Journal of Southeast University (English Edition) Vol. 25, No. 2, pp. 193 – 198 June 2009 ISSN 1003—7985

Greedy calibration method for binocular camera system via characteristic homography matrix

Liu Tianliang Jin Feiyi Luo Limin

(Laboratory of Image Science and Technology, Southeast University, Nanjing 210096, China)

Abstract: A plane-based and linear camera calibration technique without considering lens distortion is proposed in a greedy and intuitive framework for the binocular camera system. Characteristic homography matrix and consistency constraints in close range are employed in this calibration. First, in order to calculate the internal geometries of the cameras, total least-square fitting as a robust tool for the geometrical cost function is exploited to recover the accurate principal point of each camera from all the characteristic lines of the homography matrices for all model planes. Secondly, generic prior knowledge of the aspect ratio of pixel cells is incorporated into the system to obtain the exact principal length in each camera. Thirdly, extrinsic geometries are accurately computed for all planar patterns with respect to each monocular camera. Finally, the rigid displacement between binocular cameras can be obtained by imposing the consistency constraints in 3-space geometry. Both simulation and real image experimental results indicate that reasonably reliable results can be obtained by this technique. And the proposed method is sufficient for applications where high precision is not required and can be easily performed by common computer users who are not experts in computer vision.

Key words: close range; calibration; homography matrix; consistency constraint; stereovision

The binocular digital camera system has been widely applied to close range photogrammetry and machine vision with a host of stereo applications, such as robot navigation, parts inspection, image-based rendering and creation of three-dimensional models, etc. Due to high dependence of the derived object data on the camera calibration process, it is necessary to calibrate the geometry of the binocular cameras including intrinsic and extrinsic parameters when two cameras are deployed in the stereo system.

The planar pattern is employed to calibrate the binocular camera system due to its advantages of simplicity and high accuracy[1–5]. To the authors' knowledge, a number of plane-based linear calibration methods have been developed, which can be classified into traditional algebraic-based calibration[1–3, 6] and geometric-based calibration[4–5, 7]. In a more intuitive geometric framework, the characteristic line that is constrained to the internal geometry of a given camera is deduced from a planar homography matrix[7–8]. Meanwhile, generic prior knowledge about the internal geometry of the cameras(such as the unit cell size of CCD sensors from a manufacturers' specs), which is related to the planar homography matrix, should be taken into account; the influence of the slim fluctuations with respect to the unit cell size of CCD sensors can be directly ignored in the calibration process[3, 6–7]. Furthermore, there exists an additional consistency constraint between binocular cameras viewing the same scene from different positions in 3-space geometry, which can be incorporated to obtain accurate external geometry in the given stereo setup[3–4, 6]. The authors extend previous work[7–8] on single camera calibration to the binocular case. In this paper, the characteristic properties of the planar homography matrix are exploited to calibrate the binocular camera system intuitively with part of the prior knowledge about the internal geometry of the given cameras; while the consistency constraints in 3-space geometry are adopted to estimate the extrinsic parameters between binocular cameras.

1 Camera Model

The finite projective camera, which can be considered as the standard pinhole model, is used in this paper. The left and right camera $C_i(i = 1, 2)$ can be assumed to have constant intrinsic parameter matrices $\boldsymbol{K}_i$,

$$\boldsymbol{K}_i = \begin{bmatrix} \alpha_i & \gamma_i & u_{0i} \\ 0 & \beta_i & v_{0i} \\ 0 & 0 & 1 \end{bmatrix} \quad i = 1, 2 \tag{1}$$

with the principal length α_i and β_i(in pixels), the principal point (u_{0i}, v_{0i}) (in pixels) and the skew factor γ_i accounting for non-rectangular pixels. We ignore the γ_i as a normal camera in the following and set it to be zero because the pixels in the CCD array can be supposed to be rectangular in our model. Note that α_i and β_i are related to the aspect ratio k_i of the pixel cell, i. e., $\beta_i = k_i\alpha_i$. Furthermore, k_i can be simply obtained as a special value from camera manufacturers' specs in a generic prior knowledge perspective. The fluctuation of the unit cell size is small enough to be ignored in view of modern optical element manufacturing. If the pixels in the CCD array can be assumed to be square, the aspect ratio can be set to 1. The image of the absolute conic(IAC) links to calibration and metric scene reconstruction, that is $\omega_i = \boldsymbol{K}_i^{-T}\boldsymbol{K}_i^{-1}$(see Ref. [2]).

