An Evolutionary Algorithm for Camera Calibration

Philippe GUERMEUR*, Jean LOUCHET**

* ENSTA - laboratoire LEI - 32 Bd Victor - 75015 Paris - France
** INRIA Rocquencourt, Fractales project - B.P. 105 - 78153 Le Chesnay Cedex - France

Abstract: - Image calibration is the very first step in the low-level vision process, making it possible to reliably exploit geometrical information from images. In this paper, we address the problem of calculating and compensating camera lens distortion using a fast evolutionary algorithm. The advantages and limitations of this method are compared with classical calibration methods.

Key-words: calibration; evolutionary algorithm; lens distortion; collinearity; geometric invariant; optimization.

1 Introduction

Most commercial cameras and lenses deviate from the ideal pinhole model, due in particular to wide angle lens design which generates non-linear image distortion. Calibrating and correcting geometrical distortion is an essential prerequisite to the majority of 3-D reconstruction methods which are based on projective geometry [9]. New sensors with increasing image resolution enable wider viewing angles without loss in scene resolution: this opens the way to new vision applications that require wider angle lenses within the same cost constraints. Low-distortion wide angle lens technology may be extremely costly, so that using cheaper lenses associated with efficient distortion compensation is becoming an increasing economical stake.

In practice, camera and lens manufacturers do not give accurate enough geometrical information. Because of the variability of optical and geometrical characteristics in a production line, these parameters would have to be measured individually on each camera/lens system and lead to an unacceptable cost overhead [14].

In the brief review of camera calibration techniques given in section (3), we note a lack of flexibility of most existing methods, some potential problems due to their initialisation steps and their susceptibility of getting trapped into a local minimum. As an answer to these difficulties, we propose a new approach based on an evolutionary algorithm.

The implementation of our method will be presented in section (4), and calibration results will be given in section (5). These results are compared to those obtained using Tsai's calibration method [19]. Among the various intrinsic parameters we are estimating, we pay particular attention to the location of the principal point, a fundamental data in many reconstruction algorithms.

2 The calibration parameters

The parameters of an ideal pinhole camera are classically divided into extrinsic and intrinsic parameters. The extrinsic parameters represent the location and rotation of the camera in space, whereas the intrinsic parameters refer to the camera internal model. The first internal parameters are the focal length, the principal point coordinates and the horizontal/vertical scale factor.

As discussed above, this projective model is usually not accurate enough in many applications, and image distortion has to be considered. Image distortion consists of non-linear image deformations originating from optical lens imperfections, that may come from the inevitable flaws in lens design but also to optical and mechanical imperfections (glass quality, curvature, misalignment of optical elements, etc.). Typically, the image distortion model can be decomposed into two types :

- the radial distortion, whose components are expressed as:

$$d_{x_r} = kx(a_1 r^2 + a_2 r^4 + a_3 r^6)$$

$$d_{y_r} = y(a_1 r^2 + a_2 r^4 + a_3 r^6)$$

where a_1, a_2 and a_3 are the radial coefficients,

(x, y) are the centred image points coordinates, and r the radial distance from the image centre (x_0, y_0):

$$r^2 = k^2 x^2 + y^2$$

$$x = x_d - x_0 \qquad y = y_d - y_0$$

k is the horizontal scale factor.

- the tangential distortion resulting from misalignment of the lenses:

$$d_{x_t} = [p_1[r^2 + 2x^2] + 2p_2 xy](1 + p_3 r^2)$$

$$d_{y_t} = [p_2[r^2 + 2y^2] + 2p_1 xy](1 + p_3 r^2)$$

where p_1, p_2 and p_3 are the tangential coefficients. The global distortion is the sum of radial and tangential distortion:

$$d_x = d_{x_r} + d_{x_t} \qquad\qquad (1)$$

$$d_y = d_{y_r} + d_{y_t} \qquad\qquad (2)$$

3 Related works

Most calibration methods handle both intrinsic and extrinsic parameters. A difficulty with these methods derives from the non-independence of internal and external parameters, which can lead to high estimation errors on internal parameters [5,20]. Trouble also comes when there is no suitable metrology equipment available, given that these methods require accurate knowledge of 3-D coordinates of scene points. In addition, many applications do not aim at performing a complete 3D reconstruction of the scene, and in these cases there is no need to recover the extrinsic parameters. An example is the obstacle detection systems we are developing in our labs, where the goal is to obtain fast mobile robot reactions in emergency situations rather than an accurate scene reconstruction.

