Flexible and User-Centric Camera Calibration using Planar Fiducial Markers

Shreyansh Daftry
daftry@icg.tugraz.at

Michael Maurer
maurer@icg.tugraz.at

Andreas Wendel
wendel@icg.tugraz.at

Horst Bischof
bischof@icg.tugraz.at

Institute for Computer Graphics and Vision
Graz University of Technology, Austria

Abstract

The benefit of accurate camera calibration for recovering 3D structure from images is a well-studied topic. Recently 3D vision tools for end-user applications have become popular among large audiences, mostly unskilled in computer vision. This motivates the need for a flexible and user-centric camera calibration method which drastically releases the critical requirements on the calibration target and ensures that low-quality or faulty images provided by end users do not degrade the overall calibration and in effect the resulting 3D model. In this paper we present and advocate an approach to camera calibration using fiducial markers, aiming at the accuracy of target calibration techniques without the requirement for a precise calibration pattern, to ease the calibration effort for the end-user. An extensive set of experiments with real images is presented which demonstrates improvements in the estimation of the parameters of the camera model as well as accuracy in the multi-view stereo reconstruction of large scale scenes. Pixel reprojection errors and ground truth errors obtained by our method are significantly lower compared to popular calibration routines, even though paper-printable and easy-to-use targets are employed.

1 Introduction

Camera calibration is the computation of the parameters of a camera model, a necessary step in extracting 3D metric information from 2D images. Accurate intrinsic camera calibration is critical to most computer vision methods that involve image based measurements, and often a pre-requisite for multi-view stereo (MVS) tasks [27]. Some researchers have advocated the direct use of uncalibrated views for 3D reconstruction [8]. While there has been recent progress in this, we have found that accuracy of structure from motion computation is expected to be higher with a calibrated setup as shown in the work of Irschara *et al.* [16]. In particular, it considerably simplifies the problem because accurate pose is essential to obtain a good reconstruction when using dense depth techniques and it helps avoid the need to differentiate between scenes with and without dominant planar structures [24].
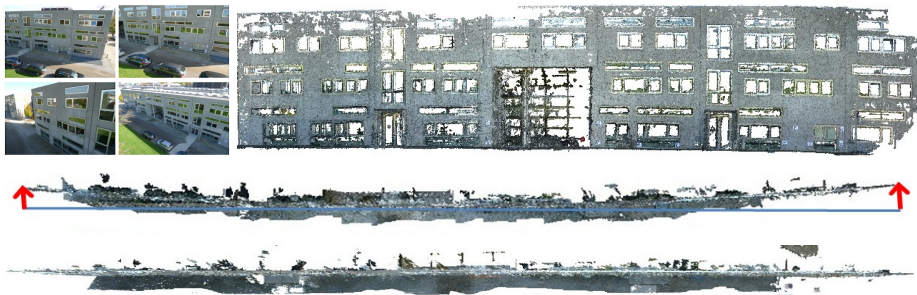
Figure 1: Reconstruction of a facade shows that though visually correct in appearance, geometric inconsistencies (significant bends along the fringes) are prevalent, when using calibration and undistortion results from OpenCV (middle). In comparison, using accurate camera parameters delivered by the proposed method results in a straight wall (bottom).

Recently, 3D vision methods demonstrate increased robustness and result in high quality models. Hence multi-view modeling appears more often in industrial projects targeted at large-scale scenes such as Google Earth and Microsoft Bing. Combined with the emergence and availability of affordable high resolution cameras and desktop computers, this development has made 3D vision methods and tools popular not only among researchers, but also lead to participation of an ever increasing population of end-users, who are not experts in computer vision. Panorama creation tools like Autostitch [5] are already well-established in the public. Now, even end-user applications for automatic structure from motion computation from a large set of images such as Photo Tourism [26] exist. Currently these software tools do not explicitly incorporate calibrated cameras, but rely partially on the focal length specification found in the image meta-data to obtain the initial metric structure. Though the advent of sophisticated multi-view stereo systems [13] used in these tools have enabled the creation of visually appealing models, the need for preserving high standards of geometric fidelity in image based modeling involving participation of end users has put growing demands on the flexibility of existing calibration procedures used to estimate the camera intrinsics.

Given its crucial role with respect to overall precision, camera calibration has been a well-studied topic over the last two decades. However, in most of the calibration literature, a strict requirement on the target geometry and a constraint to acquire the entire calibration pattern has been enforced. This has recently become a matter of question as studied in [1, 20, 28]. The issue of the robust detection of correspondences has also received little attention as most of the standard routines assume that the correspondences have already been found. The poor quality of gathered data, owing to the user's inexperience in calibration procedures (e.g. failure to cover the pattern along the corners of the image) and bad illumination are often sources of inaccuracy. It is these considerations that motivate our contribution.

