Camera matrix calibration using circular control points and separate correction of the geometric distortion field

Victoria Rudakova, Pascal Monasse

Camera Matrix Calibration Using Circular Control Points and Separate Correction of the Geometric Distortion Field

Victoria Rudakova, Pascal Monasse
Université Paris-Est, LIGM (UMR CNRS 8049), ENPC,
F-77455 Marne-la-Vallée, France
{*rudakovv, monasse*}*@imagine.enpc.fr*

***Abstract*—We achieve a precise camera calibration using circular control points by, first, separation of the lens distortion parameters from other camera parameters and computation of the distortion field in advance by using a calibration harp. Second, in order to compensate for perspective bias, which is prone to occur when using a circled pattern, we incorporate conic affine transformation into the minimization error when estimating the homography, and leave all the other calibration steps as they are used in the literature. Such an error function allows to compensate for the perspective bias. Combined with precise keypoint detection, the approach is shown to be more stable than current state-of-the-art global calibration method.**

***Keywords*-camera calibration; camera intrinsics; perspective bias; distortion bias; high precision calibration; planar homography; conic-based affine transformation;**

I. Introduction

Camera calibration is the process of finding the true parameters of a camera given an image (or video), and it is a first step towards computational computer vision. Although some information concerning the measurement of scenes can be obtained by uncalibrated cameras [1], calibration is an essential step when metric information is required.

During the past decades a lot of work had been done on the subject of camera calibration for different kinds of applications – starting from photogrammetry [2], [3], and more recently in computer vision, for example, the flexible technique by Zhang [4]. Some free software packages are available as well: the calibration toolbox by Bouguet [5], the calibration software based on bundle adjustment by Pierrot-Deseilligny and Clery [6].

Naturally, many works have focused especially on achieving high calibration accuracy and stability. These studies are mainly based on high precision control points of either 2D or 3D nature, and the accurate detection of their projections. Linear and least-square techniques for calibration are built upon Tsai [7] and Weng et al. [8], who improve the calibration accuracy by thoroughly modelling lens distortion and further optimizing with other parameters together.

For a real camera, an image of the calibration pattern is subjected to two types of transformations: a projective transformation as a result of relative 3D position, and a nonlinear transformation due to various lens distortions. The robustness of the control point detection under these two transformations is based on the combination of the pattern employed and on the detection method used. Therefore, there are two possible sources of bias in control point recovery which are named according to Mallon and Whealan [9]: *perspective bias* and *distortion bias*. The main goal would be to obtain bias free data, as this is clearly necessary for obtaining unbiased estimates for calibration algorithms.

The most famous planar calibration techniques were presented by Sturm and Maybank [10] and Zhang [4]. They only require the camera to observe the pattern shown at a few (at least three) different orientations to get a unique solution up to scale factor. Either the pattern or camera are moved and the motion needs not be known. While trying to eliminate distortion bias during optimization, the mentioned works assume that the detected image points have zero mean Gaussian error in order to correctly converge to the optimal solution. However, it is not always the case. The bias does not have the same magnitude for all types of patterns: for example, Mallon and Whealan [9] show that detected control points obtained by using centroid recovery principle can potentially be corrupted by both perspective bias and distortion bias, with the likelihood of larger distortion bias magnitude in a typical camera. They also show that the compensation of distortion bias from such circular pattern points is not possible without knowing the distortion field. At the same time, the most precise keypoints can be obtained by using circular points [11], therefore we will proceed with them by aiming at eliminating distortion and perspective bias.

Regarding the distortion bias, most common approaches mix distortion parameters with other camera parameters and optimize them simultaneously. This could potentially lead to *residual error compensation* [12], where some parameters can be adjusted to compensate the errors of some other parameters, thus decreasing the calibration stability. This error compensation cannot be eliminated in the framework of global methods, and, therefore, must be achieved separately, as a preliminary step to any calibration. Additionally, lens distortion introduces a nonlinear shape warping to the area of the conic and so it is no longer a true conic. The recently proposed *calibration harp* by Tang [12] (see Figure 1) allows

Figure 1. Calibration harp of Tang [12], used to compute the distortion field.

to solve the distortion bias problem by separate estimation of the lens distortion field. Its main idea is based on the straightness measure of stretched strings, pictures of which are taken in different orientations. Thanks to the calibration harp, we can eliminate distortion bias from calibration calculation before the main optimization.

