ACCV2002: The 5th Asian Conference on Computer Vision, 23--25 January 2002, Melbourne, Australia.

A New Linear Camera Self-Calibration Technique

Hua Li Huaifeng Zhang Fuchao Wu and Zhanyi Hu
National Laboratory of Pattern Recognition,
Institute of Automation, Chinese Academy of Sciences
P.O. Box 2728, Beijing 100080, P.R. China
Email:{hfzhang, fcwu, huzy}@nlpr.ia.ac.cn

Abstract

In this paper, a new active vision based camera self-calibration technique is proposed. The novelty of this new technique is that it can determine LINEARLY all the FIVE intrinsic parameters of a camera. The basic principle of our new calibration technique is to use the planar information in the scene and to control the camera to undergo several sets of orthogonal planar motions. Then, a set of linear constraints on the 5 intrinsic parameters is derived by means of planar homographies between images. In addition, the uniqueness of the calibration solution with respect to the configurations of the camera's motion is also investigated.

Keywords: Camera Self-Calibration, Active Vision, Homography

1 Introduction

Camera calibration is an indispensable step to obtain 3D geometric information from 2D images. With the traditional calibration method, the camera's intrinsic parameters are computed from projected images of a well structured object, called calibration grid. However, in many practical applications, a calibration grid is neither available nor desirable, thus people turned to a new paradigm, called 'self-calibration', i.e., calibration without calibration grid. Since the pioneer works in [1,2], many similar techniques have been reported in the literature [3-15]. However, almost all such techniques have to solve some nonlinear equations, which inevitably leads to either low computational speed or non-convergence. In order to overcome this difficulty, some researchers explored the possibility to constrain the camera to undergo some specially designed motions [16-22]. Ma [21] proposed an active vision based linear calibration method. In Ma's method, two different sets of camera motions, each one of which consists of 3 mutually orthogonal translations, are used to linearly determine the camera's intrinsic parameters. Yang et al. [22] improved Ma's method. In their new method, rather than two sets of 3 mutually orthogonal motions, four sets of two orthogonal planar camera motions are used. In both Ma's methods and Yang's, only 4 intrinsic parameters of camera can be linearly determined. If a full perspective camera model is used, in other words, if the skew factor is non-zero, both of their methods become invalid. In this paper, we propose a new active vision based camera calibration technique which can compute all the 5 intrinsic parameters linearly. In our new method, the planar information in the scene is used, and the camera undergoes N (N>=2) sets of three mutually orthogonal motions or N (N>=5) sets of two orthogonal planar motions.

The organization of the paper is as follows: In section 2, homographies associated with scene planes between two images are discussed. The linear constraints on camera's intrinsic parameters and the uniqueness of solution with respect to configurations of camera motions are elaborated in section 3. A new camera calibration algorithm is outlined in section 4. The experiments on simulated images and real images are reported in section 5 and section 6 respectively. Finally some conclusions are given in section 7.

2 Homography Associated with a Scene Plane Between Two Images

2.1 Camera Model

Here a full perspective camera model is assumed, then the camera intrinsic parameters matrix is

$$\boldsymbol{K} = \begin{bmatrix} f_u & s & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

where (u_0, v_0) is the principal point, f_u, f_v the focal lengths in u and v axis respectively, s the skew-factor.

2.2 Homography of a Plane between Two Images

Assuming $\boldsymbol{m} = (u, v, 1)^T$, $\boldsymbol{m}' = (u', v', 1)^T$ are the homogeneous coordinates of two corresponding points in two images. If these two corresponding points are projected from a same scene point lying on a plane $\boldsymbol{\pi}$, the following relation holds:

$$s\boldsymbol{m}' = \boldsymbol{H}\boldsymbol{m} \tag{1}$$

where matrix $\boldsymbol{H}$ is the homography between the two images induced by the plane $\boldsymbol{\pi}$, and s an unknown non-zero factor. In other words, the homography is determined up to a non-zero scale factor.

