

## Abstract

Our goal is to validate the out-of-focus camera calibration methods proposed by previous authors, and to disentangle variables such as the effects of image spatial resolution, pattern array size, and noise that affect camera calibration performance. Using synthetic and real-world experiments, we compare Phase-shifted Circular Gradient (PCG) patterns and circle grids to examine their relative calibration accuracy under out-of-focus blurred conditions. Real-world experiments are performed using a color E-Ink display as the calibration target, and target positioning is facilitated using a robotic arm. Initial results show that there is no relative advantage to the phase-shifted approaches over using a grid of small but still detectable circles, however, the difference is very small, and it is hypothesized that the E-Ink display might not work favourably for PCGs due to its inability to render grayscale patterns without dithering. This work provides a foundation for further investigation to compare real-world and synthetic cases which could offer additional clues about the effectiveness of each method and, could highlight how further improvements to accuracy can be found.

## 1 Introduction

For many 3D computer vision-based tasks, it is necessary to find relationships between 3D points in a scene and corresponding points that can be detected in a 2D image. To determine these relationships, a camera's intrinsic and extrinsic parameters need to be found through the process of camera calibration. Intrinsic parameters include focal lengths in  $x$  and  $y$  directions, and extrinsic parameters refer to the camera's location, and pose in 3D space. Applications such as Simultaneous Localization and Mapping (SLAM), Structure from Motion (SfM), 3D reconstruction, and non-contact optical measurement require frequent translations between 2D and 3D and rely heavily on high-accuracy estimations of camera parameters.

Camera calibration can be performed under different circumstances. In dynamic situations, such as with a drone, mobile robot, or vehicle, frequent estimation of the intrinsic and extrinsic parameters may be necessitated by an auto-calibration process, which can rely on natural scene features. Different auto-calibration approaches can range from analytical methods relying on numerical optimization, but more recent advances in Deep Learning (DL) have allowed researchers to explore end-to-end data-driven methods where camera parameters are directly outputted from trained networks [1–3].

In static situations, such as the manufacturing of a new camera lens system, structured patterns can be used to obtain high accuracy one-time camera parameter estimations. Over the years, various structured patterns have been devised to create 1D, 2D, and 3D calibration targets [4–9], however, 2D planar targets remain the most frequently used due to simplicity and cost-effectiveness. 2D targets can be categorized based on their ability to emit or reflect light and can be categorized as active or passive respectively. In most cases, active calibration targets consist of digital displays such as mobile phone, laptop, or desktop monitor displays and can display consecutive patterns while being held stationary for each specified pose.

Using 2D targets, the calibration process typically involves capturing several images of the structured pattern in various poses. Using these images, control points are found by detecting pattern features, and the control points are provided to an optimization process that refines an initial estimation of intrinsic and extrinsic camera parameters. Two of the most well-established and frequently referenced methods are from the work of Tsai and Zhang [7, 9] where Zhang's method originally established the use of a checkerboard calibration pattern, which is still widely used today.

As new applications demand greater camera calibration accuracy, and flexibility, improved calibration pattern and feature detection meth-

ods remain an active area of research. In particular, long focal length cameras present special challenges since focused calibration targets need to be mounted far away from the camera. This can be problematic in space-constrained environments such as a manufacturing floor or assembly line. Further to this, is the issue of decreasing accuracy of checkerboard patterns as they are used with increasing levels of defocus blur. As a workaround, researchers have been trying to find better ways of performing camera calibration from out-of-focus images [10–15], which allows the target to be placed closer to the camera.

In this work, we attempt to re-produce some of the aforementioned research to validate its effectiveness on blurred images. For comparison, we use circle grids, which are another commonly used calibration pattern, to test the effectiveness of these methods. We also attempt to disentangle calibration variables using synthetic data and real data facilitated by the use of an E-Ink display target, and the relative repeatability of poses that are provided by a collaborative robot arm.

