Self-Calibration Using a Particular Motion of Camera

A. Saaidi, A. Halli, H. Tairi and K. Satori
LIIAN, Department of Mathematics and informatics
Faculty of Sciences Dhar-Mahraz, P.O.Box 1796 Atlas-Fes, MOROCCO
htairi@yahoo.fr

Abstract:- In this article, we are interested in the camera self-calibration from three views of a 3-D scene. The originality of our method resides in the new technique used to estimate the homography of the plane at infinity by the minimization of a non-linear cost function that is based on a particular motion of the camera "translation and small rotation". Our approach also permits to calculate the camera parameters and the depths of interest points detected in the images. Experimental results demonstrate the performance of our algorithms, in term of precision and convergence.

Keywords:- Self-calibration, interest points, matching, homography of the plane at infinity.

1 Introduction

The camera calibration requires the utilization of a grid [2] or of a known scene [3], this constraint is not still present in computer vision applications, what gives new method, say self-calibration, to calculate the camera parameters while replacing the grid or the known scene by an unknown scene.

Generally, the camera self-calibration is achieved by two methods, the first uses any 3-D scene [4, 5, 6] and the second uses a planar scene [7, 8]. In this article we are interested in the first method which is more general and treats more scene than the second. In any case, the camera self-calibration is non-linear and the obtained results depend strongly on the choice of the cost function to minimize and of the initial solution.

Our approach to solve the problem of camera self calibration by a new method that simultaneously estimates the depths of the interest points, the homography of the plane at infinity and the camera parameters. Input data are the interest points between the three images that are detected by Harris algorithm and matched by *ZNCC* correlation measure. Then an initial solution of the problem is calculated while supposing that the camera has a motion "mere translation", this solution permits a non-linear optimization algorithm to minimize a cost function to get satisfactory results.

We will suppose, all along this article, that the camera intrinsic parameters are unknown constants variables.

The paper is organized as follows: The second part presents some essential preliminaries to the understanding of the problem. The third part describes our approach of self-calibration. Experimentations are presented in the fourth part and the conclusion in the fifth part.

2 Preliminaries

2.1 Notations

In this paper we denote by:
$M = (x \quad y \quad z)^T$ a scene point and $m = (u \quad v)^T$ a image point. $M_\infty = (x_\infty \quad y_\infty \quad z_\infty)^T$ a point at infinity and $m_\infty = (u_\infty \quad v_\infty)^T$ his projection in the image plan. $\overline{M} = (x \quad y \quad z \quad 1)^T$, $\overline{m} = (u \quad v \quad 1)^T$, $\overline{M}_\infty = (x_\infty \quad y_\infty \quad z_\infty \quad 0)^T$ and $\overline{m} = (u_\infty \quad v_\infty \quad 1)^T$ the homogeneous coordinates of the points, M, m, M_∞ and m_∞. $O = (0 \quad 0 \quad 0)^T$ the 3 zero vector, I_3 the 3×3 identity matrix.

2.2 Camera Model

We consider that the imaging system is based on the pinhole model. The projection from 3D space to the 2D image plane can be expressed by:

$$\alpha\,\overline{m} = K(R \quad t)\overline{M} \tag{1}$$

with:
$K = \begin{pmatrix} f & \tau = 0 & u_0 \\ 0 & af & v_0 \\ 0 & 0 & 1 \end{pmatrix}$, $(R \quad t)$ is the 3×4 matrix, α a scalar different of zero, M a scene point and m an image point. Elements of the matrix K are the camera's intrinsic parameters that will be determined by the self-calibration procedure. f is the focal length, a is the scale factors, $\overline{M}_i$ are the image coordinates of the principal point and τ is the image skew considered as equal to zero (but this can be relaxed). Elements of the rotation matrix R and the translation vector t are the extrinsic parameters

determining the relative orientation and the position of the camera in 3-D space.