Now suppose the plane $\boldsymbol{\pi}$ in 3D space is defined as: $\vec{\boldsymbol{n}}^T \boldsymbol{x} = d$, where $\vec{\boldsymbol{n}}$ is the unit normal vector of $\boldsymbol{\pi}$, d the distance from the origin of the world coordinate system to plane $\boldsymbol{\pi}$. Assume that the world coordinate system coincides with the first camera coordinate system, then for the first image, we have:

$$\lambda \boldsymbol{m} = \boldsymbol{K}\boldsymbol{x}$$

Assuming the transformation from the first camera coordinate system to the second one is $\boldsymbol{x}' = \boldsymbol{R}\boldsymbol{x} + \boldsymbol{t}$, the corresponding point in the second image can be expressed as:

$$\begin{aligned}\lambda' \boldsymbol{m}' &= \boldsymbol{K}\boldsymbol{x}' = \boldsymbol{K}\boldsymbol{R}\boldsymbol{x} + \boldsymbol{K}\boldsymbol{t} = \boldsymbol{K}\boldsymbol{R}\boldsymbol{x} + \frac{1}{d}\boldsymbol{K}\boldsymbol{t}\vec{\boldsymbol{n}}^T \boldsymbol{x} \\ &= \lambda(\boldsymbol{K}\boldsymbol{R}\boldsymbol{K}^{-1} + \boldsymbol{K}\frac{\boldsymbol{t}\vec{\boldsymbol{n}}^T}{d}\boldsymbol{K}^{-1})\boldsymbol{m}\end{aligned}$$

From (1), the homography between these two images is:

$$\boldsymbol{H} = \sigma(\boldsymbol{K}\boldsymbol{R}\boldsymbol{K}^{-1} + \boldsymbol{K}\frac{\boldsymbol{t}\vec{\boldsymbol{n}}^T}{d}\boldsymbol{K}^{-1}) \tag{2}$$

in (2), σ is an unknown non-zero factor. If the camera only undergoes a pure translation, the homography becomes:

$$\boldsymbol{H} = \sigma(\boldsymbol{I} + \boldsymbol{K}\frac{\boldsymbol{t}\vec{\boldsymbol{n}}^T}{d}\boldsymbol{K}^{-1}) \tag{3}$$

2.3 Homography Calculation and the Associated Constant Factor Determination

Since a homography can only be determined up to a scale, we can generally eliminate the scaling effect by assuming the homography has the following form:

$$\boldsymbol{H} = \begin{pmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & 1 \end{pmatrix}$$

Then $\boldsymbol{H}$ can be written as a column vector: $\overline{\boldsymbol{h}} = (h_1, h_2, h_3, h_4, h_5, h_6, h_7, h_8)^T$. From (1), a pair of corresponding points $\boldsymbol{m} = (u, v, 1)^T$, $\boldsymbol{m}' = (u', v', 1)^T$ can bring out two linear constraints on $\overline{\boldsymbol{h}}$,

$$\begin{aligned}(u, v, 1, 0, 0, 0, u'u, u'v)\overline{\boldsymbol{h}} &= u' \\ (0, 0, 0, u, v, 1, v'u, v'v)\overline{\boldsymbol{h}} &= v'\end{aligned}$$

With at least 4 pairs of corresponding points, $\boldsymbol{H}$ can be determined.

If the camera undergoes pure translations, then from (3), there exists a unique factor σ such that:

$$\boldsymbol{H} = \sigma(\boldsymbol{I} + \boldsymbol{K}\frac{\boldsymbol{t}\vec{\boldsymbol{n}}^T}{d}\boldsymbol{K}^{-1}) \tag{4}$$

$$\boldsymbol{H} - \sigma\boldsymbol{I} = \sigma\boldsymbol{K}\frac{\boldsymbol{t}\vec{\boldsymbol{n}}^T}{d}\boldsymbol{K}^{-1} \tag{5}$$

Because $rank(\boldsymbol{K}\boldsymbol{t}\vec{\boldsymbol{n}}^T\boldsymbol{K}^{-1}) = 1$, σ must be the solution of the following equations:

$$\begin{cases} det(\boldsymbol{H} - \sigma\boldsymbol{I}) = 0 \\ det([\boldsymbol{H} - \sigma\boldsymbol{I}]_{2\times2}) = 0, \quad [\boldsymbol{H} - \sigma\boldsymbol{I}]_{2\times2} \in \boldsymbol{\Omega}(\boldsymbol{H} - \sigma\boldsymbol{I}) \end{cases} \tag{6}$$

(6) contains 6 linear equations about σ, so a unique solution can be obtained. A least squares solution of these 6 linear equations is used in practice.

