,4,5,6].

In this paper, we propose a practical method for auto-calibration of zooming cameras by exploiting the scene constraints obtained from the multiple images containing a rectangle of unknown size. The motive for considering rectangles of unknown size for calibration is quite simple: rectangles of unknown size are abundant in practice. Specifically, rectangles appear very often especially in man-made environments, but the sizes of rectangles in the scene cannot be readily inferred from images due to perspective distortion.

To calibrate cameras from a rectangle of unknown size, we first parameterize a world-to-image homography of the unknown rectangle by the length ratio of the two sides of the rectangle, and then apply this parameterized homography to the equations which relate the homography to IAC (Image of Absolute Conic). This approach is inspired from the existing algorithms. For example, parameterizing a homography from an arbitrary rectangle is inspired from [7] which used the metric information on planes as the constraints for a homography. Also we should admit that our equations for calibration are derived, based on the well-known equations in plane-based camera calibration [8,9]. Nevertheless, in this paper, there are some improvements over the existing methods especially in practical aspects.

First, computation for calibration is very simple and robust to noise. In this paper, the equations for calibration are derived as the polynomials in a single variable. Hence, we can get the closed-form solutions, avoiding the initialization difficulty from which most auto-calibration algorithms formulated as multivariate non-linear problems suffer.

Second, the proposed method is able to provide reliable calibration results under the significant change of zooming. Most multi-view/mono-plane based calibration algorithms assume that the principal point is fixed while zooming since it is in general impossible to get a unique solution from a single plane when both the focal and the principal point vary over all images. But, the principal point of zoom-lens camera can be varied during zooming [10]. Furthermore, we have empirically confirmed that fixing of the principal point affects the calibration results more adversely as the change of zooming become considerable. In order to resolve this problem, we propose a technique to approximate the solution of the under-constrained equations without fixing the principal point. Experimental results show that the proposed technique provides more reliable results than the previous ones especially under the significant change of zooming.

In section 2, our camera model is introduced. In section 3, the homography for an unknown rectangle is derived. The calibration algorithm is proposed in section 4, and experimental results are presented in section 5.

2. CAMERA MODEL

For modeling cameras, we use the perspective projection. The projection from a 3D point $\mathbf{x} \in \mathrm{P}^3$ to its 2D image point $\mathbf{p} \in \mathrm{P}^2$ is represented as follows:

$$\mathbf{p} \propto \mathbf{K}[\mathbf{R} \quad \mathbf{t}]\mathbf{x}, \qquad \mathbf{K} = \begin{bmatrix} \alpha f & s(=0) & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{1}$$

where $\mathbf{R}$ is 3×3 orthogonal matrix representing camera's orientation, $\mathbf{t}$ is a 3-vector representing its position, and $\mathbf{K}$ is a 3×3 matrix containing the camera intrinsic parameters. In $\mathbf{K}$, f is an effective focal length, α is an aspect ratio, s is a skew factor, and (u_0, v_0) is the coordinate of the principal point. In this paper, we will suppose that the skew factor is zero, since it is in general very close to zero.

3. HOMOGRAPHY FOR A RECTANGLE OF UNKNOWN SIZE

In this section, we derive the world-to-image homography for the rectangle whose length ratio between two sides is τ. For convenience, let us assume that the X-axis and the Y-axis of the world reference coordinate system are respectively aligned with two sides of the rectangle. Then, image projections $p_{i=1,2,3,4} = (u_i, v_i)$ of the rectangle's four vertices satisfy:

$$\begin{pmatrix} \lambda_1 u_1 & \lambda_2 u_2 & \lambda_3 u_3 & \lambda_4 u_4 \\ \lambda_1 v_1 & \lambda_2 v_2 & \lambda_3 v_3 & \lambda_4 v_4 \\ \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 \end{pmatrix} = \widetilde{\mathbf{H}} \cdot \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \tag{2}$$

where $\widetilde{\mathbf{H}}$ is a 3×3 homography for the canonical square (size 1), and it can be decomposed as

$$\widetilde{\mathbf{H}} = \mathbf{H} \cdot \mathbf{S}. \tag{3}$$

In (3), $\mathbf{H}$ is a homography for the rectangle with the length ratio τ, and $\mathbf{S}$ is 3×3 matrix which scales the Y-axis by τ. Here, $\mathbf{S}$ is written as

$$\mathbf{S} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \tau & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{4}$$

If we measure the image projections of the rectangle's four vertices, we can compute $\widetilde{\mathbf{H}}$ from (2), and subsequently obtain $\mathbf{H}$ since $\mathbf{H} = \widetilde{\mathbf{H}} \cdot \mathbf{S}^{-1}$. Via simple algebra, we can finally express $\mathbf{H}$ as follows.

$$\mathbf{H} = \begin{bmatrix} \tilde{\mathbf{h}}_1 & \tau^{-1}\tilde{\mathbf{h}}_2 & \tilde{\mathbf{h}}_3 \end{bmatrix}, \tag{5}$$

where $\tilde{\mathbf{h}}_{i=1,2,3}$ denote the 1st, 2nd, and 3rd column of $\widetilde{\mathbf{H}}$ respectively.

