

Robust Fourier-based Checkerboard Corner Detection for Camera Calibration

Benjamin Spitschan^[0000–0001–8945–9480] and Jörn Ostermann

Institut für Informationsverarbeitung (TNT) · Leibniz Universität Hannover
Appelstr. 9A, 30167 Hannover, Germany
spitschan@tnt.uni-hannover.de

Abstract Precise localization of reference markers is crucial for the accuracy of target-based camera calibration. State-of-the art detectors, however, are sensitive to optical blur corrupting the image in many practical calibration scenarios. We propose a novel method for the sub-pixel refinement stage of common checkerboard target detectors. It uses the symmetry of checkerboard crossings and exploits the periodicity in the angular frequency domain when the origin of a polar coordinate system is centered at the crossing. The detector estimates the crossing center’s sub-pixel position by minimizing spurious frequency components that occur increasingly at ever larger distances from the crossing center.

An average localization error of 0.08 px is achieved in noisy and artificially blurred synthetic images, surpassing the state of the art by 65 %. In addition, we evaluated the detector in real-world camera calibration using a public data set, achieving an reprojection error of 0.11 px compared to 0.27 px for the state of the art.

Keywords: Optical blur · Camera calibration · Checkerboard detection

1 Introduction

1.1 Motivation

Fiducial markers with distinct appearance properties are widely used in many computer vision tasks. Accurate localization of these markers within an image determines the performance of applications such as camera calibration, photogrammetry, or augmented reality. For calibration, checkerboard crossings, or X-corners, are superior to other markers [1] and are therefore predominantly used.

Since accurate marker localization is crucial for subsequent estimation tasks, research on their robust detection is of ongoing interest. An important aspect that has been widely ignored, though, is the influence of the focussing behavior of real lenses, i. e. optical systems not observing the pin-hole camera model. Common algorithms for calibration target detection rely on images exhibiting sharp high-contrast edges. In contrast, we imagine a scenario in which a camera focussed to far distances needs to be calibrated. In this setting, the calibration would usually be carried out with targets in near-range, the camera being refocussed to obtain focussed images of the target. However, the geometric imaging properties, i. e. the calibration parameters, change when the focus is moved. It is therefore desirable to accurately detect markers in unfocussed, blurred images as well.

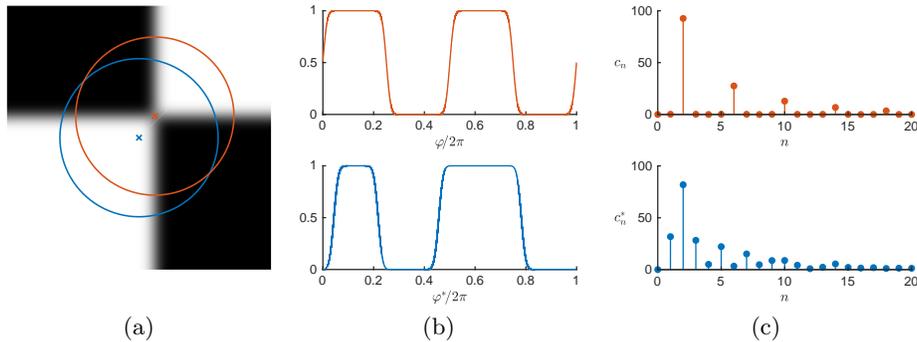


Fig. 1. (a) Decentered and blurred x-corner, (b) signal intensities along local polar angles φ (red) and φ^* (blue), and (c) corresponding magnitudes c_n and c_n^* of complex Fourier series coefficients.

1.2 Related Work

Commonly, the detection of checkerboard calibration targets in images is carried out in a two-stage hierarchical fashion. First, a coarse detection is typically implemented by Harris-type corner detection [2], or by detecting crossing lines.

In the second step, the coarse localization is refined to sub-pixel accuracy. Frequently, sub-pixel x-corner locations are estimated by fitting a polynomial surface and computing the saddle point [3, 4]. For edge approximation, gradient-based methods are used [5], or blurred step responses are fitted to the image [6]. Donn e *et al.* [7] use a convolutional neural network (CNN) to train a single-step detector under different viewing angles. Previous approaches, however, often suffer from reduced precision in the presence of strong optical blur.

In [8, 9] circular symmetry is used for crossing detection, which in part spurred the development of the proposed method. Further inspiration was drawn from [10] in which angular steerable filters are used as a matching template.

