Calibrating Freely Moving Cameras

Imran N. Junejo, Xiaochun Cao*and Hassan Foroosh
{ ijunejo | xccao | foroosh }@cs.ucf.edu
Univeristy of Central Florida, Orlando, Fl 32826

Abstract

We present a novel practical method for self-calibrating a camera which may move freely in space while changing it internal parameters by zooming. We show that point correspondences between a pair of images, and the fundamental matrix computed from these point correspondences, are sufficient to recover the internal parameters of a camera. Unlike other methods, no calibration object with known 3-D shape is required and no limitation are put on the unknown motion, as long as the camera is projective.

The main contribution of this paper is development of a global linear solution which is based on the well-known Kruppa equations. We introduce a formulation different from the Huang and Faugeras constraints. The method has been extensively tested on synthetic and real data and promising results are reported.

1 Introduction

Camera calibration is the process of extracting *intrinsic* and *extrinsic* camera parameters. The calibration is an obligatory process in Computer Vision in order to obtain a Euclidean structure of the scene (up to a global scale), and to determine rigid camera motion.

Self-calibration differs from conventional calibration where the camera internal parameters are determined from the image of a known calibration grid or properties of the scene, such as vanishing points of orthogonal directions. The prefix *self-* is added as soon as the world's Euclidean structure is unknown, which can be seen as a case of "0D" calibration. In self-calibration the metric properties of the cameras are determined directly from constraints on the internal and/or external parameters.

The first self-calibration method, originally introduced into computer vision by Faugeras *et al.* [2], involves the use of the Kruppa equations. The Kruppa equations are two-view constraints that require only the fundamental matrix to be known, and consist of two independent quadratic equations in the elements of the dual of the absolute conic. Algorithms for computing the focal lengths of two cameras given the corresponding fundamental matrix and knowledge of the remaining intrinsic parameters are provided by Hartley [4]. Mendonça [9] generalized the results in [4] for an arbitrary number of cameras and introduced a built-in method for the detection of critical motions for each pair of images in the sequence. Thorough analysis of critical motions which would result in ambiguous solutions by Kruppa-based methods are described in [12].

An alternative direct method for self-calibration was introduced by Triggs [13], which estimates the absolute dual quadric over many views. The basic idea is to transfer a constraint on the dual image of absolute conic to a constraint on the absolute dual quadric, and hence determine the matrix representing the absolute dual quadric, from which a rectifying 3D homography can be decomposed that transforms from projective to metric reconstruction. Heyden and Astrom [7] showed that metric reconstruction was possible knowing only skew and aspect ratio, and Pollefeys *et al.* [10] and Heyden and Astrom [8] showed that zero skew alone was sufficient.

Special motions can also be used for self-calibration. Agapito *et al.* [1] and Seo and Hong[11] solved the self-calibration of a rotating and zooming camera using the infinite homography constraint. Before their work, Hartley [5] solved the special case where the camera's internal parameters remain constant throughout the sequence. Frahm and Koch [3] showed it was also possible to solve the problem of generally moving camera with varying intrinsics but known rotation information.

In this paper we focus on extracting internal parameters of a freely moving camera and present a simple and novel global linear solution. We do not assume any special camera motion or known camera rotation matrix as used by [1],[11],[3],[10],[5]. The proposed method relies only on point correspondences between different views from a single camera. We test our method on synthetic as well as on real data and present encouraging results.

A brief introduction to the concepts related to a pinhole camera are presented in Section 2. Cameras are self-

*Currently working with ObjectVideo Inc.

The 18th International Conference on Pattern Recognition (ICPR'06)
0-7695-2521-0/06 $20.00 © 2006 IEEE

calibrated (Section 3) by estimating the fundamental matrix between different views captured from the same camera. We present experimental results (Section 4) before concluding (Section 5).