3 Linear Constraints and Camera Motion Configurations

3.1 Linear Constraints

If the camera undergoes two planar orthogonal translations $\boldsymbol{t}^{(1)}, \boldsymbol{t}^{(2)}$ i.e. $(\boldsymbol{t}^{(1)})^T\boldsymbol{t}^{(2)} = 0$, assume that $\boldsymbol{H}_1$ is the homography of a scene plane associated with the first translation, and $\boldsymbol{H}_2$ is the homography of a scene plane (either the same plane as in the first translation, or another plane) associated with the second translation, then based on (3), we have:

$$\boldsymbol{H}_1 = \sigma_1(\boldsymbol{I} + \boldsymbol{K}\frac{\boldsymbol{t}^{(1)}\vec{\boldsymbol{n}}_1^{\,T}}{d}\boldsymbol{K}^{-1}) \tag{7}$$

$$\boldsymbol{H}_2 = \sigma_2(\boldsymbol{I} + \boldsymbol{K}\frac{\boldsymbol{t}^{(2)}\vec{\boldsymbol{n}}_2^{\,T}}{d}\boldsymbol{K}^{-1}) \tag{8}$$

so,

$$\boldsymbol{K}^{-1}(\boldsymbol{H}_1 - \sigma_1\boldsymbol{I})\boldsymbol{K} = \frac{\sigma_1}{d}\boldsymbol{t}^{(1)}\vec{\boldsymbol{n}}_1^{\,T} \tag{9}$$

$$\boldsymbol{K}^{-1}(\boldsymbol{H}_2 - \sigma_2\boldsymbol{I})\boldsymbol{K} = \frac{\sigma_2}{d}\boldsymbol{t}^{(2)}\vec{\boldsymbol{n}}_2^{\,T} \tag{10}$$

Multiply (10) by the transpose of (9), then,

$$\begin{aligned}&\boldsymbol{K}^{-1}(\boldsymbol{H}_1^T - \sigma_1\boldsymbol{I})\boldsymbol{K}^{-T}\boldsymbol{K}^{-1}(\boldsymbol{H}_2 - \sigma_2\boldsymbol{I})\boldsymbol{K} \\ &= \frac{\sigma_1\sigma_2}{d^2}\vec{\boldsymbol{n}}_1(\boldsymbol{t}^{(1)})^T\boldsymbol{t}^{(2)}\vec{\boldsymbol{n}}_2^{\,T} = \boldsymbol{0}_{3\times3}\end{aligned}$$

Let

$$\boldsymbol{C} = \boldsymbol{K}^{-T}\boldsymbol{K}^{-1} = \begin{pmatrix} c_1 & c_2 & c_3 \\ c_2 & c_4 & c_5 \\ c_3 & c_5 & c_6 \end{pmatrix}$$

It is a symmetrical positive definite matrix. And the following linear constraint on $\boldsymbol{C}$ can be derived since σ_1, σ_2 can be obtained from $\boldsymbol{H}_1, \boldsymbol{H}_2$ uniquely as shown in the preceding section.

$$(\boldsymbol{H}_1^T - \sigma_1 \boldsymbol{I})\boldsymbol{C}(\boldsymbol{H}_2 - \sigma_2 \boldsymbol{I}) = \boldsymbol{0}_{3\times 3} \qquad (11)$$

is the fundamental constraint introduced in this paper.