When considering the perspective bias, the calibration accuracy can be improved by its iterative correction, see e.g. Heikkilä [13] and Kannala and Brandt [14], which describe calibration techniques using circular points. Or, it can be avoided by using the projection of the conic contour, rather than its center [15]. As an example, variation of Zhang's method is represented by Yang et al. [16] where instead of point features three or more conics are used to obtain the homography matrix and then deduce calibration parameters.

In our work the main idea for the elimination the perspective bias is based on the fact that the projection of the center of a circle does not correspond to the center of the resulting ellipse in the image. We compensate the perspective bias by taking into account circle-ellipse affine transformation and correspondence of detected keypoints with pattern keypoints. This is done by incorporating the conic affine transformation into the minimization step of the homography computation.

Section II provides a brief overview of the basic equations of the main calibration steps which are mainly based on the method of Zhang [4], including the homography estimation. The modelling of homography error function is given in detail in Section III. Section IV explains the workflow of our method, then Section V demonstrates synthetic and real experiments; finally, Section VI concludes.

II. Camera calibration basic equations

We denote a $2D$ point as $m = [u, v]^T$ and a $3D$ point as $M = [X, Y, Z]^T$. We use $\tilde{x}$ to indicate the homogeneous vector by adding 1 as the last element, i.e., $\tilde{m} = [u, v, 1]^T$ and $\tilde{M} = [X, Y, Z, 1]^T$. The camera model is the usual ideal pinhole and the relationship between a $3D$ point M and its image projection m is given by

$$s\tilde{m} = K \begin{bmatrix} R & t \end{bmatrix} \tilde{M}, \tag{1}$$

where s is an arbitrary scale factor (depending on point), (R, t) are the extrinsic parameters (rotation and translation), which relate the world coordinate system to camera coordinate system, and K is the intrinsic camera matrix given by

$$K = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix},$$

with (u_0, v_0) being the coordinates of the principal point, α and β the focal lengths expressed in pixels with respect to u and v pixel dimensions, and γ the parameter describing the skew of the image axes.

A. Homography between model plane and its image

The i^{th} column of the rotation matrix R is denoted by r_i. Assuming the world coordinate system is chosen so that the model plane is located at $Z = 0$, we can write (1) as

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}.$$

Since Z is always equal to zero, the $3D$ point M will be denoted as $M = [X, Y]^T$ and its corresponding $\tilde{M} = [X, Y, 1]$ (same holds for m and $\tilde{m}$). The relation between model point M and image point m is described by homography H:

$$s\tilde{m} = H\tilde{M}, \tag{2}$$

or, written in expanded form,

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = H \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}, \tag{3}$$

with homography H, defined up to scale factor λ,

$$\lambda H = K \begin{bmatrix} r_1 & r_2 & t \end{bmatrix}. \tag{4}$$

Using the orthomality of r_1 and r_2, we get the two equations

$$h_1^T K^{-T} K^{-1} h_1 = h_2^T K^{-T} K^{-1} h_2 \tag{5}$$

$$h_2^T K^{-T} K^{-1} h_1 = 0, \tag{6}$$

with h_1 and h_2 the first two columns of H. H being estimated independently, this gives two homogeneous linear equations on the coefficients of the symmetric matrix $K^{-T}K^{-1}$. From three views of the planar pattern, we can get six equations, allowing to determing the six parameters of $K^{-T}K^{-1}$. A Cholesly decomposition of the solution yields K^{-1} and finally K by inversion.

B. Estimation of the homography between the model plane and its image

The Direct Linear Transform (DLT) algorithm is a simple algorithm used to solve for the homography matrix H given sufficient number of point correspondences, namely 4 (the steps are explained in chapter 4.1 of [17]).

In many cases more than 4 correspondences are used to ensure a more robust solution. A single and exact solution is possible only if all of the correspondences are exact, however, in practice there will always be some noise, so there will be no exact solution. The estimate of the parameters is based on maximum likelyhood criterion. If M_i and m_i are the model and image points respectively, then they should satisfy (2), but they do not for real data because of the noise in extracted image points. If we assume m_i is corrupted by Gaussian noise with mean 0 and covariance matrix Λ_{m_i}, then the maximum likelihood estimation of H minimizes the functional

$$\sum_i (m_i - \hat{m}_i)^T \Lambda_{m_i}^{-1} (m_i - \hat{m}_i),$$

and $\hat{m}_i$ is

$$\hat{m}_i = \frac{1}{\bar{h_3}^T M_i} \begin{bmatrix} \bar{h_1}^T M_i \\ \bar{h_2}^T M_i \end{bmatrix} \tag{7}$$

with $\bar{h_i}^T$ is the i^{th} row of H. Later we will see that (7) cannot be applied for circled pattern due to perspective bias, and it will be replaced by applying the homography to a conic, rather than to its center.