2.3 Vision System

We consider a camera that moves in three positions to project a 3D scene in three images. The space reference is attached to the camera reference in the first position. Therefore a point scene M is projected in three points images m, m' and m'', Fig. 1 by the following relations:

$$\alpha\,\overline{m} = K\begin{pmatrix}I_3 & O\end{pmatrix}\overline{M} \qquad (2)$$

$$\alpha'\,\overline{m'} = K\begin{pmatrix}R & t\end{pmatrix}\overline{M} \qquad (3)$$

$$\alpha''\,\overline{m''} = K\begin{pmatrix}R_1 & t_1\end{pmatrix}\overline{M} \qquad (4)$$

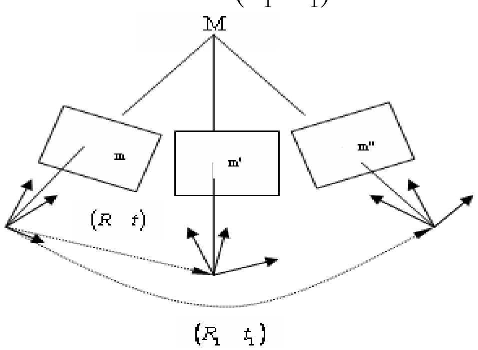

Fig. 1: vision system used.

2.4 Homography of the Plane at Infinity

The projection of a point at infinity M_∞ in the three images is given by:

$$\alpha_\infty\,\overline{m}_\infty = K\begin{pmatrix}I_3 & O\end{pmatrix}\overline{M}_\infty \qquad (5)$$

$$\alpha'_\infty\,\overline{m'}_\infty = K\begin{pmatrix}R & t\end{pmatrix}\overline{M}_\infty \qquad (6)$$

$$\alpha''_\infty\,\overline{m''}_\infty = K\begin{pmatrix}R_1 & t_1\end{pmatrix}\overline{M}_\infty \qquad (7)$$

From the relation (5) and (6) we deduct that:

$$\beta_\infty\,\overline{m'}_\infty = KRK^{-1}\overline{m}_\infty \qquad (8)$$

The matrix KRK^{-1} is called the homography of the plane at infinity between the images 1 and 2, noted H_∞, therefore:

$$H_\infty = KRK^{-1} \qquad (9)$$

The homography of the plane at infinity between the images 1 and 3 is given by:

$$H_{1\infty} = KR_1K^{-1} \qquad (10)$$

2.5 Absolute Conic, its Image and the Dual of its Image

● Absolute conic: In the projective space P^3 all spheres intersect the plan at infinity in a unique set of points $\overline{M}_\infty$, as:

$$M_\infty^T M_\infty = x_\infty^2 + y_\infty^2 + z_\infty^2 = 0 \qquad (11)$$

While avoiding the case $x_\infty = y_\infty = z_\infty = 0$, the previous equation has a solution if the coordinates x_∞, y_∞, z_∞ take some complex values. The set of points verifies the relation (11) is a conical particular of the plan at infinity, called absolute conic, associated to the matrix $\Omega \sim I_3$.

● Image of the absolute conic: According to the relation (6) we deduct that: $M_\infty = \alpha'_\infty R^T K^{-1}\overline{m'}_\infty$, and from the equation (11) we find that:

$$\overline{m'}_\infty^T\left(KK^T\right)^{-1}\overline{m'}_\infty = 0 \qquad (12)$$

Therefore the image of the absolute conic is the conical $\omega = \left(KK^T\right)^{-1}$ that is the same in all views.

● Dual of the image of the absolute conic: The matrix $\Delta = KK^T$ is the dual conic of ω that is a positive definite symmetric matrix and only depends on the camera's intrinsic parameters. The knowledge of Δ is equivalent to the matrix K, and therefore to the camera calibration [9]:

$$\Delta = KK^T = \begin{pmatrix} f^2 + u_0^2 & u_0v_0 & u_0 \\ u_0v_0 & (af)^2 + v_0^2 & v_0 \\ u_0 & v_0 & 1 \end{pmatrix} \qquad (13)$$

2. 6 Fundamental Matrix

The fundamental matrix is the correlation that transforms an image point on the epipolar line corresponding in the other image. The epipolar constraint for two points m and m' written as:

$$\overline{m'}^T F\,\overline{m} = 0 \qquad (14)$$

This equation expresses that m' is on the epipolar line corresponding to m and that is given by $l = F\overline{m}$. The matrix fundamental F can be estimated by a robust algorithm [10].