In this paper, we present and advocate the use of planar paper based fiducial markers as a target pattern to put together a flexible and reliable calibration pipeline, thereby easing the efforts for the end user. Our contribution is three-fold: First, we establish robust correspondences across multiple views with the use of simple printed markers having unique extractable ID's. This allows to robustly select only those inliers which have been precisely detected as input to the remaining calibration pipeline, and to filter out the unreliable data; a flexibility absent in the current methods. Second, we release the critical requirement to process the entire pattern of the calibration object and thereby eliminate the need for an elaborate

procedure for a user to get accurate calibration results. This also ensures that faulty images do not degrade the overall calibration. Finally, we comprehensively validate the recovery of correct calibration parameters using the proposed method with an accuracy far exceeding the standard approach [37] and demonstrate the resulting qualitative and quantitative fidelity in the multi-view stereo reconstruction of large scale scenes. Even though the robustness of typical multi-view reconstruction pipelines results in visually appealing models as can be seen in Fig. 1, the geometric fidelity mainly depends on the accuracy of the camera parameters.

2 Previous Work

Camera calibration is a mathematically well-defined problem and numerous methods have been developed to provide solutions. However, these methods make different simplifying assumptions and hold strict prior requirements as to how they should be used in practice [25]. In the initial days of computer vision, photogrammetric calibration was prevalent. It required observing a calibration object with known 3D scene structure [9]. Tsai [33, 34] introduced two methods where the pose of the target with respect to the camera is being estimated. Even though accurate calibration can be attained using these methods, they require either an expensive calibration apparatus or an elaborate setup. In contrast, some researchers [13, 31] have proposed to determine camera parameters directly from multi-view stereo and bundle adjustment. However, the joint optimization of scene structure and camera intrinsics in this technique involves many degrees of freedom, thus a strongly connected graph of observations and 3D points is necessary. In many outdoor scenes, and especially at the border of a reconstruction, this assumption is not valid and the results tend to be unstable or bent [28].

Planar calibration patterns on the other hand have gained widespread acceptance, due to their ease of use. Significant contributions in this domain were presented by Zhang [37], as well as Sturm and Maybank [29] in the late nineties. Their approach required only a few images of a planar checkerboard pattern and consisted of an initial closed form solution for camera parameters, followed by a non-linear refinement using the Levenberg-Marquardt algorithm [21]. Zhang's method is both flexible and accurate and has thus inspired standard calibration routines as implemented in OpenCV [4] and also serves as a standard benchmark to compare accuracy to nowadays.

A major drawback that still prevails in this method is the critical requirement to provide precise pattern geometry as a prior. Zhang's method, like most other existing approaches, attempt to order corners in a grid to find correspondences and thus fails to address the problem of robustness in correspondences between detected features across views. It is this small but important detail that we review here. We propose to eliminate this requirement on a fixed grid pattern by the use of printed fiducial targets, such that each individual control point can be self-identified with a high degree of confidence. The advantages of eliminating this critical requirement are manifold and will be delineated in the next section.

3 Fiducial Marker-based Calibration

Fiducial (i.e. individually identifiable) markers have become increasingly popular in recent years and are often used for ground truth estimation and precise localization when high accuracy is required. We advocate the use of such fiducial markers in a target pattern for camera calibration. The various advantages of using fiducial markers are discussed below.

Accuracy in Feature Localization and Correspondence. Typical calibration methods consist of two steps: control point localization and then solving for camera parameters. In
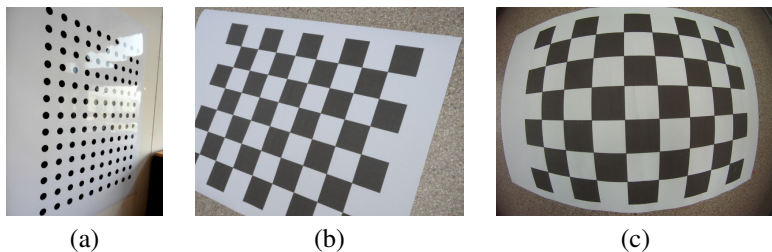
| (a) | (b) | (c) |

Figure 2: Images provided by users which *(a)* have uneven illumination or *(b)* are partially clipped serve as common failure points for calibration methods having a critical requirement on pattern geometry. *(c)* On the other hand, if images are taken correctly the critical corners of the image (highest distortion) are never covered.