As a consequence, in this paper we will be only interested in methods to estimate distortion parameters in conjunction with internal parameters. These methods are non-linear. Of course, it is still possible to use the results of intrinsic parameter estimation as a first step to further extrinsic parameters determination techniques. Unlike full calibration techniques, we do not need to measure the 3D location of reference points or camera location, and thus we are not sensitive to errors on these measurements. We only need to

recover projective geometry invariance properties, which have been lost due to distortion. The traditional approach is then to define a cost function which will represent the deviation to invariance properties, then to minimise this cost unction in order to obtain the distortion parameters. Various calibration methods have been designed, each according to one particular property. As an example, some of them use the orthogonality of 3 vanishing point vectors corresponding to 3 orthogonal lines in the scene [7]. Other methods are based on the preservation of the cross-ratio of 4 points along a line [15], others use the existence of a planar mapping between two views of a planar object [15], etc.

It is probable that the most widely used techniques are those relying on the basic invariance property which states that if there is no distortion, every straight line in the scene corresponds to a straight line in the image. An algorithm based on this property was first introduced by Duane C. Brown in 1971, as "the plumb line method". This method aims at solving an equation system whose unknowns are the q parameters of the distortion function. Every new line in the scene introduces two additional parameters (ρ, θ) which can be combined to the distortion parameters in a new expression:

$$(x + d_x) \sin\theta + (y + d_y) \cos\theta = \rho$$

where x and y are the coordinates in the ideal image. Then the two terms d_x and d_y are substituted with their expressions (1) and (2), to get a new expression which can be linearised using Taylor's expansion from an initial approximation. A set of lines enables to construct a resolvable system of q equations with q unknowns. Many recent techniques [5,16,21] can be considered as derivatives of the plumb line method. Most of them use a cost function, which gives a null value when all points are perfectly aligned, and a value increasing with line distortion. The aim is then to minimise this cost function using a non-linear least square minimisation method [1], as Lagrange's [12] or the Levenberg-Marquardt method [5, 18]. This last step can be enriched with outliers' elimination to increase robustness and accuracy of the estimation [18]. One of the drawback of the plumb line method and its relatives is that they do not yield any estimation of the focal length, and the estimation of the principal point coordinates are not very accurate. Furthermore, it is remarkable that in the

absence of information on camera parameters, results turn out to be instable [20]. One can overcome this difficulty and obtain stable and reasonably accurate results if camera motion is a pure rotation [14,15]. However, this may not always be materially possible.

These optimization methods are based on local linearisation of the non-linear problem, then reducing the calibration problem to a sequence of (possibly constrained) linear problems. Such procedures require well-chosen initial estimates in order to converge to a global minimum. Without a convenient choice, the solution may diverge or get trapped into a local minimum.

4 Calibration using an evolutionary algorithm

This section presents our alternative approach to the conventional calibration methods by using an evolutionary algorithm (EA). EAs are recognized for their capability of avoiding the risk of getting trapped into a local minimum [8]. In addition, they do not require an initial estimate to start the optimization process.

In fact, our main motivations for elaborating a calibration method based on artificial evolution are given by the following general advantages of EAs among other optimization methods:

- the calibration method is expected to be autonomous. EAs generally do not require to choose a convenient initial estimate. In fact, EA initialization consists in generating a random population of vectors.

- the method should be robust with low risk of local minimum trapping. EAs are extremely robust to complex cost landscapes with multiple local minima thanks to parallel domain exploration using a large population of vectors.

- it has now become simple, fast and easy to implement EA methods in real-world application, thanks to several public domain genetic libraries (EO, Galib, DREAM...) and, perhaps even more importantly, to recent user-friendly development tools such as EASEA [3, 4] which is able to exploit any of these libraries and produce efficient and clean source code, leaving the user with the task of writing his own application-specific cost (fitness) function

and some parameter adjustment.

- Writing this cost/fitness function turns out to be the main difficulty when designing an EA in the EASEA framework. As a consequence our calibration method is expected to be extensible. This is really an interesting property in a computer vision laboratory where everyone may have to modify the optimization problem with the application needs or when replacing the type of lens (high distortion lens, panoramic camera...).

4.1 Background

EAs have been introduced in the 60s, nearly simultaneously in the United States (John H. Holland) and in Germany (I. Rechenberg). They are efficient stochastic optimization tools based on the process of natural selection inspired by Darwin's theory on natural evolution. Basically, they consist in initializing a random population of potential solutions of the problem to be solved. The aim is to maximize (or minimize) a fitness (or cost) function by evaluating all individuals in the population and applying stochastic genetic operators, mainly crossover (recombination of two individuals) and mutation (noise on a single individual). A selection operator tends to eliminate the less performing individuals so that the population gradually concentrates into the best solutions of the problem.

