Calibration of a trinocular system formed with wide angle lens cameras

Carlos Ricolfe-Viala,* Antonio-Jose Sanchez-Salmeron, and Angel Valera

Instituto de Automática e Informática Industrial, Universitat Politècnica de València, Camino de Vera s/n, Valencia, 46022, Spain

*cricolfe@isa.upv.es

Abstract: To obtain 3D information of large areas, wide angle lens cameras are used to reduce the number of cameras as much as possible. However, since images are high distorted, errors in point correspondences increase and 3D information could be erroneous. To increase the number of data from images and to improve the 3D information, trinocular sensors are used. In this paper a calibration method for a trinocular sensor formed with wide angle lens cameras is proposed. First pixels locations in the images are corrected using a set of constraints which define the image formation in a trinocular system. When pixels location are corrected, lens distortion and trifocal tensor is computed.

References and links

1. R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. (Cambridge University Press, 2000).
2. O. Faugeras and B. Mourrain, "On the geometry and algebra of the point and line correspondences between n images," In Proc. 5th Int. Conf. Comput. Vision, 951–962 (1995).
3. R. Hartley, "Lines and points in three views and the trifocal tensor," Int. J. Comput. Vis. **22**(2), 125–140 (1997).
4. P. Torr and A. Zisserman, "Robust parameterization and computation of the trifocal tensor," Image Vis. Comput. **15**(8), 591–605 (1997).
5. O. Faugeras, Q.-T. Luong, and T. Papadopoulo, *Geometry of multiple images*. (Cambridge, MA, USA: MIT Press, 2001).
6. P. Sturm and S. Ramalingam, "A generic concept for camera calibration," in Proc. 10th ECCV, 1–13 (2004)
7. C. Ressl, "Geometry, constraints and computation of the trifocal tensor," *PhD Thesis, Instituüt für Photogrammetrie und Fernerkundung*, Technische Universität Wien, (2003).
8. W. Förstner, "On weighting and choosing constraints for optimally reconstructing the geometry of image triplets," in Proc. 6th ECCV, 669–684 (2000).
9. T. Molinier, D. Fofi, and F. Meriaudeau, "Self-calibration of a trinocular sensor with imperceptible structured light and varying intrinsic parameters," in Proc.7th Int. Conf. Quality Control by Artificial Vision, (2005).
10. D. Nistér, "Reconstruction from uncalibrated sequences with a hierarchy of trifocal tensors," in Proc. 6th ECCV, 649–663 (2000).
11. C. Ricolfe-Viala, A. J. Sanchez-Salmeron, and E. Martinez-Berti, "Accurate calibration with highly distorted images," Appl. Opt. **51**(1), 89–101 (2012).
12. C. Ricolfe-Viala and A. J. Sanchez-Salmeron, "Lens distortion models evaluation," Appl. Opt. **49**(30), 5914–5928 (2010).
13. D. Claus and A. Fitzgibbon, "A rational function lens distortion model for general cameras," In Proc. CVPR IEEE, 213–219 (2005).
14. M. Ahmed and A. Farag, "Nonmetric calibration of camera lens distortion: differential methods and robust estimation," IEEE Trans. Image Process. **14**(8), 1215–1230 (2005).
15. F. Devernay and O. Faugeras, "Straight lines have to be straight," Mach. Vis. Appl. **13**(1), 14–24 (2001).

1. Introduction

In applications such as surveillance, 3D information helps in recognition of persons and objects or in detection of human movements. Surveillance is done over large areas which have to be covered with as few cameras as possible. In consequence, retrieving accurate 3D information from a wide area represents a non trivial problem which can be solved with a

network of cameras. From the point of view of the number of cameras, two cameras could be sufficient to resolve a stereo problem. However, accuracy will be poor since cameras are located far away from the inspected area (normally in the roof) and the number of stereo cameras will be high since a small area is covered with each pair of cameras. To reduce the number of cameras wide angle lens cameras are used although images are distorted significantly. To increase the accuracy of 3D information a trifocal sensor is used. In three views the possibility of removing the effects of measurement errors increase since six measurements of a unique 3D point exist. Moreover, if a scene line which has four degrees of freedom is reconstructed with a trinocular sensor; six measurements arise to reconstruct them. This means that a scene line is over-determined and can be estimated by a suitable minimization over measurements errors. If we have a stereo system with two views only, the number of measurements equals the number of degrees of freedom [1].

