ELSEVIER

Pattern Recognition Letters 22 (2001) 1393–1404

Pattern Recognition Letters

www.elsevier.com/locate/patrec

Self-calibration with varying focal length from two images obtained by a camera with small rotation and general translation

J.-S. Liu, J.-H. Chuang *

Department of Computer and Information Science, National Chiao Tung University, 1001 Ta Hsueh Rd., Hsinchu 30050, Taiwan, ROC

Received 26 July 2000; received in revised form 5 January 2001

Abstract

The stratified self-calibration approach based on the absolute conic or its dual, the absolute dual quadric, has the merit of allowing the intrinsic camera parameters to vary while being retrieved from an image sequence. In this paper, we show that for a camera with small rotation and general translation, a new linear equation resulted from the infinity homography can be added to a system of linear equations to compute the absolute dual quadric. Experiments with both synthetic and real images show that satisfactory results can be obtained with the proposed linear approach. It is possible to further improve the calibration result by adopting some nonlinear optimization schemes, e.g., a suitable *LM*-like algorithm, to enforce the absolute dual quadric constraints using the linear solutions as an initial guess. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Self-calibration; Absolutee conic; Infinity homography

1. Introduction

Self-calibration of a camera from images has been an important research topic on computer vision over the last few years since it may reduce the need of off-line calibration and increases on-line flexibility. It is shown in (Faugeras, 1992; Hartley et al., 1992) that general projective reconstructions, i.e., the simplest type of self-calibration, can be obtained easily using two or more uncalibrated projective images. Recently more and more researchers pay their attention to possible ways of upgrading these reconstructions from projective to metric. Faugeras et al. (1992) proposed a robust self-calibration method using the Kruppa equations to impose constraints on the fixed internal parameters obtained from the fundamental matrix. A number of approaches based on similar concepts to self-calibration have been developed (Armstrong et al., 1996; Faugeras, 1995; Hartley, 1994a,b; Heyden and Åaström, 1996).

* Corresponding author. Tel.: +886-35-712121; fax: +886-35-721490.
E-mail address: jchuang@cis.nctu.edu.tw (J.-H. Chuang).

PII: S0167-8655(01)00064-2

Instead of using the Kruppa equations, some widely accepted approaches (Pollefeys and Van Gool, 1998; Quan and Triggs, 2000) are based on the absolute quadric which is a concise parameterization of the absolute conic introduced by Triggs (1997). By means of this parametric representation, it is shown that the self-calibration can be done even if the camera intrinsic parameters are allowed to vary while generating a sequence of images. On the other hand, based on the infinity homography, approaches of stratified reconstruction for projective, affine, and finally Euclidian space have also been widely adopted in the last few years (Pollefeys and Van Gool, 1999; Pollefeys et al., 1996; Zisserman et al., 1995). In fact, some other researchers have dealt with the calibration problem for some special motions (Apapito et al., 1998; Faugeras et al., 1998), in which the work in (Apapito et al., 1998) has involved both the absolute quadric and the infinity homography. A good analysis of what has been done in the filed of self-calibration can also be found in (Pollefeys, 1999).

It is shown in (Pollefeys and Van Gool, 1998) that, under the condition that the intrinsic camera parameters, except for the focal length, are known, a linear solution of the varying focal length together with the location of a particular affine structure can be obtained. The linear solution can be then used to initialize the corresponding nonlinear optimization procedure. As a special case of having only two images obtained with varying focal lengths, only one modulus constraint proposed in (Pollefeys and Van Gool, 1999) for the infinity homography exists that the plane at infinity cannot be determined. In order to handle the two-image case, Pollefeys and Van Gool (1999) add the scene constraint obtained from vanishing points to solve the self-calibration problem.

In this paper, we propose a linear approach to the self-calibration problem with varying focal length for a camera with small rotation and general translation. The approach is based on the infinity homography and the absolute quadric and does not require an additional vanishing point constraint as in (Pollefeys and Van Gool, 1999). Such camera motions can often be seen in stereo vision applications in which a small rotation between two cameras can be found. The paper is organized as follows. In Section 2, some background geometry and notation are introduced. Section 3 describes the general self-calibration problem based on the absolute quadric and the infinity homography. Then in Section 4, the linear solution for the special case of two images, obtained from a camera with small rotation and general translation, based on the infinity homography constraints is introduced. Following that, the associated 3D metric reconstruction procedure is summarized in Section 5, and some experimental results are given in Section 6. Finally, we draw conclusions in Section 7.

2. Background geometry and notation

In this section, a brief review is given for the classical projective geometry notions of infinity homography, plane at infinity, absolute conic, and their relationships to camera calibration.

2.1. Projection matrix and infinity homography

A basic projection procedure of scene points onto an image by a perspective camera can be described as

$$m \propto PM, \tag{1}$$

where $\propto$ denotes the equality up to a scaling factor, P is the 3×4 projection matrix, $M = [X\ Y\ Z\ 1]^{\mathrm{T}}$ and $m = [x\ y\ 1]^{\mathrm{T}}$ represent the homogeneous coordinates of a 3D world point and an image point, respectively. Due to the stratums of space, the projection process should be represented by means of its corresponding projection matrix in the space under consideration.

