Robust Camera Calibration using Inaccurate Targets

Andrea Albarelli
albarelli@unive.it

Emanuele Rodolà
rodola@dsi.unive.it

Andrea Torsello
torsello@unive.it

Dipartimento di Informatica
Università Ca' Foscari di Venezia
Venice, Italy

Abstract

Accurate intrinsic camera calibration is essential to any computer vision task that involves image based measurements. Given its crucial role with respect to precision, a large number of approaches have been proposed over the last decades. Despite this rich literature, steady advancements in imaging hardware regularly push forward the need for even more accurate techniques. Some authors suggest generalizations of the camera model itself, others propose novel designs for calibration targets or different optimization schemas. In this paper we take a completely different route by directly addressing one of the most overlooked problems in practical calibration scenarios. Specifically, we drop the assumption that the target is known with enough precision and we adjust it in an iterative way as part of the whole process. This is in fact the case with the typical target used in most of the calibration literature, which is usually printed on paper and stitched on a flat surface. In the experimental section we show that even with such a cheaply crafted target it is possible to obtain a very accurate camera calibration that outperforms those obtained with well-known standard techniques.

1 Introduction

The growing availability of cheap and accurate digital imaging devices has led to the wide adoption of computer vision based methods in measurement applications. Typical scenarios range from the use of a single camera for quality check on planar profiles to the adoption of complex camera networks and structured light for precise 3D reconstructions. Crucial to these applications is the ability to know in a very accurate fashion the relation between physical points in the observed scene and the associated image points captured by the camera itself. This knowledge is usually attained by defining a parametric function from $\mathbb{R}^3$ to $\mathbb{R}^2$ and by supplying a technique for finding the most suitable set of parameters for a given camera. The function responsible for the transformation between world and image coordinates is usually referred to as the (direct) *camera model*. The procedure used to search for the correct model parameters (with respect to some observations) is the *calibration method*. Obviously, camera models and calibration methods are strongly related: basic models require less parameters that can be estimated with fewer observations and simpler optimization procedures;
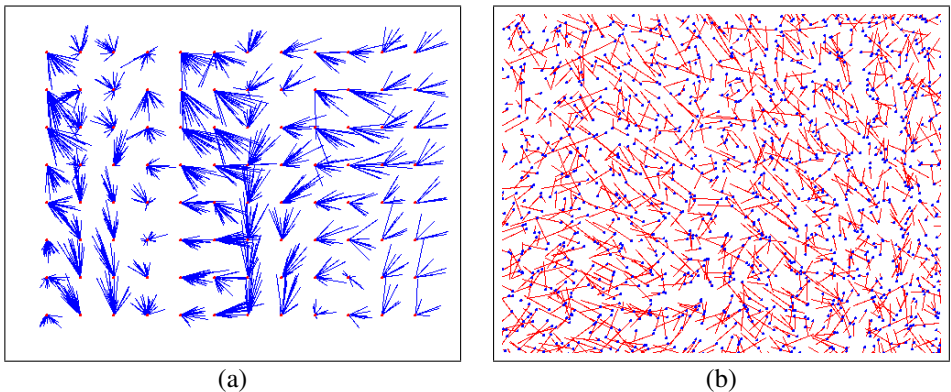
Figure 1: The strongly polarized distribution of the reprojection error with respect to target points obtained in our calibration experiments (a) suggests that the measurements on images were not subject to zero-mean error. On the other hand, the distribution of the error on the image plane (b) is isotropic, assessing the good compensation of radial and tangential distortions and the reliability of the corner detector. This supports our hypothesis that the discrepancies are due to systematic printing errors (errors magnified by a factor of 100).

by contrast, more sophisticated models describe more accurately the optical behaviour of real cameras and lenses, and thus are preferred when higher precision is required.