4. CALIBRATION ALGORITHM

In this section, the proposed calibration algorithm is presented. We begin this section by describing the basic equations for calibration.

4.1. Basic equations

A world-to-image homography $\mathbf{H}$ has a homogeneous linear relationship with IAC ($\boldsymbol{\omega} \propto \mathbf{K}^{-T}\mathbf{K}^{-1}$) as follows [8,9]:

$$\mathbf{h}_1{}^T \boldsymbol{\omega} \mathbf{h}_2 = 0, \qquad \mathbf{h}_1{}^T \boldsymbol{\omega} \mathbf{h}_1 = \mathbf{h}_2{}^T \boldsymbol{\omega} \mathbf{h}_2, \tag{6}$$

where $\mathbf{h}_1$ and $\mathbf{h}_2$ denote the 1st and 2nd column of $\mathbf{H}$. Here, incorporating equation (5) into equation (6), we can get the basic equations for calibration:

$$\tilde{\mathbf{h}}_1{}^T \boldsymbol{\omega} \tilde{\mathbf{h}}_2 = 0, \qquad \tau^2 \tilde{\mathbf{h}}_1{}^T \boldsymbol{\omega} \tilde{\mathbf{h}}_1 = \tilde{\mathbf{h}}_2{}^T \boldsymbol{\omega} \tilde{\mathbf{h}}_2. \tag{7}$$

4.2. Calibration of a camera with fixed parameters

When all the parameters are fixed, the number of unknowns is five: four camera parameters and one length ratio of the rectangle. Since two constraints are obtained from a single image (see equation (7)), minimal three images are required to get a closed-form solution. In this minimal case, we can calibrate the camera and recover the length ratio by solving simultaneously two quadratic polynomial equations in a single variable. For want of space, the detailed procedure of deriving the equations is left out, but similar procedure will be presented in detail in the next subsection.

If more than three images are available, we can recover the camera parameters by computing the linear equation system since one linear equation for each image is obtained from the first equation in (7). The length ratio is subsequently computed by applying the recovered parameters to the second equation in (7).

4.3. Calibration of cameras with varying focal lengths

When only the focal length varies, the number of unknowns to be solved, for n input images, is n+4: n focal lengths, principal point (two), one aspect ratio, and one length ratio of the rectangle. Thus, minimal four images ($2n \geq n+4$) are required to get a closed-form solution. In this minimal case, we can solve the problem by computing an 8-order polynomial equation in a single variable. The detailed procedure of deriving the equation is as follows.

A. In the case that only the focal length varies, each IAC for four images is written as: $\boldsymbol{\omega}_{i=1,2,3,4} \propto \begin{pmatrix} \omega_1 & 0 & \omega_2 \\ 0 & \omega_3 & \omega_4 \\ \omega_2 & \omega_4 & \omega_{5i} \end{pmatrix}$.

B. Using the first equation in (7) obtained from each input image, we can express ω_{5i} (i=1,2,3,4) as the linear combinations of ω_1 , ω_2 , ω_3 , and ω_4 , and then incorporating these results into the second equations in (7), we can obtain the four equations with the form as $(a_i x + b_i)\omega_1 + (c_i x + d_i)\omega_2 + (e_i x + f_i)\omega_3 + (g_i x + h_i)\omega_4 = 0$. Here, a_i , b_i , c_i , d_i , e_i , f_i , g_i , and h_i (i =1,2,3,4) are scalar values, and x denotes τ^2 .

C. Four equations derived in step B are reduced to an 8-order polynomial equation in a single variable x. The roots of the polynomial are obtained by finding the eigenvalues of its companion matrix [11].

D. Incorporating the obtained x into the basic equations in (7), eight linear equations are provided, and thus we can compute $\boldsymbol{\omega}$ by solving the linear equation system.

If we use five (resp. six) images, we can obtain the results by solving simultaneously two (resp. three) 8-order polynomial equations in a single variable. In addition, in case of using more than six images, we can compute the camera parameters linearly by over-parameterization.