2.7 Self-Calibration Equations

From the relation (2) we find that $M = \alpha K^{-1}\overline{m}$ and $\alpha = z$ therefore:

$$M = zK^{-1}\overline{m} \qquad (15)$$

From the relation (3) and (15) we find that:

$$\alpha'\overline{m'} = zKRK^{-1}\overline{m} + Kt \qquad (16)$$

We replace the vector Kt by t', according to relations (9) and (16) we find that:

$$\alpha'\,\overline{m'} = zH_\infty\overline{m} + t' \qquad (17)$$

In (17), we put $\overline{m'} = \begin{pmatrix}u' & v' & 1\end{pmatrix}^T$ and we eliminate α', we find the following equations:

$$\frac{\left(zH_\infty\overline{m} + t'\right)_{11}}{u'} = \frac{\left(zH_\infty\overline{m} + t'\right)_{21}}{v'} = \left(zH_\infty\overline{m} + t'\right)_{31} \qquad (18)$$

Among the three precedent equations, two that are independent, we eliminate one, we find that:

$$\begin{cases} \dfrac{\left(zH_\infty \overline{m} + t'\right)_{11}}{u'} = \left(zH_\infty \overline{m} + t'\right)_{31} \\ \dfrac{\left(zH_\infty \overline{m} + t'\right)_{21}}{v'} = \left(zH_\infty \overline{m} + t'\right)_{31} \end{cases} \quad (19)$$

These equations are non-linear, therefore, they require an optimization procedure.

3 Camera Self-Calibration

3.1 Matching

To achieve the camera self-calibration, the matching of certain points is an important stage that is determined in three procedures; the first consists in localizing corners points by the Harris detector, the second aims at detecting the interest points by a correlation measure. Finally, to eliminate the false matching, a part of regularization by the fundamental matrix between the two images is done in the third procedure.

3.1.1 Harris Detector

Harris [11] developed the Moravec method to calculate the local maxima in images by a matrix N:

$$N = \begin{pmatrix} \left(\dfrac{\partial I}{\partial u}\right)^2 & \left(\dfrac{\partial I}{\partial u}\right)\left(\dfrac{\partial I}{\partial v}\right) \\ \left(\dfrac{\partial I}{\partial u}\right)\left(\dfrac{\partial I}{\partial v}\right) & \left(\dfrac{\partial I}{\partial v}\right)^2 \end{pmatrix} \quad (20)$$

where $I(u \quad v)$ is the grey level intensity. If at a certain point the tow eigenvalues of the matrix N are large, then a small motion in any direction will cause an important change of grey level. This indicates that the point is a corner. The corner response function is given by:

$$r = \det(N) - \gamma(\mathrm{trce}(N))^2 \quad (21)$$

where $\gamma = 0.04$ (Harris parameter response). Corners are defined as local maxima of the cornerness function. It is necessary to smooth the images with a Gaussian filter to avoid corners due to image noise.

3.1.2 Interest Points

Interest points are Harris points detected previously matched by the correlation measure ZNCC [12, 1] that is invariant to the local linear changes of luminance. The correlation measure $\mathrm{ZNCC}(m, m')$ between two Harris points m and m' detected in the images 1 and 2, is given by the following formula:

$$\mathrm{ZNCC}(m, m') = \frac{\sum_i a_i b_i}{\sqrt{\sum_i a_i^2 \sum_i b_i^2}} \quad (22)$$

with:

$a_i = I(m+i) - \bar{I}(m)$

$b_i = I'(m'+i) - \bar{I}'(m')$

$\bar{I}(m)$ and $\bar{I}'(m')$ are means of pixel luminance on a 11×11 window centered respectively in m and m'.

3.1.3 Regularization

To eliminate the false matching detected in the previous part, we regularize all interest points by the estimation of a fundamental matrix F between images 1 and 2 while using the RANSAC method [13]. If $\overline{m'}^T F \overline{m} \le s$, with s a threshold fixed to the departure, and $(m \quad m')$ two interest points, then the couple $(m \quad m')$ is admitted, otherwise the couple $(m \quad m')$ is omitted.

3.2 Estimation of Homography of the Plane at Infinity

Several methods have been proposed to improve the evaluation of the homography of the plane at infinity, such as the techniques that are based on a translation camera motion [14] or rotation [15]. Our approach is in the same axis, but with a particular motion "translation and small rotation", this displacement type permits to initialize a non-linear cost function by a solution that guarantees the convergence toward an exact solution.