For Euclidean space, the projection matrix can be represented as

$$P_{\text{euc}} = KP_0T = \begin{bmatrix} f_x & s & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0_3^{\mathrm{T}} & 1 \end{bmatrix}, \tag{2}$$

where T represents the transformation of coordinate systems from world to the camera-centered system, P_0 denotes the perspective projection and K is the camera matrix consisting of the intrinsic parameters of camera. In the camera matrix, f_x and f_y are the focal lengths measured in width and in height of the pixels in the image, respectively, s is a factor measuring the skew of the two image axes, and u_0 and v_0 are the image coordinates of the principal point.

Consider the projective space. The projection matrix can be represented as

$$P_{\text{proj}} = [H|e_r], \tag{3}$$

where e_r is the epipole, and H, the Homography, describes the projection from a particular reference plane to the image plane, as discussed next.

Given a reference plane $\Pi = [\pi^{\mathrm{T}}\ 1]^{\mathrm{T}} \triangleq [\pi_1\ \pi_2\ \pi_3\ 1]^{\mathrm{T}}$ in the 3D space, a point $M_\Pi = [m_\Pi^{\mathrm{T}}\ 1]^{\mathrm{T}}$ is said to lie on this plane if and only if $\Pi^{\mathrm{T}} M_\Pi = \pi^{\mathrm{T}} m_\Pi + 1 = 0$. Specifically, since $\pi^{\mathrm{T}} m_\Pi = -1$, the relationship can be represented as

$$M_\Pi = \begin{bmatrix} m_\Pi \\ 1 \end{bmatrix} = \begin{bmatrix} m_\Pi \\ -\pi^{\mathrm{T}} m_\Pi \end{bmatrix} = \begin{bmatrix} I_{3\times 3} \\ -\pi^{\mathrm{T}} \end{bmatrix} m_\Pi. \tag{4}$$

Hence, the projection process which maps the 3D point M_Π to its image point $\tilde{m}_\Pi$ by the projective projection matrix can be described as

$$\tilde{m}_\Pi \propto P_{\text{proj}} M_\Pi = [H|e_r] \begin{bmatrix} I_{3\times 3} \\ -\pi^{\mathrm{T}} \end{bmatrix} m_\Pi \tag{5}$$

or

$$\tilde{m}_\Pi \propto [H - e_r \pi^{\mathrm{T}}] m_\Pi. \tag{6}$$

Thus, $[H - e_r\pi^{\mathrm{T}}]$ in fact represents the homography between M_Π and $\tilde{m}_\Pi$. Or, more precisely it can be written as

$$H' = H - e_r \pi^{\mathrm{T}}. \tag{7}$$

If the plane Π is chosen to be $[0\ 0\ 0\ 1]^{\mathrm{T}}$, the corresponding homography is simply given by H. This is the homography denoted in the projective projection matrix (3). On the other hand, the infinity homography is denoted as another special homography which describes the transformation from the plane at infinity to the image plane

$$H^\infty = H - e_r \pi_\infty^{\mathrm{T}}, \tag{8}$$

where π_∞ is the vector consisting of the first three elements of $[\pi_{\infty_1}\ \pi_{\infty_2}\ \pi_{\infty_3}\ 1]$ which represents the location of the plane at infinity, Π_∞, with respect to the reference plane of H. The details about the plane at infinity are given in the next subsection.

2.2. Plane at infinity and absolute conic

According to the notation used in (Zisserman et al., 1998), a point in 3-space is represented by homogeneous 4-vector $X = [X_1\ X_2\ X_3\ X_4]$, and the plane at infinity, or the infinity plane, Π_∞, is the

plane expressed as $X_4 = 0$ in an affine frame and is setwise invariant under Euclidean motions, i.e., any rigid motion of a camera will not change the camera's relative position and orientation with respect to Π_∞.

The absolute conic, Ω, is a point conic on Π_∞ represented as $X_1^2 + X_2^2 + X_3^2 = 0$ and $X_4 = 0$, containing only imaginary points (Semple and Kneebone, 1979). As its dual, the absolute dual quadric is denoted as Ω^*. A special property associated with the absolute conic is that if camera parameters do not change, then the image of the absolute conic, ω, and its dual, ω^*, will also stay the same for all views. In particular, for Euclidean representation of the world, such a property expressed with ω^* can be realized as

$$\omega^* \propto P_{\text{euc}} \Omega^*_{\text{euc}} P^{\text{T}}_{\text{euc}} = K[R^{\text{T}}|t] \begin{bmatrix} I_{3\times 3} 0_3 \\ 0_3^{\text{T}} 0 \end{bmatrix} \begin{bmatrix} R \\ t \end{bmatrix} K^{\text{T}} = KK^{\text{T}}. \tag{9}$$

In cases which allow variable intrinsic camera parameters, there is a particularly useful property of the dual image of the absolute conic such that

$$\omega_i^* = K_i K_i^{\text{T}} \propto P_i \Omega^* P_i^{\text{T}} \tag{10}$$

is satisfied for all views (i's). According to (10), constraints on the intrinsic camera parameters associated with K_i can therefore be transformed to constraints to those on elements of ω_i^*. This actually provides a basis for the self-calibration.