1.3 Contributions

To enable accurate checkerboard-crossing detection even in severely blurred images, we propose a new specialized detector comprising the following advancements:

- We restate that the projective transformation of a single checkerboard crossing is affine and derive the resulting angles at which it is perspectively distorted;
- We provide a Fourier analysis of the intensity signal of perspectively distorted x-corners subject to Gaussian blur given in polar coordinates;
- We formulate an optimization problem in the frequency domain for locating the center of an x-corner, exploiting angular periodicities;
- And we present comparative experimental results for the localization error on a set of *synthetic* images, and for reprojection errors and estimation variances of state-of-the-art camera calibration on a set of *real* images.

Structure This paper is organized as follows: Following the formulation of the checkerboard pattern in Section 2, we derive its geometric distortion under perspective transformation and its blurred intensity signal. In Section 3, the crossing-center detection is discussed. We present a twofold experimental evaluation in Section 4 using synthetic images and by applying the method to camera calibration, and conclude with a discussion (Section 5).

2 Image Formation

We assume the calibration target plane to be centered in the xy -plane of a 3-D world coordinate system, i. e. $z = 0$, aligned with the x - and y -axes. Let intensity values $\{0, 1\}$ represent black and white, respectively. The 2-D intensity signal $s(x, y)$ of a single black/white crossing centered at the origin is given by

$$s(x, y) = \frac{1}{2} + \frac{1}{2} \operatorname{sgn}(x) \operatorname{sgn}(y), \quad (1)$$

where $\operatorname{sgn}(\cdot)$ denotes the signum function. In 3-D homogeneous coordinates, its geometry is given by two orthogonal lines

$$\mathbf{L}_1 \equiv (\tau_1, 0, 0, 1)^\top \quad \text{and} \quad \mathbf{L}_2 \equiv (0, \tau_2, 0, 1)^\top, \quad \tau_1, \tau_2 \in \mathbb{R}. \quad (2)$$

2.1 Geometric Projection

Subsequently, the common model of a projective pin-hole camera is adopted, neglecting non-linear lens distortions. Here, a projective camera is assumed with rotation $\mathbf{R} \equiv (r_{ij}) \in \mathbb{R}^{3 \times 3}$ and translation $\mathbf{T} \equiv (t_x, t_y, t_z)^\top$ (extrinsic parameters), and camera calibration matrix $\mathbf{K} \in \mathbb{R}^{3 \times 3}$ (intrinsic parameters). We employ a simplified calibration matrix with focal length f , vanishing skew coefficient, and a centered principal point. Hence, the projections of \mathbf{L}_1 and \mathbf{L}_2 by a projection matrix $\mathbf{P} \equiv \mathbf{K} [\mathbf{R} | \mathbf{T}]$ are

$$\mathbf{l}_1 = \mathbf{P} \mathbf{L}_1 = \begin{pmatrix} fr_{11}\tau_1 + ft_x \\ fr_{21}\tau_1 + ft_y \\ r_{31}\tau_1 + t_z \end{pmatrix} \quad \text{and} \quad \mathbf{l}_2 = \mathbf{P} \mathbf{L}_2 = \begin{pmatrix} fr_{12}\tau_2 + ft_x \\ fr_{22}\tau_2 + ft_y \\ r_{32}\tau_2 + t_z \end{pmatrix}, \quad (3)$$

or, after normalization to Euclidean 2-D coordinates of the image plane,

$$\hat{\mathbf{l}}_1 = \frac{1}{r_{31}\tau_1 + t_z} \begin{pmatrix} fr_{11}\tau_1 + ft_x \\ fr_{21}\tau_1 + ft_y \end{pmatrix} \quad \text{and} \quad \hat{\mathbf{l}}_2 = \frac{1}{r_{32}\tau_2 + t_z} \begin{pmatrix} fr_{12}\tau_2 + ft_x \\ fr_{22}\tau_2 + ft_y \end{pmatrix}. \quad (4)$$