The rotation matrix and the translation vector between the camera reference frame $I_i(i = 1, 2)$ and the world reference frame $\pi^j(j = 1, 2, \ldots, J)$ are $\boldsymbol{R}_i$ and $\boldsymbol{t}_i$, respectively. A 3D point $P(X_w, Y_w, Z_w)$, which lies on a special plane($Z_w = 0$), projects onto image planes of both cameras. In a projective framework, we can obtain

$$s\begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \boldsymbol{K}_i\boldsymbol{R}_t^i\begin{bmatrix} X_w \\ Y_w \\ 1 \end{bmatrix} = \boldsymbol{H}^i\begin{bmatrix} X_w \\ Y_w \\ 1 \end{bmatrix} \quad i = 1, 2 \tag{2}$$

Received 2008-09-08.

Biographies: Liu Tianliang (1980—), male, graduate; Luo Limin (corresponding author), male, doctor, professor, luo. list@ seu. edu. cn.

Citation: Liu Tianliang, Jin Feiyi, Luo Limin. Greedy calibration method for binocular camera system via characteristic homography matrix[J]. Journal of Southeast University (English Edition), 2009, 25(2): 193 – 198.

where $\boldsymbol{R}_{\mathrm{t}}^{i} = [\boldsymbol{r}_1^i \quad \boldsymbol{r}_2^i \quad \boldsymbol{t}_i]$, $\boldsymbol{H}^i = \boldsymbol{K}_i\boldsymbol{R}_{\mathrm{t}}^{i} = \boldsymbol{K}_i[\boldsymbol{r}_1^i \quad \boldsymbol{r}_2^i \quad \boldsymbol{t}_i] = [\boldsymbol{h}_1^i \quad \boldsymbol{h}_2^i \quad \boldsymbol{h}_3^i]$, s is a scale factor, $\boldsymbol{h}_q^i (q = 1, 2, 3)$ is the q-th column of $\boldsymbol{H}^i$, i. e. , $\boldsymbol{h}_i = [h_{1q}^i \quad h_{2q}^i \quad h_{3q}^i]^{\mathrm{T}}$.

Note that $\boldsymbol{R}_i = [\boldsymbol{r}_1^i \quad \boldsymbol{r}_2^i \quad \boldsymbol{r}_3^i]$, $\boldsymbol{r}_q^i (q = 1, 2, 3)$ is the q-th column of $\boldsymbol{R}_i$; $\boldsymbol{t}_i = [t_1^i \quad t_2^i \quad t_3^i]^{\mathrm{T}}$. And $\boldsymbol{H}^i$ is the plane-to-image homography matrix, which reflects a point on the special planar pattern being mapped from the world coordinates to the image coordinates of the respective camera i.

2 Calibration Method

2.1 Planar homography matrices estimation

If J model planes with different poses are adopted to calibrate a given binocular system, the planar homography matrix of each pose should be estimated accurately. Each model plane π^j with a given 2D reference pattern is used in this algorithm. Given $n \geqslant 4$ point correspondences with plane-to-image mapping, the direct linear transformation (DLT) algorithm can be used to determine the linear solution of $\boldsymbol{H}_j^i$ related to π^j and its image on $I_i (i = 1, 2)$, which can be an initial guess for the nonlinear refinement process. Owing to the non-invariance of the DLT algorithm, both 2D point sets must be normalized first to obtain more accurate estimation[9].

2.2 Linear solution for intrinsic parameters from the planar homography matrices

2.2.1 Characteristic lines construction

Because the intrinsic parameters are assumed to be constant, they can be obtained in a greedy scheme from the planar homography matrix. A useful characteristic of the planar homography matrix such as the characteristic line can be employed separately for each camera in the calibration approach[8]. For a given planar homography matrix $\boldsymbol{H}_j^i$, characteristic line $\boldsymbol{L}_j^i$ with the homogeneous coordinate vector $\{E_j, -1, F_j\}$ can be expressed by Eq. (3). This constraint is derived from $\boldsymbol{H}_j^i$ in terms of intrinsic parameters and planar scene orientation. And the detailed deduction of Eq. (3) for a given camera can be seen in Refs. [7 – 8]. The principal point (u_{0i}, v_{0i}) of the camera i satisfying the planar homography matrix $\boldsymbol{H}_j^i$ must lie on the corresponding characteristic line.

$$E_j u_{0i} - v_{0i} + F_j = 0 \qquad i = 1, 2; j = 1, 2, \ldots, J \tag{3}$$

where

$$E_j = k_i^2 \frac{h_{11}^i h_{32}^i - h_{31}^i h_{12}^i}{h_{31}^i h_{22}^i - h_{21}^i h_{32}^i}$$

$$F_j = \frac{h_{21}^i h_{31}^i + h_{22}^i h_{32}^i}{h_{31}^i h_{31}^i + h_{32}^i h_{32}^i} - \frac{h_{11}^i h_{31}^i + h_{12}^i h_{32}^i}{h_{31}^i h_{31}^i + h_{32}^i h_{32}^i} E_j$$

