Conic Based Camera Re-calibration after Zooming

Iuri Frosio[1], Cristina Turrini[2], and Alberto Alzati[2]

[1] Computer Science Dept., University of Milan, Italy
[2] Mathematics Dept., University of Milan, Italy
{iuri.frosio,cristina.turrini,alberto.alzati}@unimi.it

Abstract. We describe here a method to compute the internal parameters of a camera whose position and orientation are known. The method is based on the observation of at least three conics on a known plane; these can be easily extracted in a real scenario from a tiled floor or other regular structures. The method estimates the principal point and focal length using a unique image of the conics when these are observed by an additional calibrated camera. Differently from other methods, no assumption is made on the conics used for calibration. The experimental results demonstrate that the accuracy of the method is comparable to that of more traditional (and time consuming) approaches. It can find applications in systems of Pan-Zoom-Tilt (PZT) or traditional cameras, that are nowadays widely employed, for instance in the surveillance domain, and require frequent re-calibration.

Keywords: camera calibration, conics, computer vision, surveillance.

1 Introduction

Camera calibration is fundamental in computer vision to recover 3D information from the scene. Several approaches to calibration have been proposed, involving 3D objects of known shape, planar (2D) patterns and 1D objects [1]. Calibration generally involves the following steps: the calibration object is moved in the scene while the camera takes images; features are extracted from the images of the calibration object and the camera internal and external parameters are estimated from the positions of the features in the acquired images.

Lines and points have been widely employed to compose the calibration pattern because of the simplicity to identify them in the images [1, 2]. The Zhang's method [2, 3], one the most diffused approach, uses a checkerboard pattern and it is based on the homographies relating the positions of the checkerboard corners in 3D space and in the acquired images. On the other hand, several authors noticed that conics convey more information and consequently produce a more accurate estimate of the camera parameters [4–8]. When conics are used for calibration, the common idea is that a planar conic in 3D space, described by a 3×3 matrix $\mathbf{C}$, is projected into a conic with matrix $\mathbf{H}^T\mathbf{C}\mathbf{H}$ onto the camera image plane, where $\mathbf{H}$ is the homography that relates the conic and the image

A. Petrosino (Ed.): ICIAP 2013, Part I, LNCS 8156, pp. 361–370, 2013.

plane [4–6]. Putting in relation two ore more confocal or co-axial conics, or conics in known position, **H** is estimated and camera parameters are derived from its decomposition [1].

Because of the noise on the measured positions of the features, lots of images are necessary to get a reliable estimate of the camera parameters [1]. On the other hand, when systems composed by more than one camera are used (e.g. like in many surveillance systems, where frequent camera re-calibration is necessary) or only partial camera calibration is needed (e.g. after zooming), less parameters have to be estimated and/or information from other cameras can be profitably used to obtain a reliable calibration from few or even just one image[4, 9]. In this context, we describe a method to compute the internal parameters of a camera after zooming. The algorithm is based on the observation of at least three degenerate conics (i.e. pair of lines) from the camera itself and from a second, calibrated camera, but it can be used without any modification for any kind of conic. Differently from other approaches, our algorithm minimizes a cost function which involves directly the parameters of the observed conics. The experimental results demonstrate that only one image is sufficient to re-calibrate the camera with an accuracy similar to that of Zhang's method [2, 3], which requires a larger number of images and therefore more time to be performed.

2 Method

2.1 Preliminaries and Aim of the Method

The present method is aimed at estimating the internal parameters (principal point, focal length) of a camera after zooming, under the following hypothesis:

- two cameras observe at least three conics lying on a known plane (e.g. six lines extracted from a tiled floor can be used as a triplet of degenerate conics with no lines in common each other); the equation of the floor plane is supposed to be known, as this can be estimated from a set of features lying on the floor before camera zooming;
- the external parameters (camera orientation and position of the optical center) of the camera to be calibrated are known (e.g. from a previous calibration performed at a different zoom level, supposing that the camera position and orientation do not change after zooming);
- the internal (focal length, principal point) and external (position, orientation) parameters of the second camera are known.

No knowledge on the observed conics is necessary: they could be ellipses, hyperbolas or degenerate conics whose parameters are in any case unknown and unnecessary to the method described here.

