AUTOMATIC CALIBRATION OF STEREO-CAMERAS USING ORDINARY CHESS-BOARD PATTERNS

A. Prokos¹, I. Kalisperakis², E. Petsa², G. Karras¹

¹ Department of Surveying, National Technical University of Athens (NTUA), GR-15780 Athens, Greece ² Department of Surveying, Technological Educational Institute of Athens (TEI-A), GR-12210 Athens, Greece

e-mail: anthpro@central.ntua.gr, ikal@teiath.gr, petsa@teiath.gr, gkarras@central.ntua.gr

KEY WORDS: Automation, Bundle, Calibration, Stereoscopic, Distortion, Extraction, Non-Metric

ABSTRACT

Automation of camera calibration is facilitated by recording coded 2D patterns. Our toolbox for automatic camera calibration using images of simple chess-board patterns is freely available on the Internet. But it is unsuitable for stereo-cameras whose calibration implies recovering camera geometry and their true-to-scale relative orientation. In contrast to all reported methods requiring additional specific coding to establish an object space coordinate system, a toolbox for automatic stereo-camera calibration relying on ordinary chess-board patterns is presented here. First, the camera calibration algorithm is applied to all image pairs of the pattern to extract nodes of known spacing, order them in rows and columns, and estimate two independent camera parameter sets. The actual node correspondences on stereo-pairs remain unknown. Image pairs of a textured 3D scene are exploited for finding the fundamental matrix of the stereo-camera by applying RANSAC to point matches established with the SIFT algorithm. A node is then selected near the centre of the left image; its match on the right image is assumed as the node closest to the corresponding epipolar line. This yields matches for all nodes (since these have already been ordered), which should also satisfy the 2D epipolar geometry. Measures for avoiding mismatching are taken. With automatically estimated initial orientation values, a bundle adjustment is performed constraining all pairs on a common (scaled) relative orientation. Ambiguities regarding the actual exterior orientations of the stereo-camera with respect to the pattern are irrelevant. Results from this automatic method show typical precisions not above ¼ pixels for 640x480 web cameras.

1. INTRODUCTION

Estimation of the camera geometry parameters represents a fundamental task in photogrammetry and computer vision. Camera calibration approaches (reviewed in Clarke & Fryer, 1998, Salvi *et al*., 2002, Villa-Uriol *et al*., 2004) differ widely e.g. regarding number of images involved, used camera models and algorithms or type of observed features. Although camera calibration is, indeed, possible without *a priori* object information (simple point matches on >2 frames from the same camera allow self-calibration), the use of reliable external control ensures calibration data which also satisfy such object space constraints. Since, furthermore, in close-range applications it is often preferable to pre-calibrate cameras via suitable image networks (Remondino & Fraser, 2006), most approaches are based on targeted test-fields and target-image correspondences. However 3D test-fields may well be replaced by simpler 2D patterns, typically of the chess-board type, imaged on multiple views. If, to quote Fiala & Shu (2005), cameras should, ideally, be automatically calibrated via rapidly taken images, coded 2D patterns are particularly suitable for the purposes of automation. Thus one finds freely available tools relying on chess-board patterns, recorded in different perspective views, for determining interior and exterior camera orientation.

Such tools have been inspired by the "plane-based calibration" approach (Sturm & Maybank, 1999; Zhang, 1999), which relies on the homographies between a plane of known metric structure and its images. The linear system in the basic camera elements provided by these transformations results in a closed-form solution, usually followed by a non-linear refinement step. The best known among such tools is the *Camera Calibration Toolbox for Matlab*[®] of J.-Y. Bouguet (implemented also in C++ and included in the Open Source Computer Vision library distributed by Intel). This algorithm, initialized by manual pointing of the four chess-board corners on all images and *a priori* knowledge of the number of nodes per row and column, identifies the nodes on all images with sub-pixel accuracy (for strong lens distortion input

of approximations may also be required). With initial values for the unknown parameters given by the closed-form plane-based calibration algorithm, an iterative bundle adjustment refines the solution for camera and pose elements. Similar approaches may be found on the cited Bouguet website. Particular reference is to be made to the *DLR CalDe*–*DLR CalLab®* calibration toolbox (see cited website), whose stereo-camera calibration procedure runs fully automatically if the chess-board includes three special circular targets at its centre, else such points must be introduced manually. Recent publications on automatic camera calibration using chess-board patterns include de la Escalera & Armingol (2010), Kassir & Peynot (2010), Narayanan & Bijlani (2011).