2.2.2 Principal points via total least-square fitting

Due to noise, the actual characteristic lines do not intersect at a point[8]. The constraint is then solved using a linear optimization framework based on the minimization of a geometrically motivated cost function. The square of the perpendicular distance in the image space from the actual principal point(u_{0i}, v_{0i}) for each camera i to the actual characteristic line $\boldsymbol{L}_j^i$ can be formulated by $d_j^2 = \dfrac{(E_j u_{0i} - v_{0i} + F_j)^2}{E_j^2 + 1}$. By total least-square fitting for all characteristic lines intersected at a point with the corresponding model planes[10], we have

$$\begin{bmatrix} \sum_{j=1}^{J} \dfrac{E_j^2}{E_j^2 + 1} & \sum_{j=1}^{J} \dfrac{-E_j}{E_j^2 + 1} \\ \sum_{j=1}^{J} \dfrac{-E_j}{E_j^2 + 1} & \sum_{j=1}^{J} \dfrac{1}{E_j^2 + 1} \end{bmatrix} \begin{bmatrix} u_{0i} \\ v_{0i} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{J} \dfrac{-E_j F_j}{E_j^2 + 1} \\ \sum_{j=1}^{J} \dfrac{F_j}{E_j^2 + 1} \end{bmatrix} \tag{4}$$

Solving Eq. (4), the principal point (u_{0i}, v_{0i}) for the respective camera i can be obtained accurately.

2.2.3 Principal lengths from reshaping planar homographies

According to the IAC, orthonormality of rotation matrix $\boldsymbol{R}_i$, α_i and β_i related by the aspect ratio k_i, and the given planar homography matrix $\boldsymbol{H}_j^i$, we can obtain[8]

$$\left.\begin{aligned} (a_j k_i^2 + b_j)\frac{1}{\beta_i^2} &= -h_{31}^i h_{32}^i \\ (c_j k_i^2 + d_j)\frac{1}{\beta_i^2} &= -h_{32}^i h_{32}^i - h_{31}^i h_{31}^i \end{aligned}\right\} \tag{5}$$

where

$$\begin{aligned} a_j &= h_{11}^i h_{12}^i - u_{0i}(h_{31}^i h_{12}^i + h_{11}^i h_{32}^i) + u_{0i}^2 h_{31}^i h_{32}^i \\ b_j &= h_{21}^i h_{22}^i - v_{0i}(h_{31}^i h_{22}^i + h_{21}^i h_{32}^i) + v_{0i}^2 h_{31}^i h_{32}^i \\ c_j &= (h_{11}^i - u_{0i} h_{31}^i)^2 - (h_{12}^i - u_{0i} h_{32}^i)^2 \\ d_j &= (h_{21}^i - v_{0i} h_{31}^i)^2 - (h_{22}^i - v_{0i} h_{32}^i)^2 \end{aligned}$$

Stacking (5) with all model planes, we have $\boldsymbol{A}\dfrac{1}{\beta_i^2} = \boldsymbol{b}$, where $\boldsymbol{A}$ and $\boldsymbol{b}$ are both $2J \times 1$ vectors. Finally we can obtain $\beta_i = \dfrac{\| \boldsymbol{A} \|}{\sqrt{\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}}}$, $\alpha_i = \dfrac{\beta_i}{k_i}$.

2.3 Imposing consistency constraint in 3-space geometry

2.3.1 Extrinsic geometries on single respective view

After the greedy estimation procedure for the intrinsic parameters as described above, let us start to estimate the extrinsic parameters between both cameras in a physical perspective. Binocular cameras in a stereo setup are not independent, since they view the same scene from different positions in 3-space geometry. Usually this connection is exploited to construct a consistency constraint in rigid displacement between binocular cameras. First the extrinsic parameters of model planes in 3-space are calibrated in their own world coordinate system for both cameras. Once $\boldsymbol{K}_i$ is known, the initial extrinsic parameters $\boldsymbol{R}_i$ and $\boldsymbol{t}_i$ for each camera reference $I_i (i = 1, 2)$ are readily computed[1]. We have $\boldsymbol{r}_1^i = \lambda \boldsymbol{K}_i^{-1}\boldsymbol{h}_1^i$, $\boldsymbol{r}_2^i = \lambda \boldsymbol{K}_i^{-1}\boldsymbol{h}_2^i$, $\boldsymbol{r}_3^i = \boldsymbol{r}_1^i \times \boldsymbol{r}_2^i$, $\boldsymbol{t}_i = \lambda \boldsymbol{K}_i^{-1}\boldsymbol{h}_3^i$ with $\lambda = \dfrac{1}{\| \boldsymbol{K}_i^{-1}\boldsymbol{h}_1^i \|} = \dfrac{1}{\| \boldsymbol{K}_i^{-1}\boldsymbol{h}_2^i \|}$.