In practice it is assumed $\Lambda_{m_i} = \sigma^2 I$ for all i. This is reasonable if points are extracted independently with the same procedure. In this case the problem becomes a non linear least square problem:

$$H = \arg\min_H \sum_i \|m_i - \hat{m}_i\|^2 . \tag{8}$$

The nonlinear minimization is conducted by using the Levenberg-Marquardt algorithm [18]. It requires an initialization, which is estimated by DLT algorithm as mentioned previously.

III. Incorporation of conic-based transform into homography estimation as perspective bias compensation

When we take a photo of a planar circle in $3D$, it becomes an ellipse in the projection plane. Both circle and ellipse shapes are instances of conics. In $2D$ projective geometry all non-degenerate conics are equivalent under projective transformations. The equation of conics in Cartesian coordinates is

$$s_1 x^2 + s_2 xy + s_3 y^2 + s_4 x + s_5 y + s_6 = 0,$$

that is, a polynomial of degree 2. When using homogeneous coordinates and replacing $x \to \frac{x_1}{x_3}$, $y \to \frac{x_2}{x_3}$, we obtain the quadratic form

$$s_1 x_1^2 + s_2 x_1 x_2 + s_3 x_2^2 + s_4 x_1 x_3 + s_5 x_2 x_3 + s_6 x_3^2 = 0.$$

Using the notations for homogeneous coordinates $\tilde{M} = [X, Y, 1]^T$ and setting $X = x_1$ and $Y = x_2$, the matrix form of conics is

$$\tilde{M}^T S \tilde{M} = 0, \tag{9}$$

where the conic coefficient matrix S is given by

$$S = \begin{bmatrix} s_1 & s_2/2 & s_4/2 \\ s_2/2 & s_3 & s_5/2 \\ s_4/2 & s_5/2 & s_6 \end{bmatrix}.$$

The conic coefficient matrix is always symmetric and matrix S is its homogeneous representation.

Transformation of conics: Under the point transformation (2), (9) becomes

$$\tilde{m}^T H^{-T} S H^{-1} \tilde{m} = 0$$

so the transformed conic or image of conic S is the conic

$$H(S) := H^{-T} S H^{-1}. \tag{10}$$

A. Center of conic's image vs. image of conic's center

Given the equation (9) of a conic in homogeneous coordinates with S a 3×3 symmetric matrix, its center of symmetry is obtained by

$$C(S) := -S_{2\times 2}^{-1} S_3,$$

with $S_{2\times 2}$ the top left 2×2 block of S and S_3 the 2×1 top part of its last column:

$$S = \begin{pmatrix} S_{2\times 2} & S_3 \\ S_3^T & \lambda \end{pmatrix}.$$

For a conic S we can compute its image $H(S)$ as in (10) and now we wish to compare image of conic center $HC(S)$ and center of image conic $CH(S)$, as shown in Figure 2. We will see it will not be the same point and the difference will be quantified.

1) Affine transformation: Consider a specific case when we deal with a large focal length, and so H is close to an affine transformation ($h_{31} = h_{32} = 0$ and $h_{33} = 1$); we will have $HC(S)$ coincide with $CH(S)$. Indeed, we may write

$$H = \begin{pmatrix} H_{2\times 2} & H_3 \\ 0_2^T & 1 \end{pmatrix} \text{ and } H^{-1} = \begin{pmatrix} H_{2\times 2}^{-1} & -H_{2\times 2}^{-1} H_3 \\ 0_2^T & 1 \end{pmatrix}.$$

We then get

$$H(S) = \begin{pmatrix} H_{2\times 2}^{-T} S_{2\times 2} H_{2\times 2}^{-1} & -H_{2\times 2}^{-T} S_{2\times 2} H_{2\times 2}^{-1} H_3 + H_{2\times 2}^{-T} S_3 \\ \cdots & \cdots \end{pmatrix}$$

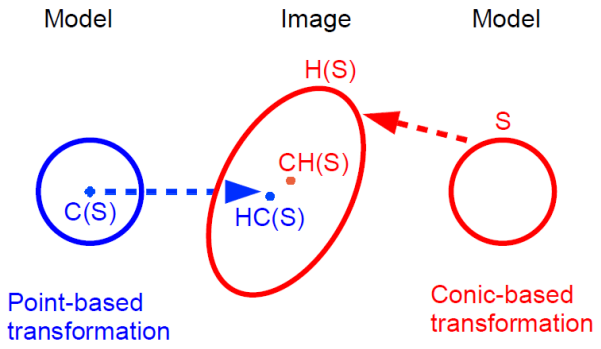

Figure 2. Difference between image of conic center and center of conic image.