3. Self-calibration

The absolute dual quadric and the infinity homography are the basis of the self-calibration since images of the former encode the camera matrix for all views while the latter encodes the camera rotation. By these parameterizations, location of the plane at infinity as well as the intrinsic camera parameters can be obtained. Thus, the projection matrix for 3D reconstructions in metric space can be obtained.

In this section, a general approach to the self-calibration problem based on the absolute dual quadric is briefly reviewed. Then, in order to overcome the difficulty induced by the special case of using only two images wherein the solution in general cannot be determined uniquely, the infinity homography constraints for camera motions with small rotation and general translation are presented in the next section.

Consider the absolute dual quadric given in (10). Starting from its Euclidean representation, such a quadric can eventually be expressed in projective space as

$$\omega^* \propto P_{\text{euc}_i} \Omega^*_{\text{euc}} P^{\text{T}}_{\text{euc}_i} = \left(P_{\text{proj}_i} T^{-1}_{\text{PM}}\right) \Omega^*_{\text{euc}} \left(T^{-\text{T}}_{\text{PM}} P^{\text{T}}_{\text{proj}_i}\right) = P_{\text{proj}_i} \left(T^{-1}_{\text{PM}} \Omega^*_{\text{euc}} T^{-\text{T}}_{\text{PM}}\right) P^{\text{T}}_{\text{proj}_i} = P_{\text{proj}_i} \Omega^*_{\text{proj}} P^{\text{T}}_{\text{proj}_i}, \tag{11}$$

where P_{proj_i} denotes the projection matrix in projective space for the ith image,

$$T_{\text{PM}} = \begin{bmatrix} K^{-1} & 0 \\ \pi_\infty & 1 \end{bmatrix} \tag{12}$$

is the transformation matrix to upgrade the geometry from projective to metric, and

$$\Omega^*_{\text{proj}} = \begin{bmatrix} KK^{\text{T}} & -KK^{\text{T}}\pi_\infty \\ -\pi_\infty^{\text{T}} KK^{\text{T}} & \pi_\infty^{\text{T}} KK^{\text{T}} \pi_\infty \end{bmatrix} \tag{13}$$

is the absolute dual quadric in projective space.

In particular, if the world frame is aligned with the first camera, $P_{\text{proj}_1} = [I|0]$, then we have

$$\omega_i^* = K_i K_i^{\text{T}} \propto P_{\text{proj}_i} \Omega^{*'}_{\text{proj}} P^{\text{T}}_{\text{proj}_i} \tag{14}$$

with

$$\Omega^{*'}_{\text{proj}} = \begin{bmatrix} K_1 K_1^{\mathrm{T}} & -K_1 K_1^{\mathrm{T}} \pi_\infty \\ -\pi_\infty^{\mathrm{T}} K_1 K_1^{\mathrm{T}} & \pi_\infty^{\mathrm{T}} K_1 K_1^{\mathrm{T}} \pi_\infty \end{bmatrix}. \tag{15}$$

For an ideal camera, as suggested with an approximation provided in (Pollefeys and Van Gool, 1998), we have $u_0 = v_0 = 0$, $s = 0$ and $f_x = f_y \triangleq f_i$, which lead to the following camera matrix:

$$K_i = \begin{bmatrix} f_i & 0 & 0 \\ 0 & f_i & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{16}$$

Thus, (15) can be simplified as [1]

$$\Omega^{*'}_{\text{proj}} = \begin{bmatrix} f_1^2 & 0 & 0 & -f_1^2 \pi_{\infty_1} \\ 0 & f_1^2 & 0 & -f_1^2 \pi_{\infty_2} \\ 0 & 0 & 1 & -\pi_{\infty_3} \\ -f_1^2 \pi_{\infty_1} & -f_1^2 \pi_{\infty_2} & -\pi_{\infty_3} & f_1^2 \pi_{\infty_1}^2 + f_1^2 \pi_{\infty_2}^2 + \pi_{\infty_3}^2 \end{bmatrix}. \tag{17}$$

Let $a = f_1^2$, $b = -f_1^2 \pi_{\infty_1}$, $c = -f_1^2 \pi_{\infty_2}$, $d = -\pi_{\infty_3}$, and $e = f_1^2 \pi_{\infty_1}^2 + f_1^2 \pi_{\infty_2}^2 + \pi_{\infty_3}^2$, (14) becomes

$$\begin{bmatrix} f_i^2 & 0 & 0 \\ 0 & f_i^2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \propto P_{\text{proj}_i} \begin{bmatrix} a & 0 & 0 & b \\ 0 & a & 0 & c \\ 0 & 0 & 1 & d \\ b & c & d & e \end{bmatrix} P_{\text{proj}_i}^{\mathrm{T}}, \tag{18}$$

and we can obtain the following system of linear equations (details are shown in Appendix A):

$$\begin{aligned} &k_{11}a + k_{12}b + k_{13}c + k_{14}d + k_{15}e + k_{16} = 0, \\ &k_{21}a + k_{22}b + k_{23}c + k_{24}d + k_{25}e + k_{26} = 0, \\ &k_{31}a + k_{32}b + k_{33}c + k_{34}d + k_{35}e + k_{36} = 0, \\ &k_{41}a + k_{42}b + k_{43}c + k_{44}d + k_{45}e + k_{46} = 0. \end{aligned} \tag{19}$$

When only two images are available, instead of a unique solution, only a family of solutions can be determined for (19). Even if the rank 3 constraint for the absolute dual quadric (17) is imposed, one still ends up with four possible solutions (Pollefeys and Van Gool, 1998). To overcome this difficulty, Pollefeys and Van Gool (1999) and Pollefeys (1999) added a scene constraint obtained from vanishing points and thus resulted in enough linear equations. In the next section, without any additional scene constraints, we show that under special camera motions of a small rotation and a general translation it is possible to obtain a close-form solution through the properties of the infinity homography.