2.3.2 Rigid displacement between two views

Let $\boldsymbol{X}$ and $\boldsymbol{X}_i (i = 1, 2)$ be the coordinates of a 3-space point in π^j and $I_i (i = 1, 2)$, respectively, we can have

$$\boldsymbol{X}_i = \boldsymbol{R}_i \boldsymbol{X} + \boldsymbol{t}_i \qquad i = 1, 2 \tag{6}$$

Note that $\boldsymbol{R}_i^{-1} = \boldsymbol{R}_i^{\mathrm{T}}$. From Eq. (6), the relationship be-

tween $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ can be given by

$$\boldsymbol{R} = \boldsymbol{R}_1\boldsymbol{R}_2^{\mathrm{T}},\ \ \boldsymbol{t} = \boldsymbol{t}_1 - \boldsymbol{R}_1\boldsymbol{R}_2^{\mathrm{T}}\boldsymbol{t}_2 \tag{7}$$

Eq. (7) can be viewed as a rigid transformation from I_1 to I_2[6].

When $J(\geqslant 2)$ model planes can be observed by C_1 and C_2 at the same time, the transformation calculated by Eq. (7) for each model plane from I_1 to I_2 might be a little different due to noise. In this case, each $\boldsymbol{R}$ can be expressed as a 3-vector $\boldsymbol{r} = [\,r_x\ \ r_y\ \ r_z\,]^{\mathrm{T}}$, which represents the rotation axis and the magnitude is equal to the rotation angle $\theta = \|\boldsymbol{r}\|$, by the Rodrigues formula[9] as follows:

$$\boldsymbol{R} = \mathrm{e}^{[\boldsymbol{r}]_\times} = \boldsymbol{I}_3 + \frac{\sin\theta}{\theta}[\boldsymbol{r}]_\times + \frac{1-\cos\theta}{\theta^2}[\boldsymbol{r}]_\times^2 \tag{8}$$

where $[\boldsymbol{r}]_\times = \begin{bmatrix} 0 & -r_z & r_y \\ r_z & 0 & -r_x \\ -r_y & r_x & 0 \end{bmatrix}$ is an anti-symmetric matrix and $\boldsymbol{I}_3$ is a 3×3 identity matrix. The means of $\boldsymbol{r}$, θ and $\boldsymbol{t}$ can be used to estimate $[\boldsymbol{R}\ \ \boldsymbol{t}]$ uniquely.

3 Implementation and Experimental Results

First, the proposed algorithm is presented or implemented in a greedy strategy in detail. Then, the performance of this algorithm is tested through both simulation and experiments with real data on the basis of the spatial configuration displayed in Fig. 1.

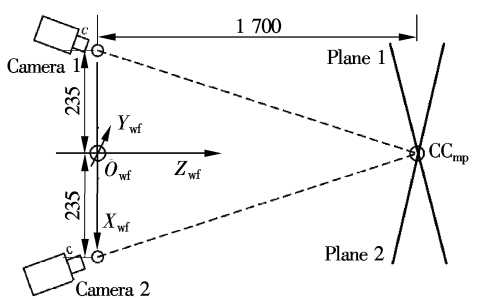


Fig. 1 Spatial configuration for binocular camera system (unit: mm)

3.1 Implementation issues

The greedy algorithm follows the problem solving metaheuristic of making the locally optimum choice at each stage with the hope of finding the global optimum. Applying the greedy strategy to the problem of calibrating the binocular camera system yields the following descriptive method with six steps in turn:

Step 1 Estimate accurately the homography matrices for all model planes with different poses w. r. t. both view cameras;

Step 2 Construct characteristic lines from the planar homography matrices just obtained;

Step 3 Calculate robust principal points for both view cameras from the achieved characteristic lines via total least-square fitting, respectively;

Step 4 Compute accurate principal lengths for both view cameras from reshaping planar homographies, respectively;

Step 5 Generate precisely extrinsic geometries on the given single view by means of the computed initial parameters and the planar homographies, respectively;

Step 6 Compute accurate rigid displacement between the two views by combining the extrinsic geometries of one view camera with that of the other one.

An advantage of this proposed technique is that the intrinsic and extrinsic parameters in the binocular camera system can be obtained greedily at different stages under the given linear assumption.