In practice, the combination of mutation and crossover operators enables to preserve genetic diversity and enables extensive exploration of the search space. The fitness function concentrates the user's knowledge of the problem. It is a real value function, whose variables are the code of an individual.

Usually, the various parameters of the EA are set in an empirical way (initialisation, crossover rate, mutation rate, selection method...). Some recent statistical studies investigate methods to derive these parameters from extrapolation of parameters which have already been defined for similar problem types [6].

4.2 Implementation

Like many calibration methods (section 3) we choose to consider the collinearity property, which states that every straight line in the scene has to correspond to a straight line in the image. This section presents the various elements of a calibration method designed

from this property and based on an EA.

Input data

From a practical standpoint, we use a calibration pattern (Fig. 1), which enables us to extract a square grid of 5x5 regularly spaced image points, using sub-pixel detection. To this end we chose the Harris characteristic point detector for its good precision, as mentioned in comparative studies relative to interest points detection [12]. This enables us to compose sets of 5 points which should be straight if there was no distortion. These sets of points are the data which will be processed by our algorithm.

Coding

The population of our EA is composed of a population of vectors, whose components v_i are the nine distortion parameters:

$$v_i = (a_1, a_2, a_3, p_1, p_2, p_3, x_0, y_0, k)$$

These parameters are coded as real values (conforming to the "evolutionary strategy" technique [11]. The bounds on the search space are chosen large enough to cover the whole set of admissible values.

Fitness

In order to evaluate an individual, we use v_i, its vector of distortion parameters to correct the input data. To obtain a measure of the quality of the correction, we consider every set of image points which should be on a straight line. For each set of points, a straight line is approximated using a least square method and then we add the square of the distance of every point to the estimated straight line. Next, the results we obtain for each set of points are added together to obtain our cost function (distortion measure). The fitness function (to be maximised) is derived from the reverse of this cost function.

Crossover

We use the classical barycentric crossover method, and in particular we create two new vectors v_{n1} and v_{n2} from a linear combination of two parents, v_1 and v_2:

$$v_{n1} = Kv_1 + (1 - K)v_2$$
$$v_{n2} = (1 - K)v_1 + Kv_2$$

where K is an uniform random number chosen in the interval $[0, 1]$.

Mutation

To generate a new individual we add a white gaussian noise to a selected vector. The standard deviation of the gaussian distribution has a fixed, arbitrary constant value.

Overall implementation

The algorithm has been implemented using the EASEA evolutionary specification language [3], in conjunction with GAlib (Genetic Algorithm library), a public domain C++ library of optimization tools.

The various algorithm parameters (mutation and crossover probabilities, population size, selection method...) have been set in an empirical way.

To ensure a good balance of robustness and precision, we hybridise the evolutionary strategy with a steepest gradient method.

5 Experimental results

In order to study how the results depend on the generation number, many independent experiments have been done, using an arbitrary but fixed population size and an increasing number of generations. Fig.2 to Fig.5 plot every results value according to the number of generations. Our camera was equipped with a 6.5 mm focal length wide-angle lens, and a sensor size equal to 2/3 inches.

Fig.2 illustrates how the cost function decreases with the number of generations. One interesting property of this approach is that a coarse result is obtained very quickly, and gets refined with time. The results for the different distortion coefficients are presented on Fig.3 to Fig.5, Their stability seems to be very close to their effective contribution in the value of the overall distortion.

The results concerning the principal points coordinates are presented on figure 3. The plots illustrate that the results are not very stable. However, the instability range can be considerably reduced if we consider only the minima of the cost function (20% best values) which is plotted on Fig. 2. It seems that this instability phenomenon is a distinctive feature of the plumb line method, where the principal points coordinates have been proved to be closely linked to the decentring distortion coefficients [2].

As a comparison, we have measured optically the coordinates of the principal point using a laser beam autocollimation method [2]. We obtained a location $(x_0, y_0) = (370, 289)$ close enough to our results. The autocollimation method seems to give more stable results (less than one pixel), but this ought to be

confirmed experimentally with other lens types.

If we consider the quality of the distortion correction, our results are quite satisfactory: after image correction using our best parameters set, the cost function was about 1,16 and the maximal distance of every calibration point to the estimated straight line was measured as 0.64 pixel. As a comparison the results we obtained using Tsai's algorithm [19], one of the most widely used algorithm for camera calibration, correspond to a cost function equal to 6.15 (without any distortion correction the cost function was equal to 378.3).