an elaborate study by Sun and Cooperstock [30] on factors influencing calibration accuracy, it has been shown that a major source of error is the prior requirement of precise localization of control points and accurate determination of target geometry, especially when it needs to be performed by an unexperienced user. Thus, both control point localization and camera parameter estimation steps should be done as accurately as possible. Square control points are a popular choice for calibrating cameras in standard routines like OpenCV [4] and Matlab Calibration Toolbox [3]. Corners are detected using a Harris interest point detector [14] and ordered into a grid based on their geometric proximity. Precise localization of such control points are however very susceptible to lens distortion effects [7]. Some researchers have proposed circular and ring control points to improve accuracy. Methods that rely on these features [6, 17, 18] cannot handle the non-linear distortion by itself because it is not possible to distinguish if the difference in the feature parameters comes from the distortion or from projective effects. Detection can be unreliable and we found that many images had to be captured to provide enough successfully detected grids for calibration. Albarelli *et al*. [1] suggested a method for target geometry optimization for camera calibration with imperfect targets. Even though their method improves the accuracy of localized points, it still does not release the requirement of using the entire grid or address the problem of robust correspondences. The use of ARTag markers [10] have been proposed as a more robust method to find correspondences [11, 36]. Although detected and identified reliably, they are not ideal for camera calibration because the corner localization is comparatively poor [2]. In contrast, the use of our proposed fiducial markers not only provides robust correspondences across views but also ensures precise localization by sub-pixel accurate line intersection.

Sensitivity to Partial Occlusions.   Camera calibration by end-users is expected to be performed in uncontrolled environments, susceptible to bad or uneven illumination. Hence, it is difficult to anticipate accurate camera calibration results because the improper lighting condition can harm the faithful extraction of the calibration points. Fig. 2(a) shows a failure case where bad illumination can severely whitewash parts of an image and make the subsequent control points occluded. As the OpenCV calibration routine uses a method to automatically find control points and fit them in a grid, it is impossible to determine the uniqueness of each control point in an image, and thus many frames in the set become unusable. In contrast, using a fiducial marker based target allows us to process only reliably located points and hence the users do not have to take care of the visibility of the calibration target. Additionally, to satisfy the requirement of covering the entire target in standard methods (Fig. 2(b)), users often avoid moving the calibration object near the image boundary and as a result poor quality parameters are estimated.
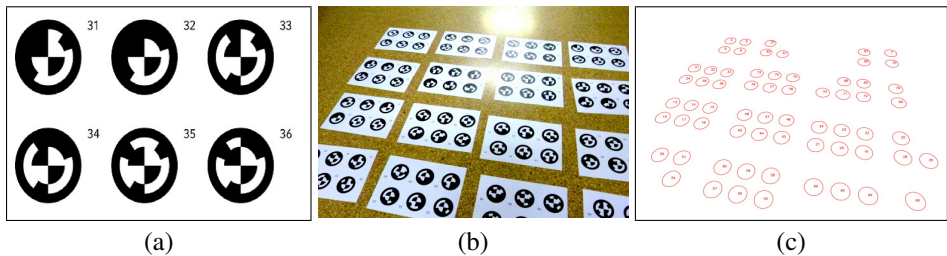
Figure 3: (a) Fiducial markers: Each marker encodes a unique ID. (b) A typical calibration image (partially corrupted due to uneven illumination) arbitrarily arranged in a 4x4 grid on the floor. (c) Reliably extracted markers with positions and ID.

Wide Angle Application. It has been observed that for wide angle lenses radial distortion can be significant, especially along the fringes of the image. Fig. 2(c) depicts a calibration image taken using a wide angle lens camera; the non-linear distortion is significant. Standard methods do not work very well for cameras with highly distorted lenses that account for the non-linear effects caused by radial distortion (i.e. bent epipolar lines instead of straight lines) [12]. The primary reason is that the target pattern is not well distributed along the fringe of the images due to the constraint to process all the points. Also, image-based city modeling generally requires an infinite focus (due to the usually large distances to the pictured object) and a wide angle setup (in order to achieve the necessary image overlap), hence the calibration pattern needs to be sufficiently large. In most cases, it is inconvenient or impossible to place a measurement target like a checkerboard pattern of sufficient size in front of the cameras for an end-user. In contrast, in our proposed method this can easily be done by printing the marker patterns on several sheets of paper and arranging them on the floor (see Fig. 3). These print-outs can be laid out arbitrarily, hence the well-known calibration method of Zhang is not applicable. The flexibility to process only a subset of points allows the control points to be well distributed even along the fringes and alleviates the need for an elaborate process or expensive calibration object for accurate results.