2.2 Camera Model

For the i-th camera of the system, the projection of a 3D point $[X/U\ Y/U\ Z/U]^T$ onto the camera image plane in homogeneous coordinates is given by [10]:

$$\left[u\ v\ w\right]^T = \mathbf{K}_i \left[\mathbf{R}_i^T\ -\mathbf{R}_i^T\mathbf{t}_i\right] \left[X\ Y\ Z\ U\right]^T \tag{1}$$

$$\mathbf{K}_i = \begin{bmatrix} f_i & 0 & c_{x,i} \\ 0 & f_i & c_{y,i} \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

where ^{T} is the transpose operator, $[uvw]^T$ are the homogeneous coordinates of the point in the image plane, f_i is the camera focal length, $[c_{x,i}\ c_{y,i}]^T$ is the camera principal point, $\mathbf{R}_i$ is a 3×3 rotation matrix describing the camera orientation and $\mathbf{t}_i$ is a 3×1 vector describing the position of the camera optical center in 3D space.

We will assume that no distortion is present on the images acquired by the cameras; this is justified with the fact that modern apparatuses generally included a lens model in their firmware: distortions are consequently automatically corrected [11, 12] and simplified distortion models can be adopted [13, 14].

2.3 Relation between Conics Observed by Two Cameras

Let us consider two cameras and a set of conics lying on the floor plane and observed by both cameras. The projection of a conic is again a conic; for the i-th camera, the j-th conic equation in homogeneous coordinates is given by:

$$\begin{bmatrix} u & v & w \end{bmatrix} \mathbf{C}_{i,j} \begin{bmatrix} u & v & w \end{bmatrix}^T = 0 \tag{3}$$

where $\mathbf{C}_{i,j}$ is a 3×3 symmetric matrix with the parameters of the conic. The cone associated to $\mathbf{C}_{i,j}$, whose vertex corresponds to the center of the i-th camera, is:

$$\begin{bmatrix} X & Y & Z & U \end{bmatrix} \mathbf{P}_i^T \mathbf{C}_{i,j} \mathbf{P}_i \begin{bmatrix} X & Y & Z & U \end{bmatrix}^T = 0 \tag{4}$$

where:

$$\mathbf{P}_i = \mathbf{K}_i \begin{bmatrix} \mathbf{R}_i^T & -\mathbf{R}_i^T \mathbf{t}_i \end{bmatrix} \tag{5}$$

is the i-th camera projection matrix. Let us suppose now that camera 2 has its optical center in $\mathbf{t}_2 = [0\ 0\ 0]^T$ and its orientation matrix $\mathbf{R}_2$ is the identity; this does not represent a limit, as it is always possible to change the reference frame such that these conditions are satisfied. Let us also suppose that the principal point of the camera 2 is in $[c_{x,2}\ c_{y,2}]^T = [0\ 0]^T$; this too can be achieved with a change of reference frame, without loosing generality. Under these hypotheses, the relation between the matrix of the conic observed by camera 1 and 2 is given by:

$$\mathbf{D}^T \mathbf{P}_1^T \mathbf{C}_{1,j} \mathbf{P}_1 \mathbf{D} = \lambda_j \mathbf{C}_{2,j} \tag{6}$$

where λ_j is an arbitrary scale factor, whereas:

$$\mathbf{D} = \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & f_2 \\ \hline p_x & p_y & p_z f_2 \end{array}\right] \tag{7}$$

and p_x, p_y and p_z are the parameters of the plane of the conic, given in the form $p_x Z + p_y Y + p_z Z = U$, in the reference of camera 2. In fact, if we eliminate

U from Eq. (4), by using the above equation of the plane, we get the equation of a cone, having vertex in $\mathbf{t}_2$, and projecting the observed conic from $\mathbf{t}_2$. By cutting such cone with the plane $Z = f_2 U$, we get the conic observed by camera 2. Hence, up to a scalar factor, the matrix of this conic is the first member of Eq. (6).