1.1   Camera model and calibration methods

In this paper we will focus on camera models based on perspective projection as they allow for a complete Euclidean scene reconstruction. The most basic model assumes an ideal pinhole camera characterized by a focal length f (the distance between the projective center and the image plane) and a principal point c (the intersection of the optical axis and the image plane). Often, the focal length is separated into a horizontal (fx) and vertical (fy) component in order to correct any possible non-squareness of the sensor elements. In real scenarios this simple model fails to deal properly with additional factors that contribute to the imaging process, such as distortions introduced by lenses or misalignment of the optical axis. In 1971 Brown [2] proposed a more sophisticated model that accounts for three radial and two decentering distortion coefficients (respectively $k1$, $k2$, $k3$, $t1$ and $t2$). While more complex models exist (accounting for affine and shear transformations of the image plane), the distortion coefficients introduced by Brown have shown to be adequate even for accurate metrology applications. Most calibration methods exploit the correspondences between 3D points of known geometry and 2D points on the image plane. Specifically, they seek a set of model parameters (intrinsic calibration) and camera orientation with respect to the world coordinate system (extrinsic calibration) that minimizes the distance from the measured 2D points and their respective projections obtained by applying the calibrated camera model to the 3D points (reprojection error). If the distortion parameters are not sought, it is possible to solve the calibration problem by using simple linear techniques, such as those suggested by Hall [5] and Faugeras [4]. Differently, it is necessary to introduce more complex approaches that usually alternate a linear technique to optimize a subset of the parameters and an iterative refinement step. One of the first calibration methods accounting for one radial distortion coefficient was proposed by Tsai [14]. Later approaches by Zhang [15] and Heikkilä [6]

are able to deal respectively with two and three radial distortion coefficients; additionally, the latter technique also estimates two coefficients of tangential distortion. The reader can further refer to [11] for a comparative review of the most adopted calibration methods.

1.2 Targets and calibration patterns

All the cited calibration methods rely both on the precise knowledge of the calibration object and on the accurate detection of its feature points on the image plane. The choice of an appropriate calibration object is often a compromise between the calibration accuracy required and the manufacturing complexity of the target itself. In general, 3D calibration objects (i.e., targets that are not composed of coplanar reference features) grant a more stable and precise calibration. This is mainly due to the lack of coupling between intrinsic and extrinsic parameters with respect to the minimization of the reprojection error. In addition, the presence of distinct depth information in the target limits the interplay between the estimation of the focal length and the lens distortion. Unfortunately, accurate 3D targets are very difficult to build, require an expensive tooling and need a remarkable amount of effort for maintenance and verification. While it is possible to use even one-dimensional objects [16], planar calibration targets are by far the most used both in experimental and industrial setups in virtue of their ease of construction. In principle, a single view of a planar target is not enough to determine all the calibration parameters at once [13]. Nevertheless, by exploiting a well planned network of independent views it is still possible to constrain the whole set of parameters and obtain an accurate estimation of the camera model [10].

Regardless of the target geometry chosen, it is equally important to select an appropriate feature pattern for the detection of the reference points. Zhang proposed to use a black square on a white background [15]; while this is an easily detectable pattern, it is biased by systematic localization errors caused by the inherent asimmetry of the marker (for instance, directional bleeding due to illumination). To reduce those biases, a better choice is to adopt symmetric markers such as circular points [6] or checkerboards [8]. Both these marker types have shown to be detectable with high subpixel accuracy, but the checkerboard schema deals better with the misplacement error introduced by radial distortion [9].

2 Camera calibration with inaccurate targets

Throughout the literature the calibration quality obtained by different authors, even when using comparable techniques, is somewhat fluctuating. Given the complex nature of the image acquisition process and the uncertainties of the target manufacturing, it is often difficult to identify the exact sources of error and how they combine. The quality measure itself, which is usually the reprojection error, is not always a correct indicator (in section 3.2 we will show that it is possible to obtain a low reprojection error and still get a bad calibration). In [4] Douxchamps suggests that a higher accuracy could be attained by increasing the total interface length of the markers themselves and thus enhance their localization. A very in-depth theoretical analysis is validated by a large amount of experiments with synthetic images; however, real-world tests exhibit errors about an order of magnitude larger, which the author explains in part with the non-planarity of the target and in part with optical artifacts not corrected by the calibration model. Similar differences between results obtained with synthetic and real images have been previously observed by Heikkilä in [6]. While we agree that this discrepancy can be attributed to imperfection of the target manufacturing, we argue