4 A User-Centric Calibration Procedure

Our calibration routine follows the basic principles of planar target based calibration [37] and thus requires simple printed markers to be imaged in several views. The marker patterns are printed on several sheets of paper and laid on the floor in an approximate grid pattern. There is no strict requirement for all markers to be visible in the captured images, but for good calibration results most markers should be well distributed in the calibration image set. The calibration procedure determines the camera intrinsics matrix K and distortion coefficients [15]. We assume a simplified camera model, where the aspect ratio is one (i.e. $\alpha = \alpha_x = \alpha_y$) and the skew factor $s_\theta = 0$. Note that this approximation is valid for almost all standard consumer and industrial digital cameras.

The first step in the calibration procedure is the detection of the fiducial markers in the images and the extraction of the unique marker ID (see Fig. 3). Ellipses are extracted via Canny edge detection, and subsequent grouping comprises the set of putative markers. The local image patch is rectified using ellipse parameters and further checked for the occurrence of a valid binary coded circular pattern. The uniqueness of each extracted marker in every image is validated robustly to eliminate any false positives in case of blurred or otherwise low-quality images, thereby ensuring that each marker is classified with a high degree of confidence. The centers of the extracted ellipses are only approximations of the true marker

centers, hence we use a ten times up-sampled image of the center area to detect Hough lines based on the Canny edge image. The second line is detected using the remaining edge image excluding the inlier set of the first line. Thus, the resulting marker positions in the images are detected with sub-pixel accuracy, even in oblique views. Matching feature points across multiple views is trivial, since unique and easily extractable ID's are available.

Since the marker images are laid out on a planar surface, corresponding feature points are related by a homography. Hence, the first estimation of lens distortion parameters attempts to minimize the reprojection error between extracted feature points with free homography and lens distortion parameters [23]. More formally, if x_i^k denotes the position of marker k in the i-th image, the initial distortion estimation determines

$$arg \min_{H_{ij}, \theta} \sum_{i,j} \left| \hat{D}(x_j^k, \theta) - \hat{D}(H_{ij}x_i^k, \theta) \right|^2, \tag{1}$$

where H_{ij} denotes the image homography from view i to j and $\hat{D}(x, \theta)$ is the inverse distortion function with coefficients θ. We use a radial distortion model denoted as

$$\hat{D}(x, \theta) = \left(x - (\hat{u}_0, \hat{v}_0)^T \right) \cdot \left(1 + k_1 r^2 + k_2 r^4 \right), \tag{2}$$

with $r = \left\| x - (\hat{u}_0, \hat{v}_0)^T \right\|$. The parameter vector $\theta = (\hat{u}_0, \hat{v}_0, k_1, k_2)$ consists of the distortion center $(\hat{u}_0, \hat{v}_0)$ and the coefficients k_1 and k_2. The center of radial distortion $(\hat{u}_0, \hat{v}_0)$ is independent from the optical principal point, thus essentially removing the need for decentering distortion parameters [33]. The initial homographies are set to the gold standard results [15] and the distortion parameters are initialized with the image center, and 0 for the coefficients k_1 and k_2, respectively. The nonlinear minimization is performed with a (sparse) Levenberg-Marquardt method [15].

After determining the initial estimate for the lens distortion, the focal length of the camera is estimated from the set of homographies. Both [31] and [22] employ a non-linear minimization technique for intrinsic parameters estimation and an initial estimate is required. We follow the approach of Irschara *et al.* [16], which utilizes a much simpler search technique to quickly determine the camera intrinsics. It is assumed that the principal point is close to the image center and that the aspect ratio and skew are one and zero, respectively. Hence, we search for a constant, but unknown focal length f determining the calibration matrix K. If the correct intrinsic matrix K is known, the image-based homographies H_{ij} can be updated to homographies between metric image planes, $\hat{H}_{ij} = K^{-1}H_{ij}K$. For a particular view i assumed with canonical pose, $\hat{H}_{ij}$ can be decomposed as $\hat{H}_{ij} = R_{ij} - t_{ij}n_i^T/d_i$, where (R_{ij}, t_{ij}) depicts the relative pose and n_i and d_i denote the plane normal and distance according to the coordinate frame of view, respectively. Note, that each $\hat{H}_{ij}$ provides its own estimate of $n_i = n(\hat{H}_{ij})$. For the true calibration matrix K, the extracted normal n_i should coincide with one common estimate of the plane normal. Hence, a low variance of the set $\{n_i\}$ indicates approximately correct calibration parameters. A slight complication is induced by the fact that decomposing $\hat{H}_{ij}$ results in two possible relative poses and plane parameters (denoted with n_i^+ and n_i^-). Let (n_0^+, n_0^-) be the most separated pair of normals from all pairs (n_i^+, n_i^-). We use n_0^+ and n_0^- as the estimates for the mean of the set $\{n_i\}$. Now, the score for K [16] is