3.2.1 Formulation of a Cost Function

In practice, there is no direct method to solve equations (19). Therefore we minimize the following non-linear cost function:

$$\min_\theta \sum_{i=1}^{n} \left(\lambda_i^2 + \beta_i^2\right) \quad (23)$$

with:
$$\begin{aligned} \lambda_i &= \left(z_i H_\infty \overline{m}_i + t'\right)_{11} - u'_i \left(z_i H_\infty \overline{m}_i + t'\right)_{31} \\ \beta_i &= \left(z_i H_\infty \overline{m}_i + t'\right)_{21} - v'_i \left(z_i H_\infty \overline{m}_i + t'\right)_{31} \\ \theta &= \left(Z, H_\infty, t'\right) \end{aligned}$$

$Z = (z_1 \dots z_i \dots z_n)^T$ is a n vector of depths of the scene points $M_i = (x_i \quad y_i \quad z_i)^T$. $m_i = (u_i \quad v_i)^T$ and $m'_i = (u'_i \quad v'_i)^T$ are projections of the point M_i in the images 1 and 2.

n (must be superior to twelve) is the number of interest points. We use the Levenberg-Marquardt algorithm (Matlab implementation) to solve this non-linear problem that must be initialized.

3.2.2 Initialization

The motion of camera is "translation and small rotation", therefore initially we suppose that the displacement is mere translation, in this case $R \sim I_3$ therefore $H_\infty \sim I_3$. This constraint permits to write equations (19) as follows:

$$\begin{pmatrix} (u_i - u'_i) & 1 & 0 & -u'_i \\ (v_i - v'_i) & 0 & 1 & -v'_i \end{pmatrix}\begin{pmatrix} z_i \\ t' \end{pmatrix} = 0 \qquad (24)$$

For n interest points, the equation (24) becomes:

$$A\ B\ = 0 \qquad (25)$$

With A as a $2n \times (n+3)$ matrix and B as a $n+3$ vector as:

$$A = (U_1 \quad V_1 \ldots\ldots U_i \quad V_i \ldots\ldots\ U_n \quad V_n \quad)^T$$

$$B = (Z^T \quad t'^T)^T$$

$$U_i = (O_{i-1}^T \quad (u_i - u'_i) \quad O_{n-i}^T \quad 1 \quad 0 \quad -u'_i)^T$$

$$V_i = (O_{i-1}^T \quad (v_i - v'_i) \quad O_{n-i}^T \quad 0 \quad 1 \quad -v'_i)^T$$

$O_i = (0 \ldots\ldots\ldots 0)^T$ is a i zero vector, O_0 is a non definite vector that must be eliminated.

The solution of the equation (25) can be obtained by the singular value decomposition method.

3.3 Intrinsic Parameters

From the relation (9) we find that $R = K^{-1} H_\infty K$ and since $RR^T = I_3$ we deduct that $KK^T = H_\infty KK^T H_\infty^T$ in (13) we have $\Delta = KK^T$ therefore:

$$\Delta = H_\infty \Delta H_\infty^T \qquad (26)$$

The equation (26) can be written as a set of nine linear equations with five unknown (elements of the matrix Δ). Among these nine equations, only five are independent, therefore two views are sufficient to determine the matrix Δ, but for reasons of numeric stability it is preferable to use two displacements (three views). Therefore the linear equation to solve is the following:

$$\begin{cases} \Delta = H_\infty \Delta H_\infty^T \\ \Delta = H_{1\infty} \Delta H_{1\infty}^T \end{cases} \qquad (27)$$

To estimate the homography of the plane at infinity $H_{1\infty}$ between the images 1 and 3, we use the interest points m and m'' detected in these two images and we apply the same procedure as in the previous parts to find H_∞. The intrinsic parameters are calculated by a Cholesky decomposition of the matrix Δ.

4 Experimentations

4.1 Simulations

We simulate a sequence of three 512×512 images of a known grid 3D in the space. The camera moves from one position to the other by motion "translation and small rotation". Points of the grid are projected in images and a calibration method is applied to find the intrinsic parameters that are $(f = 1230,\ a = 0.94,\ u_0 = 264 \text{ et } v_0 = 280)$.

The Gaussian noise with varying deviation σ is added to the image point location. Interest Points are matching by ZNCC and are regularized by the fundamental matrix after the normalization of data [15]. The initialization of the problem is obtained by the hypothesis of a motion is mere translation and the homography of the plane at infinity is estimated by the minimization of the non-linear cost function. The intrinsic parameters are calculated by equations (27). The presented tests concentrate on the comparison of the focal length estimated by our method with the classic calibration technique.

