ISSN 1392 - 1215 ELEKTRONIKA IR ELEKTROTECHNIKA. 2005. Nr. 3(59)

T 125 AUTOMATIZAVIMAS, ROBOTOTECHNIKA

Investigation of Technical Equipment in Computer Stereo Vision: Camera Calibration Techniques

P. Serafinavičius
Department of Electronics, Šiauliai University,
141 Vilniaus str., Šiauliai, Lithuania, tel. +370 41 595835, e-mail: paulius_s@splius.lt

Introduction

The sense of vision plays an important role in the life of human: it allows them to infer spatial properties of the environment. Vision is necessary to explore unfamiliar surroundings, detect and recognize distances in physical space. A visual system is a collection of devices that transform measurements of light into information about spatial and material properties of a scene. There is necessary to have photosensitive sensors (a camera), also computational mechanisms (a computer) that allows to extract information from the sensors. Certainly the pixel values recorded with a camera depend on the shape of objects in the scene: if the shape has been changed, there is noticeable a change in the image. Therefore, images depend on the geometry of the scene. However, they also depend on its photometry, the illumination and the material properties of objects. Visual measurements depend on the dynamics of the environment. In general, we do not know the shape of the scene, its material properties and motion. The goal is to infer a representation of the world from collection of images. This goal can be achieved by various tasks and algorithms, those include [1]:

- Given two images, compute matches between the images, and the 3D position of the points that generate these matches and the cameras that generate the images.
- Given three images, compute the matches between images of points and lines, and the position in 3D of these points and lines and the cameras.
- Compute the epipolar geometry of a stereo rig, and trifocal geometry of a trinocular rig, without requiring a calibration object.
- Compute the internal calibration of a camera from a sequence of images of natural scenes.

This paper is concentrated on the analysis of camera calibration techniques in computer stereo vision when given two or more images and no other information and the calibration is done with a calibration object or without it.

Classification of camera calibration techniques

Calibration of a camera is one of the most important tasks when registering a stereoscopic view. This is a necessary step in computer stereo vision in order to extract metric information from two-dimensional views. In other words, camera calibration is the process of relating the ideal model of the camera to the actual physical device and determining the position and orientation of the camera with respect to the world coordinate system. Calibration is important for accuracy in 3-D reconstruction.

Depending on the model used, there are different parameters to be determined. The pinhole camera model is broadly used. The parameters to be calibrated were classified in two groups:

- Internal (or intrinsic) parameters. Internal geometric and optical characteristics of the lenses and the imaging device.
- External (or extrinsic) parameters. Position and orientation of the camera in the world coordinate system.

Calibrating stereo cameras is usually done by calibrating each camera independently and then applying geometric transformation of the external parameters to find out the geometry of the stereo setup.

Existing camera calibration techniques can be classified by the appliance of mathematical apparatus:

- closed form solution;
- full nonlinear optimization;
- two steps methods.

The other way to classify calibration techniques is to group them into two categories depending on that a calibration object is necessary or not:

- photogrammetric calibration;
- auto-calibration (self-calibration).

Photogrammetric calibration is performed by observing a high precision 3D calibration object. The calibration object consists of two or three orthogonal to each other planes (Fig. 1) [2]. The two planes are carefully assembled so their angle is precisely known and the positions of the corners of the squares are also very precisely known. Those corners are used as reference 3-D points. The procedure then goes in the following order:

- The cameras to be calibrated observe a picture of the calibration pattern (Fig. 2).
- Polygonal approximation of the contours is computed, the corners are automatically extracted from this approximation and their pixel coordinates are computed.

- The intrinsic and extrinsic parameters can be computed.

The quality of the calibration can be qualitatively tested by verifying that the epipolar geometry is correct. These approaches require an expensive calibration apparatus and elaborate setup in most of the cases.

Auto-calibration techniques do not use any calibration object. By moving a camera in a static scene, the rigidity of the scene provides in general two constraints on the cameras' internal parameters from one camera displacement by using image information alone. Therefore, if images are taken by the same camera with fixed internal parameters, correspondences between three images are sufficient to recover both the internal and external parameters which allow as reconstructing 3-D structure up to a similarity [3]. This approach is very flexible, but it is not yet mature. Majority of the parameters couldn't be always obtained reliably.

