

Calibration of a wide angle stereoscopic system

Carlos Ricolfe-Viala,* Antonio-Jose Sanchez-Salmeron, and Enrique Martinez-Berti

Instituto de Automática e Informática Industrial, Universitat Politècnica de Valencia, Camino de Vera, Valencia, Spain 46022

*Corresponding author: cricolfe@isa.upv.es

Received May 27, 2011; revised July 5, 2011; accepted July 5, 2011;
posted July 11, 2011 (Doc. ID 148339); published August 8, 2011

Inaccuracies in the calibration of a stereoscopic system appear with errors in point correspondences between both images and inexact points localization in each image. Errors increase if the stereoscopic system is composed of wide angle lens cameras. We propose a technique where detected points in both images are corrected before estimating the fundamental matrix and the lens distortion models. Since points are corrected first, errors in point correspondences and point localization are avoided. To correct point location in both images, geometrical and epipolar constraints are imposed in a nonlinear minimization problem. Geometrical constraints define the point localization in relation to its neighbors in the same image, and epipolar constraints represent the location of one point referred to its corresponding point in the other image. © 2011 Optical Society of America

OCIS codes: 150.1488, 150.1135.

Calibration of a stereoscopic system consists of computing the geometrical relationship between images from both cameras, which is represented with the fundamental matrix. With the fundamental matrix, the projection of a point in one image defines the line where this point is projected in the other image. This fact reduces notably the searching areas of corresponding points in applications such as image matching in stereoscopic systems or scene reconstruction. The fundamental matrix represents the epipolar geometry between both images and a precise estimation is crucial to obtain an effective solution in cited applications. The fundamental matrix is computed from correspondences between image points using linear or nonlinear optimization or with robust methods.

Linear methods with data normalization are considered as a first step for other algorithms or to obtain a quick solution. Linear methods give adequate results if points are well located and the corresponding problem is solved previously. If more accuracy is desired, iterative methods are efficient with noisy data but cannot resolve the points correspondences problem. Robust methods can manage noisy data and errors in point correspondences. Mainly, errors in fundamental matrix estimation come from errors in point correspondences and point localization [1].

If wide angle lens cameras are used, errors in points correspondences and points localization increase since high distortion is present. In these cases, lens distortion is corrected first and the fundamental matrix is estimated in a second step. Images are corrected but accurate point localization is not possible since distortion models do not correct the image entirely. Also errors in points correspondence remain. Otherwise, the fundamental matrix and the lens distortion model can be computed together in one step. In these cases, coupling between both models can result in an absurd solution. To improve the fundamental matrix estimation, especially with high distorted images, a method is proposed to avoid errors in point correspondences and resolve the image distortion efficiently. First, points are corrected in both images in a nonlinear minimization problem imposing geometric and epipolar constraints. With the corrected set of points the fundamental matrix and the lens distortion model is

computed separately. Since corrected points satisfy all epipolar geometry constraints, the fundamental matrix is computed with a linear method easily. The transformation between distorted and corrected points is done using a distortion model. With the proposed method, errors in point correspondences and localization are avoided since corrected points satisfy both epipolar and geometric constraints. Correction is done before fundamental matrix is computed.

To correct the set of points in both images, geometric constraints defines the point localization in reference to its neighbors in the same image, and epipolar constraints represent the location of one point in relation to its corresponding point in the other image. Geometric constraints arise since a chessboard template is used to calibrate the wide angle stereoscopic system. In the same image, points are arranged in straight lines that are perpendicular and parallel to each other. In addition, since the same scene is taken with both cameras, epipolar constraints help to correct locations of points in one image referred to locations of corresponding points in the other image. These geometric and epipolar constraints define the rules of image formation using two cameras.

Geometric constraints referred with the image of a chessboard are shown in Fig. 1:

- Points are arranged in straight parallel lines (J_{ST}).
- Straight lines are equally distanced. Cross-ratio remains under perspective projection (J_{CR}).
- Four sets of parallel lines exist that are orthogonal to each other. Parallel lines intersect in a vanishing point. (J_{VP}).
- Four vanishing points exist ${}^1j\mathbf{q}_{vp}$, ${}^2j\mathbf{q}_{vp}$, ${}^3j\mathbf{q}_{vp}$, ${}^4j\mathbf{q}_{vp}$ that form the horizon line ${}^j\mathbf{l}_h$. (J_{HL}).