2 Some Preliminaries

The projection of a 3**D** scene point $\mathbf{X} \sim \begin{bmatrix} X & Y & Z & 1 \end{bmatrix}^T$ onto a point in the image plane $\mathbf{x} \sim \begin{bmatrix} x & y & 1 \end{bmatrix}^T$ for a perspective camera can be modeled by the central projection equation:

$$\mathbf{x} \sim \underbrace{\mathbf{K}\left[\begin{array}{c|c} \mathbf{R} & -\mathbf{RC} \end{array}\right]}_{\mathbf{P}}\mathbf{X}, \quad \mathbf{K} = \begin{bmatrix} \lambda f & \gamma & u_o \\ 0 & f & v_o \\ 0 & 0 & 1 \end{bmatrix} \tag{1}$$

where $\sim$ indicates equality up to a non-zero scale factor and $\mathbf{C} = \begin{bmatrix} C_x & C_y & C_z \end{bmatrix}^T$ represents camera center. Here $\mathbf{R}{=}\mathbf{R}_x\mathbf{R}_y\mathbf{R}_z{=}\begin{bmatrix}\mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3\end{bmatrix}$ is the rotation matrix and $-\mathbf{RC}$ is the relative translation between the world origin and the camera center. The upper triangular 3×3 matrix $\mathbf{K}$ encodes the five intrinsic camera parameters: focal length f, aspect ratio λ, skew γ and the principal point at (u_o, v_o).

The aim of camera calibration is to determine the calibration matrix $\mathbf{K}$. Instead of directly determining $\mathbf{K}$, it is common practice to compute the symmetric matrix $\mathbf{K}^{-T}\mathbf{K}^{-1}$ or its inverse. In vision literature, the matrix $\boldsymbol{\omega} = \mathbf{K}^{-T}\mathbf{K}^{-1}$ is referred to as the Image of the Absolute Conic(**IAC**) and its inverse $\boldsymbol{\omega}^*$ as the **dual IAC**, which can be decomposed uniquely using the Cholesky Decomposition to obtain $\mathbf{K}$.

For a camera moving freely in 3D space, the relation between any two views is described by the Fundamental matrix if the camera is translated between the views. The fundamental matrix maps a point from camera i to a line, an epipolar line, in cameras j and is given by:

$$\mathbf{F}_{i,j} = [e']_\times \mathbf{H}^\pi_{i,j}$$

Thus $\mathbf{F}_{i,j}$ is determined by the cross product of the epipole $\mathbf{e}'$ with the homography $\mathbf{H}^\pi_{i,j}$ induced by a plane π. Also $\mathbf{F}_{i,j} = \mathbf{K}_j^{-T}\mathbf{E}\mathbf{K}_i^{-1}$, where $\mathbf{E}$ is the essential matrix.

Fig. 1 depicts an illustration of two views from a camera. Generally, two consecutive images from a camera contain some overlapping area. This overlapping area can be used to obtain the fundamental matrix $\mathbf{F}_{i,j}$, which relates a point in image I_j to a line in image I_i. As the internal parameters change at each view, the absolute conic $\boldsymbol{\omega}$ also changes. Thus, we need to compute $\boldsymbol{\omega}$ for each image of the camera.

3 Camera Self-Calibration

Each camera is allowed to vary its internal parameters by zooming in/out and we assume zero skew, unit aspect ratio and the principle point at the center of an image. We use these general assumption to estimate the unknown varying focal length.

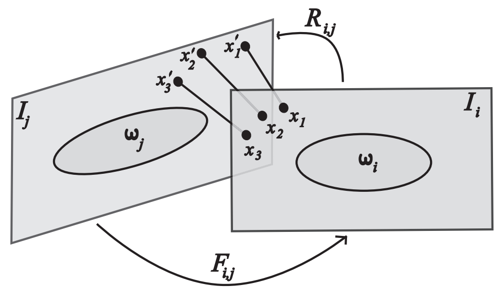


Figure 1. Illustration of two views from a camera: Two consecutive images from a camera contain an overlapping area. This overlapping area can be used to obtain the fundamental matrix $\mathbf{F}_{i,j}$, which relates a point in image I_j to a line in image I_i. As the internal parameters change at each view, the absolute conic ω also changes.