2.4 Estimation of Camera Internal Parameters

We will assume now that camera 1 has zoomed to a new focal length, thus changing both its focal length f_1 and its principal point $[c_{x,1}\ c_{y,1}]^T$ (notice that, because of the lens rotation during zooming, the position of the principal point may significantly change [13–15]). Eq. (6) poses a constraint on the parameters of the conics observed by camera 1 and 2 and it can be consequently used to estimate the internal parameters of camera 1. In fact, for the j-th observed conic, it can be rewritten as:

$$\mathbf{D}^T \left[\mathbf{R}_1^T \ -\mathbf{R}_1^T\mathbf{t}_1\right]^T \mathbf{K}_1^T \mathbf{C}_{1,j} \mathbf{K}_1 \left[\mathbf{R}_1^T \ -\mathbf{R}_1^T\mathbf{t}_1\right] \mathbf{D} = \lambda_j \mathbf{C}_{2,j} \tag{8}$$

where the unknowns f_1, $c_{x,1}$ and $c_{y,1}$ in $\mathbf{K}_1$, appear in the left side, whereas the unknown scale factor, λ_j, appears in the right side. The term $\mathbf{K}_1^T \mathbf{C}_{1,j} \mathbf{K}_1$ generates a quadratic polynomial including terms in f_1^2, $c_{x,1}^2$, $c_{y,1}^2$, $f_1 c_{x,1}$, $f_1 c_{y,1}$, $c_{x,1} c_{y,1}$, f_1, $c_{x,1}$, $c_{y,1}$ plus a scalar term.

Considering now a set of N conics, and properly rearranging the terms in Eq. (8) in a $6N \times (N+9)$ matrix $\mathbf{A}$ and in a vector $\mathbf{b}$, the following non linear relation is derived:

$$\mathbf{A} \left[f_1^2 \ c_{x,1}^2 \ c_{y,1}^2 \ f_1 c_{x,1} \ f_1 c_{y,1} \ c_{x,1} c_{y,1} \ f_1 \ c_{x,1} \ c_{y,1} \ \lambda_1 \ \lambda_2 \ \ldots \ \lambda_N\right]^T = \mathbf{b}; \tag{9}$$

which describes the equivalence (apart from the scale factors $\{\lambda_j\}_{j=1..N}$) of the parameters of the conics observed by camera 2 with those of the conics observed by camera 1, projected onto the conic plane (i.e. the floor) and then projected again onto the image plane of camera 2.

Eq. (9) can be solved in a least squares sense through an iterative optimization algorithm like the Levenberg-Marquardt method, implemented by *lsqnonlin* in Matlab. To this aim it is however necessary to provide an initial guess of the unknowns f_1, $c_{x,1}$, $c_{y,1}$, λ_1, $\ldots$, λ_N. This is obtained solving the linear system $\mathbf{Ax} = \mathbf{b}$ in a least squares sense, without considering the non linear constraints (e.g., $x_1 = x_7^2$ should hold, where x_k is the k-th element of $\mathbf{x}$) that relate the elements of $\mathbf{x}$ [4]. In particular, $\mathbf{x}$ is obtained as:

$$\mathbf{x} = \left(\mathbf{A}^T\mathbf{A}\right)^\dagger \mathbf{A}^T\mathbf{b} \tag{10}$$

where † indicate the Moore-Penrose pseudo inverse, that has been adopted here since the rank of $\mathbf{A}$ can be not maximal when N is small [16]. We have verified experimentally that, at least for N small, the values in $\mathbf{x}$ provided by Eq. (10) are not reliable; on the other hand, a reliable estimate is achieved considering the following ratios between elements of $\mathbf{x}$:

$$\tilde{f}_1 = x_1/x_7; \ \tilde{c}_{x,1} = x_4/x_7; \ \tilde{c}_{y,1} = x_5/x_7. \tag{11}$$

Similarly, initial guesses for $\{\lambda_j\}_{j=1..N}$ are obtained. Eq. (9) is finally solved in a least squares sense, starting from this initial guess, through the *lsqnonlin* Matlab function.