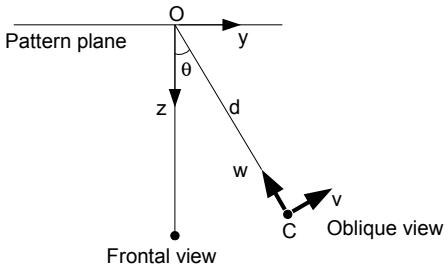

Figure 3. View of pattern plane along an angle θ.

and therefore

$$\begin{aligned} &CH(S) \\ &= -(H_{2\times2}^{-T}S_{2\times2}H_{2\times2}^{-1})^{-1}(-H_{2\times2}^{-T}S_{2\times2}H_{2\times2}^{-1}H_3 + H_{2\times2}^{-T}S_3) \\ &= H_3 - H_{2\times2}S_{2\times2}^{-1}S_3, \end{aligned}$$

which equals $HC(S)$.

At first order, the homography can be approximated by an affine transform, so the difference we get is second order, which still needs to be quantified.

2) Change of viewpoint transformation: Suppose the calibration pattern is planar and the camera points at it from an angle $\theta \in [0, \pi/2)$ as on Fig. 3. Take as origin of the world coordinate system the intersection O of the principal ray of the camera with the pattern plane. Suppose also the camera has square pixels. Take the v-axis of the image in the plane determined by the principal ray and the normal vector to the pattern plane. We also take as origin in the image the principal point. Let us write d the distance of the optical center to the pattern plane along the principal ray.

In the front view at the same distance, we can write the projection matrix

$$P_0 = \begin{pmatrix} f & & \\ & f & \\ & & 1 \end{pmatrix} \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & d \end{pmatrix}$$

and in the oblique view

$$P = \begin{pmatrix} f & & \\ & f & \\ & & 1 \end{pmatrix} \begin{pmatrix} 1 & & & \\ & \cos\theta & -\sin\theta & \\ & \sin\theta & \cos\theta & d \end{pmatrix}.$$

Since the plane has equation $z = 0$, the homography between front view and oblique view is

$$\begin{aligned} H &= K \begin{pmatrix} 1 & & \\ & \cos\theta & 0 \\ & \sin\theta & d \end{pmatrix} \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1/d \end{pmatrix} K^{-1} \\ &\sim \begin{pmatrix} 1 & & \\ & \cos\theta & \\ & \frac{\sin\theta}{f} & 1 \end{pmatrix}. \end{aligned} \tag{11}$$

Given a circle with center $M = (X\,Y)^T$ and radius r in the front view, its matrix representation is

$$S = \begin{pmatrix} 1 & 0 & -X \\ 0 & 1 & -Y \\ -X & -Y & \|M\|^2 - r^2 \end{pmatrix}.$$

After computations, we get

$$\begin{aligned} CH(S) &= \frac{1}{(1 + Y\frac{\sin\theta}{f})^2 - r^2\frac{\sin^2\theta}{f^2}} \\ &\times \begin{pmatrix} X(1 + Y\frac{\sin\theta}{f}) \\ Y\cos\theta(1 + Y\frac{\sin\theta}{f}) - r^2\frac{\cos\theta\sin\theta}{f} \end{pmatrix}, \end{aligned}$$

whereas

$$HC(S) = \frac{1}{1 + Y\frac{\sin\theta}{f}} \begin{pmatrix} X \\ Y\cos\theta \end{pmatrix}.$$

Notice that the term $1 + Y\frac{\sin\theta}{f}$ vanishes for points in front view that map to the line at infinity of the plane (horizon).

Numerical results: for small r, the distance is proportional to r^2. If $r = 0$ (point), we get coinciding points, as for $\theta = 0$. Distance in pixels between $HC(S)$ and $CH(S)$ for different values of θ are displayed on Fig. 4. Note that even though the distance remains moderates, it increases significantly with the view angle.