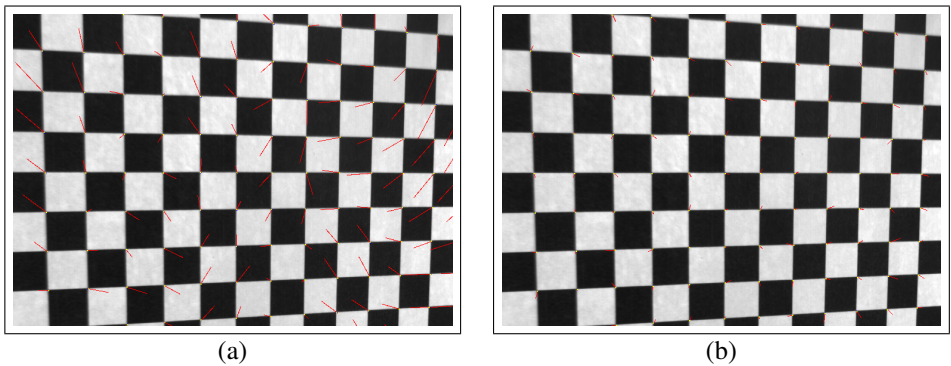(a)   (b)

Figure 2:   Reprojection errors superimposed over the corresponding image point obtained with a direct application of the calibration procedure with the theoretical checkerboard (a) and with the more accurate estimated model (b).

that the non-planarity is not the main error source. Specifically, we built a planar target by stitching a checkerboard printed on inextensible plotter paper onto a 6mm thick highly planar float glass. Using this target we intrinsically calibrated a test camera with the Zhang method and, beside studying the reprojection error with respect to image points, we also plotted the same oriented error bars applied to the theoretical model of the checkerboard. The results of this experiment are shown in Fig. 1. It is immediate to observe that the measurements are not subject to an uniformly distributed error. From a general point of view this means that the optimization was not able to fit the theoretical checkerboard model with a zero-mean Gaussian error, which in turn highlights the presence of some systematic error source. Basically, this source can be correlated with three causes: a localization bias introduced by the subpixel corner detector, the inability of the adopted camera model to fully capture the real image formation model (namely lens distortions), or some unknown discrepancy between the theorical checkerboard model and the printed one. We are not inclined to think that the bias in corner detection is significant since, if this was the case, we should spot a higher coherence in error orientations as all the corner features present roughly the same orientation, scale and illumination conditions (see Fig. 2). To rule out the deficiency of the camera model we plotted the same error measurements on the image plane space (i.e., the camera CCD sensor). By observing part (b) of Fig. 1 it is quite evident that the error distribution is isotropic and this is a strong indication that no systematic error is amendable by using a more sophisticated camera model. Those considerations suggest that the printing process itself could be inaccurate enough to represent a significant source of systematic error in the calibration process. For this reason, we decided to investigate the nature of such printing errors and to tailor a calibration procedure able to correct them. It should be noted that the influence of printing error was already observed in literature. Recently Strobl [12] described a similar scenario, albeit he suggested that the printing procedure introduces just two types of systematic biases: an error in the global scale of the checkerboard and a non predictable aspect ratio. While this simple formulation allows for an elegant calibration procedure that introduces just two additional parameters, we think that it is not general enough. In fact, the distribution of the error shown in Fig. 1 can not be justified by scale and aspect ratio transformations. In order to deal with the most general scenario we propose to lift any direct constraint on the target geometry and iteratively run a three-step procedure (see Fig. 3). In the first step we assume to know precisely the calibration target and we perform a standard

---

Algorithm 2.1: CALIBRATE CAMERA-TARGET($target, images, minError, maxIter$)

$originalTarget \leftarrow target$
$reprojectionError \leftarrow +\infty$
$currentIteration \leftarrow 0$
while $reprojectionError > minError$ **and** $currentIteration < maxIter$

$\textbf{do} \begin{cases} cameraParameters \leftarrow calibrateCamera(target, images) \\ reprojectionError \leftarrow bundleAdjust(target, cameraParameters, images) \\ target \leftarrow hornAdjust(target, originalTarget) \\ currentIteration++ \end{cases}$

return ($cameraParameters$)

Figure 3: The simple, yet effective, algorithm proposed to iteratively estimate the intrinsic camera calibration and the real target geometry.

calibration procedure with the currently trustable target model. Once plausible camera parameters are obtained, we assume them to be correct and we evaluate a more accurate target geometry by using a bundle adjustment technique to estimate only camera poses and scene (i.e. target). At this point a further step is needed, specifically we need to rescale the newly created target to fit the original one. This is necessary since the bundle adjustment step does not guarantee scale invariance. This adjustment step is performed by using the robust closed form point set alignment technique by Horn [7]. The procedure is stopped when the reprojection error falls below a fixed threshold or a maximum number of iterations is reached. Note that since the estimated target is scaled towards the theoretical one at each step, the final camera calibration could be subject to an absolute scaling error that is not avoidable as the real measures of the target are not known. Still, this error is averaged over the printing error of each corner and in practice this has shown to be very small (see Sec. 3).