$$\sum_{i,j} \min \left(\angle(n_i^+, n_0^+), \angle(n_i^-, n_0^+), \angle(n_i^+, n_0^-), \angle(n_i^-, n_0^-) \right). \tag{3}$$

where $\angle$ denotes the angle between the normal vectors. This score is evaluated for potential choices of f, *e.g.* $f \in [0.3, 3]$ in terms of normalized pixel coordinates. Fig. 4 depicts the
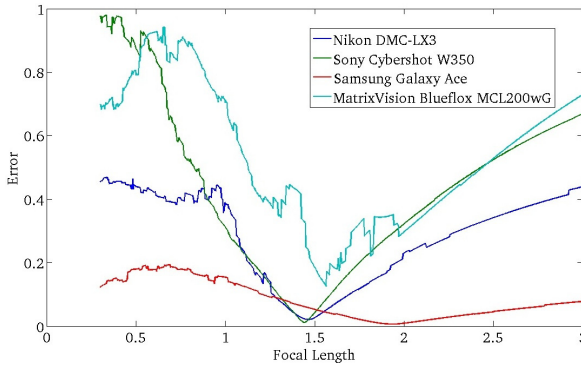
Figure 4: Distortion error function (Eq. 1) with respect to the focal length in normalized pixel coordinates, evaluated for different cameras. Note the distinct global minimum in all four cases. Best viewed in color.

exhaustively evaluated error function over a range of focal lengths for different cameras. The objective function comprises many local minima, but our experiments confirm that a robust global minimum can be found for all cameras using this method. The value of f with the lowest score is used as initial estimate for the focal length. This procedure is both simple and very fast, and yields accurate focal lengths. With the knowledge of the focal length, an initial metric reconstruction based on two appropriate views (large baseline but many correspondences) is generated. The remaining views are added by estimating their absolute poses [19]. The final bundle adjustment [32] procedure is applied to perform a non-linear optimization of the intrinsics (f, u_0 and v_0) and distortion parameters (θ) to optimize the only approximately planar 3D structure and camera poses.

5 Performance Evaluation

In this section we analyze the performance of our proposed method for camera calibration. We use a comprehensive set of off-the-shelf cameras and perform our experiments in a un-controlled environment to simulate conditions similar to that experienced by a common user. The OpenCV Camera Calibration toolbox is used as our primary reference as it represents the most popular method for calibration outside of computer vision and photogrammetry.

Consistency of Calibration Results. To check for the consistency across a varied set of views, we test our calibration routine on a randomly selected subset of 15 out of 20 acquired images, and perform 100 such trials. The evaluation was done ensuring that they have similar coverage and illumination conditions, and that the processing requirements for each method are fulfilled. The results of this experiment are shown in Table 1. Since we do not have ground truth data, the quality of every calibration is evaluated by computing the reprojection error. Mean and standard deviation values for the reprojection error clearly show the significant improvement in calibration accuracy using our method. The RMS error drops from 0.72 pixels to 0.23 pixels for the first camera, a reduction of 67%. More interesting however is the fact that our method not only recovers the calibration parameters accurately but are also highly consistent across different set of views. The σ values for the RMS error are significantly lower than for the standard method, demonstrating the robustness of the system. Note that the use of partial but accurately detected patterns is mainly responsible for the results presented here. Additionally, the variance for the different parameters of the camera model

	Nikon DMC-LX3 (3648x2736)				Samsung Galaxy Ace (1280x960)				MatrixVision Bluefox MCL200wG (752x480)			
	OpenCV		Proposed		OpenCV		Proposed		OpenCV		Proposed	
	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ
f	2638.52	13.67	2613.29	1.12	1238.23	17.72	1244.15	2.12	505.72	31.76	441.69	2.31
u_0	1830.25	6.53	1823.79	0.61	637.21	4.56	645.41	1.64	369.06	5.83	375.67	1.13
v_0	1339.67	10.21	1366.27	1.72	481.25	6.14	474.38	2.05	238.14	4.79	243.81	1.21
k_1	-0.018	0.0091	-0.027	0.0007	-0.0723	0.081	-0.1555	0.007	-0.2664	0.062	-0.2943	0.003
k_2	-0.0002	0.0001	-0.0017	0.0009	-0.0012	0.0005	-0.0424	0.0001	-0.0004	0.0001	-0.0761	0.0001
RMS	0.7251	0.095	0.2371	0.004	0.2987	0.063	0.1792	0.004	0.3213	0.072	0.1934	0.003
Time	189 s		143 s		17 s		38 s		4 s		32 s	