## 2 Related Work

To overcome the accuracy limitations of checkerboard patterns when working with out-of-focus images, Wang et al. [10] proposed the novel use of phase-shifted circular grating (PCG) arrays using a digital display as an active calibration target. The proposed method works by creating three separate patterns containing grids of grayscale PCG circles where each gradient is shifted by  $2\pi/3$ . In experiments, an LCD display was set up and captured in a variety of poses. For each pose, three images were captured from an SLR camera where each image corresponded to a respective PCG array pattern. Using a three-step phase-shifting algorithm, a wrapped phase image was generated which contains abrupt changes in grayscale intensity that form pronounced circles (Fig. 1). To account for perspective distortion, an ellipse-fitting technique was used to recover the circle centers. The circle centers then underwent sub-pixel refinement and subsequently were used to calibrate the camera. Wang et al. also demonstrated that their method was largely invariant to blur when subjected to synthetically blurred images.

Building on this work, Cai et al. [11] proposed a similar method using PCGs that generate two wrapped phase image discontinuities. Using this approach, two concentric ellipses are fitted which are used to compute the imaged center using a theoretically more accurate geometric-based and algebraically derived representation.

To eliminate the requirement for capturing three consecutive grayscale PCG patterns, Wang et al. [12] later proposed combining the three grayscale PCG arrays into a single colour array. To do this, the pixel intensities for each grayscale image are stored in each of the three RGB colour channels to create one colour PCG image which can be separated and reconstructed with the same three-step phase shifting algorithm.

To determine where possible accuracy improvements to out-of-focus calibration can be made, we begin by implementing Wang's originally proposed method [10], followed by Cai [11], to carefully evaluate their claims, and better understand the subtleties leading to improved calibration results.

## 3 Methods

### 3.1 PCG Theory

Our work has reproduced the method originally proposed by Wang et al. [10] with additional synthetic data generation, and real world experimentation using an E-ink display and collaborative robot arm.

From Eq. 1 Three grey-scale phase-shifted circular gradient (PCG) patterns are created based on pixel intensities  $I_1$  to  $I_3$ :

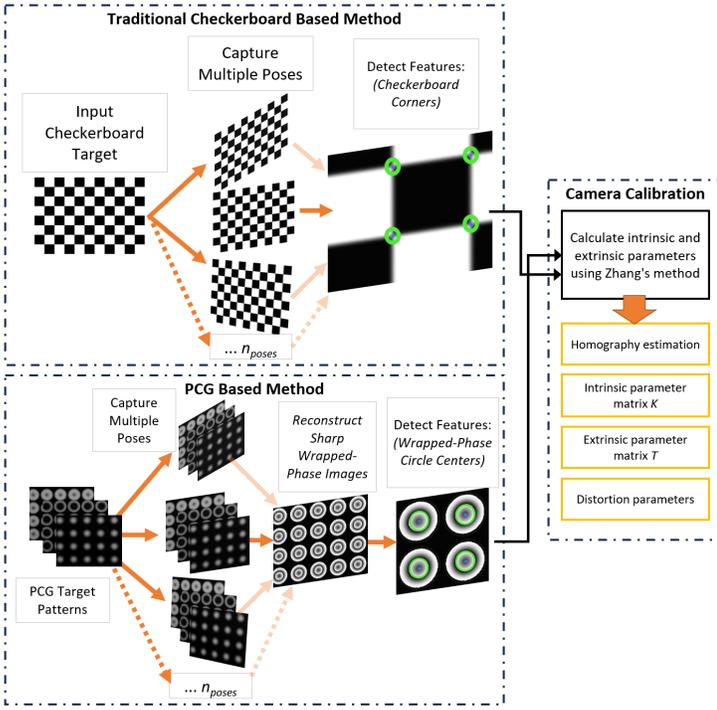


Fig. 1: PCG Camera calibration method originally proposed by Wang et al. [10] for accurate camera calibration using out-of-focus images vs. traditional method using checkerboards

$$\begin{cases} I_1(x,y) = A + B \cos[\phi(x,y) - 2\pi/3], \\ I_2(x,y) = A + B \cos[\phi(x,y)], \\ I_3(x,y) = A + B \cos[\phi(x,y) + 2\pi/3], \end{cases} \quad (1)$$

Where Eq.2 provides the unwrapped phase, where  $T$  is the period, and  $(x,y)$  is the Euclidean distance from any pixel to the circle center.

$$\phi(x,y) = \frac{2\pi r(x,y)}{T}, \quad (2)$$

PCG images 1 to 3 are repeated in a grid to create three separate pattern images. The pattern images are captured consecutively for each pose, and multiple poses with the consecutive pattern images are captured. Using Eq.3, each set of three images can be combined to create a wrapped phase image for each pose (Fig. 2).

$$\psi(u,v) = \arctan\left(\sqrt{3} \frac{J_1 - J_3}{2J_2 - J_1 - J_3}\right), \quad (3)$$

Once the wrapped-phase images are generated for each pose, feature detection is completed. In this case, the circle centers need to be identified with sub-pixel accuracy so they can be used as control points for calibration. Based on Cai et al. [11], to compensate for perspective distortion caused by pose rotation, we intentionally set the period  $T$  of Eq.2 to a value (proportional to the pattern size) that will generate two wrapped-phase discontinuity circles. Using these two discontinuity circles, inner and outer ellipses are fitted.