3 Experimental results

We performed all of our experiments using a pair of Basler Scout scA1300-32gc Gigabit Ethernet monochrome cameras with a resolution of 1280 x 960 pixels. We printed our checkerboard pattern with one laser and two different inkjet printers. Some rudimentary dimensional controls, made with a millimeter accurate ruler, showed that the laser printer is surprisingly inaccurate as it produces prints that suffer from a global error (measured between the first and last check of the A4 page) from one to three millimeters. This error was also very variable between subsequent prints and also the aspect ratio was not repeatable (as previously observed by Strobl [12]). The inkjet printers were more accurate, showing printing errors greatly below the measurement threshold of the ruler (i.e. sub-millimeter) both with respect to the global checkerboard size and its aspect ratio. Thus, we decided to build our test target by printing a large checkerboard on an inextensible A3 plotter sheet with the inkjet printer we felt to be more precise, and to stitch it on a thick planar glass with all the needed care to guarantee the best possible adhesion. The camera calibration step was performed by using
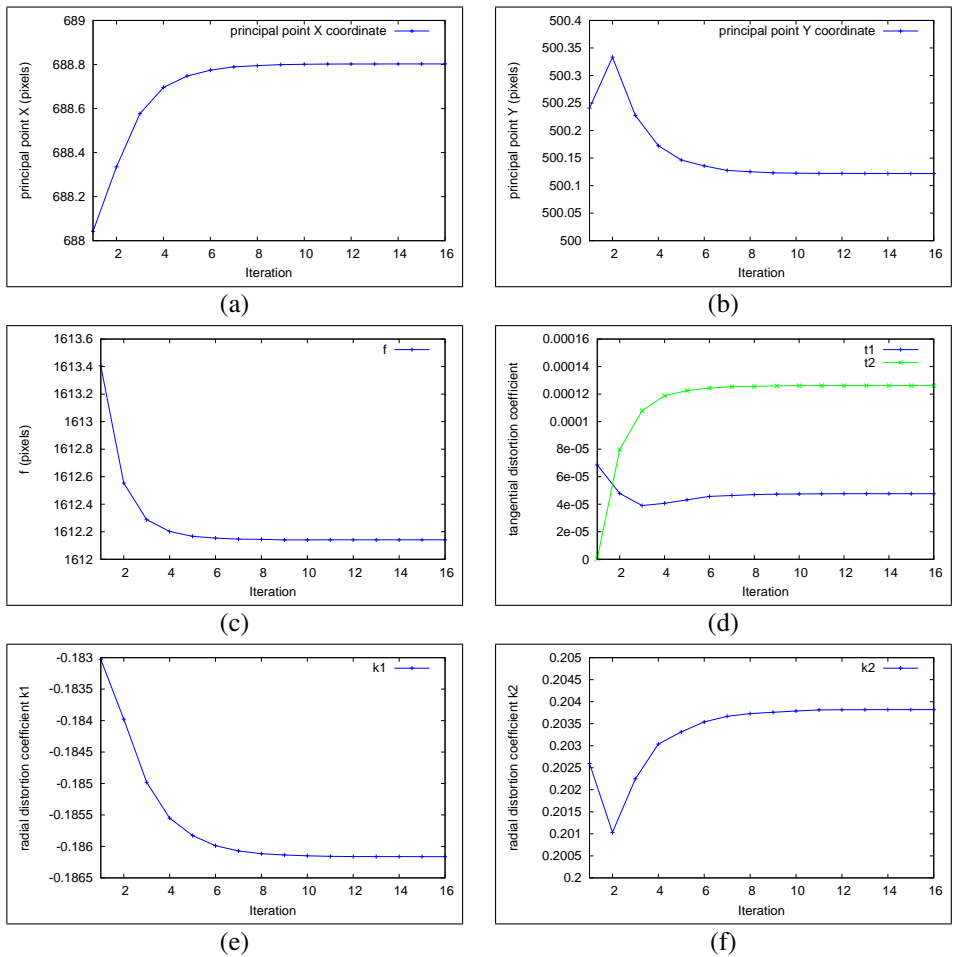
Figure 4: Convergence of both internal and distortion parameters with respect to subsequent iterations of our method.