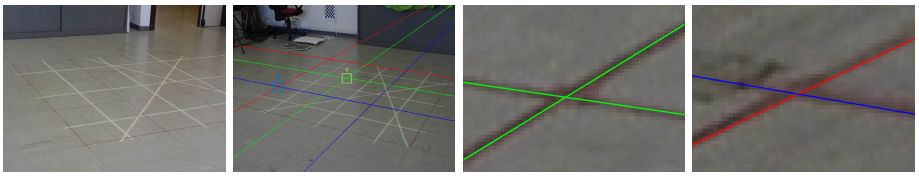
3 Experimental Setup and Results

3.1 Experimental Setup and Line Fitting

The method has been tested using two reflex digital cameras, a Nikon D3100 equipped with a Nikkor 18-55mm zoom lens, acquiring images of 4608×3072 pixels, and a Pentax K-r equipped with a Pentax SMC 18-55mm zoom lens, acquiring images of 4288×2848 pixels. The cameras were positioned on two tripods, at a height of approximately $1.5m$ and at $3.5m$ of distance one from the other. The cameras were tilted about 20° low; their optical axes were convergent with an angle of approximately 70°. Each camera looked at the floor, where squared tiles with a side of $0.6m$ were present; each camera could observe at the maximum focal length ($55mm$) an area of 3×3 tiles. The lines among the tiles were marked through a red pencil to highlight their visibility; moreover, six lines of white adhesive tape, approximately parallel to the red lines, were added to increase the number of available lines (see Fig. 1). Since each pair of lines represent a degenerate conic, a total of 7 degenerate conics with no lines in common were available to test the method (the tilted lines of white adhesive tape in Fig. 1 were not considered here).

We calibrated the cameras at different focal lengths using the Camera Calibration Toolbox for Matlab by J.-V. Bouguet [3], mainly inspired to the Zhang calibration procedure based on observation of a checkerboard planar pattern [2]. For each calibration, 30 images of the checkerboard were considered. The focal lengths considered were respectively {45mm, 38mm, 31mm, 24mm} and {25mm, 48mm} for cameras 1 and 2. Calibration parameters are reported in Table 1: notice that the principal point of camera 1 varies significantly (more than 100 pixels) for different focal lengths. Although the focal lengths changed, the position and orientation of both cameras remained unaltered between different experiments.

For each focal length and for each camera, we identified the red and white segments on the floor through the following procedure. For the red lines, we got an image $i_r(x, y)$ with enhanced red areas subtracting the green and blue channels from the red one and filtering then the image with a Gaussian filter of size 15×15 and $\sigma = 4$ pixels. For the white lines, we filtered in the same manner the sum of the three color channels to get $i_w(x, y)$. Then we clicked on the acquired images to roughly identify the area including the set of 4×4 red segments in Fig. 1; from this initial guess, the 8 red segments and the 6 white segments in the images were estimated by maximizing the average value of $i_r(x, y)$ and $i_w(x, y)$ along the segments. Typical results are shown in Fig. 1; the typical fitting errors observed on the available images were limited to 4-5 pixels in the worst cases (Fig. 1d).



(a) (b) (c) (d)

Fig. 1. (a) A typical image acquired by the Nikon camera at a focal length of 45mm, the longest one considered here. (b) Three degenerate conics (each with a different color) identified on the image of the Pentax camera. (c) Zoom of the area c in panel (b). (d) Zoom of the area d in panel (b): notice that in this case the red line is skewed by some pixels with respect to the red line on the floor.

3.2 Results

For each focal pair we built sets of 3 to 7 degenerate conics from the red and white lines identified in the images. Then we estimated the internal parameters of the Nikon and Pentax cameras for each possible pair of focal lengths. In a first experiment, we considered only degenerate conics constituted by two orthogonal lines, like those in Fig. 1; in a second experiment, we inserted two degenerate conics of parallel lines in the set of conics used for calibration. Tables 2 and 3 report the errors on the estimated focal lengths ($e_f = 100\frac{\tilde{f}_1 - f_1}{f_1}$, where $\tilde{f}_1$ is the estimated focal length and e_f defines the percent error) and principal point ($e_p = \| [(\tilde{c}_{x,1} - c_{x,1}) \ (\tilde{c}_{y,1} - c_{y,1}])\|$, where $[\tilde{c}_{x,1} \ \tilde{c}_{y,1}]$ is the estimated principal point and e_p is its distance from the ground true principal point in pixels) through the linear estimate provided by Eqs. (10) and (11) and after the non linear minimization of the quadratic cost function associated to Eq. (9). The same tables also report the rank of the matrix $\mathbf{A}$ in Eq. (8) with respect to its maximum admissible rank for the given number of conics.

When the conics are constituted of orthogonal lines and $N \leq 5$, the estimate of the camera parameters performed with the linear method (Eqs. (10) and (11)) is biased and poorly accurate, with an uncertainty always larger than 4.5% on f_1 (Table 2), which is significantly larger than the focal uncertainty obtained with the Zhang's method [2] (approximately equal to 1% in the worst case, see Table 1). Similar unsatisfying results are obtained for the estimate of the principal point, with typical uncertainties of hundreds of pixels. Notice that, for $N < 5$, the rank of $\mathbf{A}$ is not maximum and therefore the solution of the linear system in Eq. 10 is not uniquely defined.