To fit the ellipses, Gaussian blur is first applied followed by Canny Edge Detection and contour detection. Contours with five points or more are then used to generate ellipses. Ellipses that exceed size and eccentricity criteria are rejected, and the centers of closely overlapping ellipses are averaged. With two sets of fitted ellipses (one inner ellipse and one outer), we compute a corrected center point using eigenvalue decomposition based on the projection relationship described in [11], where the last eigenvalue represents the true circle center.

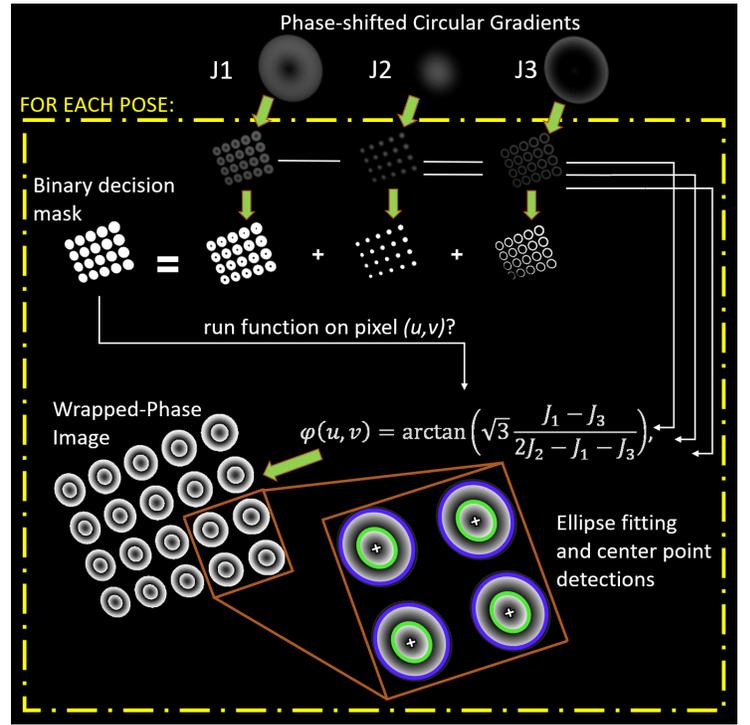


Fig. 2: Phase-shifting circular gradient method

### 3.2 Camera Calibration Using Synthetic Images

For initial validation of [10], images were rendered using Blender Open-Source 3D animation software. Using 4x5 PCG grids, the images were rendered using three copies of the same animation sequence where each of the three grayscale PCG patterns  $I_1$  to  $I_3$  move through a series of repeated poses. From the sequences of images, Eq.3, was used to generate wrapped phase images. Camera calibration was performed using OpenCV based on [9].

Using Blender's perspective camera model with a 50mm focal length, horizontal sensor fit with 36mm width sensor, and F-Stop = 1, the series of three image sequences were rendered multiple times with focus distances varying between 4m and 100m to introduce increasing amounts of defocus blur. All images were rendered as square images with a resolution of 1080 x 1080, which was needed to allow the  $f_x$  and  $f_y$  ground truth focal lengths to be the same. To verify this, Blender's internal Python scripting was used to query the camera model and retrieve the target focal lengths of the virtual camera in the Blender scenes. In the case of all synthetic experiments, the nominal focal lengths are 1500px for both  $f_x$ , and  $f_y$ .