the OpenCV [1] library, which implements the Zhang [15] method and allows to estimate, by non-linear optimization, three radial and two tangential distortion parameters. In our experiments we estimated only two radial distortion parameters since the used lens set was not wide-angle and a preliminary set of experiments showed that the third parameter was both irrelevant and unstable. Further, we fixed the camera aspect ratio to 1 (i.e. $fx=fy$), as guaranteed by the manufacturer. To perform the bundle adjustment a suitable sparse implementation of the Levenberg-Marquardt optimization algorithm was used. For each camera we took a set of 200 shots of the target: 100 shots were taken with small angles between the target and the optical axis (from 0 to 15 degrees), while the others exhibit a larger range of orientations (up to 30 degrees). These two sets have been used to study the influence of the target orientation in both the quality of the calibration and the reprojection error measure (see Sec. 3.2). Finally, the two cameras were mounted on a rigid support to form a stereo pair which was used to shot another set of 100 images of the same checkerboard, and thus perform a stereo calibration and target reconstruction as an additional validation step.
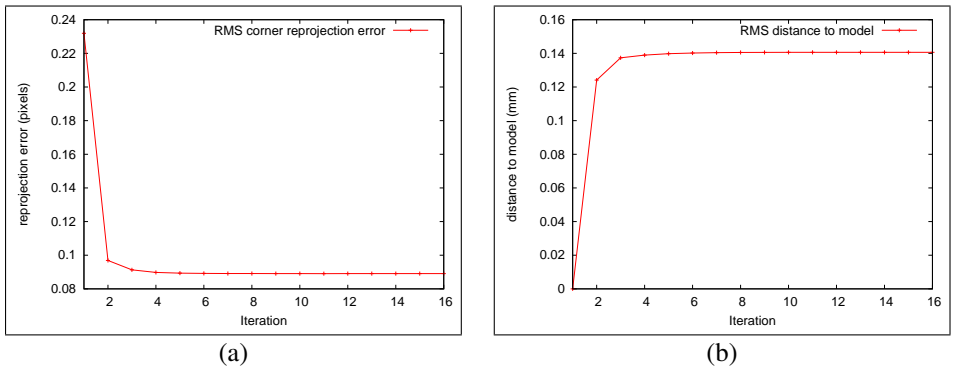
Figure 5: Convergence of both reprojection error (a) and distance of the estimated calibration object with respect to the theoretical checkerboard (b) for subsequent iterations of the proposed method.

3.1 Effects of target estimation on camera calibration

In our first set of experiments we analyzed the convergence behaviour of the proposed algorithm. This was done by applying the calibration technique to the full set of shots taken by the first camera. By observing the plots presented in Fig. 4, we can see that the position of the principal point (plots a and b) changes significantly during the first 8 iterations. In particular, its horizontal location moves by almost 1 pixel. In general, with a good quality calibration setup, this value is expected to be recovered with at most some fraction of pixel of estimation error.

We observe a similar behaviour for the focal length (plot c), which moves away from the initial approximation by more than 1 pixel. In the last three graphs of Fig. 4 we see the distortion parameters variation. Again, the distances from the first rough estimate are similar in percentage to those observed for the internal parameters.

To give an idea of how these variations relate with the reprojection error, in Fig. 5 we reported the RMS between observed and reprojected points for the same iterations shown in part (a) of Fig. 4. It is interesting to note that the reprojection error dramatically decreases after just one iteration, then it becomes quite stable and after as few as 4 iterations it seems to have reached convergence. Indeed, this does not correspond to the behaviour observed for the camera parameters, which continues to move after 4 iterations. We think that this is due both to the misleading nature of the reprojection error (which we discuss in Sec. 3.2) and to the fact that most of such initial large error is to be attributed to the very rough first approximation of the checkerboard model.