In this context, we have presented a fully automatic toolbox for camera calibration (Douskos *et al*., 2009), which is freely available on the Internet (FAUCCAL, 2009). It relies on images of standard chess-board patterns, under the single assumption that the light and dark squares are of equal size. Among extracted interest points only those are kept which may be ordered in two groups of lines referring to the main orthogonal directions of the planar pattern. To establish point matches among views pattern regularity is exploited: the lowest line in each image is assumed as the X-axis; the pattern line on the far left serves as the pattern Y-axis. The fact that, obviously, homologous image points thus determined do not necessarily refer to the same physical pattern node introduces ambiguity in rotation, translation and scale; but this affects only image exterior orientations, which are totally irrelevant in this case. Using approximations of parameter values drawn from the information embedded in the image vanishing points, the final bundle adjustment allows estimating the camera geometry parameters in a fully automatic manner. It has been shown that the method gives accurate camera calibration results.

Stereo-cameras, namely a camera pair in fixed relative position, are now often used, mainly for 3D reconstruction purposes. Yet, since no reference system in object space is available, it is clear that the above-mentioned approach is unsuitable for calibrating

stereo-cameras. Calibration of such two-camera systems means not only calibrating of both cameras but also determining the 6 parameters of their true-to-scale relative orientation. To find the latter in an automatic mode, approaches reported in literature require additional coding or targets on the chess-board pattern to fix the object space coordinate system. This, for instance, is the case of the *DLR CalDe*–*DLR CalLab®* toolbox already referred to, but also of the 3D scanner reported in Prokos et al. (2010, 2011), where the colour of one chess-board square is changed to allow automatic calibration of the stereo-camera system used.

We present here an automatic calibration toolbox for stereo-cameras based on ordinary chess-board patterns, i.e. with no extra coding or targets. First the calibration algorithm is applied separately for each camera to images of a simple chess-board pattern taken with the stereo-camera. Pattern nodes are extracted, then ordered in rows/columns and finally used for finding the two independent camera parameter sets. Unknown remain, of course, the actual node correspondences on the stereo pairs, which are indispensable for relative orientation. The missing piece of information is contributed by an image pair of some "reasonably" textured 3D scene which is also acquired with the stereo-camera. Image matches are then established automatically using the SIFT operator, which allows determination of the fundamental matrix of the stereo-camera. This knowledge of 2D epipolar geometry enables the algorithm to establish correct node correspondences for all stereo-pairs of the chess-board pattern, which finally are introduced into an adjustment for full stereo-camera calibration.

2. SINGLE CAMERA CALIBRATION

Since the main features of the camera calibration toolbox have been reported in detail in Douskos *et al.* (2009) and documented in FAUCCAL (2009), a brief outline will suffice here.

2.1 Initialization

• *Corner extraction*. The Harris corner operator with sub-pixel precision (made available in the website of Bouguet) is applied to grayscale images with equalized histograms. Standard errors of bundle adjustments support the claim that a precision of ~ 0.1 pixel is generally feasible.

• *Node selection and ordering.* On each image, the feature point closest to the median coordinates of all extracted feature points is chosen as starting point. Criterion as to whether this point and its closest neighbour represent valid nodes is the difference in gray value between either sides of the linear segment defined by these points, which should be large (this also avoids pattern diagonals). Founded on this simple idea, the algorithm extracts the two main pattern directions, and then extends its search until all possible pattern lines have been identified. It is noted that the algorithm may also accommodate "gaps", namely missing nodes or even rows and columns. Next, pattern lines are ordered. The line through the original starting point forming the smaller angle with the image x-axis establishes the rows; the line in the other direction fixes the columns. Rows and columns are then sorted. Certain precautions are taken to eliminate possible blunders and ensure convergence of bundle adjustment. Initially, for instance, only extracted points which belong to both a row and a column are accepted as valid chess-board nodes in the calibration adjustment; however, all other valid pattern nodes thus discarded are "regained" by the algorithm in a next step.