Table 1: Camera calibration results for images taken by different camera types. Compared to the traditional approach of OpenCV, the proposed fiducial marker based approach leads to a significant improvement in calibration accuracy and stability.
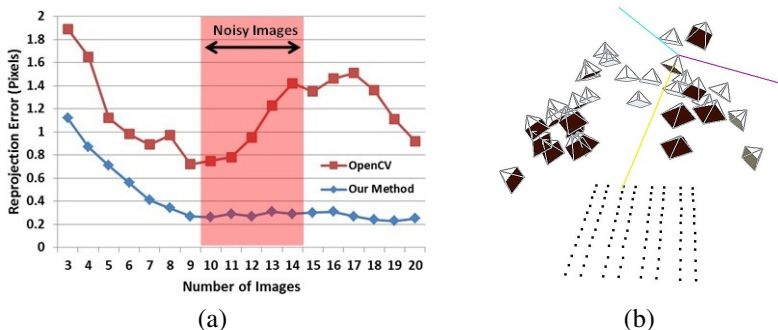

Figure 5: *(a)* Evaluation of the numerical stability of the proposed method w.r.t. the number of images selected. *(b)* Calibration results show the successful recovery of camera poses w.r.t the target pattern and a planar reconstruction of markers as expected.

obtained using OpenCV is very high and thus the final calibration is highly susceptible to the quality of the data. In contrast, a very low variance using our method indicates that the parameters are estimated with high degree of confidence.

Influence of Corrupted Images. A practical question asked by any user wishing to calibrate a camera is how many images are necessary. Even though most researchers have tried to address this in their methods and shown that the accuracy increases for an initial set of 8-10 images and then saturates [30, 31, 38], the fidelity of the complete image set has seldom been considered. We thus investigate the performance of the different calibration methods on a set of images, some of which have been corrupted by adding Gaussian noise with zero mean and $\sigma = 1$ pixel. The results are shown in Fig. 5. Since the entire grid needs to be processed for each view in the case of OpenCV, we observe that it is possible for a set of corrupted images to completely disturb the overall calibration results. The error increases almost linearly with the number of bad images in the calibration set and disturbs the overall results. In contrast our method processes only the reliably extracted information from each view and hence is robust to corrupted images and prevents the error from propagating.

Geometric Fidelity of Reconstruction. We extend our evaluation to qualitatively assess the benefits of increased calibration accuracy when applied to multi-view reconstruction of large-scale scenes. We perform 3D reconstruction of a facade using state of the art methods [16, 35]. Even though the robustness of the multi-view pipeline result in visually appeal-
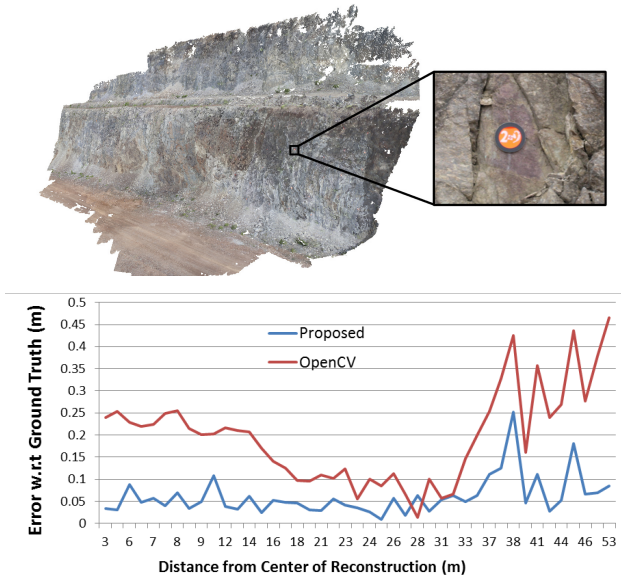
Figure 6: Ground truth evaluation. (Top) The dense multi-view reconstruction of a quarry wall showing an exemplary numbered prism target. (Bottom) The geometric error between the GCPs and the reconstructed position with respect to the distance from the center of reconstruction highlights the benefits of proper undistortion and calibration.

ing models as can be seen in Fig. 1, the geometric fidelity mainly depends on the accuracy of the camera parameters. As a ground truth, we know that the reconstructed wall of the facade should be straight. However on a detailed inspection, we can clearly see that the reconstructed wall suffers from significant bending using results from OpenCV calibration, owing mainly to the distortion parameters. In contrast, the freedom to widely distribute and match control points in our calibration allows an accurate estimation of distortion parameters and consequently reconstruction of a straight wall. Hence we can derive that accurate camera calibration using our method prevents error propagation in reconstruction.