The relative error on the focal length is increased linearly according to noise with a very reasonable precision Fig. 2.

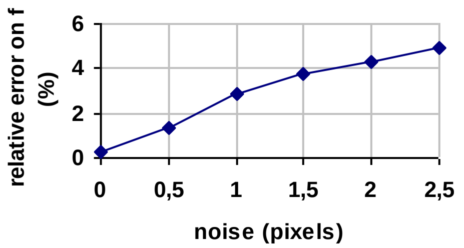

Fig. 2: The relative error on f according to noise.

The relative error on the focal length is decreased when the interest points are added Fig. 3.

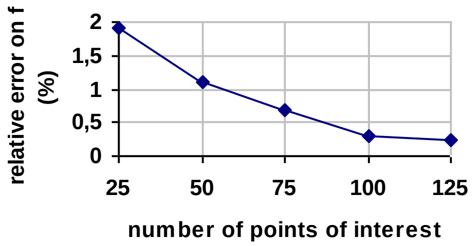

Fig. 3: The relative error on f according to number of interest points.

4.2 Real Data

To show the good quality studied algorithms in this article, three 512×512 images are acquired Fig. 4 by a numeric camera (Canon EOS 400D Digital) of which its intrinsic parameters are kept constant.

Fig.4, The three images used for the camera self-calibration.

The interest points are detected and the homography the plane at infinity is estimated, then the camera parameters are determined table 1.

	f	a	u_0	v_0
Intrinsic parameters	1232	0.87	262	310

Table 1: Camera's intrinsic parameters.

5 Conclusion

We dealt with the problem of the camera self-calibration, from the estimation of the homography of the plane at infinity by a simple technique to manipulate, which consists in displacing the camera so that its motion is "translation and small rotation", with the constraint that at least twelve scene points remain in the camera visual field between two consecutive views. The hypothesis of a motion is mere translation only used to initialize parameters to be estimated, before applying a non-linear resolution method.

References

[1] A. Saaidi, H. Tairi and K. Satori. Fast Stereo Matching Using Rectification and Correlation Techniques. ISCCSP, *Second International Symposium on Communications, Control and Signal Processing*, Marrakech, Morocco, March 2006.

[2] Z. Zhang. A Flexible New Technique for camera Calibration. *Technical Report MSR-TR-98-71, Microsoft Research*, December 1998.

[3] M. Wilczkowiak, E. Boyer, P. Sturm. Camera Calibration and 3D Reconstruction from Single Images Using Parallelepipeds. ICCV, 2001.

[4] Manolis I.A. Lourakis and R. Deriche. Camera self-calibration using the singular value decomposition of the fundamental matrix: From point correspondences to 3D measurements. *Technical Report 3748, INRIA Sophia-Antipolis*, Aug. 1999.

[5] Q.-T. Luong and O. Faugeras. Self-calibration of a moving camera from point correspondences and fundamental matrices. *The International Journal of Computer Vision*, 1997.

[6] C. Zeller and O. Faugeras. Camera Self-Calibration from Video Sequences: the kruppa Equations Revisited. *Research Report 2793, INRIA*, February 1996.

[7] B. Triggs. Autocalibration from planar sequences, *In Proceedings of 5th European Conference on Computer Vision*, Freiburg, Allemagne, Juin 1998.

[8] E. Malis and R. Cipolla, Camera self-calibration from unknown planar structures enforcing the multi-view constraints between collineations, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Septembre 2002.

[9] Q.-T. Luong and T Viéville. Canonical representation for the geometries of multiple projective views. *Computer Vision and Image Understanding*, 1996.

[10] A. Bartoli and P. Sturm, “Nonlinear Estimation of the fundamental Matrix with Minimal Parameters”, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol 26 N° 4 pp 1-7, April 2004.

[11] C. Harris and M. Stephens. A combined corner and edge detector. *In Alvey Vision Conference*, 1988.

[12] M. Lhuillier and L. Quan. Quasi-dense reconstruction from image sequence. *ECCV*'02.

[13] M. A. Fischler et R. C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Graphics and Image Processing*, Juin 1981.

[14] C. Zeller and O. Faugeras. Applications of non-metric vision to some visual guided tasks. *Research Report 2308, INRIA*, July 1994.

[15] R. I. Hartley. In Defense of the Eight-Point Algorithm, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Juin 1997.