Let $\bar{c} = (c_1, c_2, c_3, c_4, c_5, c_6)^T$, (11) can be re-written as

$$\boldsymbol{A}_{9\times 6}\bar{\boldsymbol{c}} = \vec{\boldsymbol{0}}_9 \qquad (12)$$

Although (12) contains 9 linear constraints on $\boldsymbol{C}$, we can easily prove that only one constraint is useful, all the other constraints are dependent. In other words, matrix $\boldsymbol{A}$ is of rank one. Simple illation is as below: From (9) and (10), $(\boldsymbol{H}_1^T - \sigma_1 \boldsymbol{I})$ and $(\boldsymbol{H}_2 - \sigma_2 \boldsymbol{I})$ are of rank one. So equation (11) can produce only one linear constraint on $\boldsymbol{C}$. And we can conclude that (12) can also produce only one linear constraint on $\boldsymbol{C}$. In order to uniquely obtain $\boldsymbol{C}$ in the sense of up to a scale factor, at least 5 constraints as defined in (12) are needed.

3.2 The Configurations of Camera Motions and the Uniqueness of Solution

3.2.1 Two Sets of Three Mutually Orthogonal Translations (TMOT) Each pair of translations among the 3 translations in one TMOT can produce one constraint on $\boldsymbol{C}$. So one TMOT can produce three constraints on $\boldsymbol{C}$. Then at least 2 sets of TMOTs are needed to determine $\boldsymbol{C}$. Concerning the uniqueness of the solution of $\boldsymbol{C}$, we have the following proposition:

Proposition : $\boldsymbol{\Gamma}_i = \{\boldsymbol{t}^{i1} \quad \boldsymbol{t}^{i2} \quad \boldsymbol{t}^{i3}\}$, $i = 1,2$ are two sets of TMOTs, if these two sets are independent, then $\boldsymbol{C}$ can be determined uniquely modulo a scale factor from the following constraints:

$$\begin{cases} (\boldsymbol{H}_{i1} - \sigma_{i1}\boldsymbol{I})^T \boldsymbol{C}(\boldsymbol{H}_{i2} - \sigma_{i2}\boldsymbol{I}) = \boldsymbol{0}_{3\times 3} \\ (\boldsymbol{H}_{i2} - \sigma_{i2}\boldsymbol{I})^T \boldsymbol{C}(\boldsymbol{H}_{i3} - \sigma_{i3}\boldsymbol{I}) = \boldsymbol{0}_{3\times 3}, \qquad i = 1,2 \\ (\boldsymbol{H}_{i3} - \sigma_{i3}\boldsymbol{I})^T \boldsymbol{C}(\boldsymbol{H}_{i1} - \sigma_{i1}\boldsymbol{I}) = \boldsymbol{0}_{3\times 3} \end{cases}$$

where $\boldsymbol{H}_{ij}$ is the homography associated with the j^{th} translation in the i^{th} TMOT.

Here, by "two sets of TMOTs being independent", we mean that all the following 4×3 matrices must be of rank 3,

$[\boldsymbol{t}^{i1} \quad \boldsymbol{t}^{j1} \quad \boldsymbol{t}^{k2} \quad \boldsymbol{t}^{l2}]^T \quad i, j, k, l = 1,2,3; i \neq j, k \neq l$.

In other words, 4 vectors ($\boldsymbol{t}^{i1}, \boldsymbol{t}^{j1}, \boldsymbol{t}^{k2}, \boldsymbol{t}^{l2}$) are not coplanar ones.

Due to the limited space, the proof is omitted here.

3.2.2 Five Sets of Two Planar Orthogonal Translations (TPOT) As shown in the previous section, each set of TPOT can produce one linear constraint on $\boldsymbol{C}$. Hence in general, 5 sets of TPOTs can produce 5 linear constraints on $\boldsymbol{C}$. It is evident that in some cases, 5 TPOTs can not produce 5 independent linear constraints, in other words, the corresponding motion configuration is a degenerated one. Concerning the uniqueness of the solution of $\boldsymbol{C}$ for given five TPOTs, we have the following conjecture:

Conjecture: Among the five motion planes, if no two or more planes are parallel, the matrix $\boldsymbol{C}$ can be determined uniquely modulo a scale factor.