4. A new linear method

Consider the case of two images. Eq. (8) becomes

$$H_{12}^{\infty} = H_{12} - e_r \pi_\infty^{\mathrm{T}} \propto K_2 R_{12} K_1^{-1}, \tag{20}$$

[1] This matrix is essentially the same of that shown in (Pollefeys and Van Gool, 1998), except for some minor changes in notation.

where e_r is the epipole in the second image, $H_{12} = [e_r]_\times F$ denotes the homography, [2] H_{12}^{∞} represents the corresponding infinity homography, and R_{12} is the rotation from the first camera to the second one. Note that the last term of (20) is indeed the infinity homography represented in Euclidean space (see Luong and Vieville, 1993 for details).

For a small rotation, we have

$$K_2 R_{12} K_1^{-1} = \begin{bmatrix} \frac{f_2}{f_1} r_{11} & \frac{f_2}{f_1} r_{12} & f_2 r_{13} \\ \frac{f_2}{f_1} r_{21} & \frac{f_2}{f_1} r_{22} & f_2 r_{23} \\ \frac{f_2}{f_1} r_{31} & \frac{f_2}{f_1} r_{31} & r_{33} \end{bmatrix} \approx \begin{bmatrix} \frac{f_2}{f_1} & \frac{f_2}{f_1}\omega_z & -f_2\omega_y \\ -\frac{f_2}{f_1}\omega_z & \frac{f_2}{f_1} & f_2\omega_x \\ \frac{\omega_y}{f_1} & -\frac{\omega_x}{f_1} & 1 \end{bmatrix}, \tag{21}$$

where ω_x, ω_y and ω_z are the rotation angles with respect to the x-axis, the y-axis and the z-axis, respectively. Thus, (20) can be simplified as

$$\begin{bmatrix} h_{11} - e_{r_1}\pi_{\infty_1} & h_{12} - e_{r_1}\pi_{\infty_2} & h_{13} - e_{r_1}\pi_{\infty_3} \\ h_{21} - e_{r_2}\pi_{\infty_1} & h_{22} - e_{r_2}\pi_{\infty_2} & h_{23} - e_{r_2}\pi_{\infty_3} \\ h_{31} - e_{r_3}\pi_{\infty_1} & h_{32} - e_{r_3}\pi_{\infty_2} & h_{33} - e_{r_3}\pi_{\infty_3} \end{bmatrix} \propto \begin{bmatrix} \frac{f_2}{f_1} & \frac{f_2}{f_1}\omega_z & -f_2\omega_y \\ -\frac{f_2}{f_1}\omega_z & \frac{f_2}{f_1} & f_2\omega_x \\ \frac{\omega_y}{f_1} & -\frac{\omega_x}{f_1} & 1 \end{bmatrix}. \tag{22}$$

Consider the three diagonal elements of the both sides of (22), we have

$$\pi_{\infty_1} = \frac{h_{11} - h_{22} + e_{r_2}\pi_{\infty_2}}{e_{r_1}}, \tag{23}$$

$$f_2 = \frac{h_{11} - e_{r_1}\pi_{\infty_1}}{h_{33} - e_{r_3}\pi_{\infty_3}} f_1 \tag{24}$$

$$= \frac{h_{22} - e_{r_2}\pi_{\infty_2}}{h_{33} - e_{r_3}\pi_{\infty_3}} f_1. \tag{25}$$

From (23), b can be expressed as a function of a and c as

$$b = -f_1^2\pi_{\infty_1} = \frac{-h_{11} + h_{22}}{e_{r_1}} \cdot f_1^2 - \frac{e_{r_2}}{e_{r_1}} \cdot f_1^2\pi_{\infty_2} \triangleq k_{b_1} \cdot a + k_{b_2} \cdot c. \tag{26}$$

With (19) and (26), the five variables defined in (18) and thus the four parameters f_1, π_{∞_1}, π_{∞_2} and π_{∞_3}, can be solved. Moreover, once f_1 is obtained, the value of f_2 can be calculated using either (24) or (25). [3]

5. 3D metric reconstruction

A stratified approach to self-calibration, as outlined in (Pollefeys, 1999; Pollefeys and Van Gool, 1999; Zisserman et al., 1995), includes a step-by-step procedure of projective reconstruction, affine reconstruction, and finally the desired metric reconstruction from multiple images. A similar stratified approach but using only two images, according to the linear algorithm proposed in this paper can be summarized as follows:

[2] $[e_r]_\times$ is an antisymmetric 3×3 matrix representing the vectorial product with e (see Pollefeys, 1999) and F is the fundamental matrix relating the two images.