• *Point correspondences*. Final outcome of preceding steps is a set of points coded according to the respective chess-board rows and columns with which they have been associated. As already mentioned, the lower row appearing on each image is arbitrarily considered as the object X-axis; the column to the far left is associated with the object Y-axis. Hence, thanks to the symmetric nature of the pattern, it is assumed that point correspondences among frames, as well as their correspondences with the chessboard nodes, have been fully fixed. This answers the problem of point matching for the purpose of camera calibration. Point correspondences among views as established here will not necessarily refer to identical physical nodes as all images refer to their own object systems, which may differ by in-plane shifts and rotations. In fact, in a camera calibration process with 2D control (note that the pattern spacing is also given an arbitrary size) it is the perspective image distortions which really matter, i.e. their relation to the planar object and not to a system fully fixed in object space.

• *Initial parameter values.* Estimation of approximate values for the unknowns is based here on the vanishing points of the two principal chess-board directions which are found by line-fitting to points already classified in pencils of converging image lines (this also allows estimating the coefficients of the lens distortion polynomial). Vanishing points near infinity are also accommodated. Details on estimation of initial values for interior and exterior orientation are given in Douskos *et al.* (2008). An alternative approach for finding initial values also implemented here estimates camera constant via the vanishing points and adopts a von Gruber parameterization (Bender, 1971) for estimating the remaining 8 interior and exterior orientation parameters from the homographies between images and the chess-board plane.

2.2 Camera calibration adjustment

• *Mathematical model.* With established image-to-pattern point matches and initial parameter values, an iterative bundle adjustment using the collinearity equations is next performed to estimate camera geometry. A typical camera matrix is used: besides camera constant c and principal point location (x_0, y_0) it incorporates image aspect ratio (equivalently camera constants c_x , c_y) and image skewness. Together with coefficients k_1 , k_2 for radial symmetric lens distortion, decentering distortion coefficients p_1 , p2 may also participate, although in current digital cameras their effect appears as negligible compared to sensor resolution, thus representing a source of instability (Zhang, 1999).

• *Refinement through back-projection.* In the initial adjustment only points identified on both a row and a column of the pattern are involved (to 'double-check' the validity of identified nodes). Discarded valid nodes are chiefly situated on the outer rows and columns; thus image bundles are "narrowed". A remedy is to recover such valid nodes by back-projecting XY pattern node coordinates onto the images using the information gained from the initial bundle adjustment. Points identified on at least one image are projected on all other images to detect missing nodes via a search within a window around the projected point; it is checked whether these points also belong to a column or row. Significant portions of columns and rows may be "regained" in this fashion. Further, three additional rows and columns on either side of the identified chess-board edges are back-projected onto all images. This is intended to "widen" the bundles of rays by identifying acceptable outer rows or columns of the pattern which may have been missed. Concluding, a final bundle adjustment for camera calibration is carried out using all identified points.

The above method is applied also in the case of stereo-cameras in order to calibrate the two cameras independently, since this information will be used in the following.

3. STEREO-CAMERA CALIBRATION

3.1 Mathematical model

The described algorithm extracts the chess-board nodes, orders them and determines all camera geometry parameters. But since here, next to the calibration data of the two cameras, the true-toscale relative orientation of stereo-camera is also required, input to the modified algorithm must be synchronized image pairs of a chess-board pattern of known grid spacing. Of course the conventional collinearity equation used for the first camera

$$
\begin{bmatrix} x - x_{o1} \\ y - y_{o1} \\ -c_1 \end{bmatrix} = \lambda \cdot R_1 \begin{bmatrix} X - X_{o1} \\ Y - Y_{o1} \\ Z - Z_{o1} \end{bmatrix}
$$

is modified for the second camera to accommodate the matrix of relative rotations \mathbf{R}_{12} and the three base components B_x , B_y , B_z :