3.2 Computer simulation results

Simulation experiments are performed in a stereo system as shown in Fig. 1. Two cameras C_1 and C_2 are symmetrically placed along a straight line X_{wf} which passes through the origin O_{wf} of the world reference system at 235 mm intervals, and simultaneity are oriented toward the frontage and focused on to the same fixation point CC_{mp} with the coordinate $(0, 0, 1\,700)$ (mm) in the world reference system. Simulations have been achieved using a chain of stereo images, with 640×480 resolutions, taken from different positions and orientations for both cameras. The constant intrinsic parameters can be set to common values: $\alpha_i = \beta_i = 1\,194.26$, $u_{0i} = 319.5$, $v_{0i} = 239.5$, $k_i = 1$ and $\gamma_i = 0$. Each model plane is painted with $18 \times 26 = 468$ reference points at 14 mm intervals. The center of each model plane is fixed at CC_{mp}, and each pose of the model plane can be represented by different Eulerian angles (i. e. the pitch angle θ_{wx}, the yaw angle θ_{wy} and the roll angle θ_{wz}) from the given plane to the world reference system. Different poses of planar patterns are selected symmetrically by the strategy proposed in Ref. [8].

We use six model planes by changing θ_{wy} ($-30°$, $-25°$, $-20°$, $20°$, $25°$, $30°$), $\theta_{wx} = -25°$, $\theta_{wz} = 0°$. The Gaussian noise with zero-mean and standard deviation σ are added to the projected image points. The estimated camera parameters are then compared with the ground truth. We measure the relative error of α_i and β_i, and the absolute error of u_{0i} and v_{0i} for the internal geometry of each camera; while we measure the absolute error of the rotation angles $\boldsymbol{\theta}$ and the translation $\boldsymbol{t}$ between both cameras. The noise level is varied from 0.1 pixels to 1.5 pixels for which 500 simulations are done. The results shown in Fig. 2 are the average compared with the closed-form solution from Zhang[1]. As we can see, our approach is obviously better in performance w. r. t. the noise level than Zhang's method with only the closed-form. The reason for the large amount of fluctuation of the latter method is that Zhang's method with only the closed-form[1] may be purely based on minimizing an algebraic distance which is not physically meaningful for all intrinsic parameters. However, our method is exploited to obtain each intrinsic parameter with optimal solutions via minimizing some physically meaningful distances by stages in a geometric and intuitive perspective with part prior knowledge.

3.3 Results with real data

The proposed algorithm has been tested on real data as well. The binocular system comprised of two DH-HV3102UC CCD cameras with lens of 8 mm focal length is calibrated. Both cameras have resolutions of $2\,048 \times 1\,536$, with a unit cell size of the CCD sensor being $3.2\ \mu\mathrm{m} \times 3.2$

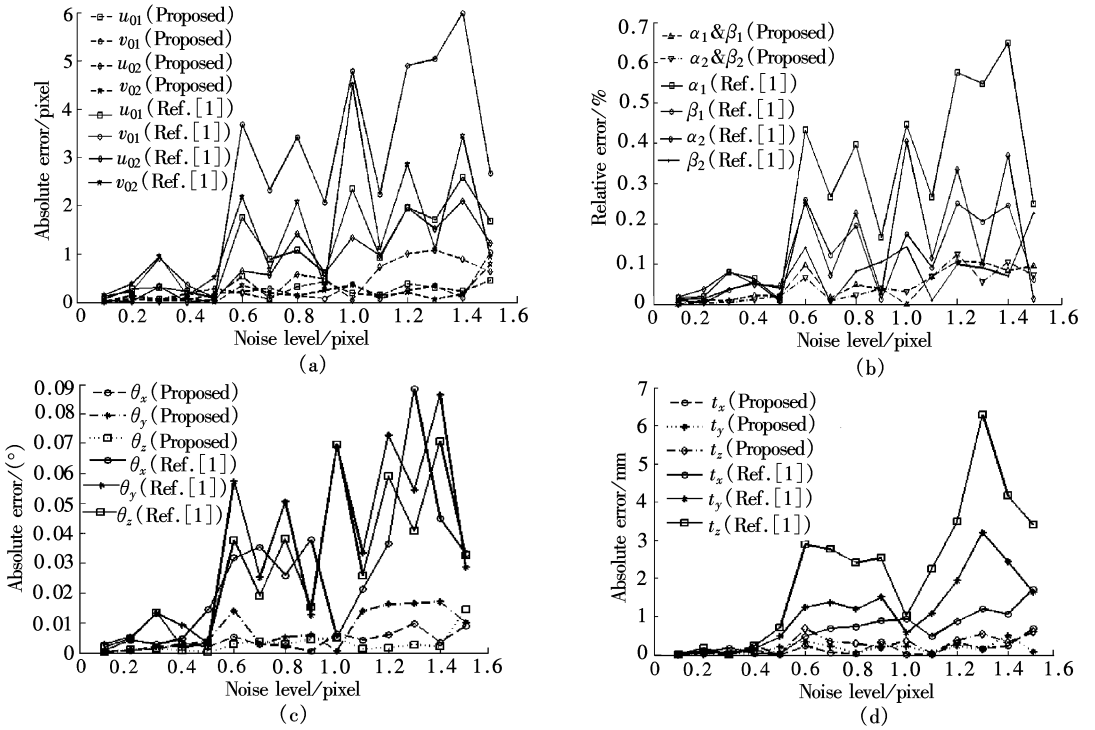


Fig. 2 Errors vs. noise level of simulated image points. (a) Principal point; (b) Principal length; (c) Rotation angle; (d) Translation

