

An Intrinsic Parameters Self-Calibration Technique for Active Vision System

Changjiang YANG 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: yangcj@prlsun7.ia.ac.cn, huzy@prlsun5.ia.ac.cn

Abstract

This paper presents a new camera intrinsic parameters self-calibration technique for ordinary active vision system. By controlling a pan-tilt-translation camera platform to do a sequence of specially designed motions (called a camera motion configuration here), we rigorously proved that the camera intrinsic parameters can be determined linearly under such two configurations: (1)regulating the camera's orientation by 3 tilts, at each camera's orientation, controlling the camera to translate twice along 2 orthogonal directions; (2)regulating the camera's orientation by 1 pan and 2 tilts, at each camera's orientation, controlling the camera to translate twice along 2 orthogonal directions. Furthermore, based on extensive simulations of stability analysis, it is shown that the configuration 2 is robust, whereas the configuration 1 is numerically unstable and sensitive to noise. Experiments with real data were carried out and the calibration results have been verified by a stereo vision experiment. A comparison with other camera calibration approaches is also reported here.

1. Introduction

Camera calibration is a preliminary step toward 3D-computer vision. The classical methods use a structure known object (called calibration reference) to determine the camera intrinsic parameters by matching the image features to those on the calibration reference[1-3]. In addition to the involved complicated computational aspects, the main drawback of the classical methods is their requirement of a calibration reference which is not always available in practice. The drawback can be overcome by self-calibration technique, which directly uses the image of the environment instead of the calibration reference.

Since the pioneering work by Maybank and Faugeras [4], a quite number of self-calibrations techniques have been proposed [4-8]. In [5], Ma proposed a technique which allows linear computation of camera's intrinsic parameters by controlling the camera to perform some specially designed motions. Camera self-calibration based on known camera motions is also investigated in [6,7].

The method proposed here is a major improvement to

Ma's method in [5]. Ma's method requires the camera to translate in arbitrary directions in 3D space that an ordinary pan-tilt-translation camera platform cannot perform. Our method requires only the camera being able to pan, tilt, and translate in a horizontal plane, which can be carried out by an ordinary pan-tilt-translation camera platform. In addition, a preliminary assessment on the robustness aspects with respect to the different motion configurations is also presented in our work. Compared to the classical methods in the literature, our method is a linear one, and the uniqueness of solution can be guaranteed. Besides, although our method requires the camera to do some specially designed motions, we do not need to know the parameters of these motions, which is different from those in [6,7].

2. Camera intrinsic parameters from 4 sets of orthogonal translations

In a head-eye active vision system, the camera is rigidly mounted on a pan-tilt-translation platform. The platform can be controlled to pan (rotation about a vertical axis), tilt (rotation about a horizontal axis), or translate in a horizontal plane, and the motion parameters can be read out from the controller.

By controlling the platform to translate twice along 2 orthogonal directions, the 2 corresponding FOE (*focus expansion*) can be detected from the images [5]. Repeat this procedure four times while regulating the camera's orientation by pan or tilt, we can get 4 sets of FOE, F_{i1} , F_{i2} ($i=1\sim 4$), correspondingly 4 sets of the vectors OF_{i1} , OF_{i2} ($i=1\sim 4$). Using the technique in [5], we can obtain following linear equations:

$$\begin{cases} (u_{11} + u_{12} - u_{21} - u_{22})x + (v_{11} + v_{12} - v_{21} - v_{22})y - (v_{11}v_{12} - v_{21}v_{22})z = u_{11}u_{12} - u_{21}u_{22} \\ (u_{21} + u_{22} - u_{31} - u_{32})x + (v_{21} + v_{22} - v_{31} - v_{32})y - (v_{21}v_{22} - v_{31}v_{32})z = u_{21}u_{22} - u_{31}u_{32} \\ (u_{31} + u_{32} - u_{41} - u_{42})x + (v_{31} + v_{32} - v_{41} - v_{42})y - (v_{31}v_{32} - v_{41}v_{42})z = u_{31}u_{32} - u_{41}u_{42} \end{cases} \quad (1)$$

where $x = u_0$, $y = v_0 f_x^2 / f_y^2$, $z = f_x^2 / f_y^2$, (u_{ij}, v_{ij}) are the image coordinates of F_{ij} ($i=1\sim 4, j=1,2$), (u_0, v_0) is the principal point of the camera. We call u_0, v_0, f_x, f_y the camera intrinsic parameters.