### 3.3 Camera Calibration Using Real-world Images

Real world experiments were conducted in a lab using a Basler acA1920-150uc camera with with a C23-5028-5M 50mm focal length lens. Camera settings were left unchanged, but it should be noted that the default camera gain created some noticeable image noise which affected the captured images of all patterns consistently. The lab lighting was tuned to eliminate as much glare as possible on the display target. Images were captured with a resolution of 1920x1200 px and the target was a 3200 x 1800 E-Ink display which displayed several consecutive patterns per pose including grayscale PCGs and circle grids. All patterns were arranged in a 4x6 grid which was found to be a good match for the screen's aspect ratio. The E-Ink display was mounted onto a Franka Emika Panda 7-DOF collaborative robot, and configured as the end-effector to compensate for gravity. The arm was controlled by ROS interfacing with the Panda's API to move the display target to it's planned poses.

The Panda arm was programmed to translate and rotate in-plane across the camera's field of view and stop at specified poses. For each pose, the display was set to cycle through each of the calibra-

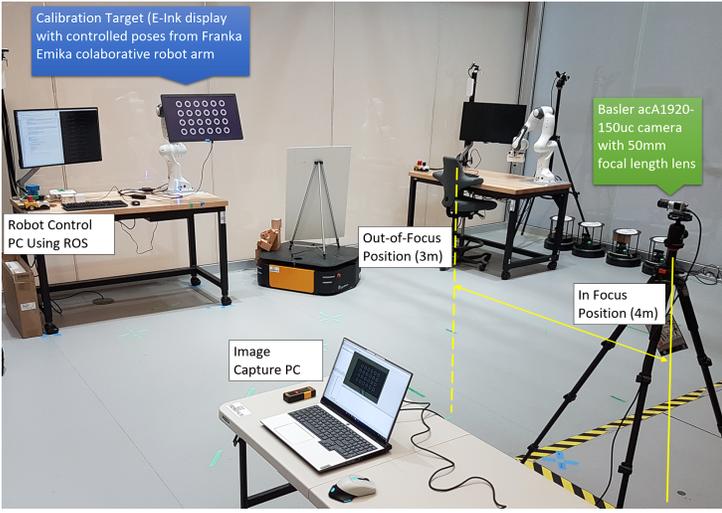


Fig. 3: Lab setup for real world PCG experiment using E-Ink display and Franka Emika collaborative robot arm

| Synthetic Image Results |        |                                                              |                |                |                       |                       |                      |                      |
|-------------------------|--------|--------------------------------------------------------------|----------------|----------------|-----------------------|-----------------------|----------------------|----------------------|
| Focus Distance (m)      | F-Stop | Nominal Focal Length (F <sub>x</sub> & F <sub>y</sub> ) (Px) | F <sub>x</sub> | F <sub>y</sub> | Δ F <sub>x</sub> (Px) | Δ F <sub>y</sub> (Px) | Δ F <sub>x</sub> (%) | Δ F <sub>y</sub> (%) |
| 4                       | 1      | 1500                                                         | 1498.385       | 1498.342       | 1.615                 | 1.65819               | 0.108%               | 0.111%               |
| 10                      | 1      | 1500                                                         | 1499.407       | 1499.55        | 0.5927                | 0.4499                | 0.040%               | 0.030%               |
| 15                      | 1      | 1500                                                         | 1498.743       | 1498.888       | 1.25676               | 1.11229               | 0.084%               | 0.074%               |
| 20                      | 1      | 1500                                                         | 1500.558       | 1500.694       | -0.55769              | -0.69376              | 0.037%               | 0.046%               |
| 25                      | 1      | 1500                                                         | 1500.776       | 1500.844       | -0.77567              | -0.84371              | 0.052%               | 0.056%               |
| 30                      | 1      | 1500                                                         | 1500.001       | 1500.068       | -0.00126              | -0.06827              | <b>0.000%</b>        | <b>0.005%</b>        |
| 100                     | 0.5    | 1500                                                         | 1502.181       | 1502.146       | -2.18149              | -2.14609              | <b>0.145%</b>        | <b>0.143%</b>        |

Table 1: Table of calibration results using synthetic images generated in Blender where 4m is in-focus and > 4m is out-of-focus

tion patterns, and the camera captured the respective patterns. The camera was mounted on a tripod and positioned at a distance of 4m for in-focus image capture and 3m for out-of-focus capturing which coincided with the physical limits of the target motion sequence remaining in the camera's field of view. For both distance cases, no changes were made to the camera's focus ring or aperture, allowing the camera system (lens and sensor) to remain with the same effective focal length. Only the tripod and camera were moved to create the out-of-focus effect. An example of the lab setup can be seen in Fig. 3.