μm. Model plane with $8 \times 6 = 48$ grid points at 30 mm intervals is roughly placed like Fig. 1, while changing θ_{wy} about ($-30°$, $-25°$, $-20°$, $-15°$, $15°$, $20°$, $25°$, $30°$), $\theta_{wx} = -20°$, $\theta_{wz} = 0°$ to obtain eight different poses. The estimated intrinsic and extrinsic parameters with respect to all combinations of seven images from the available eight images are listed in Tab. 1 and Tab. 2, respectively. And some statistical results shown in Tab. 3 and Tab. 4 are compared in the same real data with that of the closed-form solution of Zhang's method with or without the generic prior constraint of the aspect ratio of pixel cells[1]. Since ground truth is not available, in order to validate our parameters we use the sample mean μ as the ground truth with sample standard deviation σ; and we also show the relative difference w. r. t. the mean of principal length $f_i (f_i = \alpha_i = \beta_i)$ and the absolute difference w. r. t. the mean of the principal point u_{0i}, v_{0i}, the rotation angles $\boldsymbol{\theta}$, and the translation $\boldsymbol{t}$. As shown in Tab. 1 and Tab. 2, the overall results except for the last term in Tab. 2 are remarkably reliable. The reason for the term t_z with a bit large variation may be that there exist few constraints in the z-direction. As we can see in Tab. 3 and Tab. 4, our results excluding the first two terms in Tab. 3 are in total comparably steady in the view of the sample standard variation, and, on the whole, the results are also very close in the performance of the sample mean when comparing the closed-form solution of Zhang's method with or without the prior knowledge of the aspect ratio[1]. An epipolar line rectification method is performed to validate our calibration results[11]. From Fig. 3, we can see that the epipolar lines of the image pair, which are marked with horizontal lines for the model plane in 3-space with orientation about $\theta_{wx} = -20°$, $\theta_{wy} = 30°$, $\theta_{wz} = 0°$, are aligned in the row after rectification with the estimated mean system parameters.

Tab. 1 Experimental results on intrinsic parameters for real data pixel

Image set	f_1 (Relative error/%)	u_{01} (Absolute error)	v_{01} (Absolute error)	f_2 (Relative error/%)	u_{02} (Absolute error)	v_{02} (Absolute error)
(1 2 3 4 5 6 7)	3 743. 18	972. 01	810. 88	3 764. 97	1 086. 57	812. 18
	(−0. 10)	(−4. 18)	(4. 98)	(0. 05)	(−2. 19)	(5. 92)
(1 2 3 4 5 6 8)	3 747. 73	974. 08	807. 92	3 771. 17	1 088. 86	806. 27
	(0. 03)	(−2. 10)	(2. 02)	(0. 21)	(0. 10)	(0. 01)
(1 2 3 4 5 7 8)	3 752. 25	976. 57	805. 61	3 767. 30	1 088. 80	806. 39
	(0. 15)	(0. 38)	(−0. 29)	(0. 11)	(0. 04)	(0. 12)
(1 2 3 4 6 7 8)	3 758. 99	983. 38	802. 35	3 765. 36	1 089. 56	805. 29
	(0. 33)	(7. 19)	(−3. 55)	(0. 06)	(0. 80)	(−0. 98)
(1 2 3 5 6 7 8)	3 745. 64	977. 37	807. 27	3 752. 41	1 094. 27	808. 48
	(−0. 03)	(1. 18)	(1. 37)	(−0. 29)	(5. 51)	(2. 21)
(1 2 4 5 6 7 8)	3 742. 61	974. 91	803. 89	3 753. 60	1 093. 48	809. 39
	(−0. 11)	(−1. 27)	(−2. 01)	(−0. 25)	(4. 72)	(3. 12)
(1 3 4 5 6 7 8)	3 743. 68	975. 46	804. 48	3 760. 16	1 087. 67	805. 22
	(−0. 08)	(−0. 73)	(−1. 42)	(−0. 08)	(−1. 09)	(−1. 05)
(2 3 4 5 6 7 8)	3 740. 05	975. 70	804. 77	3 770. 24	1 080. 95	796. 90
	(−0. 18)	(−0. 49)	(−1. 13)	(0. 19)	(−7. 81)	(−9. 37)
μ	3 746. 76	976. 19	805. 90	3 763. 15	1 088. 76	806. 27
σ	6. 17	3. 32	2. 70	7. 13	4. 14	4. 46