Applying the re-parametrizations

$$\tau'_1 = - \left(\frac{t_z^2}{r_{31}\tau_1} + t_z \right)^{-1} \quad \text{and} \quad \tau'_2 = - \left(\frac{t_z^2}{r_{32}\tau_2} + t_z \right)^{-1} \quad (5)$$

to $\hat{\mathbf{l}}_1$ and $\hat{\mathbf{l}}_2$, respectively, we obtain

$$\hat{\mathbf{l}}_1 = \begin{pmatrix} f \left(t_x - \frac{r_{11}}{r_{31}} t_z \right) \tau'_1 + f \frac{t_x}{t_z} \\ f \left(t_y - \frac{r_{21}}{r_{31}} t_z \right) \tau'_1 + f \frac{t_y}{t_z} \end{pmatrix} \quad \text{and} \quad \hat{\mathbf{l}}_2 = \begin{pmatrix} f \left(t_x - \frac{r_{12}}{r_{32}} t_z \right) \tau'_2 + f \frac{t_x}{t_z} \\ f \left(t_y - \frac{r_{22}}{r_{32}} t_z \right) \tau'_2 + f \frac{t_y}{t_z} \end{pmatrix}, \quad (6)$$

which is a 2-D affine transformation of the straight lines

$$\mathbf{L}'_1 \equiv (\tau'_1, 0, 1)^\top \quad \text{and} \quad \mathbf{L}'_2 \equiv (0, \tau'_2, 1)^\top \quad (7)$$

in the image plane. We see that for the pair of orthogonal lines \mathbf{L}_1 and \mathbf{L}_2 , the projective transform reduces to an affine one.

Using Eq. (6) we can calculate the angles β_1 and β_2 at which the projections $\hat{\mathbf{l}}_1$ and $\hat{\mathbf{l}}_2$, respectively, intersect the x -axis of the image plane, as

$$\beta_1 = \arctan \frac{r_{31}t_y - r_{21}t_z}{r_{31}t_x - r_{11}t_z} \quad \text{and} \quad \beta_2 = \arctan \frac{r_{32}t_y - r_{22}t_z}{r_{32}t_x - r_{12}t_z}, \quad (8)$$

and see that both lines intersect with the angle $\beta = \beta_2 - \beta_1$, $0 < \beta < \pi$, at $\mathbf{x}_0 \equiv (x_0, y_0)^\top$. Note that throughout this paper, the quadrant-aware arc tangent is used. The projected intensity at $\hat{\mathbf{x}} \equiv (\hat{x}, \hat{y}, 1)^\top = \mathbf{A}(x, y, 1)^\top$ is then

$$\hat{s}(x, y) = s(\hat{x}, \hat{y}) \quad (9)$$

or, using the backprojection $\mathbf{A}^{-1}\hat{\mathbf{x}}$ and the substitutions Eq. (8),

$$\begin{aligned} \hat{s}(x, y) &= \frac{1}{2} + \frac{1}{2} \operatorname{sgn}((x - x_0) \sin \beta_2 - (y - y_0) \cos \beta_2) \\ &\quad \times \operatorname{sgn}(-(x - x_0) \sin \beta_1 + (y - y_0) \cos \beta_1). \end{aligned} \quad (10)$$

We can simplify $\hat{s}(x, y)$ further by dropping the translation \mathbf{x}_0 and rotating by $-\beta_1$ such that $\hat{\mathbf{l}}_1$ coincides with the x -axis, and obtain the aligned intensity signal

$$\hat{s}(x, y) = \frac{1}{2} + \frac{1}{2} \operatorname{sgn}(x - y \cot \beta) \operatorname{sgn}(y). \quad (11)$$

2.2 Optical Blur

The blur introduced by the optical system is commonly assumed to be Gaussian, i. e. the projected 2-D image intensity signal $\hat{s}(x, y)$ is subject to a convolution

$$\hat{s}_{\text{blur}}(x, y) = \hat{s}(x, y) * g(x, y) \quad (12)$$

with a separable isotropic Gaussian blur kernel

$$g(x, y) = \frac{1}{2\sigma^2\pi} e^{-\frac{x^2+y^2}{2\sigma^2}}. \quad (13)$$

Case 1: x-corner without projective transformation ($\beta = 90^\circ$). First, we consider a simple undistorted checkerboard crossing at the origin (Eq. (1)). Carrying out the convolution yields the 2-D step response of the Gaussian kernel

$$s_{\text{blur}}(x, y) = \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{x}{\sqrt{2}\sigma}\right) \operatorname{erf}\left(\frac{y}{\sqrt{2}\sigma}\right), \quad (14)$$

where $\operatorname{erf}(\cdot)$ is the Gaussian error function [11, Eq. (7.2.1)]. We change to polar coordinates (r, φ) with $x = r \cos \varphi$, $y = r \sin \varphi$ according to

$$s_{\text{blur}}(r, \varphi) = \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{r \cos \varphi}{\sqrt{2}\sigma}\right) \operatorname{erf}\left(\frac{r \sin \varphi}{\sqrt{2}\sigma}\right) \quad (15)$$

and develop the Fourier series expansion

$$s_{\text{blur}}(r, \varphi) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos n\varphi + b_n \sin n\varphi). \quad (16)$$

Apart from the constant offset, Eq. (15) is odd in φ , hence we immediately find $a_n = 0$ for $n = 1, 2, 3, \dots$, and $a_0 = 1$.