Other experiments have been conducted using a calibration pattern enabling the extraction of a grid composed of 10x10 characteristic points. This increasing number of input data did not result in any improvement on the results, whereas the computation time increased drastically (from a few seconds to many minutes).
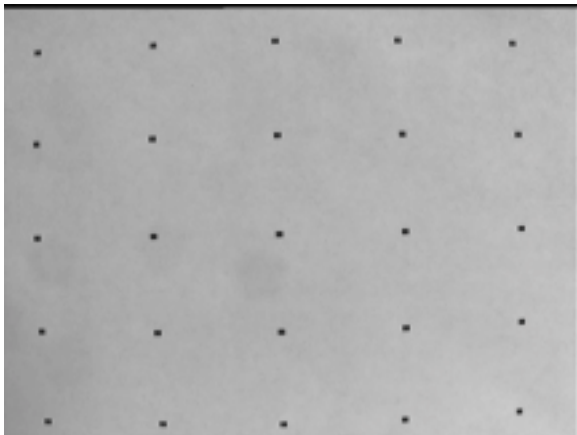

Figure 1 - *The calibration pattern is a 5x5 grid of squares whose centres are the characteristic points.*

6 Conclusion

We have presented a calibration procedure based on an evolutionary algorithm (evolutionary strategy) to perform camera calibration from a single image. The theoretical advantages of the proposed method: autonomy, robustness and simplicity have been experimentally confirmed, and the quality of the distortion correction turns out to be quite satisfactory. Further work will consist of experimenting the extensibility of the algorithm using other lens types.
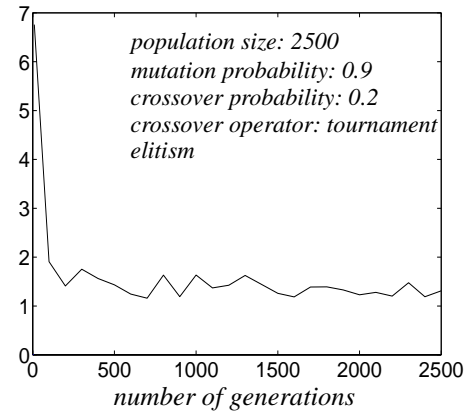
cost function

number of generations

Figure 2 -   *Evolution of the cost function in different independent experiments. The x-axis represents the number of generations.*
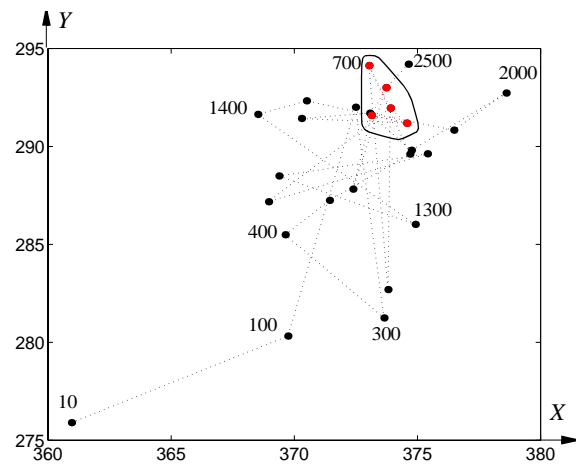

Figure 3 -   *Measured location for the principal point. The surrounded values refer to the minima of the cost function plotted on Fig.2.*

References:

[1] S. S. Beauchemin, R. Bajcsy and G. Givaty, "A Unified Procedure for Calibrating Intrinsic Parameters of Sperical Lenses," Vision Interface 99, pp. 272-279, Trois-Rivières, Canada, 1999.

[2] T. A. Clarke, J. G. Fryer, X. Wang, "The Principal Point and CCD Cameras," *Photogrammetric Record*, vol. 16, pp. 293-312, 1998.

[3] P. Collet, E. Lutton, M. Schoenauer, Jean Louchet, "Take it EASEA," Parallel Problem Solving from Nature VI, vol 1917, Springer pp. 891-901, 2000.

[4] P. Collet, "Easea Specification for Evolutionary Algorithms," *User Manual, and Reference Manual*, INRIA-Ecole Polytechnique-ENSTA, 2001.