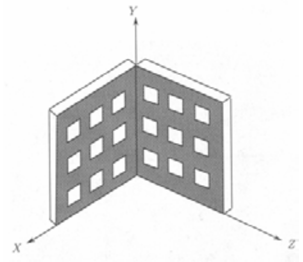

Fig. 1. A calibration pattern: calibration points are at the intersection of horizontal and vertical lines, which can be easily detected using simple algorithms

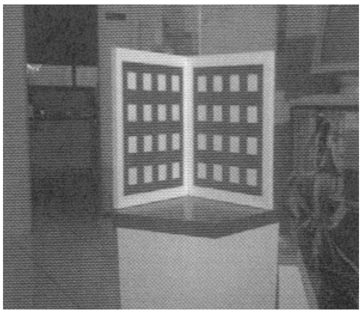

Fig. 2. An image of the calibration pattern seen from one of the cameras

Photogrammetric calibration techniques

A well known method for calibrating a camera has been proposed by Tsai [4]. The method is based on the knowledge of the position of some points in the world and the correspondent projections on the image. It required the camera to be pointed to a nonplanar calibration grid (that must be accurately prepared). The first step is to solve a linear equation to find the external parameters rotation angles for the transformation between the world and camera coordinates R, translation components for the transformation between the world and camera coordinates T and the scale factor a_x, and then a nonlinear optimization is performed over T_z, focal length of camera f, radial lens distortion coefficient k. Finally, coordinates of centre of radial lens distortion O_x, O_y are determined with another nonlinear optimization step. The procedure can be iterated in order to improve accuracy.

Every calibration technique for radial distortion, which is based on the knowledge of the position of points in the world, needs to search for the overall set of parameters at the same time, even though it is possible to separate the parameters in two groups, so that nonlinear optimization is performed only on a subset of the parameters [4].

The calibration procedure requires the user to

- use a calibration grid;
- detect the projections of calibration points in the image.

The other famous calibration technique developed by Zhang [5] is easy to use and flexible. It advances computer stereo vision one step from laboratory environments to real world use. The calibration procedure is as follows:

1) Print a pattern and attach it to a planar surface.
2) Take a few images of the model plane under different orientations by moving either the plane or the camera.
3) Detect the feature points in the images.
4) Estimate the five intrinsic parameters and all the extrinsic parameters using the closed-form solution.
5) Estimate the coefficients of the radial distortion by solving the linear least-squares.
6) Refine all parameters by minimizing.

The proposed technique requires the camera to observe a planar pattern shown at least two different orientations. The pattern can be printed on a laser printer and attached to planar surface. Either the camera or the planar pattern can be moved by hand. The motion need not be known. A camera is modeled by the usual pinhole: the relationship between a 3-D point *M* and its image projection *m* is given by

$$s\tilde{m} = A[R \quad t]\tilde{M}\,; \tag{1}$$

where *s* is an arbitrary scale factor, the augmented vector is denoted by adding 1 as the last element ($\tilde{m} = [u, v, 1]^T$ and $\tilde{M} = [X, Y, Z, 1]^T$), (R; t), called the extrinsic parameters, is the rotation and translation which relates the world coordinate system to the camera coordinate system, and A, called the camera intrinsic matrix, is given by

$$A = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}; \tag{2}$$

with (u_0; v_0) the coordinates of the principal point, α and β the scale factors in image u and v axes, and γ the parameter describing the skewness between two image axes.

Camera calibration process starts with closed-form solution. A model point *M* and its image point *m* is related by a homography *H*. There are two basic constraints (3, 4) on the intrinsic parameters, given one homography.

$$h_1^T A^{-T} A^{-1} h_2 = 0\,, \tag{3}$$

$$h_1^T A^{-T} A^{-1} h_1 = h_2^T A^{-T} A^{-1} h_2\,. \tag{4}$$

Because a homography has 8 degrees of freedom and there are 6 extrinsic parameters (3 for rotation and 3 for translation) it is possible to obtain only 2 constraints on the intrinsic parameters. A matrix B (5) describes the image of the absolute conic [2]

$$B = A^{-T} A^{-1} . \tag{5}$$

B is symmetric, defined by a 6 dimmensions vector

$$b = [B_{11}, B_{12}, B_{22}, B_{13}, B_{23}, B_{33}]^T . \tag{6}$$

The i^{th} column vector of H is $h_i = [h_{i1}, h_{i2}, h_{i3}]^T$. Then we have

$$h_i^T B h_j = v_{ij}^T b . \tag{7}$$

If n images of the model plane are observed we have

$$Vb = 0 ; \tag{8}$$

where V is a $2n \times 6$ matrix. When b is estimated, all camera intrinsic parameters (matrix A) can be computed. When A is known, the extrinsic parameters for each images can be computed.