To better explain the interplay between the estimation of the "true" calibration target and the associated error reduction, we plotted (in part (b) of Fig. 4) the RMS error between the corresponding points of the theoretical and estimated target. We can see that the target estimation process requires a few more steps to stabilize, albeit the most significant variation happens during the first few bundle adjustments. The final distance from the theoretical model settles around one-tenth of a millimeter, which is compatible with the qualitative observation obtained by using a physical ruler. While this error magnitude can be considered small for a normal printing operation, it can be cumbersome when seeking for an accurate calibration, especially considering that high-quality specifically crafted industrial calibration targets guarantee a location error below one-hundredth or one-thousandth of a millimeter.
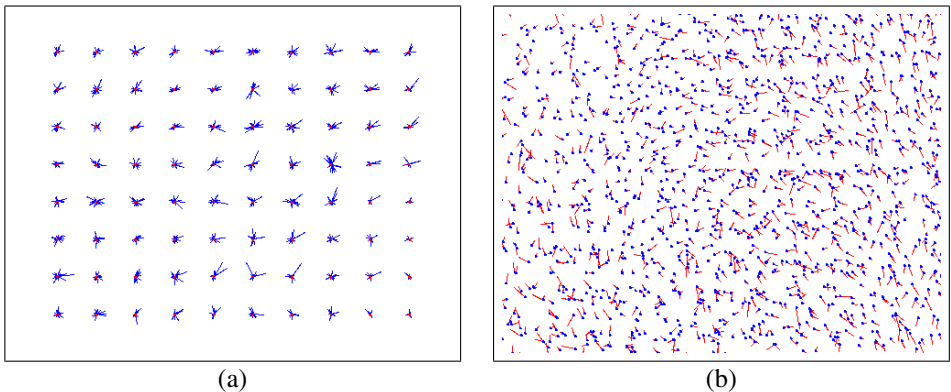
Figure 6:   Reprojection error (magnified by a factor 100) with respect to the model points (a) and the sensor plane (b) after our method comes to convergence.

In Fig. 6 we finally show the distribution of the reprojection error with respect to (a) the target points, and (b) the image plane after the application of the proposed calibration method. This figure is meant to be directly compared with Fig. 1. Regarding the measured error applied to the target points, the large reduction in magnitude is very noticeable. In addition, the error sticks appear to exhibit an isotropic distribution. This suggests that most of the systematic error has been eliminated and only Gaussian noise (due to image noise and corner detection) is measured.

3.2   Validation of the calibration by stereo triangulation

The reprojection error obtained in Sec. 3.1 is below one-tenth of pixel. While this is not a bad result overall, it is a bit large with respect to common values declared in literature by recent methods, which usually reach a reprojection error as low as 0.05 pixels even using self-printed targets. To better understand the relation between the reprojection error and the calibration accuracy we performed and independent check based on stereo reconstruction. We calibrated a pair of cameras with the described technique and then we fixed them to a rigid bar and calibrated the stereo rig using a new set of stereo shots and the target geometry previously learnt. Since the correct geometry was assumed to be known, this latter step was performed simply by using the appropriate stereo calibration function of the OpenCV library. The stereo calibration was then used to recover, in the 3D space and for each stereo shot, the position of two checkerboard corners that are 10 cm apart. Our idea is that by analyzing the measured distance between the reference points under different conditions, we should be able to assess the overall quality of the two camera calibrations and of the target estimation (as the "learnt" target is used to perform the stereo step). Specifically, we tested three calibration setups: the first is meant to be a comparison with standard calibration techniques and thus uses only the theoretical checkerboard, the second uses the estimated target and all the images, finally the last setup uses only a reduced set of images that feature angles with the optical axis lower than 15 degrees. In Fig. 7 we present a scatter plot of the different measures obtained with the three setups described. Since the target is moved among the stereo shots, we reconstructed the pair of control points in different positions in space. Of course, the better the calibration is the more consistent is the measure between different shots, since if the calibration were perfect, the only measure error would be due to the subpixel corner
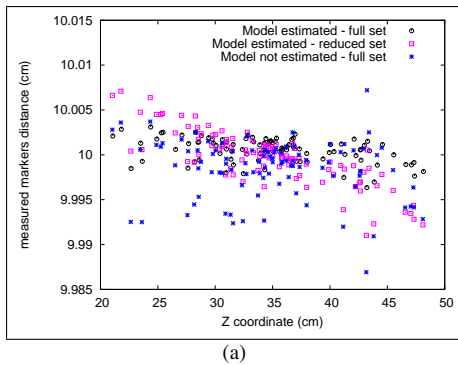
	Reproject Error	Distance to Model	Measure Error
Model full set	0.08907	0.14062	0.00145
Model small set	0.05814	0.11645	0.00317
No model full set	0.23191	0	0.00362

(a)   (b)

Figure 7: Verification of the calibration quality by means of a stereo measuring check (see the text for details).