4 Algorithm

Suppose the camera observes a scene plane and the correspondence of image points is established beforehand, then our new self-calibration algorithm is as follows:

(1) Control the camera to undergo N (N>=5) sets of two planar orthogonal translations (or N (N>=2) sets of three mutually orthogonal translations).

(2) Compute the homographies $\boldsymbol{H}_{i1}, \boldsymbol{H}_{i2}$ associated with the scene plane in each set of two planar orthogonal translations.

(3) Determine the scale factor σ_i associated with each $\boldsymbol{H}_i$, as shown in section 2.

(4) Write the linear constraints on $\bar{\boldsymbol{c}}$ in the form: $\boldsymbol{A}\bar{\boldsymbol{c}} = \vec{\boldsymbol{0}}_9$.

(5) Compute the least squares solution for $\boldsymbol{A}\bar{\boldsymbol{c}} = \vec{\boldsymbol{0}}_9$.

(6) Construct $\boldsymbol{C}$, then decompose $\boldsymbol{C}^{-1}$ as: $\boldsymbol{C}^{-1} = \boldsymbol{V}\boldsymbol{V}^T$ by Cholesky factorization, then decompose $\boldsymbol{V}$ as: $\boldsymbol{V} = \boldsymbol{K}\boldsymbol{Q}$ by RQ factorization, and finally normalize $\boldsymbol{K}$ to make $k_{33} = 1$. Then the normalized $\boldsymbol{K}$ is just the matrix of the camera's intrinsic parameters.

5 Experiments with Simulated Images

As shown above, there are two factors which largely affect the performance of our algorithm. These two factors are noise level and the orthogonality of two camera translations within a same motion set. In order to assess their influences, the following experiments have been done.

5.1 Noise Influence

Here the size of simulated images is 1024*1024 pixels, and at each image point, a random noise is added. The camera's setup is: $f_u = 1000, f_v = 1000, \quad s = 0.20$, $u_0 = 0, v_0 = 0$. Noise unit: pixel. 20 points are used for determining $\boldsymbol{H}$. With different magnitude of random noise, the algorithm is run for 100 times, and then means and RMS errors of the intrinsic parameters are computed. The results are shown in Table1 and Table2. From these two tables, we know that with linear increases of noise level, RMS also increases linearly, which is quite satisfactory.

Table 1. Means of the estimated intrinsic parameters at different noise level

Noise	f_u	f_v	s	u_0	v_0
0.1	999.620	999.705	0.535	0.108	0.117
0.2	999.524	999.627	-0.103	0.694	-0.164
0.3	997.178	998.711	-0.145	0.646	-2.838
0.4	1000.014	999.425	1.089	-0.521	0.972
0.6	997.965	998.322	-0.965	1.072	-2.249
0.8	999.079	994.869	1.826	5.065	7.403
1.0	996.211	993.771	-1.928	5.782	4.021
1.5	1025.636	986.579	7.314	19.065	65.364

Table 2. RMSs of the estimated intrinsic parameters at different noise level

Noise	$\lvert\Delta f_u\rvert$	$\lvert\Delta f_v\rvert$	$\lvert\Delta s\rvert$	$\lvert\Delta u_0\rvert$	$\lvert\Delta v_0\rvert$
0.1	4.420	1.419	1.629	1.227	6.690
0.2	11.247	3.839	3.025	3.036	15.045
0.3	16.233	5.129	4.891	4.594	23.082
0.4	19.662	6.172	5.843	5.909	27.659
0.6	31.783	10.069	11.292	8.081	41.883
0.8	45.845	16.051	12.733	12.589	63.160
1.0	51.048	17.803	15.346	15.013	70.404
1.5	82.179	29.546	26.543	31.442	141.292

5.2 Orthogonality Influence

The image size is: 1024*1024 pixels. The camera's setup is: $f_u = 1000, f_v = 1000, s = 0.20, u_0 = 0, v_0 = 0$. In this case, image points do not contain any noise. But the included angle between the two camera translations is allowed to vary at random within a given error bound[1]. At each given error bound and for each given number of motion sets, 100 runs are done. The final results are shown in Table 3 and Table 4. In Table 3 and Table 4, "X-Y" stands for the included angle of the two translations for each planar motion set.