The coefficient matrix of Eqs.(1) is

$$\mathbf{A} = \begin{bmatrix} u_{11} + u_{12} - u_{21} - u_{22} & v_{11} + v_{12} - v_{21} - v_{22} & -(v_{11}v_{12} - v_{21}v_{22}) \\ u_{21} + u_{22} - u_{31} - u_{32} & v_{21} + v_{22} - v_{31} - v_{32} & -(v_{21}v_{22} - v_{31}v_{32}) \\ u_{31} + u_{32} - u_{41} - u_{42} & v_{31} + v_{32} - v_{41} - v_{42} & -(v_{31}v_{32} - v_{41}v_{42}) \end{bmatrix}$$

Hence Eqs.(1) has a unique solution if and only if matrix \mathbf{A} is invertible, or equivalently $\det(\mathbf{A}) \neq 0$. About coefficient matrix \mathbf{A} , we have the following proposition:

Proposition 1: A Euclidean transformation in the image plane preserves the orthogonality of the camera's two translation vectors as well as the invertibility of the matrix \mathbf{A} .

A proof is given in [10]. The proposition 1 implies that the orthogonality of the two translation vectors is independent of any specific camera coordinates system.

According to proposition 1, by a Euclidean transformation to the image coordinate system, we can transform the origin of the image system into the principal point and make u axis parallel to the camera's translation plane while preserving the invertibility of matrix \mathbf{A} . In the new image coordinate system, $v_i = v_{i2} (i=1\sim 4)$. For convenience seek, we adopt *normalized coordinate system* of camera[8] in which $f=1$. Denoting $u_i = u_{i1} + u_{i2}$, $v_i = v_{i1} = v_{i2} (i=1\sim 4)$, then matrix \mathbf{A} becomes:

$$\mathbf{A} = \begin{bmatrix} u_1 - u_2 & 2(v_1 - v_2) & -(v_1^2 - v_2^2) \\ u_2 - u_3 & 2(v_2 - v_3) & -(v_2^2 - v_3^2) \\ u_3 - u_4 & 2(v_3 - v_4) & -(v_3^2 - v_4^2) \end{bmatrix}.$$

So the 4 sets of camera translations can be described by the relationship between the optical axis OC and the camera's translation plane OF_1F_2 as shown in Fig.1, where $\theta \in (-\pi/2, \pi/2)$ $\phi \in (0, \pi/2)$ are the camera's tilt and pan angles respectively. The camera's orientation at each set of camera translations can be regulated by changing the tilt or/and pan angles of the camera. There exist 4 possible ways to regulate the camera's orientations for the 4 sets of camera translations, which can be listed respectively as follows:

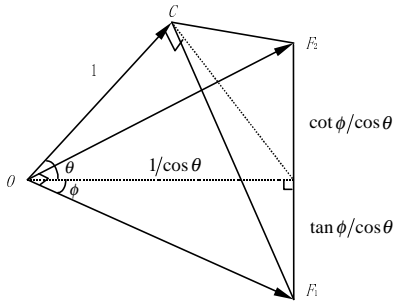


Fig.1: The orientation of the optical axis OC with respect to the camera translation plane in the normalized camera coordinates system. Plane OF_1F_2 is the camera translation plane. OF_1, OF_2 are the two translation directions.

Configuration 1: 3 pans. In this case, $\theta_1 = \theta_2 = \theta_3 = \theta_4$, $\phi_1, \phi_2, \phi_3, \phi_4$ are any 4 distinct angles, we have $v_1 = v_2 = v_3 = v_4 = \tan \theta_1$, then matrix \mathbf{A} becomes

$$\mathbf{A} = \begin{bmatrix} u_1 - u_2 & 0 & 0 \\ u_2 - u_3 & 0 & 0 \\ u_3 - u_4 & 0 & 0 \end{bmatrix}.$$

Evidently, $\det(\mathbf{A})=0$, hence Eqs.(1) has infinitely many solutions. In other words, the camera intrinsic parameters cannot be determined under such a configuration.