On the other hand, when $N > 5$, even the simple linear procedure produces an estimate of the camera parameters that fit inside the confidence intervals provided by the Camera Calibration Toolbox for Matlab [3]. Finally, the least squares solution of Eq. (9) obtained with non linear optimization does produce a solution which is always in the confidence interval provided by the Camera Calibration Toolbox for Matlab [3], for any number of conics, with precision and accuracy that increase with N.

Table 1. Internal parameters (with uncertainties) of the cameras used during the experiments, computed through the Camera Calibration Toolbox for Matlab [3]

camera	focal	f_i [pixel]	$c_{x,i}$ [pixel]	$c_{y,i}$ [pixel]
1 (Nikon)	45mm	8711±65	2374±30	1624±38
1 (Nikon)	38mm	7743±51	2324±25	1505±26
1 (Nikon)	31mm	6089±28	2366±14	1605±18
1 (Nikon)	24mm	4723±19	2340±12	1533±15
2 (Pentax)	35mm	6314±64	2161±29	1481±33
2 (Pentax)	48mm	8219±68	2168±27	1485±25

Table 2. Errors (mean ± standard deviation) on the estimated focal and principal point for different set of N degenerate conics, each constituted by two orthogonal lines

	Linear estimate		Non linear estimate			
N	e_f [%]	e_p [pixels]	e_f [%]	e_p [pixels]	rank($\mathbf{A}$)	max rank
3	-5.66±16.11	336.20±722.22	0.16±0.69	4.98±2.58	10	12
4	-0.60±6.39	104.86±138.86	0.12±0.58	4.36±1.55	12	13
5	0.63±4.56	59.04±88.85	0.13±0.59	3.89±1.47	14	14
6	0.57±1.48	5.08±4.10	0.08±0.58	3.36±1.60	15	15
7	0.53±1.47	5.56±4.92	0.12±0.56	3.74±2.47	16	16

When two of the conics considered by the method are constituted of parallel lines (Table 3), the estimate of the camera parameters provided by the linear method becomes less stable, although it still represents a valid initial guess for the minimization of Eq. (9) through the Levemberg-Marquardt algorithm; an accurate result for f_1 and $[c_{x,1}\ c_{y,1}]$ is obtained in this case only for $N = 7$. After the non linear minimization, however, the estimate of the camera parameters have a precision and accuracy comparable to those reported in Table 2 and therefore to that typical of Zhang's method [2]. It is however to be noticed that, in this case, for $N = 3$ the present method fails to provide and accurate estimate of f_1 and $[c_{x,1}\ c_{y,1}]$ for all the cases considered (see Discussion).

Overall, the non linear method described in the previous section does provide an estimate of the camera internal parameters which generally fits into the confidence intervals provided by the Camera Calibration Toolbox for Matlab [3]. If the number of conics is sufficently high ($N > 6$), even the simple linear method associated to Eqs. (10) and (11) is capable to provide reliable results.

4 Discussion and Conclusion

The method described here permits to re-calibrate a camera after zooming, under the assumption that it does not change its position ($\mathbf{t}_1$) and orientation ($\mathbf{R}_1$) during zooming. This could be for instance the case of a surveillance camera installed with a fixed angle of view, or a reflex camera mounted on a tripod. Although assuming that the camera image plane does not rotate during zooming

Table 3. Errors (mean ± standard deviation) on the estimated focal and principal point for different set of N degenerate conics, each constituted by two orthogonal lines but two constituted by parallel lines

	Linear estimate		Non linear estimate			
N	e_f [%]	e_p [pixels]	e_f [%]	e_p [pixels]	rank($\mathbf{A}$)	max rank
3	-2.17±14.19	216.65±408.07	-9.62±31.19	123.09±372.50	10	12
4	-0.71±2.99	86.16±102.12	0.13±0.44	4.27±2.25	12	13
5	23.16±69.48	448.01±1158.01	0.12±0.48	3.98±2.09	14	14
6	-11.03±36.83	132.19±391.47	0.05±0.50	3.29±1.58	15	15
7	0.34±1.45	6.29±5.03	0.08±0.50	4.04±2.49	16	16

appears to be reasonable, assuming that $\mathbf{t}_1$ does not change could be more critical; in fact, in our camera model (Eq. (1)), $\mathbf{t}_1$ represents the center of the optical system of the camera [10], which actually goes forward and backward when the zoom lens elongates or shortens to modify its focal length. Despite this geometrical aspect was neglected in Eq. (9), the results reported in Tables 2 and 3 demonstrates that the accuracy achieved by the proposed method in estimating f_1, $c_{x,1}$ and $c_{y,1}$ is comparable to that of the Zhang's method [2, 3], that on the other hand does estimate also the external parameters ($\mathbf{R}_1$, $\mathbf{t}_1$) of the camera, as well as the distortion parameters of the lens.