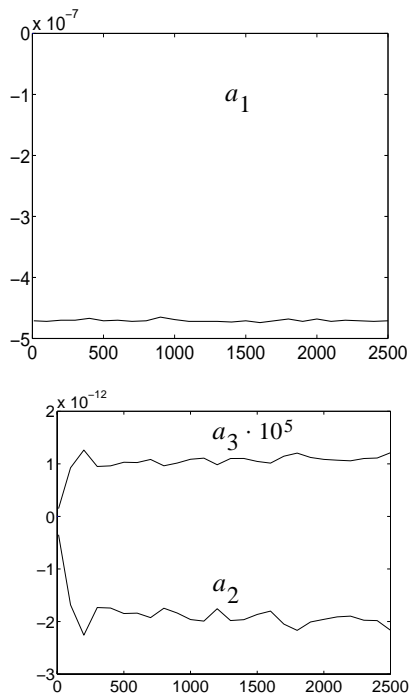[5] F. Devernay, Olivier D. Faugeras, "Straight lines have to be straight," *Machine Vision and*
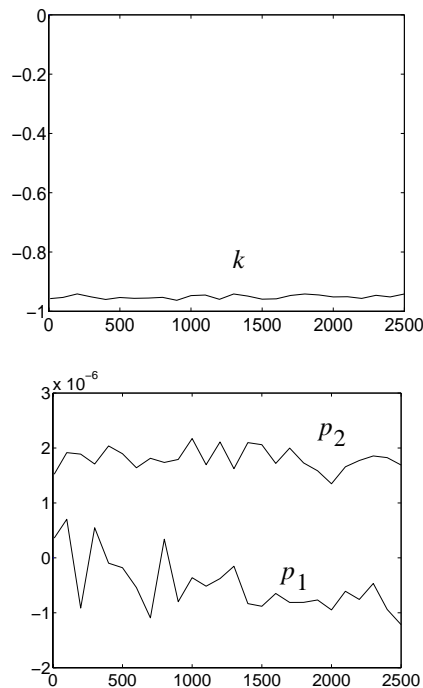
Figure 4 - Radial distortion coefficients.

Figure 5 - Scale factor k and tangential coefficient distortion.

Applications, volume 13, n° 1, pp. 14-24, 2001.

[6] O. François, C. Lavergne, "Design of Evolutionary Algorithms, a Statistical Perspective," *IEEE Transactions on Evolutionary Computation*, vol. 5, n°2, avril 2001.

[7] Mengxiang Li, "Camera Calibration of the KTH Head-Eye System," *Report from Computational Vision and Active Perception Laboratory (CVAP)*, pp. 1-26, 1996.

[8] E. Lutton, P. Collet, J. Louchet, M. Sebag, C. Fonlupt, "Evolution Artificielle," *ENSTA lectures*, 2000.

[9] J. Mundy, A. Zisserman, "Geometric Invariance in Computer Vision," The MIT Press, 1992.

[10] T. Pajdla, T. Werner, V. Hlavac, "Correcting Radial Lens Distorsion without Knowledge of 3-D Structure," *Research Report, Czech Technical University*, n° K335-CMP-1997-138, 1997.

[11] I. Rechenberg, Evolutions stategie: Optimierung Technicher System nach Prinzipien der Biologischen Evolution. *Fromman Holzboog*, Stuttgart, 1973.

[12] C. Schmid, "Appariement d'images par invariants locaux de niveaux de gris," Ph.D. Thesis, Institut National Polytechnique de Grenoble, 1996.

[13] S. Shah, J. K. Aggarwal, "Intrinsic Parameter Calibration Procedure for a (High-Distortion) Fish-Eye Lens Camera with Distortion Model and Accuracy Estimation," *Pattern Recognition*, vol. 29, pp. 1775-1788, 1996

[14] G .P. Stein, "Internal Camera Calibration using Rotation and Geometric Shapes," *Master of Science Thesis,* MIT, pp. 1-47, 1993.

[15] G. P. Stein, "Accurate Internal Camera Calibration using Rotation, with Analysis of Sources of Error," *International Conference on Computer Vision*, 1995.

[16] D. Stevenson, M. Fleck, "Nonparametric Correction of distortion," *Proc. of IEEE Workshop on Applications of Computer Vision*, pp. 214-219, 1996.

[17] Rahul Swaminathan, Shree K. Nayar, "Non-Metric Calibration of Wide Angle Lenses," *DARPA Image Understanding Workshop*, pp. 1079-1084, 1998.

[18] T. Thormählen, H. Broszio, I. Wassermann, "Robust Line-Based Calibration of Lens Distortion from a Single View," *Proceedings of Mirage 2003*, INRIA Rocquencourt, France, pp. 105-112, 2003.

[19] 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. 3, pp. 323-344, 1987.

[20] J. Weng, P. Cohen et M. Herniou, "Camera Calibration with High distortion Models and Accuracy Evaluation," *IEEE PAMI*, vol. 14, n°10, pp. 965-981, october 1992.

[21] R. J. Valkenburg, "Classification of camera calibration techniques," *SPIE Proceeding*, San Jose, USA , vol. 3641, pp. 152-163, 1999.