Table 3. Means of the estimated intrinsic parameters at different error bound

X－Y		5 sets	8 sets	10 sets	15 sets
89｜91	f_u	976.234	1001.389	998.881	998.661
	f_v	1018.556	1000.040	999.786	1000.005
	s	-14.591	-7.004	1.510	-0.893
	u_0	-24.988	0.954	-3.237	-0.165
	v_0	26.342	-1.387	0.668	-0.808
88｜92	f_u	975.939	999.004	1005.938	1003.623
	f_v	993.423	998.746	1006.312	1003.812
	s	-23.649	-3.687	-9.525	2.134
	u_0	10.699	9.810	2.975	3.891
	v_0	-21.761	-1.292	10.187	1.484
87｜93	f_u	972.576	995.750	1024.539	999.390
	f_v	994.415	1004.771	1017.879	1004.243
	s	-31.717	-5.566	-19.352	-0.644
	u_0	15.463	5.320	3.799	6.204
	v_0	-3.369	-6.304	17.905	-1.356
86｜94	f_u	944.180	1000.055	1008.938	1001.645
	f_v	1024.595	1014.221	1009.758	1017.907
	s	4.277	-17.269	2.719	-13.755
	u_0	13.534	31.311	-6.051	4.980
	v_0	-55.905	-5.713	-2.246	1.592
85｜95	f_u	928.851	990.974	1041.977	1019.730
	f_v	1008.974	1039.206	1038.543	1032.747
	s	-54.663	-40.048	-28.706	-27.762
	u_0	-42.923	34.950	10.556	21.341
	v_0	-10.703	-13.128	41.784	13.602

Table 4. RMSs of the estimated intrinsic parameters at different error bound

X－Y		5 sets	8 sets	10 sets	15 sets
89｜91	$\lvert\Delta f_u\rvert$	68.675	22.476	14.833	12.223
	$\lvert\Delta f_v\rvert$	95.449	20.111	15.430	12.137
	$\lvert\Delta s\rvert$	89.769	24.096	15.750	12.298
	$\lvert\Delta u_0\rvert$	79.402	20.059	17.572	11.282
	$\lvert\Delta v_0\rvert$	93.204	23.037	15.660	12.774
88｜92	$\lvert\Delta f_u\rvert$	76.504	36.093	28.332	24.695
	$\lvert\Delta f_v\rvert$	74.637	36.193	33.237	19.735
	$\lvert\Delta s\rvert$	76.964	40.716	34.825	25.863
	$\lvert\Delta u_0\rvert$	76.272	37.052	36.785	29.878
	$\lvert\Delta v_0\rvert$	96.948	44.650	32.760	27.553
87｜93	$\lvert\Delta f_u\rvert$	116.353	56.076	51.361	33.932
	$\lvert\Delta f_v\rvert$	117.717	54.311	42.966	37.355
	$\lvert\Delta s\rvert$	137.267	69.961	49.458	40.116
	$\lvert\Delta u_0\rvert$	124.863	63.049	41.249	36.142
	$\lvert\Delta v_0\rvert$	131.662	69.823	50.069	33.646
86｜94	$\lvert\Delta f_u\rvert$	134.559	74.162	62.798	42.071
	$\lvert\Delta f_v\rvert$	157.014	77.104	61.539	54.406
	$\lvert\Delta s\rvert$	157.637	93.322	66.006	58.155
	$\lvert\Delta u_0\rvert$	147.420	101.423	77.876	58.146
	$\lvert\Delta v_0\rvert$	181.170	88.721	81.984	51.644
85｜95	$\lvert\Delta f_u\rvert$	146.971	82.832	88.394	64.677
	$\lvert\Delta f_v\rvert$	193.981	100.698	87.937	77.262
	$\lvert\Delta s\rvert$	186.325	111.342	104.461	68.707
	$\lvert\Delta u_0\rvert$	199.289	98.310	106.240	81.448
	$\lvert\Delta v_0\rvert$	163.040	99.643	94.143	78.437

[1] Remember ideally the two translation should be orthogonal

From these tables, we know that the more the number of the motion sets and the smaller the error bound, the better the estimated results. Hence in order to obtain satisfying results, if possible, the two camera translations within a same motion set should keep as orthogonal as possible and more images should be used. Fortunately, these two conditions can be generally satisfied in practice with an active vision system.