[3] In the implementation, values of f_2 obtained from (24) and (25), respectively, are found to be quite similar. The average of the two values are used as a robust estimation of f_2 in the simulation.

The 3D metric reconstruction procedure:

- *Stage* 1. Find the two projective projection matrices:

$$P_{\text{proj}_1} = [I_{3\times 3}|0_3], \quad P_{\text{proj}_2} = [H_{12}|e_r]. \tag{27}$$

- *Stage* 2. Obtain T_{PM}; i.e., find the four parameters f_1, π_{∞_1}, π_{∞_2} and π_{∞_3} by solving (19) and (26).
- *Stage* 3. Derive the projection matrix for Euclidean space

$$P_{\text{euc}_i} = P_{\text{proj}_i} T_{\text{PM}}^{-1}, \quad i = 1, 2. \tag{28}$$

- *Stage* 4. Obtain the metric structure through the SVD-based 3D reconstruction method given in (Rothwell et al., 1995).

6. Experiment results

In this section, calibration results obtained with the proposed linear method are presented. For a stratified self-calibration approach, the accuracy of the perspective reconstruction affects the results remarkably. The modified eight-point algorithm is adopted in the implementation, which usually gives a satisfactory fundamental matrix, F, even under noisy conditions (Hartley, 1997). Using the derived fundamental matrix, the metric reconstruction is then conducted.

The performance of the proposed approach is examined using both synthetic and real images. In the former, statistical results such as means and standard deviations are provided for f_1, f_2, Π_∞ and 3D reconstruction error, respectively, under different noise conditions. In the latter, on the other hand, we measure the parallelism and orthogonality, as well as the 2D reprojection errors, of 3D structures obtained from the metric reconstruction.

6.1. Experiments using synthetic data

The simulations are carried out on pairs of images obtained from a synthetic scene consisting of 50 3D points. As shown in Fig. 1, these points are generated randomly in a $100 \times 100 \times 100$ cube centered at

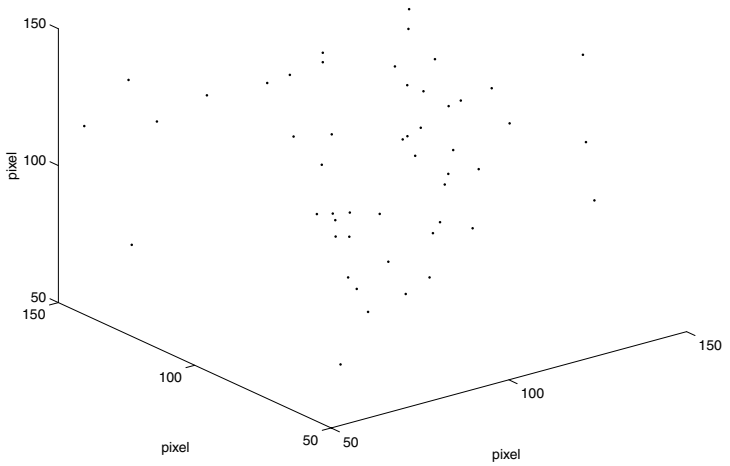

Fig. 1. The 3D structure for the simulation.

$(100, 100, 100)$. Then, two images of this synthetic scene with size of 1400×1400 pixels are generated using two Euclidean projection matrices, $P_{\text{euc}_1} = K_1[R_1|t_1]$ and $P_{\text{euc}_2} = K_2[R_2|t_2]$, with $K_1 = \text{diag}[550, 550, 1]$, $R_1(\omega = 0°,\ \phi = 0°,\ \kappa = 0°)$, $t_1 = [0, 0, 0]$, $K_2 = \text{diag}[600, 600, 1]$, $R_2(\omega = 2°,\ \phi = 4.5°,\ \kappa = 7.6°)$ and $t_2 = [1, -20, -3]$. The three parameters ω, ϕ, and κ denote the rotation angles around the x-axis, the y-axis and the z-axis, i.e., the tilt, pan, and swing angles, respectively.

Fig. 2 shows the mean and standard deviation of the estimation of the first focal length (f_1) for zero mean Gaussian noise, with standard deviation ranging from 0 to 1 pixel, added to locations of image points. For each noise level, the statistics are obtained from a total of 100 trials. Note that in this figure and also the next few figures, the half longitudinal short line denotes the standard deviation of experimental results obtained with the 100 trials for each noise level and the middle of this short line denotes the mean of these. While their standard deviations have the general trend to increase with the noise level, the means of these estimations approximately give the correct value $f_1 = 550$. Similar results can also be observed for the estimations of the second focal length, $f_2 = 600$, as shown in Fig. 3.

To evaluate the accuracy of the estimation of π_∞, the real π_∞ is first obtained with Eq. (20) using the real rotation R_2 and the fundamental matrix F obtained from the image pair before adding noises. The angular error is then defined as the angle between the estimated π_∞ and the real π_∞. Fig. 4 shows the mean and standard deviation of the angular error for various noise levels.