B. Recovering homography by conic transform cost function

Considering (7), we can re-write it in the context of conic transformation, therefore we get

$$\begin{aligned} [\hat{u}, \hat{v}, \hat{w}]^T &= H\tilde{M}_i = HC(S_i) \\ \hat{m}_i &= \left[\tfrac{\hat{u}}{\hat{w}} \quad \tfrac{\hat{v}}{\hat{w}}\right]^T \end{aligned}$$

for the pattern point M_i which is a center of circle S_i. The detected 2D keypoint m_i which corresponds to the center of projected circle is

$$m_i = CH_0(S_i),$$

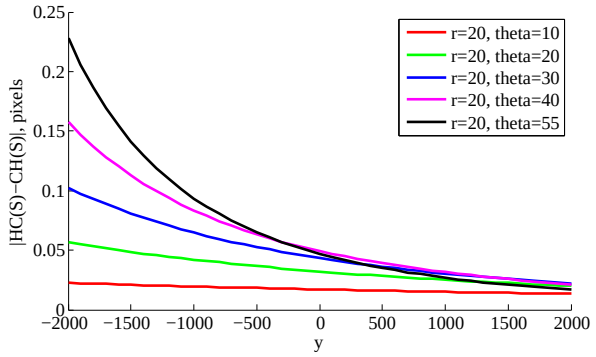

Figure 4. Difference in pixels between $HC(S)$ and $CH(S)$ for $f = 4000$, $x = 0$, $r = 20$ as a function of y for different angles θ (in degrees).

where H_0 is the ground truth homography. Therefore, (8) becomes

$$H_{\texttt{point}} = \arg\min_H \sum_i \|CH_0(S_i) - HC(S_i)\|^2 . \quad (12)$$

In order to compensate for the perspective bias, we have to take into account conic transform, that is to minimize with respect to the center of projected circle, not to projection of the circle center:

$$H_{\texttt{conic}} = \arg\min_H \sum_i \|CH_0(S_i) - CH(S_i)\|^2 . \quad (13)$$

As it can be seen, $H_{\texttt{conic}}$ does not require conic contour extraction, but only the operator which allows to extract conic center. In case of synthetic data, the performance can be evaluated by error function

$$E = \|H_0 C(S_i) - HC(S_i)\|^2 . \quad (14)$$

The Levenberg-Marquardt algorithm is used to calculate the homography matrices. For simplicity of code, our implementation calculates each new estimate of the Jacobian using finite-difference. The error function for conic based transform method is

$$e_i = \sqrt{\|m_i - CH(S_i)\|^2}.$$

IV. Calibration workflow

Since the distortion correction is performed separately from calculation of other camera parameters, the sequence of main steps is as follows:

I Lens distortion correction
   a. Take pictures of the calibration harp under different orientations
   b. Take pictures of the calibration pattern
   c. Obtain the lens correction model based on pictures of calibration harp (use method of Tang [12])
   d. Apply lens correction model to pictures of calibration pattern

II Intrinsics calculation
   a. Use as input the distortion-compensated pictures of calibration pattern
   b. Obtain the homographies using conic transformation function
   c. Extract calibration matrix

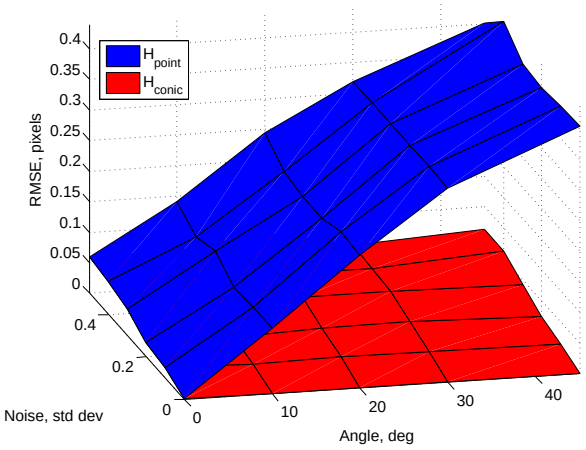

Figure 5. Comparison of two methods for homography estimation for different levels of noise and view angles. Blue color stands for $H_{\texttt{point}}$, the red color graph is for $H_{\texttt{conic}}$. Note the increasing difference of final RMSE due to perspective bias for point method when the angle increases.