Tab. 2 Experimental results on extrinsic parameters for real data

Image set	Rotation angle/(°)			Translation/mm		
	θ_x	θ_y	θ_z	t_x	t_y	t_z
(1 2 3 4 5 6 7)	0.652 (−0.007)	−17.951 (0.024)	0.340 (−0.025)	453.67 (0.47)	−10.14 (0.32)	42.48 (−2.92)
(1 2 3 4 5 6 8)	0.689 (0.030)	−17.955 (0.020)	0.363 (−0.002)	454.72 (1.52)	−10.42 (0.04)	42.18 (−3.22)
(1 2 3 4 5 7 8)	0.642 (−0.017)	−17.974 (0.001)	0.354 (−0.011)	453.60 (0.40)	−10.70 (−0.24)	45.81 (0.41)
(1 2 3 4 6 7 8)	0.562 (−0.098)	−18.022 (−0.047)	0.347 (−0.018)	452.01 (−1.19)	−12.04 (−1.58)	49.78 (4.38)
(1 2 3 5 6 7 8)	0.622 (−0.037)	−17.937 (0.038)	0.362 (−0.003)	452.53 (−0.66)	−11.16 (−0.70)	49.77 (4.37)
(1 2 4 5 6 7 8)	0.600 (−0.059)	−17.930 (0.045)	0.360 (−0.005)	453.21 (0.01)	−9.84 (0.62)	48.09 (2.69)
(1 3 4 5 6 7 8)	0.682 (0.023)	−17.980 (−0.005)	0.378 (0.013)	452.72 (−0.47)	−9.69 (0.77)	45.54 (0.14)
(2 3 4 5 6 7 8)	0.819 (0.160)	−18.049 (−0.074)	0.415 (0.050)	453.11 (−0.09)	−9.70 (0.76)	39.51 (−5.89)
μ	0.659	−17.975	0.365	453.20	−10.46	45.40
σ	0.077	0.042	0.023	0.83	0.82	3.77

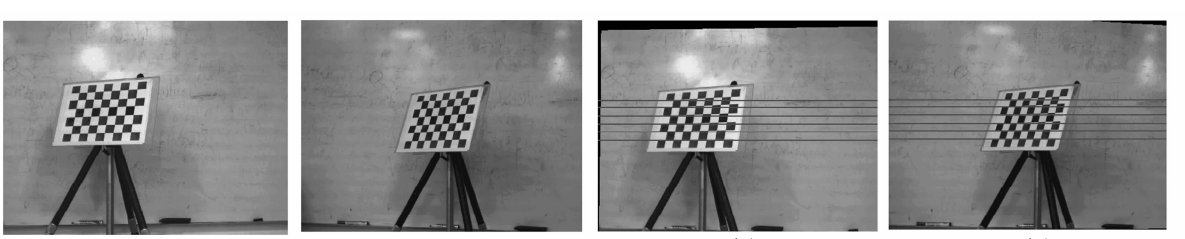
Note: Each term in brackets denotes the absolute error related to the current parameter.

Tab. 3 Comparing intrinsic parameters of our results with that of the closed-form solutions[1] with or without assumed aspect ratios for the same real data pixel

Parameters		α_1	β_1	u_{01}	v_{01}	α_2	β_2	u_{02}	v_{02}
Our results	μ	3746.76	3 746.76	976.19	805.90	3 763.15	3 763.15	1 088.76	806.27
	σ	6.17	6.17	3.32	2.70	7.13	7.13	4.14	4.46
Zhang's[1] ($k_i \neq 1$)	μ	3 758.00	3 755.90	991.74	787.59	3 748.50	3 746.70	1 108.75	821.53
	σ	5.55	3.44	1.66	4.13	17.05	11.58	7.13	15.89
Zhang's[1] ($k_i = 1$)	μ	3 757.00	3 757.00	991.74	787.59	3 747.60	3 747.60	1 108.75	821.53
	σ	4.49	4.49	1.66	4.13	14.31	14.31	7.13	15.89

Tab. 4 Comparing extrinsic parameters of our results with that of the closed-form solutions[1] with or without assumed aspect ratios for the same real data