The next step is to estimate the radial distortion coefficients k_1 and k_2 by solving (9)

$$k = (D^T D)^{-1} D^T d ; \tag{9}$$

where $Dk = d$, $k = [k_1, k_2]^T$, when given m points in n images. The equation for each point in each image

$$\breve{u} = u + (u - u_0)[k_1(x^2 + y^2) + k_2(x^2 + y^2)^2] , \tag{10}$$

$$\breve{v} = v + (v - v_0)[k_1(x^2 + y^2) + k_2(x^2 + y^2)^2] ; \tag{11}$$

where (x, y) and $(\breve{x}, \breve{y})$ are the ideal and real (distorted) image coordinates.

The next step is to refine all parameters using Maximum Likelihood Estimation. It is done by minimizing the following functional:

$$\sum_{i=1}^{n} \sum_{j=1}^{m} \| m_{ij} - \breve{m}(A, k_1, k_2, R_i, t_i, M_j) \|^2 , \tag{12}$$

where $\breve{m}(A, k_1, k_2, R_i, t_i, M_j)$ is the projection of point M_j in image i. This is nonlinear minimization task, which is solved using the Levenberg-Marquardt algorithm.

Auto-calibration techniques

Auto-calibration is the process of determining internal camera parameters directly from multiple uncalibrated images. Auto-calibration avoids the onerous task of calibrating cameras using special calibration objects. This gives great flexibility since camera can be calibrated directly from an image sequence despite unknown motion and changes in some of the internal parameters.

For example, we have a set of images acquired by a camera with fixed internal parameters, and a projective reconstruction is computed from point correspondences across the image set [1]. The reconstruction computes a projective camera matrix P_i for each view. Our constraint is that for the actual cameras the internal parameter matrix K is the same for each view but unknown. In general the calibration matrix will differ for each view. It's possible to vary the projective reconstruction by transforming the camera matrices by a homography. The reconstruction transformed by homography is within a similarity transformation of the actual cameras and scene, so a metric reconstruction can be achieved.

There are variuos kinds of auto-calibration. They differ in the constraints used and the methods whereby the homography is determined.

Liebowitch and Zisserman [6] described a technique of metric rectification for perspective images of planes using metric information such as a known angle, two equal though unknown angles, and a known length ratio. They mentioned that calibration of the internal camera parameters is possible by providing at least three such rectified planes.

There was developed by Bill Triggs [7] a technique for camera auto-calibration and scaled Euclidean structure from three or more views taken by a moving camera with fixed but unknown intrinsic parameters. The motion constancy of these is used to rectify an initial projective reconstruction. Euclidean scene structure is formulated in terms of the absolute quadric – the singular dual 3D quadric (4×4 rank 3 matrix) giving the Euclidean dot-product between plane normals. It encodes both affine and Euclidean structure, and projects very simply to the dual absolute image conic which encodes camera calibration. Requiring the projection to be constant gives a bilinear constraint between the absolute quadric and image conic, from which both can be recovered nonlinearly from $m \geq 3$ images, or quasi-linearly from $m \geq 4$. Calibration and Euclidean structure follow easily. The nonlinear method is more stable, faster, more accurate and more general than the quasi-linear one. It is based on a general constrained optimization method – sequential quadratic programming.

The full algorithm for auto-calibration and scaled Euclidean reconstruction is as follows [7]:

1) Standardize all image coordinates.
2) Find the projections P_i by projective reconstruction.
3) Find the absolute quadric Ω and image conic ω by solving 15m bilinear quadric projection constrains $\omega \wedge (P_i \Omega P_i^T) = 0$ (nonlinear and quasi-linear methods).
4) Recover the camera calibration K by Choleski decomposition of $\omega = KK^T$.
5) Find a 4×4 Euclidenizing homography T by eigendecomposition of Ω.
6) Perturb $K^{-1} P_i T^{-1} \sim R_i (I - t_i)$ to be exactly Euclidean.
7) Recover Euclidean structure by $x \to Tx$ or backprojecting with the corrected projections.

Standardization rescales image pixel coordinates. It is absolutely necessary. Their numerical conditioning is terrible and severe floating point truncation error leads to further loss of precision. This is the major reason for the observed instability of some previous auto-calibration approaches. Standardization (preconditioning) is essential whenever there is an implicit least squares trade-off, particularly with equations of high degree.

Discusion and Conclusion

There are many camera calibration techniques developed by variuos researchers. All of them have advantages and disadvantages. The main disadvantage is that not all of them takes into account camera‘s radial lens distortion.