A trinocular sensor is represented with the trifocal tensor. The trifocal tensor is considered the generalization of the fundamental matrix for three views since contains all projective geometric relationships between the three views. Given the trifocal tensor and the location of a point or a line in two images, the location in the third one is computed easily. This reduces the searching area and the point matching problem between images notably. Accurate estimation of the trifocal tensor is critical to resolve the cited application efficiently. To estimate the trifocal tensor any knowledge of intrinsic and extrinsic camera parameters are not required, neither of their relative positions. Trifocal tensor is computed using correspondences between points in the three images with linear, non linear or robust methods.

Usually, linear methods give a quick solution which is valid if points are distributed over the image and the problem of correspondences is solved properly [2]. Accuracy increases if methods which include a non linear minimization step are used [3]. In this case, solution of a linear method is used as starting point of the non linear searching. Iterative methods solve the problem of noisy data but if corresponding points between images are mismatched, a deviated solution can result. Robust methods can deal with noisy and errors in point correspondences [4]. Errors in calibration arise with errors in point localization and correspondences [1,5–8].

With wide angle lens cameras, images are high distorted and errors in point correspondences and point localization increase significantly. To calibrate a trinocular system composed with three wide angle lens cameras, first lens distortion is corrected and second the calibration algorithm is solved. Since images are not undistorted totally, accurate pixel localization is complicated even though images have been corrected previously. Therefore, errors in point correspondences remain. Also, trifocal tensor and lens distortion models can be calibrated together. Since both models are coupled, computed solution can be absurd [9,10].

In this paper a novel method for calibrating a trinocular sensor with three wide angle lens cameras is described. The proposed method corrects the image distortion efficiently and avoids errors in point correspondences. First, pixels locations which are going to participated in the calibration process are corrected. Pixel locations in the distorted image are corrected by using a set of constraints which define the image formation and a set of geometric relations which identify a trinocular sensor. This correction is done before trifocal tensor and lens distortion is calibrated. When pixels locations are corrected, the trifocal tensor and the lens distortion models are calibrated separately with some of existing methods. Trifocal tensor can be computed with a linear method since corrected points satisfy geometric constraints of points captured with a trinocular sensor. Errors in point correspondences are avoided since corrected points satisfy all geometric constraints which govern the image formation with a trinocular sensor.
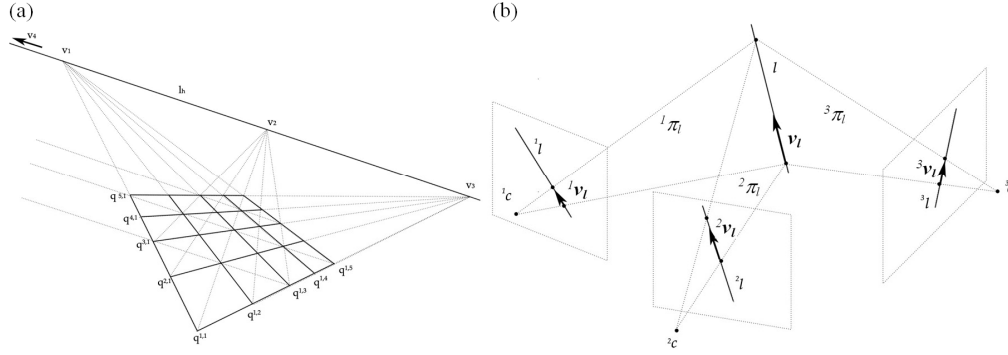
Fig. 1. (a) Geometric constraints for correcting points locations detected in the image. Cross ratio guaranties that parallel lines remain parallels under perspective projection. Points are corrected to belong to straight lines. All parallel lines meet in a vanishing point in an image. All vanishing points are in the horizon line. (b) Trifocal tensor constraints used for pixel location correction in distorted images. Planes which contain the lines in the images meet in the 3D line of the calibration template. Pixels locations in one image are corrected in reference with their corresponding points in the remaining two images.

2. Pixel location correction

To correct pixels location in the distorted image a calibration template is used. Since the three wide angle lens cameras have to see the calibration template simultaneously, a planar calibration template has been chosen to calibrate the trinocular sensor. If the distance between cameras were small a chessboard can be used to calibrate the trinocular system. Since we assume that cameras are installed to see large areas, a chessboard is too small to calibrate the sensor. Therefore, figs. in the floor of the room where cameras are installed, are going to be used to calibrate the trinocular sensor. Figures 3(a),(b),(c) shows one image of the room where cameras are installed. Straight lines of the floor will help us to calibrate the trinocular sensor accurately. If corners of the tiles are used as calibrating points, images of calibrating points from three cameras should accomplish all projective geometric relationships under which images are formed.