V. Experiments

A. Homography estimation precision

We aim to compare the performance of $H_{\texttt{point}}$ (12) with $H_{\texttt{conic}}$ (13) against factors such as noise and view angle. To generate synthetic data, we use ground truth homography obtained by (11) with $f = 4000$. The model plane is represented by circle pattern and it consists of 10×14 circles, each of radius 20 pixels, and that gives us 140 keypoints in total. The pattern is always rotated over axis y on a given angle view $[0, \cdots, 45]^\circ$. No physical image is generated, we only deal with keypoint coordinates which are obtained based on camera settings, for example, for circle S_i of the pattern, its projected image is $H(S_i)$, therefore, the extracted keypoint has coordinates $CH(S_i)$. Gaussian noise with 0 mean and standard deviation $[0, 0.5]$ pixels is added to the projected image keypoints $CH(S_i)$. For each noise level, 25 independent trials are performed and the results are displayed as an average.

Fig. 5 shows a 3D comparison graph, where the error measure is a root mean square of distance described by (14). From now on, we denote conic based minimization as 'conic method' and standard minimization as 'point method'. As it can be seen from the graph, the conic method is invariant with respect to the change of view angle, which indicates that it does not suffer from perspective bias; on the contrary, the point method is prone to perpective bias.

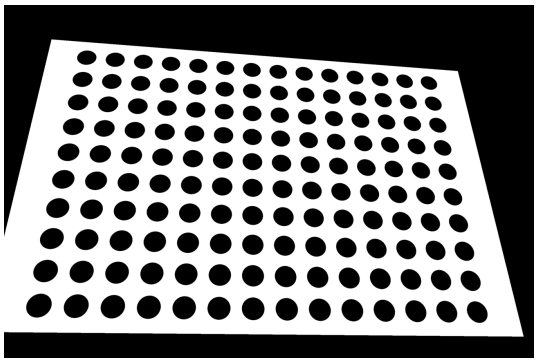

Figure 6. One view of the calibration pattern

B. Calibration matrix stability for synthetic data

Pattern and camera views synthesis: The image resolution is set to 1296×864 pixels. The number of circles is 10×14, each of radius 1cm, and consecutive circles have 3cm separation between each other. The pattern, see Figure 6, has resolution $42\text{cm} \times 30\text{cm}$. The synthetic camera has following parameters: $\alpha = 1250$, $\beta = 1250$, $\gamma = 1.1$, $u_0 = 648$, $v_0 = 432$. For the high quality images, we first generate high resolution pattern image and then subject it to the geometric displacement (all distortion is eliminated), Gaussian blurring and then down-sampling to the image resolution. Geometric image re-sampling is carried out by mapping from the transformed image to the original pattern. This involves calculating for every pixel in the transformed image, the corresponding pixel coordinate in the original image, which requires an inverse mapping. The transformed image intensity is then calculated based on the standard linear interpolation around the corresponding coordinate of the original pattern.

Pattern positioning: In order to test calibration matrix stability, we generated 5 sets of images, each set included 5 images (different views of the pattern). This allows to extract 5 calibration matrices so as to see stability of its parameters along the sets. The generated image views are simulated by using pseudo-randomly generated homographies which consist of $3D$ rotation and translation whose values are drawn randomly from a specific range. This range limit ensures that the transformed image lies roughly within an image window. Meanwhile, there is always variance of rotations and translation along the sets. The rotation angles always lie within the range $[15°, 45°]$. The keypoint detection is held the same way as in [19] which allows to achieve the detection precision of less than 0.05 pixels.

K matrix stability: To compare our method with a state-of-the-art method, we chose Software Package for Precise Camera Calibration [20] which is built upon Matlab Calibration Toolbox [5] with the difference that circle pattern can be used for the calibration, plus some improvements of the original software. The generated images are treated as distortion free for both software. Table I provides the standard deviation results for both methods (we denote state-of-art method as 'SoA'). As we had set both noise and distortion to zero, we can clearly see the improvement in calibration stability based on having more precise homography matrices for our method.

Table I
STANDARD DEVIATIONS OF CAMERA CALIBRATION PARAMETERS FOR FIVE IMAGE SETS – COMPARISON OF THE STATE-OF-ART METHOD WITH OURS. ANY NOISE AND DISTORTION ARE ELIMINATED. THANKS TO IMPROVED PRECISION OF HOMOGRAPHY CALCULATION, WE NOTICE LESS DEVIATION FOR OUR METHOD.