Since $s_{\text{blur}}(r, \varphi)$ is π -periodic with respect to φ (cf. Fig. 1 (b), top panel), its Fourier expansion is non-zero only for the doubled odd n , i. e. $n = 2, 6, 10, \dots = 2(2k + 1)$, $k \in \mathbb{N}_0$, and equals $b_n = 0$ otherwise (cf. Fig. 1 (c), top panel).

Case 2: x-corner subject to projective transformation ($0 < \beta < 180^\circ$).

For $\beta \neq 90^\circ$, $\hat{s}_{\text{blur}}(r, \varphi)$ is not odd any more, cf. Fig. 1 (b), bottom panel. Due to the signal nevertheless being π -periodic, all odd coefficients still vanish, i. e. $a_n, b_n = 0$ for $n = 2k + 1$. However, the even cosine coefficients differ from zero in this case, i. e. $a_n \neq 0, n = 2k$ (cf. Fig. 1 (c), bottom panel).

3 Corner Detection

Now we assume the origin of the local polar coordinate system be offset by \mathbf{x}_0 (cf. Fig. 1 (a)). Using the decentered polar coordinates

$$r^*(x, y) \equiv r(x - x_0, y - y_0) = \sqrt{(x - x_0)^2 + (y - y_0)^2} \quad (17)$$

and

$$\varphi^*(x, y) \equiv \varphi(x - x_0, y - y_0) = \arctan \frac{y - y_0}{x - x_0}, \quad (18)$$

we obtain the complex Fourier coefficients

$$a_n^* + j b_n^* = \frac{1}{2\pi} \int_0^{2\pi} \hat{s}_{\text{blur}}(r^*, \varphi^*) e^{jn\varphi^*} d\varphi^*. \quad (19)$$

Since we saw that the Fourier components c_n are non-zero only for even n due to the π -periodicity of \hat{s}_{blur} , the task of localizing the center of the x-corner, viz. the estimation of the offset $\mathbf{x}_0 = (x_0, y_0)^\top$, is simply the problem of minimizing the magnitudes $c_n^* = |a_n^* + j b_n^*|$, $n = 2k + 1$, of the odd components along the offset polar angle φ^* with respect to \mathbf{x}_0 . This can be expressed as the non-linear least-squares minimization problem

$$\mathbf{x}_0^{\text{est}} \equiv (x_0^{\text{est}}, y_0^{\text{est}})^\top = \arg \min_{x_0, y_0} \sum_{k=0}^{\infty} |a_{2k+1}^* + j b_{2k+1}^*|^2, \quad (20)$$

or, inserting Eq. (19),

$$\mathbf{x}_0^{\text{est}} = \arg \min_{x_0, y_0} \sum_{k=0}^{\infty} \left| \frac{1}{2\pi} \int_0^{2\pi} \hat{s}_{\text{blur}}(r^*, \varphi^*) e^{j(2k+1)\varphi^*} d\varphi^* \right|^2 \quad (21)$$

for a particular radius r^* . To exploit image information at all radii, we integrate along r^* , wherein we account for the Jacobian determinant r^* of polar coordinates. Moreover, we introduce a weighting term $r^* \exp\left(-\frac{r^{*2}}{2\alpha^2}\right)$ such that the influence at very small ($r \ll \alpha$) and large radii ($r > \alpha$) is attenuated for improved robustness, and obtain an improved second estimate $\mathbf{x}_0^{\text{est}_2} \equiv (x_0^{\text{est}_2}, y_0^{\text{est}_2})^\top$ as

$$\mathbf{x}_0^{\text{est}_2} = \arg \min_{x_0, y_0} \sum_{k=0}^{\infty} \left| \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} r^{*2} \exp\left(-\frac{r^{*2}}{2\alpha^2}\right) \hat{s}_{\text{blur}}(r^*, \varphi^*) e^{j(2k+1)\varphi^*} dr^* d\varphi^* \right|^2. \quad (22)$$