3.1 Linear Solution with varying focal length

Consider an image sequence of n frames and let $\mathbf{K}_i$ be the intrinsic parameters for a camera at i^{th} frame, then the calibration matrix is of the form:

$$\mathbf{K}_i = \begin{bmatrix} f_i & 0 & 0 \\ 0 & f_i & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

where $\gamma = 0, \lambda = 1, (u_o = 0, v_o = 0)$. Assuming a known principal point(u_o, v_o), images can be transformed so that the principle points lies at $(0, 0)$. The principle point for a camera remains unchanged even when the focal length changes.

For a freely moving camera, the fundamental matrix is the only information that can be easily obtained and is thus frequently used for self-calibration, as in the case of Kruppa equations [2]. Therefore, after computing the fundamental matrix between two different views i and j from a camera, we can use the relation:

$$\mathbf{F}_{i,j}\boldsymbol{\omega}_i^*\mathbf{F}_{i,j}^T \sim [e']_\times\boldsymbol{\omega}_j^*[e']_\times, \tag{2}$$

where ω_i^* and ω_j^* represent **dual IAC** for two different views, i and j, respectively. If the intrinsic parameters remain constant over different views then $\omega_i^*{=}\omega_j^*$ and Eq (2) can be expressed as $\mathbf{F}_{i,j}\omega_i^*\mathbf{F}_{i,j}^T = -[e']_\times\omega_i^*[e']_\times$.

Eq (2) amounts to 3 linearly independent equations with an unknown scale, allowing for the symmetry and rank deficiency. Eq (2) is not in a form that can be easily applied and traditional methods cross multiply to eliminate the unknown scale ([6, 9]). Instead of taking this approach, we solve for the unknown scale involved in the 3 equations directly.

For a camera with unknown focal length, $\boldsymbol{\omega}^*$ for an j^{th} frame is given as:

$$\boldsymbol{\omega}_j^* = \begin{bmatrix} W_j & 0 & 0 \\ 0 & W_j & 0 \\ 0 & 0 & \alpha \end{bmatrix},$$

where $W_j = \alpha f_j^2$ and α, the unknown scale, is different for every image pair. For $\boldsymbol{\omega}_i^*$, the left hand side of Eq (2), α is normalized to 1. Hence the three unknowns for a pair of images are α, W_i and W_j.

Since any $\mathbf{K}_i$ has $\mathbf{K}_{i_{(12)}} = \mathbf{K}_{i_{(13)}} = \mathbf{K}_{i_{(23)}} = 0$, Eq (2) gives us only three equations to solve for the three unknowns, which we formulate as:

$$\mathbf{A}_{i,j}\mathbf{Y}_{i,j}=\boldsymbol{b}_{i,j} \;\; where \;\; \mathbf{Y}_{i,j}= \begin{bmatrix} W_i & W_j & \alpha \end{bmatrix}^T$$

and $\mathbf{A}_{i,j}$ is a 3×3 matrix containing coefficients of W_i, W_j and α; and $\boldsymbol{b}_{i,j}$ contains the known $\mathbf{F}_{i,j}$ and $[e']_\times$. From the solution vector $\mathbf{Y}_{i,j}$, the intrinsic parameters for each view can be obtained as:

$$f_i=\sqrt{\mathbf{Y}_{i,j_{(1)}}}, \quad \alpha=\sqrt{\mathbf{Y}_{i,j_{(3)}}}, \quad f_j=\sqrt{\mathbf{Y}_{i,j_{(2)}}}/\alpha,$$