$$
\begin{bmatrix} x - x_{o2} \\ y - y_{o2} \\ -c_2 \end{bmatrix} = \lambda \cdot R_{12} \cdot R_1 \cdot \left(\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} - \left(\begin{bmatrix} X_{o1} \\ Y_{o1} \\ Z_{o1} \end{bmatrix} + R_1^T \cdot \begin{bmatrix} B_x \\ B_y \\ B_z \end{bmatrix} \right) \right)
$$

with λ : scale factors, c: camera constants, (x_0, y_0) : image coordinates of principal points, **R**: image rotation matrices, (X_0, Y_0, Z_0) : space coordinates of the projection centers. The camera model is the same as before. The calibration adjustment yields the interior orientation parameters of both cameras along with the 6 parameters defining the relative rotation matrix $\mathbf{R}_{12}(\omega,\varphi,\kappa)$ of the two cameras and the base components (B_x, B_y, B_z) of the stereocamera. But this procedure pre-supposes that nodes among the images of stereo-pairs have already been correctly matched.

3.2 Establishment of correct node matches

The missing piece of information is obtained here with the help of one or more image pairs of a (sufficiently textured) 3D scene recorded with the stereo-camera. The SIFT operator is applied to such image pairs in order to extract points whose descriptors allow establishing image point homologies (Lowe, 2004). Relying on such point correspondences, one may recover the 2D epipolar geometry of the stereo-camera as represented by its fundamental matrix **F**. Point homologies are thus refined with the help of the RANSAC algorithm (Fischler & Bolles, 1981; Hartley & Zisserman, 2000) to satisfy the epipolar geometry of the stereo-pair, i.e. only the inlying correspondences of the fundamental matrix of the image pair are accepted (the precision of this process is strengthened thanks to the correction of lens distortions known from the previous camera calibration step).

The algorithm then uses this information on epipolar geometry (represented by the fundamental matrix) to correctly match on the stereo-pairs nodes of the chess-board pattern which are to be used for calibration. A node is initially selected near the centre of the left image. Its match on the right image is assumed at the node closest to its corresponding epipolar line (see Fig. 1). As a consequence, matches for all pattern nodes of image pairs are produced automatically since nodes have already been ordered (no room is left for an ambiguity of $\pm 90^\circ$ in roll angle κ because of the fixed configuration of the a stereo-camera). Of course, all paired nodes should also satisfy the epipolar constraint. Therefore, if the RMS distance of all nodes from their corresponding epipolar lines exceeds a threshold, the algorithm proceeds to selecting the node second closest to the homologue epipolar line, and so on. Actually, the algorithm performs this for 15 nodes and chooses that with the smaller RMS distance from epipolar lines, provided that this value is not above a threshold (an empirical value of 5 pixels has been set here to allow for the uncertainty in the estimation of **F**); if this is not the case, the particular stereo-pair will not take part in the solution. Other measures need also to be taken for avoiding instances of mismatching. A danger of false matching arises, for example, if recorded pattern lines run nearly parallel to epipolar lines. Hence, it is generally important to acquire stereo-pairs whose base will not be parallel to the plane of the pattern. Such precautions have proved to be sufficient in all tests performed so far. Of course, further elaboration is possible for checking the relative position of pattern and epipolar lines.

Figure 1. Node on the left image, corresponding epipolar line on the right image and the six nodes closest to the line. The green point is among them the closest and defines the correct match.

After this procedure has been successfully carried out for a sufficient number of stereo pairs of the chess-board, a final bundle adjustment is performed under the constraint that all stereo-pairs share an identical true-to-scale relative orientation. Initial values for the exterior orientations of the two cameras are obtained as in the case of single camera calibration, from which values for the 6 relative orientation parameters may be approximated. The main output of the adjustment is the interior orientations of the two cameras and their (correctly scaled) relative orientation. As long as the ambiguity in scale has been removed, the remaining ambiguity of the in-plane translation and rotation with respect to the pattern at each position of the stereo-camera is unimportant. The results of the automatic approach described above are considered as satisfactory, with typical *a posteriori* standard errors generally not exceeding ¼ pixel when employing low resolution web cameras.

4. EXAMPLE OF APPLICATION

A total of 8 stereo-pairs of a chess-board pattern were recorded using a pair of a 640x480 web cameras fixed with convergent optical axes. Pattern nodes were ordered and the parameters of interior orientation were determined independently for the two web cameras by means of the single camera calibration toolbox FAUCCAL. The results are seen in Table 1.