On one hand, some constraints arise taking into account distribution of points over the calibration template. Corners of the tiles are located in straight lines which are equally distanced perpendicular and parallel each other. Also, two parallels lines in the scene meet in a vanishing point in the image and several vanishing points gives the horizon line in the image. Therefore, the cross-ratio, straight lines, vanishing points and horizon line can be imposed as constraints to correct point location of distorted image. Constraints which affect to one image alone are shown in Fig. 1(a). Equations (1)-(4) arise from each constraint. (1) is from cross-ratio, (2) from straight lines, (3) from vanishing points and (4) form horizon line. $CR(p_1, p_2, p_3, p_4)$ is the cross ratio of four collinear points defined when the calibration template is built. $q^{k,l} = (u^{k,l}, v^{k,l}, w^{k,l})$ represents a point k in a line $l = (a_l, b_l, c_l)$. In [11] an exhaustive definition of each Eq. is described.

$$J_{CR} = \sum_{l=1}^{n} \sum_{k=1}^{m-3} \left(CR(q^{k,l}, q^{k+1,l}, q^{k+2,l}, q^{k+3,l}) - CR(p_1, p_2, p_3, p_4) \right)^2 \qquad (1)$$

$$J_{ST} = \sum_{l=1}^{n} \sum_{k=1}^{m} d(q^{k,l}, l) = \sum_{l=1}^{n} \sum_{k=1}^{m} \frac{|a_l \cdot u^{k,l} + b_l \cdot v^{k,l} + c_l \cdot w^{k,l}|}{\sqrt{a_l^2 + b_l^2}} \qquad (2)$$

$$J_{VP} = \sum_{i=1}^{4} \sum_{l=1}^{t} d(q_{vp\,i}, l_l) = \sum_{i=1}^{4} \sum_{l=1}^{t} \frac{|a_l \cdot u_{vp\,i} + b_l \cdot v_{vp\,i} + c_l \cdot w_{vp\,i}|}{\sqrt{a_l^2 + b_l^2}} \qquad (3)$$

$$J_{HL} = \sum_{i=1}^{4} d(q_{vp\,i}, l_h) = \sum_{i=1}^{4} \frac{\left| a_h \cdot u_{vp\,i} + b_h \cdot v_{vp\,i} + c_h \cdot w_{vp\,i} \right|}{\sqrt{a_h^2 + b_h^2}} \tag{4}$$

On the other hand, pixels locations in one image can be corrected taking into account the location of their corresponding points in the remaining two images according with the trifocal tensor constraint. If a line is taken with tree cameras, a geometric relationship exists between the three images which are show in Fig. 1(b) which shows the trinocular sensor constraint. A 3D line in the calibration template projects in three lines in the images. Also, the planes which contain the lines in the images meet in the 3D line of the calibration template. This constraint is called geometric incidence in a trinocular system and is imposed as a constraint in the pixel location correction process. In consequence, pixels locations in one image are corrected in reference with their corresponding points in the remaining two images.

$^{i}a_l$, $^{i}b_l$, $^{i}c_l$ represent the set of parameters which defines the straight line ^{i}l in the image i, and $^{i}v_l$ is the vector in the direction of the same line, $i = 1..3$. The vector with the direction of the corresponding line in the scene l is called v_l. Since the line l and the projection in the image ^{i}l are in the same plane $^{i}\pi_l$, vectors $^{i}v_l$ and v_l have to be perpendicular to the normal vector of plane $^{i}\pi_l$. If $^{i}n_l$ is the normal vector of the plane $^{i}\pi_l$ following Eq. is true.

$$\begin{bmatrix} v_l \\ ^{i}v_l \end{bmatrix} \cdot {^{i}n_l} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{5}$$

In a matrix form we have $^{i}V_l \cdot {^{i}n_l} = 0$. Line l is projected in the three images ^{1}l, ^{2}l, ^{3}l and a set of s lines exist, where s is the number of lines in the calibration template. If Eq. (5) is satisfied by all straight lines following Eq. arise:

$$J_{TS} = \sum_{i=1}^{3} \sum_{l=1}^{s} {^{i}V_l} \cdot {^{i}n_l} \tag{6}$$

Joining the trifocal tensor constraint and constraints of Eqs. (1)-(4), the following Eq. measures how a set points from a trinocular sensor, correspond to the points of a "*chessboard*" template.