Parameters		Rotation angle/(°)			Translation/mm		
		θ_x	θ_y	θ_z	t_x	t_y	t_z
Our results	μ	0.659	−17.975	0.365	453.20	−10.46	45.40
	σ	0.077	0.042	0.023	0.83	0.82	3.77
Zhang's[1] ($k_i \neq 1$)	μ	0.471	−17.956	0.332	451.86	−0.718	60.90
	σ	0.140	0.047	0.065	1.24	3.41	6.98
Zhang's[1] ($k_i = 1$)	μ	0.470	−17.953	0.330	451.73	−0.733	60.86
	σ	0.155	0.043	0.064	1.27	3.08	6.98



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

Fig. 3 The epipolar line rectification results for the image pair in the binocular cameras. (a) Left image; (b) Right image; (c) Rectified left image; (d) Rectified right image

4 Conclusion

A new plane-based camera calibration method, simple and intuitive, is presented in a greedy framework for the binocular camera system in close range. The calibration procedure relies on an excellent characteristic of the planar homography matrix and on the assumed constraint with generic prior knowledge about the unit cell size of the CCD sen-

sor, as well as the inherent consistency constraints in 3-space geometry. And several predefined parameters, which represent the intrinsic and extrinsic geometries in such a stereo system, are estimated optimally step by step in a linear optimization framework via minimizing the geometrical or physical cost functions. Both computer simulation and real data are exploited to test this approach, and reasonably satisfactory results are obtained. Therefore, the given algorithm is sufficient for applications where high accuracy is not required and can be easily performed by common computer users who are not experts in computer vision, when compared with the traditional algorithm. In our future work, we plan to observe the technique with lens distortion or global optimization for obtaining more reliable results.

References

[1] Zhang Zhengyou. A flexible new technique for camera calibration [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, **22**(11): 1330 - 1334.

[2] Sturm P F, Maybank S J. On plane-based camera calibration: a general algorithm, singularities, applications [C]// *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Fort Collins, CO, USA, 1999: 432 - 437.

[3] Ueshiba T, Tomita F. Plane-based calibration algorithm for multi-camera systems via factorization of homography matrices [C]//*Proc of the* 9*th IEEE International Conference on Computer Vision*. Nice, France, 2003: 966 - 973.

[4] Menudet J F, Becker J M, Fournel T, et al. Plane-based camera self-calibration by metric rectification of images [J]. *Image and Vision Computing*, 2008, **26**(7): 913 - 934.

[5] Lucchese L. Geometric calibration of digital cameras through multi-view rectification [J]. *Image and Vision Computing*, 2005, **23**(5): 517 - 539.

[6] Zhang Liuxin, Li Bin, Jia Yunde. A plane-based calibration for multi-camera systems [C]//*Proc of the* 10*th IEEE International Conference on Computer Aided Design and Computer Graphics*. Beijing, China, 2007: 365 - 370.

[7] Gurdjos P, Crouzil A, Payrissat R. Another way of looking at plane-based calibration: the centre circle constraint [C]// *Proc of the* 7*th European Conference on Computer Vision*. Copenhagen, Denmark, 2002: 252 - 266.

[8] Wang Jianhua, Liu Yuncai. Characteristic line of planar homography matrix and its applications in camera calibration [C]//*Proc of the* 18*th International Conference on Pattern Recognition*. Hong Kong, China, 2006: 147 - 150.

[9] Hartley R, Zisserman A. *Multiple view geometry in computer vision* [M]. Cambridge: Cambridge University Press, 2004: 88 - 109.

[10] Forsyth D A, Ponce J. *Computer vision*: *a modern approach* [M]. Pearson Education International, 2003: 333 - 335.

[11] An Luping, Jia Yunde, Wang Jing, et al. An efficient rectification method for trinocular stereovision [C]//*Proc of the* 17*th International Conference on Pattern Recognition*. Cambridge, England, 2004: 56 - 59.

基于特征单应矩阵的双目摄像机系统贪婪标定方法

刘天亮　金飞逸　罗立民

(东南大学影像科学与技术实验室,南京 210096)

摘要:在贪婪及直觉的框架下,针对双目摄像机系统提出一种基于平面模板且未考虑光学畸变的线性标定方法.利用单应矩阵的特性及近距离的双目一致性约束进行标定.首先,为计算双摄像机内部几何特性,根据所有模板平面各自的单应矩阵具有特征线的属性,构造具有几何意义的成本函数;采用鲁棒的总误差最小二乘拟合策略,恢复精确的主点坐标.其次,引入像素单元纵横比的先验信息,求取精确的主轴长.然后,精确地获取所有平面模板关于各自相机的外部几何.最后,利用强加于三维几何空间中的一致性约束来计算双目摄像机之间的刚体变换关系.仿真及真实图像实验表明,所提算法能获得较可靠的标定结果,满足精度要求不是很高的应用需求,且对计算机视觉不是很内行的普通用户,能较容易地实现.

关键词:近距离;标定;单应矩阵;一致性约束;立体视觉

中图分类号:TP391.41