A global solution for computing intrinsic parameters for a varying focal length camera over k frames is given by cascading the above equation into:

$$\underbrace{\begin{bmatrix} \mathbf{A}_{i,j} & 0 & \cdots \\ 0 & \mathbf{A}_{i+1,j+1} & \cdots \\ \vdots & \ddots & \vdots \\ 0 & 0 & \mathbf{A}_{i+k,j+k} \end{bmatrix}}_{\mathcal{A}} \underbrace{\begin{bmatrix} \mathbf{Y}_{i,j} \\ \mathbf{Y}_{i+1,j+1} \\ \vdots \\ \mathbf{Y}_{i+k,j+k} \end{bmatrix}}_{\mathcal{Y}} = \underbrace{\begin{bmatrix} \boldsymbol{b}_{i,j} \\ \boldsymbol{b}_{i+1,j+1} \\ \vdots \\ \boldsymbol{b}_{i+k,j+k} \end{bmatrix}}_{\mathcal{B}} \quad (3)$$

Eq (3) computes a linear least square solution for an entire image sequence.

3.2 Linear solution for varying focal length and unknown λ

In the previous section we assumed that the aspect ratio(λ) is unity. λ remains constant for any single camera. Eq (2) can be extended to solve for an unknown λ by selecting a reference frame q. Therefore, three images i.e. two instances of Eq (2) are sufficient to solve for six unknowns. Eq (2) for an image j from a camera with respect to the reference frame q can be expressed as:

$$\mathbf{F}_{q,j}\begin{bmatrix} \lambda W_q & 0 & 0 \\ 0 & W_q & 0 \\ 0 & 0 & 1 \end{bmatrix}\mathbf{F}_{q,j}^T=[e']_\times \begin{bmatrix} \lambda W_j & 0 & 0 \\ 0 & W_j & 0 \\ 0 & 0 & \alpha \end{bmatrix}[e']_\times$$

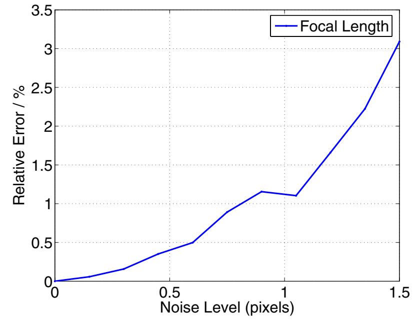


Figure 2. Performance of the self-calibration method VS. noise level in pixels.

Thus the first pair introduces four unknowns $(\lambda, W_q, W_j, \alpha)$ and every subsequent frame introduces only 2 unknowns (unknown scale and new focal length). Once λ is determined non-linearly, it is substituted into Eq (2) for improving the estimated focal length. Eq (2) can not be used to solve for any more unknown intrinsics parameters (see [6]).

An obvious advantage of the above linear solution is its simplicity and computational efficiency, making it suitable for many real time applications.

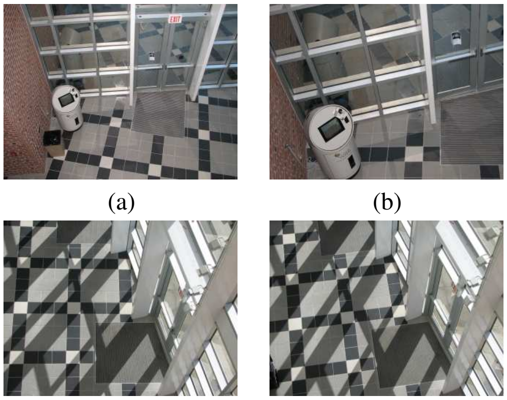
4 Experiments & Results

4.1 Synthetic Data

In order to validate the robustness of the proposed self-calibration method, we generate a synthetic camera with $f = 1000, \lambda = 1, \gamma = u_o = v_o = 0$. Gaussian noise with zero mean and standard deviation of $\sigma \leq 1.5$ was added to the data points used for computing the fundamental matrix. The relative difference with respect to the focal length rather than the absolute error is a more geometrically meaningful error measure. Therefore, we measure the relative error of estimated f with respect to true f while varying the noise level from 0.01 to 1.5 pixels. For each noise level, we performed 500 independent trials and the results are shown in Fig.2. The error increases almost linearly with respect to the noise level. For a minimum noise of 1.5 pixels, we found that the error was under 3.1%.