A textured 3D scene was also captured with the stereo-camera. As described in the previous section, points were extracted and matched using the SIFT operator. These point correspondences were then filtered with the RANSAC algorithm, which was used in the computation of the fundamental matrix. In Fig. 2 the final valid point matches are presented.

Figure 2. The stereo-pair of the auxiliary 3D scene and all SIFT point matches involved in the computation of the fundamental matrix using the RANSAC algorithm.

By exploiting the epipolar constraint as outlined above, all possible node correspondences were then established between the frames of all stereo-pairs (the image pair seen in Fig. 1 is one of the stereo-pairs used). In most, but not all, cases the first node closest to the corresponding epipolar line proved to provide the correct match. RMS distances of nodes from homologue epipolar lines for the 8 stereo-pairs were in the range 1.7–3.4 pixels. Fig. 3 shows four examples of matched nodes in stereo-pairs.

Figure 3. Node matches in four stereo-pairs of the chess-board pattern which were established thanks to the epipolar constraint and then involved in the stereo-camera calibration adjustment.

Based on these node correspondences, the final bundle adjustment resulted in the parameter values presented in Table 2.

The standard errors of the adjustment and the parameter values appear to be satisfactory. Compared to results from independent solutions shown in Table 1, the standard error σ_0 of the adjustment is here slightly higher; differences also appear in camera parameter values. Such differences are basically attributed to the additional constraint imposed by relative orientation, which allows a one-step adjustment and thus introduces additional correlations of camera parameter values. Highest are differences in the location of the camera principal point; according to Ruiz *et al*. (2002), however, the variability of its estimations is generally considered as higher compared to other camera elements, in particular if small to moderate fields of view are involved (here the field of view of the cameras used is 45°).

But it is not always trivial to evaluate the precision of a camera calibration procedure (Ruiz *et al*., 2002, refer to the controversy as regards the precision of the camera parameters required for obtaining acceptable reconstructions). A more straightforward criterion for the quality of stereo-camera calibration would be its effect on a 3D reconstruction. The calibrated stereo-camera system was, thus, combined with a hand-held laser plane in the 3D slit-scanner approach as presented in detail in Prokos *et al*. (2010). Homologue points are found on corresponding epipolar lines as intensity peaks of the laser trace on the surface, and are then used for its automatic reconstruction. A cylindrical plumbing tube of nominal diameter 125 mm was scanned from one position. A stereo-pair used for reconstruction is seen in Fig. 4.

Figure 4. Stereo-pair used in the reconstruction of the cylinder.

The collected surface points covered \sim 35% of its perimeter. The cylinder interpolated to the 9104 XYZ values of the point cloud showed a standard error of 0.22 mm. The diameter was approximated as 124.92 ± 0.02 mm. These values are practically coincident with those in Prokos *et al*. (2010), where stereo-camera calibration had been based on special coding of the chess-board pattern and surface-fitting to 3670 XYZ points had resulted in a standard error of 0.20 mm.

5. CONCLUSION

Camera calibration, which is an indispensable intermediate step in several photogrammetric and computer vision tasks, may be conveniently performed in a fully automatic mode using simple coded 2D patterns, usually of the ordinary chess-board type. If, however, information regarding the position and orientation of cameras in 3D space is needed, the common answer is additional coding or targets fixed on the pattern. In this contribution it was demonstrated that it is indeed also possible to calibrate automatically a stereo-camera system (i.e. estimate the two parameter sets of the cameras and 6 parameters defining their true-to-scale relative orientation) using ordinary chess-board patterns. This is based on exploiting the fixed epipolar geometry of the system to establish correct correspondences between pattern points on the images of the pair. This geometric relation, expressed through the fundamental matrix, is found by using a stereo-pair of some 3D scene taken with the camera system. Thus, the required input includes: a chess-board pattern of equal and accurately known spacing; a number of image pairs of this pattern under varying perspectives; a stereo-pair of a textured 3D scene. The actual exterior orientations of cameras will still remain unknown after calibration; but scale is recovered, which allows full calibration of the stereo-camera system. Results from a practical test have indicated that this approach can produce precise results. As has been mentioned, a possible future elaboration will be to check angles formed between pattern lines and epipolar lines, in order to further minimize the possibility of ambiguous node matching.