## 4 Preliminary Experimental Results

### 4.1 Synthetic Image Results

Table 1 shows the synthetic image calibration results using grayscale PCGs. Each focus distance case represents a sequence of 10 poses where 4m is in-focus and every distance greater than 4m contains increasing out-of-focus blur. However, since focal length and focus distance can be set independently in the Blender scene environment, the camera remains the same distance from the target while the focal length is unchanged, and the virtual focus distance is decoupled from the target allowing for all focus distance levels to have differing degrees of blur, but with the same size target. This is a unique characteristic that is only possible in the simulated environment, and removes the variable of spatial resolution from comparison, since all target patterns are the same size, regardless of being blurred or in-focus.

In table 1, as expected, the greatest relative % difference between the nominal focal lengths of 1500px (for  $f_x$  and  $f_y$ ), and the returned calibration results exists when the focus distance is 100m. The smallest relative % difference is at 30m, although this is somewhat counter intuitive since the in-focus (4m) case would have been expected to

| Real Image Results                 | In-Focus       |                | Out-of-Focus   |                | Δ F <sub>x</sub> (Px) | Δ F <sub>y</sub> (Px) | Δ F <sub>x</sub> (%) | Δ F <sub>y</sub> (%) |
|------------------------------------|----------------|----------------|----------------|----------------|-----------------------|-----------------------|----------------------|----------------------|
|                                    | F <sub>x</sub> | F <sub>y</sub> | F <sub>x</sub> | F <sub>y</sub> |                       |                       |                      |                      |
| <b>Small Circles</b>               | 9903.62        | 9903.63        | 9926.62        | 9926.94        | <b>22.996</b>         | <b>23.309</b>         | <b>0.232%</b>        | <b>0.235%</b>        |
| <b>Large Circles</b>               | 9888.72        | 9887.93        | 9929.18        | 9929.73        | 40.461                | 41.804                | 0.407%               | 0.421%               |
| <b>PCG Inner Ellipses</b>          | 9886.01        | 9886.43        | 9923.07        | 9923.06        | 37.066                | 36.629                | 0.374%               | 0.369%               |
| <b>PCG Outer Ellipses</b>          | 9849.41        | 9850.55        | 9908.76        | 9909.46        | 59.353                | 58.912                | 0.599%               | 0.595%               |
| <b>PCG Averaging All Ellipses</b>  | 9867.74        | 9868.61        | 9916.96        | 9917.39        | 49.219                | 48.779                | 0.496%               | 0.492%               |
| <b>PCG Center Point Correction</b> | 9898.48        | 9898.69        | 9927.69        | 9927.45        | <b>29.213</b>         | <b>28.759</b>         | <b>0.294%</b>        | <b>0.290%</b>        |

Table 2: Table of calibration results using images captured from lab setup

return the smallest difference. This difference, although small, could be explored by adding poses and / or grid control points, and could potentially reveal further limitations of the calibration optimizer.

### 4.2 Real-World Image Results

For camera calibration results based on real world images, 20 poses were captured for each corresponding set of PCG grayscale patterns from  $I_1$  to  $I_3$ . Additionally, symmetric grids of small and large circles were also captured and used for comparison. Since it is not possible to accurately know the ground truth of the camera system, a measure of success is in the degree of minimizing focal length differences between in and out-of-focus images.

Table 2 shows the calibration results using small and large circles as well as PCG patterns. Several PCG pattern cases are compared where different selections of ellipses were used from all of the detected ellipses in order to provide calibration control points. The most accurate case overall appears to be using the grid of small circles where the circles have a radius of 13 pixels. The second most accurate case is when using the center point correction method on PCG circles, however, the results are close, since there is only a 0.063% difference between the differences of the two cases in and out-of-focus focal lengths.

## 5 Conclusions

This work sets out to compare and validate previously devised methods for camera calibration from out-of-focus blur, and also, provides a foundation for determining which entangled variables affect the overall accuracy of these methods. The clearest way that this was demonstrated, was in synthetic experiments where it was possible to decouple focus distance and focal length, which effectively preserves spatial resolution and is not possible in real life. Further work is needed to confirm the effects of pattern array size on the sim-to-real gap. However, as expected, using synthetic data, PCG methods have been validated to work well, and have proven to be largely invariant to out-of-focus blur with the most observable sim-to-real difference being the lack of real world noise. Under real world conditions, the results show that small circles are marginally more accurate, however, the difference is small, and the dithering effect of the E-Ink display target may have disadvantaged the PCG method relative to the circle grid. The results provide valuable baselines for comparing these methods and to further validate the hypothesis of E-Ink display and pattern incompatibility. Further work can be completed to compare against synthetic cases that more closely mirror the real world case, as well as to eliminate real world issues such as dithering and noise.