6 Experiments with Real Images

In real image experiments, a CCD camera is used. After the calibration, the calibrated intrinsic parameters are used to reconstruct a calibration grid to verify whether the calibrated parameters are reliable.

6.1 Calibrating the Intrinsic Parameters

Here a 3D scene including a plane is used. The camera undergoes 16 sets of two orthogonal translations, and we get 16 groups of images, each group contains 3 images (one image before translation, and two images after each translation). The image size is 384*288 pixels. One of the image groups is shown in Fig1.

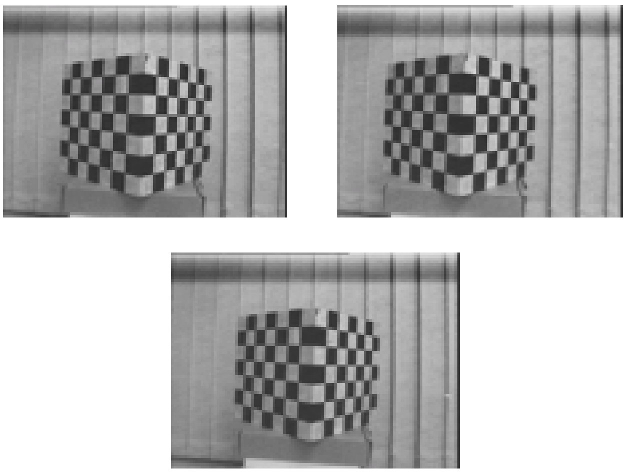


Fig 1 A group of images taken by a CCD camera

With our new algorithm, the calibrated intrinsic parameters of the CCD camera are listed in Table5.

Table 5. The estimated intrinsic parameters

f_u	f_v	u_0	v_0	s
524.6731	256.9008	172.0141	190.6486	-0.8220

6.2 Verification via Reconstruction

Here a standard stereo vision technique is used to reconstruct the 3D scene to verify whether the calibrated intrinsic parameters are reliable. Fig 2 shows a pair of images used for the reconstruction of a calibration grid. The highlighted points are the corresponding points selected from two planes which are orthogonal to each other. Fig 3 are the reconstructed planes with different view directions.

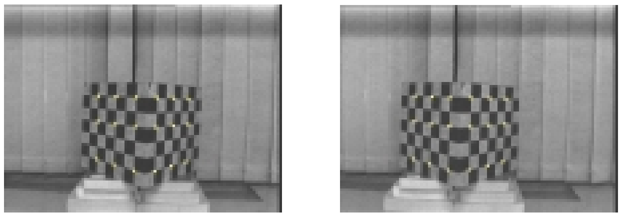
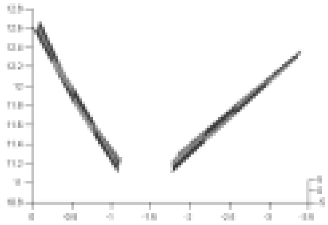
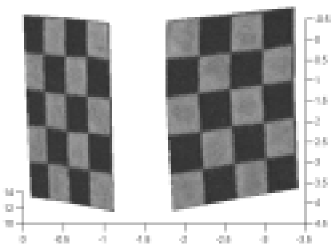


Fig 2: The two images used for the reconstruction



(1)



(2)

Fig 3: Reconstructed two planes with different view

In Fig 3, (1) is the top view, (2) is the side view. The included angle of the two reconstructed planes is 90.45 degrees, which is quite close to its real value of 90 degrees. Based on the fairly good reconstruction results, we can reasonably think that the calibrated intrinsic parameters are reliable. Besides, it is worth noting that the results from our linear calibration technique can be used as the initial values for a non-linear optimization method for further refining.