As for the 3D reconstruction errors, since real locations of all 3D points are given for the simulation, we can directly measure the average distance between the true 3D structure and the corresponding structure recovered in metric space. The distance measurement procedure can be described as follows. First, the two structures are moved to have their centroids located at the origin of the world

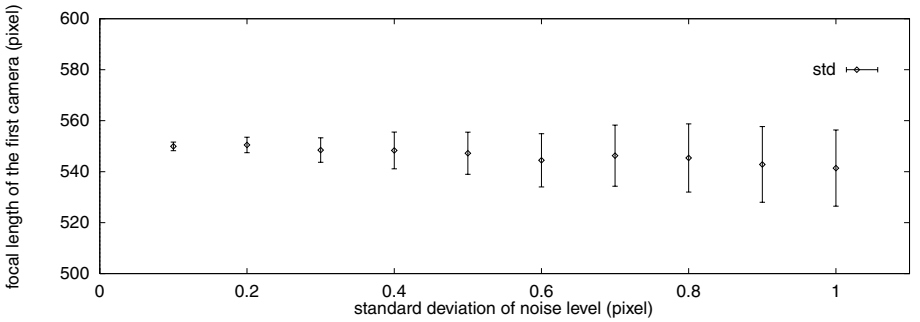

Fig. 2. The estimated first focal length (f_1) for various noise levels.

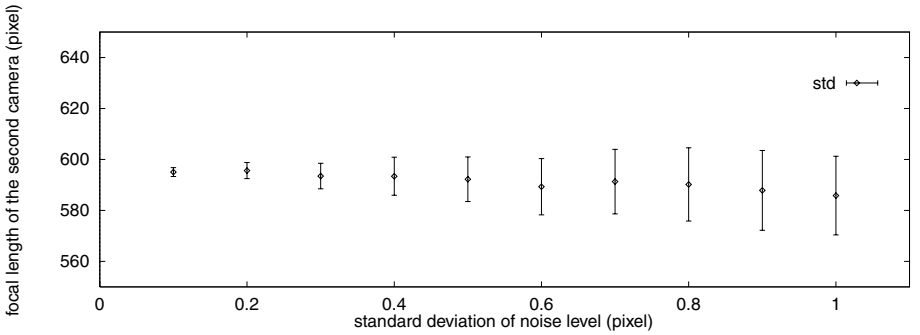

Fig. 3. The estimated second focal length (f_2) for various noise levels.

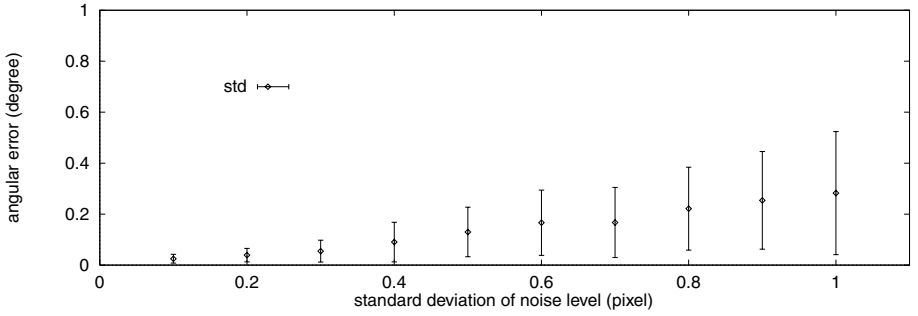

Fig. 4. Angular error for the estimates of π_∞ for various noise levels.

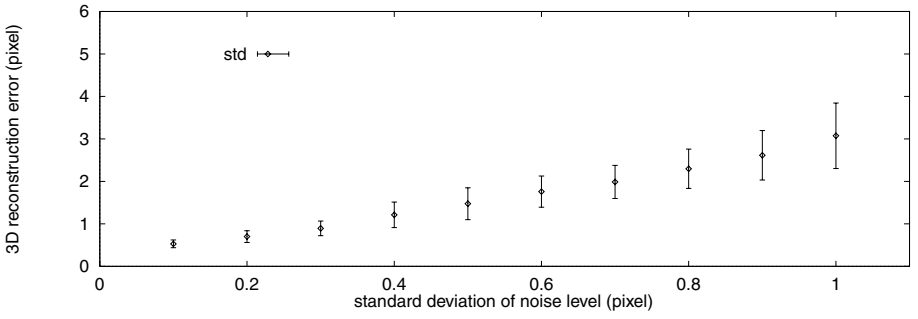

Fig. 5. The 3D metric reconstruction error for various noise levels.

coordinate system. Secondly, a size normalization operation for both structures is carried out. Thirdly, pointwise Euclidean distances between the two structures are calculated for corresponding 3D feature points. And finally, we average these distances to obtain an estimation of the 3D reconstruction error. Fig. 5 shows the mean value and the corresponding standard deviation of the 3D reconstruction error.

From these simulation results, one can observe that the proposed linear method is capable of solving the self-calibration problem, in terms of accuracy and robustness, under conditions of small rotation and general translation of the camera.

6.2. Experiments based on real images

In this subsection, results obtained with the proposed linear method for pairs of real images are presented. The images are obtained from the CMU image sequence (CIL Stereo Dataset, 1998). Fig. 6 shows a typical image in the image sequence with 36 reference points. A subset of these point features is connected with 10 line segments (marked with 0–9 in Fig. 7) to facilitate the parallelism and orthogonality measurement of the recovered 3D metric structure.