## Acknowledgments

We would like to thank our anonymous industry partner who has helped to support this research and we would also like to thank the University of Waterloo's RoboHub for providing additional lab time, expertise, and equipment.

## References

- [1] Y. Hold-Geoffroy, K. Sunkavalli, J. Eisenmann, M. Fisher, E. Gambaretto, S. Hadap, and J.-F. Lalonde, "A Perceptual Measure for Deep Single Image Camera Calibration," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, 2018, pp. 2354–2363, iSSN: 1063-6919.
- [2] M. Lopez, R. Mari, P. Gargallo, Y. Kuang, J. Gonzalez-Jimenez, and G. Haro, "Deep Single Image Camera Calibration With Radial Distortion," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 2019-. IEEE, 2019, pp. 11 809–11 817, iSSN: 1063-6919.
- [3] S. Workman, C. Greenwell, M. Zhai, R. Baltenberger, and N. Jacobs, "DEEPCAL: A method for direct focal length estimation," in *2015 IEEE International Conference on Image Processing (ICIP)*, vol. 2015-. IEEE, 2015, pp. 1369–1373, iSSN: 1522-4880.
- [4] Z. Zhang, "Camera calibration with one-dimensional objects," *IEEE Transactions on Pattern Analysis and Machine Intelligence; IEEE Trans Pattern Anal Mach Intell*, vol. 26, no. 7, pp. 892–899, 2004.
- [5] A. Duda and U. Frese, "Accurate Detection and Localization of Checkerboard Corners for Calibration." in *BMVC*, 2018, p. 126.
- [6] H. Ha, M. Perdoch, H. Alismail, I. S. Kweon, and Y. Sheikh, "Deltille Grids for Geometric Camera Calibration," in *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 5354–5362.
- [7] R. 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*, vol. 3, no. 4, pp. 323–344, 1987.
- [8] M. Rufli, D. Scaramuzza, and R. Siegwart, "Automatic detection of checkerboards on blurred and distorted images," in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2008, pp. 3121–3126.
- [9] Z. Zhang, "A flexible new technique for camera calibration," 2000, iISBN: 1939-3539 Issue: 11 Pages: 1330-1334 Publication Title: - IEEE Transactions on Pattern Analysis and Machine Intelligence Volume: 22.
- [10] Y. Wang, X. Chen, J. Tao, K. Wang, and M. Ma, "Accurate feature detection for out-of-focus camera calibration," *Applied Optics*, vol. 55, no. 28, pp. 7964–7971, 2016. [Online]. Available: <https://opg.optica.org/ao/abstract.cfm?uri=ao-55-28-7964>
- [11] B. Cai, Y. Wang, K. Wang, M. Ma, and X. Chen, "Camera Calibration Robust to Defocus Using Phase-Shifting Patterns," 2017, issue: 10 Publication Title: Sensors Volume: 17.
- [12] Y. Wang, B. Cai, K. Wang, and X. Chen, "Out-of-focus color camera calibration with one normal-sized color-coded pattern," *Optics and Lasers in Engineering*, vol. 98, pp. 17–22, 2017.
- [13] B. Cai, Y. Wang, J. Wu, M. Wang, F. Li, M. Ma, X. Chen, and K. Wang, "An effective method for camera calibration in defocus scene with circular gratings," *Optics and Lasers in Engineering*, vol. 114, pp. 44–49, 2019.
- [14] X. Chen, X. Song, J. Wu, Y. Xiao, Y. Wang, and Y. Wang, "Camera calibration with global LBP-coded phase-shifting wedge grating arrays," *Optics and Lasers in Engineering*, vol. 136, p. 106314, 2021.
- [15] J. Huo, Z. Meng, H. Zhang, S. Chen, and F. Yang, "Feature points extraction of defocused images using deep learning for camera calibration," *Measurement*, vol. 188, p. 110563, 2022.