Since the signal \hat{s}_{blur} is only given on a Cartesian sampling grid, we now change to discrete positions (i, j) in a finite window of interest W of size $N \times N$ and find

$$\mathbf{x}_0^{\text{est}_2} = \arg \min_{x_0, y_0} \sum_{k=0}^{\infty} \left| \sum_{(i,j) \in W} \frac{r^*(i,j)}{N^2} \exp\left(-\frac{r^{*2}(i,j)}{2\alpha^2}\right) \hat{s}_{\text{blur}}(i,j) e^{j(2k+1)\varphi^*(i,j)} \right|^2 \quad (23)$$

with $r^*(i, j) = \sqrt{(i-x_0)^2 + (j-y_0)^2}$ and $\varphi^*(i, j) = \arctan \frac{j-y_0}{i-x_0}$. By computing the Jacobian which is omitted here, it can be seen that Eq. (23) is locally convex within the sub-pixel search range $\|\mathbf{x}_0\| \leq 1/\sqrt{2}$ px.

4 Experimental Results

4.1 Synthetic Images

Grayscale images of blurred and perspectively distorted x-corners for varying blur parameter σ , distortion angle β , and sub-pixel offset (x_0, y_0) were generated. They were subsequently corrupted by additive Gaussian noise of reasonable variance $\sigma_n^2 = 25$ (w. r. t. an intensity value range of 0...255), corresponding to a PSNR of 34.1 dB, and quantized to 8 bit. Experiments were repeated 20× with different realizations of random noise in each iteration, with the results averaged.

In all experiments, we assumed that the first coarse localization with integer-pixel accuracy has already been done, and only evaluated the sub-pixel detection stage. The proposed method Eq. (23) was realized using MATLAB's `lsqnonlin()`, which implements the trust-region reflective algorithm. The analytically derived Jacobian was provided as input. An offset of $\mathbf{x}_0 = (0, 0)$ was used as initialization, and the lower and upper bounds $\xi_l = (-0.5, -0.5)$ and $\xi_u = (0.5, 0.5)$, respectively, were imposed. A window size of $N = 81$ was used to accommodate for high blur levels. Average execution times of 0.42 s were measured on a 3.5 GHz CPU.

We benchmarked against OpenCV v3.4.1's `cornerSubPix()` [5] based upon Bouguet's CALTECH toolbox [3], and `detectCheckerboardPoints()` from MATLAB's CV System Toolbox (v8.1, R2018a) based upon Geiger *et al.* [12].

For distortion angles $\beta = 90^\circ$, $(90 \pm 45)^\circ$, and $(90 \pm 60)^\circ$, the absolute localization error ε averaged over all sub-pixel offsets is depicted in Fig. 2 (a) as a function of the blur parameter σ . We can see that the proposed method outperforms the state of the art increasingly with the amount of blur. It was

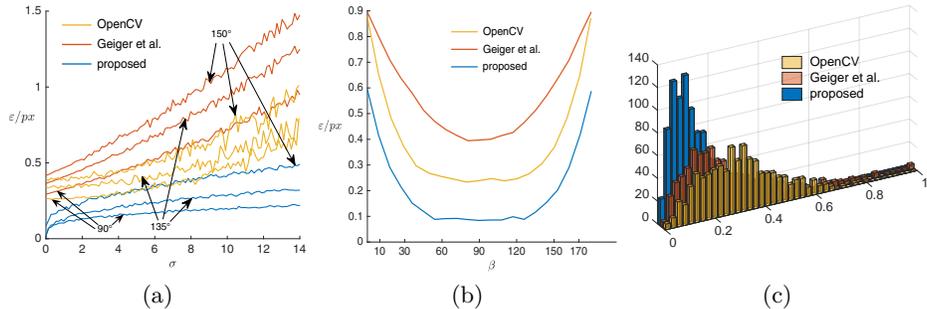


Fig. 2. *Synthetic x-corners:* Absolute localization error (a) w. r. t. blur parameter σ , and (b) w. r. t. distortion angle β at a typical blur level of $\sigma = 6$ px. (c) *Real images:* Distribution of reprojection errors e_{repr} after camera calibration.

noted that the detector still performs well in blur ranges of $\sigma > 15$ px where the state-of-the-art methods fail to return meaningful results.

For $\sigma = 6$ px which was empirically determined from the Dima data set (cf. Section 4.2 and [13]) as a typical value for real-world applications, we varied the distortion angle β in a range from 10° to 170° (Fig. 2 (b)). We find that the localization error of our method is invariant for distortions between 50° and 130° .

4.2 Application: Calibration with Checkerboard Targets

To evaluate the proposed detector for practical applications, we applied it to camera calibration with checkerboard calibration rigs. For that purpose, we chose the CALTECH Camera Calibration Toolbox for MATLAB [3], which implements the popular method by Zhang [14], and substituted the three aforementioned sub-pixel detectors for the toolbox-native one. We used the data set by Dima *et al.* [13], comprising 18-Mpx images of a target taken with a DSLR at 17 viewpoints.