Parameter	SoA	ours
$\mathtt{dev}(\alpha)$	0.10	0.008
$\mathtt{dev}(\beta)$	0.11	0.008
$\mathtt{dev}(u_0)$	0.02	0.006
$\mathtt{dev}(v_0)$	0.14	0.014

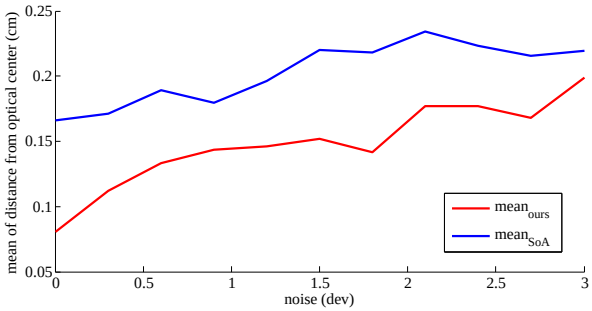

Figure 7. Euclidean distance from the ground truth optical center, in cm, to the obtained center for our (red) and state-of-art (blue) methods: experiments are displayed against different noise level introduced in the synthetic images.

Deviation from camera center: Calculated homography and camera matrix allow to obtain 3D coordinates of the camera center. If the ground truth is known, then it is possible to get an error measure in centimeters (the pattern unit for our synthetic tests). For this, we need to extract rotation and translation parameters using homography matrix (steps are described by Zhang [4]). Given ground truth rotation R_0 and translation T_0, we deduce ground truth camera center C_0 by inverting them as

$$C_0 = -R_0^{-1} T_0. \tag{15}$$

For each method, we extract the R and T parameters from the homography, and then we use eq (15) to calculate camera centers for the two methods, so as to compare the magnitude of the deviation from the ground truth. The average of Euclidean distances from C_0 to the obtained centers for each camera view is displayed on Fig. 7. As expected, the graphs show that our method has lower magnitude residual than for the state-of-art global calibration method.

Table II
STANDARD DEVIATIONS OF THE CAMERA CANON EOS5D PARAMETERS FOR REAL DATA – COMPARISON BETWEEEN A STATE-OF-THE-ART METHOD WITH OURS.

Parameter	f18mm		f27mm		f55mm	
	SoA	ours	SoA	ours	SoA	ours
$\mathtt{dev}(\alpha)$	0.583	0.245	2.394	1.727	8.810	2.668
$\mathtt{dev}(\beta)$	0.547	0.175	2.360	1.721	7.644	2.641
$\mathtt{dev}(u_0)$	0.494	0.135	1.732	0.840	3.565	1.297
$\mathtt{dev}(v_0)$	0.787	0.332	1.273	0.935	2.845	1.314

C. Calibration matrix stability for real data

The experiments for real data were performed for the Camera Canon EOS40D with lens Canon EF 18-55mm. Several tests were performed for different focal lengths. For comparison we use same state-of-the-art software as in synthetic tests. In order to treat the distortion for our method, it was necessary to use two different patterns – calibration harp and circled pattern under the same camera settings. Distortion-compensated images for our method were obtained based on the method of [12]. The input images for state-of-the-art method remained the same as input images for our method for the extraction of calibration matrix, but not distortion-compensated. For each focal lengths we took 6 datasets, each contained 5 images of circled pattern under different camera orientation. K matrix was extracted for each set and then standard deviation of each parameter was taken so as to measure the result stability. The comparison of both methods is shown in Table II. As it can be observed, our method was able to achieve more stable results since deviation is smaller than for the state-of-the-art method. More precise results are achieved by elimination of perspective bias when calculating the homography and distortion bias before the main optimization. They also validate an assumption that separation of distortion from other camera parameters helps avoiding residual error compensation by leading to more precise camera calibration.

VI. CONCLUSIONS AND FUTURE WORK

Two main contributions were made in order to improve the precision of camera calibration. First, we evaluate successively the lens distortion and the other camera parameters by using the recently developed calibration harp, instead of incorporating them all in a global evaluation. As an advantage, it helps to avoid the residual error compensation inherent for global calibration methods. The drawback is that it requires building and using an additional calibration instrument - namely, the calibration harp. However, after the distortion is calculated for the fixed camera settings, we can treat the processed images as distortion-compensated, and this allows to calculate other camera parameters separately from distortion.