Epipolar constraints connect the two images captured with the stereoscopic system. They are shown in Fig. 2:

- Each image has an epipole ${}^j\mathbf{q}_{ep}$ and each point in each image belongs to an epipolar line ${}^i\mathbf{l}_{ep}$ that goes through the epipole ${}^j\mathbf{q}_{ep}$ (J_{EL}).
- One point of the scene forms two epipolar lines, one in each image ${}^i\mathbf{l}_{ep}$, ${}^2i\mathbf{l}_{ep}$ and both lines form an epipolar plane ${}^i\pi_{ep}$ (J_{EP}).

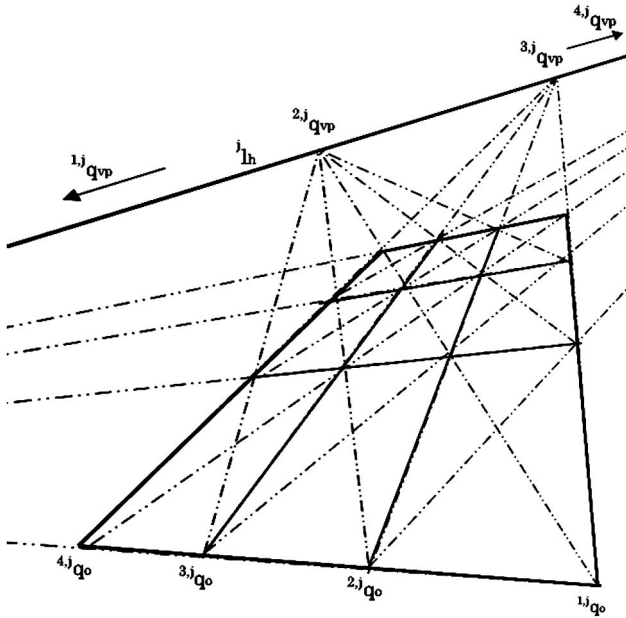


Fig. 1. Geometric constraints that represent the image formation of a set of parallel and perpendicular straight lines.

– All epipolar planes intersect in one line defined by the focal points of both images l_{12} (J_{IL}).

The correct point location is defined by its location in the image of a chessboard template and the location of its corresponding point in the other image. J_{xx} represents the function that measures the corresponding constraint with a set of points. Joining all proposed constraints, the following function measures how well the undistorted points in both images correspond to the points of a “chessboard” template taken with a stereoscopic system.

$$J = J_{ST} + J_{CR} + J_{VP} + J_{HL} + J_{EL} + J_{EP} + J_{IL} \quad (1)$$

and ${}^{ij}q_d$ represent the location of detected points in both images, $i = 1 \dots n$, where n is the number of points in the chessboard template and j is the image 1 or 2. Expression (1) evaluated with points ${}^{ij}q_d$ is nonzero since images are distorted because wide angle lens are mounted in the cameras. Also, noise is present in the point detection step and erroneous point correspondences occur. ${}^{ij}q_o$ represent the points that accomplish geometric and epipolar

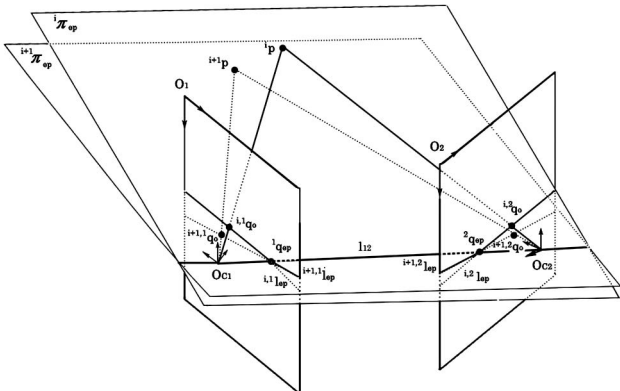


Fig. 2. Epipolar constraints that connect the location of one point in the scene with its location in two images.

constraints perfectly. Expression (1) evaluated with points ${}^{ij}q_o$ is equal to zero. The aim of the nonlinear minimization process is to start with detected points ${}^{ij}q_d$ and to end with ${}^{ij}q_o$, which represent the correct points locations in both images. They represent the correct location of each point in each image since they satisfy both geometric and epipolar constraints.