In the experiment, the principal points given in (CIL Stereo Dataset, 1998) are utilized as known intrinsic parameters in addition to $f_x/f_y = 1$ and $s = 0$. The proposed linear method is applied to 10 image pairs $(1\text{–}2, 1\text{–}3, \ldots, 1\text{–}11)$ for the corresponding 3D metric reconstruction. For the recovered 3D structure, Table 1 shows the true and estimated angles between 11 selected pairs of line segments. Each of the esti-

Fig. 6. A typical 576 × 384 image with 36 reference points in the CMU image sequence.

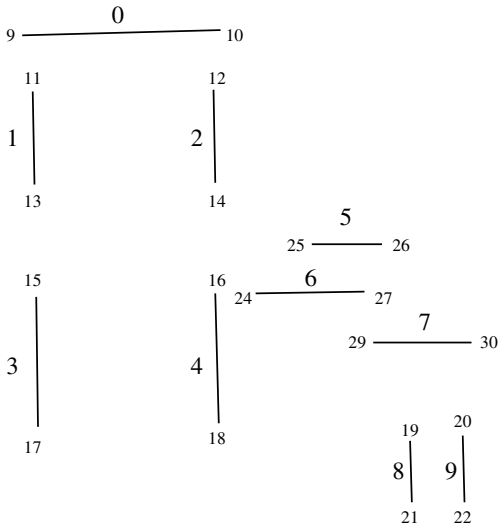

Fig. 7. Ten line segments connecting a subset of point features.

mated angles is obtained by averaging the corresponding angles calculated from the 10 image pairs. It is readily observable from Table 1 that the proposed approach preserve the parallelism and orthogonality [4] of the recovered 3D structure satisfactorily.

[4] The parallelism and orthogonality measure, which shows how the shape of the reconstructed 3D structure is preserved, is the main performance measure of a metric reconstruction approach and has been adopted in (Pollefeys and Van Gool, 1999; Louraskis and Driche, 2000).

Table 1
True and estimated angles between 11 selected pairs of line segments

Line segment pair	True angle (deg.)	Estimated angle (deg.)	Standard deviation (deg.)
1–2	1.2071	0.8955	0.1281
3–4	0.9094	0.8177	0.0669
5–6	2.5379	1.1341	0.0667
6–7	1.8160	2.9217	1.1020
8–9	1.5674	0.6351	0.1471
1–5	88.8999	89.3040	0.1398
3–7	89.3059	89.1086	0.0383
4–6	90.1267	90.1520	0.0222
5–9	89.4330	89.3644	0.2756
7–8	90.5845	90.1790	0.0373
0–1	90.2291	90.5843	0.0380

Table 2
The reprojection errors for the 3D metric reconstruction

Image	Mean error (pixel)	Standard deviation (pixel)
First	5.1806×10^{-7}	2.3773×10^{-7}
Second	1.8433×10^{-2}	6.6671×10^{-3}

Besides the parallelism and orthogonality measurement, the reprojection error, which measures the differences between the original image points and those obtained by reprojecting the recovered 3D structure back to the image plane, provides another performance measurement of the metric reconstruction. A small value of such an error implies that coplanarity and collinearity constraints associated with the projective geometry are well satisfied. In this study, because we assume $P_{\text{proj1}} = [I|0]$, i.e., the first image is perfect, the computation errors are propagated to the second image in the reconstruction process. It is shown in Table 2 that the reprojection error of the recovered 3D structure for the first image is much lower than that for the second; both are reasonably small for typical applications.

Note that the 3D reconstruction error of about 0.7 or 0.8 pixels is obtained in Fig. 5 for synthetic image data with standard deviation of the noise level equal to 0.2. For real images, e.g., the one shown in Fig. 6, the noise level could be much higher, and so will be the 3D reconstruction error. On the other hand, fairly small reprojection errors are obtained in Table 2 for real image points since they are mainly used as the aforementioned constraints in the 3D reconstruction. Such constraints should be satisfied even if the reconstructed 3D structure is not very accurate.

7. Conclusion

In this paper, we present a linear method that can solve the self-calibration problem with only two images if the images are obtained by a camera with small rotation and general translation, possibly having a varying focal length. Experiment results for synthetic simulations as well as real images show that the proposed approach performs satisfactorily. Moreover, besides image pairs, our method can also handle the case of multiple images since such a method is based on the absolute dual quadric and is designed to solve the self-calibration problem associated with image sequence. On the other hand, as part of an optimal self-calibration method, the linear solution may be good enough for an additional nonlinear optimization procedure, which can be adopted to improve the calibration result further.