Techniques with the closed form solution cannot consider lens distortion. Full nonlinear optimization can be very difficult to use because the total number of parameters to be calibrated is at least 11. Two steps methods are more efficient and accurate.

The calibration technique proposed by Zhang only requires the camera to observe a planar pattern from a few (at least two) different orientations. The camera or the planar pattern can be moved. The motion does not need to be known. There were affirmed by the author that both computer simulation and real data have been used to test technique, and very good results have been obtained. Compared with classical techniques, the technique gains more flexibility.

The calibration technique proposed by Triggs is suitable for autocalibrating a moving camera with fixed but unknown intrinsic parameters, moving arbitrarily in an unknown scene. An initial projective reconstruction is rectified to give calibration and scaled Euclidean structure and motion. There were affirmed by the author that the results were stable and accurate for generic camera motions, and the formalism clarifies the reasons for autocalibration of intrinsic parameters. A major practical advantage of the technique is the ease with which it incorporates any further constraints that may be available, significantly reducing the problems of calibrating intrinsic parameters.

References

1. **Hartley R., Zisserman A.** Multiple View Geometry in Computer Vision. Second Edition ISBN 0521 54051 8 Cambridge University Press, 2003. – P. 458-463
2. **Faugeras O.** Three-Dimensional Computer Vision. A Geometric Viewpoint. ISBN 0-262-06158-9 The MIT Press, 1993. – P. 18-20, 230-232
3. **Hartley R.** An Algorithm for Self Calibration from Several Views // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1994. – P. 908-912
4. **Tsai R**. An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision. // Proceedings of CVPR'86, 1986. – P. 364-374
5. **Zhang Z.** A flexible new technique for camera calibration // IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000. – 22(11) – P. 1330-1334
6. **Liebowitz D., Zisserman A.** Metric Rectification for Perspective Images of Planes // Proceedings of the IEEE Conference on Computer Vision and Patern Recognition, 1998. – P. 482-488
7. **Triggs B.** Autocalibration and the Absolute Quadric // Proceedings of Conference on Computer Vision and Pattern Recognition, 1997. – P. 609-614

Pateikta spaudai 2005 03 14

P. Serafinavičius. Erdvinio vaizdo registravimo techninės įrangos analizė // Elektronika ir elektrotechnika. – Kaunas: Technologija, 2005. – Nr. 3(59). – P. 24–27.

Šiame straipsnyje aprašomas vienas svarbiausių žingsnių kompiuteriniame erdviniame regėjime – vaizdo kameros kalibravimas. Ištirti kalibravimo metodai sukurti įvairių pasaulio mokslininkų. Čia pateikiama vidinių ir išorinių kalibravimo parametrų analizė, kalibravimo metodų klasifikavimas pagal naudojamą matematinį aparatą (uždaros formos sprendimas, netiesinis optimizavimas, dviejų žingsnių metodai), kalibravimo procesai (fotogrametrinis, auto-kalibravimas), jų privalumai ir trūkumai. Il. 2, bibl. 7 (anglų kalba; santraukos lietuvių, anglų ir rusų k.).

P. Serafinavičius. Investigation of Technical Equipment in Computer Stereo Vision: Camera Calibration Techniques // Electronics and Electrical Engineering. – Kaunas: Technologija, 2005. – No. 3(59). – P. 24–27.

This paper describes one of the most important steps in computer stereo vision – camera calibration. Techniques developed by various researchers have been investigated. Analysis of internal and external calibration parameters, classification of the calibration techniques by the mathematical apparatus (closed-form solution, nonlinear optimization, two steps methods), calibration processes (photogrammetric, auto-calibration), advantages and disadvantages of them are provided here. Ill. 2, bibl. 7 (in English; summaries in Lithuanian, English and Russian).

П. Серафинавичюс. Исследование технических устройсв системы регистрирования стерео изображения // Электроника и электротехника. – Каунас: Технология, 2005. - № 3(59). – С. 24–27.

В этой статъе описан один из важнейших шагов в компютерном стерео зрении – калибрация видео камеры. Разные методы калибрации были исследованы. Здесь представлен анализ внутренних и внешних параметров калибрации, класификация методов калибрации, учитывая используемый математический апарат (решение закрытой формы, нелинейная оптимизация, двух-шаговые методы), процесы калибрации (фотограмметрический, авто-калибрация), их достоинства и недостатки. Илл. 2, библ. 7 (на английском языке; рефераты на литовском, английском и русском яз.).