The nonlinear minimization process has as inputs the cross-ratio value of the template points and the set of points in the image or images ${}^{ij}q_d$. For a given set of points, straight lines parameters, vanishing points ${}^{1,j}q_{vp}$, ${}^{2,j}q_{vp}$, ${}^{3,j}q_{vp}$, ${}^{4,j}q_{vp}$, horizons lines ${}^j l_h$, epipoles in both images ${}^j q_{ep}$, epipolar lines ${}^j q_{ep}$, and epipolar planes ${}^j \pi_{ep}$ are computed. With the computed parameters, function (1) is evaluated and points locations ${}^{ij}q_d$ are corrected to minimize (1). This process is repeated until (1) is zero. When (1) is zero, distorted points detected in both images ${}^{ij}q_d$ have been undistorted to ${}^{ij}q_o$. The algorithm is a Levenberg-Marquardt nonlinear minimization that starts with the set of distorted points ${}^{ij}q_d$ and ends with undistorted points ${}^{ij}q_o$ which satisfy all constraints.

To improve the condition of the nonlinear minimization process, point coordinates are referred to the center of the image and not to the left top corner. If this nonlinear minimization process is compared with the nonlinear minimization step of any well-established fundamental matrix estimation method, this process can be considered better conditioned notably. Nonlinear minimization looks for points coordinates in pixels only. This means that all variables produce equal alteration of error function value when a variation of one unit is done with any of them. The nonlinear minimization step of any of the existing methods looks for values of the fundamental matrix elements that have different magnitudes. With the proposed method, risks of finishing the nonlinear minimization step in a local minimum decrease. On the contrary, the number of variables increases and the minimization time increases significantly. Since this is an off-line process, this fact does not represent any difficulty.

When point locations in both images have been corrected, the fundamental matrix and the lens distortion models are computed separately. One distortion model is computed for each camera. The fundamental matrix is computed with any of the existing linear methods. Although a linear method is used, calibration residual will be zero since corrected points ${}^{ij}q_o$ are used and they satisfy all epipolar constraints. The eight point algorithm proposed in [2] computes the fundamental matrix with the eigen vector associated to the small eigen value of a matrix A . Matrix A is composed with the coordinates of corresponding points in both images. In this case, the small eigen value of A will be zero since all points that compound matrix A satisfy all epipolar constraints. See [2] for details. Lens distortion models are computed using the method proposed in [3]. In this case, lens distortion models represent the transformation between the distorted points ${}^{ij}q_d$ detected in the image and the corrected ones ${}^{ij}q_o$ of both cameras. In [3] several lens distortion models are calibrated and the rational function lens distortion model presented by Claus and Fitzgibbon [4] obtains better results when high distortion is present. See [3] for details.

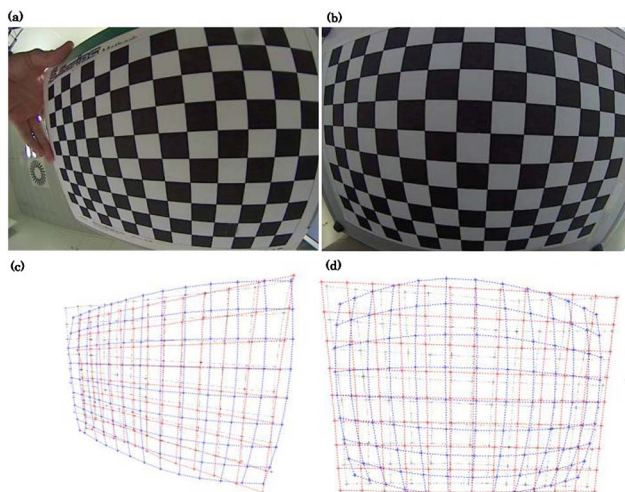


Fig. 3. (Color online) Images from a wide angle stereoscopic system build up with two IP cameras AXIS M3204. Blue dots are detected points in captured distorted images. Red dots represent blue dots corrected with the proposed method. Gray dots are detected points of the image corrected with method 1.