REFERENCES

Bender, L.U., 1971. Analytical Photogrammetry: A Collinear Theory. Final Technical Report RADC-TR-71-147, Rome Air Development Center, New York.

Bouguet J.-Y., *Camera Calibration Toolbox for Matlab*: http://www.vision.caltech.edu/bouguetj/calib_doc/

Clarke T.A., Fryer J.G., 1998. The development of camera calibration methods and models. Photogrammetric Record, 16(91), pp. 51-66.

de la Escalera A., Armingol J.M., 2010. Automatic chessboard detection for intrinsic and extrinsic camera parameter calibretion. Sensors, 10(3), pp. 2027-2044.

DLR CalLab – CalDe Camera Calibration Software: http://www.dlr.de/rm/desktopdefault.aspx/tabid-4853/6084_read-9197/

Douskos V., Kalisperakis I., Karras G., Petsa E., 2008. Fully automatic camera calibration using regular planar patterns. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, 37(5), pp. 21-26.

Douskos V., Grammatikopoulos L., Kalisperakis I., Karras G., Petsa E., 2009. FAUCCAL: an open source toolbox for fully automatic camera calibration. XXII CIPA Symposium, October 11- 15, Kyoto, Japan.

FAUCCAL (2009): http://portal.survey.ntua.gr/main/labs/photo/staff/gkarras/fauccal.html

Fiala, M., Shu, C., 2005. Fully automatic camera calibration using self-identifying calibration targets. Technical Report no. 48306/ ERB-1130, NRC Canada, pp. 26.

Fischler, Μ., Bolles, R., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the Association for Computing Machinery, 24(6), pp. 381-395.

Hartley, R., Zisserman, A., 2000. Multiple View Geometry in Computer Vision. Cambridge University Press.

Kassir A., Peynot T., 2010. Reliable automatic camera-laser calibration. Proc. Australasian Conference on Robotics & Automation, Brisbane, 1-3 December.

Lowe D.G., 2004. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2), pp. 91-110.

Narayanan A.S., Bijlani K., 2011. A practical solution to checker pattern detection based on contour validation. International Conference on Image Processing, Computer Vision, and Pattern Recognition, July 18-21, Las Vegas, USA.

Prokos A., Karras G., Petsa E., 2010. Automatic 3D surface reconstruction by combining stereovision with the slit-scanner approach. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 38(5), pp. 505-509.

Prokos A., Kalisperakis I., Karras G., 2011. Automatic point cloud generation and registration with a stereovision slit-scanner. $4th$ International Workshop on 3D Virtual Reconstruction $\&$ Visualization of Complex Architectures (3D-ARCH 2001), 2-4 March, Trento, Italy.

Remondino F., Fraser C., 2006. Digital camera calibration methods: considerations and comparisons. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, 36(5), pp. 266-272.

Ruiz A., López-de-Teruel P.E., García-Mateos G., 2002. A note on principal point estimability. Proc. $16th$ International Conference on Pattern Recognition, Quebec, vol. 2, pp. 304- 307.

Salvi J., Armangué X., Batlle J., 2002. A comparative review of camera calibrating methods with accuracy evaluation. Pattern Recognition, 35, pp. 1617-1635.

Sturm P.F., Maybank S.J., 1999. On plane-based camera calibration: a general algorithm, singularities, applications. IEEE Conference on Computer Vision & Pattern Recognition (CVPR '99), Fort Collins, USA, pp. 432-37.

Villa-Uriol M.-C., Chaudhary G., Kuester F., Hutchinson T. C., Bagherzadeh N., 2004. Extracting 3D from 2D: selection basis for camera calibration. IASTED Computer Graphics & Imaging, Kauai, USA, pp. 315-321.

Zhang Z., 1999. Flexible camera calibration by viewing a plane from unknown orientations. IEEE International Conference on Computer Vision (ICCV '99), Corfu, pp. 666-673.