4.4. Calibration of cameras with varying focal lengths and varying principal points

In the case that both the focal length and the principal point undergo a change over all images, the number of unknowns to be solved, for n input images, is $3n$+2: n focal lengths, 2 n principal points, one aspect ratio, and one length ratio of a rectangle. This means that the number of unknowns is always more than the number of constraints ($3n$+2>2n). Therefore, the problem to be solved is under-constrained, that is, we cannot get a unique solution. To resolve this under-constrained condition, we propose a technique to approximate the camera parameters without fixing the principal point, and this approximation is made by applying the center line constraint proposed in [12]. Emphasizing once again, the obtained solution are not a closed-form, but just an approximation. However, we have been able to experimentally confirm that the proposed technique provides more reliable results than the existing methods which fix the principal point, especially when the considerable change of zooming occurs.

The center line constraint. Gurdjos *et al*. exhibited that for a given world-to-image homography, the principal point restricts its location to a line segment, and they referred to this line segment as the center line [12]. Letting $\mathbf{H}_{ij}$ be the *(i,j)* element of the world-to-image homography **H**, the center line $\mathbf{l}_c$ can be written as:

$$\mathbf{l}_c = \begin{bmatrix} l_1 & l_2 & l_3 \end{bmatrix} = \begin{bmatrix} -\varphi_1\alpha^2, & -\varphi_2, & \varphi_3\alpha^2 + \varphi_4 \end{bmatrix}^T ,$$

$$\varphi_1 = (\mathbf{H}_{31}^2 + \mathbf{H}_{32}^2)(\mathbf{H}_{31}\mathbf{H}_{12} - \mathbf{H}_{11}\mathbf{H}_{32}),$$

$$\begin{aligned} \varphi_2 &= (\mathbf{H}_{31}^2 + \mathbf{H}_{32}^2)(\mathbf{H}_{31}\mathbf{H}_{22} - \mathbf{H}_{21}\mathbf{H}_{32}), \\ \varphi_3 &= (\mathbf{H}_{31}\mathbf{H}_{12} - \mathbf{H}_{11}\mathbf{H}_{32})(\mathbf{H}_{11}\mathbf{H}_{31} + \mathbf{H}_{12}\mathbf{H}_{32}), \\ \varphi_4 &= (\mathbf{H}_{31}\mathbf{H}_{22} - \mathbf{H}_{21}\mathbf{H}_{32})(\mathbf{H}_{21}\mathbf{H}_{31} + \mathbf{H}_{22}\mathbf{H}_{32}), \end{aligned} \quad (8)$$

Camera calibration algorithm. To approximate camera parameters, we make the following two assumptions: 1) the aspect ratio is one. 2) the principal point locates near the center of image. From the first assumption, we can write the coefficients l_1, l_2, and l_3 as the form of $a_i x^3 + b_i x$ (i=1,2,3) by incorporating equation (5) into equation (8), where a_i , b_i (i=1,2,3) are scalar values, and x denotes τ^{-1} . Using this result together with the second assumption, we can obtain the following inequality:

$$\frac{\left|l_1 c_1 + l_2 c_2 + l_3\right|}{\sqrt{l_1^2 + l_2^2}} \le t \Leftrightarrow px^4 + qx^2 + r \le 0, \quad (9)$$

where (c_1, c_2) is the center of image, t is the threshold distance between the center of image and the center line, and p,q, and r are scalar values. Equation (9) reveals that we can estimate the rectangles to provide the center lines whose distances from the center of image are less than a threshold t by computing the doubly quadratic inequality in a single variable $x(=\tau^{-1})$. Staring with this inequality, we calibrate the cameras and estimate the length ratio of the rectangle. The detailed calibration algorithm is presented in the following.

A. For n input images, we derive n doubly quadratic inequalities from (9), and then we compute the intervals of $x(=\tau^{-1})$ satisfying the system of these inequalities. The obtained interval will be a single one, since the interval is the solution of the system of doubly quadratic inequalities.

B. Over the obtained interval $x_0 \le x \le x_1$, we compute the $x_{\min}$ which minimizes the cost function:

$$\sum_{i=1}^{n} \frac{\left|l_{1i} c_1 + l_{2i} c_2 + l_{3i}\right|}{\sqrt{l_{1i}^2 + l_{2i}^2}} (i=1,\ldots,n).$$

Note that the minimization is simple and robust to noise since the cost function has only a single variable as well as the reliable initial guess ($x_0 \le x \le x_1$) is given.

C. For the $x_{\min}$, we determine the center line for each image from equation (8), and then we approximate the principal point of each image by picking the closest point to the center of image on each center line.