Configuration 2: 2 pans + 1 tilt. Without loss of generality, we assume $\theta_1 = \theta_2 = \theta_3 \neq \theta_4$, ϕ_1, ϕ_2, ϕ_3 are any 3 distinct angles, $\phi_3 = \phi_4$, then $v_1 = v_2 = v_3 = \tan \theta_1$, matrix \mathbf{A} becomes:

$$\mathbf{A} = \begin{bmatrix} u_1 - u_2 & 0 & 0 \\ u_2 - u_3 & 0 & 0 \\ u_3 - u_4 & 2(v_3 - v_4) & -(v_3^2 - v_4^2) \end{bmatrix}.$$

Similarly as that under Configuration 1, the camera intrinsic parameters cannot be determined under such a configuration.

Configuration 3: 1 pan + 2 tilts. Without loss of generality, we assume $\theta_1 = \theta_2$, $\theta_1, \theta_3, \theta_4$ are any 3 distinct angles, $\phi_1 \neq \phi_2 = \phi_3 = \phi_4$, then as proved in [10], $\det(\mathbf{A}) \neq 0$, hence the camera intrinsic parameters can be uniquely determined under such a configuration.

Configuration 4: 3 tilts. In this case, $\phi_1 = \phi_2 = \phi_3 = \phi_4$, furthermore, if we make the absolutes of $\theta_1, \theta_2, \theta_3, \theta_4$ to be different from each other, as proved in [10], $\det(\mathbf{A}) \neq 0$, hence the camera intrinsic parameters can also be determined under such a configuration.

3. Numerical analysis of stability

In the preceding section, it is theoretically proved that the camera intrinsic parameters can be uniquely determined under the Configuration 3 or 4. However, theoretical correctness does not mean numerical robustness. In [9], it is illustrated that even a small perturbation of the input data will cause a significant error of the solution if the equations are ill-posed. Since the image noise is inevitable in practice, it is indispensable that the solution of Eqs.(1) to be numerically stable.

To measure the stability of a solution, the condition number of matrix is introduced here. For equations $\mathbf{A}\mathbf{x}=\mathbf{b}$, suppose that $\|\cdot\|$ is a given matrix norm, $\kappa(\mathbf{A}) = \|\mathbf{A}^{-1}\| \|\mathbf{A}\|$ is called the condition number of matrix \mathbf{A} with respect to the matrix norm $\|\cdot\|$. It is obvious that the greater the $\kappa(\mathbf{A})$, the more unstable the solution, and the more ill-posed the equations. When the norm used is the spectral norm, we have

$$\kappa_2(\mathbf{A}) = \|\mathbf{A}^{-1}\|_2 \|\mathbf{A}\|_2 = \sqrt{\lambda_{\max}(\mathbf{A}^T \mathbf{A}) / \lambda_{\min}(\mathbf{A}^T \mathbf{A})} \quad (2)$$

By Eq.(2), we carried out a series of simulations to calculate the condition number of the coefficient matrix \mathbf{A} under Configuration 3 and Configuration 4. Some of the simulation results are presented in Fig.2. From the simulations, we conclude that the equations under Configuration

3 are well-posed as long as the tilt angles θ and pan angles ϕ are not too close to each other, while those under Configuration 4 are ill-posed. The reason why the equations under Configuration 4 are ill-posed could be due to the fact that the elements in the third column of matrix A under Configuration 4 are always far greater than those in the first two columns.

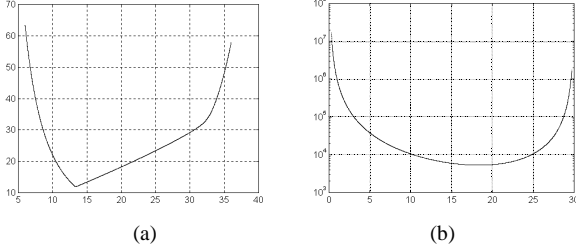


Fig2: Condition number of matrix A to $\Delta\theta$ (degree) under: (a) Configuration 3, when $\theta_3=0$, $\theta_1=\theta_2=\Delta\theta$, $\theta_4=-\Delta\theta$, $\phi_1=0$, $\phi_2=\phi_3=\phi_4=\Delta\theta$; (b) Configuration 4, where $\theta_1=0$, $\theta_2=\theta_1+\Delta\theta$, $\theta_3=\theta_2+\Delta\theta$, $\theta_4=\theta_3+\Delta\theta$, $\phi_1=\phi_2=\phi_3=\phi_4=60^\circ$.