$$J = J_{CR} + J_{ST} + J_{VP} + J_{HL} + J_{TS} \tag{7}$$

The non linear searching has as inputs the cross ratio value of the template points CR(p_1, p_2, p_3, p_4) and the set of points in the image or images q_d. The minimization process is as follows. For a given set of points, straight lines parameters (a_l, b_l, c_l), vanishing points $^{1,j}q_{vp}$, $^{2,j}q_{vp}$, $^{3,j}q_{vp}$, $^{4,j}q_{vp}$, horizons lines $^{j}l_h$, and normal vector of planes $^{i}\pi_l$ are computed. With the computed parameters, function (7) is evaluated and points locations q_d are corrected to minimize (7). This process is repeated until (7) is zero. When (7) is zero, distorted points extracted from the image q_d have been undistorted to q_o. To minimize (7) a Levenberg-Marquardt non linear minimization algorithm is used which starts with the set of points in the image or images q_d and ends with the set of undistorted points q_o. Figure 2(a) shows the original distorted points and the corrected ones. To improve the condition of the non-linear searching process points coordinates are referred to the center of the image and not to the left top corner.

3. Computing the trifocal tensor

Once distorted pixel locations are corrected the trifocal tensor is computed. with corrected points q_o and calibration template points p. Any of existing linear methods can be used since point location in the image has been corrected. With the point location correction errors in point correspondences does not exist and the noise is reduced.

The normalized linear algorithm described in [1] has been chosen. With a minimum number of 7 image point correspondences across 3 images, or at least 13 lines, or a mixture of points and lines correspondences, the trifocal tensor is computed. First images are normalized

and the trifocal tensor is computed using singular value decomposition of a matrix composed with points coordinates and lines parameters. (see [1] for details)
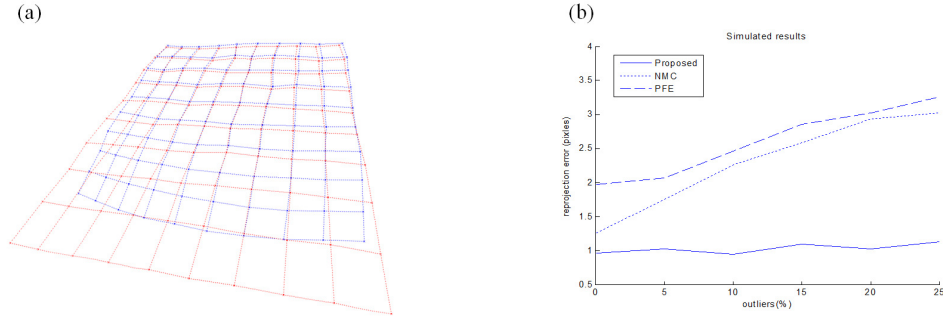
(a)

(b)

Fig. 2. (a) Blue dots represent detected points in the distorted image of Fig. 3(b). Red dots represent undistorted points corrected with the proposed method. Pixels are moved to accomplish all projective geometric and trifocal sensor constraints. (b) The number of errors in points correspondences between images (outliers) is modified. Results with the proposed method do not depend on the number of outliers since point localizations are corrected before trifocal tensor is computed.

4. Experimental results

Summarizing, for calibrating a trifocal tensor with distorted images, first distortion is corrected and second, trifocal tensor is calibrated with undistorted images. The proposed method corrects pixels locations in the distorted images using geometric and trifocal tensor constraints. Second, trifocal tensor is calibrated with corrected points. If any other existing method for distortion correction were used, images are undistorted and the trifocal tensor is calibrated with undistorted images in a second step. Several methods for lens distortion correction exists [12,13]. To test the efficiency of the proposed method the trifocal tensor is calibrated also with two existing lens distortion correction methods, the *non metric calibration of lens distortion* (NMC) proposed by Ahmed in [14] and the *polynomial fish-eye lens distortion correction* (PFE) presented by Devernay and Faugeras in [15]. Once images are undistorted, the trifocal tensor is calibrated with the method proposed in subsection 3. The calibration template is the tiles on the floor of the room of 5m x 5m approximately. Tiles on the floor form a checkerboard with 100 corners points (10x10), similar to that shown in Fig. 3. Calibration error is given with the reprojection error since the three camera matrices may be computed from the computed trifocal tensor separately. Camera matrices give the independent 3D projective transformation for each camera and lens distortion models are also computed for each camera alone. The reprojection error measures the discrepancy between pixel coordinates, $*q = (*u, *v)$ projected from measured world coordinates by the estimated camera matrix and undistorted pixel coordinates $*q_o = (*u_o, *v_o)$ computed with the estimated lens distortion model.