7. Conclusion

In this paper a new active vision based self-calibration technique is proposed. Our new technique can LINEARLY determine all the FIVE intrinsic parameters. To our knowledge, in the literature there have been no method which can linearly calibrate a full perspective camera. The experiments with simulated data and real images validate our new camera calibration technique.

Acknowledgments

This work was supported by "973" Program (G1998030502) and the National Science Foundation of China (60033010, 69975021).

Reference

[1] R. I. Hartley, Estimation of relative camera positions for uncalibrated cameras, in Proc. ECCV'92, pp.579-387, 1992.

[2] O. D. Faugeras, What can be seen in three dimensions with an uncalibrated stereo rig, in Proc. ECCV'92, pp.563-578, 1992.

[3] S. J. Maybank and O. D. Faugeras, A Theory of Self-Calibration of a Moving Camera, Inter. Journal of Computer Vision 8(2), pp.123-151, 1992.

[4] Q. T. Luong and O. D. Faugeras, The Fundamental Matrix: Theory, Algorithms, and Stability Analysis, Inter. Journal of computer Vision 17, pp.43-75, 1996

[5] O. D. Faugeras et al., 3-D reconstruction of urban scenes from image sequences, Computer Vision and Image Understanding, Vol.69, No. 3, March pp.392-309, 1998.

[6] R. I. Hartley, Self-Calibration of Stationary Cameras, Inter. Journal of Computer Vision 22(1), pp.5-23, 1997.

[7] R. I. Hartley, Line and Points in three Views and the Trifocal Tensor, Inter. Journal of Computer Vision 22(2), pp.125-140, 1997.

[8] R. Koch, M. Pollefeys, and L. Van Gool, Multi viewpoint stereo from uncalibrated video sequences, in Proc. ECCV'98, Vol.I, pp.55-71, 1998.

[9] M. Pollefeys, R. Koch and L. Van Gool, Self-Calibration and Metric Reconstruction in spite of Varying and Unknown Internal Camera Parameters, in Proc. ICCV'98, pp.90-95, 1998.

[10] M. Pollefeys, L. Van Gool, M. Proesmans, Euclidean 3D reconstruction from image sequences with variable focal lengths, in Proc. ECCV'96

[11] A. Shashua, Projective Structure from Uncalibrated Images: Structure From Motion and Recognition, IEEE-Trans. PAMI 16:8, pp.778-790, 1.

[12] S. Avidan and A. Shashua, Threading fundamental matrices, in Proc. ECCV'98, Vol.-I, pp.124-140

[13] A. Heyden, A common framework for multiple view tensors, in Proc. ECCV'98, Vol.I, pp.3-19, 1998.

[14] L. Quan and T. Kanade, Affine structure from line correspondences with uncalibrated affine cameras, IEEE-Trans. PAMI, Vol.19, No.8, pp.834-845, 1997.

[15] B. Triggs, Autocalibration from planar scenes, in Proc. ECCV'98, pp.89-105, 1998

[16] L. Dron. Dynamic Camera Self-Calibration of from Controlled Motion Sequences, in Proc.CVPR'93, pp.501-506,1993.

[17] A. Basu. Active Calibration: Alternative Strategy and Analysis, in Proc.CVPR'93, pp.405-500, 1993

[18] F. Du and M. Brady, Self-calibration of the Intrinsic Parameters of Cameras for Active Vision Systems, in Proc.CVPR'93, pp.477-482, 1993.

[19] F. X. Li, Michael Brady, Charles Wiles, Fast Computation of the Fundamental Matrix for an Active Stereo Vision System, in Proc. ECCV'96, Vol-I. pp.156-166, 1996.

[20] M. X. Li, Camera Calibration of a Head-Eye System for Active Vision, in Proc. ECCV'94, pp.543-554, 1994.

[21] S. D. Ma, A self-calibration technique for active vision systems, IEEE Trans on Robotics and Automation, 12(1): pp.114-120, 1996.

[22] C. J. Yang, W. Wang, and Z. Y. Hu, A active vision based self-calibration technique. Chinese Journal of Computers, Vol 5, pp.428-435, 1998.