Accuracy Comparison to Ground Truth. In order to give a full quantitative evaluation of the influence of our camera calibration on reconstruction accuracy we compare different calibrations using ground truth 3D data. The dataset used for this experiment has been acquired in a stone quarry using a 10 Mpx consumer camera. The recorded quarry wall with a height of about 24 m and a length of about 100 m contains a set of prism targets, the positions of which have been accurately localized using a total station. The targets are numbered and can be clearly distinguished from the rock by color. We thus able to recognize and manually label the markers in the densified point cloud. Fig. 6 (top) depicts the distribution of the markers on the reconstructed quarry wall. We thus recover the absolute reconstruction accuracy for the Ground Control Points (GCP) in scene coordinates. Fig. 6 (bottom) visualizes the absolute errors between GCPs and reconstructed positions in the model. The visualization clearly shows that using an OpenCV calibration with inaccurate distortion parameters the reconstruction bends, and errors at both small and large distances (MSE 18.71 cm) occur due to the least-square alignment. In fact, the error on the boundaries is thus even larger than it appears here (about 70 cm). In contrast, our method not only results in a lower mean squared error of 5.87 cm which is comparable to the prism target positioning accuracy, but also in a well distributed error along the entire reconstruction.

6 Conclusion

In this paper we have introduced a very flexible and user-centric, but at the same time highly effective method to calibrate a camera with a fiducial marker-based target. While even more accurate calibration is clearly possible in a laboratory, we aim to overcome the difficulties faced by most unexperienced users when using conventional methods that require precision patterns. Our technique helps to establish robust correspondences, releases strict requirements on target geometry, and eliminate the need for an elaborate procedure and controlled environment to obtain accurate camera calibration. We have presented results on real images outlining the consistency in estimation of camera calibration parameters, and the robustness to images corrupted by noise or human error. In addition, we show both qualitative and quantitative benefits of increased camera calibration towards multi-view reconstruction of large scale scenes. Facilitated with an easy to use GUI with user-feedback, the proposed calibration method has been made publicly available[1] so that a wide audience can benefit from our findings.

References

[1] Andrea Albarelli, Emanuele Rodolà, and Andrea Torsello. Robust camera calibration using inaccurate targets. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 31(2):376–383, 2009.

[2] Bradley Atcheson, Felix Heide, and Wolfgang Heidrich. Caltag: High precision fiducial markers for camera calibration. In *Vision, Modeling and Visualization*, 2010.

[3] Jean-Yves Bouguet. Camera calibration toolbox for Matlab, http://www.vision.caltech.edu/bouguetj/calib_doc/. 2004–2013.

[4] Gary Bradski, Adrian Kaehler, and Vadim Pisarevsky. Learning-based computer vision with intel's open source computer vision library. *Intel Technology Journal*, 9(2):119–130, 2005.

[5] Matthew Brown and David G Lowe. Automatic panoramic image stitching using invariant features. *International Journal of Computer Vision (IJCV)*, 74(1):59–73, 2007.

[6] Qian Chen, Haiyuan Wu, and Toshikazu Wada. Camera calibration with two arbitrary coplanar circles. In *European Conference on Computer Vision (ECCV)*, pages 521–532. 2004.

[7] Ankur Datta, Jun-Sik Kim, and Takeo Kanade. Accurate camera calibration using iterative refinement of control points. In *IEEE International Conference on Computer Vision (ICCV) Workshops*, pages 1201–1208, 2009.

[8] Olivier Faugeras, Luc Robert, Stéphane Laveau, Gabriella Csurka, Cyril Zeller, Cyrille Gauclin, and Imad Zoghlami. 3-d reconstruction of urban scenes from image sequences. *Computer vision and image understanding*, 69(3):292–309, 1998.

[9] Olivier D Faugeras and Giorgio Toscani. The calibration problem for stereo. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 86, pages 15–20, 1986.

---

[1]http://aerial.icg.tugraz.at

[10] Mark Fiala. Artag, a fiducial marker system using digital techniques. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 590–596, 2005.

[11] Mark Fiala and Chang Shu. Self-identifying patterns for plane-based camera calibration. *Machine Vision and Applications*, 19(4):209–216, 2008.

[12] Andrew W Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2001.