4. Experimental results

We have verified our method under Configuration 3 by experiments with real image data. In our experiments, a CCD camera (resolution 512×512) is mounted on a pan-tilt-translation platform. The origin is set to $x=0$, $y=0$, and $\phi_1 = \phi_2 = \arctan(x/y) = 38.66^\circ$, $\phi_3 = \phi_4 = 51.34^\circ$. The positions (in the image coordinates system) of the 4 sets of FOE points are listed in Table.1. The camera intrinsic parameters obtained by our method are listed in Table.2. Using the method in [5], we calibrated the extrinsic parameters of the camera which are also listed in Table.2.

In order to evaluate the validation of our method, we first compared Tsai's method in [11] with ours, the results are listed in Table.2. Furthermore, we verified the camera parameters calibrated from our method by a stereo vision experiment. Using the camera parameters from our method, we reconstruct a calibration reference which is placed in front of the camera at a distance of about 2.3m. The actual and reconstructed lengths of the 4 edges of the calibration reference are shown together in Table.3. The average error is less than 5mm.

Table.1: The positions (in the image coordinates system) of the 4 set of FOE points

Positions of FOE	Set 1		Set 2		Set 3		Set 4	
	1	2	1	2	1	2	1	2
u_{ij} (pixel)	714.86	-506.27	699.96	-478.77	954.94	-219.60	715.40	-504.63
v_{ij} (pixel)	462.65	464.98	256.20	257.86	257.51	257.76	48.66	51.13

Table.2: Comparison of camera parameters (extrinsic and intrinsic) determined by Tsai' method and ours

Methods	Rotation(degree)	Translation(mm)	Intrinsic parameters			
	roll-pitch-yaw	t_1, t_2, t_3	u_0, v_0	f_x, f_y	k'_1	k
Ours	-1.02, 10.49, 0.14	-62.6, -132.2, 123.1	243.91, 256.03	574.08, 774.36	—	0.741
Tsai	-0.99, 10.76, -0.25	-62.2, -132.8, 124.4	242.37, 256.75	572.31, 773.59	8.192e-04	0.740

Table 3: Comparison of actual lengths and reconstructed lengths of the calibration reference

Actual (mm)	280	280	280	280
Reconstructed (mm)	277.63	282.38	278.20	280.75

5. Conclusions

A novel camera intrinsic parameters self-calibration technique was proposed. We investigated possible motion configurations, and showed that the configuration 3 (1 pan + 2 tilts) is a robust one. Compared with other self-calibrations techniques, The main characteristics of our method are: Firstly, all the computations involved are linear. Secondly, it can be easily realized by a pan-tilt-translation camera platform. Finally, theoretical analysis and experiments with real images showed that our method is robust. A stereo vision experiment also showed that the results obtained by our method are rather faithful and stable.

References

- [1] R.Y.Tsai, An efficient and accurate camera calibration technique for 3D machine vision, In *Proc. CVPR'86*, pp. 364-374, 1986.
- [2] G.Q.Wei and S.D.Ma, Implicit and explicit camera calibration: Theory and experiments, *PAMI*, 16(5): 469-480, 1994.
- [3] O.Faugeras and G.Toscani, The calibration problem for stereo, In *Proc. CVPR'86*, pp.15-20, 1986.
- [4] S.J.Maybank and O.Faugeras, A theory of self-calibration of a moving camera, *IJCV*, 8(2): 123-151, 1992.
- [5] S.D.Ma, A self-calibration technique for active vision systems, *IEEE Robotics and Automation*, 12(1): 114-120, 1996.
- [6] L.Dron, Dynamic Camera Self-Calibration from Controlled Motion Sequences, In *Proc. CVPR'93*, pp.501-506, 1993.
- [7] Fenglei Du and Michael Brady, Self-calibration of the Intrinsic Parameters of Cameras for Active Vision Systems, In *Proc. CVPR'93*, pp.477-482, 1993.
- [8] O.Faugeras, *Three Dimensional Computer Vision: A Geometric Viewpoint*, Cambridge, MA: MIT Press, 1993.
- [9] Roger A. Horn and Charles R. Johnson, *Matrix Analysis*, Cambridge, MA: Cambridge University Press, 1985.
- [10] C.J.Yang, Z.Y.Hu, An intrinsic parameters self-calibration technique for active vision system, Tech. Report, NLP, Inst. of Automation, CAS, P.R.China, 1997.
- [11] R.Y.Tsai, A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses, *IEEE Journal of Robotics and Automation*, Vol. RA-3, No. 4, pp.323-344, 1987.