The high accuracy achieved by the proposed method is even more noticeable considering that only one image (containing 3 ore more conics) is necessary for camera re-calibration. Such result is possible because the method does not utilize only data measured from a unique camera, like in the Zhang's method. Instead, it makes use of other information beyond $\mathbf{R}_1$ and $\mathbf{t}_1$, consisting in the parameters of the conic plane in the reference frame of camera 2 (p_x, p_y and p_z), that have to be estimated before the zooming of camera 1; and the matrices $\{\mathbf{C}_{2,j}\}_{j=1..N}$, measured by camera 2. The usage of such *a-priori* information makes the method also robust with respect to noise on the measured conics. Fig. 1 highlights in fact that the line fitting procedure adopted with our experimental setup is not characterized by a sub-pixel accuracy. Despite of this, the parameters estimated by the proposed method well fit in the interval of confidence provided by the Camera Calibration Toolbox [3] that represents our ground true.

A general approach to camera re-calibration in a real scenario would actually require to develop an algorithm to identify the conics in the images (for instance, the Hough transform could be used to identify the lines). Then, correspondence between the features observed by camera 1 and camera 2 should be solved. Since these problems have been widely studied and a wide literature exists on them [10], they have not been assessed here.

It is to be noticed that the present algorithm has been tested here only using pairs of degenerate conics, as these are naturally identifiable on any tiled floor; on the other hand, the approach can be applied to any kind of conic including

circles, ellipses and hyperbolas without requiring any modification (furthermore, different kind of conics can actually appear at the same time in Eq. (9)). In particular, circles and ellipses could be interesting in practical applications, as this shapes occur quite frequently in real scenes [7].

Depending on the kind of conics appearing in Eq. (9), however, different kind of numerical issues could raise. For instance, for the case considered here of degenerate conics constituted of two lines with equations $xc\theta_1 + ys\theta_1 = \rho_1$ and $xc\theta_2 + ys\theta_2 = \rho_2$ (where $c\theta$ and $s\theta$ indicate the sin and cosine of θ), the corresponding conic matrix $\mathbf{C}$ is given by:

$$\mathbf{C} = \begin{bmatrix} c\theta_1 c\theta_2 & (c\theta_1 s\theta_2 + s\theta_1 c\theta_2)/2 & -(\rho_1 c\theta_2 + \rho_2 c\theta_1)/2 \\ (c\theta_1 s\theta_2 + s\theta_1 c\theta_2)/2 & s\theta_1 s\theta_2 & -(\rho_1 s\theta_2 + \rho_2 s\theta_1)/2 \\ -(\rho_1 c\theta_2 + \rho_2 c\theta_1)/2 & -(\rho_1 s\theta_2 + \rho_2 s\theta_1)/2 & \rho_1 \rho_2 \end{bmatrix}. \quad (12)$$

The elements of the 2×2 upper left block of $\mathbf{C}$ are always between -1 and 1; on the other hand, the third element of the diagonal of $\mathbf{C}$ is proportional to the product of the distances of the two lines from the point $[0\ 0]^T$ of the image and, given the size of the images considered here, it can easily exceed $1,000,000$, making the matrix $\mathbf{A}$ in Eq. (9) badly conditioned because of the presence of largely spread numbers. A different interpretation of badly conditioning in this situation is obtained considering that the least squares solution of Eq. (9) provides the estimate of f_1, $c_{x,1}$ and $c_{y,1}$ which minimizes the quadratic distances between the elements of any j-th conic matrix $\mathbf{D}^T\mathbf{P}_1^T\mathbf{C}_{1,j}\mathbf{P}_1\mathbf{D}$ and $\lambda_j\mathbf{C}_{2,j}$ in Eq. (6). If one of the elements of the conic matrix is dramatically larger than the others, high importance is given in the quadratic cost function to this element whereas other elements appearing in the matrices of the conics are practically neglected. To avoid this, we have therefore normalized the pixel space dividing by 1,000 the pixel coordinates (and therefore also the camera focal lengths appearing in $\mathbf{K}_1$ and $\mathbf{K}_2$) before applying the proposed method and finally multiplying by 1,000 the estimate of $(f_1, c_{x,1}, c_{y,1})$ obtained with Eqs. (11) and (9). Some experimental results, not reported here, have demonstrated that such normalization procedure is actually fundamental to the success of the method: without it, the precision and accuracy decrease dramatically, making the obtained estimate of the internal parameters of camera 1 practically unusable.