Appendix A

The coefficients of the system of linear equations given in (19) are listed in detail as follows: [5]

$k_{11} = p_{11}^2 + p_{12}^2 - (p_{21}^2 + p_{22}^2)$	$k_{12} = 2(p_{11}p_{14} - p_{21}p_{24})$	$k_{13} = 2(p_{12}p_{14} - p_{22}p_{24})$	$k_{14} = 2(p_{13}p_{14} - p_{23}p_{24})$	$k_{15} = p_{14}^2 - p_{24}^2$	$k_{16} = p_{13}^2 - p_{23}^2$
$k_{21} = p_{11}p_{21} + p_{12}p_{22}$	$k_{22} = p_{14}p_{21} + p_{11}p_{22}$	$k_{23} = p_{14}p_{22} + p_{12}p_{24}$	$k_{24} = p_{14}p_{23} + p_{13}p_{24}$	$k_{25} = p_{14}p_{24}$	$k_{26} = p_{13}p_{23}$
$k_{31} = p_{11}p_{31} + p_{12}p_{32}$	$k_{32} = p_{14}p_{31} + p_{11}p_{32}$	$k_{33} = p_{14}p_{32} + p_{12}p_{34}$	$k_{34} = p_{14}p_{33} + p_{13}p_{34}$	$k_{35} = p_{14}p_{34}$	$k_{36} = p_{13}p_{33}$
$k_{41} = p_{21}p_{31} + p_{22}p_{32}$	$k_{42} = p_{24}p_{31} + p_{21}p_{32}$	$k_{43} = p_{24}p_{32} + p_{22}p_{34}$	$k_{44} = p_{24}p_{33} + p_{23}p_{34}$	$k_{45} = p_{24}p_{34}$	$k_{46} = p_{23}p_{33}$

References

Apapito, L., Hayman, E., Reid, I., 1998. Self-calibration of a rotating camera with varying intrinsic parameters. In: Proc. British Machine Vision Conf.

Armstrong, M., Zisserman, A., Hartley, R., 1996. Self-calibration from image triplets. In: Proc. ECCV, pp. 3–16.

CIL Stereo Dataset, 1998. http://www.cs.cmu.edu/afs/cs/project/cil/ftp/html/cil-ster.html.

Faugeras, O.D., 1992. What can be seen in three dimensions with an uncalibrated stereo rig? In: Proc. ECCV, pp. 563–578.

Faugeras, O.D., 1995. Stratification of 3D vision: projective, affine, and metric representation. J. Opt. Soc. Am. 12 (3), 465–484.

Faugeras, O.D., Luong, Q.-T., Maybank, S.J., 1992. Camera self-calibration: theory and experiments. In: Proc. ECCV, pp. 321–334.

Faugeras, O.D., Quan, L., Strum, P., 1998. Self-calibration of a 1D projective camera and its application to the self-calibration of a 2D projective camera. In: Proc. ECCV.

Hartley, R.I., 1994a. Euclidean reconstruction from multiple views. In: Applications of Invariance in Computer Vision, Lecture Notes in Computer Science, Vol. 825. Springer, Berlin, pp. 237–256.

Hartley, R.I., 1994b. Self-calibration from multiple views with a rotating camera. In: Proc. ECCV, pp. 471–478.

Hartley, R.I., 1997. In defense of eight-point algorithm. IEEE Trans. Pattern Anal. Mach. Intell. 19 (6), 580–593.

Hartley, R.I., Gupta, R., Chang, T., 1992. Stereo from uncalibrated cameras. In: Proc. CVPR, pp. 761–764.

Heyden, A., Åaström, K., 1996. Euclidean reconstruction from constant intrinsic parameters. In: Proc. ICPR, pp. 339–343.

Louraskis, M., Driche, R., 2000. Camera self-calibration using the singular value decomposition of the fundamental matrix. In: Proc. ACCV, pp. 403–408.

Luong, Q.-T., Vieville, T., 1993. Canonic representations for the geometries of multiple projective views. Technique Report UCB/CSD-93-772.

Pollefeys, M., 1999. Self-calibration and metric 3D reconstruction from uncalibrated image sequences. Ph.D. Thesis, Katholieke Universtieit Leuven, Belgium.

Pollefeys, M., Van Gool, L., 1998. Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters. In: Proc. ICCV, pp. 90–95.

Pollefeys, M., Van Gool, L., 1999. Stratified self-calibration with the modulus constraint. IEEE Trans. Pattern Anal. Mach. Intell. 21 (8), 707–724.

Pollefeys, M., Van Gool, L., Proesmans, M., 1996. Euclidean 3D reconstruction from image sequences with variable focal lengths. In: Proc. ECCV, pp. 31–42.

Quan, L., Triggs, B., 2000. A unification of autocalibration methods. In: Proc. ACCV, pp. 917–922.

Rothwell, C., Faugeras, O., Csurka, G., 1995. Different paths towards projective reconstruction. In: Europe–China Workshop of Invariance and Geometric Modelling.

Semple, J., Kneebone, G., 1979. Algebraic Projective Geometry. Oxford University, UK.

Triggs, B., 1997. Autocalibration and the absolute quadric. In: Proc. CVPR, pp. 609–614.

Zisserman, A., Beardsley, P., Reid, I., 1995. Metric calibration of a stereo rig. In: Proc. IEEE Workshop Representation of Visual Scenes, pp. 93–100.

Zisserman, A., Liebowitz, D., Armstrong, M., 1998. Resolving ambiguities in auto-calibration. Philos. Trans. Roy. Soc. London, Series A 56 (1740), 1193–1211.

[5] Herein, p_{ij}s denote the elements of P_{proj_i}.