[13] Yasutaka Furukawa and Jean Ponce. Accurate camera calibration from multi-view stereo and bundle adjustment. *International Journal of Computer Vision (IJCV)*, 84(3): 257–268, 2009.

[14] Chris Harris and Mike Stephens. A combined corner and edge detector. In *Alvey vision conference*, volume 15, page 50, 1988.

[15] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*, volume 2. Cambridge Univ Press, 2000.

[16] Arnold Irschara, Christopher Zach, and Horst Bischof. Towards wiki-based dense city modeling. In *IEEE International Conference on Computer Vision (ICCV)*, 2007.

[17] Guang Jiang and Long Quan. Detection of concentric circles for camera calibration. In *IEEE International Conference on Computer Vision (ICCV)*, pages 333–340, 2005.

[18] Jun-Sik Kim, Pierre Gurdjos, and In-So Kweon. Geometric and algebraic constraints of projected concentric circles and their applications to camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 27(4):637–642, 2005.

[19] Laurent Kneip, Davide Scaramuzza, and Roland Siegwart. A novel parametrization of the perspective-three-point problem for a direct computation of absolute camera position and orientation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011.

[20] Jean-Marc Lavest, Marc Viala, and Michel Dhome. Do we really need an accurate calibration pattern to achieve a reliable camera calibration? In *European Conference on Computer Vision (ECCV)*. 1998.

[21] Manolis Lourakis and Antonis Argyros. The design and implementation of a generic sparse bundle adjustment software package based on the levenberg-marquardt algorithm. Technical report, 340, Institute of Computer Science-FORTH, Heraklion, Crete, Greece, 2004.

[22] Ezio Malis and Roberto Cipolla. Camera self-calibration from unknown planar structures enforcing the multiview constraints between collineations. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 24(9):1268–1272, 2002.

[23] Tomás Pajdla and Tomás Werner. Correcting radial lens distortion without knowledge of 3-d structure. Technical report, Center for Machine Perception, Czech Technical University, 1997.

[24] Marc Pollefeys, Frank Verbiest, and Luc Van Gool. Surviving dominant planes in uncalibrated structure and motion recovery. In *European Conference on Computer Vision (ECCV)*, pages 837–851. 2002.

[25] Joaquim Salvi, Xavier Armangué, and Joan Batlle. A comparative review of camera calibrating methods with accuracy evaluation. *Pattern recognition*, 35(7):1617–1635, 2002.

[26] Noah Snavely, Steven M Seitz, and Richard Szeliski. Photo tourism: exploring photo collections in 3d. In *ACM transactions on graphics (TOG)*, volume 25, pages 835–846, 2006.

[27] Christoph Strecha, Wolfgang von Hansen, Luc Van Gool, Pascal Fua, and Ulrich Thoennessen. On benchmarking camera calibration and multi-view stereo for high resolution imagery. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, 2008.

[28] Klaus H Strobl and Gerd Hirzinger. More accurate pinhole camera calibration with imperfect planar target. In *IEEE International Conference on Computer Vision (ICCV) Workshops*, 2011.

[29] Peter F Sturm and Stephen J Maybank. On plane-based camera calibration: A general algorithm, singularities, applications. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, 1999.

[30] Wei Sun and Jeremy R Cooperstock. An empirical evaluation of factors influencing camera calibration accuracy using three publicly available techniques. *Machine Vision and Applications*, 17(1):51–67, 2006.

[31] Bill Triggs. Autocalibration from planar scenes. In *European Conference on Computer Vision (ECCV)*, pages 89–105. 1998.

[32] Bill Triggs, Philip F McLauchlan, Richard I Hartley, and Andrew W Fitzgibbon. Bundle adjustment - a modern synthesis. In *Vision algorithms: theory and practice*, pages 298–372. 2000.

[33] Roger Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Journal of Robotics and Automation*, 3(4):323–344, 1987.

[34] Roger Y Tsai. An efficient and accurate camera calibration technique for 3d machine vision. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1986.

[35] Andreas Wendel, Arnold Irschara, and Horst Bischof. Natural landmark-based monocular localization for MAVs. In *IEEE International Conference on Robotics and Vision (ICRA)*, 2011.

[36] Hao Yongtao and Xu Jingwei. An improved camera calibration method using the fiducial marker system. In *International Conference on Digital Manufacturing and Automation (ICDMA)*, volume 1, pages 195–201, 2010.

[37] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 22(11):1330–1334, 2000.

[38] Zijian Zhao, Yuncai Liu, and Zhengyou Zhang. Camera calibration with three non-colinear points under special motions. *IEEE Transactions on Image Processing*, 17 (12):2393–2402, 2008.