To test the performance of the proposed method, the estimation of the fundamental matrix and the lens distortion models is compared with two out of the several methods presented in [5] where a survey of 19 of the most used methods is done. Two techniques have been chosen. Method 1: the Least-Median-Squares [6] obtain the best results when outliers are present. Distortion is corrected previously. In this case the lens distortion is corrected using the principle of straight lines should be straight presented in [7]. Method 2: the fundamental matrix is computed together with the distortion model with the method proposed in [8]. Point correspondences between images are resolved with the algorithm proposed by Zhang [9]. With methods 1 and 2, corresponding points are normalized with the method proposed in [10]. Experiments with simulated and real data have been done. Simulated data is generated with two pin-hole cameras where distortion, noise and erroneous point correspondences are added. Real data are acquired with a stereoscopic system that has been built up with two IP wide angle lens cameras AXIS M3204. Two images of 640×480 pixels are captured. Figure 3 shows images from the two cameras. Detected points are in blue and corrected points using the lens distortion model computed with the proposed method are in red. The corrected points using the lens distortion model computed with method 1 are shown in gray. Results using method 2 are quite similar to method 1. Gray dots show deficiencies of method 1 and 2 since points should form straight lines. When distortion is corrected they are still distorted. Red dots show that the proposed method corrects the distortion accurately. Numerical results are shown in Table 1. Every cell shows the mean and standard deviation in pixels of the discrepancy between points and epipolar lines and straight lines measured with (1). Results with simulated data show that the computed fundamental matrix and the lens distortion models are accurate if the proposed method is used. Since epipolar geometrical constraints are used in the calibration process, the computed models represent the wide angle stereoscopic system accurately.

Table 1. Experimental Results with Simulated and Real Data

Noise—outliers	Method 1	Method 2	Proposed
1 pixel—0%	0.75 ± 0.06	0.83 ± 0.02	0.02 ± 0.008
2 pixels—0%	1.67 ± 0.27	1.56 ± 0.64	0.45 ± 0.062
0 pixels—15%	2.75 ± 0.62	3.68 ± 0.79	0.24 ± 0.075
1 pixel—15%	3.64 ± 0.95	4.98 ± 1.64	0.68 ± 0.045
2 pixels—15%	5.25 ± 2.35	7.55 ± 3.56	0.69 ± 0.035
Real data	1.15 ± 0.15	1.47 ± 0.24	0.15 ± 0.075

Method 1 obtains better results than method 2 when noise or outliers are present. Since method 2 computed both the fundamental matrix and the lens distortion models in one step, coupling between both models creates an ill conditioned nonlinear minimization process and a local minimum can be reached easily. The proposed method uses the nonlinear minimization to compute the correct location of corresponding points in both images. Since only points locations are searched the process is better conditioned.

In conclusion an accurate method for calibrating a wide angle stereoscopic system has been defined. Wide angle lens distort images and errors in point correspondences and localization arise. With the proposed method detected points are corrected previously using both geometric and epipolar constraints. With corrected points errors in point localization and correspondences are avoided and the fundamental matrix can be computed using any linear method easily. Distortion model is adjusted to map from points detected in the images to the corrected ones. This method improves the nonlinear minimization process since points coordinates are computed only.

The authors would like to thank to the Universidad Politécnica de Valencia (PAID 2010-2431 and PAID 10 017), Generalitat Valenciana (GV/2011/057), and the Spanish government and the European Community under the project DPI2010-20814-C02-02 (FEDER-CICYT) and DPI2010-20286 (CICYT).

References

- Z. Zhang, H. Ma, T. Guo, S. Zhang, and J. Chen, *Opt. Lett.* **36**, 1257 (2011).
- H. C. Longuet-Higgins, *Nature* **293**, 133 (1981).
- C. Ricolfe-Viala and A. Sanchez-Salmeron, *Appl. Opt.* **49**, 5914 (2010).
- D. Claus and A. Fitzgibbon, *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition* (2005).
- X. Armangué and J. Salvi, *Image and Vision Comput.* **21**, 25 (2003).
- Z. Zhang, *International Journal of Computer Vision* **27-2** (1998).
- F. Devernay and O. Faugeras, *Machine Vis. Apps.* **13**, 14 (2001).
- Z. Zhang, *International Conference on Pattern Recognition* (1996).
- Z. Zhang, R. Deriche, O. Faugeras, and Q. T. Luong, *Institut National de Recherche en Informatique et Automatique* (1994).
- R. Harley, *Proceedings of the Eighth International Conference on Computer Vision* (1995).