The second addressed aspect was a correction of the perspective bias for the circled pattern, which was achieved by the incorporation of the conic affine transformation into homography estimation. The function serves as a compensator and at the same time it allows avoiding the use of conic contour detection; the main feature still remains a conic centroid. Of course, this would not be possible without advance distortion compensation, since distortion bias has much higher magnitude than the perspective bias, and elimination is only possible if the distortion field is known in advance, which becomes feasible using the calibration harp. The numerical results for both undistorted synthetic and distorted real data demonstrated that our method allows to get more stable results for camera calibration parameters, meaning lesser magnitude of parameters variance.

The possible directions for future work include using more broad dataset, which could be obtained with different DSLR cameras. Also, the camera calibration is only the first step in precise 3D reconstruction chain, therefore, the comparison of precision of retrieved 3D data could be done with state-of-art methods and against the groundtruth, which would help to evaluate the final precision gain when using our method in context of 3D reconstruction.

ACKNOWLEDGMENT

Part of this work was funded by the Agence Nationale de la Recherche (ANR), project STEREO (program ASTRID 2012).

REFERENCES

[1] R. I. Hartley, "Euclidean reconstruction from uncalibrated views," in *Proceedings of the Second Joint European - US Workshop on Applications of Invariance in Computer Vision*. London, UK, UK: Springer-Verlag, 1994, pp. 237–256. [Online]. Available: http://dl.acm.org/citation.cfm?id=647302.760233

[2] D. C. Brown, "Close-range camera calibration," *Photogrammetric Engineering*, vol. 37, no. 8, pp. 855–866, 1971.

[3] W. Faig, "Calibration of close-range photogrammetry systems: mathematical formulation," *Photogrammetric Engineering and Remote Sensing*, vol. 41(12), pp. 127–140, 1975.

[4] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000. [Online]. Available: http://dx.doi.org/10.1109/34.888718

[5] J.-Y. Bouguet. (2000) Matlab camera calibration toolbox. [Online]. Available: http://www.vision.caltech.edu/bouguetj/calib_doc/

[6] M. Pierrot Deseilligny and I. Clery, "Apero, an open source bundle adjusment software for automatic calibration and orientation of set of images," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXXVIII-5/W16, pp. 269–276, 2011. [Online]. Available: http://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XXXVIII-5-W16/269/2011/

[7] R. Y. Tsai, "A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses," *IEEE Journal on Robotics and Automation*, vol. RA-3, pp. 323–44, 1992. [Online]. Available: http://dl.acm.org/citation.cfm?id=136913.136938

[8] J. Weng, P. Cohen, and M. Herniou, "Camera calibration with distortion models and accuracy evaluation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 10, pp. 965–980, Oct. 1992. [Online]. Available: http://dx.doi.org/10.1109/34.159901

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

[10] P. F. Sturm and S. J. Maybank, "On plane-based camera calibration: A general algorithm, singularities, applications." in *CVPR*. IEEE Computer Society, 1999, pp. 1432–1437.

[11] J.-N. Ouellet and P. Hebert, "A simple operator for very precise estimation of ellipses," in *Canadian Conference on Computer and Robot Vision*, 2007, pp. 21–28.

[12] Z. Tang, "High accuracy measurement in 3d stereo reconstruction," Ph.D. dissertation, ENS Cachan, France, 2011.

[13] J. Heikkilä, "Geometric camera calibration using circular control points," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 10, pp. 1066–1077, Oct. 2000. [Online]. Available: http://dx.doi.org/10.1109/34.879788

[14] J. Kannala and S. S. Brandt, "A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses," *IEEE Trans. Pattern Analisys and Machine Intelligence*, vol. 28, pp. 1335–1340, 2006.

[15] J.-N. Ouellet, F. Rochette, and P. Hebert, "Geometric calibration of a structured light system using circular control points," in *3D Data Processing, Visualization and Transmission*, 2008, pp. 183–190.

[16] C. Yang, F. Sun, and Z. Hu, "Planar conic based camera calibration," in *ICPR*, 2000, pp. 1555–1558.

[17] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*, 2nd ed. New York, NY, USA: Cambridge University Press, ISBN 0521540518, 2004.

[18] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *SIAM Journal on Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.

[19] V. Rudakova and P. Monasse, "Precise correction of lateral chromatic aberration in images," in *PSIVT*, 2013.

[20] M. Higuchi, A. Datta, and T. Kanade. (2012) Software package for precise camera calibration. [Online]. Available: http://www.ri.cmu.edu/research_project_detail.html?project_id=617&menu_id=261