localization. We can observe that by using our method with all the shots (thus with angles up to 30 degrees) we obtain the most reliable results: the measure has the lowest variance and it is quite uniform throughout the depth of the view field. By contrast, using the reduced set of images (those that are more fronto-parallel) not only leads to a larger variance, but in addition a clear gradient appears in the scatter plot; that is, the measure tends to decrease as the target gets far from the stereo pair. This effect is probably due to the fact that images with weak angulation are not able to fully constrain all the calibration parameters, especially the extrinsic ones (i.e. camera poses). As expected, when not estimating the target model we obtain a larger variance, albeit we do not get any gradient along the z coordinate as we use the full set of images. Note that the average in not exactly 10 cm even for the best technique: this is to be expected, as the real distance is not known and we just know that the two reference points are separated by 10 squares allegedly long 1 cm in the print. In table (b) of Fig. 7 we show the different reprojection errors attained by the three setups. We can see that, notwithstanding the large average measurement error, the setup that uses the fronto-parallel images obtains a very low reprojection error. This apparent nonsense is indeed due to the lower error introduced by the corner detector as, when the model is correctly estimated, this zero-mean Gaussian error does not hinder a correct camera calibration, yet it contributes to the magnitude of the reprojection RMS error. On the other hand, the use of weakly angulated images poses less constraints to parameters and thus leads to both a less accurate mono calibration and stereo reconstruction. In this sense, the reprojection error is not always a good indicator since, when the calibration setup is good, it tends to be dominated by localization errors.

4   Conclusions

In this paper we introduced a very simple but effective method to precisely calibrate a camera with a cheap self-built target. While this approach is not meant to substitute the use of professional-made calibration objects, it is very useful since printed targets are still widely used both by researchers and practitioners. The effects of an accurate learning of the real target geometry are shown with a set of experiments that explore both the convergence behaviour of the algorithm and the quality of the obtained calibration.

References

[1] G. Bradski and A. Kaehler. *Learning OpenCV: Computer Vision with the OpenCV Library*. O'Reilly Media, Inc., 1st edition, October 2008. ISBN 0596516134.

[2] D. C. Brown. Close-range camera calibration. *Photogrammetric Engineering*, 37(8): 855–866, 1971.

[3] D. Douxchamps and K. Chihara. High-accuracy and robust localization of large control markers for geometric camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.*, 31 (2):376–383, 2009.

[4] O. D. Faugeras and G. Toscani. The calibration problem for stereo. In *Proceedings, CVPR '86 (IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Miami Beach, FL, June 22–26, 1986)*, pages 15–20. IEEE, 1986.

[5] E. L. Hall, J. B. K. Tio, C. A. McPherson, and F. A. Sadjadi. Measuring curved surfaces for robot vision. *Computer*, 15(12):42–54, 1982.

[6] J. Heikkilä. Geometric camera calibration using circular control points. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(10):1066–1077, 2000.

[7] B. K. P. Horn. Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America. A*, 4(4):629–642, Apr 1987.

[8] L. Lucchese and S. K. Mitra. Using saddle points for subpixel feature detection in camera calibration targets. In *APCCAS*, pages 191–195. IEEE, 2002.

[9] J. Mallon and P. F. Whelan. Which pattern? biasing aspects of planar calibration patterns and detection methods. *Pattern Recogn. Lett.*, 28(8):921–930, 2007.

[10] F. Remondino and C. Fraser. Digital camera calibration methods: considerations and comparisons. In Isprs, editor, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences,*, volume Vol. XXXVI, Dresden, Germany, 2006.

[11] J. Salvi, X. Armangue, and J. Batlle. A comparative review of camera calibrating methods with accuracy evaluation. *Pattern Recognition*, 35(7):1617–1635, 2002.

[12] K. H. Strobl and G. Hirzinger. More accurate camera and hand-eye calibrations with unknown grid pattern dimensions. In *ICRA*, pages 1398–1405. IEEE, 2008.

[13] P. F. Sturm and S. J. Maybank. On plane-based camera calibration: A general algorithm, singularities, applications. In *CVPR*, pages 1432–1437, 1999.

[14] Y. R. Tsai. An efficient and accurate camera calibration technique for 3D machine vision. In *Proc. CVPR*, 1986.

[15] Z. Zhang. A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(11):1330–1334, 2000.

[16] Z. Zhang. Camera calibration with one-dimensional objects. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(7):892–899, 2004.