$$e_d = \frac{1}{n}\sum_{i=1}^{n}\sqrt{\left(*u - *u_o\right)^2 + \left(*v - *v_o\right)^2} \tag{8}$$

We have used three IP cameras Axis 212 PTZ with 2.7 mm lens mounted which gives 85° field of view. Figure 3 shows three acquired images of 640x480 pixels with considerable distortion. Images 3(d),(e),(f) shows the undistorted images with the proposed method and the set of corresponding points which are used to calibrate the trifocal tensor using the linear method described in subsection 3. In this case, corresponding lines and corresponding points are used to calibrate the trifocal tensor. Figure 2(a) shows the points locations of the original distorted and the undistorted image. Corrected points are moved to accomplish all projective geometric and trifocal sensor constraints.
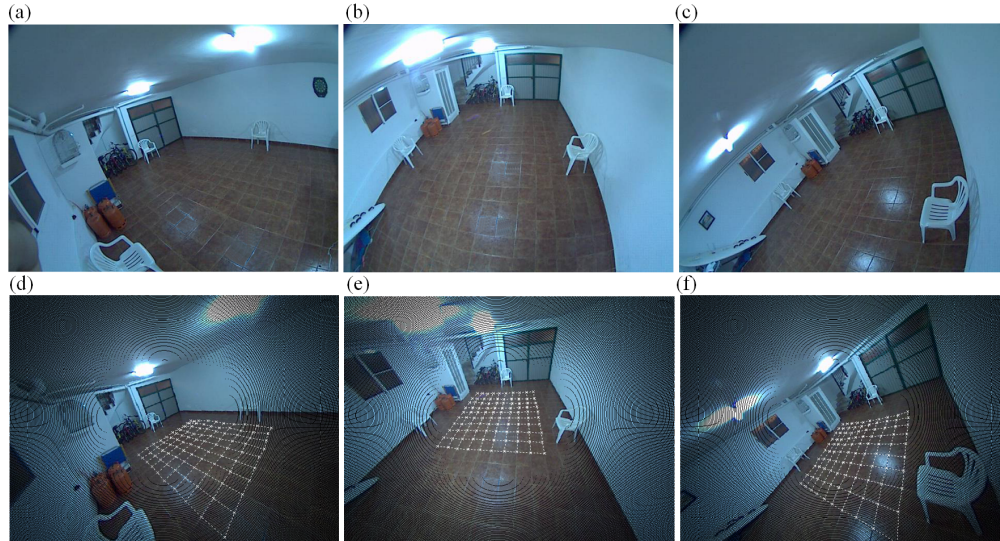
Fig. 3. Results with real data. (a)(b)(c) show 640x480 captured images with an Axis 212 PTZ with 2.7 mm lens mounted which gives 85° field of view. (d)(e)(f) show images corrected with the proposed method. White dots are the set of corresponding points which are used to calibrate the trifocal tensor.

An experiment changing the number of outliers from 0% to 25% has been done. Computed values are shown in Fig. 2(b). The reprojection error with the proposed method does not depend on the number of outliers since pixels location has been corrected before trifocal tensor is computed. Moreover, if we compare the reprojection error when the number of outliers is zero, the proposed method obtains better result since the noise is reduced when pixels locations are corrected. With the NMC and PFE methods, images are not absolutely undistorted and accurate pixel localization is complicated.

5. Conclusion

A method for calibrating a trinocular system formed with wide angle lens cameras has been defined. A planar calibration template with parallel and perpendicular straight lines is used. Pixels location in the distorted image are corrected before trifocal tensor and lens distortion model are computed. Correction is done by using geometric and trifocal tensor constraints which govern the images formation with a trinocular system. Distortion model is adjusted to map from points detected in the images to the corrected ones. Trifocal tensor is computed with corrected points and points from calibration template using a linear method. Proposed method avoids errors in point correspondences between images which influences significantly the result of a trinocular system calibration. Also, noise is reduced notably.