In this setup, no ground truth data were available. We therefore followed the common approach to assess the detector by analysing the parameter estimation variances. The key results are presented in Table 1. It can be seen that the enhanced accuracy demonstrated in Section 4.1 significantly reduced the standard deviation of the estimated camera parameters in most of the cases compared to the state of the art, and reduced the mean absolute reprojection error e_{repr} by 58 %. Histograms of e_{repr} are depicted in Fig. 2 (c).

5 Conclusions

In this paper, we presented a novel method for sub-pixel accurate estimation of the position of checkerboard crossings in images. It exploits the π -periodicity of their intensity signal in local polar coordinates, which are analyzed in the frequency domain to derive a non-linear cost function for least-squares optimization. To this end, we restated the transformation of x-corners under perspective projection.

The proposed detector was evaluated with synthetic images of blurred checkerboard crossings and compared against two widely-used state-of-the-art methods. The results show an improvement of the localization error by 65 % at minimum.

Table 1. *Real images.* Impact of detection method on calibration accuracy.

| Method | OpenCV [5] | Geiger <i>et al.</i> [12] | proposed |
|--|-------------------|---------------------------|-------------------------------------|
| Position/mm (SD) | 5.73 | 11.21 | 4.21 |
| Principal point/px (SD) | 8.87 | 12.15 | 7.09 |
| Radial coeff. κ_1 (SD) | 0.014 | 0.004 | 0.005 |
| $e_{\text{repr}}/\text{px}$ (MAE \pm SD) | 0.301 ± 0.215 | 0.265 ± 0.190 | 0.112 ± 0.060 |

In addition, in order to demonstrate its practical advantages, our method was used as a drop-in replacement in a common checkerboard-based camera calibration toolchain, and again evaluated against both state-of-the-art detectors using a real calibration image set. We found better or—in one case—comparable results for the estimation variance, and a 58 % lower reprojection error over all experiments. We conclude that our method for detecting checkerboard crossings with very high sub-pixel accuracy will be greatly beneficial for real-world applications.

References

- Mallon, J. and Whelan, P. F., “Which pattern? Biasing aspects of planar calibration patterns and detection methods,” *PRL*, vol. 28, no. 8, pp. 921–930, Jun. 2007.
- Harris, C. and Stephens, M. J., “A combined corner and edge detector,” in *Proc. Alvey Vision Conf.*, 1988, pp. 147–151.
- Bouguet, J.-Y., *Camera Calibration Toolbox for Matlab*, Jun. 1, 2017. [Online]. Available: http://www.vision.caltech.edu/bouguetj/calib_doc.
- Fürsattel, P. *et al.*, “OCPAD – occluded checkerboard pattern detector,” in *Proc. WACV*, IEEE, Mar. 2016.
- OpenCV, Open Source Computer Vision library*, version 3.4.1, Feb. 27, 2018. [Online]. Available: <http://opencv.org/>.
- Ha, H. *et al.*, “Accurate camera calibration robust to defocus using a smartphone,” in *Proc. ICCV*, IEEE, 2015.
- Donné, S. *et al.*, “MATE: Machine learning for adaptive calibration template detection,” *Sensors*, vol. 16, no. 11, p. 1858, Nov. 2016.
- Bennett, S. and Lasenby, J., “CHESS – quick and robust detection of chess-board features,” *Computer Vision and Image Understanding*, vol. 118, pp. 197–210, 2014.
- Bok, Y. *et al.*, “Automated checkerboard detection and indexing using circular boundaries,” *PRL*, vol. 71, pp. 66–72, Feb. 2016.
- Mühlich, M. *et al.*, “Design and implementation of multisteerable matched filters,” *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 34, no. 2, pp. 279–291, Feb. 2012.
- NIST *Handbook of Math. Functions*. Cambridge, U. K.: Cambridge Univ. Pr., 2010.
- Geiger, A. *et al.*, “Automatic camera and range sensor calibration using a single shot,” in *Proc. ICRA*, IEEE, May 2012.
- Dima, E. *et al.*, “Assessment of multi-camera calibration algorithms for two-dimensional camera arrays relative to ground truth position and direction,” in *Proc. 3DTV-CON*, IEEE, Jul. 2016.
- Zhang, Z., “A flexible new technique for camera calibration,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, 2000.