4.2 Real Data

In the first data set, two cameras, labeled l and r, are located on the second and third floor of a building monitoring a lobby entrance, respectively. The cameras are zooming in/out while translating and rotating at the same time. The height and motion of each camera is subjectively selected to allow observation of the specified area. We compared our method to the standard three parameter estimation technique using three orthogonal vanishing points [6]. Results obtained from the two methods are compared in Table 1 and



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

Figure 3. (a) and (b) are views taken from camera l; c and d from camera j, while looking at a lobby entrance. See text for more details.

View	Our Method	Compared Method
Fig. 3a	3048.77	3290.36
Fig. 3b	3000.35	3350.17
Fig. 3c	1590.24	1766.74
Fig. 3d	2598.47	2482.24

Table 1. Computed focal length from our method compared with vanishing points based calibration technique.

the images used are shown in Fig 3. The results obtained from the two methods are comparable to each other.

The second data set consist of a zooming in/ou video taken from a road while looking at some houses. Fig 4 depicts four such instances from the sequences taken from a camera different from the one used in above data set. The focal length for each instance is shown below each image of Fig 4.

The errors could be attributed to several sources. Besides noise, non-linear distortion and imprecision of the extracted features, one source is the causal experimental setup using minimal information, which is deliberately targeted for a wide spectrum of applications. Despite all these factors, our experiments indicate that the proposed algorithms provides good results.

5 Conclusions

We have successfully demonstrated a novel global linear solution approach to recovering the intrinsic parameters of a camera. Each camera is assumed to undergo a general motion. The method is very efficient and requires only point correspondences across fames to determine the fundamental matrix. Once the fundamental matrix is determined, we solve for the internal parameters linearly. We also provide a non-linear solution to extracting the aspect-ratio for each camera. Experiments are carried out on several sequences. Self-calibration is tested on synthetic as well as on real data. Encouraging results indicate the applicability of the proposed system.

$f = 1146.01$ $f = 843.92$ $f = 1184.21$ $f = 1723.93$

Figure 4. Four instance from a video sequence taken from a road while looking at some houses.

References

[1] L. D. Agapito, E. Hayman, and I. Reid. Self-calibration of rotating and zooming cameras. *Int. J. Comput. Vision*, 45(2):107–127, 2001.

[2] O. Faugeras, T. Luong, and S. Maybank. Camera self-calibration: theory and experiments. In *Proc. of ECCV*, pages 321–334, 1992.

[3] J. Frahm and R. Koch. Camera calibration with known rotation. In *Proc. IEEE ICCV*, pages 1418–1425, 2003.

[4] R. I. Hartley. Estimation of relative camear positions for uncalibrated cameras. In *Proc. ECCV*, pages 579–587, 1992.

[5] R. I. Hartley. Self-calibration of stationary cameras. *Int. J. Comput. Vision*, 22(1):5–23, 1997.

[6] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, 2003.

[7] A. Heyden and K. Astrom. Euclidean reconstruction from image sequences with varying and unknown focal length and principal point. In *Proc. IEEE CVPR*, pages 438–443, 1997.

[8] A. Heyden and K. Astrom. Flexible calibration: Minimal cases for auto-calibration. In *Proc. IEEE ICCV*, pages 350–355, 1999.

[9] P. R. S. Mendonça. *Multiview Geometry: Profiles and Self-Calibration*. PhD thesis, University of Cambridge, Cambridge, UK, 2001.

[10] M. Pollefeys, R. Koch, and L. V. Gool. Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters. *Int. J. Comput. Vision*, 32(1):7–25, 1999.

[11] Y. Seo and K. Hong. About the self-calibration of a rotating and zooming camera: Theory and practice.

[12] P. Sturm. Critical motion sequences for monocular self-calibration and uncalibrated euclidean reconstruction. In *Proc. IEEE CVPR*, pages 1100–1105, 1997.

[13] B. Triggs. Autocalibration and the absolute quadric. In *Proc. IEEE CVPR*, pages 609–614, 1997.