D. Since the principal points are given over all images, the focal lengths can be linearly acquired from equation (7).

5. EXPERIMENTS

5.1. Auto-calibration of a fixed camera

In this experiment we took three images of a calibration target using a SONY digital still camera F505 (7.1mm~35.5mm zoom-lens) without zooming. For a comparison, the camera was calibrated with the standard Faugeras-Toscani method [13]. The standard method utilized all the grid points of the calibration target, while the proposed method used only the four corner points on the target's top plane. Table 1. shows the results for the focal length, giving a comparison with those of standard method. In addition, the recovered length ratio of the rectangle was 1.002 (true value is 1).

Focal length	Standard	Proposed	Error
f_1	1601.8	1630.6	1.8%
f_2	1593.1	1626.6	2.1%
f_3	1595.3	1555.4	2.5%

Table 1. Results of focal lengths without zooming

We can find out that the obtained results are accurate since the maximal error rate of the focal length is 2.5% with respect to the results of the standard calibration.

5.2. Auto-calibration of a camera undergoing the significant change of zooming

In this experiment we took four images of a calibration target at the largely different zooming positions. Figure 1. shows four images to be taken. To have ground-truth values, the camera was calibrated with the standard Faugeras-Toscani method. In addition, for a comparison with the existing method, we calibrated the camera using the Sturm-Maybank plane-based method [8] where the principal point was fixed. Sturm-Maybank method used the four corner points on the target's top plane as ours, but the coordinates of the corner points were given *a priori*. The obtained results are summarized in Table 2.

Compared with the ground-truth values, the proposed technique provided the trustworthy results, and the 3D reconstruction of the grid points of the target can be done with sufficient accuracy. In contrast, the existing method produced the unreliable results in some cases, even giving the complex values for the focal lengths in the second and third images.

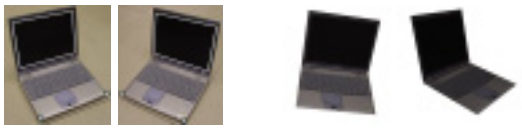
5.3. Application to the image-based modeling

Two images of a notebook computer were taken by the same camera. For a calibration, we used four corner points of the LCD panel. By reconstructing the 3D coordinates of six point matches, we were able to obtain the partial 3D model of the notebook computer. The two input images and the reconstructed 3D model are shown in Figure 2.

Figure 1: 4 Images of a calibration target

Focal length	Faugeras-Toscani	Sturm-Maybank	Proposed method
f_1	1345.1	1106.1	1413.7
f_2	1700.5	515.4i(complex)	1780.5
f_3	1911.4	630.5i(complex)	1865.1
f_4	2213.8	1625.1	2312.6

Table 2. Results of focal lengths of a camera undergoing a significant change of zooming



(a) (b)

Figure 2. 3D reconstruction of the notebook computer (a) two input images (b) the reconstructed model

7. REFERENCES

[1] M. Polleyfeys, R. Koch, and L. VanGool, "Self-calibration and metric reconstruction in spite of varying and unknown intrinsic parameters," *IJCV*, 32(1), pp. 7-25, August 1999.

[2] R. Hartley, Self-calibration from multiple views with a rotating camera, In proc. *ECCV*, pp. 471-478, 1994.

[3] D. Liebowitz and A. Zisserman, "Combining scene and auto-calibration constraints," In proc. *ICCV*, pp. 293-300, 1999

[4] E. Malis, and R. Cipolla, "Multi-view constraints between collineations: Application to self-calibration from unknown planar structures," In proc. *ECCV*, vol. 2, pp. 610-624, 2000.

[5] P. Gurdjos and P. Sturm, "Methods and geometry for plane-based self-calibration," In proc. *CVPR*, pp. 491-496, 2003.

[6] B. Triggs, "Autocalibration from planar scenes," In proc. *ECCV*, pp. 89-105, 1998.

[7] D. Liebowitz and A. Zisserman, "Metric rectification for perspective images of planes," In proc, *CVPR*, pp. 482-488, 1998.

[8] P. Sturm and S. Maybank, "On plane-based camera calibration: A general algorithm, singularities, applications," In proc. *CVPR*, pp. 432-437, 1999.

[9] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," In proc. *ICCV*, pp.666-673, 1999.

[10] M. Li and J. Lavest, "Some aspects of zoom lens camera calibration," *PAMI*, 18(11), pp. 1105-1110, November 1996.

[11] J. Goldberg, "*Matrix theory with applications*," McGraw-Hill, Inc. pp. 276-280. 1991.

[12] P. Gurdjos, A. Crouzil, and R. Payrissat, "Another way of looking at plane-based calibration: the center circle constraint," In proc. *ECCV*, vol. 4, pp. 252-266, 2002.

[13] O. Faugeras and G. Toscani, "Camera calibration for 3D computer vision," *Int. Workshop on Machine Vision and Machine Intelligence*, pp. 240-247, 1987.