An additional remark to the results presented in Tables 2 and 3 is to be done. We have in fact verified that large averages or standard deviations of the error are generally caused by only one very bad estimate of f_1, $c_{x,1}$ and $c_{y,1}$, among the ten experiments considered to compute each element of the tables. In practice, especially when N is low and the linear estimate is used, the method may fail to produce a reliable estimate of the camera parameters because of some particular geometrical configurations that induce numerical instability. We are going to investigate such conditions more in detail in the future.

Furthermore, we have obtained some analytic, preliminary results showing that Eq. (8) can be manipulated to obtain the estimate of f_1, $c_{x,1}$ and $c_{y,1}$ through the solution of a linear system. The accuracy and stability of such method still has to be evaluated and compared to the results obtained with the iterative solution of Eq. (9), but it promises to generate a more elegant framework

where the linear estimate of the parameters can be analytically analyzed to identify geometrically critical situations; this would also permit to study noise propagation from the measured conics to the estimated parameters, for instance through covariance analysis.

Overall, the method described here permits to re-calibrate a camera after zooming using a unique image and features that are easily extracted from real world scenes, achieving an accuracy similar to that of more traditional and time consuming calibration procedures.

References

1. Zhang, Z.: Camera Calibration. In: Medioni, G., Kang, S.B. (eds.) Emerging Topics in Computer Vision, ch. 2, pp. 4–43. Prentice Hall Professional Technical Reference (2004)
2. Zhang, Z.: Flexible Camera Calibration by Viewing a Plane from Unknown Orientations. In: ICCV 1999 (1999)
3. Bouguet, J.-Y.: Camera Calibration Toolbox for Matlab, `http://www.vision.caltech.edu/bouguetj/calib_doc/index.html/`
4. Abado, F., Camahort, E., Vivó, R.: Camera Calibration Using Two Concentric Circles. In: International Conference on Image Analysis and Recognition, pp. 688–696 (2004)
5. Ying, X., Zha, H.: Camera calibration using principal-axes aligned conics. In: Yagi, Y., Kang, S.B., Kweon, I.S., Zha, H. (eds.) ACCV 2007, Part I. LNCS, vol. 4843, pp. 138–148. Springer, Heidelberg (2007)
6. Yang, C., Sun, F., Hu, Z.: Planar conic based camera calibration. In: The 15th International Conference on Pattern Recognition, vol. 1, pp. 555–558 (2000)
7. Frosio, I., Alzati, A., Bertolini, M., Turrini, C., Borghese, N.A.: Linear pose estimate from corresponding conics. Pattern Recognition 45, 4169–4181 (2012)
8. Kahl, F., Heyden, A.: Using Conic Correspondences in Two Images to Estimate the Epipolar Geometry. In: Int. Conf. on Computer Vision, pp. 761–766 (1998)
9. Borghese, N.A., Colombo, F.M., Alzati, A.: Computing Camera Focal Length by Zooming a Single Point. Pattern Recognition 39, 1522–1529 (2006)
10. Faugeras, O.: Three-dimensional computer vision: a geometric viewpoint. MIT Press (1993)
11. Nikon website, `http://www.nikon.com/`
12. Canon website, `http://www.canon.com/`
13. Fraser, C.S., Ajlouni, A.S.S.: Zoom-dependent camera calibration in digital close-range photogrammetry. Photogrammetric Engineering & Remote Sensing 72(9), 1017–1026 (2006)
14. Sun, X., Sun, J., Zhang, J., Li, M.: Simple zoom-lens digital camera calibration method based on exif. In: Three-Dimensional Image Capture and Applications VI, vol. 79 (2004)
15. Calore, E., Frosio, I.: Perspective correction in digital photography using on board accelerometer. Submitted to Image and Vision Computing (2013)
16. Adi, B.-I., Greville, T.N.E.: Generalized inverses. Springer, New York (2003)